



Analisis Data MULTIVARIAT

PENYUSUN :

**Prof. Dr. Purnomo, S.T., M.Pd. | Dr. Eddy Sutadji, M.Pd.
Wargijono Utomo, M.Kom. | Okta Purnawirawan, M.Pd,
Rachmat Farich, M.Pd. | Sulistianingsih A.S., M.Pd.
Ratna Fajarwati M., M.Pd. | Annisa Carina, M.Pd.
Najih Gilang R., M.Pd.**

Analisi Data Multivariat

PENULIS

Prof. Dr. Purnomo, S.T., M.Pd; Dr. Eddy Sutadji, M.Pd, dkk

Hak cipta dilindungi oleh Undang-Undang

All Right Reserved

Hak cipta © Prof. Dr. Purnomo, S.T., M.Pd; Dr. Eddy Sutadji, M.Pd, dkk. 2022

PENYUNTING

Wahda Nur L.

ILUSTRATOR SAMPUL

Haidar Hilmy

PENATA LETAK

Dinar Hardini

Diterbitkan oleh Omera Pustaka

Anggota Ikapi

Alamat Kantor: Ajibarang Kulon, Banyumas, Jawa Tengah

Surel: omeracreative@gmail.com

Cetakan I, Februari 2022

Ukuran Buku: 14 x 21 cm

Halaman: 322 Hal.

ISBN: 978-623-5883-31-1

Sanksi Pelanggaran Pasal 113 Undang-undang Nomor 28 Tahun 2014 Tentang Hak Cipta

1. Setiap orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam pasal 9 ayat (1) huruf i untuk penggunaan secara komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan atau pidana denda paling banyak Rp 100.000.000,00 (seratus juta rupiah).
2. Setiap orang yang dengan tanpa hak dan atau tanpa izin pencipta atau pemegang hak cipta melakukan pelanggaran hak ekonomi pencipta sebagaimana dimaksud dalam pasal 9 ayat (1) huruf c, huruf d, huruf f, dan atau huruf h, untuk penggunaan secara komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan atau pidana denda paling banyak Rp 500.000.000,00 (lima ratus juta rupiah).
3. Setiap orang yang dengan tanpa hak dan atau tanpa izin pencipta atau pemegang hak melakukan pelanggaran hak ekonomi pencipta sebagaimana dimaksud dalam pasal 9 ayat (1) huruf a, huruf b, huruf e, dan atau huruf g, untuk penggunaan secara komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan atau pidana denda paling banyak Rp 1.000.000.000,00 (satu miliar rupiah).
4. Setiap orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan atau pidana denda paling banyak Rp 4.000.000.000,00 (empat miliar rupiah).

KATA PENGANTAR

Alhamdulillah, puji syukur kehadirat Allah Swt. atas limpahan rahmat dan karunia-Nya sehingga buku ajar untuk mata kuliah Analisis Data Multivariat dapat diselesaikan tepat waktu. Buku ini merupakan edisi pertama yang dibuat oleh penulis sebagai sumber belajar alternatif bagi dosen, mahasiswa, dan masyarakat umum tentang analisis data multivariat.

Buku ini membahas hal-hal yang berkaitan dengan regresi ganda, analisis jalur, ANOVA dua jalan, analisis diskriminan, korelasi kanonikal, MANOVA, analisis faktor, analisis kluster, dan *multidimensional scaling*. Buku ini juga dilengkapi dengan tutorial menggunakan aplikasi SPSS untuk memudahkan pembaca memahami materi sehingga dapat menambah pemahaman tentang teori yang telah dikuasai.

Terima kasih kami sampaikan kepada Prof. Dr. Purnomo, S.T., M.Pd. dan Dr. Eddy Sutadji, M.Pd. serta rekan-rekan mahasiswa S-3 Pendidikan Kejuruan Fakultas Teknik Universitas Negeri Malang yang ikut andil dalam penyusunan buku ini. Tim penulis menyadari masih terdapat banyak kekurangan dalam buku ini. Oleh karena itu, kritik dan saran terhadap penyempurnaan buku ini sangat diharapkan oleh kami. Semoga buku ini dapat memberi manfaat, khususnya bagi mahasiswa S-3 Pendidikan Kejuruan FT UM dan bagi semua pihak yang membutuhkan.

Malang, 30 Desember 2021

Tim Penulis

Daftar Isi

Kata Pengantar	iii
Daftar Isi	v
Bab 1 Analisis Regresi Linier Berganda	1
Bab 2 Analisis Jalur	29
Bab 3 ANOVA Dua Jalur	67
Bab 4 Analisis Diskriminan	99
Bab 5 Korelasi Kanonikal	129
Bab 6 Analisis Varian Multivariat (MANOVA)	155
Bab 7 Analisis Faktor	205
Bab 8 Analisis Kluster	251
Bab 9 <i>Multidimensional Scalling</i>	285
Tentang Penulis	317

BAB I

ANALISIS REGRESI LINIER BERGANDA

A. Pengertian Analisis Regresi Linier Berganda

Menurut Paiman (2009), analisis regresi digunakan untuk mengetahui hubungan linier antara dua variabel atau lebih. Satu variabel yang berperan sebagai variabel terikat (dependen) biasanya dilambangkan dengan notasi “Y”, sedangkan yang lainnya berperan sebagai variabel bebas (independen) dan dilambangkan dengan notasi “X”. Umumnya, analisis regresi digunakan untuk melakukan prediksi atau ramalan, sedangkan hubungan variabel tersebut bersifat fungsional yang diwujudkan dalam suatu model matematis. Selain itu, analisis regresi juga dipakai untuk memahami variabel yang berhubungan dengan variabel terkait agar lebih mengetahui bentuk-bentuk hubungan tersebut. Secara umum, manfaat penggunaan analisis regresi untuk mengetahui variabel-variabel kunci yang memiliki pengaruh terhadap suatu variabel bergantung pemodal, pendugaan, atau peramalan. Analisis regresi memiliki beberapa manfaat dalam berbagai bidang keilmuan, yakni untuk meramalkan nilai rata-rata variabel bebas yang didasari nilai variabel bebas di luar jangkauan sampel, menguji hipotesis karakteristik dependensi, serta membuat estimasi rata-rata dan nilai variabel berdasarkan nilai variabel bebas (Westriningsih, 2012). Secara umum, persamaan regresi berbentuk sebagai berikut.

$$Y = a + b_1.X_1 + b_2.X_2 + \dots + b_n.X_n$$

Keterangan:

Y = variabel terikat (dependen)

b_1, b_2, \dots, b_n = koefisien regresi

X_1, X_2, \dots, X_n = variabel bebas (independen)

a = konstanta/perpotongan antara sumbu tegak Y dan garis fungsi linier nilai Y

Adapun rumus matriks untuk regresi linier berganda dapat dilihat sebagai berikut.

$$\begin{aligned}
 b_1 &= \frac{(\sum x_2^2) \cdot (\sum x_1 y) - (\sum x_1 x_2) \cdot (\sum x_2 y)}{(\sum x_1^2) \cdot (\sum x_2^2) - (\sum x_1 x_2)^2} & \sum x_1^2 &= \sum X_1^2 \cdot \frac{(\sum x_1)^2}{n} & \sum x_1 y &= \sum X_1 Y \cdot \frac{(\sum x_1) \cdot (\sum Y)}{n} \\
 b_2 &= \frac{(\sum x_1^2) \cdot (\sum x_1 y) - (\sum x_1 x_2) \cdot (\sum x_1 y)}{(\sum x_1^2) \cdot (\sum x_2^2) - (\sum x_1 x_2)^2} & \sum x_2^2 &= \sum X_2^2 \cdot \frac{(\sum x_2)^2}{n} & \sum x_2 y &= \sum X_2 Y \cdot \frac{(\sum x_2) \cdot (\sum Y)}{n} \\
 a &= \frac{\sum Y}{n} - b_1 \cdot \left(\frac{\sum X_1}{n} \right) - b_2 \cdot \left(\frac{\sum X_2}{n} \right) & \sum y^2 &= \sum Y^2 \cdot \frac{(\sum Y)^2}{n} & \sum x_1 x_2 &= \sum X_1 X_2 \cdot \frac{(\sum x_1) \cdot (\sum x_2)}{n}
 \end{aligned}$$

Seperti halnya uji parametrik lainnya, regresi linear juga memiliki syarat atau asumsi klasik yang harus terpenuhi agar model prediksi yang dihasilkan bersifat BLUE (*Best Linear Unbiased Estimation*). Skala data seluruh variabel terutama variabel terikat adalah interval atau rasio. Asumsi ini tidak perlu diuji, cukup kita pastikan bahwa data yang digunakan adalah data interval atau rasio (numerik atau kuantitatif). Asumsi klasik pada regresi linear berganda, antara lain

1. Uji Pemenuhan Syarat (Uji Asumsi Klasik)

Menurut Janie (2012), pada saat melakukan analisis regresi berganda perlu dipenuhi beberapa asumsi, misalnya asumsi klasik yang terdiri dari uji multikolinieritas, uji autokorelasi, uji heteroskedastisitas, uji normalitas, dan uji linieritas.

a. Multikolinieritas

Menurut Ghozali (2016), “*Pengujian multikolinieritas bertujuan untuk mengetahui apakah model regresi terdapat korelasi antarvariabel independen atau variabel bebas.*” Efek dari multikolinieritas menyebabkan tingginya variabel pada sampel. Jika nilai standar eror kurang dari satu dan

nilai koefisien beta juga kurang dari satu, dapat dikatakan bahwa nilai standar eror rendah dan multikolinieritas tidak terdeteksi. Selain itu, apabila nilai VIF kurang dari 10 dan/atau nilai tolerance lebih dari 0,01, dapat disimpulkan dengan tegas bahwa tidak terdapat masalah multikolinieritas, begitu pun sebaliknya.

b. Autokolerasi

Menurut Ghozali (2016), “Autokorelasi dapat muncul karena observasi yang berurutan sepanjang waktu yang berkaitan satu sama lain.” Permasalahan ini muncul karena residual tidak bebas pada satu observasi ke observasi lainnya. Model regresi yang baik adalah model regresi yang bebas dari autokolerasi. Cara menentukan atau kriteria pengujian autokorelasi berdasarkan nilai DW adalah sebagai berikut.

1) Deteksi autokorelasi positif

Jika $dw < dL$ maka terdapat autokorelasi positif.

Jika $dw > dU$ maka tidak terdapat autokorelasi positif.

Jika $dL < dw < dU$ maka pengujian tidak meyakinkan atau tidak dapat disimpulkan.

2) Deteksi autokorelasi negatif

Jika $(4 - dw) < dL$ maka terdapat autokorelasi negatif.

Jika $(4 - dw) > dU$ maka tidak terdapat autokorelasi negatif.

c. Heterokedastisitas

Heterokedastisitas bertujuan untuk melakukan uji apakah pada sebuah model regresi terjadi ketidaknyamanan varian dari residual dalam satu pengamatan ke pengamatan lainnya. Apabila varian berbeda disebut **heteroskedastisitas**. Salah satu cara untuk mengetahui ada tidaknya heteroskedastisitas pada suatu model regresi linier berganda adalah dengan melihat grafik *scatter plot* atau dari nilai prediksi variabel terikat, yaitu

SRESID dengan residual eror (ZPRED). Apabila tidak terdapat pola tertentu dan tidak menyebar di atas maupun di bawah angka nol pada sumbu Y, dapat disimpulkan tidak terjadi heteroskedastisitas. Model penelitian yang baik adalah yang tidak terdapat heteroskedastisitas (Ghozali, 2016).

d. Normalitas

Uji normalitas digunakan untuk mengetahui apakah populasi data berdistribusi normal atau tidak. Uji ini biasanya digunakan untuk mengukur data berskala ordinal, interval, ataupun rasio. Jika analisis menggunakan metode parametrik maka persyaratan normalitas harus terpenuhi, yaitu data berasal dari distribusi yang normal. Jika data tidak berdistribusi normal atau jumlah sampel sedikit dan jenis data adalah nominal atau ordinal, metode yang digunakan adalah statistik nonparametrik. Dalam pembahasan ini, akan digunakan uji *one sample* Kolmogorov-Smirnov dengan menggunakan taraf signifikansi 0,05. Data dinyatakan berdistribusi normal jika signifikansi lebih besar dari 5% atau 0,05.

e. Linieritas

Variabel bebas berhubungan dengan variabel terikat. Asumsi linieritas diuji dengan uji linieritas regresi, misalnya dengan kurva estimasi. Menurut Sugiyono (2015), uji linieritas dapat dipakai untuk mengetahui apakah variabel terikat dengan variabel bebas memiliki hubungan linear atau tidak secara signifikan. Uji linieritas dapat dilakukan melalui *test of linearity*. Kriteria yang berlaku adalah jika nilai signifikansi pada *linearity* $\leq 0,05$, dapat diartikan bahwa antara variabel bebas dan variabel terikat terdapat hubungan yang linear.

2. Uji Kelayakan Model

Metode analisis ini menjadi salah satu analisis yang banyak digunakan karena mudah dan memiliki kekuatan yang cukup dalam menjelaskan suatu

pengaruh suatu variabel bebas ke variabel terikatnya (Narimawati, 2008). Ada banyak sekali kondisi yang dapat kita uji dengan analisis regresi linier. Adapun uji kelayakan model pada analisis regresi berganda dapat dijabarkan sebagai berikut.

a. Uji Keterandalan Model (Uji F)

Uji F bertujuan untuk mengetahui ada atau tidaknya pengaruh simultan (bersama-sama) yang diberikan variabel bebas (X) terhadap variabel terikat (Y).

b. Uji Koefisien Regresi (Uji t)

Uji t bertujuan untuk mengetahui ada atau tidaknya pengaruh parsial (sendiri) yang diberikan variabel bebas (X) terhadap variabel terikat (Y).

c. Koefisien Determinasi

Koefisien determinasi berfungsi untuk mengetahui beberapa persen pengaruh yang diberikan variabel X secara simultan terhadap (Y).

B. Analisis Regresi Linier Berganda dengan SPSS

Dalam kasus ini, untuk menambah pemahaman mengenai analisis regresi berganda kita perlu melakukan uji coba pengujian regresi berganda dengan SPSS. Berikut tahap-tahap yang dilakukan dalam regresi linier berganda menggunakan aplikasi SPSS.

1. Persiapan Data Penelitian

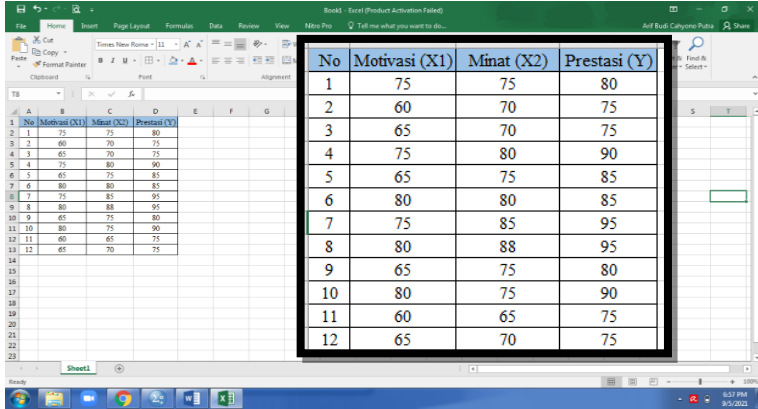
Kita ambil salah satu contoh jika data yang kita masukkan adalah data fiktif. Pembuatan data dapat menggunakan aplikasi Microsoft Excel terlebih dahulu. Adapun data yang digunakan adalah data kuantitatif dengan skala data interval atau rasio, yaitu untuk menguji:

- a. H1: terdapat pengaruh motivasi (X1) terhadap prestasi (Y);
- b. H2: terdapat pengaruh minat (X2) terhadap prestasi (Y); dan

- c. H3: terdapat pengaruh motivasi (X1) dan minat (X2) secara simultan terhadap prestasi (Y) dengan tingkat kepercayaan 95%, $\alpha = 0,05$.

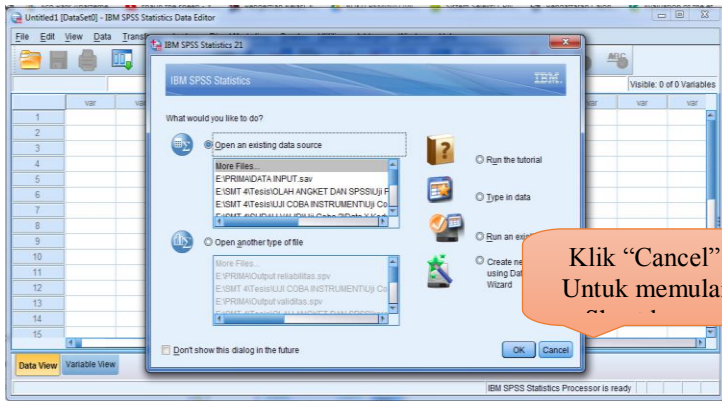
Langkah-Langkah:

- a. *Input* data di Microsoft Excel.

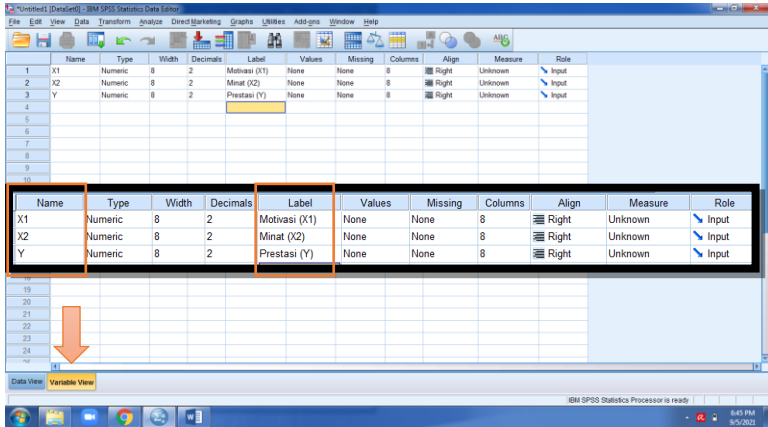


No	Motivasi (X1)	Minat (X2)	Prestasi (Y)
1	75	75	80
2	60	70	75
3	65	70	75
4	75	80	90
5	65	75	85
6	80	80	85
7	75	85	95
8	80	88	95
9	65	75	80
10	80	75	90
11	60	65	75
12	65	70	75

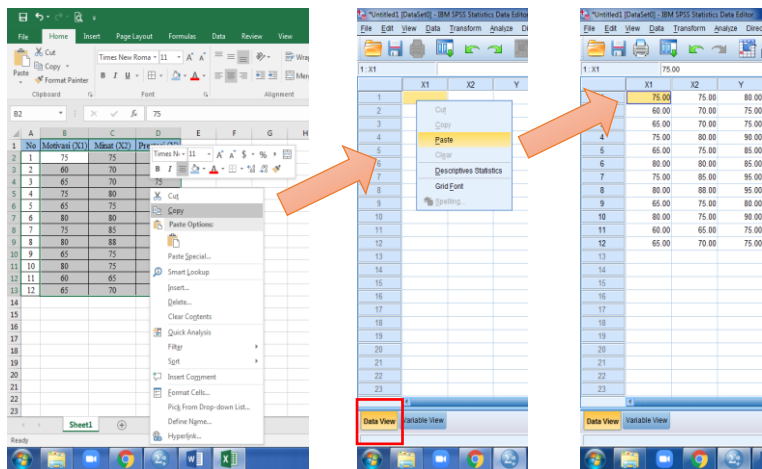
- b. Buka aplikasi SPSS.



- c. Pada Sheet Variable View, isi dengan nama variabel Anda di kolom “Nama” dan kolom “Label”. Jangan rubah isi kolom lainnya.



- d. Salin data yang sudah Anda buat di Microsoft Excel ke dalam SPSS pada Sheet Data View.



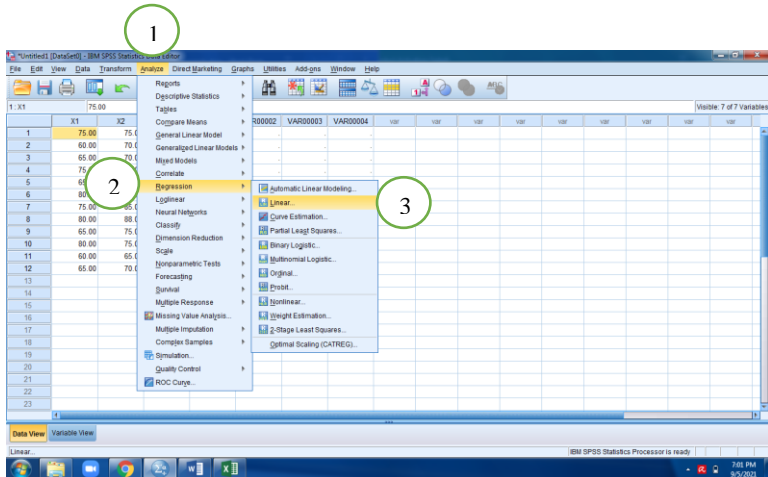
2. Estimasi Model Regresi Linier



Estimasi model dilakukan secara sekaligus dengan pengujian asumsi klasik (normalitas, multikolinieritas, heteroskedastisitas, dan autokorelasi) agar *output* yang dihasilkan dari pengolahan data dapat digunakan untuk uji

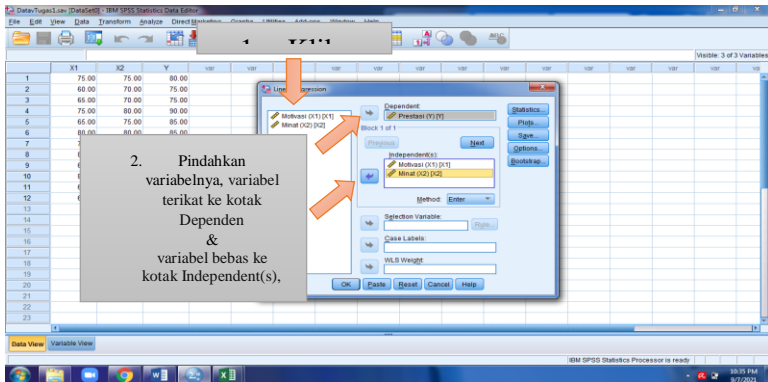
asumsi klasik dan uji kelayakan model. Adapun langkah-langkahnya adalah sebagai berikut.

a. Estimasi regresi linier.

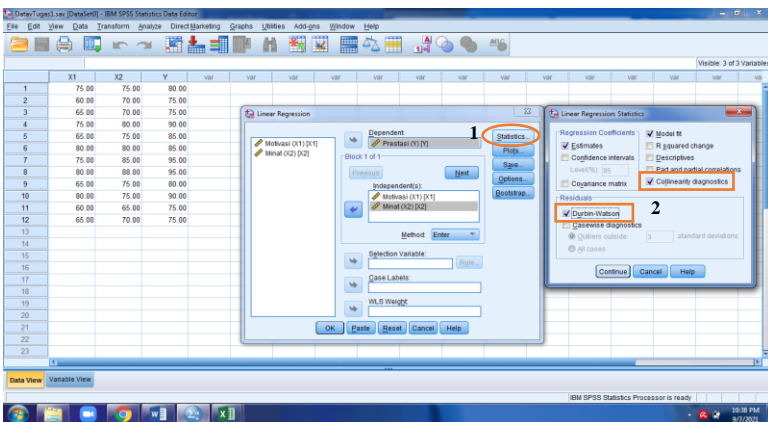
Klik Analyze lalu Regression lalu Linier.



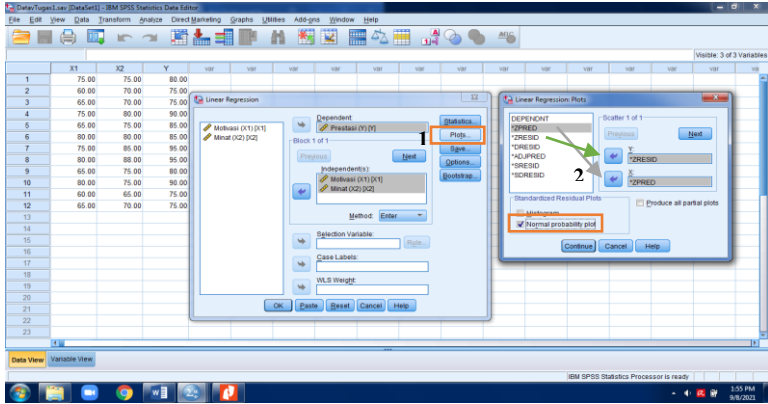
- b. Dari perintah sebelumnya akan muncul tampilan “Linier Regression”.
- c. Pindahkan variabel terikat (dependen) kotak sebelah kiri ke kotak dependen sebelah kanan. Lalu pindahkan variabel bebas (independen) kotak sebelah kiri ke kotak independen(s) sebelah kanan. Caranya adalah tekan variabel yang akan dipindahkan dari kotak kanan lalu tekan tanda panah (). Jika sudah dipindahkan, jangan klik  karena *output* yang diperoleh hanya uji kelayakan modelnya saja, tidak termasuk *output* uji asumsi klasik. Untuk menampilkan *output* uji asumsi klasik, perhatikan langkah berikutnya.



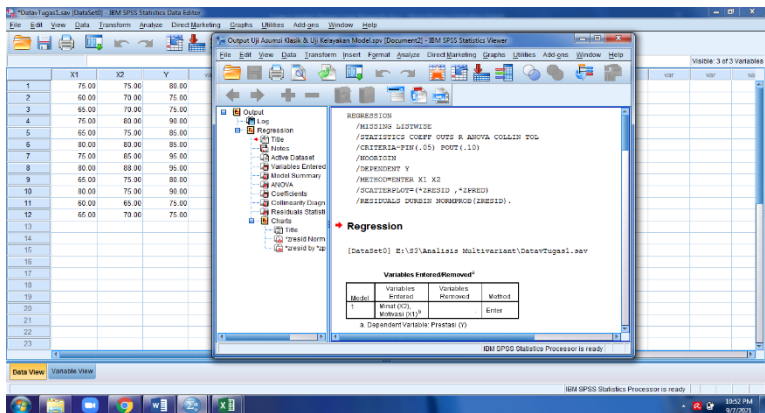
- d. Memunculkan *output* uji asumsi klasik, yaitu normalitas menggunakan Normal PP-Plot; multikolinieritas menggunakan VIF dan tolerance; heteroskedastisitas menggunakan Scatter Plot ZPRED dan ZRESID; dan autokorelasi menggunakan Durbin-Waston. Multikolinieritas dan autokorelasi ada pada tombol Statistics. Centang Collinearity diagnostics untuk memunculkan hasil uji multikolinieritas, kemudian centang Durbin-Waston untuk memunculkan hasil uji autokorelasi. Setelah itu, klik Continue.



- e. Heterokedastisitas dan normalitas ada pada tombol Plot. Untuk memunculkan hasil uji heterokedastisitas, pindahkan *ZPRED ke kotak X dan *ZRESID ke kotak Y. Pindahkan dengan menggunakan tanda panah (↔). Untuk uji normalitas, centang Normal probability plot. Setelah selesai, klik Continue.



- f. Setelah semua tombol perintah di atas sudah selesai, klik OK pada kotak Linear Regression. Dengan demikian, hasil *output* akan ditampilkan pada *file output* seperti pada gambar di bawah ini. Untuk menyimpan data, sebaiknya disimpan keduanya, yaitu *file Data* dan *file Output*.



3. Pengujian Asumsi Klasik

Tahap pengujian asumsi klasik diambil dari *output* SPSS hasil estimasi model regresi linear seperti yang dijelaskan sebelumnya. Selanjutnya, hasil *output* tersebut diinterpretasi dari masing-masing hasil pengujian, yaitu uji asumsi klasik yang terdiri dari uji multikolinieritas, uji autokorelasi dan uji heteroskedastisitas, serta uji normalitas; dan uji kelayakan model yang terdiri dari uji keterandalan model (uji F), uji koefisien regresi (uji t), serta koefisien determinasi.

a. Multikolinieritas

Tabel 1.1 Coefficients^a.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	2.971	11.753		.253	.806		
1 Motivasi (X1)	.212	.214	.213	.992	.347	.380	2.634
Minat (X2)	.864	.251	.740	3.441	.007	.380	2.634

Dependent Variable: Prestasi (Y).

Pada Tabel 1.1, didapat nilai standar error < 1 , yaitu untuk variabel motivasi (X1) = 0.214 dan variabel minat (X2) = 0.251 dengan nilai keduanya < 1 . Nilai koefisien beta juga < 1 , yaitu untuk variabel motivasi (X1) = 0.212 dan variabel minat (X2) = 0.864. Dapat dikatakan, nilai standar error rendah dan multikolinieritas tidak terdeteksi.

Nilai VIF untuk variabel motivasi (X1) dan minat (X2) sama-sama 2.634, sedangkan nilai tolerance sebesar 0.380. Karena nilai VIF dari kedua tidak ada yang lebih besar dari 10 atau 5, dapat dikatakan tidak terjadi multikolinieritas pada kedua variabel bebas tersebut (Gujarati, 2012).

Berdasarkan syarat asumsi klasik regresi linier, model regresi linier

yang baik adalah yang terbebas dari adanya multikolinearitas. Dengan demikian, model di atas telah terbebas dari adanya multikolinearitas.

b. Autokorelasi

Data yang digunakan untuk mengestimasi model regresi linier merupakan data *time series* sehingga diperlukan uji asumsi terbebas dari multikolinearitas. Hasil uji autokorelasi dapat dilihat pada Tabel 1.2.

Tabel 1.2 Model Summary^b.

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.918 ^a	.842	.807	3.42127	2.115

a. Predictors: (Constant), Minat (X2), Motivasi (X1)

b. Dependent Variable: Prestasi (Y)

Nilai Durbin-Waston yang tertera pada *output* SPSS disebut DW hitung. Angka ini dibandingkan dengan kriteria penerimaan atau penolakan yang akan dibuat dengan nilai d_L . d_U ditentukan berdasarkan jumlah variabel bebas dalam model regresi (k) dan jumlah sampelnya (n). Nilai d_L dan d_U dapat dilihat pada tabel DW dengan tingkat signifikansi (error) 5% ($\alpha = 0,05$).

Jumlah variabel bebas (X) : $k = 2$
 Jumlah sampel : $n = 12$

Tabel Durbin-Watson (DW), $\alpha = 5\%$

n	k=1		k=2	
	d_L	d_U	d_L	d_U
6	0.6102	1.4002		
7	0.6996	1.3564	0.4672	1.8964
8	0.7629	1.3324	0.5591	1.7771
9	0.8243	1.3199	0.6291	1.6993
10	0.8791	1.3197	0.6972	1.6413
11	0.9273	1.3241	0.7580	1.6044
12	0.9708	1.3314	0.8122	1.5794
13	1.0097	1.3404	0.8612	1.5621
14	1.0450	1.3503	0.9054	1.5507

Hasil dari tabel Durbin-Waston menunjukkan bahwa nilai $d_L = 0.8122$ dan $d_U = 1.5794$. Di sisi lain, pada *output* SPSS diperoleh nilai $DW = 2.115$ sehingga dapat ditentukan kriteria terjadi atau tidaknya autokorelasi sebagai berikut.

1) *Deteksi autokorelasi positif*

Jika $2.115 > 1.5794$ maka tidak terdapat autokorelasi positif.

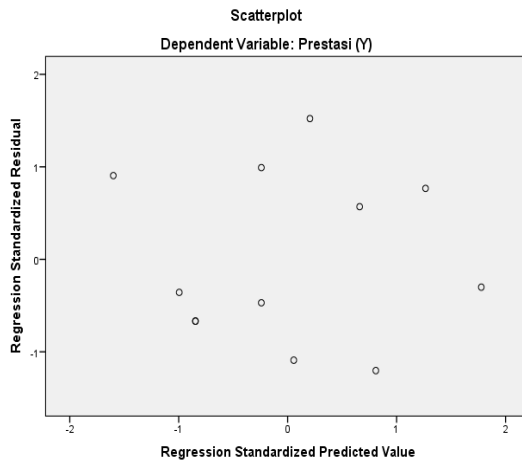
2) *Deteksi autokorelasi negatif*

Jika $(4 - 2.115) > 1.5794$ maka tidak terdapat autokorelasi negatif.

Oleh karena itu, dapat disimpulkan bahwa model regresi linier di atas tidak ada masalah autokorelasi.

c. Heteroskedastisitas

Pengujian heteroskedastisitas dilakukan dengan membuat *scatterplot* (alur sebaran) antara residual dan nilai prediksi dari variabel terikat yang telah distandarisasi. Hasil uji heteroskedastisitas dapat dilihat pada gambar *scatterplot* seperti pada gambar di bawah ini.

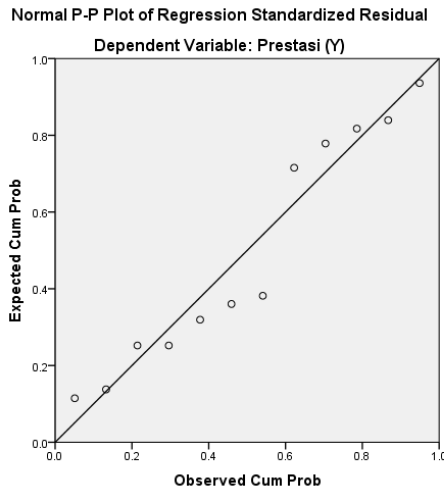


Gambar 1.1 *Scatter Plot.*

Berdasarkan gambar di atas, terlihat bahwa sebaran titik tidak membentuk suatu pola atau alur tertentu sehingga dapat disimpulkan tidak terjadi heteroskedastisitas. Uji asumsi klasik dalam model regresi linier di atas yaitu terbebas dari heteroskedastisitas.

d. Normalitas

Hasil uji normalitas dapat dilihat dari gambar Normal P-P Plot di bawah ini. Kriteria sebuah data residual terdistribusi normal atau tidak dengan menggunakan pendekatan Normal P-P Plot dapat dilakukan dengan melihat sebaran titik yang ada pada gambar. Apabila sebaran titik tersebut mendekati atau rapat pada gari lurus (diagonal), artinya data residual terdistribusi normal. Namun, jika sebaran titik-titik tersebut menjauhi garis maka tidak terdistribusi normal.



Gambar 1.2 Normal P-P Plot of Regression Standardized Residual.

Sebaran titik-titik dari gambar di atas relatif mendekati garis lurus sehingga dapat disimpulkan bahwa data residual terdistribusi normal.

e. *Linieritas*

Tabel 1.3 ANOVA Table.

		Sum of Squares	df	Mean Square	F	Sig.
Prestasi (Y) * Motivasi (X1)	(Combined)	431.250	3	143.750	4.885	.032
	Between Groups					
	Linearity	422.704	1	422.704	14.364	.005
	Deviation from	8.546	2	4.273	.145	.867
Within Groups		235.417	8	29.427		
Total		666.667	11			

Berdasarkan hasil uji linieritas, diketahui nilai Sig. Deviation from Linierity sebesar $0,867 > 0,05$. Artinya, terdapat hubungan yang linier antara variabel motivasi (X1) dan prestasi (Y).

Tabel 1.4 ANOVA Table.

		Sum of Squares	df	Mean Square	F	Sig.
Prestasi (Y) * Minat (X2)	(Combined)	585.417	5	117.083	8.646	.010
	Between Groups					
	Linearity	549.801	1	549.801	40.601	.001
	Deviation from	35.616	4	8.904	.658	.643
	Linearity					
Within Groups		81.250	6	13.542		
Total		666.667	11			

Berdasarkan hasil uji linieritas, diketahui nilai Sig. Deviation from Linierity sebesar $0,643 > 0,05$. Artinya, terdapat hubungan yang linier antara variabel minat (X2) dan prestasi (Y).

4. Uji Kelayakan Model

a. Uji Keterandalan Model (Uji F)

Menurut Yuliara (2016), uji keterandalan atau yang biasa disebut dengan Uji F merupakan tahap awal mengidentifikasi model regresi yang diestimasi layak atau tidak. Layak merupakan model diestimasi layak digunakan untuk menjelaskan pengaruh variabel-variabel bebas (X) terhadap variabel terikat (Y). Hasil uji F dapat dilihat pada Tabel 1.5 di bawah ini. Nilai *prob.* F hitung terlihat pada kolom terakhir (Sig.).

Tabel 1.5 ANOVA^a.

	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	561.321	2	280.660	23.978	.000 ^b
	Residual	105.346	9	11.705		
	Total	666.667	11			

a. Dependent Variable: Prestasi (Y)

b. Predictors: (Constant), Minat (X2), Motivasi (X1)

Nilai *prob.* F hitung (Sig.) pada tabel di atas nilainya 0,000 lebih kecil dari tingkat signifikansi 0,05 sehingga dapat disimpulkan bahwa model regresi linier yang diestimasi layak digunakan untuk menjelaskan pengaruh variabel bebas (X) terhadap variabel terikat (Y).

Uji F berdasarkan perhitungan manual menggunakan F tabel:

- 1) Jika nilai Sig. < 0,05, atau F hitung > F tabel, maka terdapat pengaruh variabel X secara simultan terhadap variabel Y.
- 2) Jika nilai Sig. > 0,05, atau F hitung < F tabel, maka tidak terdapat pengaruh variabel X secara simultan terhadap variabel Y.

k = jumlah variabel bebas
 n = jumlah sampel

Keterangan:

F tabel = F (k ; $n-k$)

= F (2 ; 12-2)

= F (2; 10) → lihat tabel F sehingga diperoleh F tabel =

4,10

Berdasarkan *output* SPSS di atas, diketahui nilai signifikansi untuk pengaruh X1 dan X2 secara simultan terhadap Y adalah sebesar $0,000 < 0,05$ dan nilai F hitung $23,978 > F$ tabel 4,10.

Tabel Nilai $F_{0,05}$

	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18,5	19,0	19,2	19,2	19,3	19,3	19,4	19,4	19,4	19,4	19,4	19,4	19,4	19,5	19,5	19,5	19,5	19,5	19,5
3	10,1	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,74	8,70	8,66	8,64	8,62	8,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,91	5,86	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,68	4,62	4,56	4,53	4,50	4,46	4,43	4,40	4,37
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,00	3,94	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,57	3,51	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,28	3,22	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,07	3,01	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,91	2,85	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,79	2,72	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,69	2,62	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,13	3,03	2,92	2,83	2,77	2,71	2,67	2,60	2,53	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,53	2,46	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,48	2,40	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,42	2,35	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,38	2,31	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,34	2,27	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,31	2,23	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,28	2,20	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,25	2,18	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,23	2,15	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,20	2,13	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,18	2,11	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,16	2,09	2,01	1,96	1,92	1,87	1,82	1,77	1,71
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,09	2,01	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,00	1,92	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,92	1,84	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,18	2,09	2,02	1,96	1,91	1,83	1,75	1,66	1,61	1,55	1,50	1,43	1,35	1,22
3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,75	1,67	1,57	1,52	1,46	1,39	1,32	1,22	1,00	

Sumber: Aplikasi Analisis Multivariate Dengan Program SPSS (Dr. Imam Ghozali)

Gambar 1.3 Tabel Nilai $F_{0,05}$.

Oleh karena itu, dapat disimpulkan bahwa model regresi linier tersebut diterima, yang berarti terdapat pengaruh variabel bebas motivasi (X1) dan variabel bebas minat (X2) secara simultan terhadap variabel terikat

prestasi (Y). Artinya, hiipotesis **H3 diterima atau terdapat pengaruh motivasi (X1) dan minat (X2) secara simultan terhadap prestasi (Y).**

b. Uji Koofisien Regresi (Uji t)

Uji t dalam regresi linier berganda dimaksudkan untuk menguji apakah parameter yang diduga untuk mengestimasi persamaan atau model regresi linier berganda merupakan parameter yang tepat atau belum, atau dapat menjelaskan perilaku variabel bebas dalam mempengaruhi variabel terikatnya. Hasil pengujian dapat dilihat pada Tabel 1.6 seperti gambar di bawah ini.

Tabel 1.6 Coefficients^a.

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
	B	Std. Error	Beta			Tolerance	VIF
(Constant)	2.971	11.753		.253	.806		
1 Motivasi (X1)	.212	.214	.213	.992	.347	.380	2.634
Minat (X2)	.864	.251	.740	3.441	.007	.380	2.634

Dependent Variable: Prestasi (Y).

Kesimpulan pada Uji t, apabila nilai *prob. t* hitung < 0,05 artinya variabel bebas berpengaruh signifikan terhadap variabel terikatnya. Nilai *prob. t* hitung variabel bebas motivasi (X1) sebesar 0,347 yang lebih besar dari 0,05 sehingga variabel bebas motivasi (X1) **tidak berpengaruh** terhadap variabel terikat prestasi (Y). Sementara itu, nilai *prob. t* hitung pada variabel bebas minat (X2) sebesar 0,007 yang lebih kecil dari 0,05 sehingga variabel bebas motivasi (X1) **berpengaruh** terhadap variabel terikat prestasi (Y).

Uji t berdasarkan perhitungan manual menggunakan t tabel:

- 1) Jika nilai Sig. < 0,05, atau t hitung > t tabel, maka terdapat pengaruh variabel X terhadap variabel Y.

- 2) Jika nilai Sig. > 0,05, atau t hitung < t tabel, maka tidak terdapat pengaruh variabel X terhadap Y.

k = jumlah variabel bebas
n = jumlah sampel

Keterangan:

t tabel = $t (\alpha/2 ; n-k-1)$
 = $t (0,05/2 ; 12-2-1)$
 = $t (0,025 ; 9) \rightarrow$ lihat tabel t sehingga diperoleh t tabel = 2,262

Tabel Nilai t

d.f	$t_{0,10}$	$t_{0,05}$	$t_{0,025}$	$t_{0,01}$	$t_{0,005}$	d.f
1	3,078	6,314	12,706	31,821	63, 657	1
2	1,886	2,920	4,303	6,965	9,925	2
3	1,638	2,353	3,182	4,541	5,841	3
4	1,533	2,132	2,776	3,747	4,604	4
5	1,476	2,015	2,571	3,365	4,032	5
6	1,440	1,943	2,447	3,143	3,707	6
7	1,415	1,895	2,365	2,998	3,499	7
8	1,397	1,860	2,306	2,896	3,355	8
9	1,383	1,833	2,262	2,821	3,250	9
10	1,372	1,812	2,228	2,764	3,169	10
11	1,363	1,796	2,201	2,718	3,106	11

Gambar 1.4 Tabel Nilai t.

Diketahui, nilai Sig. untuk pengaruh variabel bebas motivasi (X1) terhadap variabel terikat prestasi (Y) adalah $0,347 > 0,05$ dan nilai t hitung $0,992 < t$ tabel 2,262 sehingga dapat disimpulkan **ditolak**, yang berarti tidak terdapat pengaruh variabel bebas motivasi (X1) terhadap variabel terikat prestasi (Y).

Di sisi lain, nilai Sig. untuk pengaruh variabel bebas minat (X2) terhadap variabel terikat prestasi (Y) adalah $0,007 < 0,05$ dan nilai t hitung

3,441 > t tabel 2,262 sehingga dapat disimpulkan **diterima**, yang berarti terdapat pengaruh variabel bebas minat (X2) terhadap variabel terikat prestasi (Y).

Dengan kata lain, hipotesis **H1 ditolak atau tidak terdapat pengaruh motivasi (X1) terhadap prestasi (Y), sedangkan H2 diterima atau terdapat pengaruh minat (X2) terhadap prestasi (Y).**

c. Koefisien Determinasi

Koefisien determinasi menjelaskan variasi pengaruh variabel-variabel bebas terhadap variabel terikatnya atau dapat pula dikatakan sebagai proporsi pengaruh seluruh variabel bebas terhadap variabel terikat. Nilai koefisien determinasi regresi linier berganda dapat diukur oleh nilai **Adjusted R-Square**, yaitu pada saat jumlah variabel bebas lebih dari satu.

Tabel 1.7 Model Summary^b.

Model	R	R-Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.918 ^a	.842	.807	3.42127	2.115

a. Predictors: (Constant), Minat (X2), Motivasi (X1)

b. Dependent Variable: Prestasi (Y)

Jika dilihat dari nilai R-Square yang besarnya 0.807, menunjukkan bahwa proporsi pengaruh variabel bebas motivasi (X1) dan variabel minat (X2) terhadap variabel terikat prestasi (Y) sebesar 80.7%. Artinya, motivasi dan minat memiliki proporsi pengaruh terhadap prestasi siswa sebesar 80.7%, sedangkan sisanya dipengaruhi oleh variabel lain.

d. Analisis Regresi Berganda Menggunakan Manual Microsoft Excel

- 1) *Input data.*

No	Motivasi (X1)	Minat (X2)	Prestasi (Y)
1	75	75	80
2	60	70	75
3	65	70	75
4	75	80	90
5	65	75	85
6	80	80	85
7	75	85	95
8	80	88	95
9	65	75	80
10	80	75	90
11	60	65	75
12	65	70	75

2) Buat kolom pertolongan sesuai yang diminta dari rumus regresi berganda.


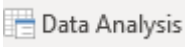
$$\begin{aligned} \sum x_{\epsilon}^2 &= \sum X_{\epsilon}^2 - \frac{(\sum X_2)^2}{n} \\ \sum x_{\epsilon}^2 &= \sum X_{\epsilon}^2 - \frac{(\sum X_2)^2}{n} \\ \sum x_{\epsilon}^2 &= \sum X_{\epsilon}^2 - \frac{(\sum Y)^2}{n} \\ \sum x_1 y &= \sum X_1 Y - \frac{(\sum X_1)(\sum Y)}{n} \\ \sum x_2 y &= \sum X_2 Y - \frac{(\sum X_2)(\sum Y)}{n} \\ \sum x_2 y_2 &= \sum X_1 Y_2 - \frac{(\sum X_1)(\sum X_2)}{n} \end{aligned}$$

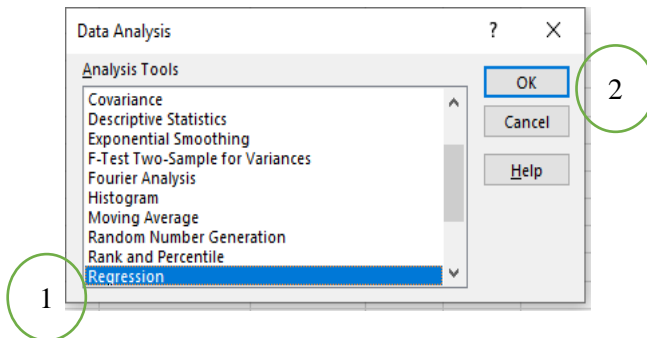
$$\begin{aligned} b_1 &= \frac{(\sum x_2^2)(\sum x_1 y) - (\sum x_1 x_2)(\sum x_2 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} \\ b_2 &= \frac{(\sum x_1^2)(\sum x_2 y) - (\sum x_1 x_2)(\sum x_1 y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1 x_2)^2} \\ \alpha &= \frac{\sum Y}{n} - b_1 \left(\frac{\sum X_1}{n} \right) - b_2 \left(\frac{\sum X_2}{n} \right) \end{aligned}$$

No	Motivasi (X1)	Minat (X2)	Prestasi (Y)	X1'2	X2'2	Y'2	X1Y	X2Y	X1X2
1	75	75	80	5625	5625	6400	6000	6000	5625
2	60	70	75	3600	4900	5625	4500	5250	4200
3	65	70	75	4225	4900	5625	4875	5250	4550
4	75	80	90	5625	6400	8100	6750	7200	6000
5	65	75	85	4225	5625	7225	5525	6375	4875
6	80	80	85	6400	6400	7225	6800	6800	6400
7	75	85	95	5625	7225	9025	7125	8075	6375
8	80	88	95	6400	7744	9025	7600	8360	7040
9	65	75	80	4225	5625	6400	5200	6000	4875
10	80	75	90	6400	5625	8100	7200	6750	6000
11	60	65	75	3600	4225	5625	4500	4875	3900
12	65	70	75	4225	4900	5625	4875	5250	4550
Σ	845	908	1000	60175	69194	84000	70950	76185	64390

3) Masukkan perhitungan sesuai dengan rumus regresi linier berganda.

$\sum x_{\varepsilon}^2 = \sum X_{\varepsilon}^2 - \frac{(\sum X_2)^2}{n}$	$\sum x1^2$	672,9167
$\sum x_{\varepsilon}^2 = \sum X_{\varepsilon}^2 - \frac{(\sum X_2)^2}{n}$	$\sum x2^2$	488,6667
$\sum x_{\varepsilon}^2 = \sum X_{\varepsilon}^2 - \frac{(\sum Y)^2}{n}$	$\sum y^2$	666,6667
$\sum x_1y = \sum X_1Y - \frac{(\sum X_1)(\sum Y)}{n}$	$\sum x1y$	533,3333
$\sum x_2y = \sum X_2Y - \frac{(\sum X_2)(\sum Y)}{n}$	$\sum x2y$	518,3333
$\sum x_2y_2 = \sum X_1Y_2 - \frac{(\sum X_1)(\sum X_2)}{n}$	$\sum x1x2$	451,6667
$b_1 = \frac{(\sum x_2^2)(\sum x_1y) - (\sum x_1x_2)(\sum x_2y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2}$	b1	0,212357
$b_2 = \frac{(\sum x_1^2)(\sum x_2y) - (\sum x_1x_2)(\sum x_1y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2}$	b2	0,864431
$\alpha = \frac{\sum Y}{n} - b_1 \left(\frac{\sum X_1}{n} \right) - b_2 \left(\frac{\sum X_2}{n} \right)$	a	2,971227

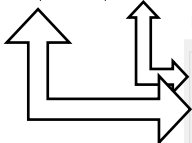
- 4) Bandingkan hasil menggunakan analisis regresi di Microsoft Excel, langkahnya klik  lalu klik  maka akan keluar perintah sebagai berikut.



- 5) Jika sudah keluar perintah Data Analysis lalu pilih Regression dan OK. Maka akan keluar perintah sebagai berikut.

No	Motivasi (X1)	Minat (X2)	Prestasi (Y)
1	75	75	80
2	60	70	75
3	65	70	75
4	75	80	90
5	65	75	85
6	80	80	85
7	75	85	95
8	80	88	95
9	65	75	80
10	80	75	90
11	60	65	75
12	65	70	75

Tempatkan di kolom yang



? X

Regression

Input

Input Y Range: ↑

Input X Range: ↑

Labels Constant is zero

Confidence Level: %

Output options

Output Range: ↑

New Worksheet Ply:

New Workbook

Residuals

Residuals Residual Plots

Standardized Residuals Line Fit Plots

Normal Probability

Normal Probability Plots

6) Hasil analisis regresi linier berganda akan keluar seperti berikut ini.

SUMMARY OUTPUT									
Regression Statistics									
Multiple R	0,917595181								
R Square	0,841980916								
Adjusted R Square	0,806865563								
Standard Error	3,421274231								
Observations	12								
ANOVA									
		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression		2	561,321	280,66	23,9776	0,00025			
Residual		9	105,346	11,7051					
Total		11	666,667						
		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
Intercept		2,971227344	11,7528	0,25281	0,80609	-23,6155	29,5579	-23,6155	29,5579
X Variable 1		0,212356888	0,21406	0,99204	0,3471	-0,27188	0,6966	-0,27188	0,6966
X Variable 2		0,86443139	0,25119	3,44128	0,00738	0,29619	1,43267	0,29619	1,43267

7) Jika dibandingkan hasilnya akan seperti di bawah ini.

$$\sum x_g^2 = \sum X_g^2 - \frac{(\sum X_g)^2}{n}$$

$$\sum x_g^2 = \sum X_g^2 - \frac{(\sum X_g)^2}{n}$$

$$\sum x_g^2 = \sum X_g^2 - \frac{(\sum Y)^2}{n}$$

$$\sum x_1y = \sum X_1Y - \frac{(\sum X_1)(\sum Y)}{n}$$

$$\sum x_2y = \sum X_2Y - \frac{(\sum X_2)(\sum Y)}{n}$$

$$\sum x_2y_2 = \sum X_1Y_2 - \frac{(\sum X_1)(\sum X_2)}{n}$$

SUMMARY OUTPUT										
Regression Statistics										
Multiple R	0,917595181									
R Square	0,841980916									
Adjusted R Square	0,806865563									
Standard Error	3,421274231									
Observations	12									
ANOVA										
	df	SS	MS	F	Significance F					
Regression	2	561,321	280,66	23,9776	0,00025					
Residual	9	105,346	11,7051							
Total	11	666,667								
Coefficients		Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95,0%	Upper 95,0%		
b1	0,212357	Intercept	2,971227344	11,7528	0,25281	0,80609	-23,6155	29,5579	-23,6155	29,5579
b2	0,864431	X Variable 1	0,212356888	0,21406	0,99204	0,3471	-0,27188	0,6966	-0,27188	0,6966
a	2,971227	X Variable 2	0,86443139	0,25119	3,44128	0,00738	0,29619	1,43267	0,29619	1,43267

$$b_1 = \frac{(\sum x_1^2)(\sum x_2y) - (\sum x_1x_2)(\sum x_1y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2}$$

$$b_2 = \frac{(\sum x_2^2)(\sum x_1y) - (\sum x_1x_2)(\sum x_2y)}{(\sum x_1^2)(\sum x_2^2) - (\sum x_1x_2)^2}$$

$$\alpha = \frac{\sum Y}{n} - b_1 \left(\frac{\sum X_1}{n} \right) - b_2 \left(\frac{\sum X_2}{n} \right)$$

Tabel 1.8 Coefficients^a.

	Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
	(Constant)	2.971	11.753		.253	.806		
1	Motivasi (X1)	.212	.214	.213	.992	.347	.380	2.634
	Minat (X2)	.864	.251	.740	3.441	.007	.380	2.634

Dependent Variable: Prestasi (Y).

8) Dengan demikian, diperoleh kesimpulan bahwa hasil perhitungan menggunakan manual Excel, analisis data regresi di Microsoft Excel dan analisis menggunakan SPSS adalah **sama**.

E. Kesimpulan

- Berdasarkan uji asumsi klasik regresi linier, model regresi linier di atas yang dianalisis menggunakan SPSS telah terbebas dari adanya multikolinieritas, tidak ada masalah autokorelasi, terbebas dari

heteroskedastisitas, dan data residual terdistribusi normal sehingga memenuhi syarat dalam uji asumsi klasik untuk agresi linier.

2. Berdasarkan uji kelayakan model pada uji F dapat disimpulkan bahwa model regresi linier tersebut diterima, yang berarti terdapat pengaruh variabel bebas motivasi (X_1) dan variabel bebas minat (X_2) secara simultan terhadap variabel terikat prestasi (Y). Sementara itu, pada uji t tidak terdapat pengaruh variabel bebas motivasi (X_1) terhadap variabel terikat prestasi (Y) akan tetapi terdapat pengaruh variabel bebas minat (X_2) terhadap variabel terikat prestasi (Y). Pada koefisien determinasi, motivasi dan minat memiliki proporsi pengaruh terhadap prestasi siswa sebesar 80.7%, sedangkan sisanya dipengaruhi oleh variabel lain.

Daftar Rujukan

- Ghozali, I. (2016). *Aplikasi Analisis Multivariete dengan Program IBM SPSS* 23. Edisi 8. Semarang: Badan Penerbit.
- Gujarati. (2012). *Dasar-Dasar Ekonometrika*. Jakarta: Salemba Empat.
- Janie, Dyah N.A. (2012). *Statistik Deskriptif & Regresi Linier Berganda dengan SPSS*. Semarang: Semarang University Press.
- Narimawati, Umi. (2008). *Metodologi Penelitian Kualitatif dan Kuantitatif, Teori, dan Aplikasi*. Bandung: Agung Media.
- Paiman. (2019). *Teknik Analisis Korelasi dan Regresi Ilmu-Ilmu Pertanian*. Yogyakarta: UPY Press.
- Sugiyono. (2015). *Metode Penelitian Kombinasi (Mix Methods)*. Bandung: Alfabeta.
- Westriningsih. (2012). *Solusi Praktis dan Mudah Menguasai SPSS 20 untuk Pengolahan Data*. Yogyakarta: CV Andi Offset.
- Yuliara, I Made. (2016). *Modul Regresi Linier Berganda*. Bali: Universitas Udayana. Fakultas Matematika dan Ilmu Pengetahuan Alam.

BAB 2

ANALISIS JALUR

A. Definisi *Path Analysis* (Analisis Jalur)

Path analysis atau yang dikenal dengan analisis jalur dikembangkan pertama kali pada tahun 1920 oleh seorang ahli genetika, yaitu Sewall Wright (Joreskog dan Sorbom, 1996; Johnson & Wicherb, 1992 dalam Riduwan dan Kuncoro, 2014). Menurut Wright, *path analysis* (analisis jalur) merupakan pengembangan korelasi. Lebih lanjut, analisis jalur mempunyai kedekatan dengan regresi berganda. Dengan kata lain, analisis regresi berganda merupakan bentuk khusus dari analisis jalur (*regression is special case of path analysis*) (Raltan dan Renhard, 2014; serta Sugiyono, 2015).

Teknik ini juga dikenal sebagai model sebab akibat (*causing modeling*). Penamaan ini disebabkan oleh analisis jalur memungkinkan pengguna dapat menguji proporsi teoritis mengenai hubungan sebab dan akibat tanpa memanipulasi variabel. *Path analysis* diartikan oleh Bohnrstedt (1974 dalam Kusnendi, 2005) bahwa “*A technique for estimating the effect a set of independent variables has on a dependent variable from a set of observed correlations, given a set of hypothesized causal asymmetric relation among the variable.* Tujuan utama dari *path analysis* adalah

“*A method of measuring the direct influence along each separate path in such a system and thus of finding the degree to which variation of a given effect is determined by each particular cause. The method depend on combination of knowledge of degree of correlation among the variable in a system with such knowledge as may possessed of the causal relations*” (Maruyama, 1998 dalam Riduwan dan Kuncoro,

2014).

Jadi, model *path analysis* digunakan untuk menganalisis pola hubungan antarvariabel guna mengetahui pengaruh langsung maupun tidak langsung seperangkat variabel bebas (eksogen) terhadap variabel terikat (endogen). Model *path analysis* yang dibicarakan adalah bila hubungan sebab akibat atau “*A set of hypothesized causal asymmetric relation among the variables*”. Oleh sebab itu, rumusan masalah penelitian dalam kerangka *path analysis* berkisar pada: 1) Apakah variabel eksogen (X_1, X_2, \dots, X_k) berpengaruh terhadap variabel endogen Y ? dan 2) Berapa besar pengaruh kausal langsung, kausal tidak langsung, kausal total, maupun simultan seperangkat variabel eksogen (X_1, X_2, \dots, X_k) terhadap variabel endogen Y ?

Riduwan dan Kuncoro (2014) menyebutkan teknik analisis jalur digunakan untuk menguji besarnya sumbangan (kontribusi) yang ditunjukkan oleh koefisien jalur pada setiap diagram jalur dari hubungan kausal antarvariabel X_1, X_2 , dan X_3 terhadap Y terhadap Z . Analisis korelasi dan regresi yang merupakan dasar perhitungan koefisien jalur.

David dan Garson (dalam Raltan dan Renhard, 2014) mendefinisikan analisis jalur sebagai model perluasan regresi yang digunakan untuk menguji keselarasan matriks korelasi dengan dua atau lebih model hubungan sebab akibat yang dibandingkan oleh peneliti. Modelnya digambarkan dalam bentuk gambar lingkaran dan anak panah. Pada diagram jalur, digunakan dua macam anak panah, yaitu 1) anak panah satu arah yang menyatakan pengaruh langsung dari sebuah variabel eksogen (variabel penyebab (X)) terhadap sebuah variabel endogen (variabel akibat (Y)), misalnya: $X_1 \rightarrow Y$ dan 2) anak panah dua arah yang menyatakan hubungan korelasi antarvariabel eksogen, misalnya $X_1 \leftrightarrow X_2$.

Regresi dikenakan pada masing-masing variabel dalam suatu model sebagai variabel tergantung (pemberi respons), sedangkan yang lain sebagai penyebab. Pembobotan regresi diprediksikan dalam suatu model yang dibandingkan dengan matriks korelasi yang diobservasi untuk semua variabel dan dilakukan juga perhitungan uji keselarasan statistik. Dari beberapa definisi yang telah dipaparkan, dapat ditarik kesimpulan bahwa sebenarnya *path analysis* (analisis jalur) merupakan kepanjangan dari analisis regresi berganda.

B. Prinsip Dasar *Path Analysis* (Analisis Jalur)

Menurut Raltan dan Renhard, (2014) terdapat beberapa prinsip dasar yang sebaiknya dipenuhi dalam analisis jalur sebagai berikut.

1. Adanya linieritas (*linierity*). Hubungan antarvariabel bersifat linier.
2. Adanya aditivitas (*additivity*). Tidak ada efek-efek interaksi.
3. Data berskala interval. Semua variabel yang diobservasi dalam bentuk data berskala interval (*scaled values*). Jika data belum dalam bentuk skala interval, sebaiknya data diubah dengan menggunakan metode suksesi interval (MSI) terlebih dahulu.
4. Semua variabel residual (yang tidak diukur) tidak berkorelasi dengan salah satu variabel dalam model.
5. Istilah gangguan (*disturbance terms*) atau variabel residual tidak boleh berkorelasi dengan semua variabel endogen dalam model. Jika dilanggar maka akan berakibat hasil regresi menjadi tidak tepat untuk mengestimasi parameter-parameter jalur.
6. Sebaiknya hanya terdapat multikolinieritas yang rendah. Maksud multikolinieritas adalah dua atau lebih variabel bebas (penyebab) mempunyai hubungan yang sangat tinggi. Jika terjadi hubungan yang

tinggi maka kita akan mendapatkan standar eror yang besar dari koefisien beta (β) yang digunakan untuk menghilangkan varian biasa dalam melakukan analisis korelasi secara parsial.

7. Adanya rekursivitas. Semua anak panah mempunyai satu arah, tidak boleh terjadi pemutaran kembali (*looping*).
8. Spesifikasi model sangat diperlukan untuk menginterpretasikan koefisien-koefisien jalur. Kesalahan spesifikasi terjadi ketika variabel penyebab yang signifikan dikeluarkan dari model. Semua efisien jalur akan merefleksikan varian baru bersama dengan semua variabel yang tidak diukur dan tidak akan dapat diinterpretasikan secara tepat dalam kaitannya dengan akibat langsung dan tidak langsung.
9. Terdapat masukan korelasi yang sesuai. Artinya, jika kita menggunakan matriks korelasi sebagai masukan maka korelasi *pearson* digunakan untuk dua variabel berskala interval; korelasi *polychoric* untuk dua variabel berskala ordinal; *tetrachoric* untuk dua variabel dikotomi (berskala nominal); *polyserial* untuk satu variabel interval dan lainnya ordinal; serta biserial untuk satu variabel berskala interval dan lainnya nominal.
10. Terdapat ukuran sampel yang memadai. Untuk memperoleh hasil yang maksimal sebaiknya digunakan sampel di atas 100.
11. Sampel sama dibutuhkan perhitungan regresi dalam model jalur.

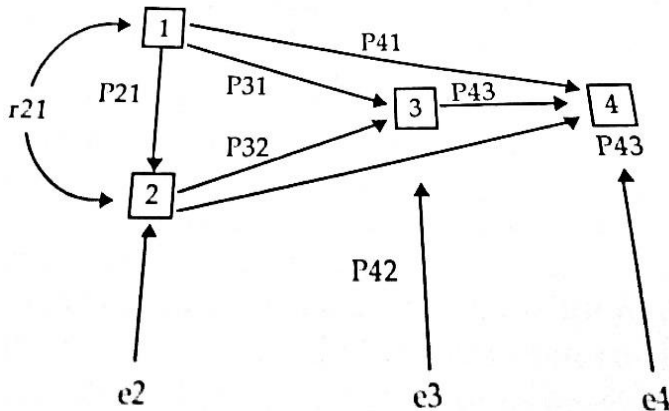
C. Manfaat *Path Analysis* (Analisis Jalur)

Sugiyono (2015) menyebutkan, melalui analisis jalur akan dapat ditemukan jalur mana yang paling tepat suatu variabel independen menuju variabel dependen yang terakhir. Riduwan dan Kuncoro (2014) memaparkan manfaat dari penggunaan *path analysis* adalah untuk:

1. Penjelasan (*explanation*) terhadap fenomena yang dipelajari atau permasalahan yang diteliti.
2. Prediksi nilai variabel terikat (Y) berdasarkan nilai variabel bebas (X) dan prediksi dengan *path analysis* (analisis jalur) ini bersifat kualitatif.
3. Faktor determinan, yaitu penentuan variabel bebas (X) mana yang berpengaruh dominan terhadap variabel terikat (Y) juga dapat digunakan untuk menelusuri mekanisme (jalur-jalur) pengaruh variabel bebas (X) terhadap variabel terikat (Y).
4. Pengujian model menggunakan *theory trimming*, baik untuk uji reliabilitas (uji keajegan) konsep yang sudah ada ataupun uji pengembangan konsep baru.

D. Konsep dan Istilah *Path Analysis* (Analisis Jalur)

Pada *path analysis*, dikenal beberapa konsep dan istilah dasar. Dengan gambar model berikut ini akan diterangkan konsep dan istilah dasar tersebut.



Gambar 2.1 Model *path analysis* (analisis jalur).

1. Model Jalur

Model jalur adalah suatu diagram yang menghubungkan antara variabel bebas, perantara, dan tergantung. Pola hubungannya menggunakan anak panah. Anak panah tunggal menunjukkan hubungan sebab akibat antara variabel eksogen dengan satu variabel tergantung atau lebih. Anak panah juga menghubungkan kesalahan (*variabel residue*) dengan semua variabel endogen masing-masing. Anak panah ganda menunjukkan korelasi antara pasangan variabel-variabel eksogen.

2. Jalur Penyebab untuk Suatu Variabel yang Diberikan

Meliputi jalur-jalur arah dari anak panah menuju ke variabel tersebut dan jalur-jalur korelasi dari semua variabel endogen yang dikorelasikan dengan variabel-variabel lain yang mempunyai anak panah-anak panah menuju ke variabel yang sudah ada tersebut.

3. Variabel Eksogen

Variabel eksogen dalam suatu model jalur adalah semua variabel yang tidak ada penyebab-penyebab eksplisitnya atau dalam diagram tidak ada anak panah yang menuju ke arahnya, selain pada bagian kesalahan pengukuran. Jika antara variabel eksogen dikorelasikan maka korelasi tersebut ditunjukkan dengan anak panah berkepala dua yang menghubungkan variabel-variabel tersebut.

4. Variabel Endogen

Variabel endogen adalah variabel yang mempunyai anak panah-anak panah menuju ke arah variabel tersebut. Variabel yang termasuk di dalamnya mencakup semua variabel perantara dan tergantung. Variabel perantara endogen mempunyai anak panah yang menuju ke arahnya dan dari arah variabel tersebut dalam suatu model. Adapun variabel tergantung hanya mempunyai anak panah yang menuju ke arahnya.

5. Koefisien Jalur atau Pembobotan Jalur

Koefisien jalur merupakan koefisien regresi standar atau disebut dengan ‘beta’, yang menunjukkan pengaruh langsung dari suatu variabel bebas terhadap variabel tergantung dalam suatu model tertentu. Oleh karena itu, jika suatu model mempunyai dua atau lebih variabel-variabel penyebab maka koefisien-koefisien regresi parsial yang mengukur besarnya pengaruh satu variabel terhadap variabel lain dalam suatu model jalur tertentu yang mengontrol dua variabel lain sebelumnya dengan menggunakan data yang sudah distandarkan atau matriks korelasi sebagai masukan.

6. Variabel-Variabel Eksogen yang Dikorelasikan

Jika semua variabel eksogen dikorelasikan maka sebagai penanda hubungannya ialah anak panah dengan dua kepala yang dihubungkan di antara variabel-variabel dengan koefisien korelasinya.

7. Istilah Gangguan

Gangguan atau *residue* mencerminkan adanya varian yang tidak dapat diterangkan atau pengaruh dari semua variabel yang tidak terukur ditambah dengan kesalahan pengukuran.

8. Aturan Multiplikasi Jalur

Nilai dari suatu jalur gabungan adalah hasil semua koefisien jalurnya.

9. Dekomposisi Pengaruh

Koefisien jalur dapat digunakan untuk mengurangi korelasi dalam suatu model ke dalam pengaruh langsung dan tidak langsung, yang berhubungan dengan jalur langsung dan tidak langsung yang direfleksikan dengan anak panah-anak panah dalam suatu model tertentu. Ini didasarkan pada aturan bahwa dalam suatu sistem linier, pengaruh penyebab total suatu variabel “I” terhadap variabel “j” adalah jumlah semua nilai jalur dari “i” ke “j”.

10. Signifikansi dan Model Keselarasan dalam Jalur

Untuk melakukan pengujian koefisien jalur secara individual, dapat menggunakan t standar atau pengujian F dan angka-angka keluaran regresi, sedangkan untuk melakukan pengujian model dengan semua jalurnya dapat menggunakan uji keselarasan dari program. Jika suatu model sudah benar, di antaranya mencakup semua variabel yang tidak sesuai, maka jumlah nilai-nilai jalur dari “i” ke “j” akan sama dengan koefisien regresi untuk “j” yang akan diprediksi dan didasarkan pada “i”, yaitu data yang dengan koefisien korelasi. Dengan demikian, jumlah semua koefisien (standar) akan sama dengan koefisien korelasi.

11. Anak Panah dengan Satu Kepala dan Dua Kepala

Jika ingin menggambarkan penyebab dengan menggunakan anak satu kepala yang menunjukkan satu arah. Adapun untuk menggambarkan korelasi, menggunakan anak panah yang melengkung dengan dua kepala yang menunjukkan dua arah. Ada kalanya hubungan sebab akibat menghasilkan angka negatif. Untuk menggambarkan hasil yang negatif, digunakan garis putus-putus.

12. Pola Hubungan

Dalam analisis jalur, tidak digunakan istilah variabel bebas ataupun tergantung. Sebagai gantinya, digunakan istilah variabel eksogen dan endogen.

13. Model *Recursive*

Model penyebab mempunyai satu arah. Tidak ada arah membalik (*feed back loop*) dan tidak ada pengaruh sebab akibat (*reciprocal*). Dalam model ini, satu variabel tidak dapat berfungsi sebagai penyebab dan akibat dalam waktu yang bersamaan.

14. Model *Non-Recursive*

Model penyebab dengan disertai arah yang membalik (*feed back loop*) atau adanya pengaruh sebab akibat (*reciprocal*).

15. *Direct Effect*

Pengaruh langsung yang dapat dilihat dari koefisien jalur dari satu variabel ke variabel lainnya.

16. *Indirect Effect*

Urutan jalur melalui satu atau lebih variabel perantara.

E. Perhitungan *Path Analysis* (Analisis Jalur)

Riduwan dan Kuncoro (2014: 116) memaparkan langkah-langkah *path analysis* (analisis jalur) sebagai berikut.

1. Merumuskan Hipotesis dan Persamaan Struktural

Perumusan persamaan ini sesuai dengan hipotesis yang diajukan.

Contoh struktur:

$$Y = \rho_{yx1} X_1 + \rho_{yx2} X_2 + \rho_y \varepsilon_1$$

- Menghitung koefisien jalur yang didasarkan pada koefisien regresi.
- Gambarkan diagram jalur lengkap, tentukan sub-sub strukturnya dan rumuskan persamaan strukturalnya yang sesuai dengan hipotesis yang diajukan.

Hipotesis: Naik turunnya variabel endogen (Y) dipengaruhi secara signifikan oleh variabel eksogen (X_1 dan X_2).

- Menghitung koefisien regresi untuk struktur yang telah dirumuskan.

Hitung koefisien regresi untuk struktur yang telah dirumuskan.

Persamaan regresi ganda: $Y = a + b_1 x_1 + b_2 x_2 + \varepsilon_1$

Pada dasarnya, koefisien jalur (*path*) adalah koefisien regresi yang distandarkan, yaitu koefisien regresi yang dihitung dari basis data yang telah

diset dalam angka baku atau *Z-score* (data yang diset dengan nilai rata-rata = 0 dan standar deviasi = 1). Koefisien jalur yang distandarkan (*standardized path coefficient*) digunakan untuk menjelaskan besarnya pengaruh (bukan memprediksi) variabel bebas (eksogen) terhadap variabel lain yang diberlakukan sebagai variabel terikat (endogen).

Khusus untuk program SPSS menu analisis regresi, koefisien *path* ditunjukkan oleh *output* yang disebut *coefficient* yang dinyatakan sebagai *standarized coefficient* atau dikenal dengan **beta**. Jika ada diagram jalur sederhana mengandung satu unsur hubungan antara variabel eksogen dengan variabel endogen maka koefisien *path*-nya adalah sama dengan koefisien korelasi *r* sederhana.

2. Menghitung Koefisien Jalur secara Simultan (Keseluruhan)

Uji secara keseluruhan hipotesis statistik dirumuskan sebagai berikut.

$$H_a : \rho_{yx1} = \rho_{yx2} = \dots = \rho_{yxk} \neq 0$$

$$H_0 : \rho_{yx1} = \rho_{yx2} = \dots = \rho_{yxk} = 0$$

- a. Kaidah pengujian signifikansi secara manual yaitu menggunakan tabel F.

$$F = \frac{(n-k-1)R_{yxk}^2}{k(1-R_{yxk}^2)}$$

Jika $F_{hitung} \geq F_{tabel}$, maka tolak H_0 artinya signifikan dan jika $F_{hitung} \leq F_{tabel}$, maka tolak H_0 tidak signifikan dengan taraf signifikansi (α) = 0,05

Carilah nilai F tabel dengan menggunakan rumus:

$$F_{tabel} = F_{\{(1-\alpha)(dk=k), (dk=n-k-1)\}} \text{ atau } F_{\{(v1=k), (v2=n-k-1)\}}$$

Cara mencari F_{tabel} : nilai $(dk=k)$ atau v_1 disebut nilai pembilang, nilai $(dk=n-k-1)$ atau v_2 disebut nilai penyebut.

- b. Kaidah pengujian signifikansi: program SPSS.

- 1) Jika nilai probabilitas 0,05 *lebih kecil atau sama dengan* nilai probabilitas Sig. atau $[0,05 \leq \text{Sig.}]$, maka H_0 diterima H_a ditolak, artinya tidak signifikan.
- 2) Jika nilai probabilitas 0,05 *lebih besar atau sama dengan* nilai probabilitas Sig. atau $[0,05 \geq \text{Sig.}]$, maka H_0 ditolak H_a diterima, artinya signifikan.

3. Menghitung Koefisien Jalur secara Sndividu

Hipotesis penelitian yang akan diuji dirumuskan menjadi hipotesis statistik berikut.

$$H_a : \rho_{yx1} > 0$$

$$H_0 : \rho_{yx} = 0$$

Secara individual uji statistik yang digunakan adalah uji t yang dihitung dengan rumus:

$$t_k = \frac{\rho_k}{se_{\rho_k}} ; (dk = n-k-1)$$

Keterangan:

Statistik $se_{\rho_{x1}}$ diperoleh dari hasil komputasi SPSS untuk analisis regresi setelah data ordinal ditransformasikan ke internal.

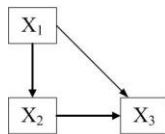
Selanjutnya, untuk mengetahui signifikansi analisis jalur bandingkan antara nilai probabilitas 0,05 dengan nilai probabilitas Sig. dengan dasar pengambilan keputusan sebagai berikut.

- 1) Jika nilai probabilitas 0,05 *lebih kecil atau sama dengan* nilai probabilitas Sig. atau $[0,05 \leq \text{Sig.}]$, maka H_0 diterima H_a ditolak, artinya tidak signifikan.
- 2) Jika nilai probabilitas 0,05 *lebih besar atau sama dengan* nilai probabilitas Sig. atau $[0,05 \geq \text{Sig.}]$, maka H_0 ditolak H_a diterima, artinya signifikan.

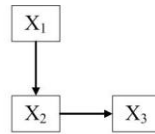
4. Menyimpulkan

F. Model *Path Analysis* (Analisis Jalur)

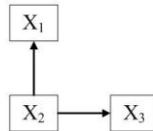
Sebelum peneliti menggunakan analisis jalur dalam penelitiannya, peneliti harus menyusun model hubungan antar variabel yang disebut **diagram jalur**. Diagram jalur tersebut disusun berdasarkan kerangka berpikir yang dikembangkan dari teori yang digunakan untuk penelitian. Sugiyono (2015) menyatakan bahwa diagram jalur dapat dibagi menjadi diagram jalur sederhana dan yang lebih kompleks. Pada diagram jalur sederhana, hanya sedikit variabel yang dilibatkan, sedangkan pada diagram jalur yang lebih kompleks melibatkan variabel yang lebih banyak sehingga banyak jalur yang dianalisis.



Gambar 2.2a

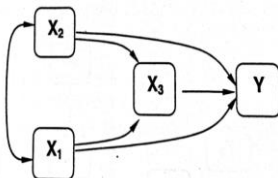


Gambar 2.2b

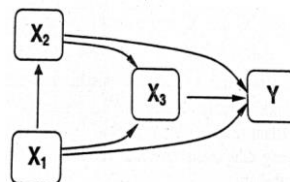


Gambar 2.2c

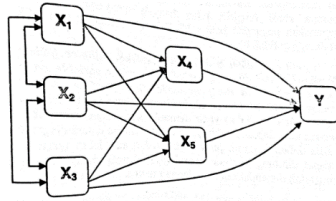
Gambar 2.2 Diagram jalur sederhana.



Gambar 2.3.a



Gambar 2.3.b



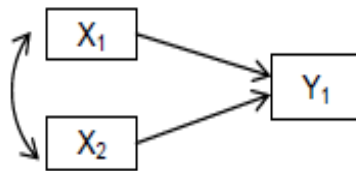
Gambar 2.3.c

Gambar 2.3 Diagram jalur kompleks.

Berikut ini model *path analysis* (analisis jalur) menurut Djudin (2013); Riduwan dan Kuncoro (2014); Sugiyono (2015); serta Raltan dan Renhard, (2014).

1. Model Regresi Berganda

Model pertama ini sebenarnya merupakan pengembangan regresi berganda dengan menggunakan dua variabel eksogen, yaitu X_1 dan X_2 dengan satu variabel endogen Y (Raltan dan Renhard, 2014). Model regresi berganda digambarkan oleh Gambar 2.4.



Gambar 2.4 Model Regresi Berganda.

Riduwan dan Kuncoro (2014: 4) menyatakan bahwa meskipun model regresi dan model *path analysis* sama-sama merupakan analisis regresi, tetapi penggunaan kedua model tersebut berbeda. Hal ini diperjelas oleh Rasyid dalam Riduwan dan Kuncoro (2014: 4) bahwa pola hubungan seperti apa

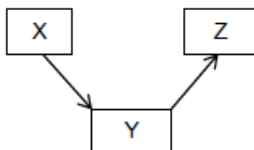
yang ingin kita ungkapkan, apakah pola hubungan yang bisa digunakan untuk meramalkan atau menduga nilai sebuah variabel-respons Y atas dasar nilai tertentu pada beberapa variabel-prediktor $X_1; X_2; \dots, X_k$ terhadap sebuah variabel akibat Y , baik pengaruh langsung secara individual maupun secara bersamaan. Telaah statistika menyatakan bahwa untuk peramalan/pendugaan nilai Y atas dasar nilai $X_1; X_2; \dots, X_k$, pola hubungan yang sesuai adalah pola hubungan yang mengikuti model regresi, sedangkan untuk tujuan hubungan sebab akibat pola yang tepat adalah model struktural. Secara matematik analisis jalur mengikuti pola struktural.

Contoh:

Dalam kasus pengaruh antara kecerdasan dan sikap siswa terhadap hasil belajar maka X_1 adalah variabel kecerdasan, X_2 adalah variabel sikap siswa, sementara Y adalah variabel hasil belajar dalam terminologi analisis jalur.

2. Model Mediasi

Model kedua adalah model mediasi atau perantara variabel Y memodifikasi pengaruh variabel X terhadap variabel Z (Raltan dan Renhard, 2014). Model mediasi digambarkan oleh Gambar 2.5.



Gambar 2.5 Model Mediasi.

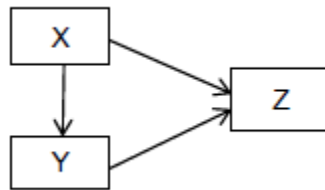
Contoh:

Hasil belajar tidak secara langsung dipengaruhi oleh kecerdasan seseorang, tetapi kepandaian seseorang akan menciptakan efisiensi belajar dan efisiensi belajarnya yang akan menentukan hasil belajar. Jika diterapkan dalam model kedua ini maka variabel X adalah kecerdasan, variabel Y adalah

variabel hasil belajar, dan variabel Z adalah variabel efisiensi belajar. Variabel kecerdasan memengaruhi variabel hasil belajar melalui variabel efisiensi belajar.

3. Model Kombinasi Regresi Berganda dan Korelasi

Model ketiga ini merupakan kombinasi antara model pertama dan kedua, yaitu variabel X berpengaruh terhadap variabel Z secara langsung dan secara tidak langsung memengaruhi variabel Z melalui variabel Y (Raltan dan Renhard, 2014). Model kombinasi digambarkan seperti pada Gambar 2.6.



Gambar 2.6 Model Kombinasi.

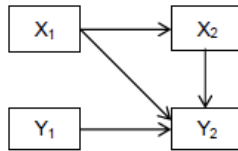
Contoh:

Kualitas layanan yang diberikan suatu perusahaan dipengaruhi oleh kinerja pegawai yang pada akhirnya akan memengaruhi tingkat kepuasan pelanggan. Dalam kasus ini, variabel X adalah kinerja pegawai, variabel Y adalah kualitas layanan, dan variabel Z adalah kepuasan pelanggan. Kinerja pegawai secara langsung memengaruhi kepuasan pelanggan, demikian pula kinerja pegawai akan memengaruhi kualitas layanan yang kemudian akan berpengaruh terhadap kepuasan pelanggan.

4. Model Kompleks

Model keempat ini merupakan model yang lebih kompleks, yaitu variabel X1 secara langsung memengaruhi Y2 dan melalui variabel X2 secara tidak langsung memengaruhi Y2, sedangkan variabel Y2 juga dipengaruhi

oleh variabel Y1 (Raltan dan Renhard, 2014). Model kompleks digambarkan sebagai oleh Gambar 2.7.



Gambar 2.7 Model Kompleks.

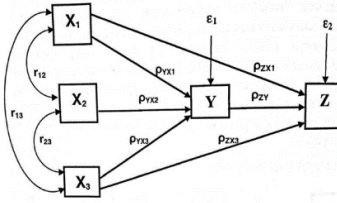
Contoh:

Contoh kasus ini diambil dari hasil penelitian Sawyer dkk. dalam masalah psikologi. Kasusnya sebagai berikut.

X1 adalah psikopatologi tahap pertama seorang ibu yang akan menjadi penentu terhadap patologi tahap kedua ibu yang bersangkutan dalam hal ini adalah variabel X2; dan mirip dengan kejadian tersebut patologi tahap pertama anaknya atau variabel Y1 akan memengaruhi patologi tahap kedua anak tersebut atau variabel Y2. Selanjutnya patologi anak tahap kedua atau Y2 juga dipengaruhi oleh patologi ibu tahap pertama, yaitu (jalur antara X1 dan Y2) dan tahap kedua, yaitu (jalur antara X2 dan Y2).

5. Model Persamaan Struktural

Persamaan struktural atau juga disebut model struktural adalah apabila setiap variabel terikat/endogen (Y) secara unik keadaannya ditentukan oleh seperangkat variabel terikat/eksogen (X) (Riduwan dan Kuncoro, 2014). Selanjutnya gambar yang meragakan struktur hubungan kausal antar variabel disebut diagram jalur (*path diagram*). Jadi, persamaan $Y = F(X1; X2; X3)$ dan $Z = F(X1; X3; Y)$ merupakan persamaan struktural karena setiap persamaan menjelaskan hubungan kausal, yaitu variabel eksogen X1; X2; dan X3 terhadap variabel endogen Y dan Z. Diagram jalur digambarkan sebagai berikut.

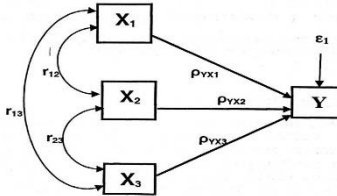


Gambar 2.8 Diagram Jalur Hubungan X1, X2, X3 dan Y dan Z.

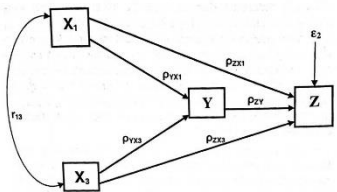
Persamaan struktural untuk diagram jalur, yaitu

$$Y = \rho_{yx1} X_1 + \rho_{yx2} X_2 + \rho_{yx3} X_3 + \epsilon_1$$

$$Z = \rho_{zx1} X_1 + \rho_{zy} Y + \epsilon_2$$



Gambar 2.9 Hubungan Kausal X1, X2, X3 ke Y.



Gambar 2.10 Hubungan Kausal X1, X2, dan Y ke Z.

Keterangan:

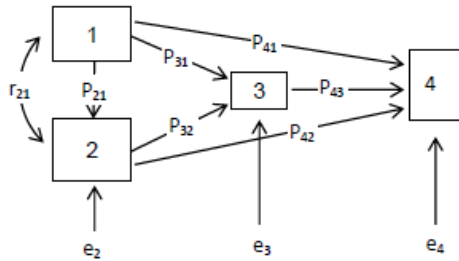
P_{ik} merupakan koefisien jalur (*path coefficient*) untuk setiap variabel eksogen k. Koefisien jalur menunjukkan pengaruh langsung variabel eksogen k terhadap variabel endogen I, sedangkan ϵ_i menunjukkan variabel atau faktor

residual yang fungsinya menjelaskan pengaruh variabel lainnya yang belum teridentifikasi oleh teori, atau muncul sebagai akibat dari kekeliruan pengukuran variabel. Pada sebuah diagram jalur, tanda panah berujung ganda (\leftrightarrow) menunjukkan hubungan korelasi dan tanda panah satu arah (\rightarrow) menunjukkan hubungan kausal atau pengaruh langsung dari variabel eksogen (X) terhadap variabel endogen (Y).

Jadi, secara sistematis *path analysis* mengikuti pola model struktural sehingga langkah awal untuk mengerjakan atau penerapan model *path analysis* adalah dengan merumuskan persamaan struktural dan diagram jalur yang berdasarkan kajian teori tertentu. Solimun dalam Riduwan dan Kuncoro (2014) menyatakan bahwa informasi diberikan apabila tujuan penelitian ingin mendapatkan model untuk kepentingan prediksi maka model yang tepat digunakan adalah model struktural. Model ini mirip dengan *path analysis* yang membedakan adalah kalau di dalam *path analysis* data yang dianalisis adalah data baku (*standarized*), sedangkan di dalam model struktural menggunakan data mentah (*raw data*). Dengan demikian, hasil analisis model struktural kurang tepat jika disajikan dalam bentuk *path diagram* dan lebih cocok disajikan dalam sistem persamaan.

Path analysis dikenal luas oleh para ahli ilmu sosial terutama sosiologi. Namun, sekarang *path analysis* bukanlah monopoli para sosiolog lagi. *Path analysis* telah menjadi model analisis para ilmuwan sosial lainnya. Bahkan, pada tahun 1970 Karl G. Joreskog dan Dag Sorbom telah mengembangkan model *path analysis* menjadi model yang sekarang dikenal dengan SEM (Structural Equation Modelling).

Jika dilihat sisi pandang sebab akibat, ada dua tipe model jalur, yaitu rekursif dan nonrekursif. Model rekursif ialah jika semua anak panah menuju satu arah seperti Gambar 2.11.



Gambar 2.11 Model Rekursif.

Model rekursif pada Gambar 2.11 dapat diterangkan sebagai berikut.

- Anak panah menuju satu arah, yaitu dari 1 ke 2, 3, dan 4; dari 2 ke 3 dan dari 3 menuju ke 4. Tidak ada arah yang terbalik, misalnya dari 4 ke 1.
- Hanya terdapat satu variabel eksogen, yaitu 1 dan tiga variabel endogen, yaitu 2, 3, dan 4. Masing-masing variabel endogen diterangkan oleh variabel 1 dan eror (e_2 , e_3 , dan e_4).
- Satu variabel endogen dapat menjadi penyebab variabel endogen lainnya, tetapi bukan ke variabel eksogen.
- Model nonrekursif terjadi jika arah anak panah tidak searah atau terjadi arah yang terbalik (*looping*), misalnya dari 4 ke 3 atau dari 3 ke 1 dan 2, atau bersifat sebab akibat (*reciprocal cause*).

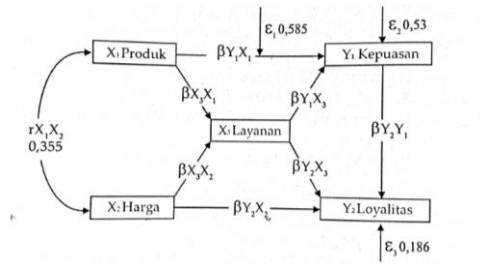
Dari berbagai model sebelumnya masih bisa dikembangkan berbagai model yang lebih kompleks tergantung dari kepentingan penelitian yang ditopang oleh teori yang relevan. Model-model analisis jalur sangat membantu melakukan analisis yang lebih teliti dan kompleks tanpa mengorbankan variabel yang benar-benar berperan dalam memengaruhi variabel endogen walaupun sifat pengaruhnya tidak langsung. Kondisi ini sangat cocok dalam mendalami kehidupan manusia, mengingat faktor sebab akibat dalam kehidupan manusia itu sangat kompleks dan tidak mungkin

faktor yang memengaruhinya itu diabaikan begitu saja.

Agar mempermudah dalam memahami analisis jalur, lazimnya digunakan model-model jalur berikut.

1. Model Persamaan Satu Jalur

Model persamaan satu jalur merupakan hubungan sebenarnya sama dengan regresi berganda, yaitu variabel bebas terdiri lebih dari satu variabel dan variabel tergangungnya hanya satu.

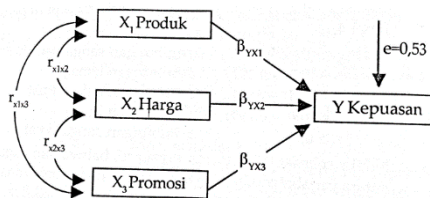


Gambar 2.12 Model Diagram Satu Jalur.

2. Model Persamaan Dua Jalur

Model ini terdiri dari tiga variabel bebas dan mempunyai dua variabel tergantung. Gambar 2.13 menggambarkan contoh model persamaan dua jalur.

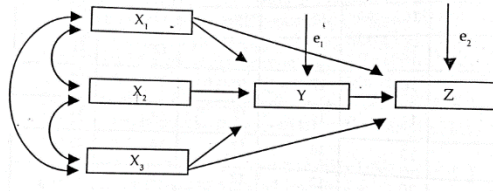
Contoh kasus: variabel X_1 adalah produk, variabel X_2 adalah harga, variabel X_3 adalah promosi, variabel Y adalah keputusan sebagai variabel *intervening*, dan Z adalah loyalitas.



Gambar 2.13 Model Diagram Dua Jalur.

3. Model Persamaan Tiga Jalur

Model ini terdiri dari tiga variabel bebas, salah satu variabel bebas menjadi variabel perantara dan mempunyai dua variabel tergantung.

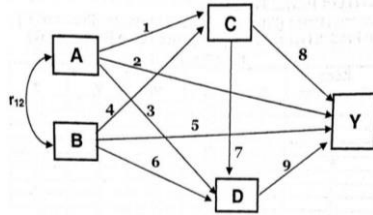


Gambar 2.14 Model Diagram Tiga Jalur.

4. Model Dekomposisi

Model dekomposisi adalah model yang menekankan pada pengaruh yang bersifat kausalitas antar variabel, baik pengaruh langsung ataupun tidak langsung dalam kerangka analisis jalur, sedangkan hubungan yang sifatnya nonkausalitas atau hubungan korelasional yang terjadi antarvariabel eksogen tidak termasuk dalam perhitungan ini. Perhitungan menggunakan analisis jalur dengan menggunakan model dekomposisi pengaruh kausal antar variabel dapat dibedakan menjadi tiga, yaitu

- Direct causal effects* (pengaruh kausal langsung) adalah pengaruh satu variabel eksogen terhadap variabel endogen yang terjadi tanpa melalui variabel endogen lain.
- Indirect causal effects* (pengaruh kausal tidak langsung) adalah pengaruh satu variabel eksogen terhadap variabel endogen yang terjadi melalui variabel endogen lain terdapat dalam satu model kausalitas yang sedang dianalisis.
- Total causal effects* (pengaruh kausal total) adalah jumlah dari pengaruh kausal langsung dan pengaruh kausal tidak langsung.



Gambar 2.15 Model Diagram Dekomposisi.

5. Model *Trimming*

Model *trimming* adalah model yang digunakan untuk memperbaiki suatu model struktur analisis jalur dengan cara mengeluarkan dari model variabel eksogen yang koefisien jalur diuji secara keseluruhan apabila ternyata ada variabel yang tidak signifikan. Walaupun ada satu, dua, atau lebih variabel yang tidak signifikan, perlu memperbaiki model struktur analisis jalur yang telah dihipotesiskan.

Cara menghitung model *trimming* adalah dengan menghitung ulang koefisien jalur tanpa menyertakan variabel eksogen yang koefisien jalurnya tidak signifikan. Menurut Riduwan dan Kuncoro (2014: 128), langkah-langkah pengujian *path analysis* model *trimming* sebagai berikut.

- a. Merumuskan persamaan struktural.
- b. Menghitung koefisien jalur yang didasarkan pada koefisien regresi.
 - 1) Gambarkan diagram jalur lengkap.
 - 2) Menghitung koefisien regresi untuk setiap sub struktur yang telah dirumuskan.
- c. Menghitung koefisien jalur secara simultan (keseluruhan).
- d. Menghitung secara individual.
- e. Menguji kesesuaian antar model analisis jalur.
- f. Merangkum ke dalam tabel.

g. Memaknai dan menyimpulkan.

G. Contoh Kasus *Path Analysis* dengan SPSS

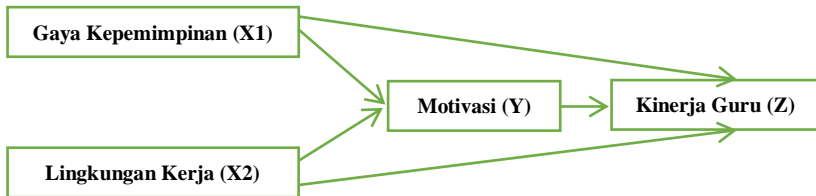
1. Contoh Kasus Uji Regresi Variabel *Intervening* (Analisis Jalur)

a. Konsep Dasar *Path Analysis*

Istilah lain dari uji regresi dengan variabel *intervening* adalah *path analysis* atau analisis jalur. *Path analysis* merupakan bagian lebih lanjut dari analisis regresi. Analisis regresi biasa digunakan untuk menguji apakah ada pengaruh langsung yang diberikan oleh variabel bebas terhadap variabel terikat. Sementara itu, analisis jalur tidak hanya menguji pengaruh langsung saja, tetapi juga menjelaskan tentang pengaruh tidak langsung yang diberikan variabel bebas melalui variabel *intervening* terhadap variabel terikat.

b. Contoh Kasus *Path Analysis*

Terdapat konsep dasar analisis jalur, berikut pemaparan beserta contoh sederhana agar kita mudah memahami. Di bawah ini adalah gambar kerangka pemikiran untuk analisis jalur.



Gambar 2.16 Kerangka pemikiran untuk analisis jalur.

Berdasarkan gambaran di atas, dapat kita rumuskan sebuah hipotesis umum yang akan diajukan dalam analisis jalur, yakni “Pengaruh gaya kepemimpinan (X1) dan lingkungan kerja (X2) terhadap motivasi (Y) serta dampaknya terhadap kinerja guru (Z)”. Sementara untuk hipotesis yang akan

diuji, antara lain

- 1) pengaruh X1 dan X2 terhadap Y;
- 2) pengaruh X1, X2 dan Y terhadap Z; dan
- 3) pengaruh X1 dan X2 melalui Y terhadap Z.

c. Langkah-Langkah Uji Analisis Jalur dengan SPSS

Untuk menguji hipotesis di atas, kita akan melakukan dua kali model regresi.

Berikut adalah data tabulasi penelitian yang akan digunakan.

X1	X2	Y	Z
40.00	55.00	50.00	50.00
41.00	50.00	43.00	40.00
42.00	38.00	31.00	32.00
41.00	50.00	40.00	40.00
45.00	55.00	41.00	44.00
32.00	44.00	31.00	33.00
40.00	35.00	39.00	45.00
36.00	35.00	42.00	39.00
33.00	35.00	35.00	32.00
37.00	48.00	47.00	48.00
39.00	54.00	49.00	47.00
44.00	55.00	49.00	50.00
39.00	51.00	50.00	50.00

- 1) Buka program SPSS versi 25 atau versi berapa pun yang Anda miliki kemudian klik Variable View dan isi kolom-kolom yang tersedia sebagaimana petunjuk di bawah ini.
 - a) Kolom Name di isi dengan X1, X2, Y dan Z.
 - b) Kolom Label untuk X1 tulis “Gaya Kepemimpinan (X1)”, X2 tulis “Lingkungan Kerja (X2)”, Y tulis “Motivasi (Y)”, dan Z tulis “Kinerja Guru (Z)”.

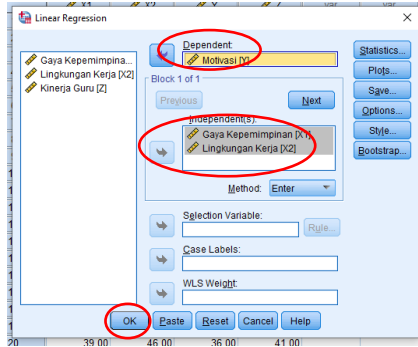
- c) Kolom Type, Width, Decimal, Value, Missing, Coloumns, Align, Meansure dibiarkan tetap *default* dan tidak perlu diubah karena sudah otomatis menyesuaikan jenis data yang dipakai.

Name	Type	Width	Decimals	Label	Values	Missing	Column:
X1	Numeric	8	2	Gaya Kepemim...	None	None	8
X2	Numeric	8	2	Lingkungan Kerja	None	None	8
Y	Numeric	8	2	Motivasi	None	None	8
Z	Numeric	8	2	Kinerja Guru	None	None	8

- 2) Langkah selanjutnya klik Data View kemudian isikan data sesuai nama variabel penelitian.

	X1	X2	Y	Z
1	40.00	55.00	50.00	50.00
2	41.00	50.00	43.00	40.00
3	42.00	38.00	31.00	32.00
4	41.00	50.00	40.00	40.00
5	45.00	55.00	41.00	44.00
6	32.00	44.00	31.00	33.00
7	40.00	35.00	39.00	45.00
8	36.00	35.00	42.00	39.00
9	33.00	35.00	35.00	32.00
10	37.00	48.00	47.00	48.00
11	39.00	54.00	49.00	47.00
12	44.00	55.00	49.00	50.00
13	39.00	51.00	50.00	50.00
14	34.00	39.00	40.00	37.00
15	34.00	32.00	31.00	32.00

- 3) Jika datanya sudah tertulis benar, klik menu Analyze lalu klik Regression dan pilih Linear sehingga muncul kotak dialog Linear Regression. Selanjutnya, klik Gaya Kepemimpinan (X1) dan Lingkungan Kerja (X2), masukkan pada kotak Independent(s). Kemudian klik Motivasi (Y) pada kotak Dependent.



- 4) Jika sudah sesuai dengan langkah-langkah di atas, klik Ok. Maka akan muncul Output SPSS (bisa disebut dengan *output regresi model 1*).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.651 ^a	.423	.413	3.85901

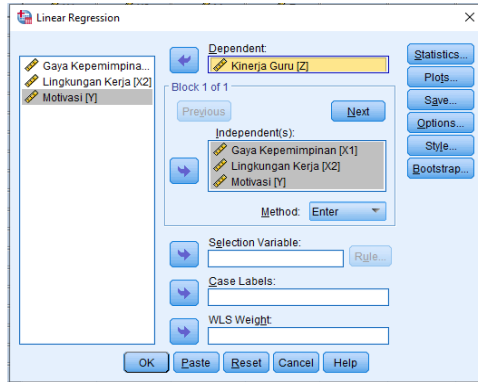
a. Predictors: (Constant), Lingkungan Kerja, Gaya Kepemimpinan

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	8.810	3.994		2.206	.029
	Gaya Kepemimpinan	.461	.114	.336	4.041	.000
	Lingkungan Kerja	.324	.067	.403	4.845	.000

a. Dependent Variable: Motivasi

- 5) Pembuatan regresi model II dilakukan dengan cara yang sama, yakni klik menu Analyze lalu klik Regression dan klik Linear. Setelah muncul kotak dialog Linear Regression, klik Reset (untuk mengeluarkan semua variabel) lalu klik Gaya Kepemimpinan (X1), Lingkungan Kerja (X2), dan Motivasi (Y) pada kotak Independent (S). Terakhir, klik Kinerja Guru (Z) pada kotak Dependent.



6) Selanjutnya, klik OK lalu akan keluar *output* SPSS.

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.064	3.245		.636	.526
	Gaya Kepemimpinan	.219	.097	.156	2.254	.026
	Lingkungan Kerja	.120	.059	.146	2.050	.043
	Motivasi	.626	.075	.612	8.335	.000

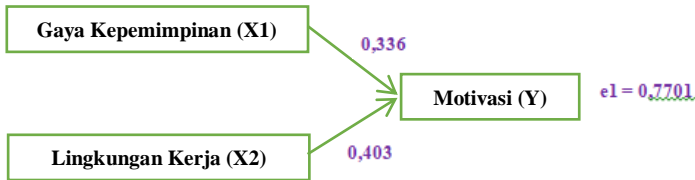
a. Dependent Variable: Kinerja Guru

d. Cara Memaknai Output Keluaran SPSS untuk Analisis Jalur

Agar lebih mudah memahami hasil di atas, pada bagian interpretasi hasil ini dibagi menjadi dua tahap interpretasi, yakni tahap menghitung koefisien jalur serta tahap uji hipotesis dan pembuatan kesimpulan.

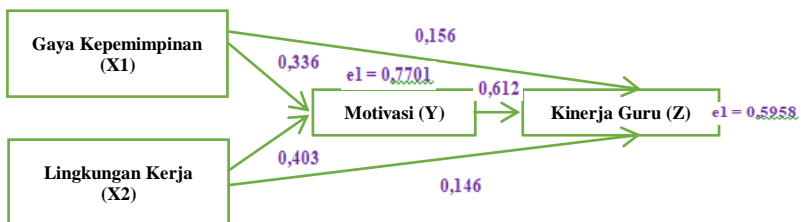
Koefisien Jalur Model I: Mengacu pada *output* regresi model I pada bagian tabel Coefficiens dapat diketahui bahwa nilai signifikansi dari kedua variable, yaitu $X1 = 0,000$ dan $X2 = 0,000$ lebih kecil dari 0,05. Hasil ini memberikan kesimpulan bahwa regresi model I, yakni variabel X1 dan X2 berpengaruh signifikan terhadap Y. Besar nilai R2 atau R-Square yang terdapat pada tabel Model Summary adalah sebesar 0,407. Hal ini

menunjukkan bahwa kontribusi atau sumbangan pengaruh X1 dan X2 terhadap Y adalah sebesar 40,7%, sementara sisanya 59,3% merupakan kontribusi dari variabel-variabel lain yang tidak dimasukkan dalam penelitian. Sementara itu, untuk nilai e_1 dapat dicari dengan rumus $e_1 = \sqrt{1 - 0,407} = 0,7701$. Dengan demikian, diperoleh diagram jalur di bawah ini.



Gambar 2.17 Diagram Jalur.

Koefisien model II: Berdasarkan *output* regresi model II pada bagian Coefficients, diketahui bahwa nilai signifikansi dari ketiga variable, yaitu X1 = 0,026, X2 = 0,043 dan Y = 0,000 lebih kecil dari 0,05. Hasil ini memberi kesimpulan bahwa regresi model II, yakni variabel X1, X2 dan Y berpengaruh signifikan terhadap Z. Besarnya nilai R² atau R-Square yang terdapat pada tabel Model Summary adalah sebesar 0,645. Hal ini menunjukkan bahwa kontribusi X1, X2 dan Y terhadap Z adalah sebesar 64,5%, sementara sisanya 35,5% merupakan kontribusi dari variabel-variabel lain yang tidak diteliti. Sementara untuk nilai $e_2 = \sqrt{1 - 0,645} = 0,5958$. Dengan demikian, diperoleh diagram jalur model struktur II sebagai berikut.



Gambar 2.18 Diagram Jalur Model Struktur II.

e. Tahap Uji Hipotesis dan Pembuatan Kesimpulan

- 1) Analisis pengaruh X1 terhadap Y: Dari analisis di atas diperoleh nilai signifikansi X1 sebesar $0,000 < 0,05$ sehingga dapat disimpulkan bahwa secara langsung terdapat pengaruh signifikan X1 terhadap Y.
- 2) Analisis pengaruh X2 terhadap Y: Dari analisis di atas diperoleh nilai signifikansi X2 sebesar $0,000 < 0,05$ sehingga dapat disimpulkan bahwa secara langsung terdapat pengaruh signifikan X2 terhadap Y.
- 3) Analisis pengaruh X1 terhadap Z: Dari analisis di atas diperoleh nilai signifikansi X1 sebesar $0,026 < 0,05$ sehingga dapat disimpulkan bahwa secara langsung terdapat pengaruh signifikan X1 terhadap Z.
- 4) Analisis Pengaruh X2 terhadap Z: Dari analisis di atas diperoleh nilai signifikansi X2 sebesar $0,043 < 0,05$ sehingga dapat disimpulkan bahwa secara langsung terdapat pengaruh signifikan X2 terhadap Z.
- 5) Analisis pengaruh Y terhadap Z: Dari analisis di atas diperoleh bahwa nilai signifikansi Y sebesar $0,000 < 0,05$ sehingga dapat disimpulkan bahwa secara langsung terdapat pengaruh signifikan Y terhadap Z.
- 6) Analisis pengaruh X1 melalui Y terhadap Z: Diketahui pengaruh langsung yang diberikan X1 terhadap Z sebesar 0,156, sedangkan pengaruh tidak langsung X1 melalui Y terhadap Z adalah perkalian antara nilai beta X1 terhadap Y dengan nilai beta Y, yaitu $0,336 \times 0,612 + 0,206$. Maka pengaruh total yang diberikan X1 terhadap Z adalah pengaruh langsung ditambah dengan pengaruh tidak langsung, yaitu $0,156 + 0,206 = 0,362$. Berdasarkan hasil perhitungan di atas, diketahui bahwa nilai pengaruh langsung sebesar 0,156 dan pengaruh tidak langsung sebesar 0,206, yang berarti bahwa nilai pengaruh tidak langsung lebih besar dibandingkan dengan nilai pengaruh langsung.

Hasil ini menunjukkan bahwa secara tidak langsung, X1 melalui Y mempunyai pengaruh signifikan terhadap Z.

- 7) Analisis pengaruh X2 melalui Y terhadap Z: Diketahui pengaruh langsung yang diberikan X2 terhadap Z sebesar 0,146, sedangkan pengaruh tidak langsung X2 melalui Y terhadap Z adalah pengaruh langsung ditambah dengan pengaruh langsung, yaitu $0,146 + 0,247 + 0,393$. Berdasarkan hasil perhitungan di atas diperoleh nilai pengaruh langsung sebesar 0,146 dan pengaruh tidak langsung sebesar 0,247 yang berarti bahwa nilai pengaruh tidak langsung lebih besar dari pada nilai pengaruh langsung. Hasil ini menunjukkan bahwa secara tidak langsung X2 melalui Y mempunyai pengaruh signifikan terhadap Z.

Dari pembahasan di atas, kita dapat menarik kesimpulan bahwa hipotesis yang berbunyi “Ada Pengaruh Gaya Kepemimpinan (X1) dan Lingkungan Kerja (X2) terhadap Motivasi (Y) serta Dampaknya terhadap Kinerja Guru (Z)”.

H. Langkah-Langkah Uji Analisis Jalur dengan Excel

1. Buka Microsoft Excel dan masukkan data tabulasi.

	A	B	C	D	E
1	VARIABEL				
2	No	X1	X2	Y	Z
3	1	40	55	50	50
4	2	41	50	43	40
5	3	42	38	31	32
6	4	41	50	40	40
7	5	45	55	41	44
8	6	32	44	31	33
9	7	40	35	39	45
10	8	36	35	42	39
11	9	33	35	35	32
12	10	37	48	47	48
13	11	39	54	49	47
14	12	44	55	49	50
15	13	39	51	50	50
16	14	34	39	40	37
17	15	34	32	31	32
18	16	40	46	44	45
...

2. Mencari nilai rata-rata dari semua variabel.
 - a. Klik pada kolom bawah variabel X1 lalu ketik dua kali **=AVERAGE**.
 - b. Blok data tabulasi X1 dari atas sampai bawah.
 - c. Terdapat **AVERAGE(B3:B117)** dan tekan Enter (terdapat hasil rata-rata dari X1).
 - d. Untuk mendapatkan keseluruhan rata-rata dari X2, Y dan Z, klik pada hasil rata-rata X1 kemudian blok ke samping sampai Z (maka akan muncul nilai rata-rata dari semua variabel).

A	B	C	D	E
100	36	41	31	37
101	37	41	44	42
102	44	55	50	50
103	41	42	42	45
104	34	39	37	37
105	36	40	39	36
106	44	53	50	50
107	41	47	43	38
108	44	54	47	47
109	40	43	40	48
110	43	42	49	48
111	40	51	49	48
112	35	38	31	33
113	45	54	50	50
114	35	46	36	36
115	32	37	35	33
	39,02609	43,94783	41,07826	41,62609

3. Mencari nilai standar deviasi.
 - a. Klik pada kolom di bawah hasil nilai rata-rata pada X1, kemudian tulis **=STDEV**.

- b. Blok data tabulasi keseluruhan X1 sehingga muncul **=STDEV(B3:B117)** lalu tekan **Enter** (maka terdapat hasil nilai standar pada kolom X1).
- c. Untuk mendapatkan hasil keseluruhan dari X2, Y dan Z, blok menyamping sampai kolom Z, lalu lepaskan (maka akan muncul hasil dari keseluruhan variabel).

104	34	39	37	37
105	36	40	39	36
106	44	53	50	50
107	41	47	43	38
108	44	54	47	47
109	40	43	40	48
110	43	42	49	48
111	40	51	49	48
112	35	38	31	33
113	45	54	50	50
114	35	46	36	36
115	32	37	35	33
	39,02609	43,94783	41,07826	41,62609
	3,669363	6,257408	5,036965	5,15456

4. Buat tabel seperti di bawah ini. Fungsi dari tabel tersebut untuk memindahkan masukan dari variabel. Terdapat huruf awalan **Z** yang dimaksudkan untuk menstandarisasi X1m maka menjadi ZX1.

VARIABEL							
X1	X2	Y	Z	ZX1	ZX2	ZY	ZZ
40	55	50	50				
41	50	43	40				
42	38	31	32				
41	50	40	40				
45	55	41	44				
32	44	31	33				
40	35	39	45				
36	35	42	39				

5. Untuk mendapatkan nilai yang dimasukkan pada kolom baru, klik pada kolom di bawah ZX1 dan tulis **=** lalu klik pada kolom angka di bawah X1 sehingga muncul **=(B3** kemudian dikurangi 39,02609 (nilai rata-rata

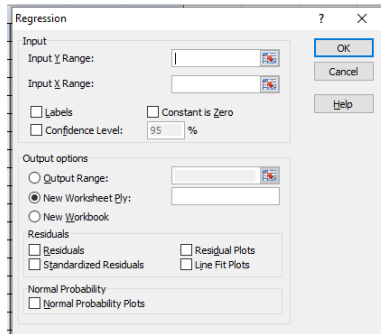
X1). Klik F4 kemudian dibagi nilai STDEV dan klik F4, tutup kurung dan Enter. Maka muncul hasil 29,36434. Langkah akhirnya adalah klik pada hasil tersebut dan blok sampai bawah sehingga muncul nilai keseluruhan untuk ZX1. Lakukan yang sama untuk mendapatkan hasil dari ZX2, Zy, dan ZZ.

The image shows two screenshots of an Excel spreadsheet. The top screenshot shows a formula bar with $= (B3-\$B\$118)/\$B\119 and a table with columns C, D, E, F, G, H, I, J. The table has a header row 'VARIABEL' and a sub-header row 'X2 Y Z'. Below this, there are 10 rows of data. The values in column G (labeled 'ZX1') are circled in red. The bottom screenshot shows a larger view of the same data table with columns D, E, F, G, H, I, J. The table has a header row 'Y Z' and a sub-header row 'ZX1 ZX2 ZY ZZ'. Below this, there are 10 rows of data. The values in column G (labeled 'ZX1') are circled in red.

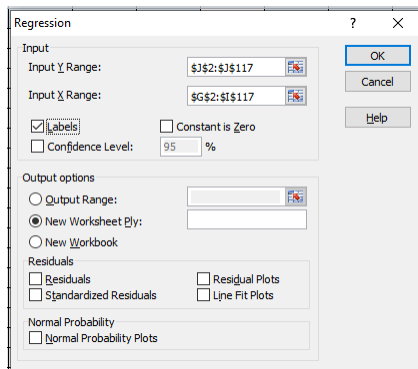
C	D	E	F	G	H	I	J
VARIABEL							
X2	Y	Z		ZX1	ZX2	ZY	ZZ
55	50	50		29,36434			
50	43	40		30,36434			
38	31	32		31,36434			
50	40	40		30,36434			
55	41	44		34,36434			
44	31	33		21,36434			
35	39	45		29,36434			
35	42	39		25,36434			
35	35	32		22,36434			
48	47	48		26,36434			
54	49	47		28,36434			

D	E	F	G	H	I	J
Y Z						
			ZX1	ZX2	ZY	ZZ
50	50		29,36434	47,97667	41,84464	41,92442
43	40		30,36434	42,97667	34,84464	31,92442
31	32		31,36434	30,97667	22,84464	23,92442
40	40		30,36434	42,97667	31,84464	31,92442
44	44		34,36434	47,97667	32,84464	35,92442
33	33		21,36434	36,97667	22,84464	24,92442
45	45		29,36434	27,97667	30,84464	36,92442
39	39		25,36434	27,97667	33,84464	30,92442
32	32		22,36434	27,97667	26,84464	23,92442
48	48		26,36434	40,97667	38,84464	39,92442
47	47		28,36434	46,97667	40,84464	38,92442
50	50		33,36434	47,97667	40,84464	41,92442
50	50		28,36434	43,97667	41,84464	41,92442

6. Semua variabel telah masuk ke data Standard Dev, kemudian kita regresikan. Yang kita regresikan adalah ZX1, ZX2, dan ZY terhadap ZZ dengan cara klik Data lalu klik Data Analisis, klik Regression, dan klik Ok.



7. Dari gambar di atas, terdapat permintaan **Input Y Range** dan **Input X Range**. Untuk Input Y Range masukkan data Standard Dev ZZ dengan cara blok keseluruhan dari atas sampai bawah, sedangkan untuk Input X Range masukkan data Standard Dev ZX1, ZX2, dan ZY dengan cara yang sama. Karena label ikut terkena blok maka Labels kita centang dan klik OK.



8. Dari langkah-langkah di atas akan keluar *output* di bawah ini. Untuk *coefficient path* terdapat pada warna biru dan nilai signifikansinya berwarna hijau. Hasil tersebut sama dengan hasil SPSS yang ada di atas.

	A	B	C	D	E	F	G	H	I
1	SUMMARY OUTPUT								
2									
3	Regression Statistics								
4	Multiple F	0,803013							
5	R Square	0,644831							
6	Adjusted R	0,635231							
7	Standard Error	3,113152							
8	Observations	115							
9									
10	ANOVA								
11		<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
12	Regression	3	1953,141	651,0471	67,17563	7,65E-25			
13	Residual	111	1075,78	9,691715					
14	Total	114	3028,922						
15									
16		<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95,0%</i>	<i>Upper 95,0%</i>
17	Intercept	2,270944	2,456736	0,924374	0,357297	-2,59724	7,139131	-2,59724	7,139131
18	ZX1	0,219165	0,09723	2,254086	0,026154	0,026497	0,411833	0,026497	0,411833
19	ZX2	0,12009	0,058582	2,049944	0,042726	0,004006	0,236173	0,004006	0,236173
20	ZY	0,626407	0,075157	8,334688	2,32E-13	0,477479	0,775335	0,477479	0,775335
21									
22									
23									

Daftar Rujukan

- Kusnendi, E.S. (2005). *Panduan Operasional Metode Penelitian Ilmiah*. Bandung: Universitas Pendidikan Indonesia.
- Kerlinger F.N. dan Pedhazur, J.P (1973), *Multiple Regression in Behavioral Research*. New York: Holt, Rinehart and Winston, Inc.
- Raltan, P. dan Renhard M. (2014). *Analisis Jalur (Path Analysis): Teori dan Aplikasi dalam Riset Bisnis*. Jakarta: Rineka Cipta.
- Riduwan dan Kuncoro, E.A. (2014). *Cara Mudah Menggunakan dan Memaknai Path Analysis (Analisis Jalur)*. Bandung: Alfabeta.
- Sugiyono. (2015). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Bandung: CV Alfabeta.

BAB 3

UJI ANOVA DUA JALUR (*TWO WAY ANOVA*)

Pada dasarnya *two way ANOVA* merupakan bagian dari statistik parametrik. Uji ANOVA terdapat beberapa jenis, di antaranya: (1) *one way ANOVA*, (2) *two way ANOVA*, dan (3) *multiple way ANOVA*. Yang dibahas pada bab ini adalah uji ANOVA dua jalur.

A. Konsep Uji ANOVA Dua Jalur/ANOVA Dua Arah

Two way ANOVA disebut juga dengan ANOVA dua jalur/ANOVA dua arah atau analisis varian dua faktor merupakan uji perbedaan parameter rata-rata satu variabel terikat untuk lebih dari dua kelompok sampel.

B. Kegunaan Uji *Two Way ANOVA*

Uji *two way ANOVA* lebih digunakan pada kelompok atau sampel dua atau lebih dari dua. Uji *two way ANOVA* digunakan untuk mengetahui apakah suatu faktor dari kedua faktor yang nanti akan dianalisis terdapat interaksi atau tidak terhadap variabel dependennya.

C. Prasyarat atau Asumsi Penggunaan Analisis ANOVA Dua Jalur

1. Sampel penelitian berasal dari kelompok yang independen.
2. Variabel faktor bersifat data kategori atau data nonmetrik atau data simbolis sifatnya data ordinal.
3. Variabel dependen adalah data berskala kuantitatif (interval/rasio).
4. Nilai residual standar berdistribusi normal (uji normalitas).
5. Varian antarkelompok harus homogen (uji homogenitas) .

D. Contoh Persoalan Kasus

Pengaruh model pembelajaran dan kemampuan awal terhadap prestasi belajar.

Faktor 1: Model pembelajaran dengan level, yaitu *inquiry* dan *jigsaw*.

Faktor 2: Kemampuan awal dengan level yaitu tinggi dan rendah.

Prestasi belajar merupakan variabel terikat.

E. Menggunakan Perhitungan Manual

Suatu penelitian bertujuan mempelajari pengaruh model pembelajaran dan kemampuan awal terhadap prestasi belajar pemrograman komputer. Untuk keperluan itu, telah diambil dua kelompok sampel acak untuk belajar menggunakan model pembelajaran *inquiry* (A1) dan *jigsaw* (A2). Setiap kelompok masing-masing dibagi dua secara acak dan dibedakan menjadi dua berdasarkan kemampuan awal, yaitu kelompok tinggi (B1) dan kelompok rendah (B2). Skor prestasi belajar pemrograman komputer dijadikan sebagai berikut.

Tabel 3.1 Skor prestasi belajar pemrograman komputer.

Model Pembelajaran (A)	Kemampuan Awal (B)	
	Tinggi (B1)	Rendah (B2)
<i>Inquiry</i> (A1)	90	80
	90	70
	90	70
	90	70
	90	70
	80	70
	80	60

	80	60
	80	60
	70	60
	90	80
	90	70
	90	70
	80	70
	80	60
	80	60
	80	60
	70	60
	70	50
	70	50
<i>Jigsaw (A2)</i>		

Lakukan pengujian hipotesis perbedaan rata-rata dengan signifikansi 5%?

1. Membuat Hipotesis Penelitian H_0 dan H_1

a. *Hipotesis Penelitian dengan Faktor Model Pembelajaran*

$H_{0(A)}$: Tidak ada perbedaan prestasi belajar pemrograman komputer antara siswa yang dibelajarkan dengan model pembelajaran *inquiry* dan *jigsaw*.

$H_{1(A)}$: Ada perbedaan prestasi belajar pemrograman komputer antara siswa yang dibelajarkan dengan model pembelajaran *inquiry* dan *jigsaw*.

b. *Hipotesis Penelitian dengan Faktor Kemampuan Awal*

$H_{0(B)}$: Tidak ada perbedaan prestasi belajar pemrograman komputer antara siswa yang memiliki kemampuan awal tinggi dan rendah.

$H_{1(B)}$: Ada perbedaan prestasi belajar pemrograman komputer antara siswa yang memiliki kemampuan awal tinggi dan rendah.

c. *Hipotesis Penelitian Interaksi Model Pembelajaran dengan Kemampuan Awal*

H_0 (AB): Tidak ada interaksi prestasi belajar pemrograman komputer antara siswa yang memiliki kemampuan awal tinggi dan rendah yang dibelajarkan dengan model pembelajaran *inquiry* dan *jigsaw*.

H_1 (AB): Ada interaksi prestasi belajar pemrograman komputer antara siswa yang memiliki kemampuan awal tinggi dan rendah yang dibelajarkan dengan model pembelajaran *inquiry* dan *jigsaw*

2. Membuat Hipotesis Statistik

a. *Hipotesis Statistik dengan Faktor Model Pembelajaran*

$$A_{0(A)} : \mu_1 = \mu_2$$

$$A_{1(A)} : \mu_1 \neq \mu_2$$

b. *Hipotesis Statistik dengan Faktor Kemampuan Awal*

$$A_{0(B)} : \mu_1 = \mu_2$$

$$A_{1(B)} : \mu_1 \neq \mu_2$$

c. *Hipotesis Statistik Interaksi Model Pembelajaran dengan Kemampuan Awal*

$$A_{0(AB)} : \mu_{11} = \mu_{12}$$

$$A_{1(AB)} : \mu_{11} \neq \mu_{12}$$

3. Menentukan Taraf Signifikansi

Taraf signifikansi $\alpha = 5\%$ atau 0,005

4. Tahap Empat Menentukan Uji yang Digunakan

Uji ANOVA dua jalur (*two way ANOVA*)

5. Tahap Lima Menentukan Kaidah Pengujian

Jika $F_{hitung} > F_{tabel}$, maka H_0 ditolak

Jika $F_{hitung} \leq F_{tabel}$, maka H_0 diterima

6. Menentukan F_{hitung} dan F_{tabel}

a. *Membuat Tabel Penolong*

Tabel 3.2 Tabel Penolong.

No	Inquiry (A1)				Jigsaw (A2)			
	Kemampuan Awal Tinggi (B1)		Kemampuan Awal Rendah (B2)		Kemampuan Awal Tinggi (B1)		Kemampuan Awal Rendah (B2)	
	X_1	$(X_1)^2$	X_2	$(X_2)^2$	X_3	$(X_3)^2$	X_4	$(X_4)^2$
1	90	8100	80	6400	90	8100	80	6400
2	90	8100	70	4900	90	8100	70	4900
3	90	8100	70	4900	90	8100	70	4900
4	90	8100	70	4900	80	6400	70	4900
5	90	8100	70	4900	80	6400	60	3600
6	80	6400	70	4900	80	6400	60	3600
7	80	6400	60	3600	80	6400	60	3600
8	80	6400	60	3600	70	4900	60	3600
9	80	6400	60	3600	70	4900	50	2500
10	70	4900	60	3600	70	4900	50	2500
Σ	840	71000	670	45300	800	64600	630	40500

Tabel 3.3 Tabel Statistika.

Tabel Statistika					Σ
n_T	$n_1 = 10$	$n_2 = 10$	$n_3 = 10$	$n_4 = 10$	40
ΣX_T	$\Sigma X_1 = 840$	$\Sigma X_2 = 670$	$\Sigma X_3 = 800$	$\Sigma X_4 = 630$	2940
$n \bar{x}$	$\frac{\Sigma X_1}{n_1} = \frac{840}{10} = 84$	$\frac{\Sigma X_2}{n_2} = \frac{670}{10} = 67$	$\frac{\Sigma X_3}{n_3} = \frac{800}{10} = 80$	$\frac{\Sigma X_4}{n_4} = \frac{630}{10} = 63$	291
$\Sigma X_{2,4}$	-	$\Sigma X_2 = 670$	-	$\Sigma X_4 = 630$	1300
$\Sigma X_{1,3}$	$\Sigma X_1 = 840$	-	$\Sigma X_3 = 800$	-	1640
$(\Sigma X_T)^2$	$\Sigma X_1^2 = 71000$	$\Sigma X_2^2 = 45300$	$\Sigma X_3^2 = 64600$	$\Sigma X_4^2 = 40500$	221400

b. Menentukan Jumlah Kuadrat (JK) untuk Beberapa Sumber Variansi

Terdiri dari total (T), antarkelompok A (A), antarkelompok B (B), antarkelompok A dan B (A, B) dan eror (E).

1) Jumlah Kuadrat Total (JK_T)

$$JK_T = \sum XT^2 - \frac{(\sum XT)^2}{nT} = 221400 - \frac{(2940)^2}{40} = 221400 - \frac{8643600}{40} = 221400 - 216090 = 5310$$

2) Jumlah Kuadrat Antarkelompok A (JK_A)

$$\begin{aligned} JK_A &= \left(\sum \frac{(\sum XA)^2}{nA} \right) - \frac{(\sum XT)^2}{nT} \\ &= \frac{(\sum X1 + \sum X2)^2 + (\sum X3 + \sum X4)^2}{2 \cdot n} - \frac{(\sum XT)^2}{nT} \\ &= \frac{(840+670)^2 + (800+630)^2}{2 \cdot 10} - \frac{(2940)^2}{40} \\ &= \frac{(1510)^2 + (1430)^2}{2 \cdot 10} - \frac{(2940)^2}{40} \\ &= \frac{2280100 + 2044900}{20} - \frac{8643600}{40} \\ &= 216250 - 216090 = 160 \end{aligned}$$

3) Jumlah Kuadrat Antarkelompok B (JK_B)

$$\begin{aligned} JK_B &= \left(\sum \frac{(\sum XA)^2}{nA} \right) - \frac{(\sum XT)^2}{nT} \\ &= \frac{(\sum X1 + \sum X3)^2 + (\sum X2 + \sum X4)^2}{2 \cdot n} - \frac{(\sum XT)^2}{nT} \\ &= \frac{(840+800)^2 + (670+630)^2}{2 \cdot 10} - \frac{(2940)^2}{40} \\ &= \frac{(1640)^2 + (1300)^2}{2 \cdot 10} - \frac{(2940)^2}{40} \\ &= \frac{2689600 + 1690000}{20} - \frac{8643600}{40} \\ &= 218980 - 216090 = 2890 \end{aligned}$$

4) Jumlah Kuadrat Antarkelompok B (JK_{AB})

$$\begin{aligned} JK_{AB} &= \left(\sum \frac{(\sum XAB)^2}{n} \right) - \frac{(\sum XT)^2}{nT} - JK_A - JK_B \\ &= \frac{(\sum X1)^2 + (\sum X2)^2 + (\sum X3)^2 + (\sum X4)^2}{n} - \frac{(\sum XT)^2}{nT} - JK_A - JK_B \end{aligned}$$

$$\begin{aligned}
&= \frac{(840)^2 + (670)^2 + (800)^2 + (630)^2}{10} - \frac{(2940)^2}{40} - 160 - 2890 \\
&= \frac{705600 + 448900 + 640000 + 396900}{10} - \frac{8643600}{40} - 160 - 2890 \\
&= \frac{2191400}{10} - \frac{8643600}{40} - 160 - 2890 \\
&= 219140 - 216090 - 160 - 2890 = 0
\end{aligned}$$

5) *Jumlah Kuadrat Error (JK_E)*

$$JK_E = JK_T - JK_A - JK_B - JK_{AB} = 5310 - 160 - 2890 - 0 = 2260$$

c. *Mencari Nilai Derajat Kebebasan*

- 1) $DK_T = n_t - 1 = 40 - 1 = 39$
- 2) $DK_A = b - 1 = 2 - 1 = 1$ (b = baris)
- 3) $DK_B = k - 1 = 2 - 1 = 1$ (k = kolom)
- 4) $DK_{AB} = (DK_A \cdot DK_B) = 1 \cdot 1 = 1$
- 5) $DK_E = n_t - (b \cdot k) = 40 - (2 \cdot 2) = 40 - 4 = 36$

d. *Menentukan Nilai Varian*

- 1) $VAR_A = \frac{JK_A}{DK_A} = \frac{160}{1} = 160$
- 2) $VAR_B = \frac{JK_B}{DK_B} = \frac{2890}{1} = 2890$
- 3) $VAR_{AB} = \frac{JK_{AB}}{DK_{AB}} = \frac{0}{1} = 0$
- 4) $VAR_E = \frac{JK_E}{DK_E} = \frac{2260}{36} = 62,778$

e. *Menghitung Nilai F_{hitung}*

- 1) $F_{hitung(A)} = \frac{VAR_A}{VAR_E} = \frac{160}{62,778} = 2,549$

$$2) F_{hitung(B)} = \frac{VAR B}{VAR E} = \frac{2890}{62,778} = 46,035$$

$$3) F_{hitung(AB)} = \frac{VAR AB}{VAR E} = \frac{0}{62,778} = 0$$

f. Menentukan Nilai F_{tabel}

Nilai F_{tabel} dapat dicari dengan menggunakan tabel F

- 1) F_{tabel} (Model Pembelajaran)

Dimana $DK_A = \text{pembilang} = 1$; $DK_E = \text{penyebut} = 36$

$$F_{A(tabel)} = F_{(\alpha)(DK A, DK E)} = F_{(0,05)(1, 36)} = 4,110$$

- 2) F_{tabel} (Kemampuan Awal)

Dimana $DK_B = \text{pembilang} = 1$; $DK_E = \text{penyebut} = 36$

$$F_{B(tabel)} = F_{(\alpha)(DK B, DK E)} = F_{(0,05)(1, 36)} = 4,110$$

- 3) F_{tabel} (Model Pembelajaran dan Kemampuan Awal)

Dimana $DK_{AB} = \text{pembilang} = 1$; $DK_E = \text{penyebut} = 36$

$$F_{AB(tabel)} = F_{(\alpha)(DK AB, DK E)} = F_{(0,05)(1, 36)} = 4,110$$

g. Membuat Tabulasi Ragam ANOVA Dua Jalur (Two Way ANOVA)

Tabel 3.4 Ragam ANOVA Dua Jalur (*Two Way ANOVA*).

Sumber Varian	JK	DK	VAR	F_{hitung}	$F_{tabel} (\alpha = 0,05)$
Antarkelompok A (Model Pembelajaran)	160	1	160	2,549	4,110
Antarkelompok B (Kemampuan Awal)	2890	1	2890	46,035	4,110
Antarkelompok AB (Model Pembelajaran dan Kemampuan Awal)	0	1	0	0	4,110
Error (E)	2260	36	62,778		
Total (T)	4230	39			

7. Menarik Kesimpulan

a. Model Pembelajaran

$$F_{hitung} = 2,549$$

$$F_{tabel} = 4,110$$

$F_{hitung} < F_{tabel}$ maka H_0 diterima dan H_1 ditolak. Artinya, tidak ada perbedaan prestasi belajar pemrograman komputer antara siswa yang dibelajarkan dengan model pembelajaran *inquiry* dan *jigsaw*.

b. Kemampuan Awal

$$F_{hitung} = 47,314$$

$$F_{tabel} = 4,110$$

$F_{hitung} > F_{tabel}$ maka H_0 ditolak dan H_1 diterima. Artinya ada perbedaan prestasi belajar pemrograman komputer antara siswa yang memiliki kemampuan awal tinggi dan rendah.

c. Interaksi Model Pembelajaran dengan Kemampuan Awal

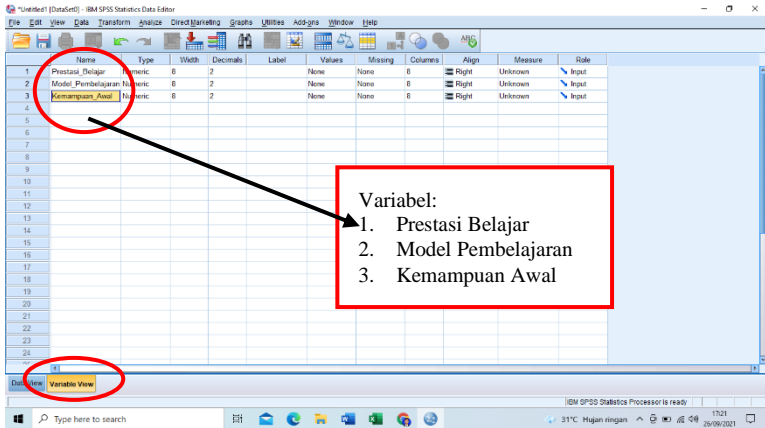
$$F_{hitung} = 0$$

$$F_{tabel} = 4,110$$

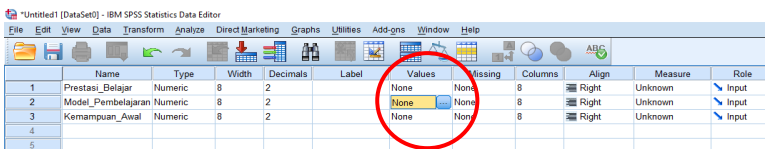
$F_{hitung} < F_{tabel}$ maka H_0 diterima dan H_1 ditolak artinya tidak ada interaksi prestasi belajar pemrograman komputer antara siswa yang memiliki kemampuan awal tinggi dan rendah yang dibelajarkan dengan model pembelajaran *inquiry* dan *jigsaw*.

F. Menggunakan SPSS

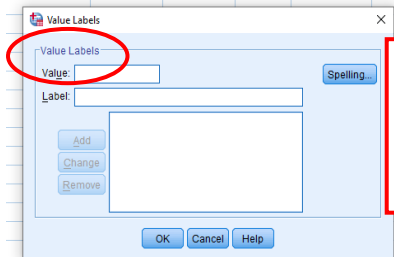
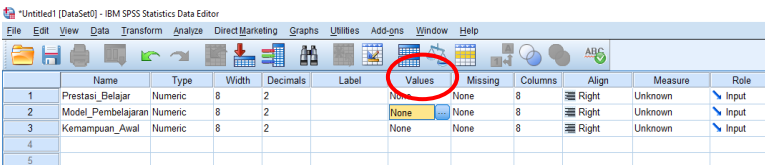
1. Buka aplikasi SPSS pada komputer.
2. Masuk pada Sheet Variabel View lalu masukkan variabelnya.



3. Atur bagian Values, klik salah satu variabel.



4. Pada bagian Values, klik kotak biru dengan tanda titik-titik (...) sehingga muncul Tab Value Labels.

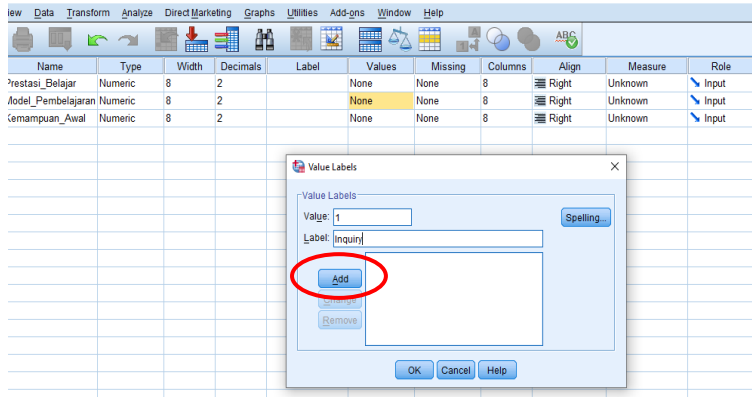


Value Labels
 Digunakan untuk memberikan label sesuai dengan kategori yang ingin kita inputkan pada masing-masing variabel

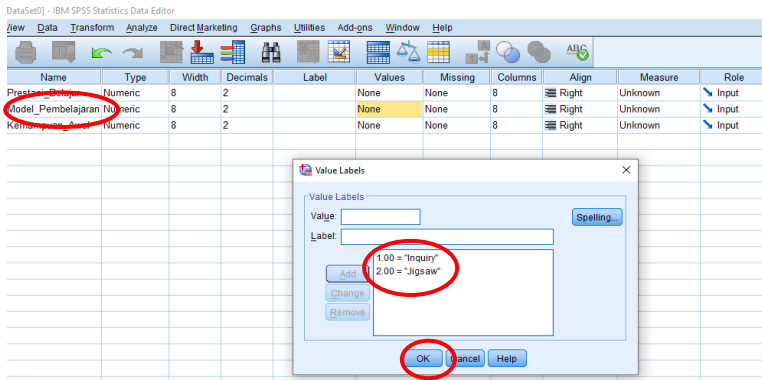
Keterangan:

Pada variabel Model Pembelajaran
Value 1 untuk label *inquiry*
Value 2 untuk label *jigsaw*

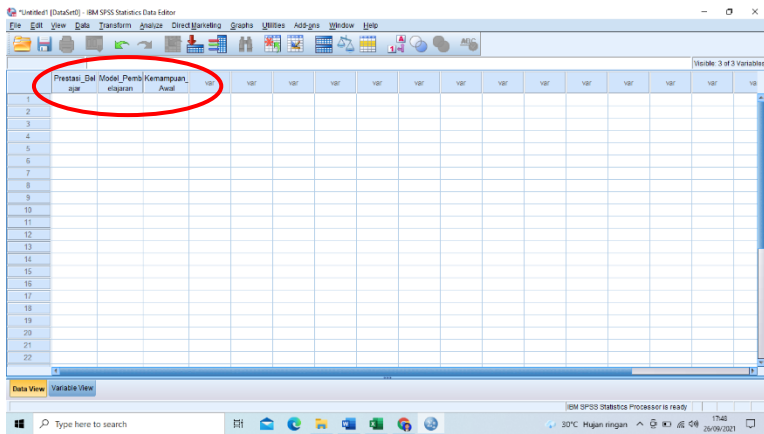
Pada variabel Kemampuan Awal
Value 1 untuk label tinggi
Value 2 untuk label rendah



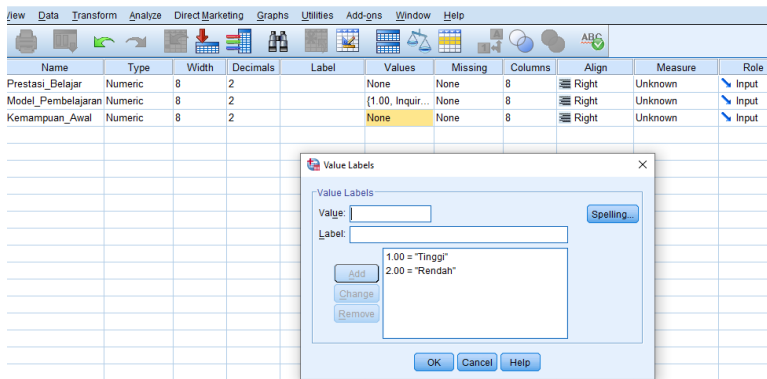
5. Setelah memasukkan Value 1, jangan lupa tekan tombol Add untuk memasukkan Value 2. Setelah masuk semuanya, tekan tombol OK untuk memasukkan variabel selanjutnya.



- Setelah klik OK lakukan hal yang sama untuk memasukkan Value 1 dan 2 pada variabel Kemampuan Awal.



- Setelah Value label sudah masuk. Langkah selanjutnya adalah menuju ke Sheet Data View yang sudah terdapat Header Prestasi Belajar, Model Pembelajaran, dan Kemampuan Awal.



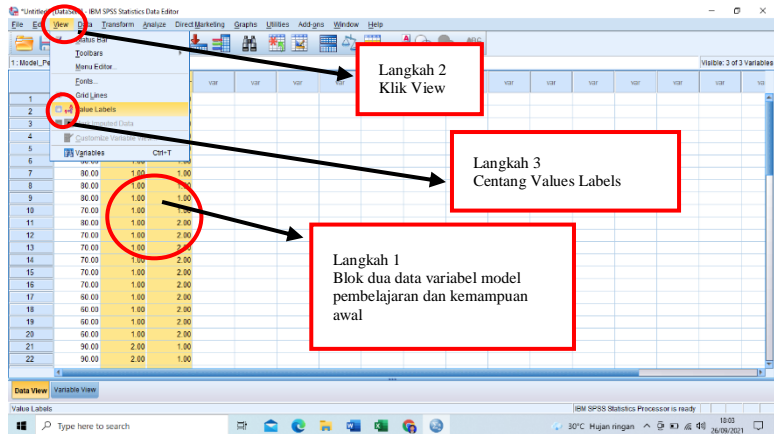
- Masukkan data-data yang akan dianalisis.

PRESTASI BELAJAR PEMROGRAMAN KOMPUTER				
No	Inquiry (A1)		Jigsaw (A2)	
	Kemampuan Awal Tinggi (B1)	Kemampuan Awal Rendah (B2)	Kemampuan Awal Tinggi (B1)	Kemampuan Awal Rendah (B2)
1	90	80	90	80
2	90	70	90	70
3	90	70	90	70
4	90	70	80	70
5	90	70	80	60
6	80	70	80	60
7	80	60	80	60
8	80	60	70	60
9	80	60	70	50
10	70	60	70	50
Inquiry = 1				
Jigsaw = 2				
Kemampuan Awal Tinggi = 1				
Kemampuan Awal Rendah = 2				

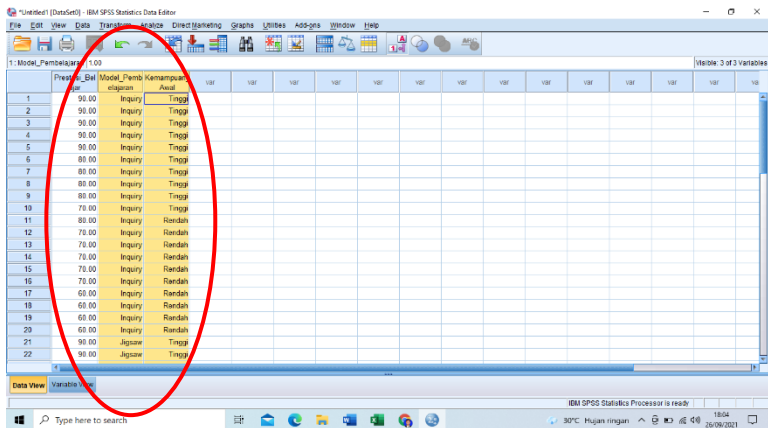
9. Data pertama, yaitu prestasi belajar, model pembelajaran *inquiry*, dan kemampuan awal tinggi. Pada Kolom Prestasi Belajar diisi sesuai data yang ada, sedangkan pada kolom Model Pembelajaran *inquiry* dan Kemampuan Awal data diisi sesuai Value, yaitu 1 karena kategori tinggi diisi sesuai Value yang sama.

	Prestasi_Belajar	Model_Pembelajaran	Kemampuan_Awal	var
1	90.00	1.00	1.00	
2	90.00	1.00	1.00	
3	90.00	1.00	1.00	
4	90.00	1.00	1.00	
5	90.00	1.00	1.00	
6	80.00	1.00	1.00	
7	80.00	1.00	1.00	
8	80.00	1.00	1.00	
9	80.00	1.00	1.00	
10	70.00	1.00	1.00	
11				

10. Masukkan data sesuai dengan kategori masing-masing.
11. Jika data telah masuk semua, untuk meyakinkan apakah value yang dimasukkan sesuai dengan kategori data, blok dua variabel kategori, yaitu Model Pembelajaran dan Kemampuan Awal lalu klik View lalu centang Value Label. Dengan demikian, akan muncul tampilan seperti di bawah ini.



Hasilnya seperti ini:



12. Jika data sudah dimasukkan semua, kita dapat melakukan uji prasyarat yang pertama, yaitu uji normalitas data.

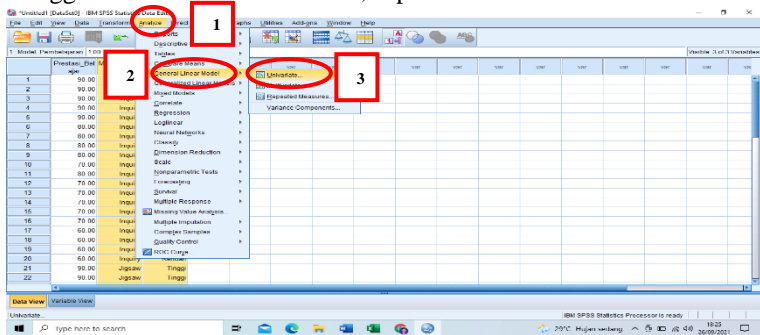
G. UJI NORMALITAS DATA

Uji normalitas data digunakan untuk menentukan apakah data yang digunakan berupa data yang berdistribusi normal atau tidak. Uji normalitas dilakukan dengan menggunakan *software* aplikasi SPSS dengan kriteria data akan dianggap normal jika skor signifikansi (Sig.) > taraf signifikansi, yaitu sebesar 0,05. Uji normalitas data dapat dilakukan dengan menggunakan beberapa metode sesuai dengan kriteria sebagai berikut.

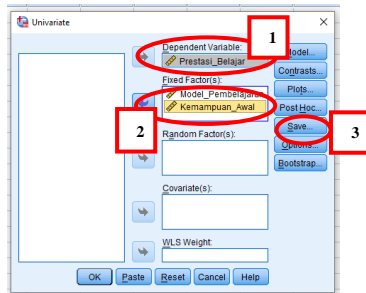
1. Uji normalitas probabilitas Shapiro-Wilk, digunakan pada jumlah data kurang dari 50 sampel.
2. Uji normalitas probabilitas Kolmogorov-Smirnova, digunakan pada jumlah data lebih dari 50 sampel.

1. Langkah-Langkah

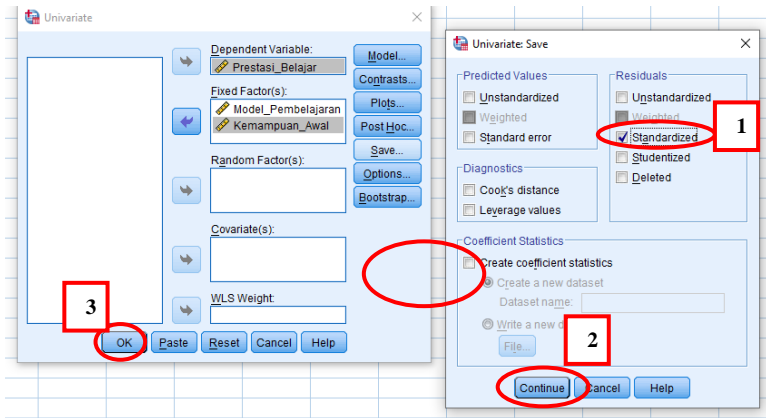
- a. Mencari nilai *standarize residual*.
- b. Klik Analyze lalu pilih General Linear Model dan klik Univariate sehingga muncul kotak Univariate, seperti berikut ini.



- c. Masukkan prestasi belajar pada kolom Dependent Variable, sedangkan model pembelajaran dan kemampuan awal pada kolom Fixed Factors(s).



- d. Klik tombol Save lalu klik/centang Standardized.
e. Klik Continue lalu klik OK.



- f. Selanjutnya beralih ke bagian Snet Data View yang menampilkan kolom nilai *standarize residual/ZRE_1*.

gg.

IBM SPSS Statistics Data Editor

File Edit View Data Transform Analyze Direct Marketing Graphs Utilities Add-ons Window Help

10 :

	Prestasi Belajar	Model Pembelajaran	Kemampuan Awal	ZRE_1	var	var	var
1	90.00	Inquiry	Tinggi	.76			
2	90.00	Inquiry	Tinggi	.76			
3	90.00	Inquiry	Tinggi	.76			
4	90.00	Inquiry	Tinggi	.76			
5	90.00	Inquiry	Tinggi	.76			
6	80.00	Inquiry	Tinggi	-.50			
7	80.00	Inquiry	Tinggi	-.50			
8	80.00	Inquiry	Tinggi	-.50			
9	80.00	Inquiry	Tinggi	-.50			
10	70.00	Inquiry	Tinggi	-1.77			
11	80.00	Inquiry	Rendah	1.64			
12	70.00	Inquiry	Rendah	.38			
13	70.00	Inquiry	Rendah	.38			
14	70.00	Inquiry	Rendah	.38			
15	70.00	Inquiry	Rendah	.38			
16	70.00	Inquiry	Rendah	.38			
17	60.00	Inquiry	Rendah	-.88			
18	60.00	Inquiry	Rendah	-.88			
19	60.00	Inquiry	Rendah	-.88			
20	60.00	Inquiry	Rendah	-.88			
21	90.00	Jigsaw	Tinggi	1.26			
22	90.00	Jigsaw	Tinggi	1.26			

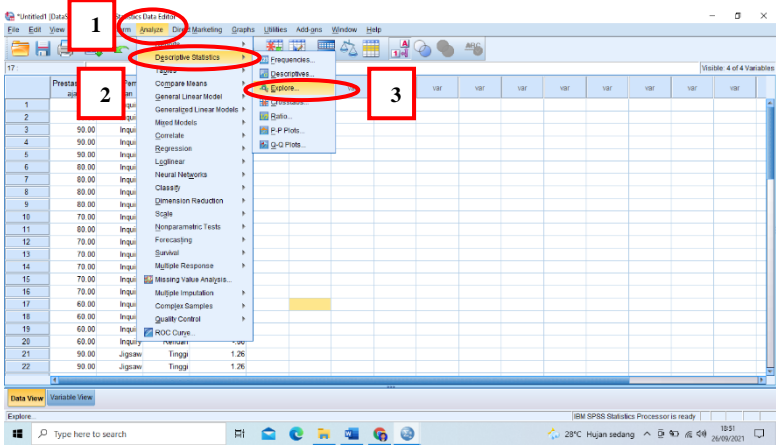
1

Data View Variable View

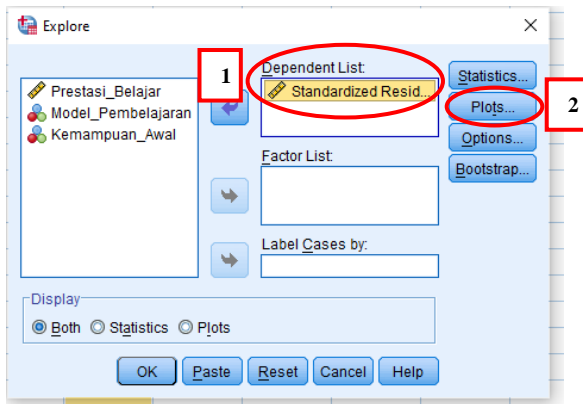
Data residual tersebut diperoleh dari hasil analisis variabel prestasi belajar, model pembelajaran dan keterampilan awal. Data residual tersebut yang nantinya akan digunakan uji distribusi datanya apakah normal atau tidak.

2. Melakukan Uji Normalitas Data

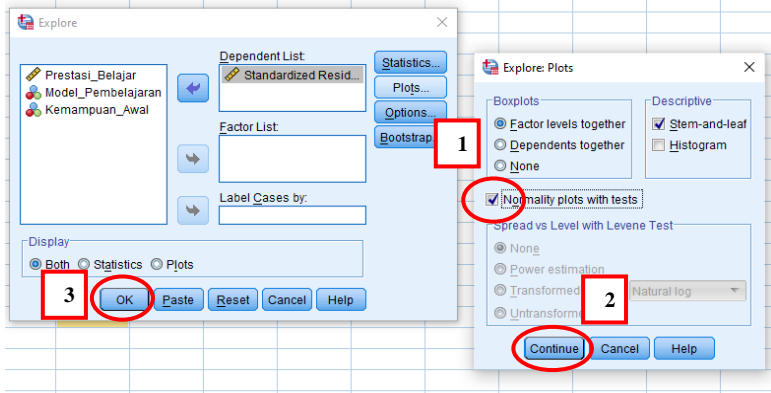
- a. Langkah yang dilakukan yaitu klik Analyze lalu pilih Descriptive Statistics, klik Explore sehingga muncul kotak Explore



- b. Muncul kotak Explore, kemudian masukkan Standardized Residu pada kolom Dependent List. Kemudian klik Plot sehingga muncul kotak Explore Plot.



- c. Setelah muncul kotak Explore Plot, klik Normality plots with test lalu klik Continue dan klik OK.



d. Selanjutnya akan muncul hasil uji normalitas pada tabel “Tests Of Normality”.

Tabel 3.5 *Tests of Normality.*

Tests of Normality						
	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual for Prestasi_Belajar	.110	40	.200 [*]	.975	40	.501

**. This is a lower bound of the true significance.*

Lilliefors Significance Correction

Karena data kurang dari 50 sampel maka yang dilihat pada kolom Shapiro-Wilk

Dasar Pengambilan Keputusan Uji Normalitas Data

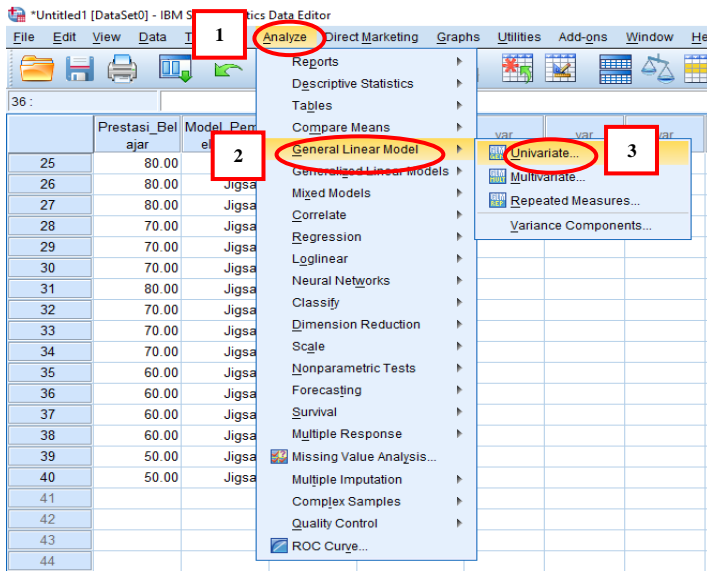
- Jika skor niali Sig. hitung \geq Sig. 5% atau 0,05 maka data berdistribusi normal.
- Jika skor niali Sig. hitung $<$ Sig. 5% atau 0,05 maka data berdistribusi tidak normal.

- c. Pada kolom Shapiro-Wilk nilai Sig. sebesar 0,501. Artinya, nilai Sig hitung \geq Sig. 5% atau 0,05 maka data yang digunakan berdistribusi normal.

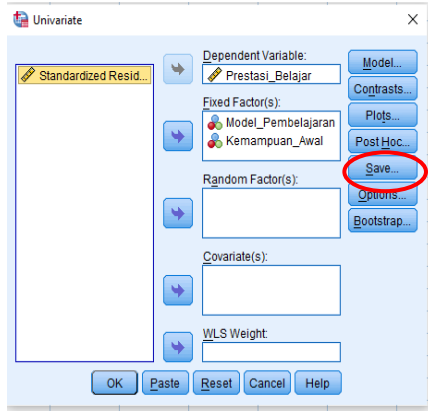
H. UJI HOMOGENITAS DATA

Uji homogenitas digunakan untuk menentukan apakah kelompok data mempunyai varian yang sama atau tidak. Jika kedua kelompok mempunyai varian yang sama maka kedua kelompok dikatakan homogen. Untuk melakukan uji homogenitas dalam menggunakan *software* aplikasi SPSS dengan kriteria data akan dianggap homogen jika skor nilai signifikansi (Sig.) pada tabel *Test of Homogeneity of Variances* > taraf signifikansi, yaitu 0,05. Langkah-langkah untuk uji homogenitas sebagai berikut.

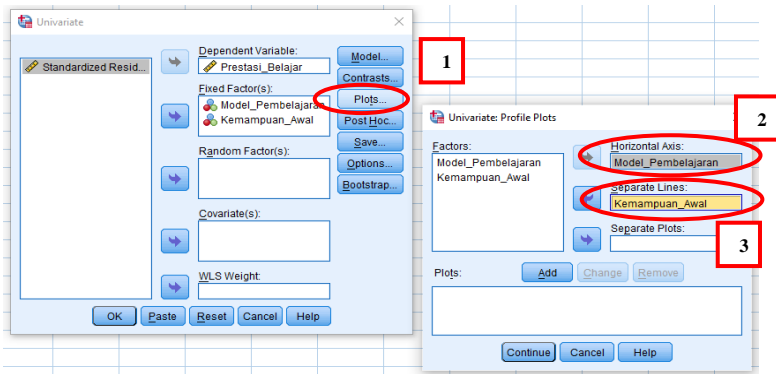
1. Klik Analyze lalu General Linear Model dan Pilih Univariate.



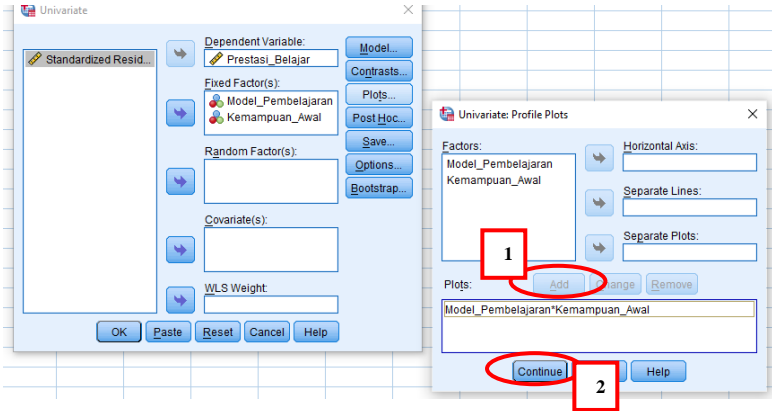
2. Selanjutnya akan muncul kotak Univariate, klik Save.



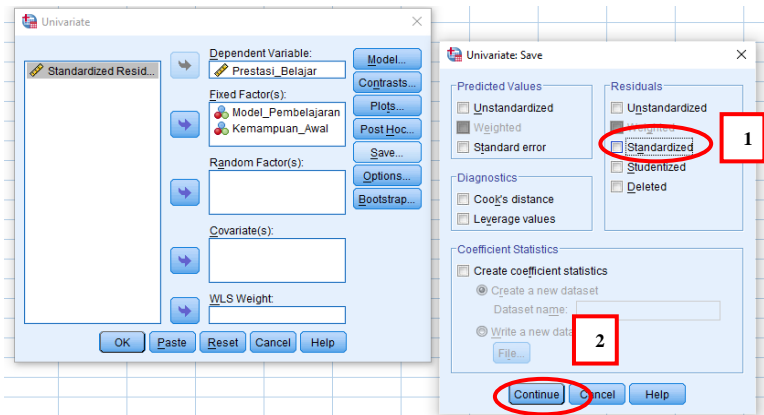
3. Hilangkan tanda centang pada bagian Standardized lalu klik Continue.
4. Kembali ke kotak Univariate lalu klik Plot sehingga muncul kotak Univariate Profile Plots. Pada kotak Univariate Profile Plots, Model Pembelajaran dimasukkan pada kolom Horizontal Axis, sedangkan Kemampuan Awal masukkan pada kolom Separate Lines.



5. Klik tombol Add lalu klik tombol Continue.

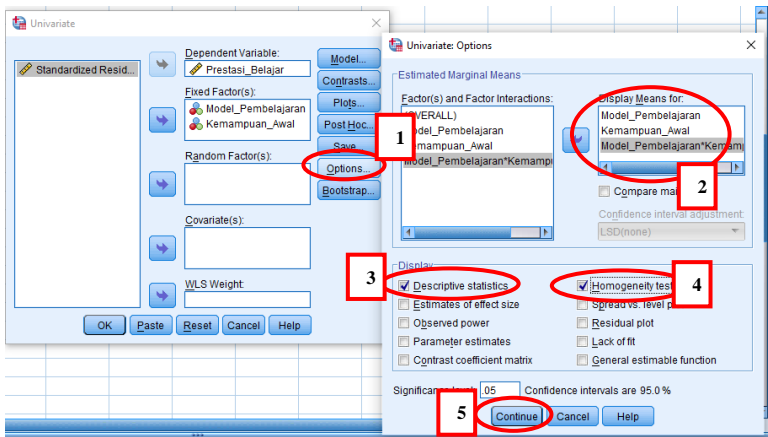


6. Selanjutnya kembali ke kotak Univariate, kemudian klik tombol Option sehingga muncul Univariate Option. Pada kotak Univariate Option, Model Pembelajaran, Kemampuan Awal, dan Model Pembelajaran*Kemampuan Awal dimasukkan pada kolom Display Means for. Pada Display klik centang Descriptive Statistics, centang Homogeneity Tests lalu klik Continue.



7. Kembali pada kotak Univariate, klik tombol Post Hock. Pada kotak Univariate Post Hock, masukkan kemampuan belajar pada kolom Post

Hoc Tests for. Kemudian klik centang Tukey's-b, klik Continue lalu klik tombol OK pada kotak Univariate.



8. Analisis hasil uji homogenitas pada tabel Levene's Test of Equality of Error Variances.

Levene's Test of Equality of Error Variances^a

Dependent Variable: Prestasi_Belajar

F	df1	df2	Sig.
.436	3	36	.728

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

Design:

Intercept + Model_Pembelajaran +
Kemampuan_Awal + Model_Pembelajaran*
Kemampuan_Awal.

Dasar pengambilan keputusan uji homogenitas:

1. Jika skor nilai Sig. hitung $\geq 0,05$ maka varian dikatakan homogen atau sama atau sejenis.
2. Jika skor nilai Sig. hitung $< 0,05$ maka varian dikatakan tidak homogen atau tidak sama atau tidak sejenis,

Pada tabel Levene's Test of Equality of Error Variances, didapatkan bahwa skor nilai Sig. hitung sebesar $0,727 \geq 0,05$ maka varian dikatakan homogen. Karena uji normalitas dan homogenitas sudah dilakukan dan hasilnya didapatkan data terdistribusi normal dan homogen maka selanjutnya melakukan uji hipotesis.

I. UJI HIPOTESIS

Untuk melakukan uji hipotesis, terdapat beberapa yang perlu diamati pada tabel Tests of Between-Subjects Effects.

Tabel 3.6 *Test of Between-Subjects Effects.*

<i>Tests of Between-Subjects Effects</i>					
Dependent Variable: Prestasi_Belajar					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	3050.000 ^a	3	1016.667	16.195	.000
Intercept	216090.000	1	216090.000	3442.142	.000
Model_Pembelajaran	160.000	1	160.000	2.549	.119
Kemampuan_Awal	2890.000	1	2890.000	46.035	.000
Model_Pembelajaran * Kemampuan_Awal	.000	1	.000	.000	1.000
Error	2260.000	36	62.778		
Total	221400.000	40			
Corrected Total	5310.000	39			

a. R Squared = .574 (Adjusted R Squared = .539)

Berdasarkan output SPSS Tests of Between-Subjects Effects.

J. CORRECTED MODEL

Pengaruh semua variabel independen (Model Pembelajaran, Kemampuan Awal dan Interaksi Model Pembelajaran dengan Kemampuan Awal atau Model Pembelajaran*Kemampuan Awal) secara bersama-sama terhadap variabel dependen (Prestasi Belajar). Apabila signifikansi (Sig.) < 0,05 (Alfa) = Signifikan. Contoh di atas 0,000 berarti model valid.

K. INTERCEPT

Nilai perubahan variabel dependen tanpa perlu dipengaruhi keberadaan variabel independen, artinya tanpa ada pengaruh variabel independen, variabel dependen dapat berubah nilainya. Apabila signifikansi (Sig.) < 0,05 (Alfa) = signifikan. Contoh di atas 0,000 berarti *intercept* signifikan.

L. HIPOTESIS

1. Model Pembelajaran

Diperoleh skor Sig. = 0,119 > 0,05 dengan nilai $F_{hitung} 2,549 < F_{tabel} 4,110$.

Jika skor Sig. = 0,119 > 0,05 dan skor $F_{hitung} < F_{tabel}$ maka H_0 diterima dan H_1 ditolak, artinya tidak ada perbedaan prestasi belajar pemrograman komputer antara siswa yang dibelajarkan dengan model pembelajaran *inquiry* dan *jigsaw*.

2. Kemampuan Awal

Diperoleh skor Sig.=0,00 < 0,05 dengan nilai $F_{hitung} 46,035 > F_{tabel} 4,110$. Jika skor Sig. = 0,00 < 0,05 dan $F_{hitung} > F_{tabel}$ maka H_0 ditolak dan H_1 diterima, artinya ada perbedaan prestasi belajar pemrograman komputer antara siswa yang memiliki kemampuan awal tinggi dan rendah.

3. Model Pembelajaran dan Kemampuan Awal

Diperoleh skor $Sig=1,000 > 0,05$ dengan nilai $0,00 < F_{tabel} 4,110$. Jika skor $Sig. = 1,000 > 0,05$, tetapi $F_{hitung} < F_{tabel}$ maka H_0 diterima dan H_1 ditolak artinya tidak ada interaksi prestasi belajar pemrograman komputer antara siswa yang memiliki kemampuan awal tinggi dan rendah yang dibelajarkan dengan model pembelajaran *inquiry* dan *jigsaw*.

M. ERROR

Nilai Error model, makin kecil maka model makin baik.

N. R-SQUARED

Nilai determinasi berganda semua variabel independen dengan dependen. Contoh di atas 0,574 di mana mendekati 1, berarti korelasi kuat.

O. POST HOC TEST

Dalam rangka interpretasi *output* Two Way ANOVA dengan SPSS, maka harus memahami tentang uji *post hoc*. Uji *post hoc* adalah uji lanjut atau uji yang menilai adanya perbedaan signifikan antarkelompok.

P. TABEL TUKEY POST HOC

Tabel di bawah ini adalah tabel *Tukey Post Hoc* digunakan untuk menilai kategori manakah dari variabel model pembelajaran dan kemampuan awal yang memiliki perbedaan signifikan:

1. Model_Pembelajaran

Dependent Variable: Prestasi_Belajar

Tabel 3.7 Model Pembelajaran.

Model_Pembelajaran	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
<i>Inquiry</i>	75.500	1.772	71.907	79.093
<i>Jigsaw</i>	71.500	1.772	67.907	75.093

2. Kemampuan_Awal

Dependent Variable: Prestasi_Belajar

Tabel 3.8 Kemampuan Awal.

Kemampuan_Awal	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
Tinggi	82.000	1.772	78.407	85.593
Rendah	65.000	1.772	61.407	68.593

3. Model_Pembelajaran*Kemampuan_Awal

Dependent Variable: Prestasi_Belajar

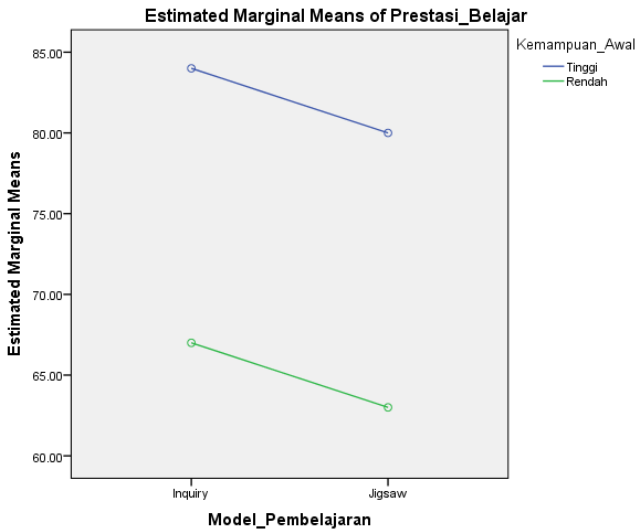
Tabel 3.9 Model_Pembelajaran*Kemampuan_Awal.

Model Pembelajaran	Kemampuan Awal	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
<i>Inquiry</i>	Tinggi	84.000	2.506	78.919	89.081
	Rendah	67.000	2.506	61.919	72.081
<i>Jigsaw</i>	Tinggi	80.000	2.506	74.919	85.081
	Rendah	63.000	2.506	57.919	68.081

Yang ada perbedaan signifikan ditandai dengan tanda bintang (*). Dari tabel di atas, tidak ada yang ditandai dengan bintang berarti semuanya tidak ada perbedaan yang signifikan. Dari hasil diagram *plot* pun juga sama.

Diagram *plot* berguna untuk menilai apakah ada interaksi efek

antarvariabel. Namun, diagram ini tidak bisa dijadikan bahan acuan yang valid, hanya sekadar memberikan gambaran saja. Apabila garis-garis tidak menunjukkan kesejajaran dicurigai ada efek interaksi.



Gambar 3.1 *Estimated Marginal Means of Prestasi_Belajar.*

Berdasarkan garis gambar di atas terdapat garis-garis yang menunjukkan kesejajaran sehingga dicurigai tidak ada efek interaksi.

1. Menganalisis Output SPSS *Between-Subjects Factors*

Tabel 3.10 *Between-Subjects Factors.*

		Value Label	N
Model_Pembelajaran	1.00	<i>Inquiry</i>	20
	2.00	<i>Jigsaw</i>	20
Kemampuan_Awal	1.00	Tinggi	20
	2.00	Rendah	20

Pada *output* SPSS Between-Subjects Factors, diketahui bahwa Jumlah sampel pada masing-masing kategori model pembelajaran dan kemampuan awal masing-masing kategori sejumlah 20 sampel hal tersebut sesuai dengan data yang kita analisis.

2. Menganalisis *Output* SPSS *Descriptive Statistics*

Tabel 3.11 *Dependent variable: prestasi belajar.*

Model_Pembelajaran	Kemampuan_Awal	Mean	Std. Deviation	N
<i>Inquiry</i>	Tinggi	84.0000	6.99206	10
	Rendah	67.0000	6.74949	10
	Total	75.5000	10.99043	20
<i>Jigsaw</i>	Tinggi	80.0000	8.16497	10
	Rendah	63.0000	9.48683	10
	Total	71.5000	12.25819	20
Total	Tinggi	82.0000	7.67772	20
	Rendah	65.0000	8.27170	20
	Total	73.5000	11.66850	40

Berdasarkan *output* SPSS descriptive statistics menunjukkan jumlah pengukuran (N), nilai rata-rata (Mean) dan standar deviasi (Std.) dari masing-masing variabel. Dari tabel di atas, kita bisa menilai rata-rata prestasi belajar berdasarkan model pembelajaran dan kemampuan awal. Sebagai contoh, nilai rata-rata prestasi belajar pada model pembelajaran *inquiry* dengan kemampuan awal tinggi sebesar 84,00; sedangkan nilai prestasi belajar model pembelajaran *inquiry* dengan kemampuan awal rendah sebesar 67,00 dan begitu seterusnya.

Daftar Rujukan

- Rumah Analisis Data. *Tutorial Video Uji Two Way ANOVA Menggunakan SPSS*. Diakses pada 26 September 2021 dari <https://www.youtube.com/watch?v=8J-hR3m9UsQ&t=377s>.
- Time 2 Study. *Tutorial Video Statistika-Two Way ANOVA (ANOVA 2 Jalur) Cara Manual*. Diakses pada 26 September 2021 <https://www.youtube.com/watch?v=KS5kPf-CMuA&t=1382s>.

BAB 4

ANALISIS DISKRIMINAN

A. Sejarah Singkat

Pembahasan analisis diskriminan dimulai pada awal tahun 1920-an. Ahli statistik Inggris, Karl Pearson (1857–1936), mengusulkan koefisien kesamaan ras (CRL), sejenis indeks jarak antarkelompok. CRL dipelajari secara ekstensif oleh GM Morant (1899–1964) pada tahun 1920-an. Pada tahun 1920-an juga, studi tentang indeks jarak lain dimulai di India yang diformalkan oleh PC Mahalanobis (1893–1972) pada tahun 1930-an. Gagasan jarak antarkelompok multivariabel diterjemahkan ke dalam komposit linier variabel yang diturunkan untuk tujuan klasifikasi dua kelompok oleh RA Fisher (1890–1962) pada tahun 1930-an. Ide komposit jarak dan variabel muncul di media cetak sebelum artikel analisis diskriminan Fisher pada tahun 1936 (*“Penggunaan beberapa pengukuran dalam masalah taksonomi,”* yang muncul di Sejarah Eugenika). Atas saran Fisher, MM Barnard menerapkan analisis diskriminan dua kelompok (prediktif) dalam penelitian tahun 1935 yang melibatkan tujuh karakter tengkorak Mesir. Perluasan klasifikasi dua kelompok menjadi beberapa kelompok diberikan oleh CR Rao pada tahun 1948. Banyak perluasan dan penyempurnaan lain dari gagasan Fisher telah muncul sejak tahun 1940-an (Huberty, 2006).

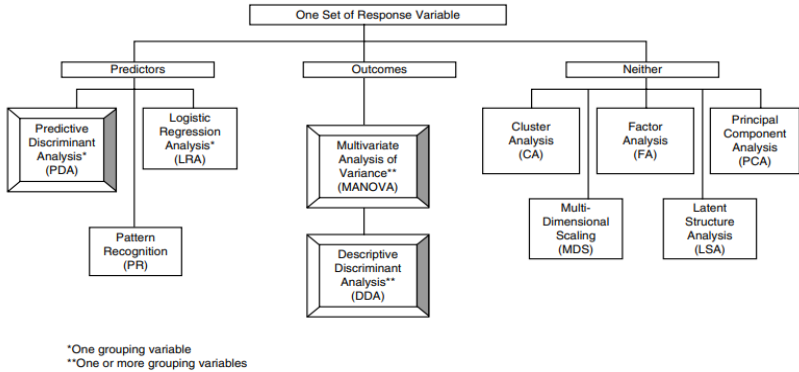
Meskipun studi awal analisis diskriminan melibatkan aplikasi dalam ilmu biologi dan kedokteran, minat yang cukup besar dibangkitkan oleh ahli statistik. Metodologi di bidang studi, seperti bisnis, pendidikan, teknik, dan psikologi. Ketertarikan ini mendorong penulisan buku teks yang membahas analisis diskriminan dalam berbagai bentuk dan dari berbagai perspektif.

Beberapa buku pra-1970 dengan cita rasa terapan adalah Rao (1952), Kendall (1957), Cooley dan Lohnes (1962), serta Rulon dkk. (1967).

Selama tiga atau empat dekade pertama, tulisan-tulisan tentang analisis diskriminan berfokus pada prediksi keanggotaan kelompok, berlabel analisis diskriminan prediktif (PDA) dalam buku saat ini. Dalam ilmu *non-behavioral*, fokus ini berlanjut hingga kini. Giri (2004: 477–482) mencantumkan 88 referensi terkait PDA, tahun publikasi berkisar dari 1921 hingga 1991. Meskipun Fisher menganggap komposit variabel linier (yaitu fungsi diskriminan linier, LDF) dari sudut pandang matematika pada tahun 1930-an, baru pada 1960-an LDF dianggap serius untuk menafsirkan efek yang diungkapkan melalui analisis multivariat varian (MANOVA) (misalnya Cooley dan Lohnes, 1968; Jones dan Bock, 1960). Aspek analisis diskriminan ini diberi label analisis diskriminan deskriptif (DDA) dalam buku saat ini. Menurut pandangan beberapa ahli metodologi, studi struktur (melalui LDFs) dalam konteks MANOVA memiliki potensi yang cukup besar untuk eksplorasi dan pengembangan teori substantif. Urgensi studi semacam itu mungkin dipertimbangkan, penggunaannya sangat terbatas dalam pengaturan penelitian terapan selama empat dekade terakhir (Huberty, 2006).

B. Konsep

Untuk mendapatkan gambaran kasar tentang bagaimana DDA dan PDA masuk ke dalam skema umum analisis data multivariat, dapat dilihat pada Gambar 4.1.



Sumber: Huberty, 2006.

Gambar 4.1 Klasifikasi Metode Multivariat.

Dalam membahas DDA dan PDA dari Gambar 4.1, referensi dibuat untuk analisis korelasi ganda (MCA) dan analisis regresi berganda (MRA). Analogi dapat dibuat sebagai berikut.

$$DDA : MCA \quad \parallel \parallel \quad PDA : MRA.$$

Artinya, DDA dan MCA dilakukan untuk tujuan “hubungan”, sedangkan PDA dan MRA untuk tujuan “prediksi” (Huberty, 2006).

C. Definisi

Analisis diskriminan merupakan teknik untuk menganalisis data dengan variabel dependen (kriteria) berkategori dan variabel independennya (*predictor*) dalam bentuk interval (matriks) (Malhotra, 2007).

Analisis diskriminan adalah metode statistik canggih yang secara bersamaan memeriksa variabel yang berbeda dan menentukan asal kelompok

suatu parameter. Metode ini membuat kombinasi linier variabel bebas untuk menguji suatu parameter tergantung pada salah satu dari dua kelompok (Lekshmi et al., 1998). Analisis diskriminan merupakan teknik statistika untuk mengelompokkan individu-individu ke dalam kelompok-kelompok yang saling bebas dengan tegas berdasarkan segugus peubah bebas (Matjik, 2004). Analisis diskriminan merupakan teknik multivariat yang termasuk dalam *dependence method*, yaitu terdapat variabel dependen dan variabel independen dengan ciri variabel dependen harus berupa data kategori, sedangkan variabel independen berupa data nonkategori. Analisis diskriminan mirip dengan analisis regresi, perbedaannya adalah analisis regresi pada variabel dependen harus data rasio, sedangkan jenis data untuk variabel independen dapat berupa data rasio atau kategori (Santoso, 2002).

Model analisis diskriminan ditandai dengan ciri khusus, yaitu data variabel dependen yang harus berupa data kategori, sedangkan data independen berupa data nonkategori. Hal ini dapat dimodelkan sebagai berikut.

$$\begin{array}{ccc}
 \text{Y}_1 & = & \text{X}_1 + \text{X}_2 + \text{X}_3 + \dots + \text{X}_n \\
 \text{Nonmetriks} & & \text{Metriks}
 \end{array}$$

Keterangan:

1. Variabel independen (X_1 dan seterusnya) adalah data metrik, yaitu data berskala interval atau rasio. Contohnya: usia seseorang, tinggi sebuah bangunan, kandungan lemak dalam tubuh, dan total pendapatan per bulan.
2. Variabel dependen (Y_1) adalah data kategorikal atau nominal, seperti level operator (kode 1), level supervisor (kode 2), level manajer (kode

3). Jika data kategorikal tersebut hanya terdiri atas dua kode saja (misal kode 1 untuk pegawai ASN dan kode 2 pegawai non-ASN), maka data kategorikal tersebut hanya terdiri dari 2 kode saja disebut ‘two-groups discriminant analysis’. Namun, apabila lebih dari 2 kategori disebut ‘multiple discriminant analysis’.

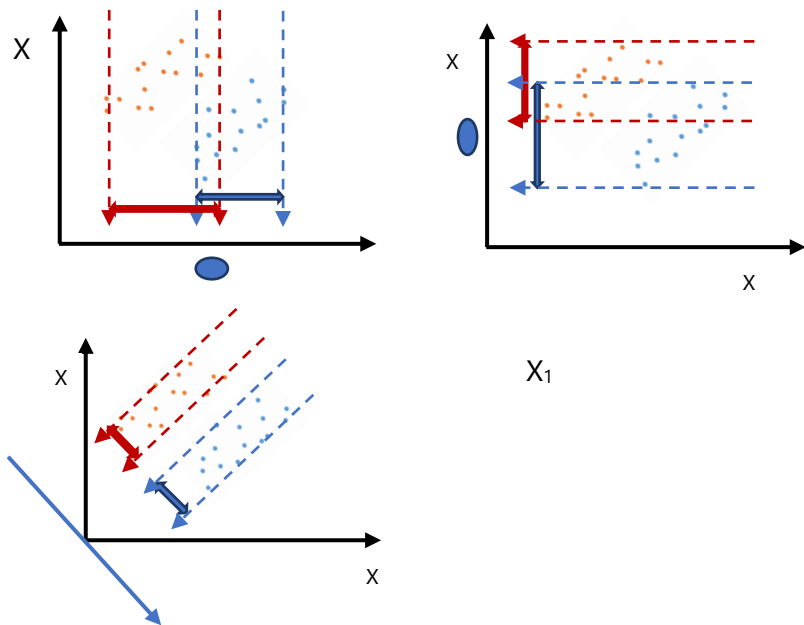
Fungsi diskriminan:

$$\widehat{D} = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + \dots b_kX_k$$

\widehat{D} = skor diskriminan

b = koefisien diskriminasi atau bobot

X = prediktor atau variabel independen



Gambar 4.2 Analogi Fungsi Diskriminan.

D. Tujuan Analisis Diskriminan

Bentuk multivariat dari analisis diskriminan adalah dependen sehingga variabel dependen adalah variabel yang menjadi dasar analisis diskriminan. Variabel dependen bisa berupa kode grup 1 atau grup 2 atau lainnya dengan tujuan diskriminan secara umum adalah

- 1) membangun fungsi diskriminan atau kombinasi linier dari independen variabel yang akan membedakannya dengan baik antara dependen variabel yang berkategori;
- 2) menguji apakah ada perbedaan yang signifikan antara kelompok dalam bentuk variabel prediktor;
- 3) menentukan manakah variabel independen yang memberikan kontribusi terbanyak terhadap kelompok yang berbeda tersebut;
- 4) mengklasifikasikan kasus-kasus dalam satu kelompok yang didasarkan pada nilai dari variabel independen; dan
- 5) mengevaluasi keakuratan dan pengklasifikasiannya.

Perbedaan dan persamaan antara analisis diskriminan dengan dua metode analisis regresi dan analisis varian (ANOVA) dapat diringkas sebagai berikut.

Tabel 4.1 Perbedaan dan persamaan analisis regresi, varian, dan diskriminan.

Pembeda	ANOVA	Regresi	Diskriminan
Persamaan			
Jumlah variabel dependen	Satu	Satu	Satu
Jumlah variabel independen	Banyak	Banyak	Banyak
Perbedaan			
Sifat variabel dependen	Matrik	Matrik	Kategori
Sifat variabel independen	Kategori	Matrik	Matrik

Sumber: Malhotra, 2007.

Ketiga jenis metode analisis tersebut melibatkan variabel dependen (kriteria) yang tunggal dan variabel independen yang banyak (ganda). Namun, sifat variabelnya yang berbeda. Pada analisis varian dan regresi, variabel dependennya adalah matriks atau skala interval, sedangkan analisis diskriminan variabel dependennya adalah dikategorikan. Dalam hal ini, analisis varian variabel independennya dikategorikan, seperti umur dan pendapatan dikategorikan dalam tinggi, sedang, dan rendah. Matriks dalam analisis regresi dan diskriminan keduanya diukur dalam skala rasio.

Tabel 4.2 Perbedaan analisis regresi linear dan diskriminan.

Analisis Diskriminan	Regresi Linear
Mengklasifikasikan suatu individu/objek.	Mengestimasi nilai Y.
Variabel respons: <i>fixed</i> (tidak disyaratkan mengikuti distribusi tertentu).	Variabel respons: distribusi normal dan homokedastisitas terpenuhi.
Variabel prediktor: berdistribusi normal, homokedastis.	Variabel prediktor: <i>fixed</i> (tidak disyaratkan mengikuti distribusi tertentu).
Memprediksi suatu objek masuk ke dalam salah satu kategori dari variabel tak bebas berdasarkan informasi dari objek tersebut.	Memprediksi atau mengestimasi nilai dari variabel tak bebas.

E. Asumsi Analisis Diskriminan

Asumsi penting yang harus dipenuhi agar model diskriminan dapat digunakan, antara lain

1. *Multivariate normality*, atau variabel independen seharusnya berdistribusi normal. Jika data tidak berdistribusi normal akan menyebabkan masalah pada ketepatan fungsi (model) diskriminan. Regresi logistik (*logistic regression*) bisa dijadikan alternatif metode jika memang data tidak berdistribusi normal.

2. Matriks *co-varian* dari semua variabel independen seharusnya sama (*equal*).
3. Tidak ada korelasi antar-variabel independen. Jika dua variabel independen mempunyai korelasi yang kuat, dikatakan terjadi multikolinieritas.
4. Tidak adanya data yang sangat ekstrem (*outlier*) pada variabel independen. Jika ada data *outlier* yang tetap diproses berakibat berkurangnya ketepatan klasifikasi dari fungsi diskriminan.

F. Proses Analisis Diskriminan

Beberapa langkah yang merupakan proses dasar dalam analisis diskriminan, antara lain

1. Memilah variabel-variabel menjadi variabel terikat (*dependent*) dan variabel bebas (*independent*).
2. Menentukan metode untuk membuat fungsi diskriminan, yaitu
 - a) *simultaneous estimation*; semua variabel dimasukkan secara bersama-sama lalu dilakukan proses diskriminan;
 - b) *step-wise estimation*; variabel dimasukkan satu per satu ke dalam model diskriminan. . pada proses ini, tentu ada variabel yang tetap ada pada model, dan ada kemungkinan satu atau lebih variabel independen yang 'dibuang' dari model.
3. Menguji signifikansi fungsi diskriminan yang terbentuk, dengan menggunakan Wilk's Lambda, Box M, F *test*.
4. Menguji ketepatan klasifikasi dari fungsi diskriminan (secara individual dengan *casewise diagnostics*).
5. Melakukan interpretasi fungsi diskriminan.
6. Melakukan uji validasi fungsi diskriminan.

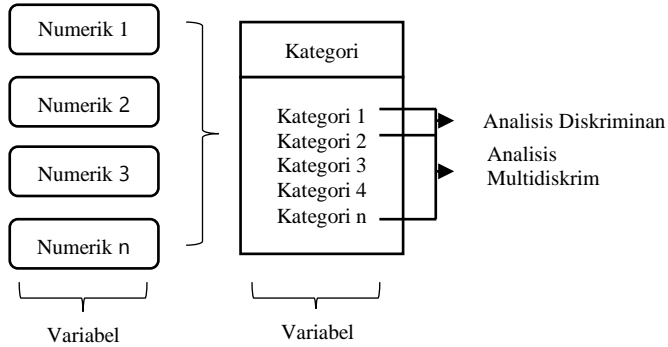
G. Analisis Diskriminan dan Analisis Multidiskriminan

1. Konsep Perbedaannya

Melalui analisis diskriminan dapat dibuat sebuah model seperti regresi, yaitu adanya satu variabel dependen, dan satu atau lebih variabel independen. Perbedaan analisis diskriminan dengan model regresi adalah pada jenis data yang digunakan. Pada analisis diskriminan, variabel dependen selalu data kategori. Prinsip diskriminan adalah ingin membuat model yang bisa secara jelas menunjukkan perbedaan (diskriminasi) antarisi variabel dependen.

Teknik analisis diskriminan dapat dikelompokkan dalam dua jenis. Jenis pertama disebut dengan “analisis diskriminan dua kelompok”, yaitu jika variabel kriteria (dependen) mempunyai dua kategori; sedangkan jenis kedua disebut “analisis diskriminan ganda”, yaitu jika melibatkan tiga atau lebih kategori (Malhotra, 2007). Perbedaan utama dua jenis analisis diskriminan tersebut adalah bahwa dalam kasus dua kelompok memungkinkan untuk menurunkan hanya satu fungsi diskriminan, sedangkan analisis diskriminan ganda lebih dari satu fungsi diskriminan harus dihitung.

Jika variabel dependen hanya dua kategori maka disebut dengan analisis diskriminan, tetapi jika variabel dependen lebih dari dua kategori maka analisis yang digunakan adalah analisis multidiskriminan. Gambar 4.3 menganalogikan perbedaan analisis diskriminan dan analisis multidiskriminan.



Gambar 4.3 Konsep perbedaan analisis diskriminan dan analisis multidiskriminan.

2. Kasus

Sebuah lembaga sertifikasi profesi sedang menganalisis hasil rekomendasi asesor kompetensi dalam pelaksanaan uji kompetensi pada peserta didik SMK di beberapa jejaring SMK yang tersebar di Jawa Timur dengan variabel berikut.

- Tipe peserta didik yang direkomendasikan kompeten, dengan kode:
Kode 1 = kompeten
- Tipe peserta didik yang direkomendasikan belum kompeten, dengan kode:
Kode 0 = belum kompeten
- Jam praktikum (jumlah jam pelajaran per minggu).
- Nilai praktikum (nilai dari uji praktik pada mata pelajaran kejuruan).
- Usia (usia peserta didik kelas XII SMK peserta uji kompetensi). Variabel nama tidak disertakan dalam proses analisis diskriminan, karena berupa data STRING (berisi karakter dan bukannya angka).

Variabel respons

- Rekomendasi belum kompeten ($Y = 0$)

b. Rekomendasi kompeten ($Y=1$)

Variabel prediktor

a. Jam Praktikum (X_1)

b. Nilai Praktikum (X_2)

c. Usia (X_3)

Untuk itu, langkah pertama pada analisis diskriminan adalah menguji apakah semua variabel independen (bebas) berbeda secara nyata berdasar variabel dependen. Terkait dengan kasus tersebut, sebagai contoh akan diuji apakah variabel jam praktikum yang berbeda, usia dan nilai praktikum akan memberikan rekomendasi kompeten dan belum kompeten pada peserta uji kompetensi.

SISWA	REKOMENDASI	JAM_PRAKTIKUM	NILAI_PRAKTIKUM	USIA
1	Belum Kompeten	7	80	15
2	Belum Kompeten	8	81	16
3	Belum Kompeten	7	82	17
4	Belum Kompeten	5	83	18
5	Belum Kompeten	4	80	19
6	Belum Kompeten	6	81	20
7	Belum Kompeten	8	82	15
8	Belum Kompeten	8	83	16
9	Belum Kompeten	6	80	17
10	Belum Kompeten	5	81	18
11	Belum Kompeten	7	82	19
12	Belum Kompeten	3	83	20
13	Belum Kompeten	7	81	15
14	Belum Kompeten	4	82	16
15	Belum Kompeten	3	83	17
16	Kompeten	12	88	18
17	Kompeten	14	89	19
18	Kompeten	15	90	20
19	Kompeten	13	91	15
20	Kompeten	15	88	16
21	Kompeten	12	89	17
22	Kompeten	17	90	18
23	Kompeten	13	91	19
24	Kompeten	15	92	20
25	Kompeten	12	89	15
26	Kompeten	12	88	16
27	Kompeten	14	90	17
28	Kompeten	16	92	18
29	Kompeten	12	92	19
30	Kompeten	13	92	20

Gambar 4.4 Tampilan data kasus

Contoh interpretasi data:

Pada baris pertama, peserta didik pertama merupakan peserta uji kompetensi yang direkomendasikan belum kompeten. Ia melaksanakan praktik pada mata pelajaran produktif selama 7 jam pelajaran dalam satu minggu, mendapatkan nilai 80 pada mata pelajaran kejuruan dan berusia 15 tahun.

Demikian seterusnya untuk konsumen-konsumen yang lain. Dari *file* diskriminan yang berisi peserta didik SMK yang melaksanakan uji kompetensi tersebut akan dianalisis diskriminan untuk mengetahui:

- a. Apakah ada perbedaan yang signifikan antara peserta didik yang direkomendasikan kompeten dengan mereka yang direkomendasikan belum kompeten?
- b. Jika ada perbedaan yang signifikan, variabel apa saja yang membuat hasil rekomendasi dari asesor kompetensi kompeten dan belum kompeten?
- c. Membuat model diskriminan dua faktor (karena hanya ada peserta didik kompeten dengan yang belum kompeten) dalam kasus ini.
- d. Menguji ketepatan model (fungsi) diskriminan.

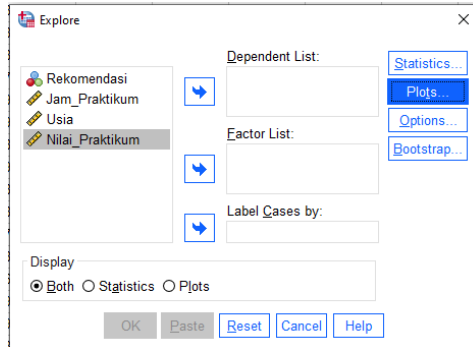
3. Langkah

Tahapan pengujian pada analisis diskriminan menggunakan SPSS:

a. Uji asumsi

Untuk memastikan variabel independen berdistribusi normal dan menghindari masalah pada ketepatan fungsi (model) diskriminan.

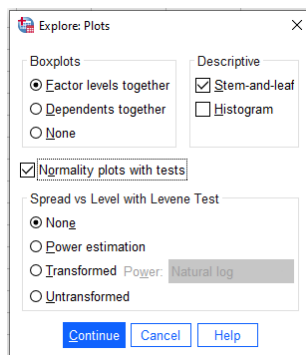
- 1) Buka file diskriminan
- 2) Menu Analyze Descriptive Statistics Explore seperti pada Gambar 4.5.



Gambar 4.5 Kotak Dialog Explore.

3) Pengisian

- a) Masukkan variabel Rekomendasi ke bagian Factor List. Hal ini berarti variabel Rekomendasi berfungsi sebagai *independent variable* (variabel bebas).
- b) Masukkan variabel Jam Praktikum, Usia, dan Nilai Praktikum ke bagian Dependent List.
- c) Klik Plots, pastikan kotak Normality plots with tests sudah ditandai lalu tekan Continue lalu OK.



Gambar 4.6 Kotak Dialog Explore Plots.

d) Cek *output* Test of Normality.

Tests of Normality							
	Rekomendasi	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
Jam_Praktikum	0	.206	15	.087	.902	15	.101
	1	.192	15	.143	.885	15	.056
Nilai_Praktikum	0	.173	15	.200 [*]	.876	15	.042
	1	.163	15	.200 [*]	.882	15	.051
Usia	0	.155	15	.200 [*]	.917	15	.175
	1	.155	15	.200 [*]	.917	15	.175

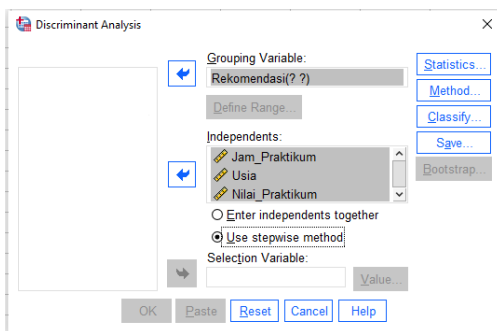
*. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Gambar 4.7 *Output* Test of Normality.

e) Kolom signifikan (Sig.) menunjukkan angka lebih besar dari 0,05 (Sig. > 0.05) artinya variabel yang digunakan berdistribusi normal.

b. Menilai Variabel yang Layak untuk Analisis

- 1) Buka *file* diskriminan.
- 2) Menu Analyze lalu Classify lalu Discriminan, seperti pada Gambar 4.8.



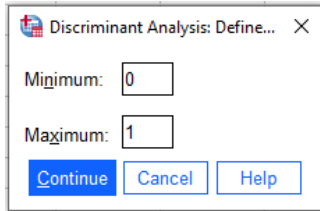
Gambar 4.8 Kotak dialog Diskriminan Analisis.

3) Pengisian:

- a) Masukkan variabel Rekomendasi ke bagian Grouping Variable. Hal ini berarti variabel Rekomendasi berfungsi sebagai *dependent*

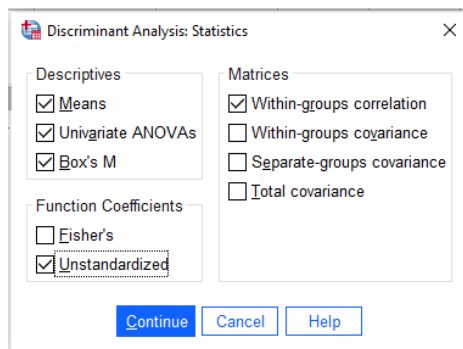
variable (variabel tergantung), yang berciri data kategori. Karena itu, SPSS meminta masukan kode kategori yang dipakai.

- b) Untuk itu, buka ikon Define Range..., hingga tampak di layar.



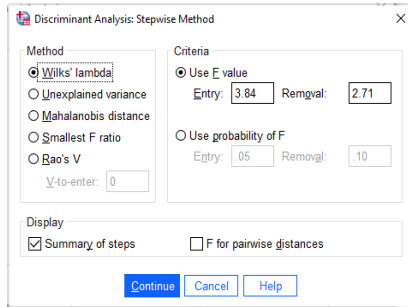
Gambar 4.9 Kotak Dialog Define Range.

- c) Sesuai kode untuk variabel Rekomendasi, maka masukkan angka 0 (nol) pada Minimum dan masukkan angka 1 (satu) pada Maximum.
- d) Tekan tombol Continue untuk kembali ke kotak dialog utama.
- e) Masukkan variabel Jam_Praktikum, Usia, dan Nilai_Praktikum ke dalam kotak Independent. Hal ini berarti, ketiga variabel tersebut berfungsi sebagai variabel bebas (independen).
- f) Pastikan Use Stepwise Method sudah ditandai lanjut.
- g) Klik ikon Statistics. Tampak di layar seperti Gambar 4.10.



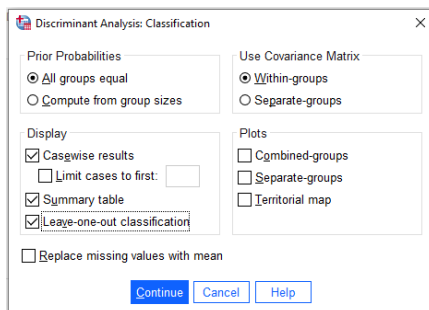
Gambar 4.10 Kotak Dialog Statistics.

- h) Pada bagian Descriptives, aktifkan pilihan Mean, Univariate ANOVAs dan Box's M. serta pastikan di bagian Function Coefficients untuk Unstandardized sudah tertandai. Abaikan bagian lain dan tekan Continue untuk kembali ke kotak dialog utama.
- i) Klik ikon Method. Tampak di layar seperti Gambar 4.11.



Gambar 4.11 Kotak Dialog Method.

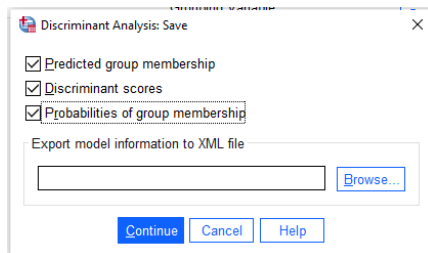
- 4) Pengisian:
- Pada bagian Method, aktifkan pilihan Wilks' Lambda.
 - Di bagian Display, aktifkan Summary of steps dan bagian lain biarkan *default*, dan tekan Continue untuk kembali ke kotak dialog utama.
 - Klik ikon Classify Tampak di layar seperti Gambar 4.12.



Gambar 4.12 Kotak Dialog Classif

5) Pengisian:

- a) Pada bagian Display, aktifkan pilihan Casewise Results (pilihan ini untuk membandingkan hasil kasus awal dengan model diskriminan), Summary Table, dan Leave-One-Out Classification (pilihan ini untuk menampilkan data yang cocok dengan hasil proses diskriminan, dan mana yang tidak cocok/*misclassified*).
- b) Biarkan bagian lain biarkan *default*, dan tekan Continue untuk kembali ke kotak dialog utama.
- c) Klik ikon Save. Tampak di layar seperti Gambar 4.13.



Gambar 4.13 Kotak Dialog Save.

- d) Beri tanda semua bagian dalam kotak dialog Save, tekan Continue untuk kembali ke kotak dialog utama.
- e) Kemudian tekan OK dan lanjutkan dengan cek *output* untuk interpretasi hasil.

c. Hasil Pengujian Variabel Bebas

Tests of Equality of Group Means					
	Wilks' Lambda	F	df1	df2	Sig.
Jam_Praktikum	.151	157.604	1	28	<.001
Nilai_Praktikum	.086	297.897	1	28	<.001
Usia	.969	.892	1	28	.353

Gambar 4.14. Output Test of Equality Group of Means.

Tabel di atas adalah hasil pengujian untuk setiap variabel bebas yang ada. Keputusan bisa diambil lewat dua cara:

1) Dengan angka Wilk's Lambda

Angka Wilk's Lambda berkisar 0 sampai 1. Jika angka mendekati 0 maka data tiap grup cenderung berbeda, sedang jika angka mendekati 1, data tiap grup cenderung sama. Dari tabel, terlihat angka Wilk's Lambda berkisar antara 0,086 sampai 0,969 (mendekati 1). Dari kolom Sig. bisa dilihat bahwa hanya variabel usia yang cenderung tidak berbeda. Hal ini berarti usia untuk mereka yang direkomendasikan belum kompeten atau kompeten ternyata tidak berbeda secara nyata. Untuk itu, pengujian dengan ANOVA, yang dibahas di bawah ini, lebih mudah dilakukan.

2) Dengan F test

Lihat angka Sig.

Jika Sig. > 0,05 berarti tidak ada perbedaan antar-grup.

Jika Sig. < 0,05 berarti ada perbedaan antar-grup.

Dari output pada Gambar 14. dapat dianalisis:

- a. Variabel jam praktikum, angka Sig. adalah di bawah 0,05 (<0,001). Hal ini berarti ada jam praktikum memengaruhi hasil rekomendasi belum

kompeten dan kompeten dalam pelaksanaan uji kompetensi responden tersebut.

- b. Variabel nilai praktikum, angka Sig. adalah di bawah 0,05 ($<0,001$). Hal ini berarti ada nilai praktikum memengaruhi hasil rekomendasi belum kompeten dan kompeten dalam pelaksanaan uji kompetensi responden tersebut.
- c. Usia angka Sig. adalah jauh di atas 0,05 (0,353). Hal ini berarti, usia peserta didik tidak memengaruhi rekomendasi pelaksanaan uji kompetensi. Kesimpulan ini sama dengan jika berpatokan pada angka Wilk's Lambda yang hampir mendekati 1 untuk variabel Usia.
- d. Pada beberapa analisis diskriminan, sebuah variabel yang tidak lolos uji tidak otomatis dikeluarkan. Seperti pada kasus di atas, variabel Usia walaupun tidak lolos uji, tetapi seharusnya tetap disertakan pada analisis diskriminan selanjutnya. Pandangan ini berdasar pada prinsip bahwa pada analisis multivariat, variabel-variabel dianggap suatu kesatuan, dan bukannya terpisah-pisah.
- e. *Hasil Pengujian Varian Tiap Variabel*

Test Results		
Box's M		2.835
F	Approx.	.872
	df1	3
	df2	141120.000
	Sig.	.455
Tests null hypothesis of equal population covariance matrices.		

Gambar 4.15. *Output Test Results Box's M.*

Jika analisis ANOVA dan angka Wilk's Lambda menguji Means (rata-rata) dari setiap variabel, maka Box's M menguji varian dari setiap variabel. Analisis diskriminan mempunyai asumsi bahwa:

- a) Varian variabel bebas untuk tiap grup seharusnya sama. Jika demikian, seharusnya varian dari responden yang direkomendasikan belum kompeten sama dengan varian dari responden yang direkomendasikan Kompeten
- b) Varian di antara variabel-variabel bebas seharusnya juga sama. Jika demikian, seharusnya varian dari Jam Praktikum sama dengan varian dari Nilai Praktikum dan sama dengan variabel Usia. Kedua pengertian di atas bisa disimpulkan, seharusnya *group covariance matrices* adalah relatif sama, yang diuji dengan alat Box's M dengan ketentuan:

- 1) Hipotesis

H_0 : *group covariance matrices* adalah relatif sama.

H_1 : *group covariance matrices* adalah berbeda secara nyata.

- 2) Keputusan dengan dasar signifikansi (lihat angka Sig.)

Jika Sig. > 0,05 berarti H_0 diterima.

Jika Sig. < 0,05 berarti H_0 ditolak.

Dari tabel terlihat bahwa angka Sig. jauh di atas 0,05 (0,455) yang berarti *group covariance matrices* adalah sama. Hal ini berarti data di atas sudah memenuhi asumsi analisis diskriminan, sehingga proses bisa dilanjutkan. Sama tidaknya *group covariance matrices* juga bisa dilihat dari tabel *output Log Determinants* berikut (ada di atas tabel Box's M) seperti *output* pada Gambar 4.16.

Log Determinants		
Rekomendasi	Rank	Log Determinant
0	2	1.323
1	2	1.778
Pooled within-groups	2	1.652

The ranks and natural logarithms of determinants printed are those of the group covariance matrices.

Gambar 4.16 *Output Log Deteminants.*

Terlihat angka Log Determinants untuk kategori belum kompeten (1,323) dan kompeten (1,778) tidak berbeda banyak, sehingga *group covariance matrices* akan relatif sama untuk kedua group. Dalam kondisi *group covariance matrices* berbeda secara nyata, maka:

1. Jika hal ini terjadi pada variabel dependen yang hanya ada dua kemungkinan, seperti pada kasus di atas (hanya ada kategori belum kompeten dan kompeten), maka proses lanjutan seharusnya tidak bisa dilakukan;
2. Namun, jika hal itu terjadi pada kasus dengan variabel dependen dengan kategori banyak (misal ada 5 kategori: Sangat Sedikit, Sedikit, Cukup Banyak, Banyak, Sangat Banyak), maka dengan melihat angka Log Determinant (lihat penjelasan terdahulu), di mana angka yang paling berbeda (misal tanda - sedang yang lain +) bisa dibuang dan proses uji diulang lagi. Jadi, sebagai contoh variabel “cukup banyak”, mungkin dikeluarkan dan proses diulang lagi. Dengan demikian, bisa saja nanti tinggal tiga kategori, namun asumsi sudah terpenuhi.

h. Hasil Pengujian Multikolinieritas

Pooled Within-Groups Matrices				
		Jam_Praktiku m	Nilai_Praktiku m	Usia
Correlation	Jam_Praktikum	1.000	.040	-.140
	Nilai_Praktikum	.040	1.000	.342
	Usia	-.140	.342	1.000

Gambar 4.17 *Output Pooled Within-Group Matrice.*

Lihat nilai korelasi, apabila ada korelasi antar variabel independen dengan nilai lebih dar 0,5, maka dicurigai ada gejala multikolinieritas. Hasil dari matriks korelasi di atas, tidak ada angka yang mencapai 0,5 atau di atasnya, sehingga bisa diidentifikasi tidak ada multikolinieritas data.

4. Menganalisis *Output Determinan*

Group Statistics					
Rekomendasi		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
0	Jam_Praktikum	5.87	1.767	15	15.000
	Nilai_Praktikum	81.60	1.121	15	15.000
	Usia	17.20	1.740	15	15.000
1	Jam_Praktikum	13.67	1.633	15	15.000
	Nilai_Praktikum	90.07	1.534	15	15.000
	Usia	17.80	1.740	15	15.000
Total	Jam_Praktikum	9.77	4.305	30	30.000
	Nilai_Praktikum	85.83	4.504	30	30.000
	Usia	17.50	1.737	30	30.000

Gambar 4.18 *Output Group Statistics.*

Tabel *Group Statistics* pada dasarnya berisi data statistik (deskriptif) yang utama, yakni rata-rata dan standar deviasi, dari kedua kelompok peserta

didik. Sebagai contoh, peserta didik yang direkomendasikan belum kompeten, telah melaksanakan jam praktikum rata-rata 5,87 jam dalam satu minggu. Sementara itu, peserta didik yang direkomendasikan kompeten telah melaksanakan jam praktikum selama 13,67 jam dalam satu minggu. Kedua angka ini tentu berbeda, tetapi apakah perbedaan tersebut nyata (signifikan)?

Demikian pula untuk variabel lainnya, semua mempunyai angka rata-rata dan standar deviasi yang berbeda untuk kedua kelompok peserta didik. Seluruhnya akan diuji untuk mengetahui variabel mana yang mempunyai perbedaan yang signifikan. Dari tabel di atas juga terlihat ada 15 responden yang direkomendasikan belum kompeten, sedangkan 15 responden lainnya direkomendasikan kompeten. Jika melihat semua variabel (jam praktikum, nilai praktikum dan usia) terisi angka 15 dan 15, maka pada kasus ini tidak ada data yang hilang (*missing*) sehingga total data untuk semua variabel adalah 30.

Variables Entered/Removed ^{a,b,c,d}									
Step	Entered	Statistic	df1	df2	Wilks' Lambda		Exact F		Sig.
					df3	Statistic	df1	df2	
1	Nilai_Praktikum	.086	1	1	28.000	297.897	1	28.000	.000
2	Jam_Praktikum	.060	2	1	28.000	211.678	2	27.000	.000

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a. Maximum number of steps is 6.
 b. Minimum partial F to enter is 3.84.
 c. Maximum partial F to remove is 2.71.
 d. F level, tolerance, or VIN insufficient for further computation.

Gambar 4.19 Output Variable Entered/Removed.

Tabel ini menyajikan variabel mana saja dari tiga variabel *input* yang bisa dimasukkan (*entered*) dalam persamaan diskriminan. Pada dua variabel pada Gambar 19, tentunya mempunyai angka Sig. di bawah 0,05 seperti angka Sig. variabel Nilai Praktikum adalah 0,000, jauh di bawah 0,05.

Dengan demikian, dari tiga variabel yang dimasukkan, hanya ada dua

variabel yang signifikan. Atau bisa dikatakan nilai praktikum dan jam praktikum memengaruhi rekomendasi pelaksanaan uji kompetensi (belum kompeten atau kompeten). Berbagai kemungkinan lain akan diuraikan pada analisis selanjutnya.

Eigenvalues				
Function	Eigenvalue	% of Variance	Cumulative %	Canonical Correlation
1	15.680 ^a	100.0	100.0	.970

a. First 1 canonical discriminant functions were used in the analysis.

Gambar 4.20 Output Eigenvalues.

Canonical correlation mengukur keeratan hubungan antara *discriminant score* dengan grup (dalam hal ini, karena ada dua rekomendasi UK, maka ada dua grup). Angka 0,970 menunjukkan keeratan yang cukup tinggi, dengan ukuran skala asosiasi antara 0 sampai 1. Kuadrat dari nilai korelasi kanonikal ($0,97^2 = 0,9401$) menyatakan kemampuan persamaan diskriminan yang melibatkan variabel bebas (jam praktikum dan nilai praktikum) dalam menjelaskan varian variabel tak bebas sebesar 94,01% sedangkan sisanya 5,99% dijelaskan oleh faktor lain.

Structure Matrix	
	Function 1
Nilai_Praktikum	.824
Jam_Praktikum	.599
Usia ^a	.195

Pooled within-groups correlations between discriminating variables and standardized canonical discriminant functions
Variables ordered by absolute size of correlation within function.

a. This variable not used in the analysis.

Gambar 4.21 Output Structure Matrix.

Gambar 4.21 menjelaskan korelasi antara variabel independen dengan fungsi diskriminan yang terbentuk. Terlihat variabel Nilai Praktikum paling erat hubungannya dengan fungsi diskriminan, kemudian jam praktikum dan usia. Hanya di sini variabel Usia tidak dimasukkan dalam model diskriminan (perhatikan tanda huruf ^(a) di dekat variabel tersebut).

Canonical Discriminant Function Coefficients	
	Function 1
Jam_Praktikum	.333
Nilai_Praktikum	.596
(Constant)	-54.452
Unstandardized coefficients	

Gambar 4.22 Output Canonical Discriminant Function.

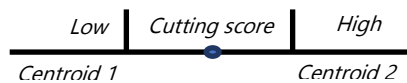
Tabel di atas mempunyai fungsi yang hampir mirip dengan persamaan regresi berganda, yang dalam analisis diskriminan disebut sebagai fungsi diskriminan.

$$\hat{D} = -54,452 + 0,333X_1 + 0,596X_2$$

Kegunaan fungsi ini untuk mengetahui sebuah *case* (dalam kasus ini adalah seorang peserta didik) masuk pada kelompok yang direkomendasikan belum kompeten atau kompeten.

5. Pembuatan Cut Off Score (Nilai Batas).

Dalam fungsi diskriminan 2 kelompok, *cutting score* digunakan untuk mengklasifikasikan pengamatan ke dalam masing-masing kelompok.



Gambar 23. Analogi *Cutting Score*

Dari tabel *Prior Probabilities for Groups*, didapat bahwa jumlah responden belum kompeten adalah 15 orang, sedangkan responden kompeten adalah 15 orang. Dengan demikian, dikaitkan dengan angka grup *centroid*:

$$(15 \times -3,826) + (15 \times 3,826) = -57,39 + 57,39 \text{ atau praktis sama dengan } 0.$$

Functions at Group Centroids	
Rekomendasi	Function
0	-3.826
1	3.826

Unstandardized canonical discriminant functions evaluated at group means

Gambar 4.23 *Output Function at group centroids.*

Z_{CU} = Angka kritis, yang berfungsi sebagai *cut off score*

n_1 dan n_2 = Jumlah sampel di kelompok 1 dan kelompok 2

c_1 dan c_2 = Angka *centroid* pada kelompok 1 dan 2

Casewise Statistics											
Original	Case Number	Actual Group	Predicted Group	Highest Group			Squared Mahalanobis Distance to Centroid	Second Highest Group		Discriminant Scores	
				p	df	P(O=g D=d)		Group	P(O=g D=d)		Squared Mahalanobis Distance to Centroid
	1	0	0	.564	1	1,000	.332	1	.000	67,690	-4,402
	2	0	0	.724	1	1,000	.125	1	.000	53,253	-3,472

Gambar 4.24 *Output Casewise Statistics.*

Tabel *Casewise* pada prinsipnya ingin menguji apakah model diskriminan yang terbentuk akan mengelompokkan dengan tepat seorang peserta UK yang direkomendasikan belum kompeten atau kompeten. sebagai contoh dari *output* pada Gambar 23, Peserta didik 1 mempunyai *score* (-

$$(n_1 c_0 + n_2 c_1) \quad (15)(-3,826) + (15)(3,826)$$

4.402). Karena $(-4.402) < 0$, maka Peserta didik 1 masuk pada kelompok 0 (belum kompeten).

6. Mengukur Ketepatan Klasifikasi

Setelah fungsi diskriminan dibuat, kemudian klasifikasi dilakukan maka selanjutnya akan dilihat seberapa jauh klasifikasi tersebut sudah tepat? Atau berapa persen terjadi kesalahan klasifikasi pada proses klasifikasi tersebut, yang akan dijelaskan berikut ini.

Classification Results ^{a,c}					
		Predicted Group Membership			Total
		Rekomendasi	0	1	
Original	Count	0	15	0	15
		1	0	15	15
	%	0	100.0	.0	100.0
		1	.0	100.0	100.0
Cross-validated ^b	Count	0	15	0	15
		1	0	15	15
	%	0	100.0	.0	100.0
		1	.0	100.0	100.0

a. 100.0% of original grouped cases correctly classified.
b. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.
c. 100.0% of cross-validated grouped cases correctly classified.

Gambar 4.25 *Output Classification Results.*

Pada bagian Original, terlihat bahwa peserta didik yang pada data awal masuk kelompok belum kompeten, dan dari klasifikasi Fungsi diskriminan tetap pada kelompok belum kompeten, adalah 15 orang. Sedangkan dengan model diskriminan, mereka yang awalnya masuk kelompok kompeten juga tetap 15 orang.

Dengan demikian, ketepatan prediksi dari model adalah:

$$(15+15)/30=1,00 \text{ atau } 100\%.$$

Karena angka ketepatan tinggi (100%) maka model diskriminan di atas sebenarnya bisa digunakan untuk analisis diskriminan. Atau penafsiran tentang berbagai tabel yang ada (lihat seluruh pembahasan di atas) valid untuk digunakan. Setelah terbukti bahwa fungsi diskriminan mempunyai ketepatan prediksi yang tinggi, maka fungsi diskriminan tersebut bisa digunakan untuk memprediksi sebuah kasus, apakah akan diklasifikasikan rekomendasi belum kompeten ataukah kompeten.

Contoh Prediksi 1

Boba memiliki karakteristik sebagai berikut: jam praktikum 7 jam; nilai praktikum 90 dan usia 17. Bagaimana prediksi rekomendasi hasil UK Boba?

$$\begin{aligned}\hat{D} &= -54,452+0,333X_1+0,596X_2 \\ &= -54,452+0,333(7)+0,596(90) \\ &= 1,519\end{aligned}$$

Karena $(1.519) > 0$, maka Boba masuk pada kelompok 1 yang diprediksi akan direkomendasikan kompeten.

Contoh Prediksi 2

Babo memiliki karakteristik sebagai berikut: jam praktikum 10 jam; nilai praktikum 85 dan usia 19. Bagaimana prediksi rekomendasi hasil UK Boba?

$$\begin{aligned}\hat{D} &= -54,452+0,333X_1+0,596X_2 \\ &= -54,452+0,333(10)+0,596(85) \\ &= -0,462\end{aligned}$$

Karena $(-0.462) < 0$, maka Babo masuk pada kelompok 0 yang diprediksi akan direkomendasikan belum kompeten.

Daftar Rujukan

- Huberty, C.J., dan Olejnik, S. (2006). *Applied MANOVA and discriminant analysis*. John Wiley & Sons.
- Lekshmi, S., Rugmini, P., dan Thomas, J. (1998). Characteristics of Defaulters in Agricultural Credit Use: A Micro Level Analysis with Reference to Kerala. *Indian Journal of Agricultural Economics*, 53, 640–647.
- Malhotra K. Naresh, (2007). *Marketing Research An Applied Orientation*. Ney Jersey. Prentice Hall International, Inc.
- Mattjik, A.A., dkk. (2004). *Modul Teori Pelatihan Analisis Multivariat*. Bandung: Departemen Statistika. FMIPA IPB
- Santoso, S. (2002). *Buku latihan SPSS Statistika Multivariat*. PT Elex Media Komptindo. Jakarta.
- Sobat Ilmiah. *Analisis Diskriminan dengan SPSS*. Diakses pada 11 Oktober 2021 dari <https://www.youtube.com/watch?v=zLPXaZfh--Y> diakses 11 Oktober 2021
- Ulil Azmi. *Analisis Diskriminan dengan SPSS, Analisis Data Finansial*. Diakses pada 11 Oktober 2021 dari <https://www.youtube.com/watch?v=PFL5c01GfvA>.
- Heri Retnawati. *Analisis Diskriminan*. Diakses pada 11 Oktober 2021 dari <https://www.youtube.com/watch?v=jntsVQDEuj8>.

BAB 5

KORELASI KANONIKAL

A. Pengertian CCA

Analisis korelasi kanonik (*canonical corelation analysis*) merupakan salah satu analisis multivariat yang pertama kali diperkenalkan oleh Harold Hotelling pada tahun 1936 sebagai suatu teknik statistika peubah ganda (multivariat) yang menyelidiki keeratan hubungan antara dua gugus variabel. Gugus yang dimaksud adalah kelompok. Satu gugus variabel diidentifikasi sebagai gugus variabel penduga (*independent variables*), sedangkan gugus variabel lainnya diperlakukan sebagai gugus variabel respon (*dependent variabel*), dan melalui ketergantungan (*dependency*) antar kedua gugus variabel tersebut dapat dijelaskan pengaruh dari satu gugus variabel terhadap gugus variabel lainnya.

Analisis korelasi kanonik adalah suatu teknik analisis statistik yang digunakan untuk melihat hubungan antara segugus variabel independen (X_1, X_2, \dots, X_p) dengan segugus variabel dependen (Y_1, Y_2, \dots, Y_q). Analisis ini dapat mengukur tingkat keeratan hubungan antara segugus variabel dependen dengan segugus variabel independen. Di samping itu, analisis korelasi kanonik juga mampu menguraikan struktur hubungan di dalam gugus variabel independen. Analisis korelasi kanonik berfokus pada korelasi antara kombinasi linier dari gugus variabel dependen dengan kombinasi linier dari gugus variabel independen. Ide utama dari analisis ini adalah menentukan pasangan dari kombinasi linier yang memiliki korelasi terbesar. Kemudian mencari pasangan dari kombinasi linier di antara pasangan yang tidak berkorelasi pada pasangan bagian di awal yang dipilih. Pasangan dari

kombinasi linier ini disebut fungsi kanonik dan korelasinya disebut korelasi kanonik.

Analisis korelasi kanonik merupakan analisis regresi ganda dengan q buah variabel tak bebas dan p buah variabel bebas, yang modelnya adalah sebagai berikut (Johnson & Wichern, 2014):

$$Y_1, Y_2, \dots, Y_q = X_1, X_2, \dots, X_p$$

B. Tujuan CCA

Tujuan dari analisis korelasi kanonik adalah mengukur tingkat keeratan hubungan antara segugus peubah tak bebas dengan segugus peubah bebas dan menguraikan struktur hubungan di dalam gugus peubah tak bebas maupun dalam gugus peubah bebas (Mattjik et al., 2011). Analisis korelasi kanonik berfokus pada korelasi antara kombinasi linear dari gugus peubah tak bebas dengan kombinasi linier dari gugus peubah bebas. Ide utama dari analisis ini adalah mencari pasangan dari kombinasi linier ini yang memiliki korelasi terbesar. Pasangan dari kombinasi linear ini disebut fungsi/peubah kanonik dan korelasinya disebut korelasi kanonik (Safitri & Indrasari, 2009).

Tujuan dari korelasi kanonikal secara dasar sama dengan korelasi sederhana atau berganda, yakni ingin mengetahui apakah ada hubungan (asosiasi) antara dua variabel ataukah tidak. Namun berbeda dengan korelasi sederhana, pada korelasi kanonik jumlah variabel dependen dan variabel independen lebih dari satu sehingga alat analisis korelasi kanonik bisa digolongkan pada multivariate. Misalnya ingin mengukur suatu hubungan linear antara himpunan variabel dependen $y_1, y_2, y_3, \dots, y_p$ yang dinotasikan dengan vektor acak y , dengan himpunan variabel independen $x_1, x_2, x_3, \dots, x_q$ yang dinotasikan dengan vektor acak x , dimana $p \leq q$. dimana y dan x adalah

nilai nilai dari himpunan variabel dependen dan independen untuk satuan pengamatan khusus.

C. Jenis-Jenis Data yang Digunakan dalam CCA

Semua data untuk analisis korelasi kanonik bertipe metrik, yakni data interval atau data rasio (Sudjana, 2002). Dengan demikian data bertipe nominal (seperti jenis kelamin) atau data bertipe ordinal sebaiknya tidak diproses dengan korelasi kanonikal (Nugroho, 2008). Apabila variabel dependennya adalah $y_1, y_2, y_3, \dots, y_p$, dan variabel independennya adalah $x_1, x_2, x_3, \dots, x_q$ maka data hasil pengamatan untuk keadaan ini adalah seperti matriks data berikut.

Tabel 4.3 *Matriks Data Canonical Correlation Analysis.*

Objek (Responden)	Variabel Independen	Variabel Dependen
1	$x_{11}, x_{21}, x_{31}, \dots, x_{q1}$	$y_{11}, y_{21}, y_{31}, \dots, y_{p1}$
2	$x_{12}, x_{22}, x_{32}, \dots, x_{q2}$	$y_{12}, y_{22}, y_{32}, \dots, y_{p2}$
.	.	.
.	.	.
n	$x_{1n}, x_{2n}, x_{3n}, \dots, x_{qn}$	$y_{1n}, y_{2n}, y_{3n}, \dots, y_{pn}$

D. Asumsi-Asumsi yang Harus Dipenuhi pada CCA

Data analisis korelasi kanonik ada beberapa asumsi yang harus dipenuhi yaitu: Ada hubungan yang bersifat linear (linieritas) antar dua variabel; Perlunya *Multivariate Normality* untuk menguji signifikansi setiap fungsi kanonik; Tidak ada Multikolinieritas antar anggota kelompok variabel, baik variabel dependen maupun variabel independen (Gudono, 2016).

1. Linieritas

Linieritas, yaitu hubungan antara himpunan variabel independen x

dengan variabel dependen y bersifat linier. Linieritas dapat dikatakan penting untuk analisis korelasi kanonik dan memengaruhi dua aspek hasil korelasi kanonik. Pertama, koefisien korelasi kanonik antara sepasang variabel kanonik adalah berdasarkan hubungan linier. Jika variabel yang berhubungan tidak linier, maka hubungan tidak akan dapat dijelaskan oleh koefisien korelasi kanonik. Kedua, analisis korelasi kanonik memaksimalkan hubungan linier antar himpunan variabel. Pengujian linieritas dapat dilihat dari tabel ANOVA. Apabila hasil uji antara variabel-variabel independen dengan variabel-variabel dependen memiliki nilai $p < 0,05$, maka model berbentuk linear.

2. Variabel Independen dan Dependen Berdistribusi Normal Multivariat

Uji normalitas pada analisis korelasi kanonik dilakukan dengan menguji normalitas terhadap setiap variabel. Dengan pengujian normalitas terhadap setiap variabel diasumsikan bahwa variabel-variabelnya telah normal secara bersama-sama. Terdapat dua cara yang digunakan dalam mengecek asumsi normal multivariat. Pertama, memeriksa asumsi kenormalan dengan membuat plot Chi Square (untuk $p \geq 2$). Langkah-langkahnya adalah sebagai berikut :

- a. Menghitung nilai $d_j^2 = (x_j - \bar{x})^2 s^{-1} (x_j - \bar{x}), j = 1, 2, \dots, n$.
- b. Mengurutkan d_j^2 sesuai dengan urutan naik $d_1^2 \leq d_2^2 \leq \dots \leq d_n^2$
- c. Plot pasangan, dengan adalah kuantil $(q_{c,p} \left(\frac{j-\frac{1}{2}}{n} \right), d_j^2)$ Chi Square $q_{c,p} \left(\frac{j-\frac{1}{2}}{n} \right)$; derajat bebas p , jika hasil plot berpuncak, maka dapat diasumsikan berdistribusi normal multivariate.

Kemudian yang kedua adalah dengan melihat banyaknya nilai $d_j^2, j = 1, 2, \dots, n$ dan kemudian membandingkannya dengan nilai kuantil χ^2 . Apabila terdapat setengah atau lebih nilai $d_j^2 \leq q_{c,p} (0,50)$, maka dapat dikatakan data

berdistribusi normal multivariate.

3. Homoskedastisitas

Homoskedastisitas menggambarkan data dimana varian dari error (e) tampak konstan melewati batas nilai dari variabel independen. Analisis korelasi kanonik menggambarkan hubungan yang baik ketika homoskedastik. Homoskedastisitas dikatakan penting karena berlawanan dengan heteroskedastisitas, dimana heteroskedastisitas merupakan korelasi antar variabel.

4. Multikolinieritas

Multikolinieritas berhubungan dengan situasi dimana ada hubungan linear baik yang pasti atau mendekati pasti diantara variabel independen. Multikolinieritas terjadi ketika beberapa variabel independen mempunyai korelasi yang tinggi dengan variabel independen yang lain. Uji Multikolinieritas bertujuan untuk menguji apakah pada model yang terbentuk ditemukan adanya korelasi antar variabel bebas. Pengujian multikolinieritas dapat dilakukan dengan dua cara, yaitu dengan melihat nilai Tolerance dan nilai VIF (*Variance Inflation Factor*). Nilai tolerance yang lebih besar dari 0,1 menunjukkan bahwa tidak terjadi multikolinieritas terhadap data yang diuji. Apabila nilai tolerance lebih kecil dari 0,1 berarti terjadi multikolinieritas pada data yang diuji. Nilai VIF yang lebih kecil dari 10 menunjukkan bahwa tidak terjadi multikolinieritas terhadap data yang diuji. Apabila nilai VIF lebih besar dari 10 berarti terjadi multikolinieritas terhadap data yang diuji.

E. Proses Dasar dari CCA

Proses korelasi knonikal, yaitu

1. Menentukan mana yang termasuk dalam kumpulan variabel dependen (*set of multiple dependent variabel*) dan mana yang termasuk dalam kumpulan variabel independen (*set of multiple independent variabel*).
2. Menentukan beberapa *cannonical functions*, yakni korelasi antara set variabel dependen dengan set variabel independen.
3. Dari beberapa *cannonical functions* yang terbentuk, akan diuji *cannonical functions* yang mana yang dapat digunakan. Pengujian dilakukan dengan uji signifikan, *cannonical relationship* serta *redudancy index*
4. Dari *cannonical functions* yang digunakan, dilakukan interpretasi hasil dengan menggunakan beberapa metode, seperti *cannonical weights*, *cannonical loading* atau *cross cannonical loading*. Pengertiannya adalah sebagai berikut:
 - a. *Canonical Weights* (Bobot Kanonik)

Bobot kanonik menggambarkan besarnya kontribusi peubah asal dalam peubah kanoniknya dalam satu kumpulan. Peubah yang memiliki angka koefisien yang besar maka memberikan kontribusi lebih pada peubah kanoniknya, begitu pula sebaliknya. Kemudian untuk peubah yang memiliki bobot yang berlawanan tanda, menggambarkan hubungan kebalikan dengan peubah kanonik lainnya dan peubah yang memiliki tanda sama memiliki hubungan langsung atau searah. Bobot kanonik memiliki beberapa kelemahan yang menjadikan jarang digunakan untuk interpretasi fungsi kanonik. Kelemahannya adalah sifat yang hanya menggambarkan besarnya kontribusi peubah-peubah asal terhadap peubah kanoniknya. Kontribusi tersebut dinilai tidak akurat dalam merefleksikan hubungan antarpeubah. Selain itu, nilai ini dikatakan tidak akurat untuk menggambarkan hubungan antar peubah karena

rentan/sensitif terhadap adanya multikolinieritas. Ia sangat tidak stabil dari satu sampel ke sampel lain.

b. *Canonical Loading* (Beban Kanonik)

Beban kanonik juga disebut sebagai korelasi struktur, mengukur korelasi linier yang sederhana antara data observasi pada peubah independen atau dependen dengan kumpulan peubah kanoniknya. Dalam SPSS nilai beban kanonik, dapat dilihat pada korelasi antara peubah dependen maupun peubah independen dengan peubah kanoniknya. Peubah asal yang memiliki nilai beban kanonik besar ($>0,5$) akan dikatakan memiliki peranan besar dalam kumpulan peubahnya. Sedangkan tanda beban kanonik menunjukkan arah hubungannya. Makin besar nilai beban kanonik maka akan makin penting peranan peubah asal tersebut dalam kumpulan peubahnya. Beban kanonik lebih baik dalam menginterpretasikan hubungan antarpeubah dari pada bobot kanonik karena kelemahan-kelemahan yang ada pada bobot kanonik.

c. *Cross Canonical Loading*

Bobot kanonik dan beban kanonik hanya melihat kontribusi dan korelasi terhadap peubah kanoniknya dalam satu kumpulan, sedangkan *cross loading* digunakan untuk melihat korelasi antar peubah asal dalam satu kumpulan dengan peubah kanonik pada kumpulan yang lainnya. Makin besar nilai ini, maka dapat menggambarkan makin erat pula hubungan antara kedua kumpulan.

5. Melakukan validasi atas hasil *output* tersebut. Validasi biasanya dilakukan dengan membagi dua bagian sampel, kemudian membandingkan kedua hasil yang ada. Jika perbedaan hasil kedua sampel tidak besar, bisa dikatakan korelasi kanonikal adalah valid.

F. Uji Signifikansi CCA

Ada dua hipotesis yang akan diujikan dalam analisis korelasi kanonik, yaitu uji korelasi secara keseluruhan dan uji secara sebagian.

1. Uji Korelasi Kanonik Secara Keseluruhan

Hipotesis:

$H_0 : r_{c1} = r_{c2} = \dots = r_{ck} = 0$ (semua korelasi kanonik tidak signifikan)

$H_1 : r_{ci} \neq 0$ (paling tidak ada satu korelasi kanonik signifikan dengan $i = 1, 2, \dots, k$)

$$\text{Statistik Uji : } F = \frac{1 - \Lambda_k^{k/t} \frac{df_2}{df_1}}{\Lambda_k^{k/t} \frac{df_2}{df_1}}$$

$$\text{dengan : } \Lambda_1 = \prod_{i=1}^k (1 - r_i^2); df_1 = pq; df_2 = wt - \frac{1}{2}pq + 1;$$

$$w = n - \frac{1}{2}(p + q + 3); t = \sqrt{\frac{p^2 q^2 - 4}{p^2 + q^2 - 5}}$$

Keterangan:

n = jumlah pengamatan

p = banyak himpunan variabel y

q = banyak himpunan variabel x

Daerah penolakan : H_0 ditolak jika $F > F_{\alpha; df_1, df_2}$ atau $\Lambda_1 \leq \Lambda_{\alpha, p, q, n-1, q}$.

2. Uji Secara Sebagian

Hipotesis:

$H_0 : r_{cj} = 0$ (korelasi kanonik tidak signifikan)

$H_1 : r_{cj} \neq 0$ (paling tidak ada satu korelasi kanonik signifikan dengan $j = 1, 2, \dots, k$)

$$\text{Statistik Uji : } F = \frac{1 - \Lambda_j^{j/t} \frac{df_2}{df_1}}{\Lambda_j^{j/t} \frac{df_2}{df_1}}$$

$$\text{dengan : } \Lambda_j = \prod_{i=j}^k (1 - r_i^2); df_1 = (p - j + 1)(q - j + 1)$$

$$df_2 = wt - \frac{1}{2}[(p - j + 1)(q - j + 1)] + 1$$

$$w = n - \frac{1}{2}(p + q + 3); t = \sqrt{\frac{(p-j+1)^2(q-j+1)^2 - 4}{(p-j+1)^2 + (q-j+1)^2 - 5}}$$

Keterangan:

n = jumlah pengamatan

p = banyak himpunan variabel y

q = banyak himpunan variabel x

Daerah penolakan : H_0 ditolak jika $F > F_{\alpha; df1, df2}$ atau $A_1 \leq A_{\alpha; p-j+1, q-k+1, n-j-q}$.

G. Analisis Redundansi

Redundansi merupakan sebuah nilai (ukuran) yang menunjukkan besarnya keragaman yang dapat dijelaskan berdasarkan korelasi antara variabel dependen dan independen dengan variabel kanonik. Besarnya keragaman untuk himpunan y yang diterangkan oleh V_1, V_2, \dots, V_k .

$$R(Y|V) = \frac{\sum_{i=1}^p (R^2_{V_i}|x)}{p}$$

Indeks redundansi y yang diterapkan oleh V_1, V_2, \dots, V_k : $RI(y|V) = R(y|V)_{ck2}$.

$$R(x|U) = \frac{\sum_{i=1}^q (R^2_{x_i}|y)}{q}$$

Indeks redundansi x yang diterangkan oleh $U_1, U_2, U_3, \dots, U_k$: $RI(x|U) = R(x|U)_{ck2}$.

H. Prosedur Melakukan CCA dengan SPSS

Langkah-langkah analisis korelasi kanonik dilakukan dengan makro SPSS:

1. Buka aplikasi SPSS.
2. Memasukkan data.
3. Melakukan analisis korelasi kanonik dengan makro SPSS, Yaitu dengan cara klik File lalu New lalu syntax.

4. Memasukkan syntax di bawah ini ke dalam Windows SPSS syntax editor:

```
MANOVA Y1 Y2 WITH X1 X2
```

```
/PRINT = ERROR (SSCP COV COR) SIGNIF
```

```
(HYPOTH EIGEN DIMENR)
```

```
/DISKRIM = RAW STAN ESTIM COR ALPHA (1.0)
```

Y1, Y2, X1, X2 disesuaikan dengan variabel yang dimasukkan ke dalam SPSS.

5. *Run all* untuk mendapatkan *output*

Hasil perhitungan (output SPSS) akan menampilkan:

- a. Matriks korelasi yang terdiri dari:
 - 1) korelasi untuk varian independen,
 - 2) korelasi untuk varian dependen, dan
 - 3) korelasi silang kedua varian.
- b. Nilai eigen (*eigen values*) dan korelasi kanonikal.
- c. Uji signifikan multivariat.
- d. Analisis redundansi.
- e. Bobot kanonikal (*canonical weights*).
- f. Muatan kanonikal (*canonical loading*).
- g. Muatan silang kanonikal (*canonical cross-loading*).

I. Contoh Kasus CCA

Pada suatu studi penelitian, analisis korelasi kanonik digunakan untuk mengidentifikasi dan mengukur tingkat keeratan hubungan linier antara variabel dependen y_1 = kualitas kerja, y_2 = loyalitas, y_3 = efektifitas kerja dan y_4 = tanggung jawab, dengan himpunan variabel independen : x_1 = *on the job training*, x_2 = *off the job training* dan x_3 = *training of trainer*, dan terdiri dari

120 responden. Sajian datanya adalah sebagai berikut.

Tabel 5.1 Contoh Kasus CCA.

No	On The Job Training	Off The Job Training	Training Of Trainer	Kualitas kerja	Loyalitas	Efektifitas Kerja	Tanggung jawab
1	59	70	53	98	83	84	84
2	63	64	52	79	82	52	82
3	75	23	25	62	51	50	59
4	50	38	33	63	59	62	68
5	38	34	36	63	69	65	63
6	30	58	43	67	71	65	70
7	35	62	40	75	67	82	76
8	57	53	27	62	53	40	46
9	39	62	42	80	78	77	78
10	44	69	51	90	74	62	78
11	53	52	39	71	65	80	73
12	36	75	52	82	86	63	87
13	44	61	48	78	85	82	79
14	67	63	47	88	74	52	67
15	40	51	42	79	74	50	61
16	56	45	23	78	49	41	42
17	54	56	33	72	56	43	61
18	38	45	30	80	57	68	58
19	55	48	35	74	49	53	62
20	30	58	45	76	74	76	84
21	17	56	38	93	71	85	80
22	33	75	59	87	96	88	93
23	74	50	37	69	70	68	69
24	18	60	49	71	83	80	82
25	53	72	47	77	87	89	79
26	33	27	38	52	78	65	76
27	51	51	31	63	58	61	64
28	73	50	34	65	71	45	72
29	58	62	52	58	84	71	87
30	97	55	54	80	79	74	80

No	On The Job Training	Off The Job Training	Training Of Trainer	Kualitas kerja	Loyalitas	Efektifitas Kerja	Tanggung jawab
31	31	69	38	78	64	77	67
32	39	70	51	52	76	63	80
33	73	60	44	71	73	66	70
34	73	46	33	68	57	77	57
35	37	42	35	76	72	80	80
36	41	46	35	67	63	40	67
37	31	26	37	75	59	57	60
38	54	66	44	75	71	72	73
39	38	63	43	77	81	57	80
40	41	44	34	56	64	47	53
41	52	46	30	70	60	57	60
42	61	77	53	81	81	68	77
43	41	62	56	82	86	67	86
44	49	77	56	94	80	70	83
45	35	76	56	87	86	80	90
46	54	75	36	58	53	67	59
47	43	40	36	71	69	60	72
48	64	74	48	83	74	87	66
49	55	69	47	82	69	57	80
50	57	72	39	66	64	59	68
51	55	74	46	94	74	56	72
52	52	73	57	96	88	78	86
53	56	52	35	60	59	66	64
54	57	77	35	65	55	78	59
55	61	39	33	65	67	42	61
56	44	72	59	76	87	77	89
57	52	61	53	71	76	72	80
58	67	82	53	77	80	86	72
59	65	65	44	81	76	63	76
60	58	56	24	65	56	69	62
61	50	66	32	52	76	56	75
62	57	70	50	80	81	86	79

No	On The Job Training	Off The Job Training	Training Of Trainer	Kualitas kerja	Loyalitas	Efektifitas Kerja	Tanggung jawab
63	27	47	32	81	55	61	71
64	50	61	42	84	64	67	71
65	38	60	34	87	52	59	58
66	54	66	51	84	80	67	87
67	70	75	50	88	77	80	79
68	27	59	45	69	81	56	65
69	28	76	38	74	64	51	59
70	36	66	50	71	76	83	74
71	32	68	47	75	83	66	72
72	68	67	40	73	67	71	60
73	60	75	28	74	60	66	67
74	45	66	58	81	88	85	84
75	38	59	50	82	77	73	82
76	26	62	42	47	85	57	88
77	67	71	32	61	72	85	61
78	54	52	35	87	67	56	63
79	22	65	41	62	60	70	74
80	47	67	31	63	54	64	55
81	57	52	24	63	57	63	55
82	38	67	43	72	63	59	68
83	39	34	22	73	67	86	64
84	42	65	40	75	74	78	67
85	65	28	14	57	74	72	59
86	34	52	24	71	54	68	48
87	35	21	35	43	66	50	71
88	52	25	41	76	50	57	64
89	51	35	29	57	56	48	57
90	51	65	49	87	80	74	82
91	54	67	35	64	86	85	71
92	49	54	23	81	61	83	66
93	63	60	11	93	74	73	58
94	41	38	20	64	57	46	49

No	On The Job Training	Off The Job Training	Training Of Trainer	Kualitas kerja	Loyalitas	Efektifitas Kerja	Tanggung jawab
95	40	68	27	74	67	74	60
96	37	73	42	77	70	54	63
97	23	23	15	58	38	48	36
98	45	54	26	66	74	60	60
99	44	59	23	76	66	81	72
100	61	63	42	78	60	74	65
101	48	50	25	74	64	61	38
102	49	50	34	56	63	54	44
103	47	70	54	78	90	76	89
104	42	49	37	59	77	66	78
105	40	54	38	66	66	62	54
106	34	57	43	87	77	50	61
107	44	31	28	75	52	30	54
108	29	63	33	80	73	45	53
109	32	59	20	53	54	68	50
110	32	73	30	76	96	89	70
111	26	38	19	63	60	45	42
112	34	34	20	68	52	66	49
113	40	39	21	70	43	61	44
114	41	38	21	81	44	64	44
115	33	66	21	89	34	71	42
116	55	69	46	83	77	83	83
117	33	48	39	68	64	70	61
118	42	77	57	72	85	68	86
119	45	26	19	52	52	59	41
120	55	20	19	70	61	57	45

Penyelesaian

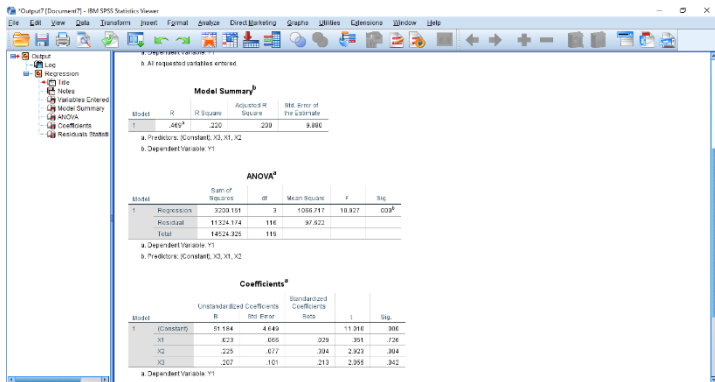
1. Pengujian Persyaratan

Sebelum pengolahan data dengan korelasi kanonikal dimulai, dilakukan uji persyaratan-persyaratan terlebih dahulu yang harus dipenuhi yaitu

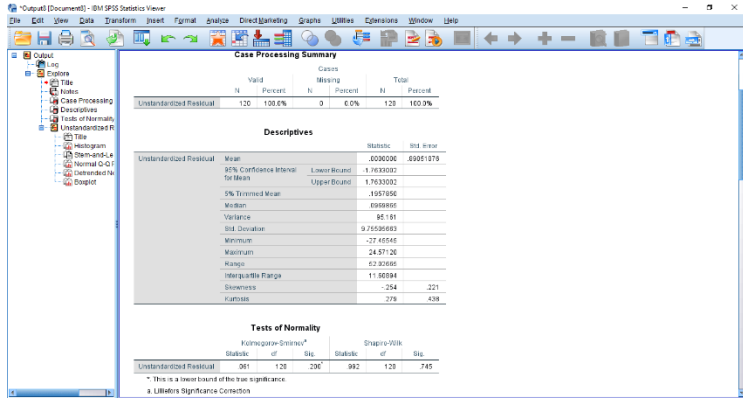
normalitas, linieritas dan multikolinier

a. Uji Normalitas

- 1) Membuat variabel residu dengan cara klik Analyze lalu regression lalu klik Linear.
- 2) Masukkan variabel x_1, x_2, x_3 ke kolom Independen dan y_1 ke kolom dependen, kemudian klik Save lalu akan muncul Linear Regression Save dan centang Residual Unstandardized. Kemudian klik continue lalu klik oke, maka akan muncul hasil seperti di bawah ini.

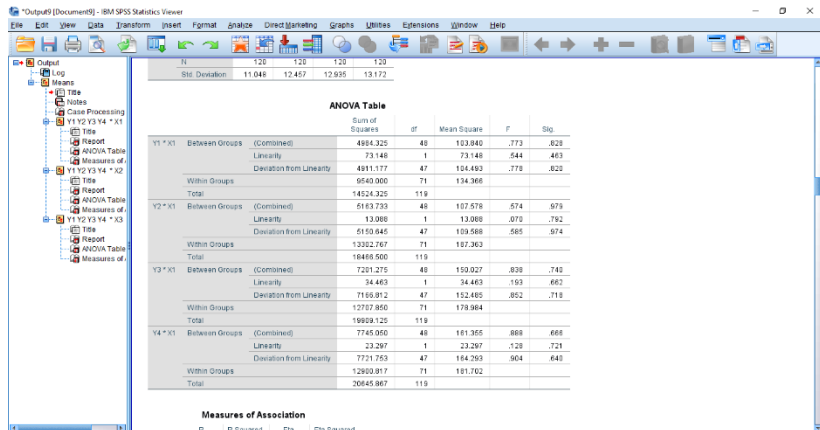


- 3) Uji normalitas dari variabel Residunya dilakukan dengan mengklik Analyze lalu Discriptive Statistics lalu Explore. Kemudian masukkan Residual Unstandardized ke kolom Dependent List, Klik Plots. Setelahnya, centang Histogram dan Normality plots with test, klik Continue lalu klik Oke, maka akan terlihat hasil sebagai berikut.



b. Uji Linieritas

Menguji apakah tiap variabel independen hubungannya linear atau tidak dengan variabel dependen, caranya adalah dengan mengklik Analize lalu Compare Means lalu klik Means, Masukkan variabel x_1, x_2, x_3 ke kolom Independen dan y_1, y_2, y_3, y_4 ke kolom Dependen. Pilih Options, centang test for linearity, kemudian Continue lalu klik Oke, maka akan mendapatkan hasil seperti di bawah ini.



Dari tebal ANOVA tersebut *deviation from linearity* nilai signification nya lebih dari 0.05. Artinya, tidak signifikan, penyimpangan dari *linearity*-nya tidak signifikan berarti asumsi normalitasnya terpenuhi, sehingga yang kita cari adalah *deviation from linearity* diatas 0.05.

c. Uji Multikolinieritas

Menguji multikolinieritas adalah dengan mengkorelasikan variabel independen dengan variabel dependen dengan cara klik Analyze lalu Correlate lalu klik Bivariate, kemudian memasukkan seluruh variabel dependen dan independen ke kolom variabel, klik oke. Maka hasilnya seperti berikut.

	X1	X2	X3	Y1	Y2	Y3	Y4
X1	Pearson Correlation	1	.606	.868	.371	.007	.443
	Sig. (2-tailed)		.328	.453	.444	.723	.652
	N	120	120	120	120	120	120
X2	Pearson Correlation	.606	1	.472 ^{**}	.407 ^{**}	.602 ^{**}	.436 ^{**}
	Sig. (2-tailed)			.889	.389	.000	.000
	N	120	120	120	120	120	120
X3	Pearson Correlation	.868	.472 ^{**}	1	.461 ^{**}	.738 ^{**}	.337 ^{**}
	Sig. (2-tailed)				.389	.000	.000
	N	120	120	120	120	120	120
Y1	Pearson Correlation	.007	.437 ^{**}	.407 ^{**}	1	.263 ^{**}	.337 ^{**}
	Sig. (2-tailed)		.441	.600		.000	.000
	N	120	120	120	120	120	120
Y2	Pearson Correlation	.007	.602 ^{**}	.738 ^{**}	.263 ^{**}	1	.458 ^{**}
	Sig. (2-tailed)		.773	.600	.888		.000
	N	120	120	120	120	120	120
Y3	Pearson Correlation	.443	.437 ^{**}	.337 ^{**}	.337 ^{**}	.458 ^{**}	1
	Sig. (2-tailed)		.652	.600	.888	.331	
	N	120	120	120	120	120	120
Y4	Pearson Correlation	.436	.437 ^{**}	.337 ^{**}	.337 ^{**}	.458 ^{**}	.436 ^{**}
	Sig. (2-tailed)		.716	.600	.888	.388	.000
	N	120	120	120	120	120	120

** Correlation is significant at the 0.01 level (2-tailed).

Variabel dikatakan multikolinieritas jika nilai R nya diatas 0.8, sedangkan nilai R pada hasil analisis di atas, di bawah nilai 0.8 sehingga tidak terjadi multikolinieritas.

2. Pengolahan Data

Dengan terpenuhinya semua persyaratan untuk uji korelasi kanonikal maka analisis data bisa dimulai. Pengolahan data korelasi kanonikal menggunakan komputer program SPSS versi 24. Langkah-langkah untuk

analisis korelasi kanonikal dengan *software SPSS* sebagai berikut.

- a. Klik Worksheet SPSS. *Input* data yang akan diolah dari Microsoft Excel. Blok data yang akan diolah kemudian klik kanan Copy kemudian masuk ke Worksheet SPSS. Pada kolom satu kita klik kanan kemudian Paste maka akan muncul sebaran data sebagai berikut.

The screenshot shows the SPSS Data Editor window with a data matrix. The columns are labeled VAR0001 through VAR0008. The data is as follows:

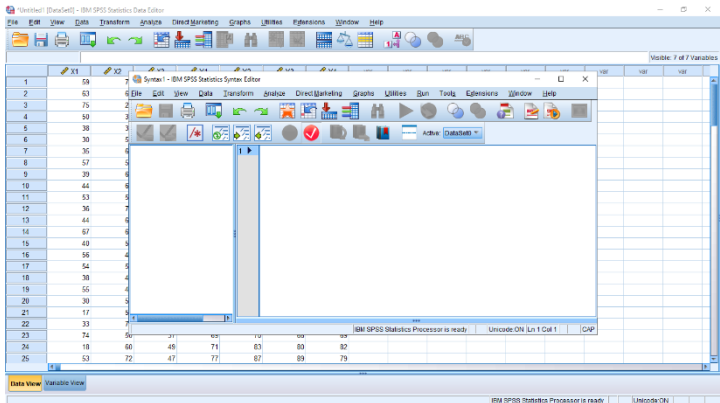
	VAR0001	VAR0001	VAR0001	VAR0001	VAR0001	VAR0001	VAR0001
1	59.00	70.00	53.00	98.00	83.00	84.00	84.00
2	63.00	64.00	52.00	79.00	82.00	52.00	82.00
3	76.00	23.00	26.00	62.00	61.00	59.00	59.00
4	50.00	38.00	33.00	63.00	59.00	62.00	68.00
5	38.00	34.00	36.00	63.00	69.00	65.00	63.00
6	30.00	58.00	43.00	67.00	71.00	65.00	70.00
7	35.00	62.00	40.00	75.00	67.00	82.00	76.00
8	57.00	53.00	27.00	62.00	53.00	40.00	46.00
9	39.00	62.00	42.00	80.00	78.00	77.00	78.00
10	44.00	69.00	51.00	90.00	74.00	82.00	79.00
11	63.00	52.00	39.00	71.00	65.00	80.00	73.00
12	36.00	75.00	52.00	82.00	86.00	63.00	87.00
13	44.00	61.00	48.00	78.00	85.00	82.00	79.00
14	67.00	60.00	47.00	88.00	74.00	52.00	67.00
15	40.00	51.00	42.00	79.00	74.00	50.00	61.00
16	56.00	45.00	23.00	78.00	49.00	41.00	42.00
17	54.00	66.00	33.00	72.00	66.00	43.00	61.00
18	38.00	45.00	30.00	80.00	57.00	68.00	58.00
19	55.00	48.00	35.00	74.00	49.00	53.00	62.00
20	30.00	58.00	45.00	76.00	74.00	76.00	84.00
21	17.00	66.00	38.00	53.00	71.00	65.00	68.00
22	33.00	75.00	59.00	67.00	96.00	88.00	93.00
23	74.00	50.00	37.00	69.00	70.00	68.00	69.00
24	18.00	60.00	49.00	71.00	83.00	89.00	82.00

- b. Pada Worksheet Variable View, nama variabel diubah menjadi X1, X, X3, Y1, Y2, Y3, Y4 dan pada kolom Desimal diubah menjadi “0”, seperti pada gambar berikut.

The screenshot shows the SPSS Variable View window. The variables are defined as follows:

Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	X1	Numeric	8	0	None	None	8	Right	Unknown	Input
2	X2	Numeric	8	0	None	None	8	Right	Unknown	Input
3	X3	Numeric	8	0	None	None	8	Right	Unknown	Input
4	Y1	Numeric	8	0	None	None	8	Right	Unknown	Input
5	Y2	Numeric	8	0	None	None	8	Right	Unknown	Input
6	Y3	Numeric	8	0	None	None	8	Right	Unknown	Input
7	Y4	Numeric	8	0	None	None	8	Right	Unknown	Input
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										
21										
22										
23										
24										
25										
26										

- c. Buat analisis korelasi kanonikal, dengan menu menu File lalu klik New lalu klik Syntax, maka akan muncul kolom dialog seperti berikut.



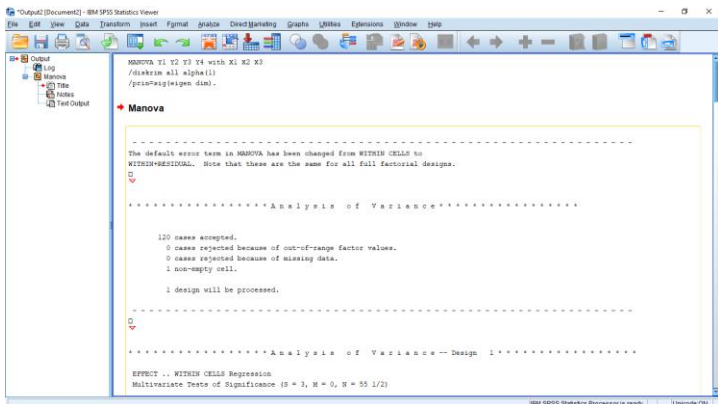
- d. Pada kolom dialog tersebut kita masukkan formula untuk uji analisis kanokal dengan cara mengetik:

MANOVA Y1 Y2 Y3 Y4with X1 X2 X3 (Enter)

/diskrim all alpha(1) (Enter)

/prin=sig (eigen dim).

Kemudian klik kanan pilih Run all. Maka akan muncul hasil seperti di bawah ini.



e. Pengujian secara Individu

Tabel 5.1 Perhitungan untuk penentuan fungsi kanonikal.

Eigenvalues and Canonical Correlations					
Root No.	Eigenvalue	Pct.	Cum. Pct.	Canon Cor.	Sq. Cor
1	2,64170	93,07799	93,07799	,85171	,72540
2	,19492	6,86794	99,94593	,40389	,16313
3	,00153	,05407	100,00000	,03914	,00153

Dimension Reduction Analysis					
Roots	Wilks L.	F	Hypoth. DF	Error DF	Sig. of F
1 TO 3	,22945	18,56323	12,00	299,26	,000
2 TO 3	,83559	3,57062	6,00	228,00	,002
3 TO 3	,99847	,08823	2,00	115,00	,916

Dengan melihat *root* ada tiga kanonik fungsi yaitu fungsi 1 korelasi kanonik 0.85171 dengan signifikansi 0.000, fungsi 2 korelasi kanonik 0.40389 dengan signifikansi 0.002, dan fungsi 3 korelasi kanonik 0.03914 dengan signifikansi 0.916. Dari hasil tersebut terlihat fungsi 1 dan 2 < 0.05 signifikan secara individual, sedangkan fungsi 3 > 0.05 maka tidak signifikan secara individual. Oleh karena itu, fungsi 1 dan fungsi 2 dapat diproses lebih lanjut. Sedangkan fungsi 3 secara individual tidak dapat diproses lebih lanjut.

f. Pengujian Kelompok

Perhitungan secara bersama-sama tampak pada tabel di bawah ini.

Tabel 5.2 Hasil perhitungan secara bersama-sama dengan empat prosedur.

EFFECT .. WITHIN CELLS Regression Multivariate Tests of Significance (S = 3, M = 0, N = 55 1/2)					
Test Name	Value	Approx. F	Hypoth. DF	Error DF	Sig. of F
Pillais	,89006	12,12797	12,00	345,00	,000
Hotellings	2,83816	26,41065	12,00	335,00	,000
Wilks	,22945	18,56323	12,00	299,26	,000
Roys	,72540				

Dengan menggunakan empat prosedur dari Pillais, Hotellings, Wilks, dan Roys, semuanya signifikan karena $< 0,05$. Dengan demikian, jika digabung secara bersama-sama, kanonikal fungsi 1, kanonikal fungsi 2, dan kanonikal fungsi 3 dapat diproses lebih lanjut. Dari hasil pengujian individu dan bersama (kolektif) terdapat perbedaan korelasi kanonik, yaitu terlihat pada table 3, dengan angka korelasi kanonik fungsi 1 = 0.85171, korelasi kanonik fungsi 2 = 0.40389, dan korelasi kanonik fungsi 3 = 0.03914. Oleh karena fungsi 1 memiliki angka korelasi kanonik yang tinggi dan signifikan, baik secara individu maupun kolektif, maka analisis selanjutnya hanya menitik beratkan pada fungsi 1.

g. Interpretasi Kanonikal Varian

Analisis ini merupakan kelanjutan dari pengujian sebelumnya yang menetapkan kanonik fungsi 1. Oleh karena itu, dalam analisis ini hanya memperhatikan kanonikal fungsi 1, tidak memperhatikan fungsi 2 dan fungsi 3. Dalam penelitian ini ada dua kanonik varian, yaitu dependen kanonik varian yang berisi kualitas kerja, loyalitas, efektifitas kerja, dan tanggung jawab dan kanonik independen varian yang berisi *on the job training*, *off the job training*, dan *training of trainer*. Analisis ini berfungsi untuk mengetahui apakah semua variabel independen dalam kanonik varian berhubungan dengan dependen varian yang diukur dengan besaran korelasi masing-masing independen variabel dengan variannya. Pengukuran dilakukan dengan dua cara yaitu kanonikal Weights dan kanonikal Loadings.

a. Canonical Weight I

Tabel 5.3 Hasil perhitungan *canonical weights* untuk *dependen variate*.

 Standardized canonical coefficients for DEPENDENT variables
 Function No.

Variable	1	2	3
Y1	,20447	-,47221	,84613
Y2	,25485	-,37396	-,96253
Y3	-,11670	-,85989	-,37320
Y4	,75310	,94546	,59327

Tabel 5.4 Hasil perhitungan *canonical weights* untuk independen varian.

 Raw canonical coefficients for COVARIATES
 Function No.

COVARIATE	1	2	3
X1	-,00176	-,00454	,07242
X2	,00602	-,08420	-,01028
X3	,08314	,07407	,00467

 Standardized canonical coefficients for COVARIATES
 CAN. VAR.

COVARIATE	1	2	3
X1	-,02432	-,06279	1,00197
X2	,08967	-1,25508	-,15319
X3	,94404	,84105	,05305

Dengan tidak memperhatikan fungsi 2 dan fungsi 3, terlihat deretan angka korelasi antara masing-masing variabelnya dengan variannya. Untuk variabel dependen ada satu angka korelasi yang tinggi ,yaitu 0,7531 (tanggung jawab), sedangkan untuk variabel independen ada satu angka korelasi yang tinggi, yaitu 0,94404 (*training og trainer*). Selain dengan *canonical weights*, interpretasi dilakukan dengan melihat besaran *canonical loading*.

b. Canonical Loading

Tabel 5.5 Hasil perhitungan *canonical loading* untuk dependen varian.

Correlations between COVARIATES and canonical variables
CAN. VAR.

Covariate	1	2	3
X1	,04901	-,11756	,99186
X2	,66585	-,74546	-,03055
X3	,99729	,06778	,02846

Tabel 5.6 Hasil perhitungan *canonical loading* untuk independen varian.

Correlations between DEPENDENT and canonical variables
Function No.

Variable	1	2	3
Y1	,48876	-,53328	,65024
Y2	,86330	-,13972	-,41326
Y3	,42070	-,72098	-,26241
Y4	,96820	,08035	-,07736

Pada Tabel 5.5 dan 5.6, hasil perhitungan *canonical loading* dengan hanya melihat fungsi 1 maka terlihat deretan angka korelasi loading masing-masing variabel dengan variabel variatnya. Untuk dependen variabel, ada dua angka *canonical loading* yang tinggi, yaitu -0,863 (*loyalitas*) dan -0,968 (*tanggung jawab*), sedangkan pada variabel independen ada dua kanonikal *loading*, yaitu 0,665 (*off the job training*) dan 0,997 (*training of trainer*).

h. Hasil Penelitian

Berdasarkan hasil penelitian di atas, maka diperoleh hal-hal sebagai berikut.

- a. Dua variabel dependen dan variabel independen memiliki hubungan yang signifikan. Dengan perkataan lain terdapat hubungan antara kualitas kerja, loyalitas, efektifitas kerja, dan tanggung jawab dengan *on the job training*, *off the job training* dan *training of trainer*, jika dilakukan pengujian secara kelompok.
- b. Dari tiga variabel independen, ada satu variabel yang memiliki hubungan yang sangat tinggi, yaitu *training of trainer*. Dengan demikian, dapat diartikan guru produktif dengan nilai ToT yang tinggi akan memiliki Kinerja yang tinggi pula.

Daftar Rujukan

- Gudono. (2016). *Analisis Data Multibariat*. Edisi 4. Yogyakarta: BPPEE.
- Johnson, R.A., dan Dean W.W. (2014). *Applied Multivariate Statistical Analysis*. Volume 6. UK: Pearson London.
- Mattjik, A.A., I S., Gusti N.A.W, dan Alfian F.H. (2011). *Sidik Peubah Ganda Dengan Menggunakan SAS*.
- Nugoho, S. (2008). *Statistika Multivariat Terapan*. Bengkulu: UNIB Press.
- Safitri D. dan Paramitra I. (2009). Analisis Korelasi Kanonik pada Perilaku Kesehatan dan Karakteristik Sosial Ekonomi di Kota Pati Jawa Tengah. *Media Statistika*, 2(1), 39–48.
- Sudjana. (2002). *Metode Statistik*. Cetakan IV. Bandung: Tarsito.

BAB 6

MULTIVARIATE ANALYSIS OF VARIANCE

(MANOVA)

A. Sejarah dan Pengertian MANOVA

Pearson dan Fisher adalah orang pertama yang memperkenalkan analisis multivariat. Mereka merupakan orang yang memiliki pemahaman paling mendasar dalam aplikasi statistika (Farrell, 1976). Analisis multivariat merupakan metode statistik yang digunakan untuk memahami struktur data dalam dimensi tinggi (Sutrisno dan Wulandari, 2018). Teknik analisis multivariat dikembangkan dari teknik analisis univariat (Simamora, 2005). Secara historis, sebagian besar teknik multivariat dilakukan dalam bidang ilmu perilaku dan biologi (Sutrisno dan Wulandari, 2018). Namun, saat ini analisis multivariat mulai banyak digunakan dalam berbagai bidang ilmu, melengkapi analisis statistik univariat dan statistik bivariat yang awalnya merupakan satu-satunya alternatif dalam analisis data (Santoso, 2010). Teknik analisis multivariat banyak disukai karena dianggap mampu memodelkan kerumitan sistem yang nyata, teknik analisis ini banyak diterapkan dalam bidang pendidikan, kimia, fisika, geologi, teknik, hukum, bisnis, sastra, agama, penyiaran publik, keperawatan, pertambangan, linguistik, psikologi, dan bidang lain (Sutrisno dan Wulandari, 2018).

Selama bertahun-tahun penerapan analisis multivariat terhalang oleh perhitungannya yang terlalu rumit sehingga tidak dapat dijangkau dengan bantuan komputer yang tersedia pada saat itu. Namun, dengan komputer modern, hampir setiap analisis yang diinginkan, tidak peduli berapa banyak variabel atau pengamatan yang terlibat, dapat dengan cepat dan mudah

dilakukan. Penggunaan secara langsung prosedur univariat pada masing-masing variabel terikat untuk data multivariat tidak lagi dimaklumi, mengingat ketersediaan teknik multivariat dan daya komputasi yang tersedia untuk menyelesaikan pekerjaan tersebut (Rencher, 2002). Berbagai macam prosedur deskriptif dan inferensial multivariat mudah diakses di berbagai *software* statistik, seperti SPSS, SYSTAT, S-Plus, R, dan Minitab (Tim, 2002).

Salah satu teknik dalam analisis multivariat, yaitu *multivariate analysis of variance* (MANOVA) atau analisis variansi multivariat. MANOVA dikembangkan sebagai konstruk teoritis oleh S.S. Wilks pada tahun 1932 dan dipublikasikan dalam *Biometrika*. MANOVA merupakan teknik statistik yang digunakan untuk menghitung pengujian signifikansi perbedaan rata-rata secara bersamaan antara kelompok untuk dua atau lebih variabel terikat. MANOVA adalah generalisasi dari ANOVA untuk situasi yang memiliki beberapa variabel terikat (Tabachnick dan Fidell, 2007). Menurut Field (2009), MANOVA memiliki kemiripan asumsi dengan ANOVA tetapi diperluas untuk kasus multivariat. Perbedaan antara ANOVA dan MANOVA terletak pada jumlah variabel dependennya. ANOVA digunakan untuk mengetahui apakah terdapat perbedaan pengaruh perlakuan terhadap satu variabel dependen, sedangkan MANOVA digunakan untuk mengetahui apakah terdapat perbedaan pengaruh terhadap lebih dari satu variabel dependen (Tabachnick dan Fidell, 2007). Berdasarkan beberapa pendapat tersebut, dapat disimpulkan bahwa MANOVA adalah uji statistik yang digunakan untuk mengukur pengaruh variabel independen.

Tabel 6.1 Perbedaan analisis ragam peubah ganda (MANOVA) dengan analisis ragam satu peubah (ANOVA).

ANOVA <i>(Univariate Analysis of Variance)</i>	MANOVA <i>(Multivariate Analysis of Variance)</i>
Hanya mengkaji berbagai pengaruh percobaan yang dilakukan terhadap respons tunggal (satu unit variabel respons).	Mengkaji pengaruh dari berbagai perlakuan yang dicobakan terhadap respons ganda (lebih dari satu respons).
Ketergantungan di antara variabel respons tidak menjadi perhatian utama karena pada dasarnya terdapat anggapan bahwa variabel-variabel respons saling bebas satu sama lain sehingga pengkajian struktur keragaman hanya dilakukan terhadap setiap variabel respons secara terpisah.	Mempertimbangkan adanya ketergantungan antar variabel respons sehingga cocok digunakan untuk pengkajian pengaruh dari berbagai perlakuan terhadap lebih dari satu respons.

Sumber : Mattjik dan Sumertajaya, 2011.

b. Tujuan MANOVA

Tujuan dari MANOVA adalah untuk menguji apakah vektor rerata dua atau lebih grup sampel diambil dari sampel distribusi yang sama. MANOVA biasa digunakan dalam dua kondisi utama. Kondisi pertama adalah saat terdapat beberapa variabel dependen yang berkorelasi, sementara peneliti hanya menginginkan satu kali tes keseluruhan pada kumpulan variabel ini dibandingkan dengan beberapa kali tes individual. Kondisi kedua adalah saat peneliti ingin mengetahui bagaimana variabel independen mempengaruhi pola variabel dependennya (Santoso, 2010).

MANOVA memiliki kelebihan bila dibandingkan ANOVA. Penggunaan MANOVA memiliki keunggulan, yaitu mampu menganalisis semua variabel terikat secara simultan, sehingga dapat memperkecil kesalahan tipe I (α) dalam pengambilan keputusan uji statistik (Steven, 2002). MANOVA mampu mendeteksi dan mengungkapkan perbedaan yang tidak ditampilkan ANOVA pada masing-masing variabel terikat secara terpisah.

MANOVA juga mampu mengoreksi hasil ANOVA palsu yang disebabkan peningkatan alpha saat melakukan beberapa tes ANOVA pada masing-masing variabel terikat (Sutrisno & Wulandari, 2018). Oleh karena itu, dapat disimpulkan dengan menggunakan MANOVA, peneliti dapat meningkatkan kesempatan untuk menemukan perubahan sebagai akibat dari perlakuan yang berbeda dan interaksinya. Dengan demikian, temuan-temuan hasil penelitian akan makin kaya dan sangat berguna bagi perkembangan ilmu pengetahuan. Dalam analisis MANOVA juga dapat diketahui *effect size* nya dan dapat dilakukan uji lanjutan.

Hair (2010) menyatakan “*Multivariate analysis of variance (MANOVA) is an extension of analysis of variance (ANOVA) to accommodate more than one dependent variable. It is a dependence technique that measures the differences for two or more metric dependent variables based on a set of categorical (nonmetric) variables acting as independent variables. ANOVA dan MANOVA can be stated in the following general forms:*

Analysis of Variance

$$Y_1 = X_1 + X_2 + X_3 + \dots + X_n$$

metric nonmetric

Multivariate Analysis of Variance

$$Y_1 + Y_2 + Y_3 + \dots + Y_6 = X_1 + X_2 + X_3 + \dots + X_n$$

metric nonmetric”.

MANOVA menguji ada tidaknya perbedaan rata-rata dari dua atau lebih variabel tak bebas secara simultan (*simultaneously*) berdasarkan kelompok-kelompok pada variabel bebas. Perlu diperhatikan bahwa pada MANOVA, variabel bebas (*independent variable*) bersifat non-metrik (terdiri dari beberapa kelompok/kategori), sedangkan variabel bebas bersifat metrik (interval atau rasio). Field (2009) menyatakan:

“ANOVA can be used only in situations in which there is one dependent variable (or outcome) and so is known as a univariate test (univariate quite obviously means 'one variable'); MANOVA is designed to look at several dependent variables (outcomes) simultaneously and so is a multivariate test (multivariate means 'many variables')”.

Pada kasus multivariat, misal terdapat sekumpulan sampel acak yang diambil dari setiap p populasi yang dilambangkan dengan y sebagai berikut.

$$Y = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1j} & \dots & y_{1p} \\ y_{21} & y_{22} & \dots & y_{2j} & \dots & y_{2p} \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ y_{i1} & y_{i2} & \dots & y_{ij} & \dots & y_{ip} \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ y_{n1} & y_{n2} & \dots & y_{nj} & \dots & y_{np} \end{bmatrix}$$

Terdapat tiga asumsi dasar yang diperlukan oleh sekumpulan sampel acak di atas, yaitu

1. $y_{11}, y_{12}, \dots, y_{1p}$ dengan $(i = 1, 2, 3, \dots, p)$ adalah sampel acak berukuran n_i dari suatu populasi dengan rata - rata μ_i .
2. Matriks kovariani antara p populasi sama.
3. Setiap populasi adalah normal multivariat.

Sebelum dilakukan analisis variansi multivariat lebih lanjut, terlebih dahulu akan diuji ketiga asumsi-asumsi dasar tersebut menyatakan bahwa dari sekumpulan data multivariat $y_{11}, y_{12}, \dots, y_{1p}$ dengan $(i = 1, 2, 3, \dots, p)$, yaitu

sampel acak berukuran n_i yang diambil dari suatu populasi dengan vektor rata-rata μ_i dan saling bebas. Pernyataan ini adalah jelas tanpa perlu diuji karena untuk tujuan uji perbedaan maka sekumpulan data multivariat dari setiap populasi harus diambil secara acak dan saling bebas satu sama lain.

c. Uji Signifikansi Multivariat

Pada analisis variansi univariat, keputusan dibuat berdasarkan satu statistika uji, yaitu uji F yang nilainya ditentukan oleh hasil bagi dari dua rata-rata jumlah kuadrat sebagai taksiran hasil bagi taksiran variansi-variansi yang bersangkutan. Pada analisis variansi multivariat, ada beberapa statistik uji yang dapat digunakan untuk membuat keputusan (Kattree dan Naik, 2000), yaitu dengan

H = Matriks varian-kovarian perlakuan pada MANOVA

E = Matriks varian-kovarian error pada MANOVA

Pada analisis varian multivariat ada beberapa statistik uji yang dapat digunakan untuk membuat keputusan, yaitu

1. *Pillai's trace*. Statistik uji ini paling cocok digunakan jika asumsi homogenitas matriks varian-kovarian tidak dipenuhi, ukuran-ukuran sampel kecil, dan hasil-hasil dari pengujian bertentangan satu sama lain yaitu jika ada beberapa vektor rata-rata yang berbeda sedang yang lain tidak. Dalam uji ini, mengasumsikan jika makin tinggi nilai statistik *pillai's trace*, maka pengaruh terhadap model makin besar. Statistik uji *pillai's trace* dirumuskan sebagai:

$$P = \sum_{i=1}^p \left(\frac{\lambda_i}{1 + \lambda_i} \right) = \text{tr} \lambda_i (1 + \lambda_i)^{-1} = \text{tr} \frac{|B|}{|B + W|} \tag{1.1}$$

Keterangan:

$\lambda_i = \lambda_1, \lambda_2, \lambda_3, \dots, \lambda_p$ adalah akar karakteristik dari $W^{-1}B$

(W) = matriks varian-kovarian galat pada MANOVA

(B) = matriks varian-kovarian perlakuan pada MANOVA

2. *Wilk's lambda*. Statistik uji digunakan jika terdapat dua atau lebih kelompok variabel independen dan asumsi homogenitas matriks varian-kovarian dipenuhi. Makin rendah nilai statistik *wilk's lambda*, pengaruh terhadap model makin besar. Nilai *wilk's lambda* berkisar antara 0-1. Statistik uji *wilk's lambda* dirumuskan sebagai berikut.

$$\Lambda^* = \prod_{i=1}^p (1 + \lambda_i)^{-1} = \frac{|W|}{|B+W|} \quad (1.2)$$

3. *Hottelling's trace*. Statistik uji ini cocok digunakan jika hanya terdapat dua kelompok variabel independen. Makin tinggi nilai statistik *hottelling's trace*, pengaruh terhadap model makin besar. Nilai *hottelling's trace* > *pillai's trace*. Statistik uji *hottelling's* dirumuskan sebagai berikut.

$$T = \sum_{i=1}^p \lambda_i = \text{tr} \lambda_i = \text{tr}(W)^{-1}(B) \quad (1.3)$$

4. *Roy's largest root*. Statistik uji ini hanya digunakan jika asumsi homogenitas varian-kovarian dipenuhi. Makin tinggi nilai statistik *roy's largest root*, pengaruh terhadap model makin besar. Nilai *roy's largest root* > *hottelling's trace* > *pillai's trace*. Dalam hal pelanggaran asumsi normalitas multivariat, statistik ini kurang *robust* (kekar) dibandingkan

dengan statistik uji yang lainnya. Statistik uji *roy's largest root* dirumuskan sebagai:

$$R = \lambda_{maks} = maks(\lambda_1, \lambda_2, \dots, \lambda_p) \quad (1.4)$$

= akar karakteristik maksimum dari $(W)^{-1}(B)$

Keempat tes multivariat tersebut menggunakan uji statistik sebagai berikut.

$H_0 : \mu_1 = \mu_2 = \dots = \mu_k$ (tidak ada perbedaan antar perlakuan)

$H_1 : \mu_1 \neq \mu_2 \neq \dots \neq \mu_k$ (setidaknya ada perbedaan antar dua perlakuan)

Kriteria pengujiannya tolak H_0 , jika $F_{hitung} > F_{tebel}$.

Banyak *software* statistik menyajikan perhitungan keempat statistik uji MANOVA tersebut dan biasanya keempat statistik uji tersebut menghasilkan kesimpulan yang sama. Dalam kasus ketika keempat statistik uji tersebut menghasilkan kesimpulan yang berbeda dalam hal menerima dan menolak hipotesis, cara yang dapat dilakukan adalah dengan menguji nilai eigen dan matriks kovariani serta mengevaluasi permasalahan kesimpulan dalam karakteristik statistik uji (Rencher, 2002).

Ketika MANOVA menunjukkan bahwa hipotesis nol ditolak, dilanjutkan dengan ANOVA pada masing-masing variabel terikat. Prosedur ini akan menjaga taraf kesalahan α sepanjang uji ANOVA dilaksanakan. Apabila ANOVA menunjukkan bahwa hipotesis nol ditolak, maka dilakukan uji *post hoc* menggunakan metode *scheffe'*. Metode ini dilakukan ketika variabel bebas memiliki lebih dari dua nilai. Jika suatu variabel bebas hanya memiliki dua nilai maka untuk melihat perbedaan antara keduanya dapat langsung dilakukan perbandingan pada masing-masing rata-rata marginalnya (Sutrisno & Wulandari, 2018).

d. Asumsi-Asumsi pada MANOVA

1. Terdapat beberapa variabel terikat yang seluruh variabelnya kontinu (bisa berupa skala interval atau rasio). Terdapat satu atau dua variabel bebas yang berupa variabel kategori (ada beberapa kelompok data).
2. Terdapat linieritas, yaitu linieritas di antara kedua variabel terikat.
3. Tidak ada multikolinieritas, di antara variabel terikat.
4. Adanya independensi. Pengamatan harus independen secara statistik. Dipenuhinya persyaratan ini dimaksudkan agar perlakuan yang diberikan kepada setiap sampel, independen antara satu dengan lainnya.
5. Sampel acak. Dalam statistika, untuk hal pengambilan sampel harus dilakukan secara random (acak) dari populasinya atau dengan kata lain menggunakan teknik probabilitas. Selain itu, data yang diukur (variabel terikat) dalam penelitian yang menggunakan teknik analisis MANOVA harus berskala berskala interval.
6. Homogenitas matriks varian kovarian. Asumsi selanjutnya yang harus dipenuhi MANOVA adalah kesamaan matriks varian co-varian antargroup variabel dependen sehingga dapat dikatakan ada homoskedastisitas data. Namun, jika matriks kovariani antargrup variabel tidak sama, dapat dikatakan bahwa terjadi heteroskedastisitas.

Pelanggaran homogenitas dari varian adalah dasar kebenaran untuk pengambilan keputusan dalam multivariat analisis varian dari pada pengulangan analisis varian. Pengujian persyaratan homokedastisitas data dapat menggunakan koefisien *Box's M*. Statistika uji diperlukan untuk menguji homogenitas matriks varian-kovarian dengan hipotesis Normalitas multivariat. Dalam kasus MANOVA diasumsikan bahwa variabel terikat (secara bersama) berdistribusi normal multivariat di dalam kelompok.

$$H_0 = \Sigma 1 = \Sigma 2 = \Sigma 3 = \dots = \Sigma n \quad (1.5)$$

H_1 = ada paling sedikit satu diantara sepasang $\Sigma 1$ yang tidak sama.

Jika dari masing-masing populasi diambil sampel acak berukuran n yang saling bebas maka penduga tak bias untuk Σ_l adalah matriks S_l sedangkan untuk Σ_0 penduga tak biasnya adalah S .

$$\text{Dengan} \quad S = \frac{1}{N} \sum_{i=1}^g (n_i - 1) S_i \quad N = \sum_{i=1}^g n_i - g \quad (1.6)$$

Untuk menguji hipotesis di atas dengan tingkat signifikansi α , digunakan kriteria uji berikut.

$$H_0 \text{ ditolak jika } MC^{-1} > \chi^2_{\left(\frac{1}{2}(g-1)p(p+1)\right)}(\alpha) \quad (1.7)$$

$$H_1 \text{ diterima jika } MC^{-1} < \chi^2_{\left(\frac{1}{2}(g-1)p(p+1)\right)}(\alpha)$$

Keterangan:

$$M = \sum_{i=1}^g (n_i - 1) \ln |S| - \sum_{i=1}^g (n_i - 1) \ln |S_i| \quad (1.8)$$

$$C^{-1} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \left(\sum_{i=1}^g \frac{1}{n_i - 1} - \frac{1}{\sum_{i=1}^g (n_i - 1)} \right) \quad (1.9)$$

Dengan bantuan program SPSS, uji homogenitas matriks varian-*co-varian* dapat dilakukan dengan Uji *Box's M*. Jika nilai sig. $> \alpha$, maka H_0 diterima sehingga dapat disimpulkan matriks varian-kovarian dari i -populasi adalah sama atau homogen. Jika ada variabel yang mengalami

heterokedastisitas, dapat dilakukan transformasi data misalnya dengan cara mengubah data kedalam bentuk logaritma atau logaritma natural.

Adapun langkah-langkah uji homogenitas varian-kovarian menggunakan program SPSS 20 adalah sebagai berikut.

- a. Dari Worksheet, entry data dilakukan melalui Variable View dan Data View.
 - b. Dari menu utama SPSS dipilih menu Analyze, kemudian submenu General Linear Mode dipilih Multivariat.
 - c. Setelah tampak dilayar tampilan Window Multivariat, kemudian melakukan Entry variabel-variabel yang sesuai pada kotak Dependent Variables dan Fixed Factor(S).
 - d. Selanjutnya Option dipilih Homogenitas Test dan Continue, terakhir OK.
7. Sebaran normalitas data. Tujuan dari uji normalitas adalah untuk mengetahui apakah distribusi satu unit data mengikuti atau mendekati distribusi normal. Data yang baik adalah data yang mempunyai pola seperti distribusi normal. Distribusi normal multivariat adalah suatu perluasan dari distrbusi normal univariat. Metode statistika multivariat MANOVA mensyaratkan terpenuhinya asumsi distribusi normalitas dengan hipotesis adalah

H_0 : Data berdistribusi normal multivariat.

H_1 : Data tidak berdistribusi normal multivariat.

Berdasarkan teorema yang menyatakan bahwa y_1, y_2, \dots, y_p berdistribusi normal multivariat maka $(Y - \mu)' \Sigma^{-1} (Y - \mu)$ berdistribusi d_p^2 . Berdasarkan sifat ini maka pemeriksaan distribusi normal multivariat dapat dilakukan pada setiap populasi dengan cara membuat *Q-Q plot* atau *scatter plot* membandingkan jarak *Mahalanobis* d_i^2 dan chi kuadrat

sentroid (q_i). Tahapan dari pembuatan $Q-Q$ plot ini adalah sebagai berikut (Johnson & Wichern, 2017)

a. Menentukan nilai rata-rata \bar{Y} dan invers dari matriks varian kovarian S^{-1} .

b. Menentukan nilai d_i^2 dengan rumus:

$$d_i^2 = (y_i - \bar{y})' S^{-1} (y_i - \bar{y}) \quad \text{---} \quad (1.10)$$

c. Mengurutkan nilai jarak mahalanobis dari yang terkecil sampai terbesar.

$$d_1^2 < d_2^2 < \dots < d_n^2 \quad (1.11)$$

d. Mencari nilai sentroid (q_i) setiap observasi yang didekati dengan.

$\chi_p^2 \left(\frac{n-i+0,5}{n} \right)$. p merupakan banyaknya variabel terikat, serta dapat dilihat pada tabel χ^2 .

e. Membuat plot antara d_i^2 dan $\chi_p^2 \left(\frac{n-i+0,5}{n} \right)$ dengan d_i^2 sebagai ordinat dan $\chi_p^2 \left(\frac{n-i+0,5}{n} \right)$ sebagai axis.

f. Data akan dikatakan berdistribusi normal multivariat apabila plot yang didapat cenderung mengikuti pola garis lurus. Kelengkungan menunjukkan penyimpangan dari distribusi normal jika digunakan kolomogorov-smirnov, kriteria pengujian adalah angka signifikansi > 0.05 maka data berdistribusi secara normal dan angka signifikansi < 0.05 maka data tidak terdistribusi secara normal.

g. Perlakuan yang memungkinkan agar data menjadi menyebar secara manual adalah menambah jumlah data, mengjhlilangkan data yang menjadi penyebab tidak normalnya distribusi dan melakukan transformasi data. Adapun langkah-langkah uji normalitas multivariat program SPSS 20 adalah sebagai berikut.

1) Menentukan hpotesis

H_0 = Data berdistribusi normal multivariat

H_1 = Data tidak berdistribusi normal multivariat.

2) Analisis data

Untuk memudahkan penghitungan peneliti menggunakan bantuan program komputer SPSS 20 dalam menganalisis data, yaitu dengan membuat *scatter plot* antara jarak mahalanobis dengan chi square.

Langkah-langkahnya sebagai berikut: masukkan data, pilih *regresi linear-save-distance*, pilih *mahalanobis* maka akan muncul variabel baru **Mah_1**. Kemudian urutkan **Mah_1** dengan *sort cases ascending*. Membuatlah variabel baru, misal beri nama "J", lalu masukkan data 1,2,3.....,n sesuai banyaknya data. Menghitung probabilitasnya melalui *compute variable*, masukkan rumus $(J-0,5)/n$, dimana n merupakan banyaknya data. Muncul variabel baru, misal diberi nama *prob_value*. Pilih *compute variable* lagi, misal diberi nama *chi*, lalu pada kolom *Numeric expression* adalah $IDF.CHISQ(prob_value,2)$. Kemudian klik OK. Setelah itu, membuat scatter plot. Pilih *grap-legacy dialog-scatter/dot*. Pilih *simple scatter*, klik *define*.kemudian masukkan Mah_1 ke Y axis dan chi ke X axis, klik OK.

Selain Uji normalitas multivariat, bisa juga dilakukan uji normalitas univariat. Hair dkk. (2010) mengemukakan meskipun normalitas univariat tidak menjamin (*does not guarantee*) normalitas multivariat, jika seluruh variabel memenuhi normalitas univariat, maka setiap penyimpangan (*any departures*) dari normalitas multivariat biasanya tidak

penting (*are usually inconsequential*). Untuk menguji asumsi normalitas populasi dari suatu sampel, dapat digunakan uji Kolmogorov-Smirnov. Hipotesis nol menyatakan sampel yang diambil berasal dari populasi yang berdistribusi normal, sedangkan hipotesis alternatif menyatakan sampel yang diambil tidak berasal dari populasi yang berdistribusi normal. Pengambilan keputusan terhadap hipotesis dapat dilakukan dengan membandingkan nilai probabilitas dari uji Kolmogorov-Smirnov terhadap tingkat signifikansi yang digunakan. Berikut aturan pengambilan keputusan terhadap hipotesis dengan pendekatan probabilitas.

Jika nilai probabilitas

\geq *tingkat signifikansi, maka H_0 diterima dan H_1 ditolak.*

Jika nilai probabilitas

$<$ *tingkat signifikansi, maka H_0 ditolak dan H_1 diterima.*

Normalitas yang diuji adalah normalitas di setiap metode mengajar dan disetiap variabel terikatnya. Langkah-langkah uji normalitas univariat adalah Analyze lalu Descriptive Statistics lalu Explore sehingga muncul kotak dialog Explore. Lalu masukkan variabel terikat ke dalam dependent list dan variabel **bebas** pada *factor list*. Selanjutnya pilih Plots, sehingga muncul kotak dialog Explore Plots. pilih/centang Normality plots with tests. kemudian pilih Continue. Pada Display pilih Plots. Selanjutnya pilih OK.

8. Data *Outlier*. Data *outlier* adalah data yang secara nyata berbeda dengan data yang lain. *Outlier* adalah kasus dengan nilai ekstrem pada kombinasi variabel yang koefisien korelasinya sangat berpengaruh terhadap nilai rata-rata dari kelompok. *Outlier* dapat ditemukan antara situasi univariat dan situasi multivariat, di antara variabel dependen dan variabel independen, serta diantara input dan output suatu analisis.

Multivariat analisis cukup sensitif dengan keberadaan data *outlier*. Oleh karena itu, data harus dianalisis apakah mengandung *outlier* atau tidak. Adanya data *outlier* pada analisis multivariat akan menyebabkan data sulit untuk ditafsirkan hasilnya. Data *outlier* bisa terjadi karena beberapa faktor yaitu kesalahan dalam pemasukan data, kesalahan pada pengambilan sampel, dan terdapat data-data ekstrem yang tidak dapat dihindarkan keberadaannya. Langkah-langkah menemukan *outlier* yaitu membuat titik pancar untuk setiap variabel, membuat diagram pancar untuk setiap variabel dan menghitung skor standar dengan rumus:

$$Z_{jk} = \frac{(Y_{jk} - Z_k)}{\sqrt{skk}} \quad (1.12)$$

Keterangan: $j = 1, 2, \dots, n$ serta setiap kolom $k = 1, 2, 3, \dots, p$. Suatu data dikatakan *outlier* jika nilai Z lebih besar dari $+2.5$ atau lebih kecil sama dengan -2.5 .

Langkah-langkah menemukan *outlier* dengan SPSS, yaitu klik Graph lalu Legacy Dialogs lalu Boxplots lalu pilih kotak Simple lalu pilih Summaries of Separate Variables lalu Define lalu setelah muncul kotak dialog, masukkan variabel terikat ke kotak Boxes Represent dan klik OK.

e. Model Analisis Multivariate Satu Arah (MANOVA Satu Arah)

Salah satu model MANOVA sebagai perluasan dari ANOVA Satu Arah adalah MANOVA Satu Arah. Model ini dengan pengaruh tetap dapat digunakan untuk menguji apakah ke- g populasi (dari satu faktor yang sama) menghasilkan vektor rata-rata yang sama untuk p variabel respon atau variabel dependent yang diamati dalam penelitian. Untuk membandingkan vektor rata-rata populasi g berdasarkan bentuk model ANOVA Satu Arah menurut Candiasa (2010) adalah sebagai berikut.

$$Y_{ij} = \mu + \tau_i + e_{ij} \quad (1.13)$$

Keterangan:

j = 1,2,3,... n dan i = 1,2,3,... g

Y_{ij} = Nilai pengamatan (respons tunggal) dari ulangan ke- j yang memperoleh perlakuan ke- i

μ = Nilai rerata (mean)

τ_i = Pengaruh dari perlakuan ke- i terhadap respons

e_{ij} = Pengaruh galat yang timbul pada ulangan ke- j dan perlakuan ke- i

Yang diasumsikan bebas dan berdistribusi $N_p(0, \Sigma)$ untuk data multivariat.

Vektor observasi dapat dikomposisi ulang sesuai model, seperti berikut.

$$Y_{ij} = y + y_i - y + y_{ij} - y_i \quad (1.14)$$

(observasi) (rata-rata (estimasi (residu e_{ij}) sampel efek keseluruhan $\bar{\mu}$) perlakuan τ_i)

Suatu vektor dari pengamatan data multivariat dianalisis berdasarkan bentuk (1.13) dan bentuk (1.14) mengacu untuk jumlah kuadrat pada model MANOVA Satu Arah. Sehingga digunakan

$$(y_{ij} - \bar{y})(\bar{y}_{ij} - \bar{y})^t \quad \text{---} \quad (1.15)$$

Dapat ditulis sebagai berikut:

$$\begin{aligned} (y_{ij} - \bar{y})(\bar{y}_{ij} - \bar{y})^t &= ((y_i - \bar{y}) + (y_{ij} - \bar{y}_i))((\bar{y}_{ij} - \bar{y}) + (\bar{y}_{ij} - \bar{y}_i))^t \\ &= (\bar{y}_i - \bar{y})(\bar{y}_i - \bar{y})^t + (\bar{y}_i - \bar{y})(y_{ij} - \bar{y}_i)^t + \\ &\quad (y_{ij} - \bar{y}_i)(\bar{y}_{ij} - \bar{y}_i)^t + (y_{ij} - \bar{y}_i)(y_{ij} - \bar{y}_i)^t \end{aligned} \quad (1.16)$$

Jumlah untuk semua pengamatan ke- i berdasarkan bentuk (1.16) dirumuskan sebagai berikut

$$\sum_{j=1}^{n_i} (y_{ij} - \bar{y})(\bar{y}_{ij} - \bar{y})^t = n_i(\bar{y}_i - \bar{y})(\bar{y}_i - \bar{y})^t + \sum_{j \neq i}^{n_i} (y_{ij} - \bar{y}_i)^t \quad (1.17)$$

dengan . Selanjutnya $\sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i) = 0$ dijumlahkan untuk semua populasi menghasilkan jumlah pengamatan total.

$$\begin{aligned} \sum_{i=1}^g \sum_{j=1}^{n_i} (y_{ij} - \bar{y})(\bar{y}_{ij} - \bar{y})^t &= \sum_{i=1}^g n_i(\bar{y}_i - \bar{y})(\bar{y}_i - \bar{y})^t \\ &\quad + \sum_{i=1}^g \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(y_{ij} - \bar{y}_i)^t \end{aligned} \quad (1.18)$$

Untuk bentuk (1.18), misalkan

$$\begin{aligned} W &= \sum_{i=1}^g \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(y_{ij} - \bar{y}_i)^t \\ &= (n_1 - 1)S_1 + (n_2 - 1)S_2 + \dots + (n_g - 1)S_g \end{aligned} \quad (1.19)$$

S_i adalah matriks kovariani sampel ke- i . Matriks tersebut mempunyai peran yang dominan dalam pengujian untuk ada tidaknya pengaruh perlakuan. Analogi pada univariat, hipotesis tanpa pengaruh perlakuan pada multivariat dapat dirumuskan dengan:

$$H_0 = \tau_1 = \tau_2 = \dots = \tau_i = \dots = \tau_g \text{ dengan } \tau_i = \begin{bmatrix} \mu_{i1} \\ \vdots \\ \mu_{ip} \end{bmatrix} \text{ dan } i = 1, 2, \dots, g. \quad (1.20)$$

Tabel 6.2 Pengamatan MANOVA Satu Arah.

Obyek	<i>Treatment</i>						Rata-Rata			
	I			...	H			\bar{Y}_1	...	\bar{Y}_p
	Y_1	...	Y_p		Y_1	...	Y_p			
1	Y_{111}		Y_{11p}		Y_{h11}		Y_{h1p}	$\bar{y}_{.11}$		$\bar{y}_{.1p}$
\vdots	\vdots	...	\vdots	...	\vdots	...	\vdots	\vdots	...	\vdots
n	Y_{1n1}		Y_{1np}		Y_{hn1}		Y_{hnp}	$\bar{y}_{.n1}$		$\bar{y}_{.np}$
Rata-rata	$\bar{y}_{1.1}$...	$\bar{y}_{1.p}$...	$\bar{y}_{h.1}$...	$\bar{y}_{h.p}$	$\bar{y}_{.1}$...	$\bar{y}_{.p}$

Sumber : Mattjik dan Sumertajaya, 2011

$k = 1, 2, 3, \dots, h$ adalah taraf dari *treatment* sebanyak h

$i = 1, 2, 3, \dots, n$ adalah banyaknya pengamatan sebanyak n

$j = 1, 2, 3, \dots, p$ adalah banyaknya variabel dependen sebanyak p

Dapat diuji kesamaan vektor rata-rata dengan mencari matriks jumlah kuadrat dan hasil kali untuk perlakuan dan sisa. Secara akuivalen, akan didapat hubungan ukuran relatif dari galat (sisa) dan total (koreksi) jumlah dari kuadrat dan hasil kali berdasarkan bentuk (1.18). Untuk perhitungan statistik uji digunakan tabel MANOVA. Tabel 6.3 menampilkan perbandingan vektor mean.

Tabel 6.3 MANOVA untuk membandingkan vektor mean.

Sumber Variansi	Matriks Jumlah Kuadrat dan Perkalian Silang	Derajat Kebebasan
Treatmenten	$B = \sum_{i=1}^g n_i (y_i - \bar{y})(y_i - \bar{y})'$ $W = \sum_{i=1}^g \sum_{j=1}^{n_i} (y_{ij} - \bar{y}_i)(y_{ij} - \bar{y}_i)'$	$g - 1$ $\sum_{i=1}^g n_i - g$
Total (Rata-rata terkoreksi)	$B + W = \sum_{i=1}^g \sum_{j=1}^{n_i} (y_{ij} - \bar{y})(y_{ij} - \bar{y})'$	$\sum_{i=1}^g n_i - 1$

Sumber : Mattjik dan Sumertajaya, 2011

Berdasarkan tabel yang telah disajikan diatas maka selanjutnya dihitung nilai koefisien *wilk's lambda* (Λ^*) dengan menggunakan rumus:

$$\Lambda^* = \frac{|W|}{|B+W|} \quad (1.21)$$

Statistik *wilk's lambda* dapat ditransformasikan ke besaran statistic *F* sehingga dapat di bandingkan dengan tabel *F*. Bentuk transformasi dari besaran *wilk's lambda* (Λ^*) ke besaran *F* untuk berbagai kombinasi jumlah peubah *P* dan derajat bebas perlakuan (*dbp*) disajikan dalam Tabel 6.4.

Tabel 6.4 Transformasi dari besaran *Wilk's Lambda* (Λ^*) ke besaran *F*

Banyak Variabel	Banyak Kelompok	Transformasi F	Derajat Bebas
1	≥ 1	$\left[\frac{1 - \Lambda}{\Lambda} \right] \left[\frac{db_p}{db_G} \right]$	$db_p; db_G$
2	≥ 1	$\left[\frac{1 - \sqrt{\Lambda}}{\sqrt{\Lambda}} \right] \left[\frac{db_G - 1}{db_p} \right]$	$2db_p; 2(db_G - 1)$
≥ 1	1	$\left[\frac{1 - \Lambda}{\Lambda} \right] \left[\frac{db_p - db_G - p}{p} \right]$	$p; (db_p + db_G - p)$

$$\geq 2 \quad 2 \quad \left[\frac{1 - \sqrt{\Lambda}}{\sqrt{\Lambda}} \right] \left[\frac{db_p - db_G - p}{p} \right] \quad 2p; (db_p + db_G - p - 1)$$

Sumber: Mattjik dan Sumertajaya, 2011

Keterangan:

p = banyak peubah respons yang diamati

db_p = derajat bebas perlakuan

db_G = derajat bebas galat

f. Analisis Varian Multivariat Dua Arah (MANOVA Dua Arah)

Uraian pada subbab sebelumnya didasarkan atas rancangan MANOVA satu arah atau satu variabel independent. Dalam hal ini diskusi diperluas menjadi MANOVA dua arah yaitu yang melibatkan dua variabel independent. Konsekuensi menerapkan MANOVA dua arah adalah jumlah hipotesis jauh lebih banyak. MANOVA dua arah ini sering disebut dengan MANOVA multifactor.

Seperti halnya analisis ragam peubah ganda satu arah, analisis peubah ganda dua arah juga merupakan pengembangan lebih lanjut dari analisis ragam satu peubah dua arah (two way ANOVA). Model yang sering digunakan dalam analisis ragam satu variabel dua arah adalah Rancangan Acak Lengkap. Persamaan yang sering digunakan adalah

$$y_{ijk} = \mu_k + \tau_{ik} + \beta_{jk} + \varepsilon_{ijk} \quad (1.22)$$

Dengan:

$$i = 1, 2, 3, \dots, a$$

$$j = 1, 2, 3, \dots, b$$

$$k = 1, 2, 3, \dots, n$$

Y_{ij} = nilai pengamatan (respons tunggal) dari ulangan ke-j yang memperoleh perlakuan ke-i

μ = nilai rerata (mean) dari respons ke k

τ_i = pengaruh dari perlakuan ke-i terhadap respons

ϵ_{ij} = pengaruh galat yang timbul pada respons ke-k dari kelompok ke-j dan memperoleh perlakuan ke-

β_{bj} = pengaruh kelompok ke-j terhadap respons ke k

Pada pengujian MANOVA perlu diketahui kesamaan rata-rata dari variabel dependen setelah dilakukan pengujian melalui dua tahap dengan rumus, Pada model dua arah, jumlah total dari table dan matriks dapat didefinisikan sebagai;

$$T = H_A + H_B + H_{AB} + E \quad (1.23)$$

dengan,

H_A = adalah jumlah prosuk yang sesuai untuk semua pasangan variable A

H_B = adalah jumlah prosuk yang sesuai untuk semua pasangan variable B

H_c = adalah jumlah prosuk yang sesuai untuk semua pasangan variable AB

1. Rata-rata variabel dependen untuk setiap pengujian.

$$h_{Arr} = nb \sum_{i=1}^a (\bar{y}_{i.r} - \bar{y}_r)^2 = \sum_{i=1}^a \frac{\bar{y}_{i.r}^2}{nb} - \frac{y_{.r}^2}{nab} \quad (1.24)$$

$$h_{Arz} = nb \sum_{i=1}^a (\bar{y}_{i.r} - \bar{y}_r)(\bar{y}_{i.s} - \bar{y}_s) = \sum_{i=1}^a \frac{\bar{y}_{i.r} \bar{y}_{i.s}}{nb} - \frac{y_{.r} \cdot y_{i.s}}{nab} \quad (1.25)$$

Dengan demikian table MANOVA dua jalur untuk membandingkan faktor mean adalah

Tabel 6.5 MANOVA dua jalur untuk membandingkan vektor mean.

Sumber Variasi	Matriks jumlah kuadrat dan perkalian silang	Derajat Kebebasan
A	$H^A = n p \sum (\underline{\lambda}^{1r} - \underline{\lambda}^{\cdot\cdot}) (\underline{\lambda}^{1r} - \underline{\lambda}^{\cdot\cdot})_t$	$a-1$
B	$H^B = n q \sum (\underline{\lambda}^{1r} - \underline{\lambda}^{\cdot\cdot}) (\underline{\lambda}^{1r} - \underline{\lambda}^{\cdot\cdot})_t$	$b-1$
AB	$H^{AB} = n \sum (\underline{\lambda}^{\hat{a}r} - \underline{\lambda}^{1r} - \underline{\lambda}^{\hat{a}\cdot} - \underline{\lambda}^{\cdot\cdot})_t$ $H^{AB} = n \sum (\underline{\lambda}^{\hat{a}r} - \underline{\lambda}^{1r} - \underline{\lambda}^{\hat{a}\cdot} - \underline{\lambda}^{\cdot\cdot})_t$	$(a-1)(b-1)$
Error	$E = \sum (\lambda^{\hat{a}r} - \underline{\lambda}^{\hat{a}\cdot}) (\lambda^{\hat{a}r} - \underline{\lambda}^{\hat{a}\cdot})_t$	$ab(n-1)$
Total	$L = \sum (\lambda^{\hat{a}r} - \underline{\lambda}^{\cdot\cdot}) (\lambda^{\hat{a}r} - \underline{\lambda}^{\cdot\cdot})_t$	$Abn-1$

Sumber : Mattjik dan Sumertajaya, 2011

2. Rata-rata keseluruhan.

$$h_{ABrr} = \sum_{ij} \frac{y_{ijr}^2}{n} - \frac{y_{\cdot r}^2}{nab} - h_{Arr} - h_{Brr}$$

$$h_{ABrs} = \sum_{ij} \frac{y_{ijr} \cdot y_{ijs}}{n} - \frac{y_{\cdot r} \cdot y_{\cdot s}}{nab} - h_{Ars} - h_{Brs}$$
(1.26)

3. Matriks error (E) merupakan hasil komputasi dari rumus (1.23) yaitu

$$T = H_A + H_B + H_{AB} + E$$

Elemen matriks E kita peroleh dari

$$e_{rr} = \sum_{ijk} y_{ijk}^2 - \frac{y_{...r}^2}{nab} - h_{Ar r} - h_{Br r} - h_{ABrr}$$

$$e_{rs} = \sum_{ijk} y_{ijk} \cdot y_{ijks} - \frac{y_{...r} \cdot y_{...s}}{nab} - h_{Ar s} - h_{Br s} - h_{ABrs}$$
(1.27)

Berdasarkan partisi matriks kovarian, matriks E didefinisikan sebagai:

$$E = \begin{bmatrix} E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{bmatrix}$$
(1.28)

Dengan, berdasarkan partisi matriks kovarian, matriks H didefinisikan sebagai:

$$H = \begin{bmatrix} H_{xx} & H_{xy} \\ H_{yx} & H_{yy} \end{bmatrix}$$
(1.29)

Adapun uji prasyarat analisis MANOVA dua arah sama dengan yang dijelaskan sebelumnya mengenai uji prasyarat secara umum.

H. Contoh Kasus dalam MANOVA

Berikut diberikan contoh kasus yang dapat diselesaikan dengan pendekatan MANOVA satu arah. Prodi **A** merencanakan akan menerapkan tiga metode mengajar. Adapun ketiga metode mengajar tersebut adalah metode mengajar *project based learning* (PjBL), *metode collaborative*

learning (CL), dan *metode problem based learning* (PBL). Ketua prodi ingin mengetahui metode mengajar manakah yang dapat membuat rata-rata nilai ujian RAB dan Manajemen Konstruksi menjadi tinggi. Untuk itu, ketua prodi tersebut akan melakukan penelitian. Berikut hal-hal yang akan diteliti.

1. Meneliti apakah ketiga metode mengajar tersebut memiliki kemampuan yang sama dalam mempengaruhi nilai ujian RAB dan Manajemen Konstruksi secara simultan (*simultaneously*) (mempengaruhi secara rata-rata). Jika tidak, berarti terdapat metode mengajar yang memiliki kemampuan berbeda dalam hal mempengaruhi nilai ujian RAB dan Manajemen Konstruksi secara simultan (*simultaneously*).
2. Seandainya terdapat metode mengajar yang memiliki kemampuan berbeda dalam hal mempengaruhi nilai ujian RAB dan Manajemen Konstruksi secara simultan (*simultaneously*), akan diselidiki lebih dalam sebagai berikut.
3. Apakah penerapan dari ketiga metode mengajar tersebut berpengaruh terhadap nilai ujian RAB (secara rata-rata).

Jika terdapat pengaruh pada penggunaan metode mengajar terhadap nilai ujian RAB, metode mengajar manakah yang membuat nilai rata-rata ujian RAB tinggi. Maka akan dilakukan hal sebagai berikut.

Membandingkan nilai rata-rata ujian RAB antara:

- a. Metode PjBL \Leftrightarrow metode CL,
 - b. Metode PjBL \Leftrightarrow metode PBL, dan
 - c. Metode CL \Leftrightarrow metode PBL.
4. Apakah penerapan dari ketiga metode mengajar tersebut berpengaruh terhadap ujian Manajemen Konstruksi.

Jika terdapat pengaruh pada penggunaan metode mengajar terhadap nilai ujian Manajemen Konstruksi, metode mengajar manakah yang membuat nilai rata-rata ujian Manajemen Konstruksi tinggi. Maka akan dilakukan hal sebagai berikut.

Membandingkan nilai rata-rata ujian Manajemen Konstruksi antara:

- a. Metode PjBL ⇔ metode CL,
- b. Metode PjBL ⇔ metode PBL, dan
- c. Metode CL ⇔ metode PBL.

Tabel 6.6 Contoh Data

No	Nama	Metode PjBL		No.	Nama	Metode CL		No.	Nama	Metode PBL	
		RA B	MK			RAB	MK			RAB	MK
1.	Udin	60	61	13.	Andi	61	61	25.	Dani	60	67
2.	Ugi	61	61	14.	Ani	61	63	26.	Darul	61	68
3.	Umar	62	62	15.	Aan	62	64	27.	Didi	62	69
4.	Uni	63	63	16.	Ado	63	65	28.	Dina	63	70
5.	Uun	64	64	17.	Adi	64	66	29.	Danu	64	71
6.	Uci	65	65	18.	Ama	65	67	30.	Dora	65	72
7.	Ulan	60	60	19.	Agung	60	62	31.	Dami	60	67
8.	Umo	61	61	20.	Anto	61	63	32.	Didin	61	68
9.	Uti	62	62	21.	Anti	62	64	33.	Desti	62	69
10.	Urdu	63	63	22.	Ahmad	63	65	34.	Damia	63	70
11.	Urza	64	64	23.	Aci	64	66	35.	Dona	64	71
12.	Uza	65	65	24.	Ali	65	67	36.	Doni	65	72
Rata-Rata		62,5	62,58	Rata-Rata		62,58	64,42	Rata-Rata		62,5	69,5

Berdasarkan data pada Tabel 6.6, dari 36 mahasiswa, 12 mahasiswa diterapkan metode mengajar PjBL, 12 mahasiswa berikutnya diterapkan metode mengajar Collaborative Learning, dan mahasiswa lainnya diterapkan

metode mengajar PBL. Setelah satu bulan kemudian, diadakan ujian RAB dan M.K dari seluruh mahasiswa tersebut. Perhatikan bahwa metode mengajar merupakan variabel bebas, sedangkan nilai ujian RAB dan M.K merupakan variabel tak bebas. Dalam contoh kasus ini, jumlah variabel tak bebas sebanyak dua, yakni nilai RAB dan M.K. Dengan menggunakan MANOVA dapat diuji ada tidaknya pengaruh (*effect*) yang signifikan secara statistik pada penggunaan metode mengajar terhadap nilai RAB dan M.K secara bersamaan (*simultaneously*).

1. Dugaan:

H₀ : terdapat perbedaan penggunaan metode mengajar terhadap nilai ujian RAB.

H₁ : tidak terdapat perbedaan penggunaan metode mengajar terhadap nilai ujian RAB.

2. Dugaan:

H₀ : terdapat perbedaan penggunaan metode mengajar terhadap nilai ujian Manajemen Konstruksi.

H₁ : tidak terdapat perbedaan pada penggunaan metode mengajar terhadap nilai ujian Manajemen Konstruksi

5. Dugaan:

H₀ : terdapat perbedaan secara simultan penggunaan metode mengajar terhadap nilai ujian RAB dan nilai ujian Manajemen Konstruksi.

H₁ : tidak terdapat perbedaan secara simultan pada penggunaan metode mengajar terhadap nilai ujian RAB dan nilai ujian Manajemen Konstruksi.

I. Penyelesaian Kasus MANOVA dengan SPSS

1. Analisis Asumsi

- a. Melakukan uji linieritas di antara kedua variabel terikat di setiap metode mengajar sehingga diperoleh 3 hasil linieritas karena ada 3 kelompok metode mengajar.
- b. Melakukan *split output*.

1

2

3

- Pilih data, Split File,
- Setelah muncul kotak dialog, pilih Compare groups lalu masukkan variabel metode ke kotak groups based on, lalu OK,
- Akan muncul keterangan *sort cases* pada layar *output*.

Current Status: Compare:Metode

OK Paste Reset Cancel Help

Sort Cases BY Metode.
Split File Layered BY Metode.

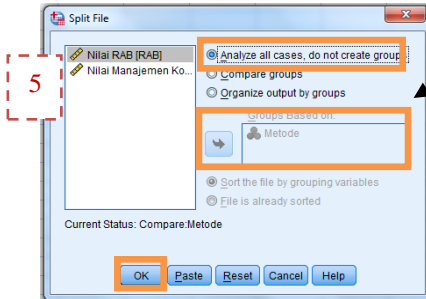
c. Membuat *scatter/dot*.

Pilih Legacy Graphs, Dialogs, lalu pilih Scatter/Dot.

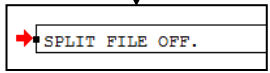
Pilih Simple Scatter lalu Define.

Masukkan variabel nilai MK ke kotak Y Axis dan variabel nilai RAB ke X Axis, lalu OK.

Setelah diperoleh Scatterplot, untuk melihat adanya linieritas antara hubungan kedua variabel terikat, yaitu nilai RAB dan nilai MK. Hubungan tersebut ditunjukkan oleh pola pada Scatterplot yang menunjukkan kecenderungan membentuk garis lurus. Dengan pola garis lurus tersebut, dapat disimpulkan bahwa nilai RAB dan nilai MK memperlihatkan hubungan linier, baik pada metode mengajar PJB, *cooperative learning*, maupun PBL.



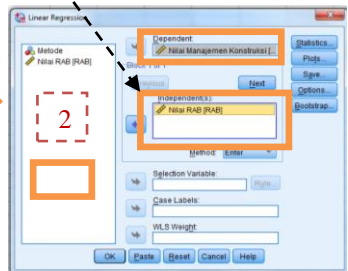
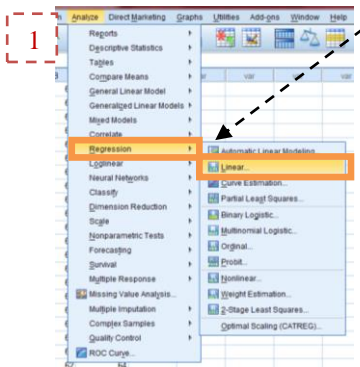
Setelah diperoleh asumsi linieritas, data kita kembalikan ke format semula dengan cara menonaktifkan Split File. Lalu akan muncul keterangan Split File Off pada layar *output*



d. Melakukan uji multikolinieritas.

Karena tidak terdapat uji multikolinieritas pada analisis MANOVA, maka kita menitipkan analisis ini pada analisis regresi.

- Pilih Analyze, Regression, Linier.
- Setelah muncul kotak dialog Linier Regression, masukkan variabel nilai MK dalam kotak dependen dan variabel nilai RAB pada kotak independen, kita bebas memilih variabel mana yang akan



Dari hasil *output*: cukup perhatikan nilai R Square. Bila terlalu tinggi, dinyatakan terjadi multikolinieritas, Nilai yang diperoleh pada kasus ini adalah 0,241 (moderat), artinya tidak terlalu tinggi dan tidak terlalu rendah.

3 *Output*

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,491 ^a	,241	,219	3,044

a. Predictors: (Constant), Nilai RAB

Hasil analisis menunjukkan bahwa nilai RAB dan nilai MK saling berhubungan dan kekuatan hubungannya pada kategori moderat ($R^2 = 0,241$). Dengan demikian, gejala multikolinieritas tidak

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	100,030	1	100,030	10,798	,002 ^b
	Residual	314,970	34	9,264		
	Total	415,000	35			

a. Dependent Variable: Nilai Manajemen Konstruksi
b. Predictors: (Constant), Nilai RAB

Jika nilai R Square terlalu rendah, dan tidak terdapat hubungan antar kedua variabel terikat, maka analisis tidak dapat dilanjutkan menggunakan analisis MANOVA, tetapi melakukan 2x ANOVA untuk masing-masing variabel terikat. Jika terdapat gejala multikolinieritas, maka analisis tidak dapat dilanjutkan menggunakan MANOVA

b. Melakukan Uji Asumsi Normalitas

a. Menentukan hipotesis.

H_0 = Data berdistribusi normal multivariat

H_1 = Data tidak berdistribusi normal multivariat.

b. Normalitas univariat.

a) Pilih Analyze, Descriptive Statistics, Explore.

b) Pada kotak dialog Explore, masukkan variabel nilai RAB dan MK pada kotak Dependent List, dan variabel metode mengajar dalam kotak Factor List, lalu OK.

c) Pilih Plots, centang Normality plots with test pada kotak dialog Explore Plots, lalu Continue.

		Tests of Normality					
		Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Metode	Statistic	df	Sig.	Statistic	df	Sig.
Nilai RAB	Metode Project Based Learning	,133	12	,200	,929	12	,372
	Metode Cooperative Learning	,161	12	,200	,933	12	,412
	Metode Problem Based Learning	,133	12	,200	,929	12	,372
Nilai Manajemen Konstruksi	Metode Project Based Learning	,161	12	,200	,933	12	,412
	Metode Cooperative Learning	,127	12	,200	,955	12	,708
	Metode Problem Based Learning	,133	12	,200	,929	12	,372

^a. This is a lower bound of the true significance.
a. Lilliefors Significance Correction

Dasar pengambilan keputusan, perhatikan nilai Sig. jika nilai Sig. > dari 0,05 maka asumsi normalitas univariat terpenuhi.

c. Normalitas multivariat. Untuk uji normalitas multivariat, ada beberapa variabel baru yang dimunculkan untuk membantu proses analisis.

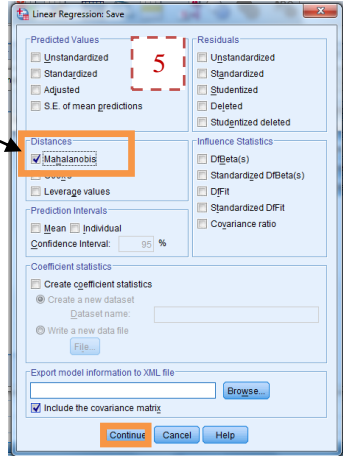
Membuat variabel jumlah antara variabel terikat 1 dan variabel terikat 2.

Pilih Analyze, Regression, Linier, untuk memunculkan variabel baru.

Masukkan variabel jumlah ke kotak Dependent, dan kedua variabel terikat ke dalam kotak Independent, pilih Plots, lalu OK.

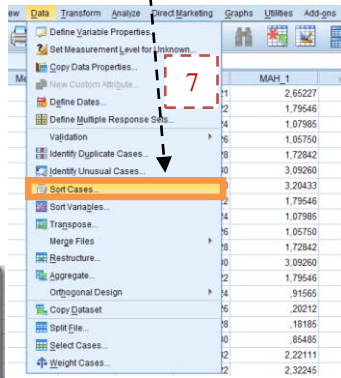
	Metode	RAB	MK	Jumlah
1	1	60	61	121
2	1	61	61	122
3	1	62	62	124
4	1	63	63	126
5	1	64	64	128
6	1	65	65	130
7	1	60	60	120
8	1	61	61	122
9	1	62	62	124
10	1	63	63	126
11	1	64	64	128
12	1	65	65	130
13	2	61	61	122
14	2	61	63	124
15	2	62	64	126
16	2	63	65	128
17	2	64	66	130
18	2	65	67	132
19	2	60	62	122
20	2	61	63	124
21	2	62	64	126
22	2	63	65	128
23	2	64	66	130
24	2	65	67	132

Setelah muncul kotak dialog Linear Regression Save, centang Mahalanobis pada kotak Distance. Lalu pilih Continue.

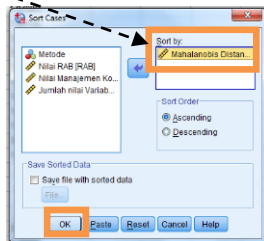


	Metode	RAB	MK	Jumlah	MAH_1
1	1	60	61	121	2.65227
2	1	61	61	122	1.79546
3	1	62	62	124	1.07985
4	1	63	63	126	1.05750
5	1	64	64	128	1.72842
6	1	65	65	130	3.09260
7	1	60	60	120	3.20433
8	1	61	61	122	1.79546
9	1	62	62	124	1.07985
10	1	63	63	126	1.05750
11	1	64	64	128	1.72842
12	1	65	65	130	3.09260
13	2	61	61	122	1.79546
14	2	61	63	124	,91565
15	2	62	64	126	,20212
16	2	63	65	128	,18185
17	2	64	66	130	,85485
18	2	65	67	132	2.22111
19	2	60	62	122	2.32245
20	2	61	63	124	,91565
21	2	62	64	126	,20212
22	2	63	65	128	,18185
23	2	64	66	130	,85485
24	2	65	67	132	2.22111

Setelah muncul variabel baru (MAH_1)/Mahalanobis, Klik data, lalu pilih Sort Cases.



Jika sudah muncul kotak dialog Sort Cases, pindahkan variabel Mahalanobis ke dalam kotak Sort by, lalu OK.



Perintah dari Sort Cases akan mengurutkan nilai variabel Mahalanobis dari yang terkecil ke yang terbesar.

Metode	RAB	MK	Jumlah	MAH_1
1	63	65	128	.18
2	63	65	128	.18
3	62	64	126	.20
4	62	64	126	.20
5	64	66	130	.85
6	64	66	130	.85
7	61	63	124	.92
8	61	63	124	.92
9	63	63	126	1.06
10	63	63	126	1.06
11	62	62	124	1.08
12	62	62	124	1.08
13	64	64	128	1.73
14	64	64	128	1.73
15	61	61	122	1.80
16	61	61	122	1.80
17	61	61	122	1.80
18	63	70	133	1.88
19	63	70	133	1.88
20	62	69	131	1.90
21	62	69	131	1.90
22	65	67	132	2.22
23	65	67	132	2.22
24	60	62	122	2.22

raphs Utilities Add-ons Window Help

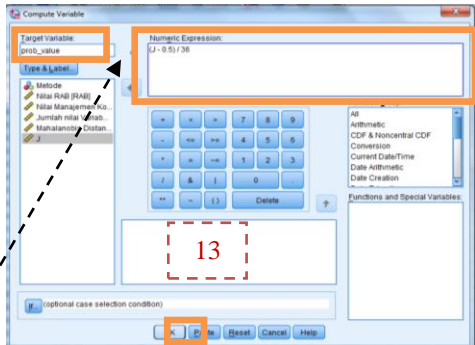
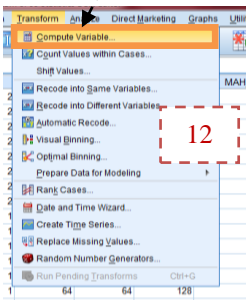
MAH_1	J	var
.18	1	
.18	2	
.20	3	
.20	4	
.85	5	
.85	6	
.92	7	
.92	8	
1.06	9	
1.06	10	
1.08	11	
1.08	12	
1.73	13	
1.73	14	
1.80	15	
1.80	16	
1.80	17	
1.88	18	
1.88	19	
1.90	20	
1.90	21	
2.22	22	
2.22	23	
2.22	24	

Nilai dari variabel "J" dapat diambil dari nomor banyaknya jumlah sampel, yaitu 36, diurutkan dari yang paling kecil ke yang paling besar, seperti tampak pada gambar 11.

Langkah selanjutnya, buat variabel baru semisal diberi nama "J"

Name	Type	Width	Decimals	Label
1 Metode	Numeric	8	0	
2 RAB	Numeric	8	0	Nilai RAB
3 MK	Numeric	8	0	Nilai Manajemen...
4 Jumlah	Numeric	8	0	Jumlah nilai Va...
5 MAH_1	Numeric	11	2	Mahalanobis Di...
6 J	Numeric	8	0	
7				

Langkah selanjutnya membuat variabel Probabilitas Value dari variabel "J", caranya: pilih Transform, lalu Compute Variable, akan muncul kotak dialog Compute Variable.



- Setelah muncul kotak dialog Compute Variable, ketik "prob_value" pada kotak Target Variable.
- Masukkan variabel "(J-0.5)/jumlah sampel" pada kotak Numeric Expression dengan menggunakan tombol yang ada pada kotak dialog Compute Variabel, lalu OK.

13

14

15

16

Setelah muncul variabel prob_value. Klik Transform, Compute Variable.

chi

Target Variable: chi

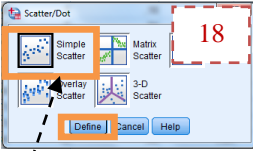
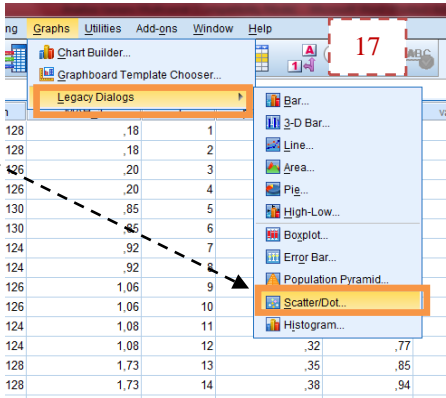
Numeric Expression: IDF.CHISQ(prob_value,2)

IDF.CHISQ(prob, df). Numeric. Returns the value from the chi-square distribution, with the specified degrees of freedom df, for which the cumulative probability is prob. For example, the chi-square value that is significant at the 0.05 level with 3 degrees of freedom is IDF.CHISQ(0.95,3).

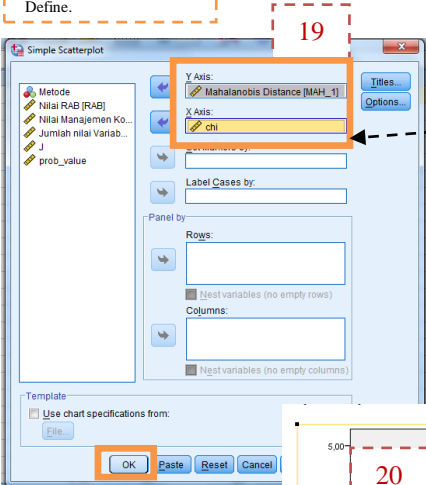
Setelah muncul kotak dialog Compute Variabel, tuliskan label variabel baru "chi" pada kotak target variabel.

- Lalu pilih Idf Chisq pada kotak Functions and Special Variables,
- Setelah muncul Idf Shisq (prob), pindahkan ke kotak Numeric Expression
- Akan muncul tanda tanya, lalu ganti tanda tanya pertama dengan variabel prob_value dan tanya tanya kedua diganti dengan jumlah variabel terikat, yaitu 2. Lalu klik OK.
- Lalu akan muncul variabel baru dengan nama "chi", seperti pada box 16.

Langkah selanjutnya pilih Graphs, Legacy Dialogs, lalu Scatter/Dot.



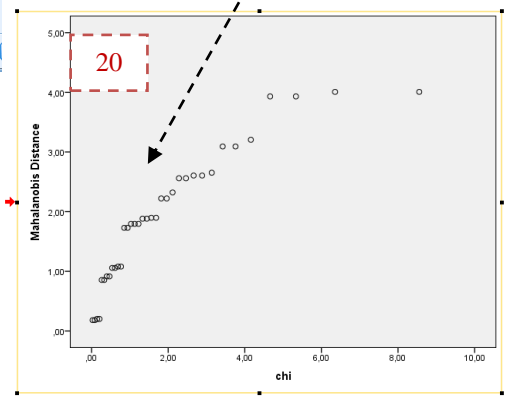
Setelah muncul kotak dialog Scatter/Dot, pilih Simple Scatter lalu Define.



Setelah muncul kotak dialog Simple Scatterplot, masukkan variabel Mahalanobis ke kotak Y axis dan variabel chi ke kotak x axis, lalu OK.

Grafik 20 menunjukkan adanya kecenderungan sebaran data membentuk garis lurus. Hal tersebut menunjukkan bahwa sebaran data variabel terikat, berdistribusi normal multivariat.

Berdasarkan grafik 20, pernyataan normalitas multivariat dapat diperkuat dengan melakukan uji korelasi, antara variabel Mahalanobis dan chi



Berdasarkan nilai pada pada tabel 21, nilai korelasi antara Mahalanobis dan chi menunjukkan nilai korelasi yang cukup tinggi, yaitu 0,918, hal tersebut makin menguatkan pernyataan bahwa sebaran data kedua variabel terikat berdistribusi normal multivariat. Karena asumsi normalitas telah terpenuhi, boleh melanjutkan pada analisis selanjutnya.

21

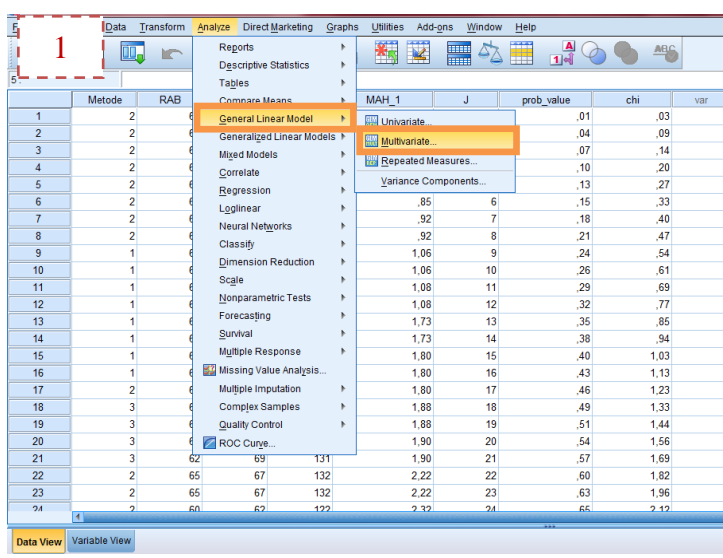
Correlations			
		Mahalanobis Distance	chi
Mahalanobis Distance	Pearson Correlation	1	,918**
	Sig. (2-tailed)		,000
	N	36	36
chi	Pearson Correlation	,918**	1
	Sig. (2-tailed)	,000	
	N	36	36

** . Correlation is significant at the 0.01 level (2-tailed).

d. Melakukan uji homogenitas matriks varian kovarian, uji tersebut dapat dilakukan bersamaan dengan analisis MANOVA. Pengambilan keputusan akan dijelaskan pada tabel no.4 pada rangkaian analisis MANOVA.

e. Melakukan uji MANOVA

- 1) Buka file MANOVA.
- 2) Pilih menu **Analyze** lalu **General Linier Model** lalu **Multivariate**, seperti pada gambar.



2

- Setelah muncul kotak dialog Multivariate, masukkan variabel nilai RAB dan nilai MK ke dalam kotak Dependent Variables.
- Lalu variabel metode mengajar dimasukkan ke dalam kotak Fixed Factor (s).
- Pilih Options.
- OK.

3

- Setelah muncul kotak dialog Multivariate Options, masukkan variabel metode ke dalam kotak Display Means for, lalu centang Compare Main Effects.
- Pada kotak Display, centang Descriptive Statistics, Estimate of effect size, dan Homogeneity tests.
- Pilih continue dan OK.
- Akan ada banyak tabel yang muncul tapi kita akan berfokus pada beberapa tabel saja seperti: Box's test of equality of covariance matrices, Multivariate tests, dan Tests of between subjects effects.

Uji Homogenitas

4

Box's Test of Equality of Covariance Matrices	
Box's M	8,321
F	2,500
df1	3
df2	87120,000
Sig.	,058

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + Metode

Tabel 4 menunjukkan output asumsi homogenitas matriks varian kovarian. Berbeda dengan analisis analisis lainnya, pengambilan keputusan didasarkan pada nilai Sig. yang batasan nilainya adalah 0,001, karena analisis ini sangatlah teliti. Prinsipnya, bila nilai Sig. yang diperoleh diatas 0,001, maka dapat disimpulkan asumsi homogenitas terpenuhi.

Nilai Box'M yang dihasilkan adalah sebesar 8,321 ($P = 0,058$). Dengan demikian matriks kovarian antarkelompok diasumsikan sama.

Uji Hipotesis MANOVA

5

Multivariate Tests ^a							
Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	,999	23139,304 ^b	2,000	32,000	,000	,999
	Wilks' Lambda	,001	23139,304 ^b	2,000	32,000	,000	,999
	Hotelling's Trace	1446,207	23139,304 ^b	2,000	32,000	,000	,999
	Roy's Largest Root	1446,207	23139,304 ^b	2,000	32,000	,000	,999
Metode	Pillai's Trace	,986	16,043	4,000	66,000	,000	,493
	Wilks' Lambda	,015	116,737 ^b	4,000	64,000	,000	,879
	Hotelling's Trace	67,790	525,374	4,000	62,000	,000	,971
	Roy's Largest Root	67,790	1118,530 ^c	2,000	33,000	,000	,985

a. Design: Intercept + Metode
 b. Exact statistic
 c. The statistic is an upper bound on F that yields a lower bound on the significance level.

Karena semua uji asumsi telah kita lakukan maka saatnya melihat hasil analisis MANOVA nya. Tabel uji MANOVA terlihat pada tabel 5. Fokus pada kotak metode, kotak Intercept dapat diabaikan. Perhatikan hasil perhitungan Wilks' Lambda nya.

Dengan $F(4, 64) = 116,737, P < 0,001$; $Wilks' \Lambda = 0,015, \eta^2 = 0,879$ ← effect size
 Berdasarkan tabel "Multivariate tests" hasil uji MANOVA menginformasikan terdapat perbedaan signifikan nilai ujian mahasiswa yang berasal dari metode mengajar yang berbeda. Dengan demikian pada Dugaan ke-3 H_0 gagal ditolak (diterima).

Uji Hipotesis univariat

Pada tabel Tests of Between Subjects Effects, yang perlu diperhatikan adalah kotak Metode dan Error.

6

Tests of Between-Subjects Effects							
Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Nilai RAB	,056 ^a	2	,028	,009	,991	,001
	Nilai Manajemen Konstruksi	308,167 ^b	2	154,083	47,595	,000	,743
Intercept	Nilai RAB	140750,028	1	140750,028	46025,608	,000	,999
	Nilai Manajemen Konstruksi	154449,000	1	154449,000	47708,115	,000	,999
Metode	Nilai RAB	,056	2	,028	,009	,991	,001
	Nilai Manajemen Konstruksi	308,167	2	154,083	47,595	,000	,743
Error	Nilai RAB	100,917	33	3,058			
	Nilai Manajemen Konstruksi	106,836	33	3,237			
Total	Nilai RAB	140851,000	36				
	Nilai Manajemen Konstruksi	154664,000	36				
Corrected Total	Nilai RAB	100,972	35				
	Nilai Manajemen Konstruksi	115,000	35				

a. R Squared = ,001 (Adjusted R Squared = -.066)
b. R Squared = ,743 (Adjusted R Squared = ,727)

Selanjutnya dengan [$F(2, 33) = 0,009, P > 0,991; \eta^2 = 0,001$] hasil univariat menunjukkan bahwa perbedaan metode mengajar **tidak** menyebabkan **perbedaan signifikan** pada hasil **ujian RAB**, **namun** perbedaan metode mengajar **menyebabkan perbedaan signifikan** pada nilai **ujian Manajemen Konstruksi** [$F(2, 33) = 47,595, P < 0,001; \eta^2 = 0,743$]

Berdasarkan tabel "tests of between subjects effects," hasil uji univariatnya menginformasikan dengan perbedaan metode mengajar, menyebabkan **perbedaan signifikan** pada nilai ujian Manajemen Konstruksi. Oleh karena itu, yang perlu dilakukan uji lanjut adalah variabel hasil ujian Manajemen Konstruksi.

7 **Tabel Uji**

Uji lanjut bisa kita lihat pada tabel Pairwise Comparisons.

Pairwise Comparisons

Dependent Variable	(I) Metode	(J) Metode	Mean Difference (I-J)	Std. Error	Sig. ^b	95% Confidence Interval for Difference ^b	
						Lower Bound	Upper Bound
Nilai RAB	Metode Project Based Learning	Metode Cooperative Learning	-,083	,714	,908	-1,536	1,369
		Metode Problem Based Learning	1,040E-014	,714	1,000	-1,452	1,452
	Metode Cooperative Learning	Metode Project Based Learning	,083	,714	,908	-1,369	1,536
		Metode Problem Based Learning	,083	,714	,908	-1,369	1,536
	Metode Problem Based Learning	Metode Project Based Learning	-1,040E-014	,714	1,000	-1,452	1,452
		Metode Cooperative Learning	,083	,714	,908	-1,369	1,536
Nilai Manajemen Konstruksi	Metode Project Based Learning	Metode Cooperative Learning	-1,833	,735	,018	-3,328	-,339
		Metode Problem Based Learning	-6,917 [*]	,735	,000	-8,411	-5,422
	Metode Cooperative Learning	Metode Project Based Learning	1,833	,735	,018	,339	3,328
		Metode Problem Based Learning	-5,083 [*]	,735	,000	-6,578	-3,589
	Metode Problem Based Learning	Metode Project Based Learning	6,917 [*]	,735	,000	5,422	8,411
		Metode Cooperative Learning	5,083 [*]	,735	,000	3,589	6,578

Based on estimated marginal means

^{*}. The mean difference is significant at the .05 level.

^b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

J. Rangkuman Uji MANOVA

- a. Berdasarkan Tabel 6.7, berikut akan diuji asumsi homogenitas matriks-matriks kovarian populasi.

Tabel 6.7 Box's M Test of Equality of Covariance Matrices.

Box's Test of Equality of Covariance Matrices	
^a	
Box's M	8,321
F	2,500
df1	3
df2	87120,000
Sig.	,058

Hipotesis nol menyatakan tidak terjadi perbedaan yang signifikan secara statistik di antara matriks-matriks kovarian populasi (*equal population*)

Nilai kritis F dengan derajat bebas pembilang adalah 3, derajat bebas penyebut adalah 87120, dan tingkat signifikansi yang digunakan $\alpha = 0,05$ adalah 2,605 sehingga diketahui nilai statistik dari uji F adalah 2,5. Berikut aturan pengambilan keputusan berdasarkan uji F .

Jika nilai statistik dari uji F

\leq nilai kritis F , maka H_0 diterima dan H_1 ditolak.

jika nilai statistik dari uji F

$>$ nilai kritis F , maka H_0 ditolak dan H_1 diterima.

Karena nilai nilai statistik dari uji F lebih kecil dibandingkan nilai kritis F , maka hipotesis nol diterima dan hipotesis alternatif ditolak. Hal ini berarti asumsi homogenitas matriks-matriks kovarian populasi dipenuhi.

- b. Berdasarkan Tabel 6.9, nilai rata-rata ujian RAB dengan metode *cooperative learning* (62,58) lebih tinggi dibandingkan metode PjBL (62,50) dan metode PBL (62,50). Diketahui juga bahwa nilai rata-rata ujian bahasa Manajemen Konstruksi dengan metode PBL (69,50) lebih tinggi dibandingkan metode PjBL (62,58) dan metode *cooperative learning* (64,42).

Tabel 6.9 *Descriptive Statistics.*

Descriptive Statistics				
	Metode	Mean	Std. Deviation	N
Nilai RAB	Metode Project Based Learning	62,50	1,784	12
	Metode Cooperative Learning	62,58	1,676	12
	Metode Problem Based Learning	62,50	1,784	12
	Total	62,53	1,699	36
Nilai Manajemen Konstruksi	Metode Project Based Learning	62,58	1,676	12
	Metode Cooperative Learning	64,42	1,929	12
	Metode Problem Based Learning	69,50	1,784	12
	Total	65,50	3,443	36

c. Berdasarkan Tabel 6.10, akan ditentukan apakah ketiga metode mengajar tersebut memiliki kemampuan yang sama dalam mempengaruhi nilai ujian RAB dan MK secara simultan (*simultaneously*) (mempengaruhi secara rata-rata) atau tidak. Untuk menentukan apakah ketiga metode mengajar tersebut memiliki kemampuan yang sama dalam mempengaruhi nilai ujian RAB dan Manajemen Konstruksi secara simultan (*simultaneously*) (mempengaruhi secara rata-rata) atau tidak, dapat dilakukan dengan membandingkan nilai *Sig.* dari *Pillai's Trace*, *Wilks' Lambda*, *Hotelling's Trace*, atau *Roy's Largest Root* terhadap tingkat signifikansi. Perhatikan bahwa, karena keempat nilai *Sig.* tersebut lebih kecil dibandingkan tingkat signifikansi $\alpha = 0,05$, maka disimpulkan bahwa terdapat metode mengajar yang memiliki kemampuan berbeda dalam hal mempengaruhi nilai ujian RAB dan manajemen konstruksi secara simultan (*simultaneously*).

Tabel 6.10 *Multivariate Tests.*

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	,999	23139,304 ^b	2,000	32,000	,000	,999
	Wilks' Lambda	,001	23139,304 ^b	2,000	32,000	,000	,999
	Hotelling's Trace	1446,207	23139,304 ^b	2,000	32,000	,000	,999
	Roy's Largest Root	1446,207	23139,304 ^b	2,000	32,000	,000	,999
Metode	Pillai's Trace	,986	16,043	4,000	66,000	,000	,493
	Wilks' Lambda	,015	116,737 ^b	4,000	64,000	,000	,879
	Hotelling's Trace	67,790	525,374	4,000	62,000	,000	,971
	Roy's Largest Root	67,790	1118,530 ^c	2,000	33,000	,000	,985

a. Design: Intercept + Metode

b. Exact statistic

c. The statistic is an upper bound on F that yields a lower bound on the significance level.

d. Berdasarkan informasi Tabel 6.11, nilai *Sig.* dari variabel nilai ujian RAB adalah 0,961. Karena nilai *Sig.* tersebut lebih besar dibandingkan

tingkat signifikansi $\alpha = 0,05$, maka asumsi kesamaan varian populasi nilai ujian RAB berdasarkan kategori-kategori pada metode mengajar dipenuhi.

Tabel 6.11 *Levene's Test of Equality of Error Variances.*

Levene's Test of Equality of Error Variances^a

	F	df1	df2	Sig.
Nilai RAB	,040	2	33	,961
Nilai Manajemen Konstruksi	,107	2	33	,899

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + Metode

Nilai Sig. dari variabel nilai ujian Manajemen Konstruksi adalah 0,899. Karena nilai Sig. tersebut lebih besar dibandingkan tingkat signifikansi $\alpha = 0,05$, maka asumsi kesamaan varian populasi nilai ujian Manajemen Konstruksi berdasarkan kategori-kategori pada metode mengajar dipenuhi.

e. **Tabel 6.12** *Tests of Between-Subjects Effects.*

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Nilai RAB	,056 ^a	2	,028	,009	,991	,001
	Nilai Manajemen Konstruksi	308,167 ^b	2	154,083	47,595	,000	,743
Intercept	Nilai RAB	140750,028	1	140750,028	46025,608	,000	,999
	Nilai Manajemen Konstruksi	154449,000	1	154449,000	47708,115	,000	,999
Metode	Nilai RAB	,056	2	,028	,009	,991	,001
	Nilai Manajemen Konstruksi	308,167	2	154,083	47,595	,000	,743
Error	Nilai RAB	100,917	33	3,058			
	Nilai Manajemen Konstruksi	106,833	33	3,237			
Total	Nilai RAB	140851,000	36				
	Nilai Manajemen Konstruksi	154864,000	36				
Corrected Total	Nilai RAB	100,972	35				
	Nilai Manajemen Konstruksi	415,000	35				

a. R Squared = ,001 (Adjusted R Squared = -,060)
b. R Squared = ,743 (Adjusted R Squared = ,727)

Berdasarkan Tabel 6.12, selanjutnya akan diperiksa hal-hal sebagai berikut.

- 1) Apakah penerapan dari ketiga metode mengajar tersebut berpengaruh terhadap nilai ujian RAB (secara rata-rata). Dengan kata lain, dari tiga metode mengajar tersebut, apakah terdapat metode mengajar yang memiliki kemampuan yang berbeda dalam hal mempengaruhi nilai ujian RAB.
- 2) Apakah penerapan dari ketiga metode mengajar tersebut berpengaruh terhadap ujian Manajemen Konstruksi. Dengan kata lain, dari tiga metode mengajar tersebut, apakah terdapat metode mengajar yang memiliki kemampuan yang berbeda dalam hal mempengaruhi nilai ujian Manajemen Konstruksi.

Pada baris metode, nilai Sig. dari variabel tak bebas nilai ujian RAB adalah 0,991. Karena nilai Sig. lebih besar dibandingkan tingkat signifikansi $\alpha = 0,05$, maka disimpulkan bahwa penerapan dari ketiga metode mengajar tersebut tidak berpengaruh terhadap nilai ujian RAB (secara rata-rata). Selanjutnya pada baris metode, nilai Sig. dari variabel tak bebas nilai ujian Manajemen Konstruksi adalah 0,000. Karena nilai Sig. lebih kecil dibandingkan tingkat signifikansi $\alpha = 0,05$, maka disimpulkan bahwa penerapan dari ketiga metode mengajar tersebut berpengaruh terhadap ujian Manajemen Konstruksi.

f. **Tabel 6.13** *Parwise Comparisons.*

Pairwise Comparisons

Dependent Variable	(i) Metode	(j) Metode	Mean Difference (i-j)	Std. Error	Sig. ^a	95% Confidence Interval for Difference ^b	
						Lower Bound	Upper Bound
Nilai RAB	Metode Project Based Learning	Metode Cooperative Learning	-.083	.714	.908	-1,536	1,369
		Metode Problem Based Learning	1,040E-014	.714	1,000	-1,452	1,452
		Metode Cooperative Learning	.083	.714	.908	-1,369	1,536
	Metode Cooperative Learning	Metode Project Based Learning	.083	.714	.908	-1,369	1,536
		Metode Problem Based Learning	.083	.714	.908	-1,369	1,536
		Metode Project Based Learning	-1,040E-014	.714	1,000	-1,452	1,452
Nilai Manajemen Konstruksi	Metode Project Based Learning	Metode Cooperative Learning	-1,833 [*]	.735	.018	-3,328	-.339
		Metode Problem Based Learning	-6,917 [*]	.735	.000	-8,411	-5,422
		Metode Cooperative Learning	1,833 [*]	.735	.018	.339	3,328
	Metode Cooperative Learning	Metode Project Based Learning	-5,083 [*]	.735	.000	-6,578	-3,589
		Metode Problem Based Learning	6,917 [*]	.735	.000	5,422	8,411
		Metode Project Based Learning	5,083 [*]	.735	.000	3,589	6,578

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Karena terdapat pengaruh yang signifikan secara statistik pada penggunaan metode mengajar terhadap nilai ujian Manajemen Konstruksi, maka akan ditentukan metode mengajar manakah yang membuat nilai rata-rata ujian Manajemen Konstruksi tinggi. Maka akan dilakukan hal sebagai berikut: Membandingkan nilai rata-rata ujian Manajemen Konstruksi antara: metode PjBL v/s metode *cooperative learning*, metode PjBL v/s metode PBL, dan metode *cooperative learning* v/s metode PBL.

Selanjutnya akan dilakukan uji perbandingan berganda (*multiple comparison test*) dengan maksud untuk mengetahui metode mengajar manakah yang membuat nilai rata-rata ujian Manajemen Konstruksi tinggi. Berdasarkan Tabel *Pairwise Comparisons*, dapat ditarik informasi sebagai berikut.

- 1) Bahwa tidak terdapat perbedaan (secara rata-rata) yang signifikan secara statistik mengenai nilai ujian RAB ketika diterapkan metode mengajar PjBL dan *cooperative learning*. Hal ini dapat dilihat dari

nilai Sig. yang lebih besar dari tingkat signifikansi yang digunakan 5% ($0,908 \geq 0,05$).

- 2) Bahwa tidak terdapat perbedaan yang signifikan secara statistik mengenai nilai ujian RAB ketika diterapkan metode mengajar PjBL dan PBL. Hal ini dapat dilihat dari nilai Sig. yang lebih besar dari tingkat signifikansi yang digunakan 5% ($1 \geq 0,05$).
- 3) Bahwa tidak terdapat perbedaan yang signifikan secara statistik mengenai nilai ujian RAB ketika diterapkan metode mengajar *cooperative learning* dan PBL. Hal ini dapat dilihat dari nilai Sig. yang lebih besar dari tingkat signifikansi yang digunakan 5% ($0,908 \geq 0,05$).
- 4) Bahwa **terdapat perbedaan** yang signifikan secara statistik mengenai nilai ujian Manajemen Konstruksi ketika diterapkan metode mengajar PjBL dan *cooperative learning*. Hal ini dapat dilihat dari nilai Sig. yang **lebih kecil** dari tingkat signifikansi yang digunakan 5% ($0,018 < 0,05$).
- 5) Bahwa **terdapat perbedaan** yang signifikan secara statistik mengenai nilai ujian Manajemen Konstruksi ketika diterapkan metode mengajar metode PjBL dan PBL. Hal ini dapat dilihat dari nilai Sig. yang **lebih kecil** dari tingkat signifikansi yang digunakan 5% ($0,000 < 0,05$).
- 6) Bahwa **terdapat perbedaan** yang signifikan secara statistik mengenai nilai ujian Manajemen Konstruksi ketika diterapkan metode mengajar metode *cooperative learning* dan PBL. Hal ini dapat dilihat dari nilai Sig. yang **lebih kecil** dari tingkat signifikansi yang digunakan 5% ($0,000 < 0,05$).

Berdasarkan uraian tersebut, disimpulkan bahwa metode mengajar PBL memberikan pengaruh yang signifikan secara statistik dalam hal membuat nilai ujian Manajemen Konstruksi tinggi (secara rata-rata).

Daftar Rujukan

- Field, A. (2009). *Discovering Statistics Using SPSS, 3rd Edition*. London: Sage.
- Hair, J.F.Jr., Anderson R.E., Babin B.J., dan Black W.C. (2010). *Multivariate Data Analysis, 7th Edition*. Pearson Prentice Hall.
- Mardani R. Retrieved Februari 24 2017, http://eprints.ums.ac.id/30733/2/BAB_I.pdf
- Mattjik, A.A. dan Sumertajaya, I.M. (2011). *Sidik Peubah Ganda dengan Menggunakan SAS*. Bogor: IPB Press.
- Oja, H. (2010). *Multivariate Nonparametric Methods with R: An Approach Based on Spatial Signs and Ranks*. New York: Springer Science+Business Media.
- Rencher, A.C. (1998). *Multivariate Statistical Inference and Applications*. Kanada: John Wiley and Sons, Inc.
- Rencher, A.C. (2002). *Methods of Multivariate Analysis*. Kanada: John Wiley and Sons, Inc.
- Richard A.J., dan Dean W.W. (2007). *Applied Multivariate Statistical Analysis*. United States of America: Pearson Education, Inc.
- Stevens, J.P. (2009). *Applied Multivariate Statistics For The Social Science, 5th Edition*. New York: Routledge.
- Sutrisno, S., & Wulandari, D. (2018). *Multivariate Analysis of Variance (MANOVA) untuk Memperkaya Hasil Penelitian Pendidikan*. *AKSIOMA: Jurnal Matematika Dan Pendidikan Matematika*, 9(1), 37-53.
- Tabachnick, B.G. dan Fidell, L.S. (2007). *Using Multivariate Statistics*. Boston: Pearson Education, Inc.

- Timm, N.H. (2002). *Applied Multivariate Analysis*. New York:
SpringerVerlag.
- Walpole, R.E. (2011). *Probability and Statistics For Engineers and
Scientist 9th Edition*. Boston : Prentise Hall
- Walpole, R.E. (1993). *Pengantar Statistika*. Jakarta: PT. Gramedia Pustaka
Utama.

BAB 7

ANALISIS FAKTOR

A. Pengenalan Analisis Faktor

1. Pengertian Analisis Faktor

Analisis faktor adalah kajian tentang saling ketergantungan antara variabel-variabel, dengan tujuan untuk menemukan himpunan variabel-variabel baru yang lebih sedikit jumlahnya dari pada variabel semula, dan yang menunjukkan yang mana di antara variabel-variabel semula itu yang merupakan faktor-faktor persekutuan (Suryanto, 1988:234). Dalam analisis faktor, variabel-variabel dalam jumlah besar dikelompokkan dalam sejumlah faktor yang mempunyai sifat dan karakteristik yang hampir sama sehingga lebih mempermudah pengolahan.

Alasan umum dibalik metode analitik faktor adalah bahwa informasi yang diperoleh tentang saling ketergantungan antara variabel yang diamati dapat digunakan nanti untuk mengurangi kumpulan variabel dalam kumpulan data. Analisis faktor umumnya digunakan dalam psikometri, teori kepribadian, biologi, pemasaran, manajemen produk, riset operasi, keuangan, dan pembelajaran mesin. Ini mungkin membantu untuk menangani kumpulan data yang memiliki sejumlah besar variabel yang diamati, yang dianggap mencerminkan sejumlah kecil variabel yang mendasari/laten. Ini adalah salah satu teknik saling ketergantungan yang paling umum digunakan ketika kumpulan variabel yang relevan menunjukkan saling ketergantungan yang sistematis untuk mengetahui faktor laten yang menciptakan kesamaan.

2. Tujuan Analisis Faktor

Tujuan utama analisis faktor adalah untuk menjelaskan struktur

hubungan di antara banyak variabel dalam bentuk faktor atau variabel laten atau variabel bentukan. Faktor yang terbentuk merupakan besaran acak (*random quantities*) yang sebelumnya tidak dapat diamati atau diukur atau ditentukan secara langsung. Selain tujuan utama tersebut, terdapat tujuan lainnya, yaitu

- a. Untuk mereduksi sejumlah variabel asal yang jumlahnya banyak menjadi sejumlah variabel baru yang jumlahnya lebih sedikit dari variabel asal, dan variabel baru tersebut dinamakan faktor atau variabel laten atau konstruk atau variabel bentukan.
- b. Untuk mengidentifikasi adanya hubungan antarvariabel penyusun faktor atau dimensi dengan faktor yang terbentuk, dengan menggunakan pengujian koefisien korelasi antar faktor dengan komponen pembentuknya. Analisis faktor ini disebut analisis faktor konfirmatori.
- c. Untuk menguji validitas dan reliabilitas instrumen dengan analisis faktor konfirmatori.
- d. Validasi data untuk mengetahui apakah hasil analisis faktor tersebut dapat digeneralisasi ke dalam populasinya sehingga setelah terbentuk faktor peneliti sudah mempunyai suatu hipotesis baru berdasarkan hasil analisis tersebut.

3. Perbedaan Analisis Faktor Eksploratori dan Konfirmatori

Analisis faktor eksploratori atau analisis komponen utama (PCA) adalah suatu teknik analisis faktor yang memiliki beberapa faktor yang akan terbentuk berupa variabel laten yang belum dapat ditentukan sebelum analisis dilakukan. Pada prinsipnya, terbentuknya faktor-faktor atau variabel laten baru dalam analisis faktor eksploratori bersifat acak, yang selanjutnya dapat diinterpretasi sesuai dengan faktor atau komponen atau konstruk yang terbentuk. Analisis faktor eksploratori persis sama dengan analisis komponen

utama (PCA). Dalam analisis faktor eksploratori, peneliti tidak atau belum mempunyai pengetahuan atau teori atau suatu hipotesis yang menyusun struktur faktor-faktornya yang akan dibentuk atau yang terbentuk. Oleh karena itu, analisis faktor eksploratori menjadi teknik untuk membantu membangun teori baru. Analisis faktor eksploratori merupakan suatu teknik untuk mereduksi data dari variabel asal atau variabel awal menjadi variabel baru atau faktor yang jumlahnya lebih kecil dari pada variabel awal. Proses analisis tersebut mencoba untuk menemukan hubungan antarvariabel baru atau faktor yang terbentuk yang saling independen sesamanya sehingga bisa dibuat satu atau beberapa kumpulan variabel laten atau faktor yang lebih sedikit dari jumlah variabel awal yang bebas atau tidak berkorelasi sesamanya. Jadi, antarfaktor yang terbentuk tidak berkorelasi sesamanya.

Analisis faktor konfirmatori adalah suatu teknik analisis faktor secara apriori berdasarkan teori dan konsep yang sudah diketahui dipahami atau ditentukan sebelumnya lalu dibuat sejumlah faktor yang akan dibentuk, serta variabel apa saja yang termasuk ke dalam masing-masing faktor yang dibentuk dan sudah pasti tujuannya. Pembentukan faktor konfirmatori (CFA) secara sengaja berdasarkan teori dan konsep dalam upaya untuk mendapatkan variabel baru atau faktor yang mewakili beberapa item atau sub-variabel yang merupakan variabel teramati atau *observable variable*. Pada dasarnya, tujuan analisis faktor konfirmatori adalah untuk mengidentifikasi adanya hubungan antar variabel dengan melakukan uji korelasi dan untuk menguji validitas dan reliabilitas instrumen. Dalam pengujian terhadap validitas dan reliabilitas, instrumen atau kuesioner untuk mendapatkan data penelitian yang valid dan reliabel dengan analisis faktor konfirmatori.

B. Analisis Faktor Eksploratori

Analisis multivariat (*multivariate analysis*) merupakan salah satu jenis analisis statistik yang digunakan untuk menganalisis data yang terdiri dari banyak variabel, baik variabel bebas (*independent variable*) maupun variabel tak bebas (*dependent variable*). Data multivariat adalah data yang dikumpulkan dari dua atau lebih observasi dengan mengukur observasi tersebut dengan beberapa karakteristik. Selanjutnya, dalam analisis dibagi menjadi dua kategori metode, yaitu metode dependensi dan interdependensi. Model pertama terdapat dua jenis variabel, yaitu variabel bebas dan tergantung; sedang model kedua hanya terdapat satu jenis variabel, yaitu variabel bebas. Metode dependensi terdiri atas beberapa teknik analisis, yaitu regresi berganda, analisis diskriminan, korelasi kanonikal dan MANOVA. Metode interdependensi terdiri atas beberapa teknik analisis, yaitu analisis faktor, analisis kluster dan *multidimensional scalling*.

1. Model Analisis Faktor

Setelah variabel ditentukan dan dipilih serta perhitungan korelasinya telah memenuhi persyaratan untuk dilakukan analisis, langkah selanjutnya adalah membentuk faktor untuk menemukan struktur yang mendasari hubungan antar variabel awal tersebut. Metode yang sering digunakan dalam analisis faktor eksploratori adalah metode *principal component*. Lebih lanjut, dalam pembahasan ini akan dibatasi pada metode *principal component* dengan rotasi ortogonal.

Secara umum analisis faktor ortogonal disusun seperti model dalam analisis regresi multivariat. Setiap variabel awal dinyatakan sebagai kombinasi linear dari faktor-faktor yang mendasari. Misalkan, vektor acak \mathbf{X} dengan banyak komponen p dan mempunyai mean $\boldsymbol{\mu}$ dan matriks kovariani Σ merupakan penyusunan model faktor. Variabel F_1, F_2, \dots, F_m merupakan faktor yang nilainya tidak terobservasi, $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ merupakan kesalahan

(error) atau faktor spesifik. Secara matematis, model analisis faktor ditulis sebagai berikut (Johnson, 2007).

$$\begin{aligned} X_1 - \mu_1 &= P_{11}F_1 + P_{12}F_2 + \dots + P_{1m}F_m + \epsilon_1 \\ X_2 - \mu_2 &= P_{21}F_1 + P_{22}F_2 + \dots + P_{2m}F_m + \epsilon_2 \\ X_p - \mu_p &= P_{p1}F_1 + P_{p2}F_2 + \dots + P_{pm}F_m + \epsilon_p \end{aligned} \quad (2.1)$$

atau dalam notasi matriks ditulis sebagai berikut:

$$X(\mathbf{p} \times \mathbf{1}) - \boldsymbol{\mu}(\mathbf{p} \times \mathbf{1}) = L(\mathbf{p} \times \mathbf{m})\mathbf{F}(\mathbf{m} \times \mathbf{1}) + \boldsymbol{\varepsilon}(\mathbf{p} \times \mathbf{1}) \quad (2.2)$$

dengan:

L : matriks faktor *loading*.

X_1 : vektor acak yang memiliki p komponen pada amatan ke- i

μ_i : rata-rata dari variabel ke- i

F_j : faktor bersama (*common factor*) yang ke ke- j atau disebut faktor-faktor umum

ϵ_i : sisaan atau eror dari variabel ke- i (*specific factor*) atau disebut faktor khusus

Asumsi yang harus dipenuhi dalam model (2.2) adalah

$$E(F) = 0_{(m \times n)}, Cov(F) = E[FF'] = I_{m \times m} \quad (2.3)$$

$$E(\epsilon) = 0_{(p \times 1)}, Cov(\epsilon) = E[\epsilon\epsilon'] = \psi_{m \times m} = \begin{bmatrix} \varphi_1 & 0 & \dots & 0 \\ 0 & \varphi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad (2.4)$$

$$F \text{ dan } \epsilon \text{ saling bebas, } Cov(\epsilon, F) = E[\epsilon, F] = 0_{p \times m} \quad (2.5)$$

Model ortogonal dari sebuah analisis faktor adalah

$$X_{(p \times 1)} = \mu_{(p \times 1)} + L_{(p \times m)}F_{(m \times 1)} + \epsilon_{(p \times 1)} \quad (2.6)$$

Model ortogonal dari analisis faktor berakibat kepada struktur kovariani untuk variabel acak X , yaitu

$$\begin{aligned}
 (\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})' &= (\mathbf{LF} + \boldsymbol{\varepsilon})(\mathbf{LF} + \boldsymbol{\varepsilon})' \\
 &= (\mathbf{LF} + \mathbf{s})(\mathbf{LF})' + \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \\
 &= \mathbf{LF}(\mathbf{LF})' + \mathbf{s}(\mathbf{LF})' + \mathbf{LF} \boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}' \quad (2.7)
 \end{aligned}$$

Berdasarkan pada persamaan (2.3) dan (2.4) maka matriks kovarianinya menjadi:

$$\begin{aligned}
 \boldsymbol{\Sigma} &= \text{Cov}(\mathbf{X}) = E(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})' \\
 &= E[\mathbf{LF}(\mathbf{LF})' + \boldsymbol{\varepsilon}(\mathbf{LF})' + \mathbf{LF} \boldsymbol{\varepsilon}' + \boldsymbol{\varepsilon} \boldsymbol{\varepsilon}'] \\
 &= \mathbf{LL}'E(\mathbf{FF}') + \mathbf{L}'E(\boldsymbol{\varepsilon} \mathbf{F}') + \mathbf{LE}(\mathbf{F} \boldsymbol{\varepsilon}') + E(\boldsymbol{\varepsilon} \boldsymbol{\varepsilon}') \\
 &= \mathbf{LL}' \cdot \mathbf{I} + \mathbf{L}' \cdot \mathbf{0} + \mathbf{L} \cdot \mathbf{0} + \boldsymbol{\psi} \\
 &= \mathbf{LL}' + \boldsymbol{\psi} \quad (2.8)
 \end{aligned}$$

atau dapat ditulis sebagai:

$$\text{Var}(X_i) = \ell_{i1}^2 + \dots + \ell_{im}^2 + \varphi_i \quad (2.9)$$

$$\text{Cov}(X_i, X_k) = \ell_{i1} \ell_{k1} + \dots + \ell_{im} \ell_{km} \quad (2.10)$$

Kovariani untuk variabel acak X dan faktor umum F adalah

$$\begin{aligned}
 (\mathbf{X} - \boldsymbol{\mu})\mathbf{F}' &= (\mathbf{LF} + \boldsymbol{\varepsilon})\mathbf{F}' \\
 &= \mathbf{LFF}' + \boldsymbol{\varepsilon} \mathbf{F}' \quad (2.11)
 \end{aligned}$$

Berdasarkan pada persamaan (2.3) dan (2.4) maka:

$$\begin{aligned}
 \text{Cov}(\mathbf{X}, \mathbf{F}) &= E(\mathbf{X} - \boldsymbol{\mu})\mathbf{F}' \\
 &= E[\mathbf{LFF}' + \boldsymbol{\varepsilon} \mathbf{F}'] \\
 &= \mathbf{LE}(\mathbf{FF}') + E(\boldsymbol{\varepsilon} \mathbf{F}') \\
 &= \mathbf{LI} + \mathbf{0} \\
 &= \mathbf{L} \quad (2.12)
 \end{aligned}$$

Secara umum dapat ditulis sebagai berikut:

$$\text{Cov}(\mathbf{X}_i, \mathbf{F}_j) = P_{ij} \text{ dengan } i = 1, 2, \dots, p \text{ dan } j = 1, 2, \dots, m \quad (2.13)$$

Variasi dari variabel ke- i yang disumbangkan oleh m common faktor disebut komunalitas ke- i . Nilai $\text{Var}(X_i) = \sigma_{ii}$ merupakan nilai komunalitas yang ditambahkan dengan nilai variansi spesifik atau uniknya. Dari persamaan (2.8) dan persamaan (2.9) diperoleh

$$\underbrace{\sigma_{ii}}_{\text{var}(X_i)} = \underbrace{\ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2}_{\text{komunalitas}} + \underbrace{\varphi_i}_{\text{variansi spesifik}} \quad (2.14)$$

Komunalitas ke- i dinotasikan sebagai h_i^2 , atau dengan menggunakan notasi di atas diperoleh persamaan

$$h_i^2 = \ell_{i1}^2 + \ell_{i2}^2 + \dots + \ell_{im}^2 \quad (2.15)$$

dan

$$\sigma_{ii} = h_i^2 + \varphi_i, i = 1, 2, \dots, p \quad (2.16)$$

Nilai komunalitas ke- i merupakan jumlah kuadrat dari faktor *loading* variabel ke- i pada m faktor.

2. Metode *Principal Component*

Tujuan khusus dari metode analisis faktor *principal component* adalah mengetahui struktur yang mendasari variabel-variabel awal dalam analisis dan melakukan penyederhanaan struktur sekumpulan variabel awal tersebut melalui reduksi data. Prosedur matematis untuk mencari struktur kovariani matriks Σ dapat dilakukan dengan menggunakan matriks dekomposisi spektral. Misal, Σ mempunyai pasangan nilai eigen dan vektor eigen (λ_i, e_i) dengan urutan nilai λ nya adalah $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, maka:

$$\Sigma = \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \dots + \lambda_p e_p e_p' \quad (2.17)$$

$$= [\sqrt{\lambda_1} e_1 | \sqrt{\lambda_2} e_2 | \dots | \sqrt{\lambda_p} e_p] \begin{bmatrix} \sqrt{\lambda_1} e_1' 1 \\ - - - \\ \sqrt{\lambda_2} e_2' 2 \\ - - - \\ \vdots \\ - - - \\ \sqrt{\lambda_2} e_2' 2 \end{bmatrix} \quad (2.18)$$

Model ini adalah gambaran struktur kovariani dari analisis faktor yang mempunyai variabel awal sama dengan jumlah faktor yang terbentuk ($m = p$) dan variansi spesifiknya $\varphi_i = \mathbf{0}$ untuk semua i . Matriks faktor *loading* pada kolom j dituliskan $\sqrt{\lambda_j} e_j$. Dalam bentuk umum dapat dituliskan sebagai berikut.

$$\Sigma_{(p \times p)} = \mathbf{L}_{(p \times p)} \mathbf{L}'_{(p \times p)} + \mathbf{0}_{(p \times p)} = \mathbf{L} \mathbf{L}' \quad (2.19)$$

Selanjutnya, faktor *loading* yang terbentuk tersebut merupakan koefisien faktor pada metode *principal component*. Dalam persamaan (2.19) belum sesuai dengan tujuan analisis faktor karena belum diperoleh jumlah faktor yang lebih sedikit dari variabel-variabel awalnya. Selain itu, beberapa variansi pada faktor spesifik ε belum dilibatkan. Untuk itu, dibuat sebuah model baru yang dapat menjelaskan struktur kovariani dengan melibatkan jumlah faktor yang lebih sedikit. Pendekatan yang digunakan dalam model ini adalah dengan menggunakan nilai eigen. Apabila $p - m$ nilai eigen terakhir mempunyai nilai eigen yang cukup kecil maka kontribusi dari $\lambda_{m+1} e_{m+1} e_{m+1}' + \dots + \lambda_p e_p e_p'$ terhadap Σ pada persamaan (2.19) dapat

diabaikan. Dengan mengabaikan kontribusi ini, diperoleh persamaan berikut.

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} e_1 & \sqrt{\lambda_2} e_2 & \cdots & \sqrt{\lambda_m} e_m \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} e'_1 1 \\ - - - \\ \sqrt{\lambda_2} e'_2 \\ - - - \\ \vdots \\ - - - \\ \sqrt{\lambda_m} e'_m \end{bmatrix} = L_{p \times m} L'_{m \times p} \quad (2.20)$$

Pada pendekatan di atas diasumsikan bahwa faktor spesifik ϵ pada (3.4) keberadaannya tidak terlalu penting dan dapat diabaikan pada pemfaktoran Σ . Akan tetapi, jika faktor spesifik tetap dilibatkan dalam model dapat dilakukan pendekatan sebagai berikut.

$$\Sigma = LL' + \Psi \quad (2.21)$$

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} e_1 & \sqrt{\lambda_2} e_2 & \cdots & \sqrt{\lambda_m} e_m \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} e'_1 1 \\ - - - \\ \sqrt{\lambda_2} e'_2 \\ - - - \\ \vdots \\ - - - \\ \sqrt{\lambda_m} e'_m \end{bmatrix} + \begin{bmatrix} \varphi_1 & 0 & \cdots & 0 \\ 0 & \varphi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \quad (2.22)$$

dengan $\varphi_i = \sigma_{ii} - \sum_{j=1}^m \ell_{ij}^2$ untuk $i = 1, 2, \dots, p$

Tujuan analisis faktor adalah menemukan struktur yang lebih sederhana maka yang diperlukan adalah $m < p$.

Matriks faktor *loading* terestimasi adalah $\{\tilde{\ell}_{ij}\}$.

$$\tilde{L} = \begin{bmatrix} \sqrt{\hat{\lambda}_1} \hat{e}_1 & \sqrt{\hat{\lambda}_2} \hat{e}_2 & \cdots & \sqrt{\hat{\lambda}_m} \hat{e}_m \end{bmatrix} \quad (2.23)$$

Estimasi untuk variansi spesifik diberikan oleh elemen diagonal dari matriks $S - \tilde{L}\tilde{L}'$ sehingga:

$$\tilde{\Psi} = \begin{bmatrix} \tilde{\varphi}_1 & 0 & \cdots & 0 \\ 0 & \tilde{\varphi}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\varphi}_p \end{bmatrix} \text{ dengan } \tilde{\varphi}_i = S_{ii} - \sum_{j=1}^m \tilde{\ell}_{ij}^2 \quad (2.24)$$

Sedangkan nilai estimasi komunalitasnya adalah

$$\tilde{h}_i^2 = \tilde{\ell}_{i1}^2 + \tilde{\ell}_{i2}^2 + \cdots + \tilde{\ell}_{im}^2 \quad (2.25)$$

3. Persyaratan dalam Analisis faktor

Untuk melakukan analisis faktor, harus ada normalitas univariat dan multivariat dalam data (Child, 2006). Penting juga bahwa tidak adanya *outlier* univariat dan multivariat (Field, 2009). Juga, faktor penentu didasarkan pada asumsi bahwa ada hubungan linier antara faktor dan variabel ketika menghitung korelasi (Gorsuch, 1983). Untuk sesuatu yang akan diberi label sebagai faktor harus memiliki setidaknya 3 variabel, meskipun hal ini tergantung pada desain penelitian (Tabachnick dan Fidell, 2007). Sebagai panduan umum, faktor-faktor yang dirotasi yang memiliki 2 variabel atau lebih sedikit harus ditafsirkan dengan hati-hati. Suatu faktor dengan 2 variabel hanya dianggap reliabel bila variabel-variabel tersebut berkorelasi tinggi satu sama lain ($r > 0,70$), tetapi cukup tidak berkorelasi dengan variabel lain.

Ukuran sampel yang direkomendasikan setidaknya 300 peserta dan variabel yang menjadi sasaran analisis faktor masing-masing harus memiliki setidaknya 5 sampai 10 pengamatan (Comrey dan Lee, 1992). Kami biasanya mengatakan bahwa rasio responden terhadap variabel harus setidaknya 10:1

dan bahwa faktor-faktor tersebut dianggap stabil dan divalidasi silang dengan rasio 30:1. Ukuran sampel yang lebih besar akan mengurangi kesalahan dalam data Anda dan EFA umumnya bekerja lebih baik dengan ukuran sampel yang lebih besar. Namun, Guadagnoli dan Velicer (1988) mengusulkan bahwa jika kumpulan data memiliki beberapa skor pemuatan faktor yang tinggi ($> 0,80$), maka ukuran kecil yang lebih kecil ($n > 150$) sudah cukup. Pemuatan faktor untuk suatu variabel adalah ukuran seberapa besar kontribusi variabel terhadap faktor tersebut. Dengan demikian, skor pemuatan faktor yang tinggi menunjukkan bahwa dimensi faktor lebih baik diperhitungkan oleh variabel.

a. Ukuran sampel

Meskipun ukuran sampel penting dalam analisis faktor, ada berbagai pendapat dan beberapa pedoman praktis dikutip dalam literatur. Kurangnya kesepakatan dicatat oleh Hogarty et al., yang menyatakan bahwa "*Rekomendasi (ukuran sampel) yang berbeda ini tidak membantu peneliti dengan baik*". Panduan umum termasuk, aturan praktis Tabachnick yang menyarankan memiliki setidaknya 300 kasus diperlukan untuk analisis faktor. Hair et al. menyarankan bahwa ukuran sampel harus 100 atau lebih besar. Sejumlah buku teks mengutip karya Comrey dan Lee dalam panduan mereka untuk ukuran sampel: 100 sebagai buruk, 200 sebagai cukup, 300 sebagai baik, 500 sebagai sangat baik, dan 1000 atau lebih sebagai sangat baik.

b. Rasio Sampel terhadap Variabel (Rasio N:p)

Seperangkat rekomendasi lain juga ada yang memberikan panduan kepada peneliti mengenai berapa banyak peserta yang diperlukan untuk setiap variabel sering disebut rasio sampel terhadap variabel, sering dilambangkan sebagai rasio N:p dengan N mengacu pada jumlah peserta dan

p mengacu pada jumlah dari variabel. Rekomendasi berbeda yang sama juga terjadi untuk rasio sampel terhadap variabel seperti yang mereka lakukan untuk menentukan ukuran sampel yang memadai. Misalnya, aturan praktis berkisar dari 3:1, 6:1, 10:1, 15:1, atau 20:1. Untuk menyoroti ambiguitas ini, peneliti seperti Hogarty et al. dan MacCallum et al. telah melakukan penelitian untuk menguji panduan ini. Hogarty et al. mencatat bahwa, "*Hasil kami menunjukkan bahwa tidak ada tingkat minimum rasio N atau N:p untuk mencapai pemulihan faktor yang baik di seluruh kondisi yang diperiksa*".

c. Asumsi-Asumsi Analisis Faktor

Dalam analisis faktor, tidak ada variabel dependen dan independen. Proses analisis faktor sendiri mencoba menemukan hubungan antarsejumlah variabel-variabel yang saling dependen dengan yang lain sehingga bisa dibuat satu atau beberapa kumpulan variabel yang lebih sedikit dari jumlah awal. Khusus untuk analisis faktor, sejumlah asumsi berikut harus dipenuhi (Santoso, 2006: 13).

- 1) Besarnya korelasi antarvariabel independen harus cukup kuat, misalnya di atas 0,5.
- 2) Besarnya korelasi parsial, korelasi antar dua variabel dengan menganggap tetap variabel lain, justru harus kecil. Pada SPSS deteksi terhadap korelasi parsial diberikan melalui pilihan *Anti Image Correlation*.
- 3) Pengujian seluruh matriks korelasi antarvariabel yang diukur dengan besaran *bartlett test of sphericity* atau *measure sampling adequacy* (MSA). Pengujian ini mengharuskan adanya korelasi yang signifikan diantara paling sedikit beberapa variabel.
- 4) Pada beberapa kasus, uji asumsi klasik antar faktor sebaiknya dipenuhi

d. Tahapan Analisis Faktor

Proses dasar analisis faktor meliputi hal-hal berikut.

1) Menentukan variabel apa saja yang akan dianalisis.

Variabel-variabel yang dipilih adalah variabel yang relevan dengan penelitian yang dilakukan dan harus didasarkan pada penelitian-penelitian terdahulu, teori dan pendapat peneliti sendiri.

2) Menghitung matriks korelasi dengan metode *bartlett test of sphericity* serta pengukuran MSA (*measure of sampling*).

Uji bartlett bertujuan untuk mengetahui hubungan antarvariabel dalam kasus multivariat. Jika variabel X_1, X_2, \dots, X_p *independent* (bersifat saling bebas) maka matriks korelasi antarvariabel sama dengan matriks identitas. Oleh karena itu, untuk menguji kebebasan antar variabel ini, uji *bartlett* menyatakan hipotesis sebagai berikut.

$$H_0 : \rho = \mathbf{I}$$

$$H_1 : \rho \neq \mathbf{I}$$

Statistik uji

$$\bar{r}_k = \frac{1}{p-1} \sum_{i=1}^p r_{ik}, \quad k = 1, 2, \dots, p$$

$$\bar{r} = \frac{2}{p(p-1)} \sum_{i < k} r_{ik}$$

$$\hat{\gamma} = \frac{(p-1)^2 [1 - (1 - \bar{r})^2]}{p - (p-2)(1 - \bar{r})^2} \quad (2.26)$$

Keterangan:

\bar{r}_k = rata-rata elemen diagonal pada kolom atau baris ke k dari matrik

R (matrik korelasi) \bar{r} = rata-rata keseluruhan dari elemen diagonal

Daerah penolakan

Tolak H_0 jika:

$$T = \frac{(n-1)}{(1-\bar{r})^2} \left[\sum_{i < k} \sum (r_{ik} - \bar{r})^2 - \hat{\gamma} \sum_{k=1}^p (\bar{r}_k - \bar{r})^2 \right] > \chi^2_{(p+1)(p-2)/2; \alpha} \quad (2.27)$$

Maka variabel-variabel saling berkorelasi hal ini berarti terdapat hubungan antar variabel. Jika H_0 ditolak maka analisis multivariat layak untuk digunakan terutama metode analisis komponen utama dan analisis faktor. Uji KMO bertujuan untuk mengetahui apakah semua data yang telah terambil telah cukup untuk difaktorkan. Hipotesis dari KMO adalah sebagai berikut.

Hipotesis:

H_0 : Jumlah data cukup untuk difaktorkan

H_1 : Jumlah data tidak cukup untuk difaktorkan

Rumus uji KMO:
$$KMO = \frac{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2}{\sum_{i=1}^p \sum_{j=1}^p r_{ij}^2 + \sum_{i=1}^p \sum_{j=1}^p a_{ij}^2} \quad (2.28)$$

$$i = 1, 2, \dots, p; j = 1, 2, \dots, p$$

Keterangan:

r_{ij} = koefisien korelasi sederhana antara variabel i dan j

a_{ij} = koefisien korelasi parsial antara variabel i dan j

Kriteria uji KMO dari matriks antara variabel:

Untuk $0,9 < \text{KMO} < 1,00$ data sangat baik,
Untuk $0,8 < \text{KMO} < 0,9$ data baik
Untuk $0,7 < \text{KMO} < 0,8$ data baik .
Untuk $0,6 < \text{KMO} < 0,7$ data lebih dari cukup
Untuk $0,5 < \text{KMO} < 0,6$ data cukup
Untuk $\text{KMO} < 0,5$ data tidak layak.

Apabila nilai KMO lebih besar dari 0,5 maka terima H_0 sehingga dapat disimpulkan jumlah data telah cukup difaktorkan.

3) Ekstraksi atau proses *factoring*.

Ekstraksi faktor adalah suatu metode yang digunakan untuk mereduksi data dari beberapa indikator untuk menghasilkan faktor yang lebih sedikit yang mampu menjelaskan korelasi antarindikator yang diobservasi. Metode yang digunakan untuk melakukan ekstraksi faktor adalah *principal component analysis*, *principal axis factoring*, *unweighted leastsquares*, *generalized least squares*, dan *maximum likelihood*. Metode ekstraksi faktor yang digunakan dalam penelitian ini adalah metode *principal component analysis* (PCA). *Principal component analysis* merupakan suatu teknik analisis untuk transformasi variabel-variabel asli yang masih saling berkorelasi satu dengan yang lain menjadi suatu variabel baru yang tidak berkorelasi. Hal ini dilakukan dengan cara menghilangkan korelasi di antara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali atau yang biasa disebut dengan *principal component*. Pendekatan *principal component analysis* jika diekstraksi dari matriks korelasi diperoleh faktor dengan beberapa kriteria sebagai berikut.

- a) *Communalities* adalah besarnya varian variabel yang disaring dengan variabel lainnya.
- b) Nilai *eigen* dengan persamaan karakteristiknya $|M_{vv} - \lambda| = 0$ dengan

>1.

4) Menentukan jumlah faktor.

Penentuan jumlah faktor didasarkan pada besarnya eigen *value* setiap faktor yang muncul. Eigen *value* merupakan jumlah varian yang dijelaskan oleh setiap faktor. Maksudnya, melakukan analisis faktor untuk mencari variabel baru yang disebut faktor yang saling berkorelasi, bebas, atau satu sama lainnya. Lebih sedikit jumlahnya daripada variabel asli, tetapi bisa menyerap sebagian informasi yang terkandung dalam variabel asli. Faktor-faktor inti yang dipilih adalah faktor yang memiliki eigen *value* > 1 .

5) Rotasi faktor-faktor

Rotasi faktor dilakukan untuk mempermudah interpretasi dalam menentukan variabel-variabel mana saja yang tercantum dalam suatu faktor karena terkadang ada beberapa variabel yang mempunyai korelasi tinggi dengan lebih dari satu faktor atau jika sebagian faktor *loading* dari variabel bernilai di bawah terkecil yang telah ditetapkan. Suatu hasil atau *output* yang penting dalam analisis faktor ialah apa yang disebut matriks faktor pola (*factor pattern matrix*). Matriks faktor memuat/berisi koefisien yang dipergunakan untuk mengekspresikan variabel yang dibakukan (*standardized*) dinyatakan dalam faktor.

Hasil dari ekstraksi faktor tidak mempunyai arti jika tidak dirotasi karena rotasi ekstraksi berguna untuk memungkinkan penafsiran dan pengulasan ilmiah. Analisis faktor mempunyai dua rotasi yaitu rotasi ortogonal dan rotasi *oblique*. Bagian dari rotasi ortogonal adalah *varimax*, *quartimax*, dan *equamax*. Dalam penelitian ini digunakan rotasi *varimax*. Rotasi *varimax* adalah salah satu jenis rotasi ortogonal yang pilihannya ada pada SPSS. Tujuan utama metode ini adalah untuk mendapatkan struktur faktor sehingga setiap variabel termuat tinggi hanya pada suatu faktor. Jadi,

setiap variabel harus mempunyai muatan tinggi pada satu faktor dan nol pada faktor lain. Struktur faktor seperti ini mengindikasikan bahwa setiap faktor menyatakan sebuah konstruk yang berbeda. Jadi, rotasi *varimax* menghilangkan faktor umum.

6) Menentukan skor faktor

Skor faktor adalah nilai-nilai untuk faktor acak yang tidak teramati. $F_j, j = 1, 2, \dots, n$. Jadi, skor faktor f_j (kasus ke- j). Pada persamaan (2.25), dari faktor yang terbentuk dapat memberikan penjelasan bahwa antara variabel di dalam faktor tertentu mempunyai hubungan yang sangat kuat. Namun terhadap variabel dalam faktor lain mempunyai hubungan yang relatif kecil. Sebagai tahap yang terakhir untuk memperoleh struktur yang berarti. Skor faktor adalah ukuran individual pada faktor yang merupakan nilai rata-rata terbobot. Komponen hipotesis yang diturunkan harus memiliki sifat-sifat sebagai berikut:

- a) Komponen hipotesis tersebut diberi nama faktor. Faktor-faktor ini membentuk *linearly independent set variabel*. Tak ada faktor yang menjadi kombinasi linear dari faktor lain, sebab faktor-faktor tersebut dibuat sedemikian rupa sehingga bebas (independen) satu sama lain.
- b) Variabel komponen hipotesis bisa dikelompokkan menjadi dua, yaitu: *common factors* dan *unique factors*. Dua komponen ini bisa dibedakan kalau dinyatakan dalam timbangan (*weight*). Di dalam persamaan linear, yang menurunkan variabel terobservasi dari variabel komponen hipotesis. Suatu *common factors* mempunyai lebih dari suatu variabel dengan timbangan yang bukan nol atau faktor *loading* yang terkait dengan faktor. Jadi hanya satu variabel yang tergantung dalam suatu faktor unik.
- c) *Common factor* selalu dianggap tidak berkorelasi dengan faktor unik.

Faktor unik biasanya juga dianggap tidak berkorelasi (*mutually uncorrelated*), tetapi *common factor* mungkin atau tidak mungkin berkorelasi satu sama lain. Umumnya dianggap bahwa jumlah *common factor* lebih sedikit dari jumlah variabel asli. Akan tetapi, banyaknya faktor unik biasanya dianggap sama dengan banyaknya variabel asli.

C. Implementasi Analisis Faktor

1. Contoh Analisis Faktor dengan Microsoft Excel

Dalam menghitung analisis faktor dengan Microsoft Excel dapat dilakukan dengan tahapan sebagai berikut.

a. Membuat Tabel Pembantu

Tabel 7.1 Tabel Pembantu

No. Res p	X1	X2	X3	X4	X5	X1 ²	X2 ²	X3 ²	X4 ²	X5 ²	X1.X2	X1.X3	X1.X4	X1.X5
1	3	2	4	3	3	9	4	16	9	9	6	12	9	9
2	4	3	1	3	3	16	9	1	9	9	12	4	12	12
3	3	4	3	3	3	9	16	9	9	9	12	9	9	9
4	3	4	4	4	3	9	16	16	16	9	12	12	12	9
5	4	4	4	5	5	16	16	16	25	25	16	16	20	20
6	5	5	3	3	3	25	25	9	9	9	25	15	15	15
7	4	4	4	4	5	16	16	16	16	25	16	16	16	20
8	4	4	5	4	4	16	16	25	16	16	16	20	16	16
9	3	3	4	3	3	9	9	16	9	9	9	12	9	9
10	4	5	3	4	4	16	25	9	16	16	20	12	16	16
11	4	5	4	5	5	16	25	16	25	25	20	16	20	20
12	4	5	4	3	4	16	25	16	9	16	20	16	12	16
13	5	4	4	5	4	25	16	16	25	16	20	20	25	20
14	3	3	3	3	3	9	9	9	9	9	9	9	9	9
15	5	4	3	4	4	25	16	9	16	16	20	15	20	20
Jml.	58	59	53	56	56	232	243	199	218	218	233	204	220	220

Cara menghitung tabel korelasi dengan bantuan Microsoft Excel dengan rumus =correl(D3:D17, C3:CC17) hasilnya adalah 0,529 seperti yang dapat dilihat pada gambar di bawah ini.

D24 : ✕ ✓ *fx* =CORREL(D3:D17,C3:C17)

	A	B	C	D	E	F	G
22		Matrik Korelasi					
23			X1	X2	X3	X4	X5
24		X1	1.000	0.529	-0.098	0.417	0.417
25		X2		1.000	0.135	0.378	0.479
26		X3			1.000	0.404	0.404
27		X4				1.000	0.776
28		X5					1.000
29							

Kemudian menghitung manual korelasi dengan cara mengambil data dalam tabel pembantu seperti di bawah ini.

Tabel 7.2 Bantu.

N=15	$\Sigma X_1^2 = 232$
$\Sigma X_1 = 58$	$(\Sigma X_1)^2 = 3,364$
$\Sigma X_2 = 59$	$\Sigma X_2^2 = 243$
$\Sigma X_1 X_2 = 233$	$(\Sigma X_2)^2 = 3,481$

$$r_{x_1 x_2} = \frac{N \Sigma X_1 X_2 - (\Sigma X_1)(\Sigma X_2)}{\sqrt{[N \Sigma_1^2 - (\Sigma X_1)^2][N \Sigma X_2^2 - (\Sigma X_2)^2]}}$$

$$rx_1x_2 = \frac{15(233) - (58)(59)}{\sqrt{[15(233) - (3,364)][15(243) - (3,481)]}}$$

$$rx_1x_2 = \frac{73}{\sqrt{(3,480) - (3,364)(3,645) - (3,481)}}$$

$$rx_1x_2 = \frac{73}{\sqrt{(116) - (164)}}$$

$$rx_1x_2 = \frac{73}{\sqrt{(137,9275)}}$$

$$rx_1x_2 = 0,529$$

Setelah menghitung, manual berikutnya di validasi dengan hitungan SPSS versi 21 hasilnya seperti yang terlihat pada gambar di bawah ini.

		Correlations				
		X1	X2	X3	X4	X5
X1	Pearson Correlation	1	.529	-.098	.417	.417
	Sig. (2-tailed)		.042	.728	.122	.122
	N	15	15	15	15	15
X2	Pearson Correlation	.529	1	.135	.378	.479
	Sig. (2-tailed)	.042		.630	.165	.071
	N	15	15	15	15	15
X3	Pearson Correlation	-.098	.135	1	.404	.404
	Sig. (2-tailed)	.728	.630		.136	.136
	N	15	15	15	15	15
X4	Pearson Correlation	.417	.378	.404	1	.776
	Sig. (2-tailed)	.122	.165	.136		.001
	N	15	15	15	15	15
X5	Pearson Correlation	.417	.479	.404	.776	1
	Sig. (2-tailed)	.122	.071	.136	.001	
	N	15	15	15	15	15

Setelah dilakukan validasi antara hitungan menggunakan Microsoft Excel dan menggunakan hitungan manual dan SPSS, maka hasilnya sama.

b. Membuat Uji Kolmogorov Smirnov

Langkah-langkah pengujian adalah sebagai berikut:

Tabel 7.3 Responden.

=COUNT(C4:C18)

	I	J	K	L	M	N	O	P	Q	R	S
		statistik	test KS		data	frekuensi	frek kumulatif	S(X)	Z-Score	F(X)	Difference
		n	15		3	5	5	0.333333	-1.166091812	0.12178866	0.2115447
		\bar{x}	3.867		4	7	12	0.8	0.17939874	0.57118769	0.2288123
		s	0.743		5	3	15	1	1.524889293	0.93635664	0.0636434
		D	0.229								
		KS Tabel	0.338								

Cara menghitung dengan rumus:

$$n = \text{COUNT}(C4:C18),$$

$$\bar{x} = \text{AVERAGE}(C4:C18)$$

$$s = \text{STDEV}(C4:C18)$$

$$D = \text{MAX}(S17:S19)$$

KS Tabel=0,338 diambil dari tabel nilai kritis uji *kolmogorov smirnov* dengan cara melihat nilai n=15 dengan tingkat signifikansi yang digunakan adalah 5% $\alpha = 0,05$.

	A	B	C	D	E	F	G
3		No. Resp	X1	X2	X3	X4	X5
4		1	3	2	4	3	3
5		2	4	3	1	3	3
6		3	3	4	3	3	3
7		4	3	4	4	4	3
8		5	4	4	4	5	5
9		6	5	5	3	3	3
10		7	4	4	4	4	5
11		8	4	4	5	4	4
12		9	3	3	4	3	3
13		10	4	5	3	4	4
14		11	4	5	4	5	5
15		12	4	5	4	3	4
16		13	5	4	4	5	4
17		14	3	3	3	3	3
18		15	5	4	3	4	4

Tabel Nilai Kritis Uji Kolmogorov-Smirnov

n	$\alpha = 0,20$	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,02$	$\alpha = 0,01$
1	0,900	0,950	0,975	0,990	0,995
2	0,684	0,776	0,842	0,900	0,929
3	0,565	0,636	0,708	0,785	0,829
4	0,493	0,565	0,624	0,689	0,734
5	0,447	0,509	0,563	0,627	0,669
6	0,410	0,468	0,519	0,577	0,617
7	0,381	0,436	0,483	0,538	0,576
8	0,359	0,410	0,454	0,507	0,542
9	0,339	0,387	0,430	0,480	0,513
10	0,323	0,369	0,409	0,457	0,486
11	0,308	0,352	0,391	0,437	0,468
12	0,296	0,338	0,375	0,419	0,449
13	0,285	0,325	0,361	0,404	0,432
14	0,275	0,314	0,349	0,390	0,418
15	0,266	0,304	0,338	0,377	0,404
16	0,258	0,295	0,327	0,366	0,392
17	0,250	0,286	0,318	0,355	0,381
18	0,244	0,279	0,309	0,346	0,371
19	0,237	0,271	0,301	0,337	0,361
20	0,232	0,265	0,294	0,329	0,352
21	0,226	0,259	0,287	0,321	0,344
22	0,221	0,253	0,281	0,314	0,337
23	0,216	0,247	0,275	0,307	0,330
24	0,212	0,242	0,269	0,301	0,323
25	0,208	0,238	0,264	0,295	0,317

Cara mencari data dengan melihat data responden X1 dengan mencari nilai yang mewakili di antaranya nilai 3 frekuensinya 5, nilai 4 frekuensinya 7 dan nilai 5 frekuensinya ada 3. Cara mencari nilai frekuensi kumulatif adalah dengan nilai pertama lalu ditambah 7 hasilnya 12 dan ditambah 3 hasilnya menjadi 15.

Kemudian mencari nilai:

$$S(X) = O17/K\$17$$

$$Z\text{-Score} = \text{STANDARDIZE}(M17, K\$18, K\$19)$$

$$F(X) = \text{NORMDIST}$$

$$\text{Difference} = \text{ABS}(R17-P17)$$

Dapat dilihat bahwa $D=0,229 < D\text{-Tabel} = 0,338$ yang berarti statistik uji D tidak berada daerah kritis jadi H_0 diterima sehingga diperoleh kesimpulan bahwa data berdistribusi normal.

c. Membuat Uji Kaiser Meyer Olkin Measure Of Sampling Adequacy

Tahapan dalam KMO sebagai berikut.

- a) Membuat *corrlation matrix*: SUMSQ(Correlation Coefficients) dan jumlah total (*correlation*).
- b) Membuat *anti image matrix* = $D^{1/2} \times R^{-1} \times D^{1/2}$

R = *correlation matrix*

R^{-1} = *inverse matrix*

$D^{1/2}$ = 1/Sqrt (diagonal) *anti image matrix* = MMULT(MMULT($D^{1/2}$, R^{-1}), $D^{1/2}$) SUMSQ (*anti image*)

Jumlah total *anti image*

- c) KMO = jumlah korelasi total/(korelasi total) + (total *anti image*)

Tabel 7.4 Tabel Responden.

No. Resp	X1	X2	X3	X4	X5
1.	3	2	4	3	3
2.	4	3	1	3	3
3.	3	4	3	3	3
4.	3	4	4	4	3
5.	4	4	4	5	5
6.	5	5	3	3	3
7.	4	4	4	4	5
8.	4	4	5	4	4
9.	3	3	4	3	3
10.	4	5	3	4	4
11.	4	5	4	5	5
12.	4	5	4	3	4
13.	5	4	4	5	4
14.	3	3	3	3	3
15.	5	4	3	4	4

Tabel 7.5 Matriks Korelasi.

L3 : \times f_x =CORREL(C3:C17,D3:D17)

	J	K	L	M	N	O	P	Q	R
2	X1	X2	X3	X4	X5				
3	X1	1.000	0.529	-0.098	0.417	0.417			
4	X2	0.529	1.000	0.135	0.378	0.479			
5	X3	-0.098	0.135	1.000	0.404	0.404			
6	X4	0.417	0.378	0.404	1.000	0.776			KMO
7	X5	0.417	0.479	0.404	0.776	1.000			0.674
8		1.638	1.671	1.354	2.082	2.169	3.91		

Tabel 7.6 Matriks Inverse.

L33 : \times f_x =MMULT(MMULT(K23:O27,K13:O17),K23:O27))

	J	K	L	M	N	O	P	Q
32	X1	X2	X3	X4	X5			
33	X1	1.000	-0.419	0.357	-0.240	-0.093		
34	X2	-0.419	1.000	-0.086	0.081	-0.255		
35	X3	0.357	-0.086	1.000	-0.229	-0.193		
36	X4	-0.240	0.081	-0.229	1.000	-0.640		
37	X5	-0.093	-0.255	-0.193	-0.640	1.000		
38		1.37	1.25	1.22	1.53	1.52	1.89	

Tabel 7.7 Diagonal Matriks

K13 : \times f_x =MINVERSE(K3:O7))

	J	K	L	M	N	O
12	X1	X2	X3	X4	X5	
13	X1	1.725	-0.692	0.557	-0.521	-0.209
14	X2	-0.692	1.582	-0.129	0.168	-0.548
15	X3	0.557	-0.129	1.412	-0.450	-0.392
16	X4	-0.521	0.168	-0.450	2.737	-1.805
17	X5	-0.209	-0.548	-0.392	-1.805	2.909

Tabel 7.8 Anti Image.

K23 : \times f_x =1/SQRT(K13)

	J	K	L	M	N	O
22	X1	X2	X3	X4	X5	
23	X1	0.761	0.000	0.000	0.000	0.000
24	X2	0.000	0.795	0.000	0.000	0.000
25	X3	0.000	0.000	0.842	0.000	0.000
26	X4	0.000	0.000	0.000	0.604	0.000
27	X5	0.000	0.000	0.000	0.000	0.586

Tabel 7.9 Uji Barlett.

	T	U	V	W	X	Y	Z	AA	AB	AC	
1											
2	Bartlett Test Statistics						Calculation	SPSS Output			
3	Chi Sqr = Minus (n-1)((-2p+5)/6)xLN R						n=	15			
4	df=[px(p-1)/2						p=	5			
5	n=Respondent						Determinant	0.144			
6	p=Variabel						2p+5/6=	2.5			
7							Approx. Chi-Square	22.271	22.271		
8							df	10	10		Compare
9							Sig	0.0047	0.014		0.0093
10											
11							KMO	0.674	0.674		

Keterangan:

n = jumlah responden

p = jumlah variabel

Determinant : 0,144 = MDETERM(K3:O7)

2p+5/6 : 2,55 = -(Z3-1-Z6)*LN(Z5)

Approx. Chi-Square : 22,271 = -(Z3-1-Z6)*LN(Z5)

df (derajat kebebasan) : 10 = Z4*(Z4-1)/2

Sig : 0,0047 = CHISQ.DIST(Z7,Z8,)

Gambar 7.1 Uji KMO dan Bartlett.

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.674
Bartlett's Test of Sphericity	Approx. Chi-Square	22.271
	df	10
	Sig.	.014

Setelah dilakukan uji KMO dan barlett secara manual dan di bandingkan

dengan uji KMO dan barlett dengan SPSS ada perbedaan tingkat signifikansi adalah 0,0093.

d. Menghitung Manual Faktor Loading

Berikut ini tahapan menghitung manual faktor *loading* dan *eigen value* dan *comunality* menurut Dr. Balamurugan sebagai berikut:

- 1) Masukan data responden ke dalam Microsoft Excel.
- 2) Buatlah tabel 1 data untuk mencari data matriks korelasi kemudian buat notifikasi korelasi negatif.

	A	B	C	D	E	F	G	H
24		Table 1						
25		Correlation (R) Matrix						
26			X1	X2	X3	X4	X5	
27		X1	1.000	0.529	-0.098	0.417	0.417	
28		X2	0.529	1.000	0.135	0.378	0.479	
29		X3	-0.098	0.135	1.000	0.404	0.404	
30		X4	0.417	0.378	0.404	1.000	0.776	
31		X5	0.417	0.479	0.404	0.776	1.000	

- 3) Buat tabel 2 untuk mengonversi semua nilai positif menggunakan **ABS Excel Function** dengan cara menjumlahkan semua kolom dan cari jumlah total. Kemudian untuk mencari Factor Loading = Column SUM/SQRT jumlah total (=C42/SQRT((\$H\$42)).

D37 :

	A	B	C	D	E	F	G	H	I
34		Table 2							
35		Q1-Absolute Values							
36			X1	X2	X3	X4	X5		
37		X1	1.000	0.529	0.098	0.417	0.417		
38		X2	0.529	1.000	0.135	0.378	0.479		
39		X3	0.098	0.135	1.000	0.404	0.404		
40		X4	0.417	0.378	0.404	1.000	0.776		
41		X5	0.417	0.479	0.404	0.776	1.000		
42			2.46	2.52	2.04	2.98	3.08	13.07	
43			0.681	0.697	0.564	0.823	0.851		factor 01

- 4) Buat tabel 3 rechecking iterasi faktor pertama dan kedua dengan cara menyilang (*cross row*) lalu jumlahkan semua data dalam kolom dan cari jumlah total atau faktor *loading* = Column SUM/SQRT jumlah total.

C49 : =C48*\$B\$49

	A	B	C	D	E	F	G	H	
46									
47			Table 3						
48			Rechecking first factor by 2nd iteration by cross row col product (R1)						
48			0.68	0.70	0.56	0.82	0.85		
49		0.68	0.46	0.48	0.38	0.56	0.58		
50		0.70	0.48	0.49	0.39	0.57	0.60		
51		0.56	0.38	0.39	0.31	0.46	0.48		
52		0.82	0.56	0.57	0.46	0.67	0.70		
53		0.85	0.58	0.60	0.48	0.70	0.72		
54			2.45	2.53	2.02	2.96	3.07	13.03	
55			0.68	0.70	0.56	0.82	0.85		

- 5) Buat tabel 4 untuk mendapatkan matriks residu yang pertama dengan membagi tabel ke-2 dan ke-3.

C61 : =C49-C37

	A	B	C	D	E	F	G	
58								
59			Table 4					
59			Residual (R2) score					
60			X1	X2	X3	X4	X5	
61		X1	-0.538	-0.053	0.283	0.141	0.161	
62		X2	-0.053	-0.510	0.257	0.196	0.116	
63		X3	0.283	0.257	-0.686	0.055	0.072	
64		X4	0.141	0.196	0.055	-0.328	-0.079	
65		X5	0.161	0.116	0.072	-0.079	-0.278	

- 6) Buat tabel 5 untuk mengonversi semua nilai positif menggunakan **ABS Excel Function** dan ulangi dengan proses yang sama.

C73 : =ABS(-0.538)

	A	B	C	D	E	F	G	H	I	
70										
71				Table 5						
72				Q2-Absolute Values						
73				X1	X2	X3	X4	X5		
74				X1	0.538	0.053	0.283	0.141	0.161	
75				X2	0.053	0.510	0.257	0.196	0.116	
76				X3	0.283	0.257	0.686	0.055	0.072	
77				X4	0.141	0.196	0.055	0.328	0.079	
78				X5	0.161	0.116	0.072	0.079	0.278	
79					1.176	1.132	1.353	0.799	0.706	
					0.52	0.50	0.60	0.35	0.31	
									5.166	
									Factor 2	

7) Tabel 6 *rechecking* iterasi faktor kedua.

C85 : =C84*B85

	A	B	C	D	E	F	G	H	
82									
83				Table 6					
84				Rechecking first factor by 2nd iteration by cross row col product					
85				0.51	0.50	0.60	0.35	0.31	
86				0.51	0.260	0.255	0.306	0.179	
87				0.50	0.255	0.250	0.300	0.175	
88				0.60	0.306	0.300	0.360	0.210	
89				0.35	0.179	0.175	0.210	0.123	
90				0.31	0.158	0.155	0.186	0.109	
91					1.2	1.1	1.4	0.8	
					0.51	0.5	0.6	0.35	
								5.15	

8) Perbandingan hasil hitung manual dengan Microsoft Excel dengan SPSS versi 21 dengan cara hasil hitung manual dikurangi hasil hitung SPSS (0,036=K26-Q26).

M26 : =K26-Q26

	K	L	M	N	O	P	Q	R	S	T	
22											
23				Hasil Hitung SPSS Factor Loading							
24				Hasil Hitung manual Factor Loading				Component Matrix(a)			
25				Compare				Component			
26				factor 01	factor 02	F1	F2		1	2	
27				0.681	0.517	0.036	1.14	reliability	0.645	-0.623	
28				0.697	0.498	-0.006	0.868	responsiv	0.703	-0.37	
29				0.564	0.595	0.114	-0.186	assurance	0.45	0.781	
30				0.823	0.352	-0.035	0.155	empathy	0.858	0.197	
31				0.851	0.311	-0.037	0.152	tangibles	0.888	0.159	
32						0.014	0.426	Extraction Method: Principal Component Analysis.			
								a. 2 components extracted.			

e. **Menghitung Manual Eigen Value dan Commuality**

Eigen value = SUMQ setiap factor loading

- 1) Menghitung total variance extracted
 - a) *percentage of variance extracted* = eigen total/jumlah variabel,
 - b) *cumulative variance* = up to current column variance,
 - c) *communality* (h^2)= SUMQ setiap variabel,

E35 : X ✓ f_x =E34/5

	D	E	F	G
26	Component Matrix(a)			
27	Component			
28	Variable	1	2	
29	X1	0.645	-0.623	
30	X2	0.703	-0.37	
31	X3	0.45	0.781	
32	X4	0.858	0.197	
33	X5	0.888	0.159	
34	Eigen of value	2.63744	1.19908	> 1
35	Percentage of variance	53%	24%	
36	Cum. Perc of Variance	53%	77%	>60%

- 2) Penggunaan rumus Microsoft Excel untuk menghitung komponen matriks sebagai berikut.

- a) *Eigen value* = 2,63744=SUMSQ(E29:E33)
- b) *Percentage of variance extracted* = 53% = +E34/5
- c) *Cumulative variance* = 77% = +F35+E36

Component	Total Variance Explained									
	Initial Eigenvalues				Extraction Sums of Squared Loadings					
	Total	% of Varia	Cumulativ	Total	% of Varianc		Cumulativ	Total	% of Varia	
	1	2.638	52.75	52.751	2.638	52.751	52.751	2.005	40.09	40.09
	2	1.198	23.97	76.72	1.198	23.966	76.717	1.831	36.627	76.717
dimension0	3	0.58	11.594	88.311						
	4	0.374	7.481	95.792						
	5	0.21	4.208	100						

Extraction Method: Principal Component Analysis.

- 3) Dari hasil perbandingan hitungan manual persentase varian dan

kumulatif dengan hitungan SPSS maka hasilnya sama.

Q41 : =SUMSQ(L41:P41)

	K	L	M	N	O	P	Q	R	S	T	U	
38	Component Matrix a											
39		Component								Communalities		
40		1	2	3	4	5	communality		Initial	Extraction		
41	reliability	.645	-.623				0.804		1.000	.804		
42	responsiveness	.703	-.370				0.631	>0.5	1.000	.631		
43	assurance	.450	.781				0.812		1.000	.812		
44	empathy	.858	.197				0.775		1.000	.775		
45	tangibles	.888	.159				0.813		1.000	.813		

Untuk menghitung *communality* dengan cara menjumlahkan lima komponen variabel menggunakan rumus =SUMSQ(L41:P41).

2. Contoh Analisis Faktor dengan SPSS

Seorang peneliti akan melakukan penelitian tentang *Analisis Faktor-Faktor yang Memengaruhi Kepuasan Pelanggan terhadap Pelayanan Mahasiswa Politeknik XYZ*. Kepuasan pelanggan akan menjadi indikator apakah pelayanan yang diberikan berbagai unit tersebut berkualitas atau tidak. Dampak dari kepuasan pelanggan pada akhirnya akan menciptakan loyalitas pelanggan. Para pelanggan tidak menilai kualitas jasa dari satu dimensi saja, tetapi dari berbagai dimensi (Parasuraman, A., Valerie Zeithaml, & L.Berry, 1988). Menurut Parasuraman dkk, ada lima dimensi kualitas jasa, yaitu

- Reliability*: Kemampuan penyedia jasa secara handal dan akurat dalam menyediakan jasa sesuai dengan yang dijanjikan atau dipromosikan.
- Responsiveness*: Kemauan untuk membantu pelanggan dan menyediakan jasa dengan segera.
- Assurance*: Pengetahuan penyedia jasa dan kemampuannya untuk membuat pelanggan percaya dan yakin bahwa akan menerima jasa yang

berkualitas.

- d) *Empathy*: Penuh perhatian dan perlakuan secara individual yang diberikan kepada pelanggan.
- e) *Tangibles*: Penampilan fasilitas fisik, peralatan, serta penampilan penyedia jasa dan material-material komunikasi di perusahaan jasa tersebut.

Berdasarkan teori Parasuraman dkk. pertanyaan kuisioner ada 5 item tentang Q1 (*reliability*), Q2 (*responsiviness*), Q3 (*assurance*), Q4 (*emphaty*), dan Q5 (*tangibles*) dengan 15 responden. Pertanyaan tersebut adalah jenis pertanyaan positif. Setiap pertanyaan memiliki lima pilihan jawaban dengan kriteria skor, sebagai berikut ini.

- a) Sangat Tidak Puas (skor 1)
- b) Kurang Puas (skor 2)
- c) Cukup Puas (skor 3)
- d) Puas (skor 4)
- e) Sangat Puas (skor 5)

Adapun data jawaban ke-15 orang responden tersebut adalah sebagai berikut.

Tabel 7.10 Kuisioner Responden.

No. Resp	Q1 (Reliability)	Q2 (Responsiviness)	Q3 (Assurance)	Q4 (Emphaty)	Q5 (Tangibles)
1.	3	2	4	3	3
2.	4	3	1	3	3
3.	3	4	3	3	3
4.	3	4	4	4	3
5.	4	4	4	5	5
6.	5	5	3	3	3
7.	4	4	4	4	5

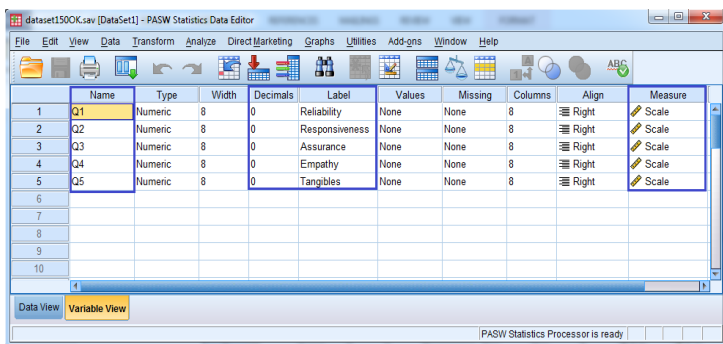
8.	4	4	5	4	4
9.	3	3	4	3	3
10.	4	5	3	4	4
11.	4	5	4	5	5
12.	4	5	4	3	4
13.	5	4	4	5	4
14.	3	3	3	3	3
15.	5	4	3	4	4

Dari data di atas, penulis akan melakukan uji normalitas terhadap kelima variabel faktor yang diteliti, adapun ke-5 variabel faktor tersebut adalah:

- a) *Reliability* (keandalan)
- b) *Responsiveness* (ketanggapan)
- c) *Assurance* (keyakinan)
- d) *Emphy* (empati)
- e) *Tangibles* (berwujud)

a. Uji Normalitas dalam Analisis Faktor dengan SPSS Versi 18

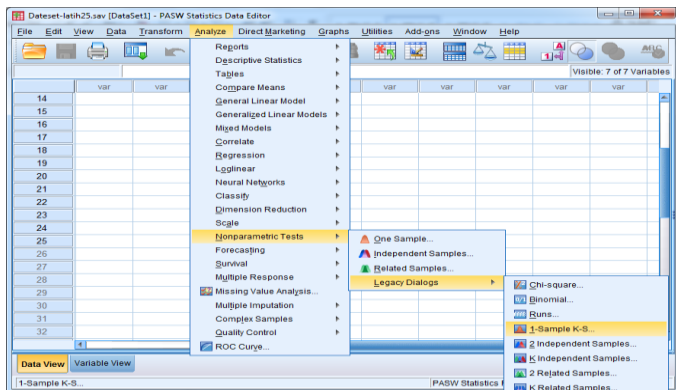
- 1) Aktifkan lembar kerja baru SPSS, kemudian klik **Variable View** untuk mengisi **Name**, **Decimals**, **Label**, dan **Measure**, dengan ketentuan sebagaimana gambar berikut.



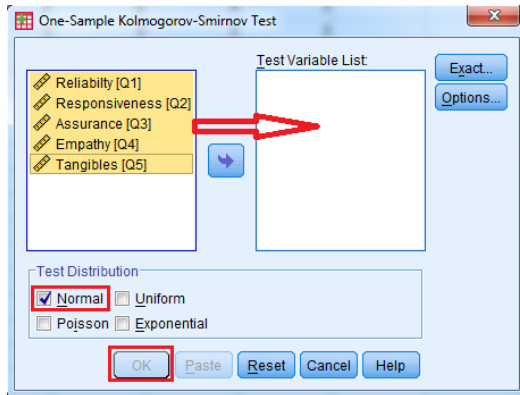
- 2) Langkah selanjutnya klik **Data View**, lalu masukkan data *reliability* ke kolom Q1, data *responsiviness* kolom Q2, data *assurance* kolom Q3, data *empathy* kolom Q4, dan data *tangibles* kolom Q5. Tampak seperti gambar berikut.

	Q1	Q2	Q3	Q4	Q5	var	var	var
1	3	2	4	3	3			
2	4	3	1	3	3			
3	3	4	3	3	3			
4	3	4	4	4	3			
5	4	4	4	5	5			
6	5	5	3	3	3			
7	4	4	4	4	5			
8	4	4	5	4	4			
9	3	3	4	3	3			
10	4	5	3	4	4			
11	4	5	4	5	5			
12	4	5	4	3	4			
13	5	4	4	5	4			
14	3	3	3	3	3			
15	5	4	3	4	4			

- 3) Setelah itu, dari menu SPSS klik **Analyze** lalu **Nonparametric Tests** lalu **Legacy Dialogs** lalu **1-Sample K-S...**



- 4) Muncul kotak dialog **One-Sample Kolmogorov-Smirnov Test** kemudian masukkan kelima variabel faktor tersebut ke kotak Variables. Pada bagian **Test Distribution** beri tanda centang (ν) pada **Normal**.



- 5) Langkah terakhir adalah klik **Ok**. Maka akan muncul *output* SPSS seperti berikut.

One-Sample Kolmogorov-Smirnov Test

		Reliability	Responsiveness	Assurance	Empathy	Tangibles
N		15	15	15	15	15
Normal Parameters ^{a,b}	Mean	3.87	3.93	3.53	3.73	3.73
	Std. Deviation	.743	.884	.915	.799	.799
Most Extreme Differences	Absolute	.238	.263	.295	.287	.287
	Positive	.229	.203	.238	.287	.287
	Negative	-.238	-.263	-.295	-.179	-.179
Kolmogorov-Smirnov Z		.921	1.020	1.142	1.113	1.113
Asymp. Sig. (2-tailed)		.364	.249	.147	.168	.168

a. Test distribution is Normal.

b. Calculated from data.

Gambar 7.2 Hasil Uji Kolmogorov Smirnov Test.

Dasar pengambilan keputusan dalam uji normalitas

- Jika nilai Asymp. Sig. (2-tailed) $> 0,05$ maka data berdistribusi normal.
- Jika nilai Asymp. Sig. (2-tailed) $< 0,05$ maka data tidak berdistribusi

normal.

- c) Keputusan uji normalitas dalam analisis faktor.

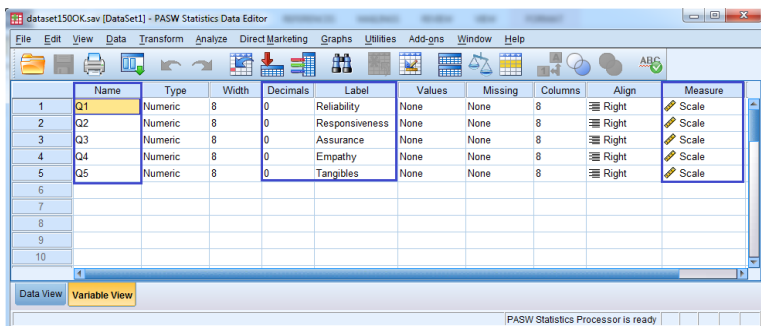
Berdasarkan *output* SPSS “One-Sample Kolmogorov-Smirnov Test” di atas diketahui nilai Asymp. Sig. (2-tailed) untuk masing-masing variabel adalah sebagai berikut.

- a) *Reliability* sebesar 0,364
- b) *Responsiveness* sebesar 0,249
- c) *Assurance* sebesar 0,147
- d) *Empathy* sebesar 0,168
- e) *Tangibles* sebesar 0,168

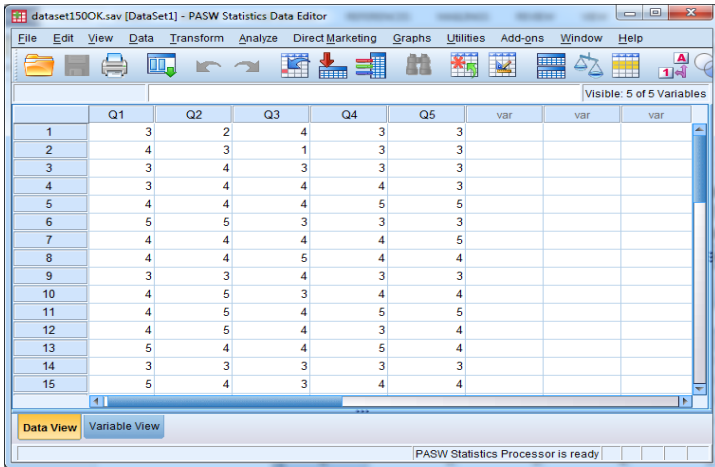
Karena nilai Asymp. Sig. (2-tailed) untuk ke-5 variabel di atas lebih besar dari 0,05 maka dapat disimpulkan bahwa data dari ke-5 variabel tersebut adalah berdistribusi normal. Dengan demikian, maka analisis faktor dalam penelitian ini dapat dilanjutkan.

b. Langkah-Langkah Analisis Faktor dengan SPSS Versi 18

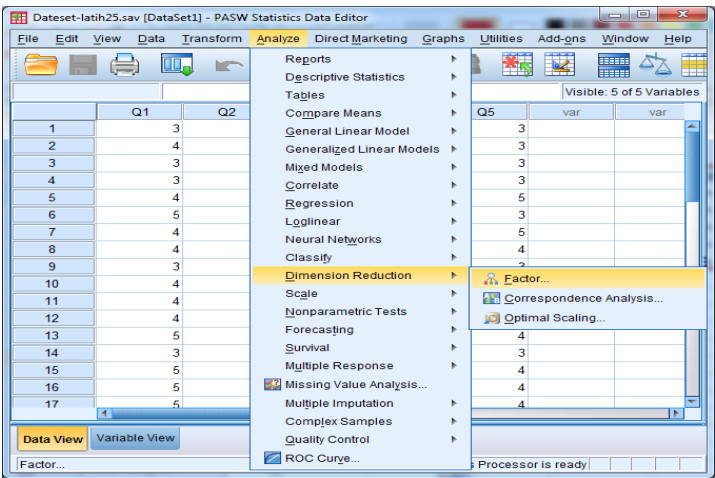
- 1) Buka lembar kerja baru SPSS, lalu klik **Variable View** untuk mengisi **Name**, **Decimals**, **Label** dan **Measure**, dengan ketentuan sebagaimana gambar di bawah ini.



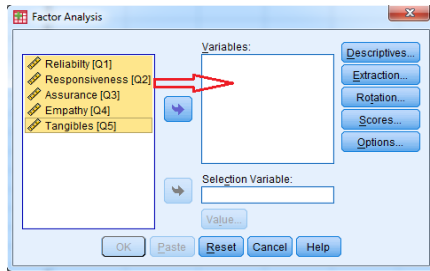
- 2) Selanjutnya klik **Data View**, kemudian masukkan data *reliability* ke kolom Q1, data *responsiviness* kolom Q2, data *assurance* kolom Q3, data *empathy* kolom Q4, dan data *tangibles* kolom Q5.



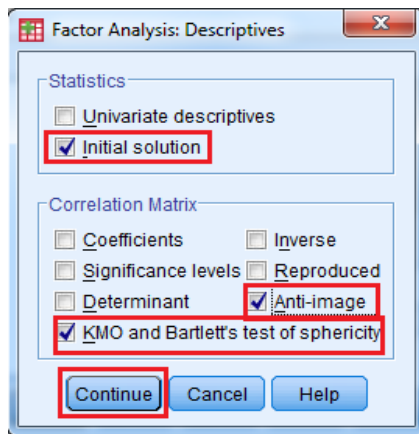
- 3) Dari menu SPSS klik Analyze lalu Dimension Reduction lalu Factor.



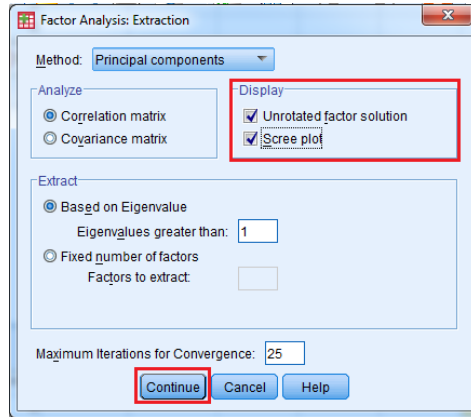
- 4) Muncul kotak dialog Factor Analysis. Selanjutnya, masukkan semua variabel ke kotak Variables, lalu klik Descriptive.



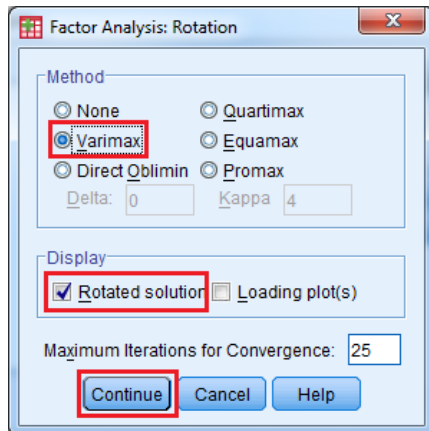
- 5) Muncul kotak dialog “*Factor Analysis: Descriptives*” kemudian berikan tanda centang (v) pada *Initial solution*, *KMO and Bartlett’s test of sphericity*, dan *Anti-image*, lalu klik *Continue*.



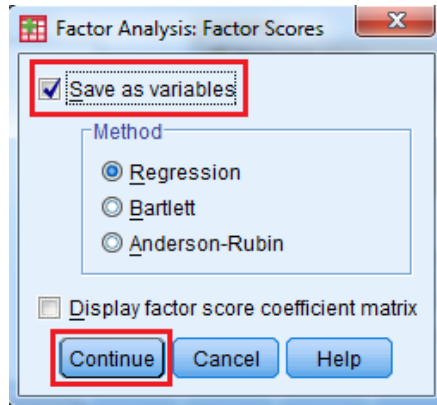
- 6) Selanjutnya klik *Extraction...* maka muncul kotak dialog “*Factor Analysis: Extraction*” berikan tanda centang (v) pada *Unrotated factor solution* dan *Scree plot*, lalu klik *Continue*.



- 7) Berikutnya klik Rotation... maka muncul kotak dialog “Factor Analysis: Rotation” lalu aktifkan Varimax dan beri tanda centang (v) pada Rotated solution, lalu klik Continue.



- 8) Kemudian klik Scores... maka muncul kotak dialog “Factor Analysis: Factor Scores” lalu aktifkan Save as Variable, lalu klik Continue



9) Terakhir adalah klik Ok. Maka akan muncul *output Factor Analysis SPSS*.

c. Interpretasi Analisis Faktor dengan SPSS Lengkap

Tabel 7.11 Output Pertama “KMO and Bartlett's Test”

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.674
Bartlett's Test of Sphericity	Approx. Chi-Square	22.271
	df	10
	Sig.	.014

Tabel output KMO and Bartlett's Test berguna mengetahui kelayakan suatu variabel, apakah dapat di proses lebih lanjut menggunakan teknik analisis faktor ini atau tidak. Caranya dengan melihat nilai KMO MSA (*Kaiser-Meyer-Olkin Measure of Sampling Adequacy*). Jika nilai KMO MSA

lebih besar dari 0,50 maka teknik analisis faktor dapat di lanjutkan. Berdasarkan output di atas diketahui nilai KMO MSA sebesar $0,674 > 0,50$ dan nilai *Bartlett's Test of Sphericity (Sig.)* $0,014 < 0,05$, maka analisis faktor dalam penelitian ini dapat dilanjutkan karena sudah memenuhi persyaratan pertama.

Tabel 7.12 Output Kedua “Anti-image Matrices”

		Anti-image Matrices				
		Reliability	Responsiveness	Assurance	Empathy	Tangibles
Anti-image Covariance	Reliability	.580	-.253	.229	-.110	-.042
	Responsiveness	-.253	.632	-.058	.039	-.119
	Assurance	.229	-.058	.708	-.116	-.095
	Empathy	-.110	.039	-.116	.365	-.227
	Tangibles	-.042	-.119	-.095	-.227	.344
Anti-image Correlation	Reliability	.633 ^a	-.419	.357	-.240	-.093
	Responsiveness	-.419	.725 ^a	-.086	.081	-.255
	Assurance	.357	-.086	.612 ^a	-.229	-.193
	Empathy	-.240	.081	-.229	.673 ^a	-.640
	Tangibles	-.093	-.255	-.193	-.640	.692 ^a

a. Measures of Sampling Adequacy(MSA)

Anti-image Matrices berguna untuk mengetahui dan menentukan variabel mana saja yang layak pakai dalam analisis faktor. Perhatikan bagian Anti-image Correlation, pada tabel tersebut terdapat kode huruf (a) yang artinya tanda untuk *Measure of Sampling Adequacy* (MSA). Diketahui nilai MSA dari masing-masing yang diteliti adalah sebagai berikut:

- 1) Reliability sebesar 0,633
- 2) Responsieness sebesar 0,725
- 3) Assurance sebesar 0,612
- 4) Empathy sebesar 0,673
- 5) Tangibles sebesar 0,692

Persyaratan yang harus terpenuhi dalam analisis faktor adalah nilai $MSA > 0,50$. Dari hasil di atas diketahui bahwa nilai MSA untuk semua variabel yang diteliti adalah $> 0,50$, maka persyaratan kedua dalam analisis faktor ini pun terpenuhi.

Tabel 7.13 Output Ketiga “Communalities”

	Initial	Extraction
Reliability	1.000	.804
Responsiveness	1.000	.631
Assurance	1.000	.812
Empathy	1.000	.775
Tangibles	1.000	.813

Extraction Method: Principal Component Analysis.

Tabel Communalities ini menunjukkan nilai variabel yang diteliti apakah mampu untuk menjelaskan faktor atau tidak. Variabel dianggap mampu menjelaskan faktor jika nilai Extraction lebih besar dari 0,50. Berdasarkan output di atas, diketahui nilai Extraction untuk semua variabel adalah lebih besar dari 0,50. Dengan demikian dapat disimpulkan bahwa semua variabel dapat dipakai untuk menjelaskan faktor.

Tabel 7.14 Output Keempat “Total Variance Explained”

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.638	52.751	52.751	2.638	52.751	52.751	2.005	40.090	40.090
2	1.198	23.966	76.717	1.198	23.966	76.717	1.831	36.627	76.717
3	.580	11.594	88.311						
4	.374	7.481	95.792						
5	.210	4.208	100.000						

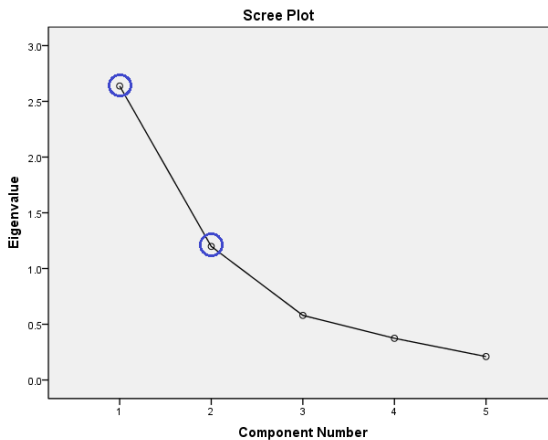
Extraction Method: Principal Component Analysis.

Tabel Total Variance Explained menunjukkan nilai masing-masing variabel yang di analisis. Dalam penelitian ini ada 5 variabel berarti ada 5

Component yang di analisis. Ada dua macam analisis untuk menjelaskan suatu varian, yaitu *Initial Eigenvalues* dan *Extraction Sums of Squared Loadings*. Pada varian *Initial Eigenvalues* menunjukkan faktor yang terbentuk. Apabila semua faktor dijumlahkan menunjukkan jumlah variabel (yaitu $2,368 + 1,198 + 0,580 + 0,374 + 0,210 = 5$ variabel). Sedangkan pada bagian *Extraction Sums of Squared Loadings* menunjukkan jumlah variasi atau banyaknya faktor yang dapat terbentuk, pada hasil output di atas ada 2 (dua) variasi faktor, yaitu 2,638 dan 1,198.

Berdasarkan tabel output *Total Variance Explained* pada bagian "*Initial Eigenvalues*", maka ada 2 (dua) faktor yang dapat terbentuk dari 5 variabel yang di analisis. Dimana syarat untuk menjadi sebuah factor, maka nilai Eigenvalue harus lebih besar 1. Nilai *Eigenvalue Component* 1 sebesar 2,638 atau >1 maka menjadi faktor 1 dan mampu menjelaskan 52,751% variasi. Sedangkan nilai *Eigenvalue Component* 2 sebesar 1,198 atau >1 maka menjadi faktor 2 dan mampu menjelaskan 23,966% variasi. Jika faktor 1 dan faktor 2 dijumlahkan maka mampu menjelaskan 76,717% variasi.

Tabel 7.15 Output Kelima "Scree Plot".



Gambar *Scree Plot* ini dapat juga menunjukkan jumlah faktor yang terbentuk. Caranya dengan melihat nilai titik *Component* yang memiliki nilai *Eigenvalue* > 1. Dari gambar *Scree Plot* di atas ada 2 titik *Component* yang memiliki nilai *Eigenvalue* >1 maka dapat diartikan bahwa ada 2 faktor yang dapat terbentuk.

Tabel 3.7 Output Keenam “Component Matrix”

Component Matrix^a

	Component	
	1	2
Reliability	.645	-.623
Responsiveness	.703	-.370
Assurance	.450	.781
Empathy	.858	.197
Tangibles	.888	.159

Extraction Method: Principal Component Analysis.

a. 2 components extracted.

Component Matrix ini menunjukkan nilai korelasi atau hubungan antara masing-masing variabel dengan faktor yang akan terbentuk. Misal: dari output di atas terlihat pada variabel *Reliability*, yakni nilai korelasi variabel ini dengan faktor 1 adalah sebesar 0,645, dan korelasi dengan faktor 2 adalah sebesar -0,623. Untuk variabel yang lain cara memaknainya sama seperti pada variabel *Reliability*.

Tabel 3.8 Output Ketujuh “Rotated Component Matrix”

Rotated Component Matrix^a

	Component	
	1	2
Reliability	.896	-.038
Responsiveness	.771	.190
Assurance	-.181	.883
Empathy	.511	.717
Tangibles	.559	.708

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

Untuk memastikan suatu variabel masuk dalam kelompok faktor mana, maka dapat ditentukan dengan melihat nilai korelasi terbesar antara variabel dengan faktor (Component) yang terbentuk. Cara membaca hasil analisis faktor model rotasi, dapat mengikuti penjelasan berikut ini.

- 1) Variabel Reliability. Nilai korelasi variabel ini dengan faktor 1 = 0,896 dan faktor 2 = -0,038, karena nilai korelasi faktor 1 > faktor 2 maka variabel Keandalan termasuk kelompok faktor 1
- 2) Variabel Responsiviness. Nilai korelasi variabel ini dengan faktor 1 = 0,771 dan faktor 2 = 0,190, karena nilai korelasi faktor 1 > faktor 2 maka variabel Ketanggapan termasuk kelompok faktor 1
- 3) Variabel Assurance. Nilai korelasi variabel ini dengan faktor 1 = -0,181 dan faktor 2 = 0,883, karena nilai korelasi faktor 2 > faktor 1 maka variabel Keyakinan termasuk kelompok faktor 2
- 4) Variabel Empathy. Nilai korelasi variabel ini dengan faktor 1 = 0,511 dan faktor 2 = 0,717, karena nilai korelasi faktor 2 > faktor 1 maka variabel Empati termasuk kelompok faktor 2
- 5) VariabelTangibles. Nilai korelasi variabel ini dengan faktor 1 = 0,559 dan faktor 2 = 0,708, karena nilai korelasi faktor 2 > faktor 1 maka variabel Berwujud termasuk kelompok faktor 2

Dengan melihat pembahasan di atas maka kesimpulan yang dapat kita ambil dalam analisis faktkor ini adalah sebagai berikut.

Tabel 3.9 Hasil Analisis Faktor

Faktor	Variabel
1	Reliability, dan Responsiveness
2	Assurane, Empathy dan Tangibles

Daftar Rujukan

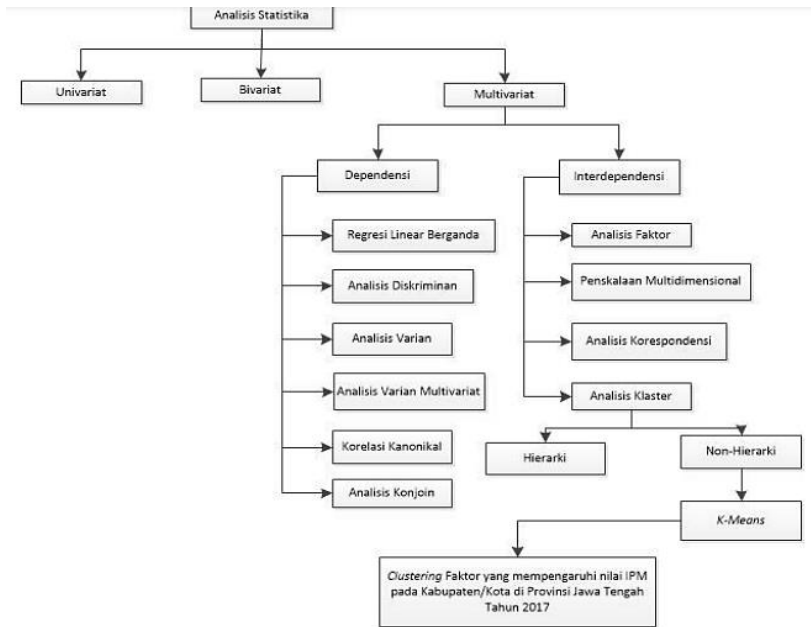
- Thomson B. (2004). *Exploratory and Confirmatory Factor Analysis Understanding Concepts and Applications*. American Psychological Association.
- Leandre R.F. dan Duane T.W. (2012) *Exploratory Factor Analysis*. Oxford University Press, Inc.
- Williams B., Onsmann A., Brown T. (2010). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care (JEPHC)*. 8(3), 990399.
- Yong A.G. dan Pearce S. (2013). A Beginner's Guide to Factor Analysis: Focusing on Exploratory Factor Analysis. *Tutorials in Quantitative Methods for Psychology*. 9(2), 79–94.
- Elpira F. (2014). *Penerapan Analisis Faktor untuk Menentukan Faktor-faktor yang Mempengaruhi Mahasiswa dalam Memilih Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Alauddin Makassar* (Skripsi, Program Studi Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Alauddin Makassar, Makassar.)
- Wirmanto. (2014). *Analisis Faktor Dan Penerapannya Dalam Mengidentifikasi Faktor-Faktor Yang Mempengaruhi Kepuasan Konsumen Terhadap Penjualan Media Pembelajaran* (Skripsi, Program Studi Matematika Jurusan Pendidikan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Yogyakarta, Yogyakarta.)

BAB 8

ANALISIS KLASTER

A. Konsep Analisis Kluster

Gambar 8.1 Analisis Statistika.



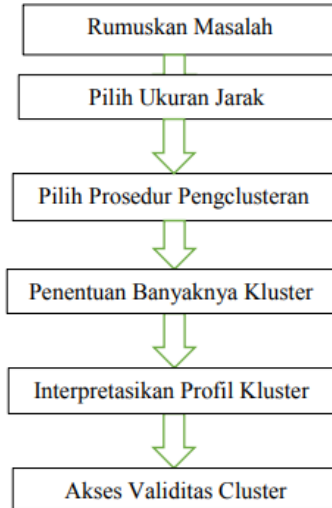
Analisis kluster adalah teknik multivariat yang mempunyai tujuan utama untuk mengelompokkan objek-objek/cases berdasarkan karakteristik yang dimilikinya. Kluster atau ‘kluster’ dapat diartikan ‘kelompok’ dan pada dasarnya analisis kluster akan menghasilkan sejumlah kluster (kelompok) (Santoso, 2014). Analisis kluster mengklasifikasi objek sehingga setiap objek

yang memiliki sifat yang mirip (paling dekat kesamaannya) akan mengelompok ke dalam satu kluster (kelompok) yang sama.

Sekilas memang terdapat kemiripan antara analisis faktor dengan analisis kluster, keduanya disebut dengan “faktor analisis”. Namun, terdapat perbedaan penting di antara keduanya. Jika analisis faktor (disebut dengan *R factor analysis*) bertujuan mereduksi sejumlah variabel menjadi satu atau beberapa faktor maka analisis kluster (disebut dengan *Q factor analysis*) lebih bertujuan untuk mengelompokkan isi variabel walaupun bisa juga disertai dengan pengelompokan variabel. Jika dikaitkan dengan pengolahan data menggunakan SPSS, analisis faktor adalah perlakuan terhadap isi dari kolom data, sedangkan analisis kluster adalah perlakuan terhadap isi dari baris data (Santoso, 2014).

Tujuan utama analisis kluster ialah mengklasifikasi objek (kasus/elemen) seperti orang, produk (barang), toko, perusahaan, ke dalam kelompok-kelompok yang relatif homogen didasarkan pada suatu set variabel yang dipertimbangkan untuk diteliti (Supranto, 2004). Menurut Santoso (2014) secara logika, kluster yang baik adalah kluster yang mempunyai homogenitas (kesamaan) yang tinggi antar-kluster dalam satu kluster (*within-kluster*) dan heterogenitas (perbedaan) yang tinggi antar-kluster yang satu dengan kluster yang lainnya (*between-kluster*).

Dari dua hal di atas, dapat disimpulkan bahwa sebuah kluster yang baik adalah kluster yang mempunyai anggota-anggota yang semirip mungkin satu dengan yang lain, tetapi sangat tidak mirip dengan anggota-anggota kluster yang lain. Di sini, 'mirip' diartikan sebagai tingkat kesamaan karakteristik antara dua data.

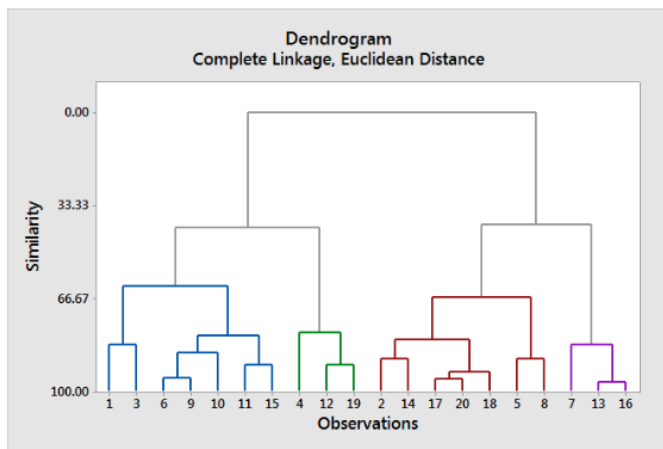


Gambar 8.2 Prosedur Analisis Kluster.

Dalam analisis kluster terdapat 2 metode yang dapat digunakan yaitu *hirarchical method* dan *non-hirarchical method*.

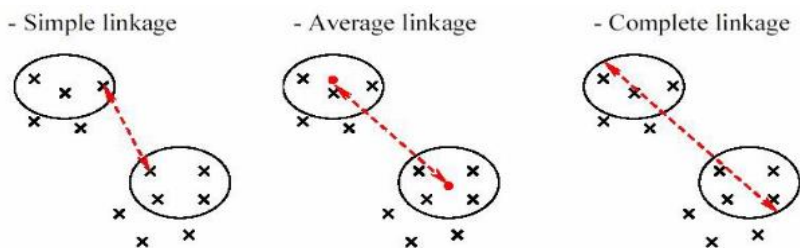
1. *Hirarchical Method*

Metode ini memulai pengelompokan dengan dengan dua atau lebih objek yang mempunyai kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang mempunyai kedekatan kedua. Demikian seterusnya hingga kluster akan membentuk semacam “pohon” yang memiliki hierarki (tingkatan) yang jelas antarobjek, dari yang paling mirip sampai paling tidak mirip. Secara logika, semua objek pada akhirnya akan membentuk sebuah kluster. Hasil dari pengklusteran dengan menggunakan metode ini dapat disajikan dengan menggunakan dendogram.



Gambar 8.3 Dendrogram.

Dendrogram adalah representasi visual dari langkah-langkah dalam analisis kluster yang menunjukkan bagaimana kluster terbentuk dan nilai koefisien jarak pada setiap langkah. Angka di sebelah kanan adalah objek penelitian, tempat objek-objek tersebut dihubungkan oleh garis dengan objek yang lain sehingga pada akhirnya akan membentuk satu kluster.



Gambar 8.4 Teknik pengelompokan pada *agglomerative hierarchical clustering*.

Metode-metode yang bisa digunakan dalam metode hierarki adalah metode agglomeratif (*agglomerative method*) dan metode defisif (*devisive method*).

a. Metode Aglomerasi (*Agglomerative Method*)

Aglomerasi adalah bagian yang sangat penting pada saat kita coba interpretasi analisis kluster hierarki ini. Proses aglomerasi ini bersifat kompleks, khususnya perhitungan koefisien yang melibatkan sekian banyak objek dan terus bertambah. Proses aglomerasi pada akhirnya akan menyatukan semua objek menjadi satu kluster. Hanya saja dalam prosesnya dihasilkan beberapa kluster dengan masing-masing anggotanya, tergantung jumlah kluster yang dibentuk. Metode aglomerasi terdiri dari beberapa macam, yaitu

1) *Metode Single Linkage (Jarak Terdekat atau Tautan Tunggal)*

Untuk menentukan jarak antar kluster dengan menggunakan metode *single linkage* dapat dilakukan dengan melihat jarak antar dua kluster yang ada kemudian memilih jarak paling dekat atau aturan tetangga dekat (*nearest neighbour rule*).

2) *Metode Complete Linkage (Jarak Rata-Rata atau Tautan Rata-Rata)*

Pada metode *complete linkage*, jarak antar kluster ditentukan oleh jarak terjauh (*farthest-neighbour*) antara dua objek dalam kluster yang berbeda.

3) *Metode Centroid*

Centroid adalah rata-rata semua objek dalam kluster. Pada metode ini, jarak antarkluster adalah jarak antar-*centroid*. *Centroid* baru dihitung ketika setiap kali objek digabungkan sehingga setiap kali anggotanya bertambah maka *centroid*-nya akan berubah.

4) *Metode Average Linkage*

Pada metode *average linkage*, jarak antara dua kluster dianggap sebagai jarak rata-rata antara semua anggota dalam satu kluster dengan semua anggota kluster lain.

5) *Metode Ward*

Metode varian bertujuan untuk memperoleh kluster yang memiliki varian internal kluster yang sekecil mungkin. Metode varian yang umum dipakai adalah metode *Ward* dengan rata-rata untuk setiap kluster dihitung. Lalu jarak euclides antara setiap objek dan nilai rata-rata itu dihitung. Pada setiap tahap, dua kluster yang memiliki kenaikan '*sum of squares* dalam kluster' yang terkecil digabungkan.

Ward merupakan suatu metode pembentukan kluster yang didasari oleh hilangnya informasi akibat penggabungan objek menjadi kluster. Hal ini diukur dengan menggunakan jumlah total dari deviasi kuadrat pada mean kluster untuk setiap pengamatan. Error Sum of Squares (SSE) digunakan sebagai fungsi objektif. Dua objek akan digabungkan jika mempunyai fungsi objektif terkecil di antara kemungkinan yang ada.

b. Metode Devisif

Proses dalam metode divisif berkebalikan dengan metode aglomerasi. Metode ini dimulai dengan satu kluster besar yang mencakup semua objek pengamatan. Selanjutnya, secara bertahap objek yang mempunyai ketidakmiripan cukup besar akan dipisahkan ke dalam kluster-kluster yang berbeda. Proses dilakukan sehingga terbentuk sejumlah kluster yang diinginkan, seperti, dua kluster, tiga kluster, dan seterusnya.

2. *Non-Hierarchical Method*

Non-hierarchical method merupakan pengelompokan berbasis sekatan atau *partitioning* yang menghasilkan partisi dari data sehingga objek

dalam kluster lebih mirip satu sama lain daripada objek yang ada dalam kluster lain (Triyanto, 2015). Berbeda dengan hierarki, prosedur pengelompokan ini tidak dilakukan secara bertahap dan jumlah kluster ditentukan terlebih dahulu (Machfudhoh, 2013). Biasanya metode nonhierarki diterapkan pada data yang memiliki jumlah sangat banyak.

Pusat kluster atau biasa disebut *centroid* yang dipilih pada metode ini merupakan pusat kluster sementara dengan terus memperbaharui pusat kluster untuk tiap iterasi sampai kriteria pemberhentian tercapai, sehingga dimungkinkan bahwa objek yang telah berada pada suatu kelompok atau kluster tertentu dapat pindah ke kluster yang lain (Safe'i, 2018). Salah satu metode pengelompokan nonhierarki adalah K-Means.

K-Means merupakan salah satu metode pengelompokan data nonhierarki yang berusaha memartisi data yang ada ke dalam bentuk dua atau lebih kluster. Metode ini memartisi data ke dalam kluster sehingga data berkarakteristik sama dimasukkan ke dalam satu kluster yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kluster yang lain. Adapun tujuan pengelompokan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi dalam suatu kluster dan memaksimalkan variasi antar-kluster (Supranto, 2004).

K-Means menggunakan “mean”, yaitu rata-rata sebagai pusat kluster-nya sehingga K-Means tidak *robust* atau tidak cukup baik terhadap data yang memiliki *outlier*. Dengan demikian, untuk mengatasi hal tersebut dapat digunakan metode K-Medoids untuk mengelompokkan data yang memiliki *outlier* (Larasati, 2017).

B. Asumsi Analisis Klaster

Menurut Santoso (2014), asumsi untuk analisis klaster, yaitu

- 1) Sampel yang diambil benar-benar bisa mewakili populasi yang ada. Memang tidak ada ketentuan jumlah sampel yang representatif, tetapi tetap diperlukan sejumlah sampel yang cukup besar agar proses penglasteran bisa dilakukan dengan benar.
- 2) Multikolinieritas, yakni kemungkinan adanya korelasi antarobjek. Sebaliknya tidak ada atau seandainya ada, besar multikolinieritas tersebut tidaklah tinggi (misal di atas 0,5). Jika sampai terjadi multikolinieritas, dianjurkan untuk menghilangkan salah satu variabel dari dua variabel yang mempunyai korelasi cukup besar.

C. Perhitungan Menggunakan Microsoft Excel Analisis Klaster Nonhierarki Menggunakan Metode K-Means

Non-hierarchical methods merupakan pengelompokan berbasis sekatan atau *partitioning* yang menghasilkan partisi dari data sehingga objek dalam klaster lebih mirip satu sama lain dari pada objek yang ada dalam klaster lain (Triyanto, 2015). Berbeda dengan hierarki, prosedur pengelompokan ini tidak dilakukan secara bertahap dan jumlah klaster ditentukan terlebih dahulu (Machfudhoh, 2013). Biasanya, metode nonhierarki diterapkan pada data yang memiliki jumlah sangat banyak.

1. Judul Penelitian

Bagaimana Hasil Pengelompokan Siswa Kelas X RPL 1 Berdasarkan Nilai Rata-Rata Mata Pelajaran Produktif Kompetensi Keahlian Pada Semester 1 Dengan Menggunakan Metode Analisis Klaster K-Mean.

Mata Pelajaran Produktif /Atribut:

- a. Sistem Komputer.
- b. Pemrograman Dasar.
- c. Teknik Komputer dan Jaringan.

Tabel 8.1 Data hasil belajar siswa X RPL 1.

No	Nama Siswa	Daftar Nilai Rata-Rata Pelajaran Produktif		
		Sistem Komputer (X)	Pemrogra man Dasar (Y)	Teknik Komputer dan Jaringan (Z)
1.	Achmad Rafi	78	80	82
2.	Adam Surya Pratama	75	83	85
3.	Aditya Fajar Mulyana	95	98	93
4.	Adrian Alam F.	83	85	87
5.	Afika Rara Rahmawati	95	92	93
6.	Ajeng Intan Kumala S.	85	88	86
7.	Ajeng Wulandari J.	88	85	78
8.	Alvin Winar Saputra	65	75	70
9.	Alvina Aulia Putri	80	86	89
10.	Alvina Windi Pertiwi	70	73	75
11.	Alyssa Dinda R.	68	78	80
12.	Amaelia Ayu R.	73	82	80
13.	Amirudin Yahya	68	78	72
14.	Ananda Sesarianti	78	80	67
15.	Angelia Hana Arsita	78	65	80
16.	Anitya Risa Amalia	69	79	70
17.	Cika Suci Dewi Utama	90	91	88
18.	Daffa Hendri W.	66	68	67
19.	Dea Nita Anggraini	69	70	67
20.	Dea Novita Putri	71	73	72
21.	Dias Agung Pratama	80	83	87

22.	Dwi Widiyanti	65	68	80
23.	Eliyana Agustin	90	85	80
24.	Endih Dwi Yuliyantha	80	82	81
25.	Fajar Aditya	82	83	87
26.	Farida Ajeng Triana	90	91	93
27.	Fica Febi Dewanti	85	87	89
28.	Galang Wicaksono	68	69	80
29.	Hefa Anisa Ilma	72	78	71
30.	Haziq Amirul Cahyo	90	89	95

2. Tahap-Tahap yang Dilakukan dalam Analisis Kluster K-Means

Berikut ini merupakan tahap-tahap untuk melakukan analisis menggunakan kluster K-Means.

- a. Menentukan jumlah banyak kelompok yang akan dibuat atau dibentuk sebagai pusat kluster.
- b. Menentukan beberapa data set yang akan dijadikan sebagai pusat kluster (*centroid*).
- c. Menghitung jarak antarobjek masing-masing *centroid*. Dengan menggunakan rumus jarak euclides (*euclidean distance*).

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

Keterangan:

d = Jarak

x_1 = Nilai pada data ke-1

x_2 = Nilai pada data *centroid* 1

y_1 = Nilai pada data ke-2

y_2 = Nilai pada data *centroid* 2

z_1 = Nilai pada data ke-3

z_2 = Nilai pada data *centroid* 3

- d. Mengelompokkan objek berdasarkan jarak terdekat dengan *centroid*.
- e. Menentukan *centroid* baru dengan menggunakan rumus:

$$C_{m(q)} = \frac{1}{n_m} \sum_{i=1}^{n_m} x_{i(q)}$$

- f. Mengulangi langkah 3 dan 4 sampai tidak ada lagi data yang berpindah klaster.

3. Iterasi Pertama

- a. Menentukan jumlah klaster yang akan dibuat sebanyak 3 klaster.
- b. Menentukan beberapa data set yang akan dijadikan sebagai pusat klaster (*centroid*).

Tabel 8.2 Data Set.

Data ke	Centroid	Nilai X	Nilai Y	Nilai Z
5	C1	95	92	93
18	C2	66	68	67
25	C3	82	83	87

- c. Menghitung jarak antarobjek masing-masing *centroid*. Dengan menggunakan rumus jarak euclides (*euclidean distance*).

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

Selanjutnya menentukan nilai minum dan menentukan masuk dalam klaster ke berapa. Hasilnya sebagai berikut ini.

Tabel 8.3 Nilai Minimum.

Data ke	Nilai X	Nilai Y	Nilai Z	C1	C2	C3	Minimum	Klaster
1	78	80	82	23,54	22,65	7,07	7,07	3
2	75	83	85	23,35	25,10	7,28	7,28	3
3	95	98	93	6,00	49,16	20,74	6,00	1
4	83	85	87	15,13	31,27	2,24	2,24	3
5	95	92	93	0,00	45,75	16,91	0,00	1
6	85	88	86	12,85	33,50	5,92	5,92	3
7	88	85	78	17,97	29,90	11,00	11,00	3
8	65	75	70	41,45	7,68	25,34	7,68	2
9	80	86	89	16,64	31,69	4,12	4,12	3
10	70	73	75	36,19	10,25	19,70	10,25	2
11	68	78	80	33,08	16,52	16,43	16,43	2
12	73	82	80	27,44	20,35	11,45	11,45	3
13	68	78	72	36,96	11,36	21,12	11,36	2
14	78	80	67	33,30	16,97	20,62	16,97	2
15	78	65	80	34,45	17,94	19,72	17,94	2
16	69	79	70	37,07	11,79	21,77	11,79	2
17	90	91	88	7,14	39,32	11,36	7,14	1
18	66	68	67	45,75	0,00	29,68	0,00	2
19	69	70	67	42,85	3,61	27,17	3,61	2
20	71	73	72	37,12	8,66	21,12	8,66	2
21	80	83	87	18,49	28,65	2,00	2,00	3
22	65	68	80	40,56	13,04	23,73	13,04	2
23	90	85	80	15,59	32,16	10,82	10,82	3
24	80	82	81	21,66	24,25	6,40	6,40	3
25	82	83	87	16,91	29,68	0,00	0,00	3
26	90	91	93	5,10	42,20	12,81	5,10	1
27	85	87	89	11,87	34,73	5,39	5,39	3
28	68	69	80	37,78	13,19	21,00	13,19	2
29	72	78	71	34,77	12,33	19,52	12,33	2

30	90	89	95	6,16	42,44	12,81	6,16	1
----	----	----	----	------	-------	-------	------	---

Perhitungan tabel di atas menggunakan rumus Microsoft Excel yang sudah dibagikan pada grup pesan WhatsApp.

- d. Memetakan *centroid* pada masing-masing kluster.

Tabel 8.4 Pemetaan *Centroid*.

Data ke	C1	C2	C3
1			1
2			1
3	1		
4			1
5	1		
6			1
7			1
8		1	
9			1
10		1	
11		1	
12			1
13		1	
14		1	
15		1	
16		1	
17	1		
18		1	
19		1	
20		1	
21			1
22		1	
23			1
24			1

25		1
26	1	
27		1
28		1
29		1
30	1	

- e. Langkah selanjutnya memastikan tidak ada perubahan data set pada iterasi selanjutnya.

4. Iterasi Kedua

- a. Menghitung nilai masing-masing *centroid* baru dengan rumus:

$$C_{m(q)} = \frac{1}{n_m} \sum_{i=1}^{n_m} x_{i(q)}$$

Centroid baru = rata-rata dari set data pada masing-masing *centroid*.

Hasilnya sebagai berikut.

Tabel 8.5 Hasil Nilai *Centroid* Baru.

Penentuan Kluster Baru	Nilai X	Nilai Y	Nilai Z
<i>Centroid</i> baru yang ke-1	92	92,2	92,4
<i>Centroid</i> baru yang ke-2	69,77	73,38	73,15
<i>Centroid</i> baru yang ke-3	81,58	84,08	84,25

Perhitungan tabel di atas menggunakan rumus Microsoft Excel yang sudah dibagikan pada grup pesan WhatsApp.

- b. Menghitung jarak antarobjek masing-masing *centroid*. Dengan menggunakan rumus jarak euclides (*euclidean distance*).

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$

Selanjutnya, menentukan nilai minum dan menentukan masuk dalam kluster ke berapa. Hasilnya sebagai berikut ini.

Tabel 8.6 Hasil Nilai Minimum *Centroid*.

Data ke	Nilai X	Nilai Y	Nilai Z	C1	C2	C3	Minimum	Kluster
1	78	80	82	21,28	13,78	5,88	5,88	3
2	75	83	85	20,70	16,13	6,71	6,71	3
3	95	98	93	6,56	40,45	21,22	6,56	1
4	83	85	87	12,73	22,40	3,23	3,23	3
5	95	92	93	3,07	37,11	17,87	3,07	1
6	85	88	86	10,37	24,71	5,48	5,48	3
7	88	85	78	16,59	22,15	9,00	9,00	3
8	65	75	70	39,07	5,94	23,68	5,94	2
9	80	86	89	13,93	22,69	5,36	5,36	3
10	70	73	75	33,99	1,90	18,51	1,90	2
11	68	78	80	30,52	8,44	15,48	8,44	2
12	73	82	80	24,88	11,47	9,80	9,80	3
13	68	78	72	34,55	5,08	19,28	5,08	2
14	78	80	67	31,46	12,22	18,09	12,22	2
15	78	65	80	33,01	13,60	19,88	13,60	2
16	69	79	70	34,71	6,49	19,68	6,49	2
17	90	91	88	4,98	30,66	11,52	4,98	1
18	66	68	67	43,67	9,00	28,27	9,00	2
19	69	70	67	40,83	7,07	25,58	7,07	2
20	71	73	72	35,01	1,73	19,62	1,73	2
21	80	83	87	16,06	19,72	3,35	3,35	3
22	65	68	80	38,32	9,93	23,49	9,93	2
23	90	85	80	14,48	24,31	9,47	9,47	3
24	80	82	81	19,44	15,51	4,17	4,17	3
25	82	83	87	14,62	20,83	2,98	2,98	3
26	90	91	93	2,41	33,37	13,97	2,41	1
27	85	87	89	9,36	25,85	6,54	6,54	3

28	68	69	80	35,61	8,32	20,74	8,32	2
29	72	78	71	32,55	5,56	17,45	5,56	2
30	90	89	95	4,58	33,62	14,51	4,58	1

Perhitungan tabel di atas menggunakan rumus Microsoft Excel yang sudah dibagikan pada grup pesan WhatsApp.

- c. Membandingkan hasil iterasi ke-1 dan ke-2.

Tabel 8.7 Perbandingan Iterasi 1 dan 2.

Data ke	Klaster pada iterasi 1	Klaster pada iterasi 2	Keterangan
1	3	3	Aman
2	3	3	Aman
3	1	1	Aman
4	3	3	Aman
5	1	1	Aman
6	3	3	Aman
7	3	3	Aman
8	2	2	Aman
9	3	3	Aman
10	2	2	Aman
11	2	2	Aman
12	3	3	Aman
13	2	2	Aman
14	2	2	Aman
15	2	2	Aman
16	2	2	Aman
17	1	1	Aman
18	2	2	Aman
19	2	2	Aman
20	2	2	Aman
21	3	3	Aman

22	2	2	Aman
23	3	3	Aman
24	3	3	Aman
25	3	3	Aman
26	1	1	Aman
27	3	3	Aman
28	2	2	Aman
29	2	2	Aman
30	1	1	Aman

Jika data set tidak terjadi perubahan (aman) maka iterasi selanjutnya dihentikan.

D. Perhitungan SPSS Analisis Kluster

Permasalahan dalam penelitian ini dapat dirumuskan dalam data sebagai berikut.

Tabel 8.9 Tabel Data.

Kabupaten	Keluar ga Pertani an (%)	Pendu duk Tidak Tamat SD (%)	Banyak nya Industr i Kecil dan Menen gah	Banyak nya Industr i Pengol aan	Pangasa Pasar Rokok Produksi Gudang Garam	Pangasa Pasar Rokok Produksi Sampoer-na	Pangasa Pasar Rokok Produksi Djarum Kudus
Banyuwangi	21	30	166	55	2.7	10	13
Cianjur	26	35	220	206	6.9	10	12
Jember	10	5.2	538	255	20	29	27
Jepara	8.4	4.5	555	119	13	27	18
Jombang	16	25	86	94	10	11	11
Kediri	5.7	2.3	468	287	19	28	24
Kudus	9.8	15	579	197	21	30	26
Kulonprogo	7.1	13	404	48	14	18	18
Lumajang	38	33	71	146	4.5	1.1	8.5
Majalengka	22	24	114	197	12	20	18

Pacitan	24	29	152	79	6	13	8.1
Pamekasan	31	30	63	40	3.2	2.9	7.1
Pasuruan	21	5.2	40	38	19	26	16
Pati	11	7.2	447	160	11	18	16
Probolinggo	8.6	5.1	214	205	11	27	15
Rembang	9.9	4.9	111	17	9.9	25	17
Serang	6.9	18	467	189	5.9	25	17
Sleman	16	3.4	419	300	18	28	26
Sukabumi	22	34	215	183	9.2	2.7	12

1. Langkah-Langkah dalam SPSS

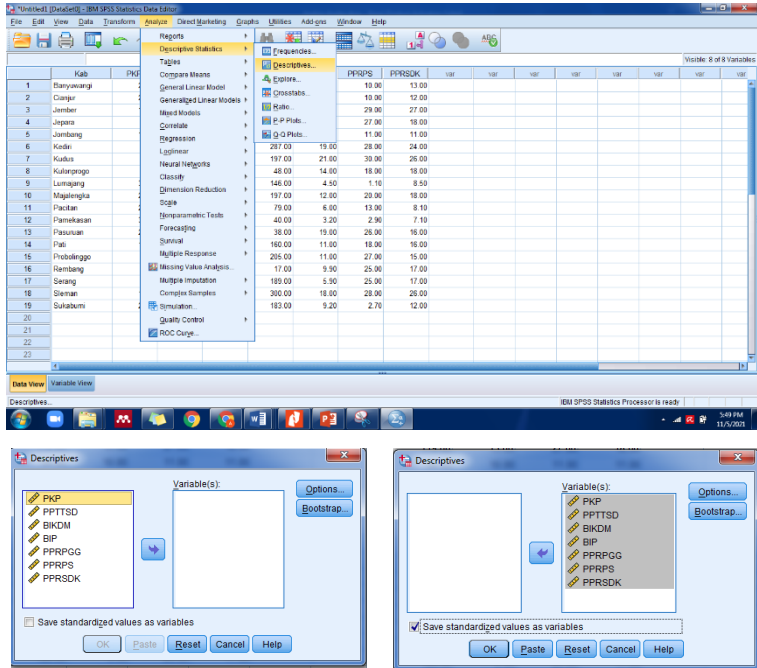
The screenshot shows the 'Data Editor' window in SPSS. The 'Data Editor' tab is active, displaying a list of variables with their properties. The variables are: Kab (String, width 11, decimals 0), PKP (Numeric, width 8, decimals 2), PPTTSD (Numeric, width 8, decimals 2), BIKDM (Numeric, width 8, decimals 2), BIP (Numeric, width 8, decimals 2), PPRPGG (Numeric, width 8, decimals 2), PPRPS (Numeric, width 8, decimals 2), and PPRSDK (Numeric, width 8, decimals 2). The 'Measure' column shows 'Nominal' for 'Kab' and 'Scale' for the other variables. The 'Role' column shows 'Input' for all variables.

	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Kab	String	11	0		None	None	11	Left	Nominal	Input
2	PKP	Numeric	8	2		None	None	8	Right	Scale	Input
3	PPTTSD	Numeric	8	2		None	None	8	Right	Scale	Input
4	BIKDM	Numeric	8	2		None	None	8	Right	Scale	Input
5	BIP	Numeric	8	2		None	None	8	Right	Scale	Input
6	PPRPGG	Numeric	8	2		None	None	8	Right	Scale	Input
7	PPRPS	Numeric	8	2		None	None	8	Right	Scale	Input
8	PPRSDK	Numeric	8	2		None	None	8	Right	Scale	Input
9											

The screenshot shows the 'Data Editor' window in SPSS, displaying a data view of the dataset. The variables are: Kab, PKP, PPTTSD, BIKDM, BIP, PPRPGG, PPRPS, and PPRSDK. The data is organized into rows and columns, with the first column representing the 'Kab' variable and the subsequent columns representing the other variables. The data is as follows:

	Kab	PKP	PPTTSD	BIKDM	BIP	PPRPGG	PPRPS	PPRSDK
1	Banyuwangi	21.00	30.00	166.00	55.00	2.70	10.00	13.00
2	Cianjur	26.00	35.00	220.00	206.00	6.90	10.00	12.00
3	Jember	10.00	5.20	538.00	255.00	20.00	29.00	27.00
4	Jepara	8.40	4.50	555.00	119.00	13.00	27.00	18.00
5	Jombang	16.00	25.00	86.00	94.00	10.00	11.00	11.00
6	Kediri	5.70	2.30	468.00	287.00	19.00	28.00	24.00
7	Kudus	9.80	15.00	579.00	197.00	21.00	30.00	26.00
8	Kulonprogo	7.10	13.00	404.00	48.00	14.00	18.00	18.00
9	Lumajang	38.00	33.00	71.00	146.00	4.50	1.10	8.50
10	Majalengka	22.00	24.00	114.00	197.00	12.00	20.00	18.00
11	Pacitan	24.00	29.00	152.00	79.00	6.00	13.00	8.10
12	Pamekasan	31.00	30.00	63.00	40.00	3.20	2.90	7.10
13	Pasuruan	21.00	5.20	40.00	38.00	19.00	26.00	16.00
14	Pati	11.00	7.20	447.00	160.00	11.00	18.00	16.00
15	Probolinggo	8.60	5.10	214.00	205.00	11.00	27.00	15.00
16	Rembang	9.90	4.90	111.00	17.00	9.90	25.00	17.00
17	Serang	6.90	18.00	467.00	189.00	5.90	25.00	17.00
18	Sleman	16.00	3.40	419.00	300.00	18.00	28.00	26.00
19	Sukabumi	22.00	34.00	215.00	183.00	9.20	2.70	12.00

- a. Lakukan transformasi atau standardisasi data tersebut. Klik menu Analyze Descriptives Statistics Descriptives. Masukkan seluruh variabel instrumen penilai (dalam hal ini variabel kabupaten tidak dimasukkan karena data bertipe *string*). Kemudian berikan centang pada “Save standardized values as variables“. Kemudian klik OK.

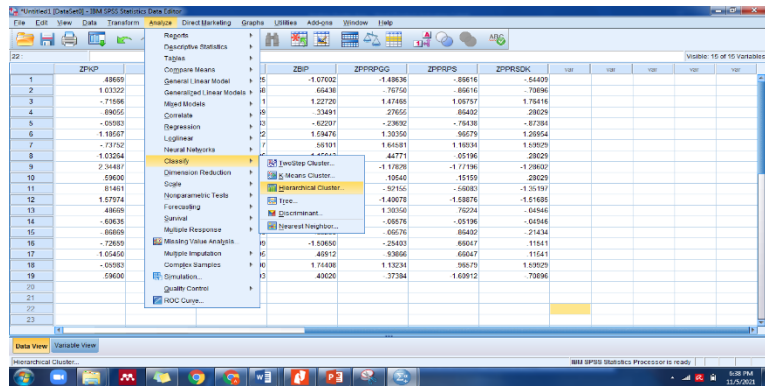
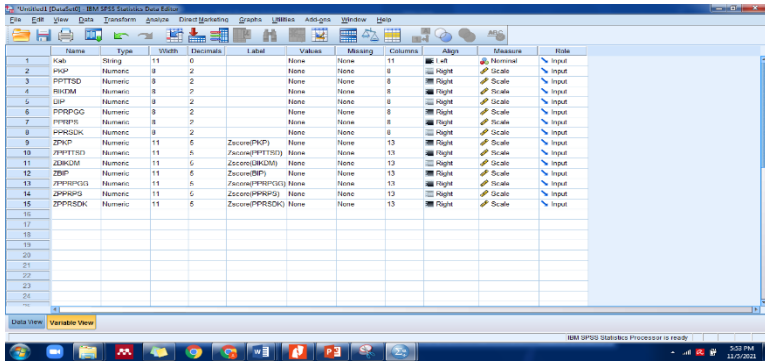


Tabel 8.10 *Descriptive Statistics.*

Descriptive Statistics					
	N	Minimum	Maximum	Mean	Std. Deviation
PKP	19	5.70	38.00	16.5474	9.14873
PPTTSD	19	2.30	35.00	17.0421	12.25834
BIKDM	19	40.00	579.00	280.4737	189.76136
BIP	19	17.00	300.00	148.1579	87.06145

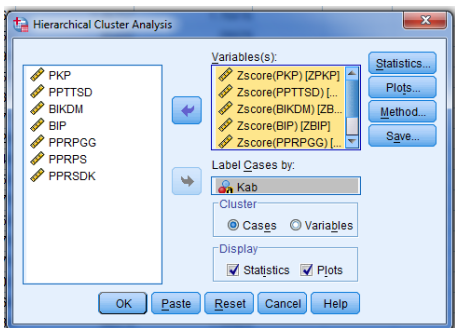
PPRPGG	19	2.70	21.00	11.3842	5.84259
PPRPS	19	1.10	30.00	18.5105	9.82558
PPRS DK	19	7.10	27.00	16.3000	6.06520
Valid N (listwise)	19				

b. Akibatnya muncul *output* deskriptif statistik. Kemudian pada data View akan terlihat juga hasil dari perhitungan *Z-score* dan hasil *Z-score* inilah yang akan dipakai sebagai dasar analisis kluster. Namun, apabila data yang terkumpul tidak mempunyai variabilitas satuan maka proses analisis kluster dapat langsung dilakukan tanpa terlebih dahulu melakukan transformasi atau standarisasi.

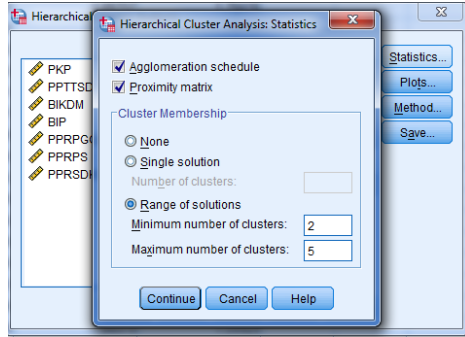


	ZPKP	ZPPTSD	ZBIKDM	ZBIP	ZPPRGG	ZPPRPS	ZPPRSK	VAR	VAR	VAR	VAR	VAR	VAR	VAR
1	48609	1.05707	-6025	-1.07029	-1.48676	-88616	-54409							
2	-1.02322	1.44465	-31848	64438	-76769	68646	-70895							
3	-71468	-96604	1.35711	1.22770	1.47465	1.66757	1.76416							
4	89055	1.02315	1.44669	-23491	27655	66402	28029							
5	-69061	60199	-1.10183	-63297	-73862	-1.16488	-87264							
6	1.18507	1.20262	98822	1.29476	1.30359	56578	1.26954							
7	-73782	-18859	1.57317	69191	1.64881	1.18934	1.59929							
8	1.03264	-20914	66996	1.10523	44771	65196	28029							
9	2.34487	1.30180	-1.10388	-62479	-1.17828	-1.77196	-1.26602							
10	59600	66700	-87729	65191	10140	15150	28029							
11	81401	37949	-87703	73426	-32155	-59092	-1.26197							
12	1.01974	1.05707	1.14604	-1.20250	1.46078	-1.08076	-1.10485							
13	48669	-96604	-1.26724	-1.26529	1.30359	76224	-94946							
14	60636	-80299	87756	13602	-96576	-65196	-94946							
15	86669	-97420	39230	65296	96576	66402	21434							
16	-72651	-39062	-89309	-1.50650	-28463	69647	11541							
17	1.05403	07914	96295	40912	93969	66647	11541							
18	-69061	-1.11268	72900	1.74988	1.12624	59578	1.59929							
19	59600	1.38338	34503	49626	37364	1.69912	70896							
20														
21														
22														
23														

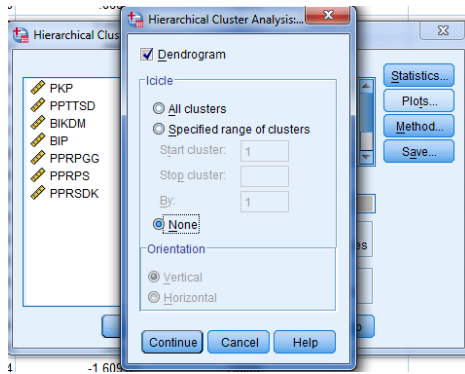
- c. Selanjutnya, klik menu Analyze lalu Classify lalu Hierarchical Kluster. Kemudian masukkan seluruh variabel yang telah distandariskan tadi (Z-score) ke dalam kotak Variable(s). pada bagian Label Cases by isi dengan variabel Kabupaten sedangkan untuk bagian Kluster pilih Cases, pada bagian Display pilih keduanya, yaitu Statistics dan Plots.



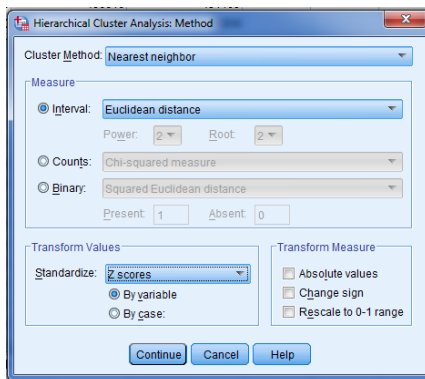
- d. Klik Statistics, berikan centang pada Agglomeration Schedule dan Proximity matrix. Kemudian tekan tombol Continue untuk kembali ke menu utama.



- e. Klik Plots, aktifkan pilihan Dendogram, kemudian pada bagian Icicle pilih None. Selanjutnya klik Continue untuk kembali ke menu utama.



- f. Kemudian klik tombol Method. Pada bagian Klaster Method pilih Nearest Neighbor. Pada Measure pilih Euclidean distance dan pada Transform Values pilih Z-score. Lalu tekan tombol Continue untuk kembali ke menu utama. Dari tampilan menu utama, tekan tombol OK. Pada Klaster Method akan digunakan beberapa metode.



Hasil Output SPSS

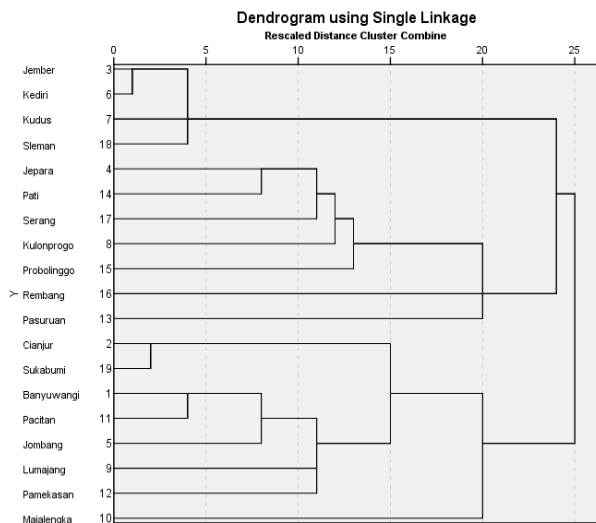
Tabel 8.11 Hasil Output SPSS.

<i>Cluster Membership</i>				
<i>Case</i>	<i>5 Cluster</i>	<i>4 Cluster</i>	<i>3 Cluster</i>	<i>2 Cluster</i>
1: Banyuwangi	1	1	1	1
2: Cianjur	1	1	1	1
3: Jember	2	2	2	2
4: Jepara	3	3	3	2
5: Jombang	1	1	1	1
6: Kediri	2	2	2	2
7: Kudus	2	2	2	2
8: Kulonprogo	3	3	3	2
9: Lumajang	1	1	1	1
10: Majalengka	4	1	1	1
11: Pacitan	1	1	1	1
12: Pamekasan	1	1	1	1
13: Pasuruan	5	4	3	2
14: Pati	3	3	3	2
15: Probolinggo	3	3	3	2
16: Rembang	3	3	3	2
17: Serang	3	3	3	2

18: Sleman	2	2	2	2
19: Sukabumi	1	1	1	1

Hasil *output* di atas menunjukkan pengelompokan objek pengamatan terhadap 5 klaster yang telah diset di awal. Tidak ada alasan khusus mengapa memilih 5 klaster. Hal ini dilakukan agar hasil lebih terpusat pada 2, 3, 4, atau 5 klaster sehingga mudah dibaca. Misalkan, jika menggunakan 2 klaster maka Banyuwangi, Cianjur, Jombang, Lumajang, Majalengka, Pacitan, Pamekasan, dan Sukabumi berada pada klaster 1, sedangkan kabupaten lainnya dikelompokkan pada klaster 2. Begitu pula untuk jumlah klaster sebanyak 3, 4 atau 5.

Untuk 5 klaster, diperoleh pengelompokan klaster 1 yang terdiri dari Banyuwangi, Cianjur, Jombang, Lumajang, Pacitan, Pamekasan, dan Sukabumi. Klaster 2 terdiri dari Jember, Kediri, Kudus, dan Sleman. Klaster 3 terdiri dari Jepara, Kulonprogo, Pati, Probolinggo, Rembang, dan Serang. Klaster 4, yaitu Majalengka dan klaster 5, Pasuruan.



Gambar 8.5 *Dendogram using single linkage.*

Hasil *output* di atas merupakan dendogram hasil analisis kluster dengan metode *single linkage* sehingga makin banyak kluster yang dipilih maka jaraknya makin kecil. Dari dendogram di atas, dengan jarak lebih dari 25 diperoleh 1 kluster (tidak terjadi pengelompokan), sedangkan jika jaraknya 25 didapat 2 kluster, yaitu kluster 1 terdiri dari Cianjur, Sukabumi, Banyuwangi, Pacitan, Jombang, Lumajang, Pamekasan, dan Majalengka, sedangkan kabupaten lainnya masuk ke kluster 2. Hasil ini serupa dengan tabel *output* sebelumnya (tabel kluster *membership*) jika yang dipilih adalah 2 kluster. Begitu pula jika kita mengambil jaraknya adalah 20 maka terdapat 3 kluster dengan kluster 1 terdiri dari Jember, Kediri, Kudus, dan Sleman; kluster 2 terdiri dari Cianjur, Sukabumi, Banyuwangi, Pacitan, Jombang, Lumajang, Pamekasan, dan Majalengka; sedangkan kabupaten sisanya masuk ke kluster 3. Begitu pula untuk jarak lainnya yang menghasilkannya jumlah kluster yang kemungkinan berbeda satu sama lain. Makin banyak kluster yang terbentuk maka jaraknya makin kecil. Artinya, makin memiliki kemiripan yang besar di dalam kluster dan memiliki ketidakmiripan antarkluster.

Dalam pengolahan analisis kluster dengan menggunakan SPSS, dapat dijadikan sebagai catatan bahwa hasil dendogram pada metode hierarki tidak ditentukan jumlah klasternya. Hasil atau *output* dendogram dengan jumlah kluster tertentu ditentukan oleh jaraknya sampai hanya terbentuk 2 kluster. Pada dendogram di atas, jarak maksimumnya adalah 25 (yaitu hingga hanya terbentuk 2 kluster). Berbeda dengan metode non-hierarki seperti K-Means yang terlebih dahulu sudah ditentukan jumlah klasternya.

E. Metode Tak Berhierarki Dengan SPSS

Metode K-Means merupakan salah satu metode analisis non-hierarki yang digolongkan sebagai metode pengklasifikasian yang bersifat *unsupervised* (tanpa arahan) karena data yang dianalisis tidak mempunyai label kelas yang berarti dalam proses pengelompokannya, analisis ini tidak mempunyai anggota kluster yang pasti. Melainkan data yang sudah masuk ke dalam kluster bisa saja berpindah ke kluster yang lain. Akan tetapi, karena peneliti sering menentukan sendiri jumlah kluster awal, baik dengan menggunakan metode tertentu atau berdasarkan pengalaman, maka metode K-Means ini disebut sebagai metode *semi-supervised classification* (Agusta, 2007).

Berbeda dengan metode hierarki, metode nonhierarki justru dimulai dengan menentukan terlebih dahulu jumlah kluster yang diinginkan (dua kluster, tiga kluster, atau yang lain). Setelah jumlah kluster diketahui, baru proses kluster dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan K-Means kluster (Santoso, 2014). Metode K-Means berusaha mengelompokkan data yang ada ke dalam beberapa kelompok yang data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain (Nasari, 2015). Berikut adalah langkah-langkahnya menggunakan SPSS.

1. Klik menu Analyze lalu Descriptives Statistics lalu Descriptives. Masukkan Seluruh variabel instrument penilai (dalam hal ini variabel kabupaten tidak dimasukkan karena data bertipe string). Kemudian berikan centang pada “Save standardized values as variables“. Kemudian klik OK.

Case	Zscore(PKPF)	Zscore(PPTSD)	Zscore(BKMD)	Zscore(BIP)	Zscore(PPRGG)	Zscore(PPRPS)	Zscore(PPRSK)
1	48969	-1.48636	88616	54409			
2	10122	-7610	88616	-8950			
3	-71566	1.47465	1.04787	1.76146			
4	89955	2.7855	88402	28129			
5	-05883	-2.9595	-74110	4.87384			
6	-1.18847	1.30350	56379	1.26554			
7	73742	1.64581	1.78384	1.95928			
8	-1.03764	4.4771	-0.1106	28075			
9	2.34487	-1.17828	1.17190	1.28602			
10	59889	60481	31310	28020			
11	81141	-59185	-46683	-1.38107			
12	1.97974	1.64870	1.58870	1.18601			
13	48969	1.30350	76254	61840			
14	88883	-88180	-61840	1.64866			
15	89959	65529	65576	85482	21424		
16	73659	-3.66655	-3.66637	66687	11541		
17	1.92450	48912	53950	66647	11541		
18	99182	1.44329	1.13238	36579	1.95929		
19	89889	48920	-3.7858	-1.68957	-78956		
20							
21							
22							
23							

Variable(s):

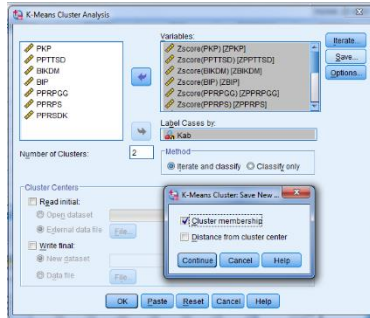
- Zscore(PKPF) [ZPKPF]
- Zscore(PPTSD) [Z]
- Zscore(BKMD) [ZB]
- Zscore(BIP) [ZBIP]
- Zscore(PPRGG) [L]
- Zscore(PPRPS) [ZP]
- Zscore(PPRSK) [Z]

Save standardized values as variables

OK Paste Reset Cancel Help

2. Selanjutnya, klik menu Analyze lalu Classify lalu K-Means Kluster. Kemudian masukkan seluruh variabel yang telah distandardisasi ke dalam kotak Variable(s). pada bagian Label Cases by isi dengan variabel Kabupaten sedangkan untuk bagian Number of Klusters masukkan angka 2 sesuai dengan perintah dari soal. Klik pilihan SAVE, beri tanda ceklist pada “Cluster Membership”. Selanjutnya pilih OK.

Case	Zscore(PKPF)	Zscore(PPTSD)	Zscore(BKMD)	Zscore(BIP)	Zscore(PPRGG)	Zscore(PPRPS)	Zscore(PPRSK)
1	48969	-1.48636	88616	54409			
2	10122	-7610	88616	-8950			
3	-71566	1.47465	1.04787	1.76146			
4	89955	2.7855	88402	28129			
5	-05883	-2.9595	-74110	4.87384			
6	-1.18847	1.30350	56379	1.26554			
7	73742	1.64581	1.78384	1.95928			
8	-1.03764	4.4771	-0.1106	28075			
9	2.34487	-1.17828	1.17190	1.28602			
10	59889	60481	31310	28020			
11	81141	-59185	-46683	-1.38107			
12	1.97974	1.64870	1.58870	1.18601			
13	48969	1.30350	76254	61840			
14	88883	-88180	-61840	1.64866			
15	89959	65529	65576	85482	21424		
16	73659	-3.66655	-3.66637	66687	11541		
17	1.92450	48912	53950	66647	11541		
18	99182	1.44329	1.13238	36579	1.95929		
19	89889	48920	-3.7858	-1.68957	-78956		
20							
21							
22							
23							



F. Interpretasi Tak Berhierarki

Tabel 8.12 Initial Cluster Centers.

	<i>Initial Cluster Centers</i>	
	<i>Cluster</i>	
	1	2
Zscore(PKP)	1.57974	-.71566
Zscore(PPTTSD)	1.05707	-.96604
Zscore(BIKDM)	-1.14604	1.35711
Zscore(BIP)	-1.24232	1.22720
Zscore(PPRPGG)	-1.40078	1.47465
Zscore(PPRPS)	-1.58876	1.06757
Zscore(PPRSDK)	-1.51685	1.76416
Zscore(PKP)	1.57974	-.71566
Zscore(PPTTSD)	1.05707	-.96604
Zscore(BIKDM)	-1.14604	1.35711
Zscore(BIP)	-1.24232	1.22720
Zscore(PPRPGG)	-1.40078	1.47465
Zscore(PPRPS)	-1.58876	1.06757
Zscore(PPRSDK)	-1.51685	1.76416

Tabel 8.13 *Final Cluster Centers.*

<i>Final Cluster Centers</i>		
	<i>Cluster</i>	
	1	2
Zscore(PKP)	.92391	-.67194
Zscore(PPTTSD)	1.05707	-.76878
Zscore(BIKDM)	-.76200	.55418
Zscore(BIP)	-.26599	.19345
Zscore(PPRPGG)	-.78248	.56908
Zscore(PPRPS)	-.98447	.71598
Zscore(PPRSDK)	-.83880	.61004
Zscore(PKP)	.92391	-.67194
Zscore(PPTTSD)	1.05707	-.76878
Zscore(BIKDM)	-.76200	.55418
Zscore(BIP)	-.26599	.19345
Zscore(PPRPGG)	-.78248	.56908
Zscore(PPRPS)	-.98447	.71598
Zscore(PPRSDK)	-.83880	.61004

Tabel 8.14 *Final Cluster Centers.*

<i>Final Cluster Centers</i>		
	<i>Cluster</i>	
	1	2
Zscore(PKP)	.92391	-.67194
Zscore(PPTTSD)	1.05707	-.76878
Zscore(BIKDM)	-.76200	.55418
Zscore(BIP)	-.26599	.19345
Zscore(PPRPGG)	-.78248	.56908
Zscore(PPRPS)	-.98447	.71598
Zscore(PPRSDK)	-.83880	.61004
Zscore(PKP)	.92391	-.67194
Zscore(PPTTSD)	1.05707	-.76878

Zscore(BIKDM)	-.76200	.55418
Zscore(BIP)	-.26599	.19345
Zscore(PPRPGG)	-.78248	.56908
Zscore(PPRPS)	-.98447	.71598
Zscore(PPRSDK)	-.83880	.61004

Tabel 8.15 Number of cases in each cluster.

Number of Cases in Each Cluster		
Klaster	1	8.000
	2	11.000
Valid		19.000
Missing		.000

Seperti yang telah dijelaskan sebelumnya, metode nonhierarki adalah metode kluster dengan jumlah klasternya diketahui. Pada kasus ini, jumlah kluster ditentukan sebanyak 2 kluster. Dari *output* di atas, terlihat bahwa banyaknya objek di kluster pertama sebanyak 8 kabupaten, sedangkan sisanya (11 kabupaten) dikelompokkan ke kluster 2.

Tabel 8.16 Cluster Membership.

Case	Cluster Membership			
	5 Cluster	4 Cluster	3 Cluster	2 Cluster
1: Banyuwangi	1	1	1	1
2: Cianjur	1	1	1	1
3: Jember	2	2	2	2
4: Jepara	3	3	3	2
5: Jombang	1	1	1	1
6: Kediri	2	2	2	2
7: Kudus	2	2	2	2
8: Kulonprogo	3	3	3	2
9: Lumajang	1	1	1	1

10: Majalengka	4	1	1	1
11: Pacitan	1	1	1	1
12: Pamekasan	1	1	1	1
13: Pasuruan	5	4	3	2
14: Pati	3	3	3	2
15: Probolinggo	3	3	3	2
16: Rembang	3	3	3	2
17: Serang	3	3	3	2
18: Sleman	2	2	2	2
19: Sukabumi	1	1	1	1

Dari hasil di atas dapat dilihat bahwa

1. Klaster 1: Banyuwangi, Cianjur, Jombang, Lumajang, Majalengka, Pacitan, Pamekasan, dan Sukabumi.
2. Klaster 2: Jember, Jepara, Kediri, Kudus, Kulonprogo, Pasuruan, Pati, Probolinggo, Rembang, Serang, dan Sleman.

Daftar Rujukan

- Agusta, Y. (2007). K-Means-Penerapan, Permasalahan dan Metode Terkait. *Jurnal Sistem dan Informatika*, 3, 47–60.
- Larasati, A. (2017). *Metode K – Medoids pada Data dengan Pencilan* (Skripsi, Fakultas Matematika dan Ilmu Pengetahuan, Universitas Gadjah Mada, Yogyakarta).
- Machfudhoh, S. dan Wahyuningsih, N. (2013). Analisis Kluster Kabupaten/Kota Berdasarkan Pertumbuhan Ekonomi Jawa Timur. *Jurnal Sains dan Seni POMITS*, 2(1).
- Nasari, F. (2015). *Penerapan K-Means Klustering Pada Data Penerimaan Mahasiswa Baru (Studi Kasus: Universitas Potensi Utama)*. Seminar Nasional Teknologi Informasi dan Multimedia, 74.
- Safe'i, A. I. (2018). *Aplikasi K-Means untuk Pengelompokan Kabupaten dan Kota Berdasarkan Produktivitas Tanaman Pangan di Provinsi Jawa Tengah Tahun 2015* (Tugas Akhir).
- Santoso, S. (2014). *Statistik Multivariat Edisi Revisi konsep dan Aplikasi dengan SPSS*. Jakarta: PT Elex Media Komputindo.
- Supranto, J. (2004). *Analisis Multivariat Arti & Interpretasi*. Jakarta: PT RINEKA CIPTA.
- Triyanto, Wiwit Agus. 2015. Algoritma K-Medoids untuk Penentuan Strategi Pemasaran Produk. *Jurnal SIMETRIS* 6(1), 183.

BAB 9

MULTIDIMENSIONAL SCALLING

A. Definisi

Ada beberapa definisi penyekalaan dimensi ganda (*multidimensional scaling*) yang diungkapkan oleh beberapa ahli, salah satunya adalah penyekalaan multidimensional (PMD) atau *multidimensional scaling* (MDS), yang merupakan suatu teknik yang bisa membantu peneliti untuk mengenali (mengidentifikasi) dimensi kunci yang mendasari evaluasi objek dari responden atau pelanggan (Supranto, 2004). Analisis *multidimensional scaling* (MDS) merupakan salah satu teknik peubah ganda yang dapat digunakan untuk menentukan posisi suatu objek lainnya berdasarkan penilaian kemiripannya. MDS berhubungan dengan pembuatan map untuk menggambarkan posisi sebuah objek dengan objek lainnya berdasarkan kemiripan objek-objek tersebut (Walundungo, 2014)

Dari definisi tersebut, *multidimensional scaling* berguna untuk menyajikan objek-objek secara visual berdasarkan kemiripan yang dimiliki. Selain itu, kegunaan lain dari teknik ini adalah mengelompokkan objek-objek yang memiliki kemiripan dilihat dari beberapa peubah atau atribut yang dianggap mampu mengelompokkan objek-objek tersebut (Masuku, dkk., 2014).

Oleh karena itu, dapat disimpulkan bahwa, *multidimensional scaling* adalah 1) kumpulan teknik-teknik statistika untuk menganalisis kemiripan dan ketidakmiripan antarobjek; 2) memberikan hasil yang berupa plot titik-titik sehingga jarak antartitik menggambarkan tingkat kemiripan atau ketidakmiripan; dan c) memberikan petunjuk untuk mengidentifikasi atribut

tidak diketahui atau faktor yang mempengaruhi munculnya kemiripan atau ketidakmiripan (Timm, 2002).

Beberapa istilah (terminologi) statistik yang penting dalam analisis MDS, antara lain sebagai berikut.

1. Analisis agregat (*aggregate analysis*), sebuah pendekatan dalam MDS tempat *perceptual map* dibuat untuk evaluasi sekelompok responden terhadap objek-objek. *Perceptual map* dapat dibuat dengan komputer maupun peneliti sendiri.
2. Penilaian kesamaan (*similarity judgement*), merupakan perangkat seluruh pasangan merek yang mungkin atau stimulan lain berdasarkan kesamaan yang dinyatakan melalui skala pengukuran (*measurement scale*) berskala nonmetrik atau semacamnya.
3. Peringkat preferensi (*preference rankings*), adalah ranking berupa urutan merek-merek mulai dari yang paling diinginkan sampai paling tidak diinginkan konsumen atau responden.
4. Stress, adalah skor yang menyatakan ketidaktepatan pengukuran (*lack offit measurement*). Makin tinggi Stress, makin tinggi ketidaktepatan.
5. R kuadrat (R-Squared), adalah indeks korelasi pangkat dua yang menyatakan proporsi varian data asli yang dapat dijelaskan MDS.
6. Peta spasial (disebut juga *perceptual map*) adalah suatu peta geometris yang menyatakan hubungan atau perbandingan antar merek atau stimulan lain berdasarkan dimensi-dimensi yang diukur.
7. Koordinat (*coordinates*), menyatakan posisi suatu merek atau stimulus lain dalam peta spasial.
8. *Unfolding*, representasi merek dan responden sebagai pola dalam ruang yang sama (Simamora, 2005).

Konsep dasar dari *multidimensional scaling* adalah jarak yang dihasilkan dalam ruang harus sesuai dengan *proximities* yang sebenarnya sehingga yang dilakukan oleh *multidimensional scaling* adalah mencari posisi dalam ruang dan koordinat untuk setiap stimulan. Diharapkan, jarak yang dihasilkan akan mendekati nilai *proximities*. *Proximity* dibagi atas dua macam, yaitu *similarity* (kemiripan), yakni jika makin kecil nilai jaraknya menunjukkan bahwa objeknya lebih mirip dan *dissimilarity* (ketidakmiripan), yakni jika makin besar nilai jaraknya menunjukkan bahwa objeknya makin tak mirip. Keberhasilan dari proses ini ditentukan oleh seberapa baik jarak yang dihasilkan (d_{ij}) dalam ruang sesuai dengan *proximities* yang sebenarnya (δ_{ij}) (Demaine et al., 2021).

B. Prinsip Dasar *Multidimensional Scalling* (MDS)

Analisis *multidimensional scalling* merupakan salah satu teknik peubah ganda yang dapat digunakan untuk menentukan posisi suatu objek lainnya berdasarkan penilaian kemiripannya. MDS disebut juga *perceptual map*. MDS berhubungan dengan pembuatan *map* untuk menggambarkan posisi sebuah objek dengan objek lainnya berdasarkan kemiripan objek-objek tersebut. MDS juga merupakan teknik yang bisa membantu peneliti untuk mengenali (mengidentifikasi) dimensi kunci yang mendasari evaluasi objek dari responden (pelanggan).

Konsep dan ruang lingkup penyekalaan multidimensional (*multidimensional scaling*/MDS) dalam riset pemasaran dan menguraikan berbagai aplikasinya; menguraikan langkah-langkah yang harus dilalui pada penyekalaan multidimensional tentang data persepsi, meliputi perumusan masalah, mendapatkan data *input*, memilih prosedur MDS, memutuskan banyaknya dimensi, memberikan interpretasi kepada konfigurasi

(*configuration*) dan memberikan penilaian (*to asses*) keandalan dan kesahihan (*reability and validity*), serta menjelaskan penyekalaan data preferensi; menjelaskan analisis korespondensi dan kebaikan serta kelemahannya; memahami hubungan antar-MDS, analisis diskriminan, dan analisis faktor.

MDS dapat menentukan:

1. Dimensi apa yang digunakan oleh responden ketika mengevaluasi objek.
2. Berapa dimensi yang akan digunakan untuk masalah yang dihadapi (sedang diteliti).
3. Kepentingan relatif dari setiap dimensi.

Bagaimana objek dikaitkan atau dihubungkan secara perseptual? Dua teknik yang terkait untuk menganalisis persepsi dan preferensi pelanggan ialah analisis penyekalaan multidimensional dan analisis konjoin (*multidimensional scaling and conjoint analysis*). Namun, pada bab ini akan ditunjukkan langkah-langkah yang diperlukan untuk melakukan analisis penyekalaan dimensional. MDS digunakan untuk mengetahui hubungan interdependensi atau saling ketergantungan antar variabel atau data. Hubungan ini tidak diketahui melalui reduksi ataupun pengelompokan variabel, melainkan dengan membandingkan variabel yang ada pada setiap objek yang bersangkutan dengan menggunakan *perceptual map*. Konsep dasar MDS adalah pemetaan. Analisis penyekalaan multidimensional ialah suatu kelas prosedur untuk menyajikan persepsi dan preferensi pelanggan secara spasial dengan menggunakan tayangan yang bisa dilihat (*a visual display*). Persepsi atau hubungan antara stimulus secara psikologis ditunjukkan sebagai hubungan geografis antara titik-titik di dalam suatu ruang multidimensional. Sumbu dari peta spasial diasumsikan menunjukkan dasar psikologis (*psychological basis*) atau dimensi yang mendasari (*underlying dimensions*)

yang digunakan oleh pelanggan/responden untuk membentuk persepsi dan preferensi untuk stimulus. Analisis penyekalan multidimensional digunakan pada pemasaran untuk mengenali hal-hal berikut.

1. Banyaknya dimensi dan sifat/cirinya yang digunakan untuk memersepsikan merek yang berbeda di pasar.
2. Penempatan (*positioning*) merek yang diteliti dalam dimensi ini.
3. Penempatan merek ideal dari pelanggan dalam dimensi ini.

Informasi sebagai hasil analisis penyekalaan multidimensional (Gudono, 2016) telah digunakan untuk berbagai aplikasi dalam pemasaran, antara lain sebagai berikut.

1. Ukuran citra (*image measurement*). Membandingkan persepsi pelanggan dan bukan pelanggan dari perusahaan dengan persepsi perusahaan sendiri.
2. Segmentasi pasar (*market segmentation*).
3. Pengembangan produk baru (*new product development*). Melihat adanya celah (*gap*) dalam peta spasial yang menunjukkan adanya peluang untuk penempatan produk baru. Juga untuk mengevaluasi konsep produk baru dan merek yang sudah ada *on a test basis* untuk menentukan bagaimana pelanggan memersepsikan/memahami konsep baru. Proporsi preferensi untuk setiap produk baru merupakan satu indikator keberhasilannya (maksudnya satu jenis produk tertentu banyak yang menyenangkannya/menggemarinya daripada produk lainnya).
4. Menilai keefektifan iklan (*assessing advertising effectiveness*). Peta spasial bisa digunakan untuk menentukan apakah iklan/advertensi telah berhasil dalam mencapai penempatan merek yang diinginkan (misalnya dari posisi nomor 3 ke nomor 2 atau dari nomor 2 ke nomor 1).

5. Analisis harga (*pricing analysis*) Peta spasial dikembangkan dengan dan tanpa informasi harga dapat dibandingkan untuk menentukan dampak yang ditimbulkan harga.
6. Keputusan saluran (*channel decisions*). Pertimbangan pada kecocokan (*compatibility*) dari merek toko dengan eceran yang berbeda dapat mengarah ke peta spasial yang berguna untuk keputusan saluran.
7. Pembentukan skala sikap (*attitude scale construction*). Teknik penyekalaan multidimensional dapat digunakan untuk mengembangkan *the appropriate dimensionality and configuration of the attitude space*.

C. Tujuan Penggunaan *Multidimensional Scalling* (MDS)

Terdapat dua macam tujuan dalam analisis data menggunakan *multidimensional scaling* (MDS) menurut Grudono (2016), di antaranya adalah

1. Untuk mengidentifikasi dimensi yang tidak diketahui yang mendasari perilaku atau fitur objek yang terukur atau tampak di permukaan. Dalam konteks ini, penggunaan MDS lebih condong untuk riset eksploratoris.
2. Untuk mendapatkan ukuran pembanding antarbeberapa objek manakala dasar untuk membandingkan objek-objek tersebut belum ada atau tidak diketahui.

MDS tidak mensyaratkan jenis data apa yang dapat digunakan untuk MDS ataupun bentuk distribusi data yang diolah harus tertentu. Namun, penggunaan MDS menurut Grudono (2016) biasanya mensyaratkan beberapa hal sebagai berikut.

1. Model telah dispesifikasi dengan tepat. Maksudnya adalah bahwa semua objek (variabel) yang relevan telah dimasukkan untuk diperbandingkan dengan MDS.

2. Telah menggunakan level pengukuran yang tepat, misalnya spesifikasi skala yang digunakan—apakah rasio, interval, atau ordinal—harus sesuai dengan teknik MDS yang digunakan. Misalnya, untuk MDS matriks sebaiknya menggunakan skala rasio atau interval.
3. Jumlah objek paling tidak sebanyak dimensi. Jika jumlah objek kurang dari jumlah dimensi maka MDS tidak akan stabil, jika jumlah objek sedikit lebih banyak jumlah dimensi maka R-Squared akan terinflasi. Jumlah objek paling tidak empat kali jumlah dimensi ditambah.
4. Skala yang digunakan setara dan jika tidak setara maka ukuran yang digunakan sebaiknya adalah ukuran yang distandardisasi (*standardized value*). Misalnya, jika variabel yang satu adalah ribuan rupiah, sedangkan yang lainnya adalah Kg berat tubuh manusia maka akan menyebabkan ketidaksetaraan skala ukuran.
5. Komparabilitas, objek yang diperbandingkan seharusnya memiliki kesamaan tertentu yang cukup berarti sehingga pantas diperbandingkan.
6. Ukuran variabel yang banyak memang tidak disyaratkan tetapi ukuran objek (variabel) minimum adalah empat.

D. Jenis-Jenis Teknik *Multidimensional Scalling* (MDS)

Pada *multidimensional scaling* (MDS), teknik algoritma untuk menemukan koordinat titik-titik data dalam ruang berdimensi p yang ditentukan berdasarkan data mengenai kedekatan (*proximity*) jarak (*distance*) titik yang satu dengan titik yang lain. Ada beberapa jenis algoritma MDS dan karena itu, jenis MDS dapat diklasifikasikan ke dalam beberapa kategori.

Pertama, MDS dapat dibedakan menjadi matriks MDS dan nonmatriks MDS. MDS matriks menggunakan nilai matriks. Tujuan MDS matriks adalah mendapat konfigurasi titik-titik data dalam ruang

multidimensi yang kedekatan (*proximity*) jaraknya menunjukkan kesamaan dengan data observasi. Jika kedekatan jarak tersebut diukur dengan jarak euclides (*euclidean distance*) maka teknik tersebut disebut *the classical matrix MDS* (Torgeson, 1965). Tujuan MDS nonmatriks adalah untuk menetapkan hubungan nonmonotonik antara jarak antar titik dengan kesamaan yang diobservasi. Keunggulan MDS nonmatriks adalah bahwa teknik ini tidak membutuhkan asumsi mengenai fungsi transformasi yang mendasarinya. Satu-satunya asumsi yang diperlukan hanyalah bahwa data yang diolah merupakan data rangking (*ordinal*). Kruskal (1964) mengusulkan ukuran untuk menggambarkan penyimpangan dari kondisi *monotonicity* antara jarak taksiran dengan jarak yang diobservasi yang disebut *Stress*.

Kedua, MDS juga dapat diklasifikasikan sebagai *weighted MDS*, *replicated MDS*, dan *single matrix MDS*. Dalam *replicated MDS* beberapa matriks kesamaan (*similarity*) dianalisis secara simultan. Matriks data tersebut diambil dari beberapa subjek atau dari satu subjek yang diambil beberapa kali serta satu solusi skala (*scaling solution*) yang merepresentasikan data matriks kesamaan tersebut. Dalam *Weighted MDS* dimensi-dimensi dalam solusi skala dapat diberi bobot berbeda (*weighted differently*) untuk memodelkan perbedaan sensitivitas untuk masing-masing dimensi.

Ketiga, algoritma MDS dapat juga dibedakan menjadi: (a) *deterministic MDS* dan (b) *probabilistic MDS*. Pada MDS deterministik, masing-masing objek direpresentasikan oleh nilai tunggal dalam ruang multidimensi (*multidimensional scaling*). Sementara itu, dalam MDS probabilistik masing-masing objek dinyatakan sebagai distribusi probabilitas dalam ruang multidimensi. Teknik MDS probabilistik bermanfaat manakala gambaran mengenai objek dianggap *noisy*.

Tipe data berdasarkan skala pengukuran dibagi menjadi 4 tipe, yaitu skala nominal, ordinal, interval, dan rasio. Berdasarkan tipe data tersebut, *multidimensional scaling* dibagi menjadi 2 jenis, yaitu penyekalaan berdimensi ganda metrik dan penyekalaan berdimensi ganda nonmetrik (Ghojogh et al., 2020).

1. *Multidimensional scaling* (MDS) Metrik

Data jarak yang digunakan dalam *multidimensional scaling* metrik adalah data rasio atau interval. MDS metrik (*classical scaling*) digunakan untuk menemukan himpunan titik dalam ruang dimensi n dan masing-masing titik mewakili satu objek. Dalam *classical scaling*, *dissimilarities* (δ_{ij}) diperlakukan sama dengan jarak (d_{ij}), yaitu $\delta_{ij} = d_{ij}$.

Tujuan dari *multidimensional scaling* adalah untuk mencari koordinat titik-titik dalam ruang euclides dari matriks jarak yang tersedia. Misalkan koordinat n titik dalam ruang euclides dimensi p adalah \mathbf{x}_i ($i = 1, 2, \dots, n$) dengan $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{ik})$ dan $\mathbf{x}_j = (x_{j1}, x_{j2}, \dots, x_{jk})$. Jarak euclides antara titik ke- i dan ke- j adalah (Mattjik dkk., 2011).

$$d_{ij}^2 = \sum_{k=1}^p (x_{ik} - x_{jk})^2 \dots \dots \dots (1)$$

Langkah berikutnya adalah menentukan matriks hasil kali dalam \mathbf{B} , dengan cara mendekomposisikan matriks \mathbf{D} melalui proses *double centering*. Matriks \mathbf{B} memiliki elemen-elemen:

$$b_{ij} = -\frac{1}{2} (d_{ij}^2 - d_i^2 - d_j^2 - d_{..}^2) \dots \dots \dots (2)$$

Dengan,

$$d_i^2 = \frac{1}{n} \sum_{j=1}^n d_{ij}^2$$

$$d_j^2 = \frac{1}{n} \sum_{i=1}^n d_{ij}^2$$

$$d_{..}^2 = \frac{1}{n} \sum_{i,j=1}^n d_{ij}^2$$

Bila ditulis dalam bentuk matriks menjadi:

$$\mathbf{B} = \frac{1}{2} \left(\mathbf{I} - \frac{1}{n} \mathbf{V} \right) \mathbf{D}^2 \left(\mathbf{I} - \frac{1}{n} \mathbf{V} \right) \dots \dots \dots (3)$$

Keterangan:

\mathbf{I} : matriks identitas dengan ukuran $n \times n$;

\mathbf{V} : matriks berukuran $n \times n$ dengan entri $V_{ij} = 1$ untuk semua i, j ;

\mathbf{D}^2 : matriks kuadrat jarak berukuran $n \times n$ dengan elemen d_{ij}^2

(Mattjik dkk., 2011).

Matriks hasil kali dalam \mathbf{B} dapat juga diekspresikan sebagai $\mathbf{B} = \mathbf{X}\mathbf{X}^T$, dengan $\mathbf{X} = [x_1, \dots, x_n]^T$ adalah matriks koordinat yang berukuran $(n \times p)$.

Rank dari matriks \mathbf{B} , $r(\mathbf{B})$ adalah

$$r(\mathbf{B}) = r(\mathbf{X}\mathbf{X}^T) = r(\mathbf{X}) = p \dots \dots \dots (4)$$

Sekarang, \mathbf{B} adalah matriks yang simetris, semi definit positif dan berpangkat (*rank*) p , sehingga memiliki p *eigen values* nonnegatif dan $n - p$ *eigen values* sama dengan 0.

Matriks \mathbf{B} kemudian ditulis dalam bentuk dekomposisi spektral:

$$\mathbf{B} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^T \dots\dots\dots(5)$$

Keterangan:

$\mathbf{\Lambda} = \text{diag} (\lambda_1, \lambda_2, \dots, \lambda_n)$, yaitu matriks diagonal dari *eigen values* $\{\lambda_i\}$ matriks \mathbf{B}

$\mathbf{V} = [\gamma_1, \dots, \gamma_n]$, yaitu matriks dengan *eigen vectors* yang dinormalkan.

Dengan demikian, koordinat matriks \mathbf{X} berisi titik konfigurasi dalam R^p diberikan oleh

$$\mathbf{X} = \mathbf{V}\mathbf{\Lambda}^{\frac{1}{2}} \dots\dots\dots(6)$$

Permasalahan berikutnya adalah menentukan jumlah dimensi yang diperlukan untuk menampilkan koefisien ketidakmiripan (δ_{ij}). Jika \mathbf{B} adalah matriks semi definit positif maka jumlah akar ciri yang tak nol menunjukkan jumlah dimensi yang diperlukan. Jika \mathbf{B} bukan matriks semi definit positif maka jumlah akar ciri yang positif menunjukkan jumlah dimensi yang tepat. Jumlah dimensi tersebut merupakan jumlah dimensi maksimal yang diperlukan, sedangkan untuk lebih praktisnya lebih baik memilih dimensi yang lebih kecil. Secara singkat algoritma *classical scaling* sebagai berikut (Mattjik dkk., 2011).

- a. Menentukan koefisien ketidakmiripan (δ_n)
- b. Mencari matriks $\mathbf{A} = \left[-\frac{1}{2} \delta_{ij}^2 \right]$
- c. Mencari matriks $\mathbf{B} = [a_{ij} - a_i - a_j + a]$
- d. Mencari *eigenvalues* $\lambda_1, \dots, \lambda_{n-1}$ dan *eigenvector* v_1, \dots, v_{n-1} yang kemudian dinormalkan sehingga $v_i^T v_i = \gamma_i \lambda_i$. Jika \mathbf{B} tidak semi definit positif (beberapa *eigen values* bernilai negatif). Maka terdapat 2 pilihan,

pilihan ke-1 adalah membuang *eigen values* yang bernilai negatif dan melanjutkan proses. Pilihan ke-2 adalah menambahkan suatu konstanta c pada koefisien ketidakmiripan sebagai berikut $\delta_{ij} = \delta_{ij} + c(1 - \delta_{ij}^2)$ dan kembali ke langkah 2.

- e. Memilih jumlah dimensi yang tepat. Dapat menggunakan:

$$\frac{\sum_{i=1}^p \lambda_i}{\sum(\text{eigenvalues positif})}$$

- f. Menentukan koordinat n titik pada ruang euclides dimensi p dengan:

g. $x_{ij} = v_{ij}(i = 1, \dots, n; j = 1, \dots, p)$

2. *Multidimensional scaling (MDS) Nonmetrik*

Multidimensional scaling nonmetrik mengasumsikan bahwa datanya adalah kualitatif (nominal dan ordinal). Program MDS nonmetrik menggunakan transformasi monoton (sama) ke data yang sebenarnya sehingga dapat dilakukan operasi aritmatika terhadap nilai ketidaksamaannya, untuk menyesuaikan jarak dengan nilai urutan ketidaksamaannya. Hasil perubahan ini disebut *disparities*. *Disparities* digunakan untuk mengukur tingkat ketidaktepatan konfigurasi objek-objek dalam peta berdimensi tertentu dengan *input* data ketidaksamaannya. Pendekatan yang sering digunakan saat ini untuk mencapai hasil yang optimal dari skala nonmetrik digunakan *kruskal's least-square monotonic transformation*, *disparities* ini merupakan nilai rata-rata dari jarak-jarak yang tidak sesuai dengan urutan ketidaksamaannya. Informasi ordinal kemudian dapat diolah dengan MDS nonmetrik sehingga menghasilkan konfigurasi dari objek-objek yang terdapat pada dimensi tertentu dan kemudian agar jarak antara objek sedekat mungkin dengan *input* nilai ketidaksamaan atau

kesamaannya. Koordinat awal dari setiap subjek dapat diperoleh melalui cara yang sama seperti metode MDS metrik dengan asumsi bahwa meskipun data bukan jarak informasi yang sebenarnya tapi nilai urutan tersebut dipandang sebagai variabel interval (Irmawati, 2017).

3. Asumsi dalam *Multidimensional Scaling* (MDS)

Multidimensional scaling tidak memiliki asumsi yang baku dalam metodologinya, tipe data, atau hubungan antarvariabel-variabelnya. MDS hanya mensyaratkan bahwa peneliti menerima beberapa prinsip mengenai persepsi yang meliputi

- a. Pembatasan *multidimensional scaling*. Diasumsikan bahwa kemiripan stimulus A dan B sama dengan stimulus B ke A. Akan tetapi, ada beberapa kasus yang asumsi ini tidak berlaku.
- b. Variasi dimensi. Tiap-tiap responden tidak akan menilai suatu stimulus dalam dimensi yang sama. Misalnya, seseorang mungkin akan menilai sebuah mobil dari tenaga dan modelnya, sedangkan yang lain tidak memperhatikan faktor ini melainkan sebuah mobil dari harga dan kenyamanannya.
- c. Variasi kepentingan. Responden tidak menilai kepentingan dimensi pada tingkat yang sama, walaupun seluruh responden menilai dimensi tersebut. Misalnya, dua orang responden menilai suatu minuman ringan dari tingkat karbonasinya. Seorang responden mungkin akan menilai bahwa faktor ini tidaklah penting, sedangkan yang lainnya menilai bahwa faktor ini penting.
- d. Variasi waktu. Pernyataan yang didapat dari stimulus-stimulus tidak bisa digunakan dalam jangka waktu yang lama. Dengan kata lain, peneliti

tidak dapat mengharapkan persepsi yang stabil dari waktu ke waktu (Lestari, 2017).

E. Validasi Model MDS

Untuk mendapatkan model MDS yang cocok, terdapat beberapa kriteria atau pedoman agar hasil yang didapatkan layak dan dapat digunakan untuk interpretasi sesungguhnya, yaitu

1. Nilai Stress (*Standardized Residual Sum of Square*)

Stress ialah ukuran ketidakcocokan (*a lack of fit measure*). Makin tinggi nilai Stress makin tidak cocok, sehingga dapat disimpulkan kalau data tidak cocok digunakan untuk proses analisis *multidimensional scaling*. Untuk menemukan tingkat ketidakcocokan tersebut, kita dapat mencari dengan menggunakan rumus Stress. Stress dapat dicari dengan menggunakan rumus sebagai berikut.

$$Stress = \sqrt{\frac{\sum_{i,j}^n (d_{ij} - \hat{d}_{ij})^2}{\sum_{i,j}^n (d_{ij} - \bar{d})^2}}$$

Keterangan:

d_{ij} = jarak kemiripan sesungguhnya

\hat{d}_{ij} = jarak yang dihasilkan dari kemiripan data

\bar{d} = jarak rata-rata pada peta $\left(\frac{\sum_{i,j}^n d_{ij}}{n}\right)$

Untuk *Kruskal* Stress, formula terdapat pedoman untuk mengidentifikasi model yang baik bila dilihat dari nilai Stress dengan menggunakan standar kriteria sebagai berikut (Wardhana & Sutanto, 2009).

Tabel 9.1 Kriteria Nilai Stress.

Stress (%)	Kriteria Model <i>Multidimensional Scaling</i>
>20 %	Jelek
20-10 %	Cukup
10-5 %	Baik
5-2,5 %	Sangat Baik
< 2,5 %	Sempurna

2. Nilai R^2

$R^2 = R$ kuadrat (R-Squared) ialah kuadrat dari koefisien korelasi yang menunjukkan proporsi varian dari skala optimal data, yang disumbangkan oleh prosedur penyekalaan multidimensional ukuran kecocokan/ketepatan (*goodness of fit measure*) (Simamora, 2005). Yang diinginkan ialah nilai R^2 yang tinggi ($R^2=1$ atau 100% model mewakili dengan sempurna), tetapi, $R^2 \geq 0.60$ (60% atau lebih) sudah bisa diterima. Artinya, bisa mewakili data *input* dengan cukup baik. R^2 dapat dicari dengan menggunakan rumus sebagai berikut (Supranto, 2004).

$$R^2 = 1 - \frac{\sum_{i,j}^n (d_{ij} - \hat{d}_{ij})^2}{\sum_{i,j}^n (d_{ij} - \bar{d})^2}$$

3. *Eigenvalue dan Eigenvector*

Vektor kolom X merupakan *eigenvector* matriks A dan adalah *eigenvalue* atau sering disebut juga *characteristic value*. Jika A adalah sebuah matriks bujur sangkar berukuran n dan X adalah suatu vektor kolom, persamaan:

$$AX = \lambda X$$

Untuk mencari nilai *eigen* matriks A yang berukuran $n \times n$, dari persamaan di atas dapat ditulis kembali sebagai suatu persamaan homogen:

$$(A - \lambda I)X = 0$$

Dengan I adalah matriks identitas yang berordo sama dengan matriks A.

$$\left[\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} - \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix} \right] \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$\begin{bmatrix} (a_{11} - \lambda_1)x_1 + a_{12}x_2 + \dots + a_{1n}x_n & = 0 \\ a_{21}x_1 + (a_{22} - \lambda_2)x_2 + \dots + a_{2n}x_n & = 0 \\ \vdots & \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + (a_{nn} - \lambda_n)x_n & = 0 \end{bmatrix}$$

Penyelesaian tersebut akan mempunyai persamaan tak trivial ini hanya jika determinan dari matriks koefisien adalah nol.

$$\begin{vmatrix} a_{11} - \lambda_1 & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda_2 & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda_n \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{vmatrix}$$

Yang dapat ditulis sebagai:

$$\text{Det}(A - \lambda I)X = 0$$

Merupakan suatu suku banyak berderajat n dalam. Akar dari persamaan suku banyak ini disebut *eigenvalue* (nilai eigen) dari atau nilai karakteristik dari matriks A. Untuk setiap *eigenvalue* (nilai eigen) akan ada penyelesaian $X \neq 0$ yang merupakan suatu penyelesaian tak trivial yang dinamakan *eigenvector* (vector eigen) atau vektor karakteristik dari nilai eigen-nya.

4. Kemiripan (*Similarity*)

Dalam beberapa metode yang berkaitan dengan kemiripan (*similarity*), jarak dimaksudkan sebagai ukuran kemiripan. Ukuran kemiripan ditentukan berdasarkan jarak (*distance*) antar titik. Ukuran jarak dalam bidang dua dimensi dapat ditentukan dengan menggunakan jarak euclidean (*euclidean distance*) adalah perhitungan jarak dari dua buah titik dalam *euclidean space*. *Euclidean space* diperkenalkan oleh Euclid, seorang matematikawan dari Yunani sekitar tahun 300 SM untuk mempelajari hubungan sudut dan jarak. Euclides ini berkaitan dengan teori Phytagoras.

Untuk menghitung nilai kedekatan jarak antar objek pada peta persepsi dapat diperoleh dengan menggunakan jarak euclides antara objek ke-i dengan objek ke-j.

$$d_{ij} = \sqrt{\sum_{k=1}^n (x_{ik} - x_{jk})^2}$$

Keterangan:

d_{ij} = jarak antar objek ke-i dan objek ke-j

x_{ik} = hasil pengukuran objek ke-i pada peubah/atribut k

x_{jk} = Hasil pengukuran objek ke-j pada peubah/atribut k

5. *Perceptual Map*

Hair dkk. (2009) mendefinisikan peta persepsi adalah sebuah representasi visual dari persepsi seorang responden terhadap beberapa objek pada dua atau lebih dimensi. Tiap objek akan memiliki posisi spasial pada peta persepsi tersebut yang merefleksikan kesamaan atau preferensi (*preference*) ke objek lain dengan melihat dimensi-dimensi pada peta persepsi.

Perceptual map juga sering disebut peta spasial (*spatial map*). Peta spasial (*spatial map*) ialah hubungan antara merek atau stimulus lain yang dipersepsikan, dinyatakan sebagai hubungan geometris antara titik-titik di alam ruang yang *multidimensional* koordinat (*coordinates*), menunjukkan posisi (letak) suatu merek atau suatu stimulus dalam suatu peta spasial (Supranto, 2010).

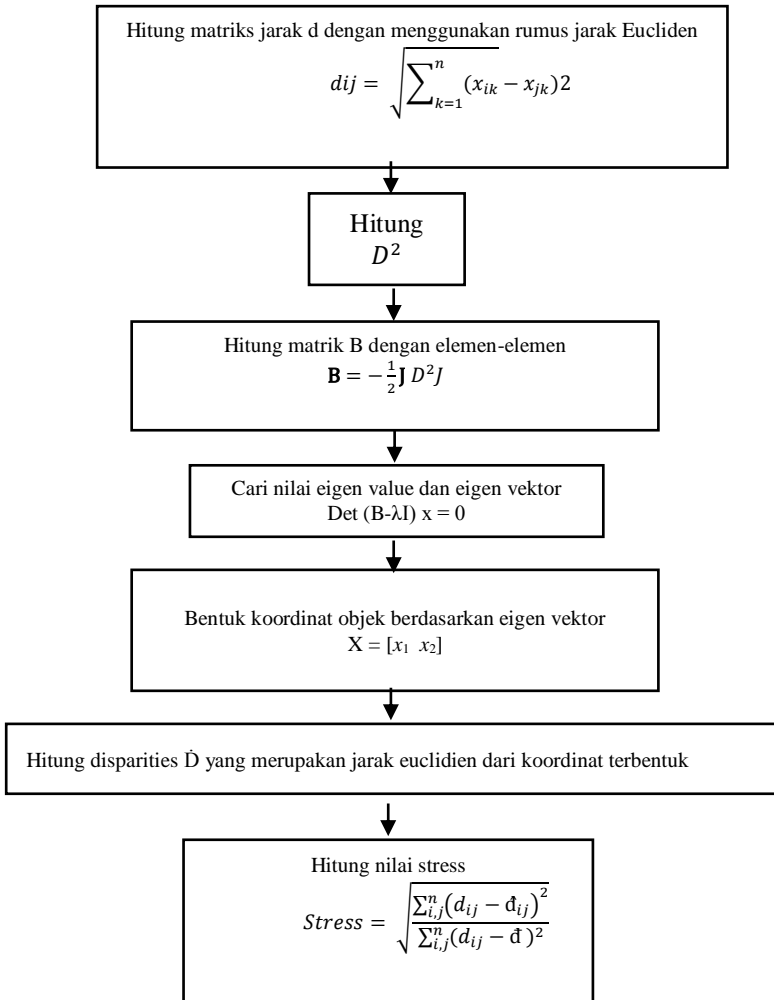
Untuk memperoleh peta persepsi, maka harus diperoleh stimulus koordinat. Algoritma *multidimensional scaling* fokus pada fakta bahwa koordinat matriks X dapat diperoleh dengan dekomposisi *eigenvalue* dari produk skalar matriks $B = XX'$. Masalah dalam mengontribusi D dari matriks *proximity* P diselesaikan dengan mengalikan kuadrat dari matriks *proximity* dengan matriks $J = I - 1/n A$. prosedur ini dinamakan *double centering*.

6. *Potitioning*

Menurut Kotler (1988), *positioning* adalah suatu tindakan merancang nilai dan kesan yang ditawarkan perusahaan sehingga segmen pelanggan memahami dan mengapresiasi apa yang dilakukan perusahaan dalam kaitan dengan para pesaingnya.

Positioning didefinisikan sebagai seni dan ilmu pengetahuan dalam mencocokkan produk atau jasa dengan satu atau lebih segmen pasar dalam rangka menetapkan bagian yang berarti dari produk atau jasa tersebut dari persaingan. *Positioning* juga merupakan upaya untuk mendesain produk agar menempati sebuah posisi yang unik di benak pelanggan. *Positioning* akan menjadi penting apabila persaingan sudah sangat sengit di segala bidang.

7. *Prosedur Perhitungan Multidimensional Scalling (MDS)*



F. Contoh Kasus Menggunakan Analisis MDS

Contoh analisis MDS dapat diterapkan pada penelitian dengan judul: *Penerapan Metode Multidimensional Scaling dalam Positioning Pemilihan Kampus di Malang Raya Berdasarkan Persepsi Calon Mahasiswa*. Tujuan penelitian ini untuk mengetahui kemiripan beberapa kampus yang ada di Malang Raya berdasarkan penilaian calon mahasiswa. Berikut data rata-rata penilaian dari 80 responden.

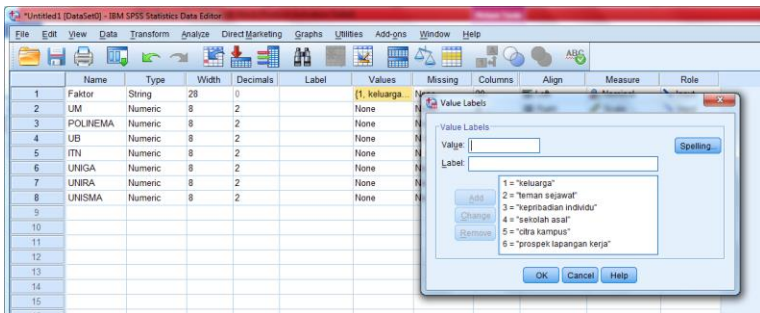
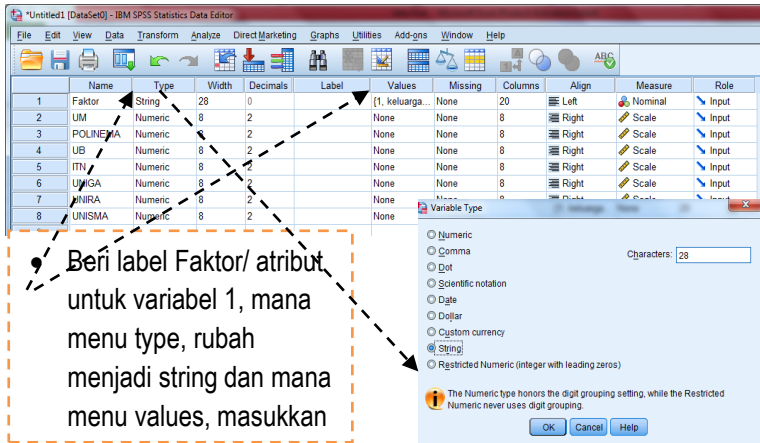
Tabel 9.2 Nilai Rata-Rata.

Faktor	Keluarga/ Orang Tua	Teman Sejawat	Kepribadian Individu	Sekolah Asal	Citra Kampus	Prospek Lapangan Kerja
UM	3,21	3,3	3,8	4	3,4	3,5
POLINEMA	3,2	3,4	4,1	4,6	4,5	4,4
UB	3,1	3	3,9	4,3	3,2	3,33
ITN	3	4	2,6	3,8	3,7	3,8
UNIGA	4,2	3,3	4,3	4,5	4,2	4
UNIRA	3,2	2,8	3,9	3,8	3,3	2,8
UNISMA	2,9	2,9	3,5	3,94	2,9	3,5

G. Analisis MDS dalam SPSS

1. Menggunakan ALSCAL

- a. Langkah pertama, meng-*input* nilai ke dalam program SPSS.

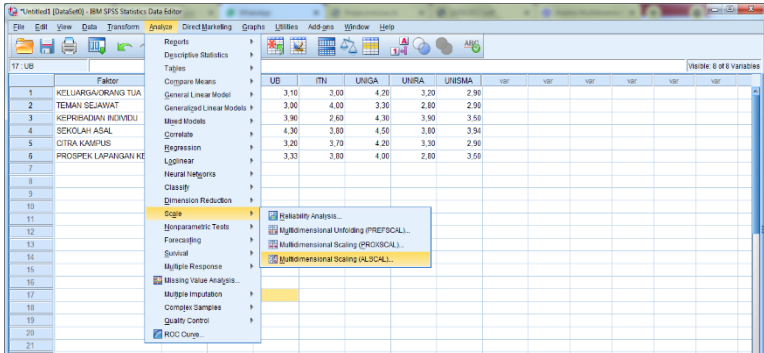


b. Nilai dari Microsoft Excel dipindahkan ke SPSS, seperti yang tampak pada gambar di bawah ini.

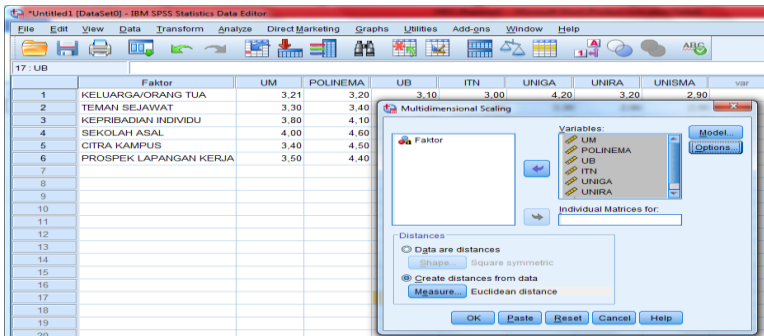
The screenshot shows the IBM SPSS Statistics Data Editor interface in data view. The main window displays a table with 7 rows and 10 columns. The first column is labeled 'Faktor' and the subsequent columns are labeled 'UM', 'POLINEMA', 'UB', 'ITN', 'UNIGA', 'UNIRA', 'UNISMA', and 'var'. The data values are as follows:

Faktor	UM	POLINEMA	UB	ITN	UNIGA	UNIRA	UNISMA	var
1 KELUARGA/ORANG TUA	3.21	3.20	3.10	3.00	4.20	3.20	2.90	
2 TEMAN SEJAWAT	3.30	3.40	3.00	4.00	3.30	2.80	2.90	
3 KEPRIBADIAN INDIVIDU	3.80	4.10	3.90	2.60	4.30	3.90	3.50	
4 SEKOLAH ASAL	4.00	4.60	4.30	3.80	4.50	3.80	3.94	
5 CITRA KAMPUS	3.40	4.50	3.20	3.70	4.20	3.30	2.90	
6 PROSPEK LAPANGAN KERJA	3.50	4.40	3.33	3.80	4.00	2.80	3.50	
7								

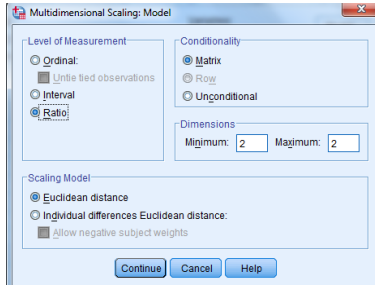
- c. Pilih Analyze, klik Scale, lalu klik Multidimensional Scaling (ALSCAL).



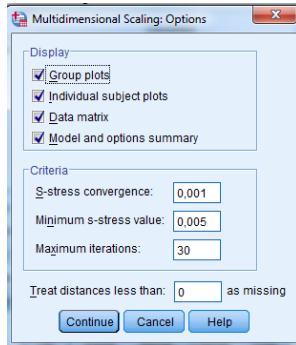
- d. Pindahkan semua variabel kampus ke dalam kotak dialog Variables.



- e. Pada menu model, pada kotak dialog Level of Measurement, pilih Ratio. Pada kotak dialog Conditionality, pilih Matrix. Pada kotak dialog Scaling Model, pilih Euclidean Distance. Lalu Continue.



- f. Pada menu Option, pada kotak dialog Display, centang semua pilihan lalu Continue dan klik OK.



2. Interpretasi MDS

Raw (unscaled) Data for Subject 1

	1	2	3	4	5	6	7
1	,000						
2	1,575	,000					
3	,521	1,771	,000				
4	1,481	2,071	1,850	,000			
5	1,540	1,145	1,717	2,366	,000		
6	,894	2,245	,769	2,081	2,037	,000	
7	,774	2,123	,677	1,667	2,178	,964	,000

Gambar 9.1 Matriks D.

Interpretasi: *Output* di atas merupakan matriks D (*distance*).

```
For matrix
Stress = ,09370    RSQ = ,97514
```

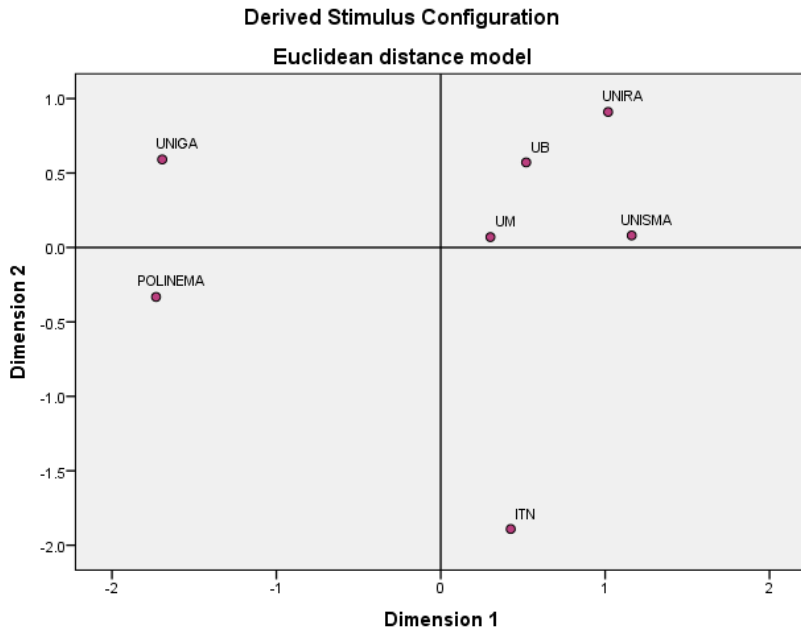
Gambar 9.2 Nilai Stress dan RSQ.

Interpretasi: *Output* di atas menunjukkan bahwa nilai Stress sebesar 0.09370 atau 9.37%, yang berarti bahwa model cukup memiliki kelayakan yang baik (fit). Demikian nilai RSQ yang cukup tinggi, yaitu sebesar 0.97514 atau 97.51% sehingga dapat kita simpulkan bahwa model penelitian yang digunakan sudah benar-benar fit.

Stimulus Coordinates			
		Dimension	
Stimulus Number	Stimulus Name	1	2
1	UM	,3018	,0699
2	POLINEMA	-1,7319	-,3321
3	UB	,5196	,5715
4	ITN	,4253	-1,8911
5	UNIGA	-1,6943	,5912
6	UNIRA	1,0182	,9101
7	UNISMA	1,1613	,0806

Gambar 9.3 Stimulus Coordinate.

Interpretasi: *Output* di atas menunjukkan letak koordinat. Terlihat bahwa letak koordinat UM di dimensi 1 dan dimensi 2 secara berturut-turut adalah 0.3018 dan 0,06. Untuk POLINEMA -1.7319 dan -0.33, dan seterusnya seperti yang terlihat pada *output* di atas.



Gambar 9.4 Grafik *Derived Stimulus Configuration*.

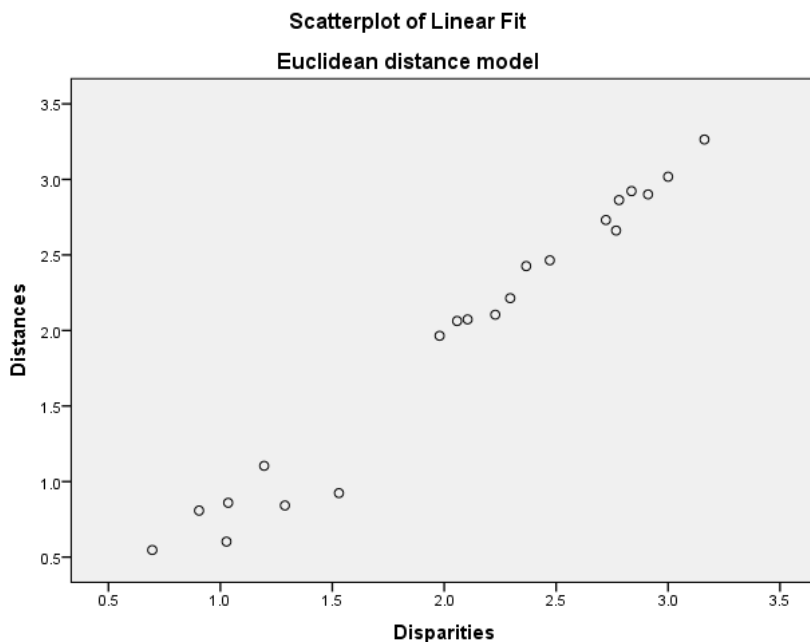
Gambar 9.4 di atas menunjukkan pengelompokan universitas. Terlihat bahwa keempat universitas yaitu UB, UNIRA, UM, dan UNISMA memiliki kemiripan karena terletak di kuadran yang sama sedangkan UNIGA, POLINEMA, dan ITN tidak memiliki kemiripan dengan universitas lain karena masing-masing berada dalam kuadran yang berbeda.

Untuk analisis dari sudut dimensi 1 (sumbu X) dan dimensi 2 (sumbu Y).

1. Dimensi 1: pada gambar terlihat bahwa makin ke kanan, angka dimensi 1 makin besar, terlihat bahwa UNISMA terletak paling dekat dengan dimensi 1 terbesar, dan memiliki kemiripan dengan tiga universitas

lainnya, yaitu UB, UNIRA dan UM. Untuk ITN tidak berada dinilai terbesar dimensi 1 dan juga tidak memiliki kemiripan dengan universitas lainnya karena berada di kuadran yang berbeda dengan Universitas lain.

2. Dimensi 2: pada gambar terlihat bahwa makin ke atas, maka angka dimensi 2 makin besar. Terlihat bahwa UNIRA terletak paling ujung dimensi 2 angka terbesar. Sedangkan ITN terletak paling ujung bawah dimensi ini. ITN mempunyai jarak terdekat dengan POLINEMA tetapi berada dalam kuadran yang berbeda.

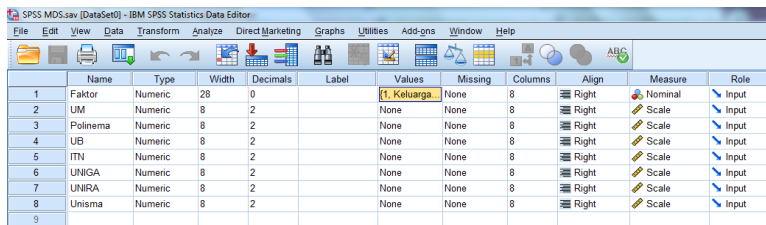


Gambar 9.5 Grafik *Scatterplot of Linear Fit*.

Grafik *Scatterplot of Linear Fit* dapat diketahui bahwa titik-titik koordinat tidak membentuk berbagai kelompok koordinat tersendiri, namun relatif ,menggerombol dan membentuk garis lurus atau pola penyebaran menuju kanan atas sehingga dapat dikatakan bahwa responden konsisten dalam memberikan penilaian terhadap universitas.

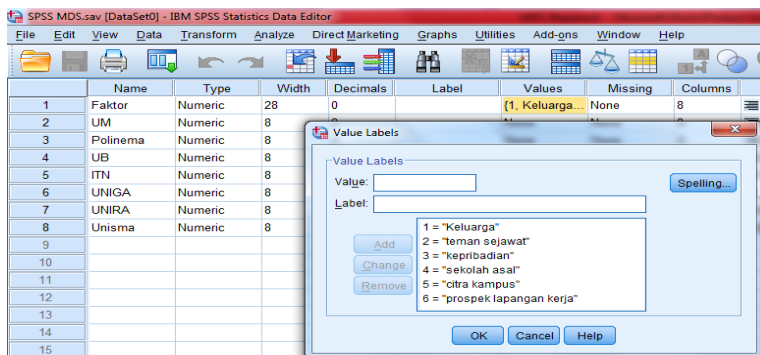
3. Menggunakan PREFSCAL

- a. Meng-*input* nilai ke dalam program SPSS, langkah pengisian identitas, sama dengan cara pertama (ALSCAL).



	Name	Type	Width	Decimals	Label	Values	Missing	Columns	Align	Measure	Role
1	Faktor	Numeric	28	0		1, Keluarga	None	8	Right	Nominal	Input
2	UM	Numeric	8	2		None	None	8	Right	Scale	Input
3	Polinema	Numeric	8	2		None	None	8	Right	Scale	Input
4	UB	Numeric	8	2		None	None	8	Right	Scale	Input
5	ITN	Numeric	8	2		None	None	8	Right	Scale	Input
6	UNIGA	Numeric	8	2		None	None	8	Right	Scale	Input
7	UNIRA	Numeric	8	2		None	None	8	Right	Scale	Input
8	Unisma	Numeric	8	2		None	None	8	Right	Scale	Input
9											

- b. Meng-*input* atribut.



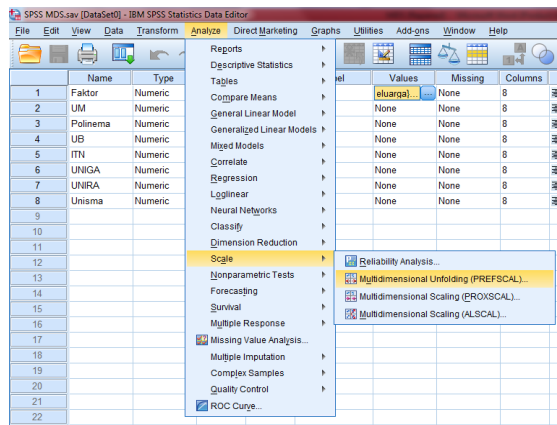
The screenshot shows the SPSS Value Labels dialog box for the variable 'Faktor'. The dialog box contains the following information:

- Value:** (Empty text box)
- Label:** (Empty text box)
- Buttons:** Add, Change, Remove, Spelling...
- List of Value-Label Pairs:**
 - 1 = "Keluarga"
 - 2 = "teman sejawal"
 - 3 = "kepribadian"
 - 4 = "sekolah asal"
 - 5 = "citra kampus"
 - 6 = "prospek lapangan kerja"
- Bottom Buttons:** OK, Cancel, Help

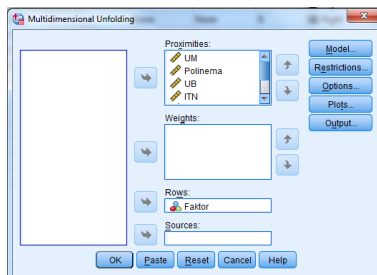
- c. Nilai dari Microsoft Excel dipindahkan ke SPSS seperti yang tampak pada gambar di bawah ini.

	Faktor	UM	POLINEMA	UB	ITN	UNIGA	UNIRA	UNISMA	var
1	KELUARGA/ORANG TUA	3,21	3,20	3,10	3,00	4,20	3,20	2,90	
2	TEMAN SEJAWAT	3,30	3,40	3,00	4,00	3,30	2,80	2,90	
3	KEPRIBADIAN INDIVIDU	3,80	4,10	3,90	2,60	4,30	3,90	3,50	
4	SEKOLAH ASAL	4,00	4,60	4,30	3,80	4,50	3,80	3,94	
5	CITRA KAMPUS	3,40	4,50	3,20	3,70	4,20	3,30	2,90	
6	PROSPEK LAPANGAN KERJA	3,50	4,40	3,33	3,80	4,00	2,80	3,50	
7									

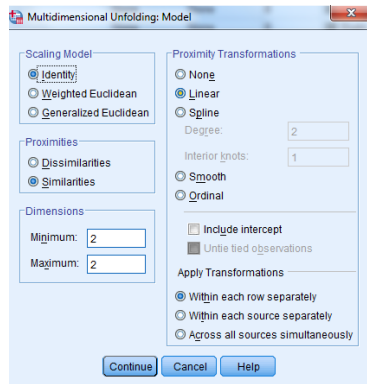
- d. Pilih Analyze lalu Scale lalu Multidimensional Unfolding (PREFSCAL).



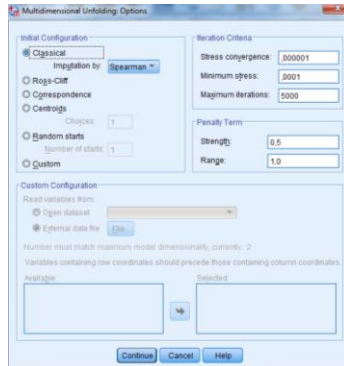
- e. Pindahkan semua variabel kampus ke dalam kotak dialog Proximities dan pindahkan faktor ke kotak Rows.



- f. Pilih kotak dialog Model. Pada menu Scaling Model, pilih Identity. Pada kotak dialog Proximities, pilih Similarities. Pada kotak dialog Proximity Transformations, pilih Linier. Klik Continue.



- g. Pilih kotak dialog Options. Pada menu Initial Configuration, pilih Classical, pada imputation by, pilih Spearman. Continue lalu OK.

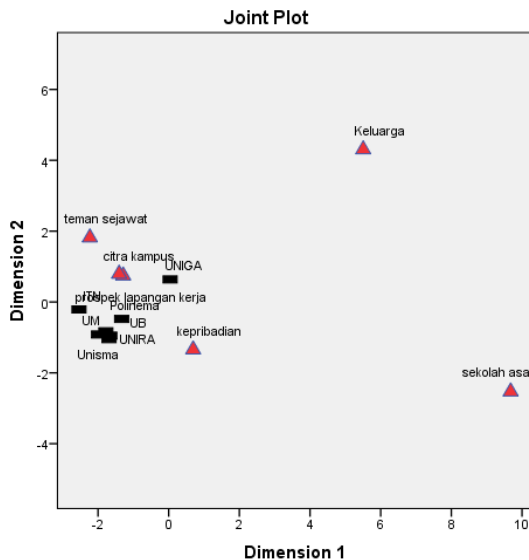


4. Interpretasi

Gambar 9.3 menunjukkan posisi antaratribut yang mempengaruhi pemilihan kampus calon mahasiswa. Dapat dilihat bahwa atribut keluarga memiliki jarak yang jauh dengan atribut lainnya, demikian juga dengan

atribut antarsekolah. Atribut teman sejawat, citra kampus, prospek lapangan kerja dan atribut kepribadian, berada dekat dengan posisi kampus-kampus yang menjadi variabel penelitian. Oleh karena itu, berdasarkan letak tersebut, dapat disimpulkan bahwa

- a. Atribut keluarga dan sekolah asal disimpulkan tidak berpengaruh signifikan terhadap pemilihan kampus untuk calon mahasiswa, hal tersebut tampak dari jarak antar atribut dengan kampus sangat jauh.
- b. Atribut teman sejawat, kepribadian, citra kampus dan prospek lapangan kerja disimpulkan memiliki pengaruh yang signifikan terhadap pemilihan kampus calon mahasiswa. Hal tersebut tampak dari jarak antara keempat atribut tersebut terhadap variabel kampus, memiliki jarak yang dekat.



Gambar 9.6 Joint Plot.

Daftar Rujukan

- Carroll, J.D., dan Arabie, P. (1998). Multidimensional Scaling. *Measurement, judgment and decision making*, 179–250.
- Cox, M.A., dan Cox, T.F. (2008). Multidimensional Scaling. In *Handbook of data visualization*, 315–347. Springer, Berlin, Heidelberg.
- Demaine, E., Hesterberg, A., Koehler, F., Lynch, J., dan Urschel, J. (2021). Multidimensional scaling: Approximation and complexity. In *International Conference on Machine Learning*, 2568–2578.
- Ghojogh, B., Ghodsi, A., Karray, F., dan Crowley, M. (2020). Multidimensional Scaling, Sammon Mapping, and Isomap: Tutorial and survey. *arXiv preprint arXiv: 2009.08136*.
- Gudono. (2016). *Analisis Data Multivariat*. Edisi 4. Yogyakarta: Universitas Gajah Mada.
- Irmawati, I. (2017). *Penerapan Analisis Multidimensional Scalling pada Pemetaan Karakteristik Kemiskinan di Provinsi Sulawesi Selatan (Doctoral Dissertation, Universitas Islam Negeri Alauddin Makassar, Makassar)*.
- Kruskal, J.B. (1978). *Multidimensional scaling*. No. 11. Sage.
- Lestari, N. (2017). *Produktivitas Padi di Provinsi Sulawesi Selatan dengan Analisis Multidimensional Scaling (MDS) (Doctoral Dissertation, Univeritas Islam Negeri Alauddin Makassar, Makassar)*.
- Machado, J. T., dan Lopes, A. M. (2020). Multidimensional Scaling Locus of Memristor and Fractional Order Elements. *Journal of Advanced Research*, 25, 147–157.
- Masuku, T. J., Paendong, M. S., dan Langi, Y. A. (2014). Persepsi Konsumen Terhadap Produk Sepatu Olahraga di Sport Station Megamall dengan

- Menggunakan Analisis Multidimensional Scaling. *Jurnal Ilmiah Sains*, 14(2), 68–72.
- Mattjik, A.A., Sumertajaya, I., Wibawa, G.N.A., & Hadi, A.F. (2011). Sidik Peubah Ganda Dengan Menggunakan SAS.
- Simamora, B. (2005). *Analisis Multivariat Pemasaran*. Jakarta: Gramedia Pustaka Utama.
- Supranto, J. (2004). *Analisis Multivariat Arti dan Interpretasi*. Jakarta: Rineka Cipta.
- Timm, N.H. (2002). *Applied Multivariate Analysis*. Springer.
- Torgerson, W.S. (1952). Multidimensional Scalling: I. Theory and Method. *Psychometrika*, 17(4), 401–419.
- Walpole, R.E. 1995. *Pengantar Statistika*. Jakarta: Gramedia Pustaka Utama.
- Walundungo, G., Paendong, M., dan Manurung, T. (2014). Penggunaan Analisis Multidimensional Scaling untuk Mengetahui Kemiripan Rumah Makan di Manado Town Square Berdasarkan Karakteristik Pelanggan. *d'CARTESIAN*, 3(1), 30–35.
- Wardhana, R.P., & Sutanto, H. T. 2009. Analisis Posisi Produk Kartu GSM dengan Metode *Multidimensional Scalling* (MDS) pada Mahasiswa FMIPA Universitas Negeri Surabaya.

Tentang Penulis



Purnomo adalah salah satu Guru Besar Fakultas Teknik Universitas Negeri Malang (UM), Pria kelahiran Kediri, 21 Juni 1959 ini, merupakan alumni S1 Pendidikan Teknik Mesin IKIP Malang yang lulus tahun 1983. Purnomo menyelesaikan S2 Pendidikan Teknologi dan Kejuruan IKIP Jakarta di IKIP Yogyakarta tahun 1991. Program Doktor S3 Pendidikan Teknologi dan Kejuruan diselesaikan di Tahun 2014 di Universitas Negeri Yogyakarta.

Profesor bidang Pendidikan Teknologi dan Kejuruan Universitas Negeri Malang ini memiliki banyak pengalaman di Pendidikan dan pelatihan bidang penjaminan mutu, maupun reviewer hibah kompetisi.

Karir beliau di bidang pendidikan dimulai sejak Tahun 1983 sebagai guru di SMA Negeri Kandat Kediri, hingga tahun 1985, kemudian sejak tahun 1985 menjadi Dosen Tetap di Universitas Negeri Malang sampai sekarang. Karir beliau di struktural, pernah menjadi Sekretaris Jurusan Pendidikan Teknik Mesin di tahun 1993 hingga 1996. Pernah menduduki jabatan sebagai Ketua Jurusan Pendidikan Teknik Mesin, FPTK IKIP Malang di periode 1996-1999, Ketua Kurusan Teknik Mesin FT UM tahun 1999-2003. Menjabat sebagai Sekretaris Badan Penjamin Mutu Universitas Negeri Malang Tahun 2005-2008, dan menjabat sebagai Koordinator Program Studi S2 dan S3 Pendidikan Kejuruan Pascasarjana Universitas Negeri Malang tahun 2015-2019.

Penghargaan dari Rektor IKIP Malang, sebagai Dosen Teladan 1 Institut, tahun 1994; Penghargaan dari Menteri Pendidikan Dan Kebudayaan Republik Indonesia, sebagai Dosen Teladan Nasional (Aditya Tridharma Nugraha); tahun 1994, dan Penghargaan dari Presiden Republik Indonesia, berupa Satyalancana Karya Satya XX Tahun, tahun 2007.

Sejak tahun 2005 beliau telah menorehkan 14 karya penelitian di bidang Pendidikan kejuruan, lebih dari 10 judul publikasi jurnal ilmiah internasional

bereputasi terindeks scopus. Sederetan pengalaman, pendidikan, pelatihan, karya dan prestasi inilah yang membawa beliau meraih gelar Profesor (Guru Besar) bidang Pendidikan Teknologi dan Kejuruan Fakultas Teknik Universitas Negeri Malang (UM).



Eddy Sutadji, Wakil Dekan II Fakultas Teknik Universitas Negeri Malang (UM), Pria kelahiran Surabaya, 13 Maret 1961 ini merupakan alumni S1 Pendidikan Teknik Mesin Universitas Negeri Surabaya yang lulus tahun 1986. Eddy Sutadji menyelesaikan S2 Teknologi Pembelajaran Universitas Negeri Malang yang lulus tahun 1997. Program Doktor S3 Penelitian dan Evaluasi Pendidikan Universitas Negeri Yogyakarta yang diselesaikan tahun 2009.

Dosen S1 Jurusan Teknik Mesin dan Dosen S2 dan S3 Prodi Pendidikan Kejuruan dan Dosen S2 Prodi Dikdas ini memiliki banyak pengalaman di bidang pendidikan.

Karir beliau di bidang pendidikan dimulai sejak Tahun 1987 sebagai Dosen S1 jurusan Teknik Mesin Universitas Negeri Malang sampai sekarang. Pada tahun 2001 menjabat sebagai kepala Lab. Teknik Mesin UM sampai tahun 2005, dari tahun 2002 sampai 2005 menjadi Direktur program Community College UM, tahun 2009 menjabat sebagai Koordinator Program Pendidikan Profesi Guru (PPG) Teknik Mesin Fakultas Teknik UM. Tahun 2010 sampai sekarang menjadi Dosen S2 dan S3 Prodi Pendidikan Kejuruan dan Dosen S2 Prodi Diknas Pascasarjana UM. Tahun 2011 sampai 2013 menjabat sebagai Dosen Program PPG Terintegrasi Dikti Kemendikbud. Tahun 2010 sampai 2014 sebagai Tim Pakar (expert) program TEQIP (Teacher Quality Improvement Program). Tahun 2010 menjabat sebagai Ketua Tim Pengembang Fakultas Teknik Fakultas Teknik UM. Tahun 2011 sampai tahun 2015 sebagai Koordinator Sarjana Mendidik di Daerah Terdepan, Terluar, dan Tertinggal (Koord. SM-3T) UM. Tahun 2000 sampai 2013 sebagai Reviewer Pengabdian Masyarakat Ditbinlitabmas Dikti Kemendikbud. Tahun 2011 sampai 2013 sebagai Asesor Badan Akreditasi Nasional Perguruan Tinggi (BAN-PT) dan Tim Monev Internal (Monevin)

Program Hibah Kompetensi Institusi (PHKI). Pada tahun 2010 mendapat penghargaan Dosen Berprestasi tingkat Fakultas. Tahun 2012 sampai 2015 Reviewer Penelitian DP2M Dikti Kemendikbud. Tahun 2013 sampai 2014 sebagai Ketua Pelaksana Program Pendidikan Profesi Guru (PPG) Pasca SM-3T. November 2012 sampai Desember 2014 menjabat sebagai Kepala Pusat Penelitian dan Pengabdian kepada Masyarakat Bidang Sains, Teknologi dan Industri serta Hak Kekayaan Intelektual (PSTIHKI) LP2M UM. Desember 2014 sampai 20 Januari 2015 menjabat sebagai Staf Ahli Wakil Rektor I UM. Februari 2015 sampai Desember 2018 menjabat sebagai sekretaris LP3 UM, Januari 2019 sampai Desember 2022 menjabat sebagai Wakil Dekan II Fakultas Teknik UM.

Sejak tahun 2009 beliau telah melakukan lebih dari 10 kegiatan penelitian di berbagai bidang, menerbitkan puluhan Jurnal Nasional, serta banyak kegiatan ilmiah (Seminar, Workshop dan lokakarya) dan kegiatan Pengabdian kepada masyarakat yang lainnya.



Annisa' Carina, lahir di Lamongan, 20 November 1990. Berdomisili di Jl. Made Karyo 3/32, Kecamatan Lamongan, Kabupaten Lamongan, Provinsi Jawa Timur. Pernah menempuh pendidikan di SD Negeri 1 Tumenggungan tahun 1997–2003, SMP Negeri 3 Lamongan tahun 2003–2006, SMA Negeri 2 Lamongan tahun 2006–2009, S-1 Pendidikan Teknik Bangunan Jurusan Teknik Sipil Universitas Negeri Malang tahun 2009–2013, S-2 Pendidikan Kejuruan Konsentrasi Teknik Sipil dan Bangunan Universitas Negeri Malang tahun 2013–2015, dan tahun 2021 sampai saat ini sedang menempuh S-3 Pendidikan Kejuruan Universitas Negeri Malang.

Dari tahun 2015 sampai sekarang bekerja di STT STIKMA Internasional Malang. Selama bekerja, Annisa pernah menjadi ketua program studi S-1 Arsitektur STT STIKMA Internasional tahun 2018–2020, menjadi Tim Penerimaan Mahasiswa Baru 2018, menjadi anggota senat 2018–sekarang, menjadi Tim Sistem Manajemen Mutu STT STIKMA Internasional bekerja sama dengan PT Asuh Institut Teknologi Sepuluh

November (ITS) tahun 2018–2019, dan penerima dana hibah Penelitian Dosen Pemula Kemenristek DIKTI tahun anggaran 2021.



Najih Gilang Ramadhan, lahir di Jember pada 14 Februari 1994. Bertempat tinggal di Desa Andongsari RT 001/ RW 004, Kecamatan Ambulu, Kabupaten Jember, Provinsi Jawa Timur. Pernah menempuh pendidikan di MI 1 Muhammadiyah Watukebo lulus pada tahun 2006, SMP 9 Muhammadiyah lulus pada tahun 2009, SMA Negeri 1 Ambulu lulus pada tahun 2012, S-1 Pendidikan Teknik Otomotif Universitas Negeri Malang lulus pada tahun 2016, S-2 Pendidikan Kejuruan Universitas Negeri Malang lulus pada tahun 2020. Pada tahun 2021, ia melanjutkan pendidikan S-3 Pendidikan Kejuruan di Universitas Negeri Malang.

Karier Gilang sebagai guru SMK dimulai pada tahun 2021, mengajar mata pelajaran produktif pada Kompetensi Keahlian Teknologi Kendaraan Ringan (TKR), seperti Pemeliharaan Mesin Kendaraan Ringan, Pemeliharaan Kelistrikan Kendaraan Ringan, Pemeliharaan Sasis dan Pemindahan Tenaga Kendaraan Ringan, serta Produk Kreatif dan Kewirausahaan. Selama bekerja, ia pernah menjadi bagian dari Tim Event Mobil Listrik di 5 Kabupaten (2021), Pengembang atau Inovasi Sepeda Listrik dan Mobil listrik (2021), Guru Terbaik kategori Pembina Siswa (2021), dan Tim Pengembang MoU antara Sekolah dengan Industri dalam Praktik Kerja lapangan (2021).



Okta Purnawirawan, lahir di Tulungagung pada 21 Oktober 1993. Bertempat tinggal di RT 002 RW 009, Dusun Melikan, Kecamatan Kedungwaru, Kabupaten Tulungagung, Provinsi Jawa Timur. Pernah mengenyam pendidikan di SD Negeri 2 Dono-Sendang lulus pada tahun 2006, SMP Negeri 2 Tulungagung lulus pada tahun 2009, SMA Negeri 1 Kauman Tulungagung lulus pada tahun 2012, S-1 Pendidikan Teknik Informatika dan Komputer

Universitas Negeri Semarang lulus pada tahun 2016, S-2 Pendidikan Kejuruan Universitas Negeri Semarang lulus pada tahun 2019. Pada tahun 2021, ia melanjutkan S-3 Pendidikan Kejuruan di Universitas Negeri Malang.

Kariernya sebagai guru SMK dimulai dengan mengajar mata pelajaran produktif pada kompetensi keahlian Rekayasa Perangkat Lunak (RPL), seperti Sistem Komputer, Dasar Desain Grafis, Pemrograman Web, Basis Data, serta Produk Kreatif dan Kewirausahaan. Prestasi yang pernah diraih, yaitu Pemakalah/Presenter Terbaik Seminar Nasional Industri dan Kerajinan Batik tahun 2021, Juara 1 Lomba Video Pembelajaran Inovatif Guru SMA/MA/SMK Tingkat Nasional tahun 2020, Juara 2 Lomba Desain Logo HUT PT Jasa Raharja ke-60 Tingkat Nasional, dan Guru Berkinerja Terbaik Tahun 2020.



Rachmat Faris di Kota Malang pada 15 Oktober 1979, ia putra pertama dari Bapak Suhadi dan Ibu Kartini. Menamatkan pendidikan dasar tahun 1992 di SDN Purwantoro 1 Kota Malang, pendidikan menengah pertama di SMPN 3 Kota Malang pada tahun 1995, dan pendidikan menengah atas di SMAN 4 Kota Malang pada tahun 1998. Pada tahun yang sama, melanjutkan pendidikan

S-1 Manajemen Fakultas Ekonomi Universitas Brawijaya Malang. Menyelesaikan studi S-2 Pendidikan Bisnis dan Manajemen di Universitas Negeri Malang, dan pada tahun 2021 melanjutkan S-3 Pendidikan Kejuruan

di Universitas Negeri Malang.



Ratna Fajarwati Meditama, lahir di Blitar pada 8 November 1990. Saat ini tinggal di Perum Villa Ken Arok Wonokoyo Kota Malang, Jawa Timur. Riwayat pendidikan Ratna dimulai dari SD Negeri 1 Karangsono, SMP N 1 Kanigoro, SMA N 1 Talun, S-1 Pendidikan Teknik Mesin Universitas Negeri Malang, S-2 Pendidikan Kejuruan Universitas Negeri Malang. Pada tahun 2021 ia melanjutkan S-3 Pendidikan Kejuruan di

Universitas Negeri Malang. Kini, Ratna bekerja sebagai dosen di Program Studi Teknik Mesin Universitas Islam Raden Rahmat Malang.



Sulistianingsih A.S., lahir di Ujung Pandang, Makassar, pada 08 Januari 1988. Bertempat tinggal di Graha Merjosari Asri Cluster Mawar No.7 RT 08/RW 12, Kelurahan Merjosari, Kecamatan Lowokwaru, Kota Malang, Provinsi Jawa Timur. Riwayat pendidikannya dimulai dari SD Negeri 612 Mallaulu lulus pada tahun 1999, SMP Negeri 2 Malili lulus pada tahun 2002, SMA Negeri 1 Malili lulus pada tahun 2005, S-1 Pendidikan Teknik Bangunan Universitas Negeri Makassar

lulus pada tahun 2010, S-2 Pendidikan Kejuruan Universitas Negeri Malang lulus pada tahun 2014. Pada tahun 2021, ia melanjutkan S-3 Pendidikan Kejuruan di Universitas Negeri Malang. Kini, ia bekerja sebagai dosen Program Studi Arsitektur Sekolah Tinggi Teknologi STIKMA Internasional, mengajar mata kuliah Teknologi Bahan, Utilitas Bangunan, Rencana Anggaran Biaya, Manajemen Konstruksi, dan Sejarah Arsitektur.



Wargijono Utomo, lulus S-1 Program Studi Teknik Industri Fakultas Teknologi Industri Universitas Persada Indonesia YAI Jakarta tahun 2003 dan lulus S-2 Program Studi Teknik Informatika STMIK Eresha Jakarta tahun 2016, saat ini sedang menempuh S-3 Program Studi Pendidikan Kejuruan Fakultas Teknik Universitas Negeri Malang. Sejak tahun 2016, ia menjadi dosen tetap Program Studi Teknik Informatika Fakultas Teknik Universitas Krisnadwipayana Jakarta.

