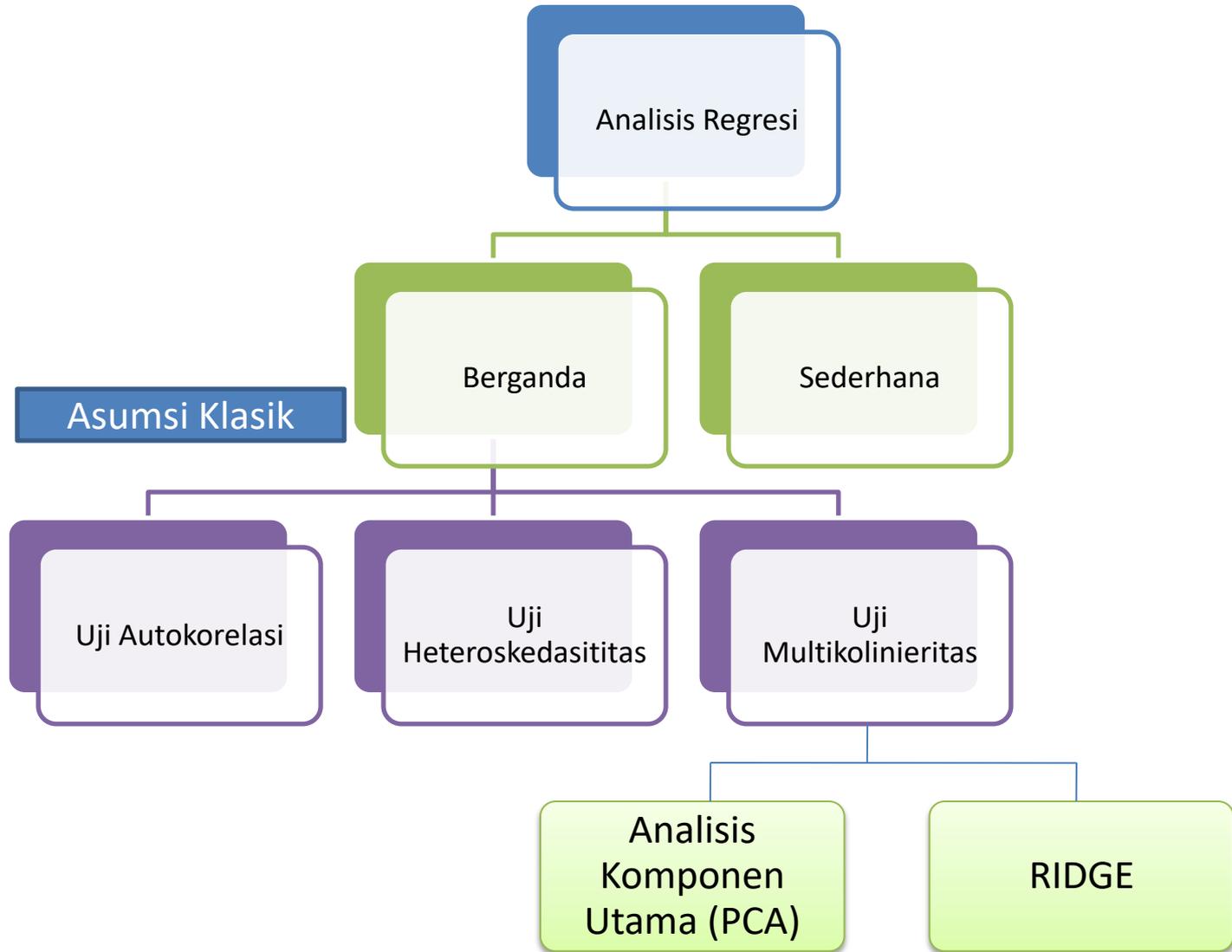


Principal Component Analysis (PCA)
Salah Cara Mengatasi
Multikolinearitas



Multikolinearitas

- Istilah multikolinearitas mula-mula ditemukan oleh Ragnar Frisch. Pada mulanya multikolinearitas berarti adanya hubungan linear yang sempurna atau pasti, diantara beberapa atau semua variabel bebas dari model regresi ganda (Gujarati, 1995:157)
- Multikolinearitas muncul ketika variabel-variabel bebasnya saling berkorelasi.
- Jika terdapat multikolinearitas dalam model regresi, hal itu dapat menyebabkan hasil estimasi menggunakan metode kuadrat terkecil menjadi tidak valid.

Cara Mendeteksi Multikolinearitas

- Ada beberapa cara untuk mengetahui keberadaan multikolinearitas dalam suatu model regresi, salah satunya dengan bantuan *software* SPSS dengan melihat nilai *Tolerance* atau VIF (*Variance Inflation Factor*) pada tabel "*coefficients*". Jika nilai *Tolerance* kurang dari 0,1 atau nilai VIF melebihi 10 maka hal tersebut menunjukkan bahwa multikolinearitas adalah masalah yang pasti terjadi antar variabel bebas (Soemartini,2008:10).

Cara mengatasi Multikolinieritas dengan Principal Component Analysis (PCA)

- Metode PCA bertujuan untuk menyederhanakan variabel yang diamati dengan cara mereduksi dimensinya. Hal ini dilakukan dengan cara menghilangkan korelasi diantara variabel bebas melalui transformasi variabel bebas asal ke variabel baru yang tidak berkorelasi sama sekali.
- Setelah beberapa komponen hasil PCA yang bebas multikolinieritas diperoleh, maka komponen-komponen tersebut menjadi variabel bebas baru yang akan diregresikan atau dianalisis pengaruhnya terhadap variabel tak bebas (Y) dengan menggunakan analisis regresi.
- Keunggulan metode PCA diantaranya adalah dapat menghilangkan korelasi secara bersih tanpa harus mengurangi jumlah variabel asal.

Langkah-langkah penggunaan PCA adalah sebagai berikut:

1. *Kaiser-Meyer-Olkin (KMO) dan Barlett Test*

- a. KMO berkisar antara 0,5 s.d1 maka analisis faktor layak digunakan
- b. *Barlett Test* digunakan untuk menguji apakah benar variabel-variabel yang dilibatkan berkorelasi.

Kriteria uji dengan melihat *p-value* (signifikan): terima H_0 jika $\text{sig.} > 0,05$ atau tolak H_0 jika $\text{sig.} < 0,05$.

2. *Anti Image Matriks*

Angka MSA (*Measure of Sampling Adequacy*) berkisar dari 0 sampai 1, dengan kriteria sebagai berikut:

- $\text{MSA} = 1$, variabel tersebut dapat diprediksi tanpa kesalahan o/ variabel lain.
- $\text{MSA} > 0,5$, variabel masih bisa diprediksi dan bisa dianalisis lebih lanjut.
- $\text{MSA} < 0,5$, variabel tidak bisa diprediksi dan tidak bisa dianalisis lebih lanjut, atau dikeluarkan dari variabel lainnya.

PCA Lanjutan

3. *Communalities*

Communalities menunjukkan berapa varians yang dapat dijelaskan oleh faktor yang terbentuk.

4. *Total Variance Explained*

Susunan *eigenvalue* selalu diurutkan dari yang terbesar sampai yang terkecil, dengan kriteria bahwa angka *eigenvalue* di bawah 1 tidak digunakan dalam menghitung jumlah faktor yang terbentuk.

5. *Componen Matriks*

Componen Matriks merupakan tabel yang berisikan *factor loading* (nilai korelasi) antara variabel-variabel analisis dengan faktor yang terbentuk

6. *Component Score Coefficient Matriks*

Setelah didapatkan faktor yang terbentuk melalui proses reduksi, maka perlu dicari persamaan sehingga dapat dihitung skor setiap faktor secara manual. Setelah komponen hasil PCA yang bebas multikolinearitas diperoleh maka komponen-komponen tersebut diregresikan atau dianalisa pengaruhnya terhadap variabel tak bebas (Y) dengan menggunakan analisis regresi linear.

PCA MENGGUNAKAN SOFTWARE SPSS

1. Pilih menu **Analyze**, kemudian submenu **Data reduction**, lalu pilih **factor**.
2. Tampak di layar *windows Factor Analysis*.
3. Masukkan semua faktor dikotak **Factor Analysis** ke dalam kotak **variables**.
4. klik tombol **Deskriptives** yang berada disebelah kiri bawah. Pilih dengan menandai **KMO and Bartlett's test of sphericity** serta **anti- image**. Klik **continue**.
5. Klik tombol **Extraction**.
6. Tampak di layar *kotak dialog Extraction*.
7. Pada kotak **Method**, pilih **Principal Components**.
8. Pada kotak **Analyze**, tetap pada pilihan **Correlation Matrix**.
9. Pada kotak **Display**, aktifkan **unrotated factor Solution**.
10. Pada kotak **Eigenvalues Over**, tetap pada angka 1.
11. Pada kotak **Maximum Iteration For Convergen**, tetap pada angka 25, tekan **Continue**.
12. Klik tombol **Rotation**.
13. Pada kotak **Method**, pilih **Varimax**

PCA SOFTWARE SPSS Lanjutan

14. Pada kotak **Display**, aktifkan **Rotated Solution** dan **Loading Plot(s)**.
15. Pada kotak **Maximum Iteration For Convergen**, tetap pada angka 25, tekan **Continue**.
16. Klik tombol **Scores**, aktifkan kotak **Save as variables**, dan secara otomatis kotak **Method** akan terbuka, pilih **Regression**, Klik **Display factor score coefficient matrix** tekan **Continue**.
17. Abaikan yang lain dan tekan **Ok**.

CONTOH KASUS

Hubungan variabel independen (X1,X2, dan X3) terhadap variable dependen (Y)

Y	X1	X2	X3
62,8	43,41	17,1	3,96
65	46,44	18,65	5,48
63,9	44,35	17,09	4,37
67,5	47,82	19,28	4,51
71,3	51,02	23,24	4,88
76,6	58,71	28,11	6,37
86,3	87,69	30,29	8,96
95,7	76,73	28,26	9,76
98,3	75,91	27,91	9,31
100,3	77,62	32,3	9,85
103,2	78,01	31,39	7,21
108,9	83,57	35,61	7,39
108,5	90,59	37,58	7,98
111,4	95,47	35,17	7,42

X1: SDM yang Kompeten; X2: Sistem Pengendalian Intern; X3: Sistem Pengolahan Data; Y: Kualitas L/K

Hasil Pengolahan Data SPSS

Coefficients^a

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations			Collinearity Statistics	
	B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
(Constant)	18,702	6,845		2,732	,021					
1 X1	,380	,312	,385	1,218	,251	,942	,360	,110	,081	12,297
X2	1,419	,720	,539	1,969	,077	,946	,529	,178	,108	9,230
X3	,533	1,400	,059	,381	,711	,769	,120	,034	,336	2,977

a. Dependent Variable: Y

Dengan nilai VIF =12,297 dan nilai toleransi 0,081 (VIF>10 dan tolenrance < 1) disimpulkan bahwa terdapat multikolinieritas pada data

Mengatasi Masalah Multikolinearitas dengan Menggunakan Metode *Principal Component Analysis* (PCA)

1. KMO and Bartlett's Test

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		,666
	Approx. Chi-Square	36,774
Bartlett's Test of Sphericity	df	3
	Sig.	,000

1. Nilai KMO = 0,666 berada pada 0,5 dan 1, maka analisis faktor layak digunakan

2. Sedangkan *Bartlett Test* digunakan untuk menguji apakah benar variabel-variabel yang dilibatkan berkorelasi.

Hipotesis:

H₀ = tidak ada korelasi antar variabel bebas

H_a = ada korelasi antar variabel bebas

Kriteria uji dengan melihat *p-value* (signifikansi). Terima jika Sig. > 0,05. Tabel **KMO and Bartlett's Test** menunjukkan bahwa nilai chi-square = 36,774 dengan derajat kebebasan 3, dan *p-value* (0,000) < 0,05, maka H₀ ditolak. Artinya terdapat korelasi antar variabel bebas.

Metode *Principal Component Analysis* (PCA) Lanjutan

2. *Anti Image Matriks* (MSA)

Berdasarkan kriteria angka MSA, pada tabel **Anti-image Matrices** terlihat bahwa semua angka MSA memiliki nilai di atas 0,5. Artinya analisis dapat dilanjutkan.

Anti-image Matrices

		X1	X2	X3
Anti-image Covariance	X1	,081	-,082	-,085
	X2	-,082	,108	,027
	X3	-,085	,027	,336
Anti-image Correlation	X1	,601 ^a	-,873	-,514
	X2	-,873	,647 ^a	,141
	X3	-,514	,141	,809 ^a

a. Measures of Sampling Adequacy(MSA)

3. *Communalitas*

Tabel **Communalities** terlihat bahwa untuk variabel X1, diperoleh nilai sebesar 0,954 = 95,4%. Hal ini berarti 95,4% variabel X1 dapat dijelaskan oleh faktor yang terbentuk. Demikian juga untuk variabel X2 dan X3.

Communalities

	Initial	Extraction
X1	1,000	,954
X2	1,000	,905
X3	1,000	,805

Extraction Method: Principal Component

Analysis.

Metode *Principal Component Analysis* (PCA) Lanjutan

4. *Total Variance Explained*

Total Variance Explained terlihat bahwa angka *eigenvalues* di bawah 1 tidak dapat digunakan dalam menghitung jumlah faktor yang terbentuk, sehingga proses *factoring* berhenti pada pada satu faktor saja. Faktor satu memiliki *eigenvalues* sebesar 2,664, artinya dengan satu faktor ini dapat menjelaskan 2,664 atau 88,79% dari total keragaman variabel asal.

Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2,664	88,787	88,787	2,664	88,787	88,787
2	,288	9,599	98,386			
3	,048	1,614	100,000			

Extraction Method: Principal Component Analysis.

Metode *Principal Component Analysis* (PCA) Lanjutan

5. *Component Matriks dan Component Score Coefficiens Matriks*

Tabel **Component Matrix** terlihat bahwa hanya satu faktor yang terbentuk dari ketiga variabel. Hal tersebut berarti bahwa satu faktor adalah jumlah yang paling optimal untuk mereduksi ketiga variabel bebas tersebut.

Dengan menggunakan tabel **Component Score Coefficient Matrix** diperoleh persamaan untuk faktor baru yang terbentuk adalah sebagai berikut:

$$F1 = 0,367X1 + 0,357X2 + 0,337X3$$

Skor-skor faktor yang dihasilkan dapat digunakan untuk menggantikan skor-skor pada variabel bebas yang asli. Setelah komponen hasil PCA yang bebas multikolinearitas diperoleh maka komponen-komponen tersebut diregresikan atau dianalisis pengaruhnya terhadap variabel tak bebas dengan menggunakan analisis regresi linear.

	Component
	1
X1	,367
X2	,357
X3	,337

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Component Scores.

Metode *Principal Component Analysis* (PCA) Lanjutan

5. *Component Matriks dan Component Score Coefficiens Matriks*

Tabel **Component Matrix** terlihat bahwa hanya satu faktor yang terbentuk dari ketiga variabel. Hal tersebut berarti bahwa satu faktor adalah jumlah yang paling optimal untuk mereduksi ketiga variabel bebas tersebut.

Dengan menggunakan tabel **Component Score Coefficient Matrix** diperoleh persamaan untuk faktor baru yang terbentuk adalah sebagai berikut:

$$F_1 = 0,367x_1 + 0,357x_2 + 0,337x_3$$

Skor-skor faktor yang dihasilkan dapat digunakan untuk menggantikan skor-skor pada variabel bebas yang asli. Setelah komponen hasil PCA yang bebas multikolinearitas diperoleh maka komponen-komponen tersebut diregresikan atau dianalisis pengaruhnya terhadap variabel tak bebas dengan menggunakan analisis regresi linear.

	Component
	1
X1	,367
X2	,357
X3	,337

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

Component Scores.

Metode *Principal Component Analysis* (PCA) Lanjutan

6. Model Regresi yang Ideal

Setelah didapatkan variabel bebas baru (F1) yang bebas multikolinearitas melalui teknik PCA, maka langkah berikutnya adalah meregresikan variabel bebas yang baru (F1) terhadap variabel tak bebas (Y). Karena variabel bebas baru (F1) yang terbentuk hanya satu, maka pada model tersebut digunakan analisis regresi linear sederhana sebagai berikut:

$$Y = \beta_0 + \beta_1 F_1 + \epsilon_i$$

Dimana: $F_1 = 0,367x_1 + 0,357x_2 + 0,337x_3$

Berdasarkan tabel **Coefficients**a diperoleh model regresi sebagai berikut:

$$Y = 87,121 + 17,568F_1$$

Model		Coefficients ^a			t	Sig.
		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta		
1	(Constant)	87,121	1,275		68,325	,000
	REGR factor score 1 for analysis 1	18,071	1,323	,969	13,656	,000

a. Dependent Variable: Y

- $Y = 87$

TERIMAKASIH

Tabel 2
 PERBANDINGAN ANTARA ANALISIS FAKTOR
 DENGAN ANALISIS KOMPONEN UTAMA

Karakteristik	Analisis Komponen Utama	Analisis Faktor
Model matematika	<ul style="list-style-type: none"> - Komponen Utama merupakan kombinasi linier dari variabel asal (X_1, X_2, \dots, X_p) - Komponen Utama merupakan fungsi dari variabel asal (X_1, X_2, \dots, X_p) 	<ul style="list-style-type: none"> - Faktor terkandung dalam variabel asal (X_1, X_2, \dots, X_p) - Variabel asal (X_1, X_2, \dots, X_p) merupakan fungsi dari faktor
Variabel asal (X_1, X_2, \dots, X_p)	Tidak memiliki atau memiliki faktor bersama	Harus memiliki faktor bersama (<i>common factor</i>)
Metode Pendugaan <i>loading</i>	Konsep <i>eigen value</i> dan <i>eigen vector</i>	<ul style="list-style-type: none"> - PCA <i>solution</i> - MLE
Model pengukuran	<ul style="list-style-type: none"> - Indikator: formatif - Variabel laten: berupa "variabel komposit" yang dibentuk oleh Indikator 	<ul style="list-style-type: none"> - Indikator: reflektif - Variabel laten: berupa "faktor" yang direfleksikan oleh Indikator