

# **ANALISIS BIPLLOT (BIPLOT ANALYSIS)**

Bahan Kuliah Secara Daring  
Mahasiswa Departemen Statistika-FMIPA-IPB  
Oleh: Dr. Ir. Budi Susetyo

# Latar Belakang (1)

- Dalam bidang penelitian tertentu, seringkali jumlah objek/individu dan jumlah peubah pencirinya tidak terlalu banyak sehingga tertarik untuk melihat bagaimana pola sebaran individu, membandingkan karakteristik antar peubah penciri dan karakteristik individu yang digambarkan oleh peubah penciri
- Sebagai contoh ingin dilihat tingkat kesejahteraan dari 34 provinsi di Indonesia berdasarkan pendidikan penduduk, kesehatan, penghasilan, dan pekerjaan. Berdasarkan penciri ini ingin dilihat provinsi mana saja yang memiliki kemiripan tingkat kesejahteraan. Selain itu juga ingin diketahui provinsi tersebut lebih dominan dicirikan oleh apa dan bagaimana struktur keragaman dan korelasi antar peubah penciri.
- Analisis Biplot merupakan tehnik eksplorasi peubah ganda yang dapat menyajikan data peubah ganda dalam peta dua dimensi, sehingga perilaku data, baik objek maupun peubahnya, mudah dilihat dan diinterpretasikan

# Latar Belakang (2)

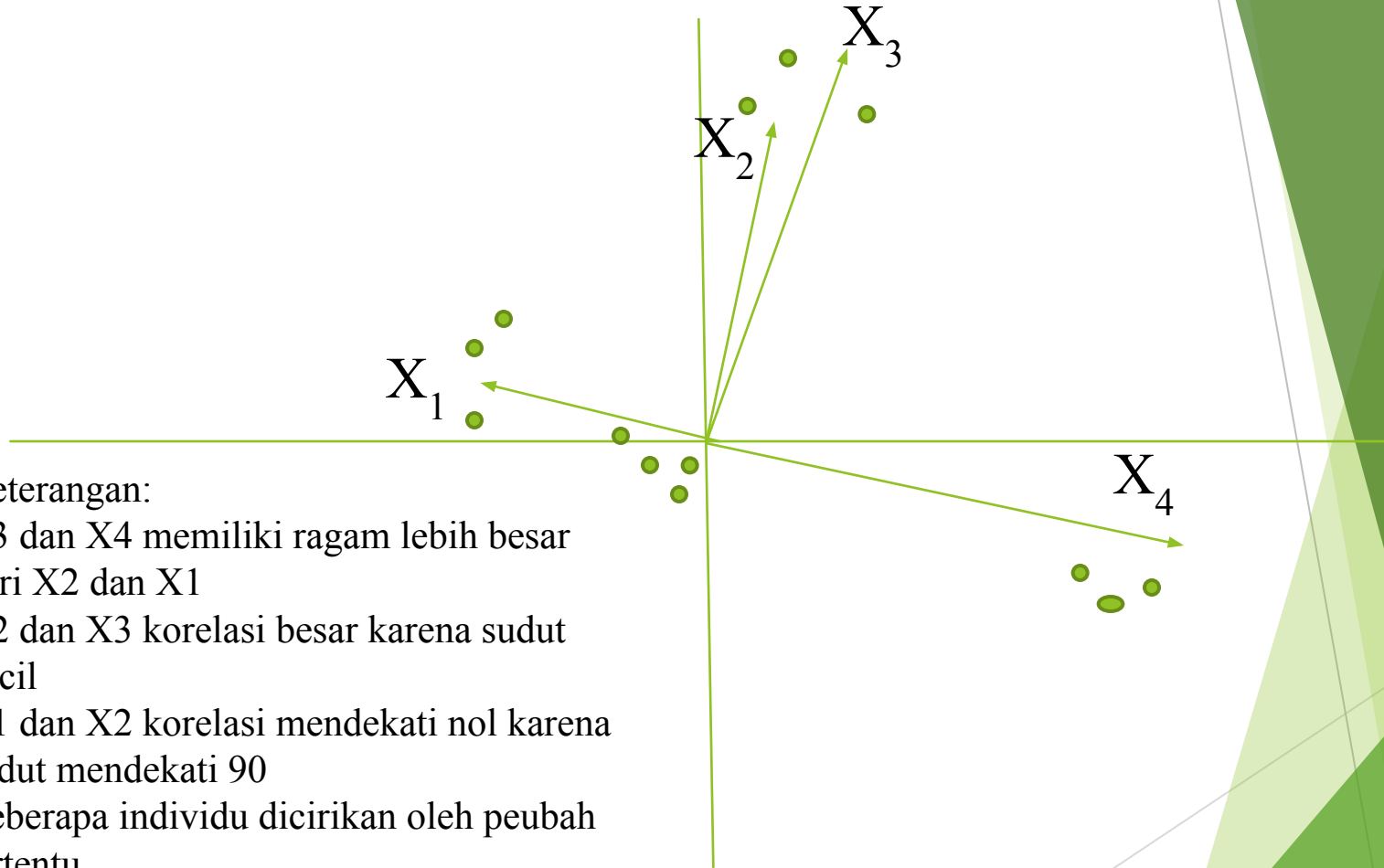
- Biplot adalah teknik statistika eksploratif yang dapat menyajikan secara visual secara simultan  $n$  obyek pengamatan dan  $p$  peubah dalam ruang berdimensi dua, sehingga ciri-ciri peubah dan obyek pengamatan serta posisi relatif antar obyek pengamatan dengan peubah dapat dianalisis.
- Analisis Biplot cocok digunakan apabila seluruh peubah amatan memiliki skala kuantitatif dan efektif apabila jumlah individu tidak terlalu banyak
- Keakuratan Analisis Biplot tergantung kepada berapa persen keragaman total yang dapat diterangkan oleh 2 komponen utama dalam biplot, karena pada prinsipnya biplot merupakan pereduksian dari ruang berdimensi besar ke ruang dimensi dua

# Informasi yang diperoleh dari Biplot

Analisis Biplot dapat menyajikan 4 jenis informasi:

1. Keragaman peubah
2. Keeratan hubungan antar peubah
3. Kemiripan relatif antar obyek pengamatan
4. Posisi relatif antar obyek pengamatan dengan peubah yang mengartikan bahwa objek/sekelompok objek dicirikan oleh peubah apa

# Bentuk Penampilan Biplot



Keterangan:

- $X_3$  dan  $X_4$  memiliki ragam lebih besar dari  $X_2$  dan  $X_1$
- $X_2$  dan  $X_3$  korelasi besar karena sudut kecil
- $X_1$  dan  $X_2$  korelasi mendekati nol karena sudut mendekati 90
- Beberapa individu dicirikan oleh peubah tertentu

# Struktur Data Amatan

Individu	Peubah				
	X1	X2	X3	...	Xp
1	x11	x12	x13		x1p
2	x21	x22	x23		x2p
3	x31	x32	x33		x3p
4	x41	x42	x43		x4p
5	x51	x52	x53		x5p
...	...	...	...	...	...
...	...	...	...	...	...
n	xn1	xn2	xn3		xnp

# Teori Biplot (1)

- Biplot merupakan teknik statistika deskriptif dimensi ganda yang berdasarkan pada penguraian nilai singular (PNS) atau *Singular Value Decomposition* (SVD).

Misalkan suatu matriks data  $\mathbf{X}$  berukuran  $n \times p$  yang berisi  $n$  pengamatan dan  $p$  peubah yang dikoreksi terhadap nilai rata-ratanya dan berpangkat  $r$ , dapat dituliskan menjadi

$$\mathbf{X} = \mathbf{U} \mathbf{L} \mathbf{A}'$$

- Matriks  $\mathbf{U}$  dan  $\mathbf{A}$  masing-masing berukuran  $(n \times r)$  dan  $(p \times r)$  sehingga  $\mathbf{U}'\mathbf{U} = \mathbf{A}'\mathbf{A} = \mathbf{I}_r$

$\mathbf{L}$  adalah matrik diagonal berukuran  $(r \times r)$  dengan unsur-unsur diagonalnya adalah akar kuadrat dari akar ciri  $\mathbf{X}'\mathbf{X}$  atau  $\mathbf{X}\mathbf{X}'$  sehingga

$$\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_r}$$

# Teori Biplot (2)

- Kolom matriks **A** adalah vektor ciri yang berpadanan dengan akar ciri  $\lambda$  dari matriks **X'X** atau **XX'**.
- Lajur-lajur matriks **U** dapat dihitung melalui :

$$U_i = \frac{1}{\sqrt{\lambda_i}} Xa_i$$

Dengan  $\lambda_i$  adalah akar ciri ke-i dari matriks **X'X** dan  $a_i$  adalah lajur ke-i matriks **A**.



□ Sehingga secara matematis SVD dapat ditulis:

$${}_n X_p = {}_n U_r {}_r L_r {}_r A_p$$

$$U = \left\{ \frac{1}{\sqrt{\lambda_1}} Xa_1, \frac{1}{\sqrt{\lambda_2}} Xa_2, \dots, \frac{1}{\sqrt{\lambda_r}} Xa_r \right\}$$

$$L = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \square & 0 \\ 0 & \sqrt{\lambda_2} & \square & 0 \\ \square & \square & \square & \square \\ 0 & 0 & \square & \sqrt{\lambda_r} \end{bmatrix}$$

$$A = [a_1, a_2, \dots, a_r]$$

□ Selanjutnya matriks  $X$  dapat ditulis:

$$\begin{aligned} \blacktriangleright X &= U L^{\alpha} L^{1-\alpha} A' \\ &= G H' \end{aligned}$$

dimana  $G = U L^{\alpha}$  serta  $H' = L^{1-\alpha} A'$

□ Unsur ke- $(i,j)$  matriks  $X$  dapat dituliskan sbb :

$$X_{ij} = g_i' h_j$$

dimana:  $i = 1, 2, 3, \dots, n$   $j = 1, 2, 3, \dots, p$

dengan  $g_i'$  dan  $h_j'$  masing-masing merupakan baris-baris matriks  $G$  dan  $H$

□ Jika  $r(X)=2$  maka  $g_i$  dan  $h_j$  digambarkan dalam ruang berdimensi 2

- ▶ Jika  $\alpha = 1$ , maka  $G = UL$  dan  $H = A$ , sehingga diperoleh hubungan:

$$\begin{aligned} XX' &= (GH')(GH')' \\ &= GH' HG' \\ &= GA' AG' \\ &= GG' \end{aligned}$$

- ▶ Jika  $\alpha = 0$  maka  $G = U$  dan  $H = AL$ , sehingga diperoleh

$$\begin{aligned} X'X &= (GH')'(GH') \\ &= HG' GH' \\ &= HU' UH' \\ &= HH' \end{aligned}$$

# Interpretasi dari Biplot

## 1. Kedekatan antar obyek.

Dua obyek dengan karakteristik sama akan digambarkan dalam posisi berdekatan.

## 2. Keragaman peubah.

Semakin panjang vector maka semakin besar keragaman suatu peubah.

## 3. Hubungan (korelasi) antar peubah

Jika sudut dua peubah  $< 90^0$  maka korelasi bersifat positif,

Jika sudut dua peubah  $> 90^0$  maka korelasi bersifat negatif

Jika sudut dua peubah  $= 90^0$  maka korelasi = nol

Semakin kecil sudutnya, maka semakin besar korelasinya.

## 3. Nilai peubah pada suatu obyek

Karakteristik suatu obyek bisa disimpulkan dari posisi relatifnya yang paling dekat dengan suatu peubah.

# Ukuran keragaman Biplot

- Besarnya keragaman yang diterangkan oleh biplot didefinisikan sebagai

$$\rho_2 = (\lambda_1 + \lambda_2) / \sum \lambda_k$$

keterangan:

$\lambda_1$  = Akar ciri terbesar pertama

$\lambda_2$  = Akar ciri terbesar kedua

$\lambda_i$  = Akar ciri terbesar ke- $i$  dari  $\mathbf{X}'\mathbf{X}$

$i = 1, 2, \dots, k.$

## Ilustrasi: Raw Data

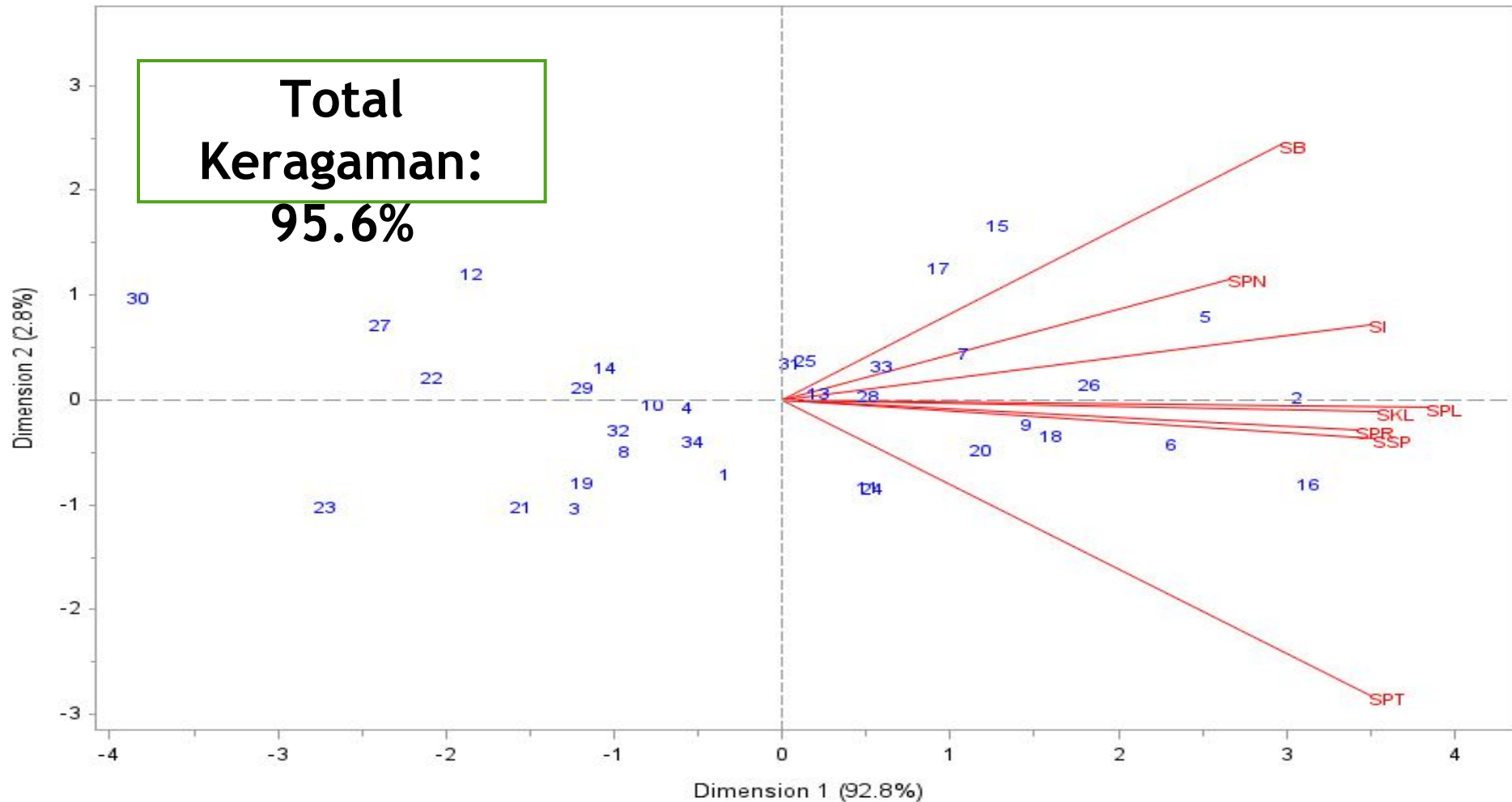
No	Provinsi	Nama Sekolah/Madrasah	Program Keahlian	SI	SPR	SKL	SPT	SSP	SPL	SB	SPN	Nilai Peringkat	Peringkat
1	Aceh	SMKN 1 lembah Sabil	Teknik komputer dan jaringan	58	63	60	60	64	58	60	59	60	C
2	Aceh	SMKS Hidayatul Anam	Budidaya Perkebunan	40	41	55	28	38	37	61	48	43	TT
3	Aceh	SMKN 1 Labuhan Haji Timur	Teknologi hasil pertanian	55	65	63	56	61	65	67	62	62	C
4	Aceh	SMKN 1 Kluet Timur	Teknik komputer dan jaringan	43	45	48	43	51	58	70	62	51	TT
...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...
8064	DI Yogyakarta	SMKS Ma'arif 2 Wates	Busana Butik	81	84	82	64	79	83	86	87	80	B
8065	DI Yogyakarta	SMKS Bina Harapan	Teknik komputer dan jaringan	87	92	76	66	74	85	89	88	82	B

# Ilustrasi

## ► Rata-rata per Provinsi

No	Provinsi	SI	SPR	SKL	SPT	SSP	SPL	SB	SPN
1	Aceh	81	80	79	76	79	75	81	82
2	Bali	94	91	91	85	89	90	92	93
3	Banten	77	76	78	73	75	74	77	79
4	Bengkulu	80	76	77	76	78	73	85	82
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...
33	Sumatera Selatan	83	82	81	79	80	79	89	86
34	Sumatera Utara	78	77	79	75	77	74	82	83

# Hasil Biplot





# Kode Provinsi

Label	Provinsi	Label	Provinsi	Label	Provinsi
1	Aceh	13	Kal. Selatan	25	Papua Barat
2	Bali	14	Kal. Tengah	26	Riau
3	Banten	15	Kal. Timur	27	Sulawesi Barat
4	Bengkulu	16	Kal. Utara	28	Sulawesi Selatan
5	DI Yogyakarta	17	Kep. B. Belitung	29	Sulawesi Tengah
6	DKI Jakarta	18	Kepulauan Riau	30	Sulawesi Tenggara
7	Gorontalo	19	Lampung	31	Sulawesi Utara
8	Jambi	20	Maluku	32	Sumatera Barat
9	Jawa Barat	21	Maluku Utara	33	Sumatera Selatan
10	Jawa Tengah	22	NTB	34	Sumatera Utara
11	Jawa Timur	23	NTT		
12	Kal. Barat	24	Papua		

# Interpretasi Hasil

- ▶ Prov Bali (2), Kalimantan Utara (16), DI Yogyakarta (5), dan DKI Jakarta (6) memiliki nilai rata-rata total skor delapan SNP yang paling tinggi untuk semua peubah
- ▶ Provinsi Sulawesi Tenggara (30), NTT (23), Sulawesi Barat (27), NTB (22), dan Kalimantan Barat (12) memiliki nilai rata-rata total skor yang rendah atau jauh di bawah nilai rata-rata total skor.
- ▶ Provinsi Kalimantan Timur (15), Kepulauan Bangka Belitung (17), Papua Barat (25), Sulawesi Utara (31), Kalimantan Selatan (13), dan Sumatera Barat (33) dicirikan oleh peubah standar pembiayaan (SB).
- ▶ Provinsi DIY(5), Gorontalo (7), Sumatera Barat (33), dan Kalimantan Selatan (13) adalah provinsi-provinsi yang dicirikan oleh peubah standar penilaian (SPN).
- ▶ Provinsi Bali (2), Riau (26), Sulawesi Selatan (28), dan Kalimantan Selatan (13) dicirikan oleh peubah standar pengelolaan (SPL) dan standar kompetensi lulusan (SKL).

# Matriks Korelasi Antar Peubah

	SI	SPR	SKL	SPT	SSP	SPL	SB	SPN
SI	1.000							
SPR	0.957	1.000						
SKL	0.936	0.969	1.000					
SPT	0.864	0.911	0.896	1.000				
SSP	0.923	0.943	0.959	0.924	1.000			
SPL	0.916	0.932	0.950	0.882	0.935	1.000		
SB	0.894	0.866	0.877	0.791	0.901	0.863	1.000	
SPN	0.912	0.922	0.944	0.832	0.936	0.942	0.933	1.000

Terimakasih