

BIPLOT DENGAN DEKOMPOSISI NILAI SINGULAR BIASA DAN KEKAR UNTUK PEMETAAN PROVINSI BERDASARKAN PRESTASI MAHASISWA

Warsito*, Hairul Saleh

Universitas Muhammadiyah Tangerang, Jl. Perintis Kemerdekaan I/33 Tangerang 151818 Indonesia

*Corresponding Author: warsito@umt.ac.id

ARTICLE INFO

ABSTRACT

Article history:

Received: 4 Desember 2022
Revised: 26 Desember 2022
Accepted: 11 Januari 2023

Keywords:

dekomposisi nilai singular, biplot, kesuaian biplot (*goodness of fit*), pencilan, data ekstrim.

Analisis Biplot dapat dikonstruksi dengan pendekatan DNS biasa dan kekar. Pendekatan DNS biasa memerlukan matriks data yang tidak ada pencilan atau data ekstrim. Apabila dalam penelitian ini ditemukan data pencilan maka biplot dari DNS biasa belum menjamin gambaran pemetaan antara objek pengamatan dan peubahnya. Untuk menyelesaikan masalah ini, diberikan pendekatan metode DNS kekar. Pasangan eigennilai dan eigenvektor dari DNS biasa dan kekar mempunyai sifat meminimalkan jarak antara matriks data dengan matriks dugaannya dengan norma Euclid atau norma L2 dan jarak blok kota (*city block*) atau norma L1. Analisis Biplot berdasarkan dari metode DNS kekar dapat memberikan gambaran objek pengamatan dan peubah yang tahan terhadap pencilan. Metode diagram kotak dari data penelitian prestasi mahasiswa IPB menunjukkan bahwa data memiliki pencilan. Analisis Biplot yang didasarkan DNS biasa dan kekar terhadap matriks data tersebut memperlihatkan grafik dan struktur data hampir sama, hal ini menunjukkan bahwa pencilan tidak berpengaruh pada perubahan struktur data. Matriks data diberi beberapa data ekstrim kemudian dianalisis dengan DNS biasa dan kekar. Biplot dari DNS biasa hasilnya berbeda jauh dengan biplot dari matriks data tanpa data ekstrim, sedangkan biplot dari DNS kekar memperlihatkan struktur data kecil atau tidak terpengaruh dengan data ekstrim.

Biplot analysis can be constructed from an ordinary and robust Singular Value Decomposition (SVD) approach. A SVD approach ordinary need data matrix that no outliers or extreme data. If a outliers data is found in this research, ordinary SVD biplot will not able to assure the mapping illustration between observation object and variable. To solve this problem, approaching method to robust SVD is applied. The couple of ordinary and robust SVD' eigenvalue couple and eigenvector have distance minimizing feature among data matrix and estimation matrix with Euclid norm or L2 norm and city block distance or L1 norm. Biplot analysis which is based on robust SVD method can give illustrations to a outliers resist observation object and variable. The boxplot of the achievement of IPB show that the data has outliers. Robust and ordinary SVD to data matrix based biplot analysis shows the data graph and structure is almost identic. It implies that outliers has no effect to data structure alteration. A couple of extreme datas are given to the data matrix and then analysed using ordinary and robust SVD. Ordinary SVD biplot result is far more different from data matrix biplot without extreme data, in

the other hand robust SVD biplot shows small data structure which means it is not influenced by extreme data.

Copyright © 2023 Universitas Siliwangi
All rights reserved

How to Cite:

Warsito & Saleh, H. (2023). Biplot Dengan Dekomposisi Nilai Singular Biasa dan Kekar Untuk Pemetaan Provinsi Berdasarkan Prestasi Mahasiswa. *Journal of Authentic Research on Mathematics Education*, 5(1), 33-45. <https://doi.org/10.37058/jarme.v5i1.6004>

1. PENDAHULUAN

Analisis Peubah Ganda (APG) merupakan bentuk lain dari aljabar linear terapan dalam matematika, misalkan \mathbf{X} matriks berukuran $n \times p$ dapat menjelaskan dengan adanya n objek dan masing-masing objek tersebut diamati p peubah (Siswadi dan Suharjo, 2007, Hardle, WK. & Simar, L., 2015; Mattjik AA, & Sumertajaya IM., 2011). Analisis biplot merupakan suatu teknik eksplorasi peubah ganda yang menyajikan posisi n objek dengan p peubah secara simultan pada dua dimensi (Jolliffe, 2002; Daigle, G, & Riverst, LP., 1992; Mattjik AA & Sumertajaya IM., 2011). Data pencilan merupakan pengamatan tampak tidak konsisten dengan data pengamatan lain (Neter, J., Wasserman, W., & Kutner, 1990). Untuk mengidentifikasi adanya pencilan salah satunya dapat dilakukan dengan boxplot. Adanya data pencilan berakibat akan mengganggu pemenuhan hasil analisis dari model yang dihasilkan tidak dapat diandalkan. Demikian juga dalam kasus multivariat, hasilnya adalah interpretasi yang tidak akurat dan kesalahan dalam pengambilan keputusan atas model yang diperoleh (Huber & Ronchetti, 2011; Utami, *et al.*, 2017). Oleh karena itu, perlu analisis lain yang kekar terhadap data pencilan.

Analisis biplot dapat dikonstruksi dengan pendekatan dekomposisi nilai singular (DNS) biasa dan kekar (Liu *et al*, 2003). Pendekatan DNS biasa memerlukan matriks data tanpa pencilan atau data ekstrim. Pasangan eigennilai dan eigenvektor dari DNS biasa mempunyai sifat meminimalkan jarak antara matriks data dengan matriks dugaannya dengan norma Euclid atau norma L2 (Johnson & Wichern, 2002)

Apabila dalam suatu penelitian ditemukan data pencilan, biplot dengan DNS biasa belum menjamin gambaran pemetaan antara objek pengamatan dan peubah oleh sebab itu perlu digunakan pendekatan DNS kekar (Liu *et al*, 2003; Hawkins, *et al*, 2001). Pendekatan ini digunakan untuk menduga sejumlah eigennilai dan eigenvektor kiri dan kanan yang tahan terhadap pengaruh pencilan. Pasangan eigennilai dan eigenvektor dari DNS kekar mempunyai sifat meminimalkan jarak antara matriks data dengan matriks dugaannya dengan norma blok kota (*city block*) atau norma L1 (Hawkins, *et al*, 2001). Analisis biplot yang dihasilkan dengan pendekatan DNS kekar diharapkan dapat memberikan gambaran objek pengamatan dan peubah yang tahan terhadap pencilan.

Mahasiswa IPB hampir mewakili 33 provinsi di Indonesia, diharapkan mampu memberikan gambaran prestasi dan pemetaan mutu pendidikan setiap daerahnya. Pencapaian prestasi tersebut salah satunya dipengaruhi oleh mutu masukan, di mana seleksi penerimaan mahasiswa baru IPB dapat melalui jalur USMI, SNMPTN, dan BUD.

Keragaman mutu yang diperoleh memungkinkan terdapat pencilan, sehingga untuk mendapatkan pemetaan provinsi berdasarkan peubah mata kuliah dan IPK mendorong untuk membandingkan analisis biplot dengan pendekatan DNS biasa dan kekar. DNS biasa akan memberikan biplot yang memvisualisasikan dari segugus objek dan peubah dalam bentuk grafik bidang datar sehingga ciri-ciri peubah dan objek pengamatan serta posisi relatif antar objek dengan peubah dapat dianalisis. Apabila dalam suatu penelitian ditemukan data pencilan, biplot dengan DNS biasa belum menjamin gambaran pemetaan antara objek pengamatan dan peubah oleh sebab itu diberikan sebuah pendekatan DNS secara iteratif yang disebut pendekatan DNS kekar.

DNS dari matriks data adalah suatu alat yang dapat digunakan untuk memahami struktur data. Misalkan \mathbf{X} adalah matriks data dengan n objek pengamatan dan p peubah yang terkoreksi terhadap rata-ratanya. Jika matriks \mathbf{X} berpangkat r dengan $r \leq \min \{n, p\}$, maka dengan menggunakan DNS biasa (Siswadi dan Suharjo, 2007; Gabriel 2002) diperoleh:

$${}_n\mathbf{X}_p = {}_n\mathbf{U}_r\mathbf{L}_r\mathbf{A}'_p$$

Matriks \mathbf{U} dan \mathbf{A} merupakan matriks ortonormal kolom, di mana $\mathbf{U}'\mathbf{U} = \mathbf{A}'\mathbf{A} = \mathbf{I}_r$. Matriks \mathbf{A} adalah matriks yang kolom-kolomnya terdiri eigenvektor \mathbf{a}_i yang berpadanan dengan eigennilai λ_i dari matriks $\mathbf{X}'\mathbf{X}$ dan matriks \mathbf{U} adalah:

$$\mathbf{U} = \left(\frac{\mathbf{x}\mathbf{a}_1}{\sqrt{\lambda_1}}, \frac{\mathbf{x}\mathbf{a}_2}{\sqrt{\lambda_2}}, \dots, \frac{\mathbf{x}\mathbf{a}_r}{\sqrt{\lambda_r}} \right).$$

Matriks $\mathbf{L} = \text{diag} (\sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_r})$, di mana $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$ dan λ_i disebut nilai singular dari matriks $\mathbf{X}'\mathbf{X}$ atau matriks $\mathbf{X}\mathbf{X}'$. DNS biasa juga dapat ditulis dalam bentuk:

$$\mathbf{X} = \sum_{i=1}^r \lambda_i \mathbf{u}_i \mathbf{a}'_i$$

Bila $r > 2$ maka pendekatan terbaik matriks data \mathbf{X} oleh \mathbf{Y} yang berpangkat s , dengan $s < r$ diperoleh dari kuadrat jarak minimum matriks \mathbf{Y} ke matriks \mathbf{X} yaitu:

$$\min \|\mathbf{X} - \mathbf{Y}\|_F^2 = \min \sum_{i=1}^n \sum_{j=1}^p (x_{ij} - y_{ij})^2$$

Dengan DNS biasa maka:

$${}_n\mathbf{X}_p = {}_n\mathbf{U}_r\mathbf{L}_r\mathbf{A}'_p.$$

Dalam Jolliffe (2002), dengan mendefinisikan $\mathbf{G} = \mathbf{U}\mathbf{L}^\alpha$ dan $\mathbf{H}' = \mathbf{L}^{1-\alpha}\mathbf{A}'$ untuk $\alpha \in [0,1]$, maka ${}_n\mathbf{X}_p = {}_n\mathbf{U}_r\mathbf{L}_r\mathbf{A}'_p$

$$= {}_n\mathbf{U}_r\mathbf{L}_r^\alpha\mathbf{L}_r^{1-\alpha}\mathbf{A}'_p = {}_n\mathbf{G}_r\mathbf{H}'_p$$

Untuk menggambarkan di ruang berdimensi $k < r$, maka pendekatan matriks berpangkat k :

$$\begin{aligned} \hat{\mathbf{X}}_{(k)} &= \mathbf{G}_{(k)}\mathbf{H}'_{(k)} \\ &= (\mathbf{U}_k\mathbf{L}_k^\alpha)(\mathbf{L}_k^{1-\alpha}\mathbf{A}'_k) \end{aligned}$$

biasanya digunakan $k = 2$, sehingga koordinat-koordinat \mathbf{G} dan \mathbf{H} dapat digambarkan dalam ruang berdimensi 2 (Lipkovich & Smith, 2002). Pengambilan α tertentu berimplikasi dalam interpretasi biplot. Jika $\alpha = 0$, maka diperoleh $\mathbf{G} = \mathbf{U}$ dan $\mathbf{H}' = \mathbf{L}\mathbf{A}'$ akibatnya:

$$\begin{aligned} \mathbf{X}'\mathbf{X} &= (\mathbf{G}\mathbf{H}')'(\mathbf{G}\mathbf{H}') \\ &= \mathbf{H}\mathbf{H}' \\ &= (n-1)\mathbf{S} \end{aligned}$$

sehingga diperoleh:

$$\checkmark \mathbf{h}'_i\mathbf{h}_j = (n-1)s_{ij}, \text{ di mana } s_{ij} \text{ adalah koragam peubah ke-}i \text{ dan ke-}j.$$

- ✓ Panjang $\|h_i\| = \sqrt{n-1} s_i$ dengan $s_i = \sqrt{s_{ii}}$, menggambarkan keragaman peubah ke- i .
- ✓ Korelasi antara peubah ke- i dan ke- j dijelaskan oleh cosinus sudut antara h_i dan h_j , yaitu: $\cos \theta = r_{ij}$.
- ✓ Jika \mathbf{X} berpangkat p , maka $(x_i - x_j)' \mathbf{S}^{-1} (x_i - x_j) = (n-1)(g_i - g_j)'(g_i - g_j)$ artinya, kuadrat jarak Mahalanobis antara x_i dan x_j sebanding kuadrat jarak Euclid antara g_i dan g_j .

Untuk $\alpha \in (0,1)$, maka interpretasi pada korelasi serta jarak Euclid dan Mahalanobis tidak berlaku, sedangkan posisi relatif g_i dan h_j masih mencerminkan mengenai besaran objek ke- i pada peubah ke- j , $x_{ij} = g_i' h_j$

Analisis biplot yang didasarkan DNS kekar, mengambil pendekatan matriks \mathbf{X} berpangkat dua yaitu :

$$\mathbf{X} \approx \mathbf{L}\mathbf{R}'$$

dengan \mathbf{L} dan \mathbf{R} adalah matriks yang terdiri dua eigenvektor kiri dan eigenvektor kanan pertama pada \mathbf{X} (Hawkins *et al.*, 2001). Kemudian, matriks \mathbf{L} dan \mathbf{R} digunakan sebagai matriks \mathbf{G} dan \mathbf{H} yang masing-masing merupakan gambarkan vektor-vektor baris dan kolom matriks \mathbf{X} .

2. METODE

Pada dasarnya, analisis ini merupakan suatu alat statistika yang menyajikan posisi relatif n objek dengan p peubah secara simultan dalam dua dimensi. Analisis ini dapat mengkaji hubungan antara objek pengamatan dan peubah. Selain itu dapat dilihat juga ciri-ciri masing-masing objek dan peubahnya. Oleh karena itu, metode penelitian ini merupakan bentuk statistik deskriptif yang untuk melihat pemetaan peubah dan objek secara bersama-sama pada demensi dua dengan menggunakan analisis Biplot. Menurut Gabriel (2002), biplot tidak hanya sebagai pendekatan matriks data \mathbf{X} dengan menggunakan matriks \mathbf{GH}' , tetapi juga hasil perkalian \mathbf{HH}' sebagai pendekatan dari matriks $\mathbf{X}'\mathbf{X}$ yang berkaitan dengan ragam-koragam dan korelasi antar peubah, dan matriks \mathbf{GG}' sebagai pendekatan bagi \mathbf{XX}' yang berkaitan dengan ukuran kemiripan antar objek. Selanjutnya Gabriel mengemukakan ukuran kesuaian biplot (*Goodness of Fit of Biplot*) sebagai ukuran pendekatan dalam bentuk sebagai berikut:

- 1) Kesuaian data: $GF(\mathbf{X}, \mathbf{GH}') = \frac{\text{tr}^2(\mathbf{X}'\mathbf{GH}')}{\text{tr}(\mathbf{X}'\mathbf{X})(\mathbf{HG}'\mathbf{GH}'')}$
- 2) Kesuaian peubah: $GF(\mathbf{X}'\mathbf{X}, \mathbf{HH}') = \frac{\text{tr}^2(\mathbf{X}'\mathbf{XHH}')}{\text{tr}(\mathbf{X}'\mathbf{XX}'\mathbf{X})\text{tr}(\mathbf{HH}'\mathbf{HH}'')}$
- 3) Kesuaian objek: $GF(\mathbf{XX}', \mathbf{GG}') = \frac{\text{tr}^2(\mathbf{XX}'\mathbf{GG}')}{\text{tr}(\mathbf{XX}'\mathbf{XX}')\text{tr}(\mathbf{GG}'\mathbf{GG}'')}$

2.1. Subjek Penelitian

Peubah dan objek yang digunakan dalam penelitian ini adalah:

Tabel 1. Nama peubah mata kuliah

No.	Peubah	Kode	No.	Peubah	Kode
1	Agama (AGM)	V1	9	Pengantar Kewirausahaan (KWR)	V9
2	Biologi (BIO)	V2	10	Pengantar Matematika (MTK)	V10
3	Ekonomi Umum (EKU)	V3	11	Olah Raga dan Seni (ORS)	V11
4	Fisika (FIS)	V4	12	Pengantar Ilmu Pertanian (PIP)	V12
5	Bahasa Indonesia (IND)	V5	13	Pengantar Kewarganegaraan (PKN)	V13
6	Bahasa Inggris (ING)	V6	14	Sosiologi Umum (SOU)	V14
7	Kalkulus (KAL)	V7	15	Indeks Prestasi Kumulatif (IPK)	V15
8	Kimia (KIM)	V8			

Nilai peubah $V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, V_{10}, V_{11}, V_{12}, V_{13}, V_{14}$ merupakan rata-rata nilai mutu mata kuliah dan V_{15} merupakan rata-rata nilai IPK mahasiswa.

Tabel 2. Nama objek pengamatan asal provinsi

Asal Provinsi	Seleksi	Kode	Asal Provinsi	Seleksi	Kode
NAD 1	Non-BUD	1	IATIM 1	Non BUD	28
NAD 2	BUD	2	IATIM 2	BUD	29
SUMUT 1	Non BUD	3	BALI	Non BUD	30
SUMUT 2	BUD	4	NTB	Non BUD	31
SUMBAR 1	Non BUD	5	NTT 1	Non BUD	32
SUMBAR 2	BUD	6	NTT 2	BUD	33
RIAU 1	Non BUD	7	KALBAR	Non BUD	34
RIAU 2	BUD	8	KALTENG 1	Non BUD	35
IAMBI 1	Non BUD	9	KALTENG 2	BUD	36
IAMBI 2	BUD	10	KALSEL 1	Non BUD	37
SUMSEL 1	Non BUD	11	KALSEL 2	BUD	38
SUMSEL 2	BUD	12	KALTIM 1	Non BUD	39
BENGKULU	Non BUD	13	KALTIM 2	BUD	40
LAMPUNG 1	Non BUD	14	SULUT	Non BUD	41
LAMPUNG 2	BUD	15	SULSEL 1	Non BUD	42
KEP.BABEL. 1	Non BUD	16	SULSEL 2	BUD	43
KEP.BABAL. 2	BUD	17	SULTRA 1	Non BUD	44
DKI IAKARTA 1	Non BUD	18	SULTRA 2	BUD	45
DKI IAKARTA 2	BUD	19	SULTENG 1	Non BUD	46
IABAR 1	Non BUD	20	SULTENG 2	BUD	47
IABAR 2	BUD	21	GORONTALO	Non BUD	48
BANTEN 1	Non BUD	22	MALUKU 1	Non BUD	49
BANTEN 2	BUD	23	MALUKU 2	BUD	50
IATENG 1	Non BUD	24	MALUT 1	Non BUD	51
IATENG 2	BUD	25	MALUT 2	BUD	52
DIY 1	Non BUD	26	PAPUA 1	Non BUD	53
DIY 2	BUD	27	PAPUA 2	BUD	54

2.2. Pengumpulan Data

Penelitian ini menggunakan data sekunder dari nilai mata kuliah dan IPK mahasiswa Tingkat Persiapan Bersama (TPB) IPB Bogor tahun 2008. Pelaksanaan penelitian ini melibatkan semua mahasiswa TPB yang terdiri 3001 mahasiswa yang dikelompokkan berdasarkan provinsi dan hasil seleksi masuk IPB, yaitu melalui jalur Biasiswa Utusan Daerah (BUD) atau non BUD. Hasil matriks data peubah ganda berukuran 54x15 yang menunjukkan 24 provinsi asal daerah mahasiswa BUD dan 30 provinsi mahasiswa non BUD. Setiap provinsi diamati rata-rata nilai mutu dari 14 mata kuliah dan nilai IPKnya.

2.3. Analisis Data

Data dianalisis dengan dua pendekatan. Pendekatan I menentukan biplot dengan DNS biasa. Pendekatan II menentukan biplot dengan DNS kekar (menggunakan paket RobustBiplotPack Versi 1.1, Ardana (2009) *software Mathematica 6.0*).

Sebagai gambaran bahwa biplot dengan DNS kekar lebih tahan terhadap data pencilan, diberikan data ekstrim pada peubah mata kuliah Biologi dari objek provinsi NAD (non BUD) dan peubah mata kuliah Pengantar Matematika dari objek provinsi PAPUA (non BUD) yang masing-masing sebesar 7.50 dan 10.50. Hasil biplot kemudian dibandingkan hasil biplot dengan metode biasa. Jika analisis data dengan metode biasa dan kekar memberikan hasil yang tidak jauh beda, maka data yang telah diberi beberapa data ekstrim dianalisis dengan pendekatan DNS biasa dan kekar, hasilnya dibandingkan dengan analisis data awal.

Pendekatan I dengan DNS biasa

- a. Transformasi matriks data ke bentuk matriks koragam.
- b. Analisis dengan menggunakan paket Biplot versi 3.2, dengan memilih nilai $\alpha = 0$, Ardana (2009).
- c. Menelusuri ketepatan biplot dengan menggunakan ukuran kesuaian dari Gabriel (2002).

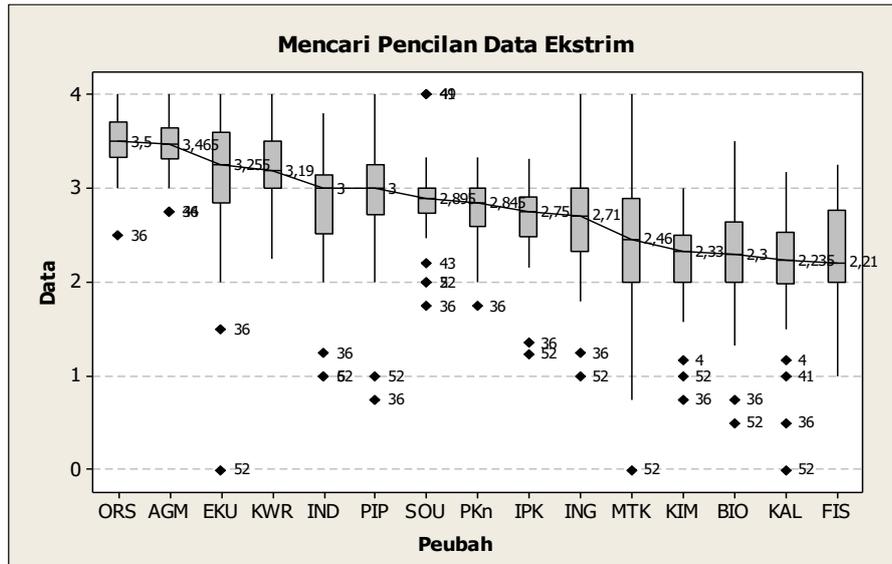
Pendekatan II dengan DNS kekar

- a. Transformasi matriks data ke bentuk matriks koragam yang terpusatkan terhadap median.
- b. Analisis dengan menggunakan paket BiplotRobustPack versi 1.0, Ardana (2009) *software Mathematica 6.0* dengan memilih nilai $\alpha = 0$.
- c. Menelusuri ketepatan biplot dengan menggunakan ukuran kesuaian dari Gabriel (2002).

3. HASIL DAN PEMBAHASAN

3.1. Hasil

Diagram kotak atau *box plot* merupakan salah teknik untuk memberikan gambaran tentang lokasi pemusatan data, rentangan penyebaran, dan kemiringan pola sebaran. Diagram kotak hasil penelitian memberikan pencilan yang disajikan pada Gambar 1.



Gambar 1. Diagram kotak data prestasi mahasiswa IPB

Berdasarkan Gambar 1 terdapat 21 pencilan yang terdiri dari 17 pencilan minor dan 4 data ekstrim. Data ekstrim ada di peubah V_3 (Ekonomi Umum), V_7 (Kalkulus) dan V_{14} (Sosiologi Umum) pada objek 52 (KALSEL), 41 (SULUT) dan 49 (MALUKU 1). Berdasarkan data asal nilai objek 52, 41 dan 49 masing-masing 0.00, 4.00, dan 4.00. Dari Gambar 1, juga diperoleh gambaran bahwa peubah V_3 (Ekonomi Umum), V_4 (Fisika) dan V_{10} (Pengantar Matematika) mempunyai keragaman yang paling besar, sedangkan peubah V_{14} (Sosiologi Umum) mempunyai keragaman yang paling kecil.

Pemetaan Prestasi Provinsi

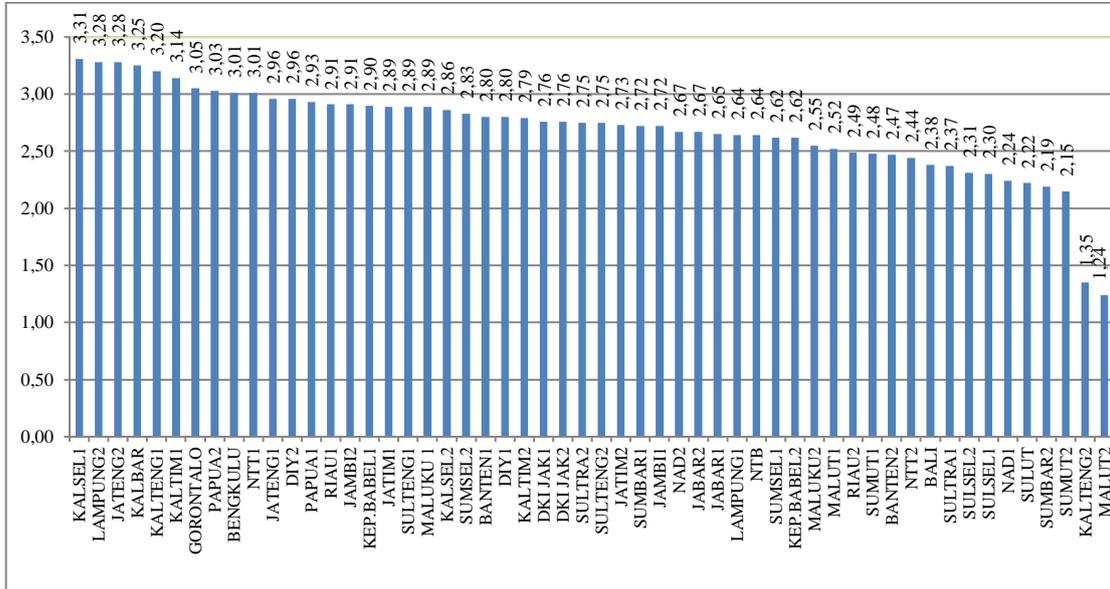
Indikator prestasi mahasiswa biasanya dikaitkan dengan pencapaian prestasi nilai mutu tiap mata kuliah dan IPK. Korelasi antar peubah dapat dilihat pada Tabel 3.

Tabel 3. Matriks Korelasi Pearson

Peubah	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	V_{11}	V_{12}	V_{13}	V_{14}	V_{15}
V_1	1														
V_2	0.51	1													
V_3	0.48	0.75	1												
V_4	0.01	0.54	0.43	1											
V_5	0.57	0.69	0.67	0.40	1										
V_6	0.40	0.70	0.59	0.62	0.74	1									
V_7	0.24	0.72	0.81	0.73	0.58	0.67	1								
V_8	0.41	0.75	0.70	0.68	0.60	0.68	0.86	1							
V_9	0.57	0.50	0.51	-	0.67	0.48	0.32	0.41	1						
V_{10}	0.41	0.70	0.82	0.02	0.57	0.57	0.85	0.81	0.38	1					
V_{11}	0.20	0.33	0.04	0.62	0.22	0.41	0.24	0.24	0.06	0.16	1				
V_{12}	0.44	0.68	0.75	0.24	0.73	0.63	0.69	0.62	0.58	0.70	0.22	1			
V_{13}	0.42	0.45	0.50	0.41	0.42	0.40	0.50	0.55	0.20	0.53	0.32	0.51	1		
V_{14}	0.50	0.54	0.42	0.41	0.53	0.51	0.34	0.44	0.68	0.24	0.30	0.55	0.23	1	
V_{15}	0.47	0.71	0.72	0.12	0.69	0.73	0.74	0.75	0.56	0.72	0.20	0.63	0.53	0.48	1

Berdasarkan Tabel 3 korelasi terbesar terjadi antara peubah V_7 (Kalkulus) dan V_8 (Kimia) yaitu sebesar 0.86. Korelasi peubah V_{15} (IPK) dengan peubah V_8 (Kimia), V_7

(Kalkulus), V_6 (Bahasa Inggris), V_{10} (Pengantar Matematika), V_3 (Ekonomi Umum), V_4 (Fisika), V_2 (Biologi) dan V_5 (Bahasa Indonesia) sebesar 0.75, 0.74, 0.73, 0.72, 0.72, 0.72, 0.71, dan 0.69. Nilai korelasi tersebut menunjukkan bahwa rata-rata IPK yang dicapai mahasiswa sangat terkait oleh nilai mata kuliah yang bersifat eksak dan bahasa.

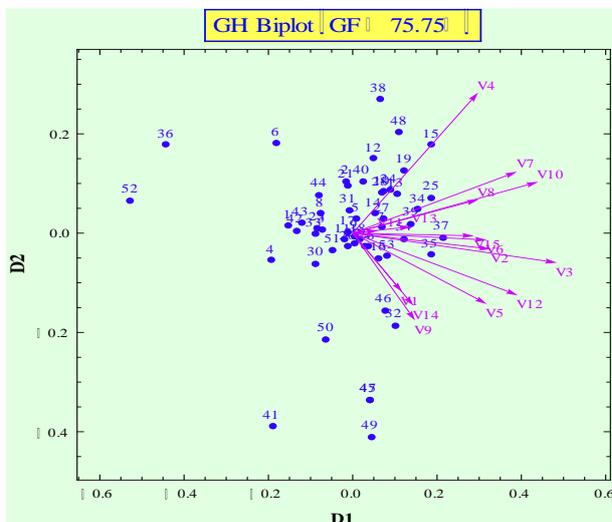


Gambar 2. Peringkat provinsi berdasarkan rata-rata IPK

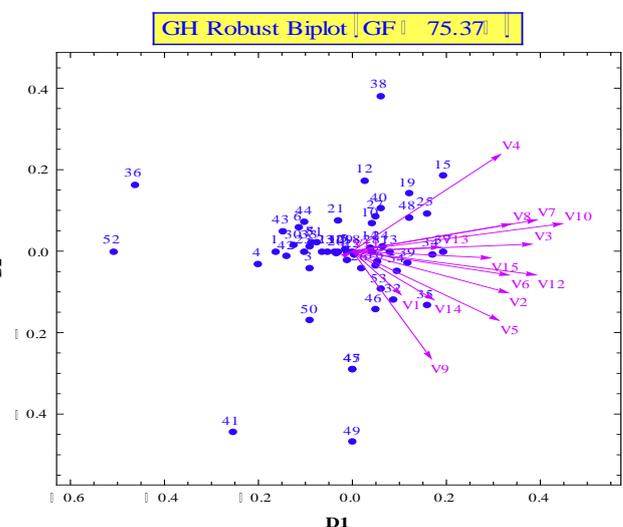
Berdasarkan Gambar 2 sepuluh besar provinsi yang mendapat peringkat IPK tertinggi didominasi oleh provinsi di luar pulau Jawa. Provinsi Kalimantan Tengah dan Maluku Utara mendapatkan peringkat terendah dari semua provinsi.

Analisis Biplot dengan DNS Biasa dan Kekar

Analisis Biplot didasarkan pada DNS biasa dan kekar dari matriks data yang sudah terkoreksi terhadap rata-rata dan median, yaitu menggunakan paket Biplot versi 3.2 dan paket Robust BiplotPack versi 1.1, Ardana (2009) *software Mathematica* 6.0 dengan memilih $\alpha = 0$. Hasil biplot yang diperoleh disajikan pada Gambar 3 dan 4.



Gambar 3. Biplot biasa



Gambar 4. Biplot kekar

Tabel 4. Ukuran kesuaian biplot (%)

Kesuaian		DNS Biasa	DNS Kekar
GF	Data Peubah	75.75	75.37
	Objek	98.16	96.32
		64.17	64.20

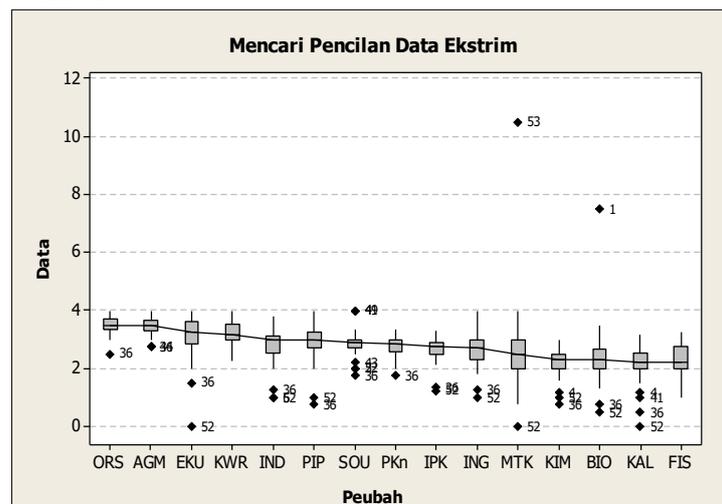
Beberapa hasil biplot biasa dan kekar pada Gambar 3 dan Gambar 4 dapat diperoleh:

- Ragam dan korelasi peubah dari hasil biplot berdasarkan DNS biasa dan kekar relatif sama dengan hasil yang diperoleh dari diagram kotak dan matriks korelasi.
- Posisi objek 37 (KALSEL) terletak paling kanan menunjukkan prestasi terbaik secara menyeluruh dalam semua mata kuliah serta nilai IPK.
- Semua objek yang mengelompok di sebelah kanan memiliki nilai IPK di atas rata-rata yaitu objek ke: 7, 10, 12, 13, 14, 15, 16, 19, 22, 24, 25, 26, 27, 28, 32, 34, 35, 37, 38, 39, 40, 46, 48, 53, dan 54.
- Objek yang mengelompok di sebelah kiri secara umum memiliki nilai di bawah rata-rata nilai IPK yaitu objek ke : 1, 2, 3, 4, 6, 8, 11, 17, 20, 21, 23, 30, 31, 33, 36, 41, 42, 43, 44, 50, 51, dan 52. Berdasarkan Gambar 3 dan 4 yaitu objek 52 (MALUT 2) dan objek 36 (KALTENG 2) adalah provinsi dengan prestasi paling rendah.
- Berdasarkan Tabel 4 dapat dilihat bahwa dengan memilih nilai $\alpha = 0$ kedua biplot mempunyai pendekatan matriks data, matriks peubah dan matriks objek yang tidak jauh berbeda.

Hasil analisis di atas menunjukkan bahwa pencilan yang ada tidak berpengaruh atau kecil pengaruhnya dalam perubahan konfigurasi data atau gambaran pemetaan.

Eksplorasi Data Ekstrim

Dengan diagram kotak pada data baru memberikan pencilan yang lebih ekstrim yang disajikan pada Gambar 5.



Gambar 5. Diagram kotak data ekstrim

Berdasarkan Gambar 5, objek 1 (NAD 1) dan 53 (PAPUA 1) pada peubah V_2 (Biologi) dan peubah V_{10} (Pengantar Matematika) menunjukkan sebagai data ekstrim, sebelumnya kedua objek tersebut bukan merupakan pencilan. Gambar 5, juga diperoleh gambaran

bahwa peubah V_3 (Ekonomi Umum), V_4 (Fisika) dan V_{10} (Pengantar Matematika) mempunyai keragaman yang paling besar, sedangkan peubah V_{14} (Sosiologi Umum) mempunyai keragaman yang paling kecil.

Tabel 5. Matriks korelasi Pearson

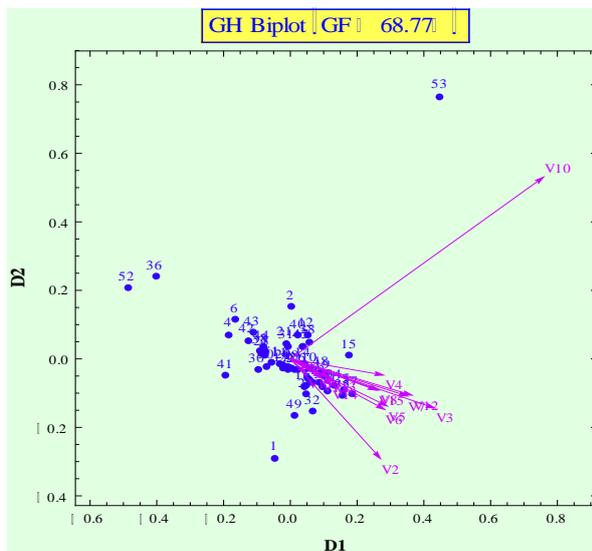
Peubah	V_1	V_2	V_3	V_4	V_5	V_6	V_7	V_8	V_9	V_{10}	V_{11}	V_{12}	V_{13}	V_{14}	V_{15}
V_1	1														
V_2	0.17	1													
V_3	0.48	0.29	1												
V_4	0.01	0.24	0.43	1											
V_5	0.57	0.35	0.67	0.40	1										
V_6	0.40	0.34	0.59	0.62	0.74	1									
V_7	0.24	0.33	0.81	0.73	0.58	0.67	1								
V_8	0.41	0.33	0.70	0.68	0.60	0.68	0.86	1							
V_9	0.57	0.27	0.51	-0.02	0.67	0.48	0.32	0.42	1						
V_{10}	0.21	0.18	0.47	0.42	0.40	0.36	0.49	0.46	0.29	1					
V_{11}	0.20	0.20	0.04	0.24	0.22	0.41	0.24	0.24	0.06	0.08	1				
V_{12}	0.44	0.28	0.75	0.41	0.73	0.63	0.69	0.62	0.58	0.49	0.22	1			
V_{13}	0.42	0.17	0.50	0.41	0.42	0.40	0.50	0.55	0.20	0.26	0.32	0.51	1		
V_{14}	0.50	0.19	0.42	0.12	0.53	0.51	0.35	0.44	0.68	0.22	0.30	0.55	0.23	1	
V_{15}	0.54	0.38	0.86	0.66	0.79	0.81	0.87	0.88	0.57	0.53	0.32	0.82	0.60	0.55	1

Berdasarkan Tabel 5 korelasi terbesar terjadi antara peubah V_8 (Kimia) dan V_{15} (IPK) yaitu sebesar 0.88. Korelasi peubah V_2 (Biologi) dan V_{10} (Pengantar Matematika) terhadap IPK turun menjadi 0.38 dan 0.53, sedangkan korelasi peubah lainnya terhadap nilai IPK meningkat dengan adanya data ekstrim.

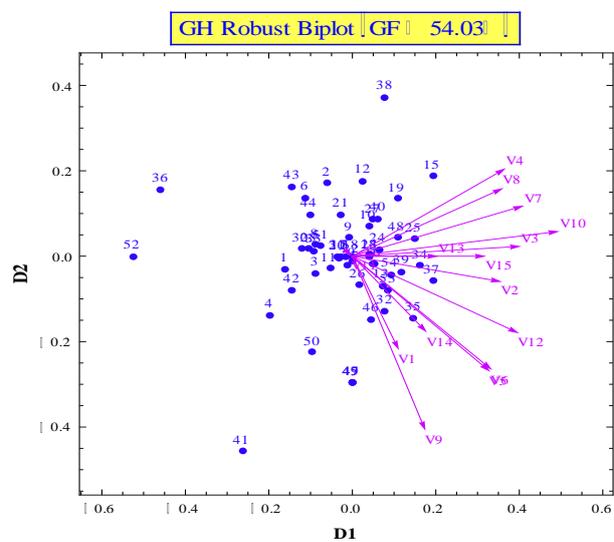
3.2. Pembahasan

Analisis Biplot Data Ekstrim dengan DNS Biasa dan Kekar

Sebagai gambaran untuk mengetahui metode yang tahan terhadap pencilon maka data yang telah diberi data ekstrim dianalisis dengan metode DNS biasa dan kekar, yaitu menggunakan paket Biplot versi 3.2, (Ardana, 2009) dan paket Robust Biplot Pack versi 1.1, dengan memilih $\alpha = 0$. Hasil biplot yang diperoleh disajikan pada Gambar 6 dan 7.



Gambar 6. Biplot biasa dengan nilai ekstrim



Gambar 7. Biplot kekar dengan nilai ekstrim

Tabel 6. Ukuran kesuaian biplot (%)

	Kesuaian	DNS Biasa	DNS Kekar
	Data Peubah	68.11	52.90
GF	Objek	90.04	69.79
		74.11	45.65

Hasil biplot biasa dan kekar pada Gambar 6 dan Gambar 7 dapat dianalisis yaitu:

- a. Ragam dan korelasi peubah dari hasil biplot berdasarkan DNS kekar berbeda dengan hasil biplot berdasarkan DNS biasa dan kekar data tanpa data ekstrim.
- b. Pada Gambar 6 objek 53 (PAPUA 1) menunjukkan prestasi IPK di bawah rata-rata, konfigurasi ini berbeda dengan data asal, hal ini menunjukkan adanya pencilan memberikan dampak terhadap model dan konfigurasi data (Huber & Ronchetti, 2011). Sedangkan pada Gambar 7 objek 53 menunjukkan prestasi IPK di atas rata-rata, konfigurasi data ini hampir sama pada Gambar 3 dan 4 biplot data asal, artinya data pencilan tidak terpengaruh terhadap hasil pemetaan data (Liu *et al.*, 2003; Hawkins, Liu, & Young, 2001; Widowati & Muzdalifah, 2017; Dwipurwani, Cahyawatin, & Susanti, 2022)
- c. Pada Gambar 6 objek 1 (NAD 1) menunjukkan prestasi nilai IPK di atas rata-rata, konfigurasi ini berbeda dengan data asal. hal ini menunjukkan adanya pencilan memberikan dampak terhadap model dan konfigurasi data (Huber & Ronchetti, 2011). Sedangkan Gambar 7 objek 1 menunjukkan prestasi nilai IPK di bawah rata-rata, konfigurasi data ini hampir sama pada Gambar 3 dan 4 biplot data asal. Hasil ini menunjukkan data pencilan tidak terpengaruh terhadap hasil pemetaan data (Liu *et al.*, 2003; Hawkins, Liu, Young, 2001; Widowati & Muzdalifah, 2017; Dwipurwani, Cahyawatin, & Susanti, 2022)
- d. Biplot berdasarkan DNS biasa pada Gambar 6 menunjukkan konfigurasi data atau pemetaan data jauh berbeda setelah diberikan data ekstrim, hal ini menunjukkan metode biasa tidak tahan terhadap data ekstrim. Sedangkan Biplot berdasarkan DNS kekar pada Gambar 7 menunjukkan konfigurasi data yang hampir sama dengan biplot biasa pada Gambar 3 dan 4, hal ini menunjukkan biplot berdasarkan DNS kekar tahan terhadap pencilan atau data ekstrim Hasil ini menunjukkan data pencilan tidak terpengaruh terhadap hasil pemetaan data (Liu *et al.*, 2003; Hawkins, Liu, & Young, 2001; Widowati & Muzdalifah, 2017; Dwipurwani, Cahyawatin, & Susanti, 2022).

4. SIMPULAN

Simpulan dari penelitian ini adalah:

- a. Dengan metode diagram kotak atau box plot, data penelitian prestasi mahasiswa IPB tahun 2008 menunjukkan adanya pencilan. Analisis biplot dengan DNS biasa dan DNS kekar memperlihatkan gambaran pemetaan data yang relatif hampir sama, hal ini menunjukkan bahwa pencilan tidak berpengaruh pada hasil pemetaan data.
- b. Sebagai gambaran metode yang tahan terhadap pencilan maka dengan adanya data ekstrim maka analisis biplot yang didasarkan dengan DNS kekar memperlihatkan

bahwa gambaran pemetaan yang diperoleh tidak terpengaruh oleh adanya data ekstrim.

REFERENSI

- Ardana, NKK & Siswadi (2009). Paket Biplot Biasa dan Kekar dengan Pemrograman Fungsional Mathematica Berbasis GUI. *JMA*. Vol. 8(2): 57-64
- Ardana NKK. (2009). *RobustBiplotPack Versi 1.1. A Mathematica Package for Multivariate Data Visualization*. Bogor: Departemen Matematika FMIPA IPB.
- Ardana NKK, & Siswadi (2005). Biplot dan Implementasinya dengan Pemrograman Fungsional Mathematica. *JMA* 4(2): 21-31.
- Aunuddin (1989). *Analisis Data*. Bahan Pengajaran. Depdikbud Ditjen Dikti, PAU Ilmu Hayat, Institut Pertanian Bogor.
- Daigle, G. & Rivest, LP. (1992). A Robust Biplot. *The Canadian Journal of Statistics*. Statistical Society of Canada.
- Dwipurwani, O., Cahyawatin, D., & Susanti, E. (2022). Analisis Biplot Robust dengan Metode Minimum Covariance Determinant dalam Mendeskripsikan Provinsi Sumantera Selatan Berdasarkan karakteristik Angkatan Kerja Menganggur Dari Aspek Gender. *EULER: Jurnal Ilmiah Matematika, sains dan Teknologi*. 10(1) 54-65.
- Gabriel KR. (2002). Goodness of Fit of Biplots and Correspondence Analysis. *Biometrika*. 89: 423-436.
- Gentle JE. (2007). *Matrix Algebra: Theory, Computations, and Applications in Statistics*. New York: Springer Science + Business Media.
- Hardle, WK & Simar, L. (2015). *Analisis Statistik Multivariate Terapan (edisi ke-4)*. Berlin: Springer.
- Hawkins DM, Liu L, & Young SS. (2001). Robust Singular Value Decomposition. *Technical Report Number 122*. NISS.
- Huber PJ. (1981). *Robust Statistics*. New York: John Wiley & Sons.
- Huber, PJ, & Ronchetti, EM. (2011). *Statistik Kuat (edisi ke-2)*. New Jersey: John Wiley & Sons.
- Johnson RA, & Wichern DW. (2002). *Applied Multivariate Statistical Analysis*. Ed ke-5. Upper Saddle River, NJ: Prentice-Hall.
- Jolliffe IT. (2002). *Principal Component Analysis*. Ed ke-2. Berlin: Springer-Verlag.
- Liu, L., Hawkins, DM, Astaga, S., & Muda, SS (2003). Analisis Dekomposisi Nilai Singular yang Kuat dari Data Microarray. *Prosiding Akademi Ilmu Pengetahuan Nasional AS*, 100, 167-172
- Lipkovich I, & Smith EP. (2002). Biplot and Singular Value Decomposition Macros for Excel®. *Journal of Statistical Software*. 7(5): 1-5.
- Neter, J., Wasserman, W., & Kutner, MH. (1990). *Model Statistik Linier Terapan*. New York: Richard D Irwin Inc.
- Mattjik AA, & Sumertajaya IM. (2006). *Perancangan Percobaan: dengan Aplikasi SAS dan Minitab*. Bogor: IPB.
- Mattjik AA. & Sumertajaya IM. (2011). *Sidik Peubah Ganda dengan menggunakan SAS*. Bogor: IPB Pres.

- Moronna RA., Martin RD, & Yohai UJ. (2006). *Robust Statistics: Theory and Methods*. West Sussex: John Wiley & Sons.
- Siswadi, & Suharjo B. (1997). *Analisis Eksplorasi Data Peubah Ganda*. Bogor: Jurusan Matematika FMIPA.
- Widowati, W. & Muzdalifah, L. (2017). Perbandingan Analisis Biplot Klasik dan Robust Biplot Pada Pemetaan Perguruan Tinggi Swasta di Jawa Timur. *JRAM: Jurnal Riset dan Aplikasi Matematika*.