PEMETAAN PROVINSI DI INDONESIA DENGAN *ROBUST BIPLOT* MENGGUNAKAN METODE *MINIMUM COVARIANCE DETERMINANT* BERDASARKAN PERILAKU PEDULI LINGKUNGAN HIDUP

(Skripsi)

Oleh

NURLELA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2023

ABSTRACT

MAPPING PROVINCES IN INDONESIA WITH ROBUST BIPLOT USING MINIMUM COVARIANCE DETERMINANT METHOD BASED ON ENVIRONMENTAL BEHAVIOUR

By

NURLELA

Outlier data can affect the results of the analysis, so we need a method that is robust for the presence of outliers, namely Robust Biplot with the Minimum Covariance Determinant (MCD) method in estimating the average and also the variance covariance matrix. The results obtained are biplot mapping in which nine adjacent provincial groups are obtained based on the characteristics of environmental care behavior. The provincial group that has better environmental care behavior than other provincial groups includes the first province group is DKI Jakarta, the second province group is Central Java, East Java, Banten, Bali, the third province group is DI Yogyakarta, the fourth province group is Aceh , North Sumatra, Riau, NTB, West Kalimantan and Central Kalimantan, and the sixth group of provinces, namely South Sulawesi. The graphical appearance of the Robust Biplot is able to describe the data characteristics of environmental care behavior in Indonesia by 70.45%.

Keywords : Robust Biplot Analysis; Minimum Covariance Determinant; Environmental Care Behavior.

ABSTRAK

PEMETAAN PROVINSI DI INDONESIA DENGAN *ROBUST BIPLOT* MENGGUNAKAN METODE *MINIMUM COVARIANCE DETERMINANT* BERDASARKAN PERILAKU PEDULI LINGKUNGAN HIDUP

Oleh

NURLELA

Data Pencilan dapat mempengaruhi hasil analisis, sehingga diperlukan metode yang bersifat kekar (*robust*) terhadap adanya pencilan yaitu *Robust Biplot dengan metode Minimum Covariance Determinant (MCD)* dalam pendugaan rata-rata dan juga matriks varian kovarian. Hasil yang diperoleh adalah pemetaan biplot yang mana diperoleh sembilan kelompok provinsi yang berdekatan berdasarkan karakteristik perilaku peduli lingkungan hidup. Kelompok provinsi yang memiliki perilaku peduli lingkungan hidup yang lebih baik daripada kelompok provinsi lain diantaranya kelompok provinsi pertama adalah DK Jakarta, Kelompok provinsi kedua yaitu Jawa Tengah, Jawa Timur, Banten, Bali, kelompok provinsi ketiga yaitu DI Yogyakarta, kelompok provinsi ke empat yaitu Aceh, Sumatera Utara, Riau, NTB, Kalimantan Barat dan Kalimantan Tengah, dan kelompok provinsi ke enam yaitu Sulawesi Selatan. Tampilan grafik *Robust Biplot* mampu menggambarkan karakteristik data perilaku peduli lingkungan hidup di Indonesia sebesar 70,45%.

Kata Kunci: Analisis *Robust Biplot*; *Minimum Covariance Determinant*; Perilaku Peduli Lingkungan Hidup

PEMETAAN PROVINSI DI INDONESIA DENGAN *ROBUST BIPLOT* MENGGUNAKAN METODE *MINIMUM COVARIANCE DETERMINANT* BERDASARKAN PERILAKU PEDULI LINGKUNGAN HIDUP

Oleh

NURLELA

Skripsi

Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Univeristas Lampung

JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
2023

Judul Skripsi : **PEMETAAN PROVINSI DI INDONESIA DENGAN *ROBUST BIPLOT* MENGGUNAKAN METODE *MINIMUM COVARIANCE DETERMINANT* BERDASARKAN PERILAKU PEDULI LINGKUNGAN HIDUP**

Nama Mahasiswa : Nurlela

Nomor Pokok Mahasiswa : 1817031005

Jurusan : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing

Prof. Drs. Mustofa Usman, M.A., Ph.D.	**Prof. Dr. La Zakaria, S.Si., M.Sc.**
NIP 195701011984031020	NIP 196902131994021001

2. Ketua Jurusan Matematika

Dr. Aang Nuryaman, S.Si., M.Si.
NIP 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua : **Prof. Drs. Mustofa Usman, M.A., Ph.D.** .....

Sekretaris : **Prof. Dr. La Zakaria, S.Si., M.Sc.** .....

Penguji
Bukan Pembimbing : **Ir. Warsono, M.S., Ph.D.** .....

2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Suripto Dwi Yuwono, S.Si., M.T.
NIP 19740705 200003 1 001

Tanggal Lulus Ujian Skripsi : **9 Januari 2023**

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama Mahasiswa : Nurlela

Nomor Pokok Mahasiswa : 1817031005

Jurusan : Matematika

Judul Skripsi : **PEMETAAN PROVINSI DI INDONESIA DENGAN *ROBUST BIPLOT* MENGGUNAKAN METODE *MINIMUM COVARIANCE DETERMINANT* BERDASARKAN PERILAKU PEDULI LINGKUNGAN HIDUP**

Dengan ini menyatakan bahwa penelitian ini adalah hasil pekerjan saya sendiri dan apabila kemudian hari terbukti bahwa skripsi ini merupakan salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku. Semua tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya tulis ilmiah Universitas Lampung.

Bandar Lampung, 30 Januari 2023

Yang Menyatakan

Nurlela

RIWAYAT HIDUP

Penulis bernama lengkap Nurlela, anak kedua dari dua bersaudara yang dilahirkan di Tanjung Raya pada tanggal 29 Juni 2000 oleh pasangan Bapak Eli Suheli dan Ibu Ita Sumiati (almh). Penulis memiliki satu orang kakak laki-laki bernama Nurdin Nasrudin.

Penulis menyelesaikan pendidikan di TK Gelora Mekar pada tahun 2004-2005 Pendidikan sekolah dasar di SD Negeri 4 Padang Tambak pada tahun 2006-2012. Pendidikan sekolah menengah pertama di SMP Negeri 1 Way Tenong pada tahun 2012-2015. Pendidikan sekolah menengah atas di SMA Negeri 1 Way Tenong pada tahun 2016-2018. Kemudian penulis diterima sebaga mahasiswa S1 Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan alam (FMIPA) Universitas Lampung (Unila) pada tahun 2018 melalui jalur SNMPTN (Seleksi Nasional Masuk Perguruan Tinggi Negeri). Pada periode 2018/2019 penulis terdaftar sebagai anggota GEMATIKA Himpunan Mahasiswa FMIPA Unila. Penulis pernah menjadi anggota bidang Kaderisasi dan Kepemimpinan Himpunan Mahasiswa Matematika periode 2019. Penulis pernah menjadi anggota Korps Muda BEM (KMB) BEM U KBM Unila pada tahun 2019. Penulis pernah menjadi sekretaris biro Kesekretariatan Himpunan Mahasiswa Matematika periode 2020.

Sebagai bentuk penerapan ilmu perkuliahan, penulis telah melakukan Kerja Praktik (KP) selama 40 hari di UPTB Pengelolaan Pendapatan Daerah Provinsi Sumatera Selatan Wilayah Oku Timur 1 pada tahun 2021. Pada tahun yang sama, sebagai bentuk pengabdian kepada masyarakat, penulis telah melaksanakan Kuliah Kerja Nyata (KKN) selama 40 hari di Kelurahan Fajar Bulan Kecamatan Way Tenong Kabupaten Lampung Barat. Penulis mengikuti kegiatan Magang dan Studi Independen Bersertifikat Di PT Progate Global Indonesia pada semester 7.

KATA INSPIRASI

"Sesungguhnya sesudah kesulitan itu ada kemudahan"

(QS. Al Insyirah: 6)

"Sesungguhnya allah tidak akan mengubah nasib suatu kaum sebelum mereka mengubah diri mereka sendiri"

(Qs. Al Ra'd;11)

"Bersabarlah kamu, sesungguhnya janji Allah adalah benar"

(QS. Ar-Rum : 60)

"Tidak masalah seberapa lambat kau berjalan asalkan kau tidak pernah berhenti"

(Confusius)

"Kalau Kita Tidak Pernah Berjuang Sampai Akhir, Kita Tidak Akan Pernah Melihatnya Walaupun Ada Di Depan Mata"

(Marshal D. Teach)

PERSEMBAHAN

Alhamdulillah,

Puji dan syukur tiada hentinya kepada Allah SWT atas segala nikmat dan karunia-Nya, sehingga skripsi ini dapat terselesaikan dengan baik.

Kupersembahkan karya sederhana ini untuk:

Kedua Orang Tuaku Tercinta Bapak Eli Suheli dan Ibu Ita Sumiati (almh)

Terimakasih atas limpahan kasih sayang, pengorbanan, doa dan seluruh motivasi di setiap langkah penulis. Karena atas doa dan ridho kalian, Allah memudahkan setiap perjalanan hidup ini.

Kakak Tercinta Nurdin Nasrudin

Terimakasih telah mendukung dan mendoakan setiap waktu untuk keberhasilan penulis

Dosen Pembimbing dan Dosen Pembahas

Terimakasih kepada dosen pembimbing dan pembahas yang sudah sangat membantu, memberikan motivasi, memberikan arahan serta ilmu yang berharga.

Teman dan Sahabat

Terimakasih kepada orang-orang baik yang telah memberikan bantuan, dukungan, dan selalu ada saat suka maupun duka

Alamamater Tercinta Universitas Lampung

SANWACANA

Alhamdulillah, segala puji bagi Allah SWT yang telah memberikan rahmat dan hidayah-nya sehingga penulis dapat menyelesaikan skripsi yang berjudul "Pemetaan Provinsi Di Indonesia dengan *Robust Biplot* Menggunakan Metode *Minimum Covariance Determinant* Berdasarkan Perilaku Peduli Lingkungan Hidup." Shalawat serta salam senantiasa terlimpah curahkan kepada Nabi Muhammad SAW yang telah menjadi suri tauladan yang baik sepanjang masa.

Terselesaikannya skripsi ini tidak lepas dari bantuan, kerjasama dan dukungan berbagai pihak. Untuk itu, penulis ingin mengucapkan terimakasih kepada:

1. Bapak Prof. Drs. Mustofa Usman, M.A., Ph.D. selaku dosen pembimbing satu, yang senantiasa membimbing dan memberikan arahan, ide, kritik dan saran serta semangat kepada penulis selama penyusunan skripsi ini.
2. Bapak Prof. Dr. La Zakaria, S.Si., M.Sc. selaku dosen pebimbing kedua, yang telah membimbing, memberi masukan, dan mengarahkan penulis selama proses penyusunan skripsi ini.
3. Bapak Ir. Warsono, M.S., Ph.D selaku dosen pembahas yang telah memberikan kritik dan saran yang membangun kepada penulis selama proses penyusunan skripsi ini.
4. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam
5. Bapak Dr. Eng. Suripto Dwi Yuwono, S.Si., M.T., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam
6. Seluruh dosen, staf dan karyawan Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam
7. Keluargaku tercinta, terutama Bapak, Ibu yang berada di surga dan kakak tercinta yang menjadi motivasi terbesar dalam hidup dan tak pernah berhenti

memberi semangat, doa, dorongan, kasih sayang dan nasihat untuk selalu berjuang setiap harinya

8. Sepupuku tercinta Pita Suliawati dan Sella Selviana yang telah memberikan doa, dukungan serta motivasi kepada penulis
9. Sahabat seperjuanganku tersayang Jessica Andriani yang senantiasa memberikan bantuan, dukungan, serta menemani suka duka penulis
10. Sahabatku Sweetheart Nadea, Teh Susan dan Mba Titis yang senantiasa memberikan semangat dan motivasi kepada penulis
11. Ilham Budiyansah seseorang yang selalu menemani, memberikan doa, menyemangati dari awal hingga akhir
12. Pimpinan HIMATIKA 2020 dan Biro Kesekretariatan atas kebersamaannya selama 1 periode yang telah memberikan banyak pelajaran dan juga pengalaman yang sangat berharga
13. Teman-teman penulis Inka, Wayan, Ramona, Freza, Virda, Nuva, Anis, Rendi, Ria dan Matematika Angkatan 2018. Terimakasih atas bantuan dan keakraban serta kebersamaan selama ini
14. Teman-teman Asrama Mahasiswa Lampung Barat
15. Orang-orang baik yang telah membantu penulis yang tidak dapat disebutkan satu persatu atas peran dan dukungannya dalam menyusun skripsi ini.

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan skripsi ini. Oleh karena itu, penulis mengharapkan masukan berupa kritik dan saran untuk dijadikan pelajaran kedepannya.

Bandar Lampung, 30 Januari 2023

Penulis,

Nurlela

DAFTAR ISI

Halaman

DAFTAR TABEL xv

DAFTAR GAMBAR xvi

I. PENDAHULUAN 1

1.1 Latar Belakang dan Masalah 1

1.2 Tujuan Penelitian 3

1.3 Manfaat Penelitian 3

II. TINJAUAN PUSTAKA 4

2.1 Analisis Multivariat 4

2.2 Analisis Biplot 4

2.2.1 Definisi Analisis Biplot 4

2.2.2 *Covariance Matrix* 5

2.2.3 *Singular Value Decomposition (SVD)* 6

2.3 Pencilan (*Outlier*) 8

2.4 Pendeteksi *Multivariat Outlier* 9

2.5 Pemeriksaan Kesesuaian Biplot 10

2.6 Interpretasi 11

2.7 Biplot Kekar (*Robust Biplot*) 12

2.8 Perilaku Peduli Lingkungan Hidup 14

III. METODOLOGI PENELITIAN 18

3.1 Waktu dan Tempat Penelitian 18

3.2 Data Penelitian 18

3.3 Metode Penelitian 19

IV. HASIL DAN PEMBAHASAN 22

4.1 Deskripsi Data Perilaku Peduli Lingkungan Hidup 22

4.2 Pengujian Pencilan Pada Data 25
4.3 Analisis *Robust Biplot* 28
4.4 Penguraian Nilai Singular 31
4.4.1 Memperoleh Matriks $\boldsymbol{U}$, Matriks $\boldsymbol{\Lambda}$ dan Matriks $\boldsymbol{V}$ 32
4.4.2 Mencari Matriks **F** dan **G** 35
4.5 Plot dan Pengelompokan Provinsi 36
4.6 Uji Kesesuaian Biplot 41
V. KESIMPULAN 42
DAFTAR PUSTAKA 44
LAMPIRAN 46

DAFTAR TABEL

Tabel Halaman

1. Variabel perilaku peduli lingkungan hidup 18
2. Persentase Variabel Peduli Lingkungan Hidup setiap Provinsi di 22
3. Nilai Standar Deviasi dari setiap variabel 25
4. Hasil Pengujian *outlier* menggunakan Jarak Mahalanobis 26
5. Matriks Acak $\boldsymbol{R_1}$ 30
6. Pengelompokan provinsi berdasarkan perilaku peduli lingkungan hidup 37

DAFTAR GAMBAR

Gambar Halaman

1. Plot Pengamatan menggunakan *Robust Distance* 27
2. Plot Pengamatan menggunakan Jarak Mahalanobis 28
3. Tampilan Pemetaan Perilaku Peduli Lingkungan Hidup berdasarkan..... 36

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Kesejahteraan manusia bergantung pada kondisi lingkungannya. Lingkungan dapat mempengaruhi kehidupan manusia. Demikian juga sebaliknya, aktivitas dan juga peilaku manusia akan mempengaruhi kondisi lingkungannya. Saling ketergantungan-nya kehidupan antara manusia dan lingkungannya menjadikan manusia sebagai sasaran utama pemanfaatan sumber daya alam untuk menunjang subsistensinya. Manusia sebagai subjek dalam pemanfaatan sumber daya alam, terkadang membuatnya lupa diri karena didorong adanya keinginan untuk memenuhi kebutuhan. Proses pembangunan yang dilakukan seringkali dilakukan semata-mata untuk mengejar pertumbuhan ekonomi tanpa mempertimbangkan aspek keberlanjutan lingkungan hidup sehingga sering menimbulkan kerusakan.

Lingkungan hidup adalah kesatuan ruang dengan semua benda, daya, keadaan dan makhluk hidup termasuk manusia dan perilakunya yang mempengaruhi alam itu sendiri, kelangsungan perikehidupan dan kesejahteraan manusia serta makhluk hidup lain. Perilaku peduli lingkungan hidup adalah tindakan atau sikap rumah tangga dalan pemenuhan kebutuhan sehari-harinya yang berupaya meminimalisir dampak negatif terhadap alam, kepentingan umum dan kepentingan generasi yang akan datang.

Pada data kita sering menemukan adanya pencilan (*outlier*) yaitu objek pengamatan yang memiliki nilai yang jauh berbeda dengan pengamatan lainnya. Dalam data perilaku peduli lingkungan hidup kemungkinan memiliki data pencilan. Metode statistika klasik sangat sensitif terhadap pencilan, maka dibutuhkan metode yang sifatnya kuat (*robust*) yaitu analisis *robust biplot*. Adapun pendekatan dapat

digunakan untuk metode *robust biplot* yaitu pendekatan *Robust Singular Value Decompotition (RSVD)* atau dengan pendekatan *Minimum Covariance Determinant (MCD)*. Pendekatan RSVD dilakukan dengan cara menduga nilai eigen dan vektor eigen matriks kiri dan matriks kanan yang *robust* pada proses *Singular Value Decompotition (RSVD)*. Sedangkan pendekatan MCD digunakan untuk menduga rataan dan matriks keragaman yang *robust*.

Analisis biplot digunakan dalam berbagai bidang penelitian, misalnya bidang sosial, pendidikan, ekonomi dan bidang lainnya. Salah satu penelitian pada bidang bencana telah dilakukan oleh Hilda Venelia dkk. (2021) tentang *Robust Biplot Analysis Of Natural Disasters In Indonesia From 2019 To 2021*. Selain itu, ada juga penelitian dari Dwipurwani dkk (2022) tentang Analisis Biplot Robust dengan *Metode Minimum Covariance Determinant* dalam Mendeskripsikan Provinsi Sumatera Selatan Berdasarkan Karakteristik Angkatan Kerja Menganggur Dari Aspek Gender.

Penelitian ini betujuan untuk melakukan pemetaan provinsi di Indonesia dengan *Robust Biplot* menggunakan metode *Minimum Covariance Determinant (MCD)* berdasarkan Perilaku Peduli Lingkungan Hidup. Perilaku Peduli Lingkungan Hidup dengan variabel yang berkaitan dengan perilaku rumah tangga dimana data yang digunakan dalam bentuk persentase. Variabel-variabel yang terlibat dalam penelitian ini diantaranya perlakuan terhadap sampah, akses air minum layak, akses hunian layak, akses sanitasi layak, penggunaan lampu hemat energi, sumber penerangan listrik dan penggunaan gas sebagai bahan bakar utama gas.

1.2 Tujuan Penelitian

Adapun tujuan penelitian ini adalah sebagai berikut:

1. Mendapatkan hasil penerapan metode *Robust Biplot* dengan metode *Minimum Covariance Determinant* pada data perilaku peduli lingkungan hidup berupa karakteristik perilaku lingkungan hidup setiap provinsi di Indonesia
2. Mendapatkan grafik pemetaan provinsi-provinsi di Indonesia berdasarkan variabel-variabel perilaku peduli lingkungan hidup.

1.3 Manfaat Penelitian

Manfaat penelitian ini adalah sebagai berikut:

1. Dapat dijadikan sebagai referensi bahan acuan apabila akan melaksanakan suatu penelitian tentang analisis *robust biplot* dengan menggunakan algoritma *Fast-MCD*
2. Menambah pengetahuan tentang cara mengaplikasikan metode *robust biplot* dengan metode *Minimum Covariance Determinant* pada perilaku peduli lingkungan hidup di Indonesia
3. Memberikan informasi terkait dengan karakteristik perilaku peduli lingkungan hidup setiap Provinsi di Indonesia dalam bentuk grafik.

II. TINJAUAN PUSTAKA

2.1 Analisis Multivariat

Analisis multivariat merupakan suatu metode statistika yang digunakan untuk menganalisis hubungan antara beberapa variabel yang dilakukan secara simultan. Tujuan yang dapat dicapai dengan metode analisis multivariat yaitu reduksi data atau penyederhanaan struktur data, sortasi dan pengelompokan, identifikasi hubungan antar variabel, prediksi serta konstruksi dan pengujian hipotesis. Struktur data dalam analisis multivariat terdiri dari n objek dan p variabel $(\boldsymbol{X} = \boldsymbol{X}_1, \boldsymbol{X}_2, ..., \boldsymbol{X}_p)$ dapat dinyatakan dalam bentuk matriks sebagai berikut (Johnson & Wichern, 2007).

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{2.1}$$

2.2 Analisis Biplot

2.2.1 Definisi Analisis Biplot

Pertama kalinya analisis biplot diperkenalkan oleh Gabriel pada tahun 1971. Menurut Gabriel (1971), analisis biplot adalah upaya untuk memberikan peragaan secara grafis dari matriks data $\boldsymbol{X}$ dalam suatu plot dengan menumpangtindihkan vektor-vektor baris matriks $\boldsymbol{X}$ yang menggambarkan objek dengan vektor-vektor yang mewakili kolom matriks $\boldsymbol{X}$ yang menggambarkan variabel.

Terdapat beberapa tipe biplot. Perbedaan tipe ini didasarkan pada nilai a digunakan. Nilai a yang digunakan dalam biplot yaitu $0 \le a \le 1$. Untuk nilai a yang sering digunakan adalah $a = 1; a = 0,5; a = 0$. Apabila a yang digunakan adalah $a = 1$ maka biplot yang dibentuk disebut dengan biplot R*ow Metric Preserving* (RMP). Biplot RMP digunakan untuk menduga jarak Euclid secara opimal. Biplot dengan $a = 1$ disebut sebagai Biplot Komponen Utama atau sering disebut sebagai Biplot Klasik (Nugroho, 2008).

Menurut Solimun dan Fernandes (2008), analisis biplot dapat digunakan untuk memperagakan antara objek dan peubah yang berada pada ruang berdimensi tinggi ke dalam ruang yang berdimensi rendah (dua atau tiga) sekaligus. Ruang dimensi rendah ini digunakan karena kemudahan dalam interpretasi secara grafis, karena jika digunakan pada grafik berdimensi tinggi (lebih dari 3), sangat sukar untuk diinterpretasikan titik-titik dalam grafis tersebut. Matriks $\boldsymbol{X}$ yang digunakan dalam analisis biplot yaitu:

$$\boldsymbol{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \tag{2.2}$$

dimana:

n = banyak objek pengamatan

p = banyak variabel yang diteliti

2.2.2 *Covariance Matrix*

Matriks kovarian berdimensi d dari vektor acak $\tilde{\boldsymbol{x}}$ adalah matriks $d \times d$

$$\Sigma_{\tilde{\boldsymbol{x}}} = E(c(\tilde{x})c(\tilde{x})^T) = \begin{bmatrix} Var(\tilde{x}[1]) & Cov(\tilde{x}[1], \tilde{x}[2]) & \cdots & Cov(\tilde{x}[1], \tilde{x}[d]) \\ Cov(\tilde{x}[1], \tilde{x}[2]) & Var(\tilde{x}[2]) & \cdots & Cov(\tilde{x}[2], \tilde{x}[d]) \\ \vdots & \vdots & \ddots & \vdots \\ Cov(\tilde{x}[1], \tilde{x}[2]) & Cov(\tilde{x}[1], \tilde{x}[d]) & \cdots & Var(\tilde{x}[d]) \end{bmatrix} \tag{2.3}$$

dimana $c(\tilde{x}) := \tilde{x} - E(\tilde{x})$

Untuk sembarang vektor acak $\tilde{x}$ dengan matriks kovarians $\Sigma_{\tilde{x}}$ dan sembarang vektor $\boldsymbol{u}$

$$Var(v^T \tilde{x}) = E((v^T \tilde{x} - E(v^T \tilde{x}))^2) \tag{2.4}$$

Pembuktian:

$$\begin{aligned} Var(v^T \tilde{x}) &= v^T \Sigma_{\tilde{x}} v \\ &= E((v^T c(\tilde{x}))^2 \\ &= v^T E(c(\tilde{x})c(\tilde{x})^T v \end{aligned} \tag{2.5}$$

2.2.3 *Singular Value Decomposition (SVD)*

Perhitungan biplot didasarkan pada dekomposisi nilai singular atau *Singular Value Decomposition (SVD)*. Kegunaan SVD dalam analisis biplot adalah menghasilkan baris matriks yang saling bebas (*orthogonal*). Menurut Seber (2009) definisi *Singular Value Decomposition* yaitu sebagai berikut:

Terdapat matriks $\boldsymbol{X}$ berukuran $n \times p$ dari *rank* r $(r \le n \le p)$ dapat dinyatakan dalam bentuk:

$$\boldsymbol{X}_{(n\times p)} = \boldsymbol{U}_{(n\times p)} \boldsymbol{\Lambda}_{(p\times p)} \boldsymbol{V}^T_{(p\times p)} \tag{2.6}$$

dimana:

$\boldsymbol{X}_{(n\times p)}$: Matriks data berukuran $n \times p$ yang dikoreksi terhadap nilai rata-ratanya

$\boldsymbol{U}_{(nxp)}$: Matriks orthogonal yaitu matriks dengan kolom orthonormal sehingga $\boldsymbol{U}^T \boldsymbol{U} = \boldsymbol{I_r}$ (matriks identitas berdimensi r).

Kolom matriks $\boldsymbol{U}$ berisi vektor eigen matriks $\boldsymbol{XX}^T$ dengan $v_i, i = 1, 2, ..., r$ adalah vektor eigen ke-i yang bersesuaian dengan nilai eigen tak nol matriks $\boldsymbol{XX}^T$. Matriks $\boldsymbol{U}$ didefinisikan sebagai:

$\boldsymbol{U} = \left[\frac{\boldsymbol{X}_{v_1}}{\sqrt{\lambda_1}}, \frac{\boldsymbol{X}_{v_2}}{\sqrt{\lambda_2}}, ..., \frac{\boldsymbol{X}_{v_r}}{\sqrt{\lambda_r}} \right]$; dengan kolom-kolom matriks $\boldsymbol{U}$ disebut vektor singular kolom matriks $\boldsymbol{X}$ dalam ruang berdimensi n

$\boldsymbol{\Lambda}_{(p\times p)}$: Matriks berukuran $p\times p$ yang merupakan matriks diagonal dengan unsur-unsur diagonalnya merupakan akar kuadrat nilai eigen $\boldsymbol{X}^T\boldsymbol{X}$ atau $\boldsymbol{X}\boldsymbol{X}^T$. Matriks $\boldsymbol{X}^T\boldsymbol{X}$ dan $\boldsymbol{X}\boldsymbol{X}^T$ merupakan matriks simetri dengan nilai eigen (λ) non negatif dan akar ciri (nilai eigen) tak nol matriks $\boldsymbol{X}^T\boldsymbol{X}$ dan $\boldsymbol{X}\boldsymbol{X}^T$ adalah sama. $\sqrt{\lambda_i}$ dengan $i=1,2,...,r$ merupakan nilai eigen terbesar ke-i sehingga $\sqrt{\lambda_1}\geq\sqrt{\lambda_2}\geq...\geq\sqrt{\lambda_r}$, oleh karena itu, matriks $\boldsymbol{\Lambda}$ didefinisikan sebagai berikut:

$$\boldsymbol{\Lambda}=\begin{bmatrix}\sqrt{\lambda_1} & 0 & \cdots & 0\\ 0 & \sqrt{\lambda_2} & \vdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \sqrt{\lambda_r}\end{bmatrix}$$

dengan unsur-unsur diagonal matriks $\boldsymbol{\Lambda}$ disebut nilai singular dari matriks $\boldsymbol{X}$.

$\boldsymbol{V}_{(p\times p)}$: Matriks $\boldsymbol{V}$ berukuran $(p\times p)$ yang merupakan matriks dengan kolom orthonormal sehingga $\boldsymbol{V}^T\boldsymbol{V}=\boldsymbol{I}_r$. Kolom-kolom matriks $\boldsymbol{V}$ adalah vektor eigen dari matriks $\boldsymbol{X}^T\boldsymbol{X}$ yang bersesuaian dengan nilai eigen. Matriks $\boldsymbol{V}$ didefinisikan sebagai:

$\boldsymbol{V}=[\boldsymbol{v}_1,\boldsymbol{v}_2,...,\boldsymbol{v}_r]$; dengan kolom-kolom matriks $\boldsymbol{V}$ disebut sebagai vektor singular baris matriks $\boldsymbol{X}$ dalam ruang berdimensi p.

Unsur diagonal dari matriks $\boldsymbol{\Lambda}$ disebut SVD dari matriks $\boldsymbol{X}$ sedangkan r merupakan *rank* dari matriks $\boldsymbol{X}$. SVD bergantung pada *rank* dari matriks $\boldsymbol{X}$ atau $r=\min(n,p)$ (Solimun dan fernandes, 2008). Biplot digambarkan sebagai grafik dua dimensi, maka r yang digunakan adalah $r=2$, sehingga diperoleh penguraian matriks sebagai berikut:

$$\boldsymbol{u}=[\boldsymbol{u}_1\quad \boldsymbol{u}_1]\qquad \mathbf{v}=[\mathbf{v}_1\quad \mathbf{v}_2]\qquad \boldsymbol{\Lambda}=\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} \tag{2.7}$$

$$\boldsymbol{u}_1=\begin{bmatrix}\boldsymbol{u}_{11}\\ \boldsymbol{u}_{12}\\ \vdots\\ \boldsymbol{u}_{1n}\end{bmatrix}\qquad \boldsymbol{u}_2=\begin{bmatrix}\boldsymbol{u}_{21}\\ \boldsymbol{u}_{22}\\ \vdots\\ \boldsymbol{u}_{2n}\end{bmatrix}\qquad \boldsymbol{v}_1=\begin{bmatrix}\boldsymbol{v}_{11}\\ \boldsymbol{v}_{12}\\ \vdots\\ \boldsymbol{v}_{1n}\end{bmatrix}\qquad \boldsymbol{v}_2=\begin{bmatrix}\boldsymbol{v}_{21}\\ \boldsymbol{v}_{21}\\ \vdots\\ \boldsymbol{v}_{2n}\end{bmatrix} \tag{2.8}$$

Gabriel (1971) mendefinisikan $\Lambda^{\alpha} = \left(\sqrt{\lambda_1^{\alpha}}, \sqrt{\lambda_2^{\alpha}}, ..., \sqrt{\lambda_r^{\alpha}}\right)$ dengan $\alpha \epsilon [0,1]$ dan memperkirakan $\boldsymbol{F} = \boldsymbol{U\Lambda^{\alpha}}, \mathbf{G} = \mathbf{V\Lambda^{1-\alpha}}$ kemudian dapat dituliskan sebagai berikut:

$$\begin{aligned} \boldsymbol{X} &= \boldsymbol{U\Lambda V^T} \\ &= \left(\boldsymbol{U\Lambda^{\alpha}}\right)\left(\boldsymbol{V\Lambda^{1-\alpha}V^T}\right) \\ &= \boldsymbol{F\,G^T} \end{aligned} \tag{2.9}$$

Oleh karena itu, elemen (i, j) dari matriks data $\boldsymbol{X}_{(n\times p)}$ dapat dinyatakan:

$$x_{(i,j))} = f_i^T g_j \tag{2.10}$$

dimana $g_i^T, i = 1, 2, \ldots, n$ dan $h_j, j = 1, 2, \ldots, p$ adalah vektor kolom dari $\boldsymbol{F}$ dan $\boldsymbol{G}$, masing-masing, dengan elemen r. Disini, n baris dari $\boldsymbol{F}$ sesuai dengan baris $\boldsymbol{X}$ dan p baris dari $\boldsymbol{G}$ sesuai dengan kolom $\boldsymbol{X}$.

2.3 Pencilan (*Outlier*)

Keragaman data sangat dibutuhkan dalam analisis statistika, namun keragaman data juga dapat menyebabkan adanya nilai pengamatan yang berbeda dengan nilai pengamatan lain. Dapat diartikan terdapat beberapa data yang berbeda dengan pola keseluruhan data. Penyebab pencilan dapat disebabkan oleh kesalahan pengamatan, pencatatan maupuan kesalahan lainnya. Pencilan merupakan data yang mempunyai karakteristik unik terlihat sangat jauh berbeda dari hasil pengamatan lain dan muncul dalam bentuk nilai ekstrim baik untuk peubah tunggal atau peubah kombinasi (Hair, dkk., 1998).

Suatu pengamatan dianggap *outlier* apabila pengamatan tersebut memiliki nilai lebih besar dari rata-rata ditambahkan dengan 3 kali standar deviasi (Kriegel, 2010). Secara matematis suatu pengamatan dianggap outlier jika:

$$y_i \geq \bar{y} + 3stdv\left(y_i\right) \tag{2.11}$$

Pencilan dapat dihilangkan atau tidak diikutsertakan dalam proses pengolahan data kecuali pencilan yang memberikan informasi penting tentang model. Ada beberapa

cara untuk mengatasi pencilan yaitu dengan transformasi data yang mana pencilan tetap dipertahankan dan juga mengurangi besarnya keragaman. Selain itu dapat digunakan teknik analisis yang dapat mengatasi pencilan seperti metode *robust* yang tidak sensitif terhadap pencilan (Soemartini, 2007).

2.4 Pendeteksi *Multivariat Outlier*

Menurut Barnett dan Lewis (1978), pencilan merupakan suatu pengamatan yang tidak mengikuti sebagian besar pola dan biasanya terletak jauh dari pusat data. Jarak Mahalonobis menyatakan bahwa pengamatan ke-*j* dikatakan pencilan apabila jaraknya lebih besar dari nilai *chi-square* tabel pada sejumlah pengamatan.

Perhitungan jarak mahalanobis sebagai berikut:

$$d^2_{MD} = (x_i - \bar{\boldsymbol{x}})^T \boldsymbol{S}^{-1} (x_i - \bar{\boldsymbol{x}}) > \chi^2_{P,(1-a)},\ j = 1, 2, ..., n \qquad (2.12)$$

dengan:

$\bar{\boldsymbol{x}}$	: vektor rata-rata
$\boldsymbol{S}$	: matriks varian kovarian
p	: banyaknya variabel pengamatan
$x_1, x_2, ..., x_n$	: observasi sampel

Untuk mendeteksi adanya pencilan tunggal dapat menggunakan jarak mahalanobis. Jika $p > 2$ akan sulit dalam melakukan pendeteksian adanya pencilan pada p variabel data. Jika terdapat pencilan tunggal dapat kita gunakan jarak mahalanobis, namun pendekatan ini kurang efektif untuk pengamatan yang pencilannya lebih dari satu. Pengaruh *masking* (penyamaran) dan *swamping* (pelimpahan) dapat menyebabkan identifikasi pencilan menjadi kurang optimal. Penyamaran terjadi ketika pengamatan pencilan tidak terdeteksi dikarenakan adanya pengamatan pencilan lain yang berdekatan, sedangkan pelimpahan terjadi saat teridentifikasinya pengamatan baik sebagai pengamatan pencilan.

Oleh karena itu, akan lebih baik jika menggunakan perhitungan jarak dengan penduga *robust* bagi sebaran multivariat. Jarak *robust* merupakan pendekatan untuk mengidentifikasi pencilan pada data multivariat, dengan menggunakan penaksir dari vektor rata-rata $(\bar{\boldsymbol{x}})$ dan matriks varian kovarian $(\boldsymbol{S})$. Sehingga, metode ini mampu meminimumkan pengaruh dari adanya efek *masking* (penyamaran) dan *swamping* (pelimpahan) dalam pendeteksian pencilan (Rencher, 2002). Salah satu penduga *robust* yang memiliki kemampuan mengukur jarak dan mendeteksi titik *leverage* (pencilan yang disebabkan oleh variabel independen) adalah *Minimum Covariance Determinant* (MCD). Suatu pengamatan x_i diidentifikasi sebagai pencilan jika jarak mahalonis *robust* sebagai berikut:

$$d^2_{RD} = (x_i - \bar{\boldsymbol{x}}_{\boldsymbol{MCD}})^T \boldsymbol{S}^{-1}_{\boldsymbol{MCD}} (x_i - \bar{\boldsymbol{x}}_{\boldsymbol{MCD}}) > \chi^2_{p,(1-a)}, i = 1, 2, ..., n \tag{2.13}$$

dimana $\bar{\boldsymbol{x}}_{\boldsymbol{MCD}}$ dan $\boldsymbol{S}_{\boldsymbol{MCD}}$ menyatakan vektor rata-rata dan matriks varian kovarian dari sebagian data X yang mempunyai determinan matriks varian kovarian terkecil (Hubert, dkk., 2008).

2.5 Pemeriksaan Kesesuaian Biplot

Analisis biplot memerlukan adanya pemeriksaan kesesuain biplot. Hal ini bertujuan untuk mengetahui seberapa besar biplot dapat menjelaskan data sebenarnya. Menurut Mattjik dan Sumerteja (2011), biplot dianggap memberikan informasi yang cukup jika telah memberikan informasi minimal 70%. *Goodness of fit* dari biplot diketahui dengan memeriksa dua nilai eigen pertama λ_1 dan λ_2 yaitu :

$$p^2 = \frac{\lambda_1 + \lambda_2}{\sum_{k=1}^{r} \lambda_k} \tag{2.14}$$

dimana:

λ_1 : nilai eigen terbesar ke-1

λ_2 : nilai eigen terbesar ke-2

λ_k : nilai eigen ke-k, dengan $k = 1, 2, ..., r$

Apabila nilai p^2 mendekati satu, maka biplot memberikan penyajian yang semakin baik mengenai informasi data yang sebenarnya.

2.6 Interpretasi

Menurut Mattjik dan Sumerterja (2011), informasi yang didapatkan dari analisis biplot yaitu:

1. Kedekatan antar objek yang diamati

   Kemiripan sifat dari dua objek menjelaskan kedekatan antar objek. Semakin dekat letak dua objek kemiripan dua objek akan semakin tinggi.
2. Keragaman peubah

   Keragaman peubah dapat dilihat berdasarkan panjang vektor peubah. Peubah yang nilai keragamannya kecil digambarkan sebagai vektor pendek. Sebaliknya peubah yang nilai keragamannya besar digambarkan sebagai vektor panjang.
3. Korelasi antarpeubah

   Dua peubah memiliki nilai korelasi positif jika digambarkan sebagai dua buah vektor yang membentuk sudut lancip. Sebaliknya dua peubah memiliki nilai korelasi negatif jika digambarkan sebagai dua buah vektor yang membentuk sudut tumpul. Apabila digambarkan sebagai dua vektor yang membentuk sudut segititga siku-siku maka dua peubah tidak memiliki nilai korelasi.
4. Nilai peubah pada suatu objek

   Objek yang letaknya searah dengan arah dari vektor peubah akan memiliki nilai diatas rata-rata. Sebaliknya objek yang letaknya berlawanan arah dari vektor peubah akan memiliki nilai dibawah rata-rata. Kegunaan nilai peubah pada suatu objek adalah untuk melihat peubah ciri dari setiap objek. Dengan melakukan proyeksi orthogonal dari objek ke vektor peubah maka dapat terlihat nilai peubah pada suatu objek.

2.7 Biplot Kekar (*Robust Biplot*)

Menurut Hawkins, dkk. (2003), analisis biplot kekar merupakan salah satu pendekatan yang dilakukan dengan cara menduga nilai eigen dan vektor eigen kiri ***(U)*** dan nilai eigen serta vektor eigen kanan $(\boldsymbol{A})$ sehingga hasil dugaannya tahan terhadap data pencilan. Metode yang dapat digunakan untuk menduga rataan dan matriks kovarian salah satunya yaitu dengan menggunakan metode MCD *(Minimum Covariance Determinant).*

Metode *Minimum Covariance Determinant* merupakan penduga *robust* untuk rata-rata dan matriks varian kovarian dengan mencari himpunan data yang menghasilkan determinan terkecil yang digunakan untuk mendeteksi adanya suatu pencilan. Tujuan dari metode ini adalah untuk mendapatkan suatu sub sampel berukuran h dari keseluruhan jumlah pengamatan n yang matriks varian kovariannya memiliki determinan terkecil diantara semua kemungkinan kombinasi data. Misalkan himpunan bagian x_h, maka terdapat C_h^n kombinasi himpunan yang harus dicari untuk mendapatkan penduga MCD dengan,

$$h = \frac{(n+p+1)}{2}, h \leq n \tag{2.15}$$

Dimana p menyatakan banyaknya variabel (Rousseeuw and Driessen, 1999).

Misalkan $\boldsymbol{X} = (x_1, x_2, ..., x_n)^T$ adalah himpunan data sejumlah n pengamatan yang terdiri dari $p-$variabel dimana $n \geq p+1$. Penduga MCD adalah pasangan sub sampel $\boldsymbol{T}$ dan $\boldsymbol{S}$ yang termasuk matriks definit positif simetri berdimensi $p \times p$ dari suatu sub sampel berukuran h pengamatan dimana $h = \frac{(n+p+1)}{2}, h \leq n$ dengan

$$\boldsymbol{T} = \frac{1}{h} \sum_{i=1}^{h} x_i, i = 1, 2, ..., n \tag{2.16}$$

$$\boldsymbol{S} = \frac{1}{h} \sum_{i=1}^{h} (x_i - T_i)^T (x_i - T_i), i = 1, 2, ..., n \tag{2.17}$$

dimana,

$$h = \frac{(n+p+1)}{2}$$

dengan:

$\boldsymbol{T}$	: vektor rata-rata
$\boldsymbol{S}$	: matriks varian kovarian
$x_1, x_2, ..., x_n$	: observasi sampel

(Rousseeuw and Driessen, 1999).

Pada populasi dengan jumlah pengamatan yang kecil, penduga MCD dapat dengan cepat dihitung. Tetapi jika jumlah pengamatannya besar, maka akan banyak kombinasi sub sampel dari H yang harus dicari dan perhitungannya akan cukup memakan waktu. Dalam mengatasi keterbatasan ini maka digunakan suatu algoritma baru untuk metode MCD yang dinamakan dengan metode *Fast-MCD* yaitu teorema *C-Step* (Dayanti, dkk, 2016).

Teorema *C-Step*

Menurut Rousseuw and Driessen (1999), misalkan $\boldsymbol{X} = (x_1, x_2, ..., x_n)^T$ merupakan himpunan sejumlah n pengamatan yang terdiri dari $p-$variabel. Misalkan $H_1 \subset \{1, 2, ..., n\}$ dengan sejumlah elemen $|H_1| = h$, tetapkan $\boldsymbol{T_1} = \frac{1}{h}\sum_{i=1}^{h} x_i$ dan $\boldsymbol{S_1} = \frac{1}{h}\sum_{i}^{h}(x_i - \boldsymbol{T_1})^T(x_i - \boldsymbol{T_1})$. Jika det $(S_1) \neq 0$ maka jarak relatifnya sebagai berikut:

$$d_1(i) = \sqrt{(x_i - \boldsymbol{T_1})^T \boldsymbol{S_1}^{-1}(x_i - \boldsymbol{T_1})}, i = 1, 2, ..., n \tag{2.18}$$

Selanjutnya, ambil H_2 sedemikian sehingga, $\{d_1(i); i \in H_2 := \{(d_1)_{1:n}, (d_1)_{2:n,...,}(d_1)_{h:n}\}$ dimana $(d_1)_{1:n} \leq (d_1)_{2:n} \leq ... \leq (d_1)_{h:n}$ merupakan orde jarak T_2 dan S_2 dihitung berdasarkan H_2. Selanjutnya, $\det(S_2) \leq \det(S_2)$ dan akan sama jika dan hanya jika $T_1 = T_2$ dan $S_1 = S_2$ (Rousseeuw and Driessen, 1999).

Berikut ini adalah Algoritma *Fast-MCD*:

1. Ambil himpunan $\boldsymbol{X}$ secara acak. Misalkan himpunan bagian tersebut sebagai H_1 dimana $h = \frac{(n+p+1)}{2}$

2. Hitung vektor rata-rata T_1 dan matriks varian kovarian S_1 dari H_1 dengan

$$\boldsymbol{T_1} = \frac{1}{h}\sum_{i=1}^{h} x_i \text{ dan } \boldsymbol{S_1} = \frac{1}{h}\sum_{i=1}^{h}(x_i - \boldsymbol{T}_i)^T(x_i - \boldsymbol{T}_i) \quad (2.19)$$

3. Hitung determinan dari matriks varian kovarian $\boldsymbol{S_1}$
4. Hitung jarak relatif dari setiap pengamatan terhadap rata-rata dan matriks varian kovarian dengan persamaan:

$$d_1(i) = \sqrt{(x_i - \boldsymbol{T_1})^T \boldsymbol{S_1}^{-1}(x_i - \boldsymbol{T_1})} \quad (2.20)$$

5. Urutkan jarak relative berdasarkan jarak mahalanobis dari yang terkecil hingga yang terbesar
6. Bentuk himpunan bagian baru dengan H_2 sedemikian sehingga

   $\{d_1(i); i \in H_2 := \{(d_1)_{1:n}, (d_1)_{2:n,\ldots,}(d_1)_{h:n}\}$ dimana $(d_1)_{1:n} \le (d_1)_{2:n} \le \ldots \le (d_1)_{h:n}$
7. Hitung vektor rata-rata $\boldsymbol{T_2}$, matriks varian kovarian $\boldsymbol{S_2}$ dan $d_2(i)$ dari H_2
8. Ulangi langkah 1 sampai dengan 6 hingga didapatkan bahwa $\det(S_2) \le \det(S_1)$

2.8 Perilaku Peduli Lingkungan Hidup

Menurut Undang-Undang Republik Indonesia Nomor 32 Tahun 2009 Tentang Perlindungan dan Pengelolaan Lingkunga Hidup pada Bab 1 Pasal 1 ayat 1 menyatakan bahwa " Lingkungan hidup adalah kesatuan ruang dengan semua benda, daya, keadaan, dan makhluk hidup, termasuk manusia dan perilakunya, yang mempengaruhi alam itu sendiri, kelangsungan perikehidupan, dan kesejahteraan manusia serta makhluk hidup lain". Pasal 1 Ayat 2 menyatakan bahwa "Perlindungan dan pengelolaan lingkungan hidup adalah upaya sistematis dan terpadu yang dilakukan untuk melestarikan fungsi lingkungan hidup dan mencegah terjadinya pencemaran atau kerusakan lingkungan hidup yang meliputi perencanaan, pemanfaatan, pengendalian, pemeliharaan, pengawasan dan penegakan hukum".

Manusia memiliki peranan yang sangat penting dalam menjaga kelestarian lingkungan hidup. Satuan terkecil setelah manusia adalah rumah tangga, yang dituntut untuk

menjaga kelestarian lingkungan, dengan cara memberikan tingkah laku dan kebiasaan hidup yang baik dan tidak mencemari lingkungan.

Berdasarkan hasil survei Perilaku Peduli Lingkungan Hidup (SPPLH) yang dilakukan oleh BPS 2012, indikator dan parameter yang digunakan untuk mengukur kepedulian masyarakat terhadap lingkungan yaitu:

1. Perilaku dalam penghematan energi

   Energi memiliki peran strategis dalam kehidupan masyarakat. Kebutuhan masyarakat akan energi cukup tinggi untuk mendukung kegiatan sosial ekonomi, karena peranan tersebut maka perilaku penghematan energi dijadikan sebagai indikator untuk mengukur perilaku peduli lingkungan. Konsumsi energy di Indonesia umumnya adalah penggunaan tenaga listrik, gas, dan kayu bakar. Penghematan energi dapat dilakukan dengan penggunaan listrik sesuai dengan keperluan sehingga dapat dimatikan setelah tidak digunakan.

2. Perilaku membuang sampah

   Sering kali terjadi masyarakat membuang sampah secara sembarangan yang mengakibatkan tersumbatnya saluran drainase dan saluran air hujan sehingga menimbulkan banjir serta dapat berpengaruh terhadap kesehatan. Cara untuk mengolah sampah diantaranya yaitu mendaur ulang barang yang masih dapat dipakai, menyediakan tempat sampah yang sesuai dan menggunakan barang ramah lingkungan.

3. Perilaku pemanfaatan air

   Krisis air bersih terjadi akibat tidak berfungsinya sumur sebagai sumber air, menurunnya debit air permukaan tanah, berkurangnya pasokan air tanah serta daerah resapan air berkurang mengakibatkan kekeringan di musim kemarau dan banjir di musim hujan. Maka dari itu, pemanfaatan dan penghematan air menjadi sangat penting terutama air bersih. Konservasi air merupakan hal penting. Konservasi air dapat dilakukan dengan cara berikut:

   a. Menyimpan air yang berlebih saat hujan untuk dapat dimanfaatkan ketika diperlukan
   b. Menghemat air dengan pemakaian yang efektif dan efisien
   c. Mengendalikan penggunaan air tanah.

4. Perilaku penyumbang emisi karbon

   Tempat tinggal berpengaruh terhadap kondisi lingkungan melalui penggunaan alat-alat rumah tangga seperti penghangat ruangan, pemanas air, lemari es, mesin cuci, AC dan alat lainnya. Berdasarkan laporan PBB, emisi karbon telah melewati batas terburuk, selain itu juga emisi karbon menyebabkan terjadinya efek rumah kaca yang terus meningkat setiap tahunnya. Emisi karbon dapat dikurangi dengan cara mengontrol perilaku sehari-hari, terutama perilaku yang berpotensi menjadi penyumbang emisi karbon.
5. Perilaku hidup sehat

   Perilaku hidup sehat adalah keputusan untuk menjaga keberlangsungan fungsi kehidupan, baik tingkat individu maupun lingkungan. Perilaku hidup sehat dapat diukur berdasarkan kebiasaan pemeliharaan di sekitar rumah, penyediaan area resapan air serta menjaga kebersihan lingkungan diantaranya lingkungan rumah, kebersihan jalan, kebersihan tempat umum dan lainnya.
6. Perilaku penggunaan bahan bakar

   Sumber daya yang tidak dapat diperbaharui yaitu bahan bakar. Contohnya penggunaan transportasi berbahan bakar berguna untuk mempermudah mobilitas penduduk dari satu tempat ke tempat lain. Pentingnya bahan bakar untuk transportasi maka berhemat dalam penggunaan bahan bakar adalah salah satu indikator perilaku peduli lingkungan, selain itu juga dapat mengurangi terjadinya polusi udara.

Variabel- variabel perilaku peduli lingkungan hidup yang digunakan dalam penelitian ini menurut Badan Pusat Statistik-Statistik Indonesia (2019):

1. Perlakuan terhadap sampah merupakan kegiatan yang dilakukan dalam rangka pengolahan terhadap sampah rumah tangga, seperti pengolahan sampah organik dan anorganik
2. Akses air minum layak merupakan sarana memperoleh air minum yang layak untuk dikonsumsi sehari-hari. Air minum yang layak memiliki ciri yaitu airnya jernih, tidak berbau dan tidak berwarna, memiliki rasa tawar, tidak meninggalkan noda, tidak mengandung zat kimia berbahaya dan juga bebas endapan

3. Akses hunian layak merupakan sarana yang tersedia sebagai tempat untuk berlindung
4. Akses sanitasi layak merupakan sarana yang tersedia untuk proses sanitasi dalam upaya agar kotoran tidak bersentuhan langsung dengan manusia. Sanitasi dapat diartikan sebagai upaya untuk membina dan menciptakan kehidupan yang baik di bidang kesehatan
5. Penggunaan lampu hemat energi adalah suatu perilaku dalam rumah tangga yang menggunakan sumber penerangan dengan lampu berdaya hemat energi seperti penggunaan lampu LED *(Light Emitting Diode)*
6. Sumber penerangan listrik adalah perilaku dalam rumah tangga yang menggunakan sumber penerangan listrik
7. Penggunaan gas sebagai bahan bakar utama memasak adalah perilaku dalam rumah tangga dalam memasak sehari-hari menggunakan bahan bakar gas.

III. METODOLOGI PENELITIAN

3.1 Waktu dan Tempat Penelitian

Penelitian ini dilakukan pada semester genap tahun ajaran 2021/2022 bertempat di Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

3.2 Data Penelitian

Data yang digunakan dalam penelitian ini adalah data Perilaku Peduli Lingkungan Hidup yang merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) tahun 2020. Berikut ini adalah variabel-variabel perilaku peduli lingkungan hidup

Tabel 1. Variabel perilaku peduli lingkungan hidup

No.	Nama Variabel	Simbol	Keterangan
1.	Perlakuan terhadap Sampah	$X1$	Persentase Rumah Tangga berdasarkan sampah yang diangkut oleh petugas menurut Provinsi
2.	Akses air minum layak	$X2$	Persentase Rumah Tangga berdasarkan produksi sampah menurut Provinsi
3.	Akses hunian Layak	$X3$	Persentase Rumah Tangga yang memiliki akses hunian layak menurut Provinsi

4.	Akses sanitasi Layak	$X4$	Persentase Rumah Tangga yang memiliki akses sanitasi layak menurut Provinsi
5.	Penggunaan Lampu Hemat Energi	$X5$	Persentase Rumah Tangga yang menggunakan lampu hemat energi menurut Provinsi
6.	Sumber Penerangan Listrik	$X6$	Persentase Rumah Tangga yang menggunakan sumber penerangan listrik menurut Provinsi
7.	Penggunaan Gas Sebagai Bahan Bakar Utama Memasak	$X7$	Persentase Rumah Tanggayang menggunakan gas sebagai sebagai bahan utama memasak menurut Provinsi

3.3 Metode Penelitian

Adapun langkah-langkah yang dilakukan pada penelitian ini adalah sebagai berikut:

1. Mengumpulkan data sekunder dari Katalog Kajian Lingkungan Hidup Tahun 2020 dan web BPS (Badan Pusat Statistik) tentang data Peduli Lingkungan Hidup
2. Mendeskripsikan data setiap variabel dalam bentuk tabel berdasarkan provinsi-provinsi di Indonesia
3. Menguji pencilan pada data
4. Memperoleh matriks kovarian dengan menggunakan metode MCD
   Tahapan Algoritma *Fast-MCD* sebagai berikut:
   a. Mengambil himpunan bagian $\boldsymbol{R_1}$ berisi r elemen secara acak dari matriks $\boldsymbol{X}$ dengan jumlah elemen sebanyak:
   $$r = \frac{(n+p+1)}{2}$$
   dimana n dan p merupakan ukuran sampel dan jumlah variabel dalam masing-masing data.

b. Menghitung vektor rata-rata $\bar{\boldsymbol{X}}_{\boldsymbol{1}}$ dan matriks kovarian $\boldsymbol{G}_{\boldsymbol{1}}$ dari himpunan bagian $\boldsymbol{R}_{\boldsymbol{1}}$ dengan menggunakan rumus (2.16) yang hasilnya sebagai berikut:

$$\left.\begin{aligned} \bar{\boldsymbol{X}}_{\boldsymbol{1}} &= \frac{1}{r}\sum_{i\in K}^{r} x_i \\ \boldsymbol{G}_{\boldsymbol{1}} &= \frac{1}{r-1}\sum_{i=1}^{r}[x_i - \bar{\boldsymbol{X}}_{\boldsymbol{1}}]^T[x_i - \bar{\boldsymbol{X}}_{\boldsymbol{1}}] \end{aligned}\right\} \qquad (3.\ 1)$$

c. Menghitung determinan matriks G_1

d. Menghitung jarak relatif dari setiap pengamatan $\bar{\boldsymbol{X}}_{\boldsymbol{1}}$ dan kovarian $\boldsymbol{G}_{\boldsymbol{1}}$ menggunakan jarak mahalanobis $d(x_i, \bar{\boldsymbol{X}}_{\boldsymbol{1}}, \boldsymbol{G}_{\boldsymbol{1}}) = \sqrt{(x_i - \bar{\boldsymbol{X}}_{\boldsymbol{1}})' G^{-1}{}_1 (x_i - \bar{\boldsymbol{X}}_{\boldsymbol{1}})}$ untuk $i = 1, 2, ..., n$. Kemudian mengurutkan jarak mahalanobis setiap pengamatan dari yang bernilai terkecil hingga yang terbesar.

e. Mengambil himpunan bagian $\boldsymbol{R}_{\boldsymbol{2}}$ berisi elemen r dari objek yang memiliki jarak mahalanobis terkecil.

f. Mengulangi langkah *(b-e)* sehingga didapat himpunan bagian yang konvergen dan memiliki determinan matriks kovarians paling kecil, yaitu $|G_{n+1}| < |G_n|$

g. Berdasarkan elemen *r*, maka selanjutnya data diboboti dengan w_i bernilai 1 jika jarak $d^2\left(x_i, \bar{\boldsymbol{X}}_1, \boldsymbol{G}_{\boldsymbol{1}}\right) \le \chi^2_{p,1-\alpha}$ dan bernilai nol untuk nilai w_i lainnya

h. Langkah terakhir, berdasarkan elemen r tersebut, diperoleh penduga *fast MCD* yaitu vektor rataan $\bar{x}_{MCD}$ dan matriks kovarian G_{MCD}, dengan determinant matriks kovarian paling kecil. Dirumuskan sebagai berikut:

$$\left.\begin{aligned} \bar{x}_{MCD} &= \frac{\sum_{i=1}^{n} w_i x_{ij}}{\sum_{i=1}^{n} w_i} \\ G_{MCD} &= \frac{\sum_{i=1}^{n} w_i (x_{ij} - \bar{x}_{MCD})^T (x_{ij} - \bar{x}_{MCD})}{\sum_{i=1}^{n} w_i} \end{aligned}\right\} \qquad (3.2)$$

Untuk memperoleh $\bar{\boldsymbol{x}}_{MCD}$ dan $\boldsymbol{G}_{MCD}$, digunakan *Software* Rstudio dengan *Robustbase Package*

5. Melakukan penguraian nilai singular atau *Singular Value Decomposition (SVD)* dengan menggunakan matriks kovarian $\boldsymbol{G}_{MCD}$ yang diperoleh melalui algoritma *Fast MCD* pada langkah ke-3
6. Membuat grafik Biplot *Robust* dan menghitung kesesuaian grafik. Grafik *Biplot Robust* dapat diperoleh dengan bantuan *software* Rstudio
7. Menginterpretasikan hasil grafik Biplot *Robust*, kemudian mengambil kesimpulan.

V. KESIMPULAN

Berdasarkan hasil dan pembahasan yang telah dijelaskan pada bagian sebelumnya tentang Pemetaan Provinsi di Indonesia berdasarkan perilaku peduli lingkungan hidup dapat diambil kesimpulan sebagai berikut:

1. Pemetaan provinsi di Indonesia mengenai perilaku peduli lingkungan hidup terbentuk sembilan kelompok provinsi berdasarkan kemiripan karakteristiknya yaitu, Kelompok pertama hanya terdiri dari satu provinsi yaitu Provinsi DKI Jakarta karena letaknya jauh dari provinsi lainnya. Artinya Provinsi DKI Jakarta memiliki karakteristik tersendiri. Karakteristik perilaku peduli lingkungan hidup yang dominan untuk provinsi DKI Jakarta adalah perlakuan terhadap sampah. Kelompok kedua terdiri dari Provinsi Jawa Tengah, Jawa Timur, Banten dan Bali. Keempat provinsi ini memiliki karakteristik perilaku peduli lingkungan hidup yang dominan yaitu perlakuan terhadap sampah lebih baik dibandingkan dengan provinsi lainnya. Hal ini dikarenakan kelompok ini memiliki letak searah vektor variabel perlakuan terhadap sampah. Kelompok ketiga terdiri dari Provinsi DI Yogyakarta memiliki karakteristik perilaku peduli lingkungan hidup yang dominan yaitu penggunaan lampu hemat energi. Kelompok keempat terdiri dari Provinsi Aceh, Sumatera Utara, Riau, NTB, Kalimantan Barat dan Kalimantan Tengah. Kelompok ini memiliki karakteristik perilaku peduli lingkungan hidup yaitu penggunaan gas sebagai bahan bakar utama memasak. Kelompok kelima terdiri dari Provinsi Sumatera Selatan, Bengkulu, Lampung, Kep. Bangka Belitung, Kepulauan Riau dan Jawa Barat. Kelompok ini memiliki karakteristik perilaku peduli lingkungan hidup yaitu penggunaan gas sebagai bahan bakar utama memasak. Kelompok keenam terdiri Provinsi Sulawesi Selatan yang memiliki karakteristik perilaku peduli lingkungan hidup yaitu akses sumber air minum layak dan akses sanitasi layak. Kelompok ketujuh terdiri dari Provinsi Sumatera Barat,

Jambi dan NTT. Ketiga kelompok ini memiliki karakteristik perilaku peduli lingkungan hidup yaitu penggunaan gas sebagai bahan bakar utama memasak. Kelompok kedelapan terdiri dari Provinsi Kalimantan Selatan, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tengah, Sulawesi Tenggara, Gorontalo, Sulawesi Barat dan Maluku Utara. Kelompok ini memiliki karakteristik perilaku peduli lingkungan hidup yaitu akses hunian layak dan Perlakuan terhadap sampah yang di angkut pada kelompok ini masih sangat rendah dibadingkan dengan provinsi yang lain. Kelompok kesembilan terdiri dari Provinsi Maluku, Papua Barat dan Papua. Ketiga Provinsi ini memiliki karakteristik perilaku peduli lingkungan hidup yaitu akses hunian yang layak.

2. Tingkat kesesuaian biplot yang diperoleh sebesar 70,45% yang artinya pemetaan Provinsi di Indonesia berdasarkan perilaku peduli lingkungan hidup dapat memberikan informasi yang cukup baik mengenai data asalnya.

DAFTAR PUSTAKA

Badan Pusat Statistik BPS-Statistics Indonesia. (2014). *Indikator Perilaku Peduli Lingkungan Hidup 2014*. Badan Pusat Statistika, Jakarta.

Badan Pusat Statistik- Statistik Indonesia. 2021. *Statistik Lingkungan Hidup Indonesia 2021*. Badan Pusat Statistik, Jakarta.

Badan Pusat Statistik. 2013. *Survei Perilaku Peduli Lingkungan Hidup 2013*. BPS, Jakarta.

Barnett, V & Lewis, T. 1978. *Outliers in Statistical Data*. John Wiley & Sons, New York.

Dayanti, N.P., Suciptawati, N.L., & Susilawati, M. 2016. Penerapan Bootstrap dalam Metode Minimum Covariance Determinant (MCD) dan Least Median Square (LMS) pada Analisis Regresi Linear Berganda. *E-Jurnal Matematika*. **5**(1): 22-26.

Dwipurwani, O., Cahyawati, D., & Susanti, E. (2022). Analisis Biplot Robust dengan Metode Minimum Covariance Determinant dalam Mendeskripsikan Provinsi Sumatera Selatan Berdasarkan Karakteristik Angkatan Kerja Menganggur Dari Aspek Gender. *Euler: Jurnal Ilmiah Matematika, Sains dan Teknologi.* 10(**1**). 54-65.

Gabriel, K. R. (1971). The biplot graphic display of matrices with application to principal component analysis. *Biometrika*. 58(**3**). 453-467.

Hair, J.F., Anderson, R.E., Tatham, R.L, & Black, W.C. 1998. *Multivariate Data Analysis. Prentice-Hall*, New Jersey.

Hubert, M., Rousseeuw, P.J., & Aelst, S.V. 2008. High-Breakdown Robust Multivariate Methods. *Statistical Science*. 23(**1**): 92-119.

Johnson, R.A. & Wichern, D.W. (2007). *Applied Multivariate Statistical Analysis*. 6th Edition. Prentice Hall, New Jersey.

Kriegel, H. P., Kröger, P., & Zimek, A. (2010). Outlier detection techniques. *Tutorial at KDD*. **10.** 1-76.

Mattjik, A.S. & Sumertajaya, I.M. 2011. *Sidik Peubah Ganda Dengan Menggunakan SAS*. IPB Press, Bogor.

Nugroho, S. 2008. *Statistika Multivariat Terapan*. Bengkulu: UNIB Press.

Rencher, A.C. 2002. *Methods of Multivariate Analysis*. 2th Edition. John Wiley & Sons, New York.

Rousseeuw, P.J. & Driessen, K.V. 1999. A Fast Algorithm for the Minimum Covariance Determinant Estimator. *Technometrics Journal*. 41(**3**). 212- 223.

Seber, G. A. (2009). *Multivariate observations*. John Wiley & Sons.

Solimun, dan Fernandes, A. A. R. (2008). *Modul Pelatihan Multivariate Analysis 6&8 Juni 2008*. Malang: LPM UB.

Undang-Undang Republik Indonesia Nomor 32 Tahun 2009 Tentang Perlindungan dan Pengelolaan Lingkungan Hidup (2009).

Soemartini. 2007. *Outlier-Pencilan*. Jurusan Statistika FMIPA Universitas Padjajaran. Jatinangor.

Venelia, H., Nisa, K., Wibowo, R. A., & Muda, M. A. (2021). Robust Biplot Analysis of Natural Disasters in Indonesia from 2019 to 2021. *Jurnal Aplikasi Statistika & Komputasi Statistik*. 13(**2**). 61-68.