

**PERBANDINGAN ANALISIS KOMPONEN UTAMA DAN ROBUST PCA
(ROBPCA)
(STUDI KASUS: PADA ANALISIS DATA RATA-RATA PENGELUARAN
PER KAPITA SEBULAN UNTUK KOMODITAS MAKANAN DI
PROVINSI SULAWESI SELATAN)**

Drajat Indra Purnama¹, Pardomuan Robinson Sihombing*²

¹Badan Pusat Statistik (BPS) Parigi Moutong

²Badan Pusat Statistik (BPS) Jakarta

email: robinson@bps.go.id

Abstract: *South Sulawesi is a province in eastern Indonesia which is the gateway to several regions or cities in Eastern Indonesia. The level of welfare of the South Sulawesi population can be measured, one of which is by calculating the expenditure to meet their daily needs, both food and non-food needs. This research will analyze the average monthly per capita expenditure in fourteen food groups. To facilitate the analysis of the large number of variables, it is necessary to simplify the number of variables using the main component analysis with the classical PCA and ROBPCA methods. The results obtained by the ROBPCA method provide better performance than classic PCA, this can be seen from the ROBPCA model is able to produce a smaller number of main components than the original variable, with only 3 components being able to explain up to 80.69 percent.*

Keyword: *Expenditure, Food, PCA, ROBPCA, South Sulawesi*

Abstrak: Sulawesi Selatan merupakan provinsi di wilayah timur Indonesia yang merupakan pintu gerbang menuju beberapa wilayah atau kota di Indonesia Timur. Tingkat kesejahteraan penduduk Sulawesi Selatan dapat diukur salah satunya dengan menghitung pengeluaran dalam memenuhi kebutuhan hidupnya baik kebutuhan makanan maupun bukan makanan. Pada penelitian ini akan dilakukan analisis terhadap rata-rata pengeluaran per kapita sebulan pada empat belas kelompok makanan. Untuk memudahkan analisis terhadap variabel yang jumlahnya banyak maka perlu dilakukan penyederhanaan jumlah variabel menggunakan analisis kompone utama dengan metode PCA klasik dan ROBPCA. Hasil yang didapat metode ROBPCA memberikan performa yang lebih baik dari pada PCA klasik, hal ini dapat dilihat model ROBPCA mampu menghasilkan jumlah komponen utama yang lebih sedikit daripada variabel asalnya, hanya dengan 3 komponen sudah mampu menjelaskan hingga 80,69 persen.

Katakunci: Makanan, Pengeluaran, PCA, ROBPCA, Sulawesi Selatan

PENDAHULUAN

Sulawesi Selatan merupakan provinsi di wilayah timur Indonesia yang merupakan pintu gerbang menuju beberapa wilayah atau kota di Indonesia Timur. Sulawesi Selatan juga merupakan pusat perekonomian dan perdagangan di Indonesia Timur. Pertumbuhan ekonomi di Sulawesi Selatan pada Tahun 2019 adalah 6,92 dengan ditopang oleh tiga sektor terbesar yaitu Pertanian, Kehutanan dan Perikanan (21,28 persen), Perdagangan (14,80 persen) dan Konstruksi (14,188 persen) (Badan Pusat Statistik Provinsi Sulawesi Selatan, 2020). Kondisi perekonomian yang baik di Sulawesi Selatan berimbas pada kesejahteraan penduduknya. Indeks Pembangunan Manusia (IPM) menunjukkan angka peningkatan di Tahun 2019 (71,66) dibanding Tahun 2018 (70,90). Sedangkan persentase angka kemiskinan turun di Tahun 2019 (8,56 persen) dibanding 2018 (8,87 persen).

Tingkat kesejahteraan penduduk Sulawesi Selatan dapat diukur salah satunya dengan menghitung pengeluaran dalam memenuhi kebutuhan hidupnya. Survei Sosial Ekonomi Nasional (Susenas) merupakan survei terhadap rumah tangga yang diselenggarakan oleh BPS sebagai salah satu sumber informasi untuk mendapatkan gambaran mengenai kondisi sosial ekonomi masyarakat. Besar pengeluaran yang diperoleh oleh Susenas merupakan pengeluaran untuk satu rumah tangga sehingga untuk memperoleh pengeluaran per kapita maka pengeluaran untuk konsumsi semua anggota rumah tangga dibagi dengan banyaknya anggota rumah tangga. Pada Susenas pengeluaran rumah tangga dibedakan menjadi dua yaitu pengeluaran komoditas makanan dan komoditas bukan makanan. Komoditas makanan terdiri dari empat belas kelompok makanan mulai dari beras sampai rokok.

Pada penelitian ini akan dilakukan analisis terhadap rata-rata pengeluaran per kapita sebulan pada empat belas kelompok makanan tersebut. Untuk memudahkan analisis terhadap variabel yang jumlahnya banyak maka perlu dilakukan penyederhanaan jumlah variabel menggunakan analisis kompone utama. Analisis komponen utama merupakan analisis umum yang menggunakan prinsip matematika dasar dengan mengubah sejumlah variabel yang mungkin berkorelasi menjadi lebih sedikit variabel yang disebut komponen utama (Mishra, et al., 2017). Dalam analisis komponen utama variabel asal yang berjumlah p akan direduksi menjadi k komponen utama. Akan tetapi metode analisis komponen utama klasik sangat dipengaruhi oleh kehadiran pencilan dikarenakan analisis komponen utama, didasarkan pada matriks kovarian yang sensitif terhadap pencilan. Hal ini dapat diatasi menggunakan metode analisis komponen utama yang kekar (*robust*) terhadap data pencilan.

Metode yang dapat digunakan untuk mengatasi data pencilan adalah *Robust Principle Component Analysis* (ROBPCA). ROBPCA menggabungkan konsep *Projection Pursuit* (PP) dan penduga kovarian *Minimum Covariance Determinant* (MCD) (Hubert, Rousseeuw, & Branden, 2005). Penelitian ini bertujuan untuk membandingkan hasil analisis komponen utama metode klasik dengan metode kekar *Robust Principle Component Analysis* (ROBPCA) pada data rata-rata pengeluaran per kapita sebulan pada empat belas kelompok makanan di Provinsi Sulawesi Selatan.

Uji Statistik *Bartlett Sphericity*

Hubungan atau korelasi antar variabel pada data multivariat dapat diuji menggunakan uji independensi (kebebasan) antar variabel dengan metode *bartlett sphericity*. Variabel X_1, X_2, \dots, X_p dikatakan bersifat independen (saling bebas) jika matriks korelasi antar variabel membentuk matriks identitas (Morrison, 1978). Hipotesis uji independensi antar variabel menggunakan uji *bartlett sphericity* adalah sebagai berikut.

$H_0 : \mathbf{R} = \mathbf{I}$ (Variabel independen tidak berkorelasi)

$H_1 : \mathbf{R} \neq \mathbf{I}$ (Variabel independen berkorelasi)

Statistik uji :

$$\chi^2 = \left\{ n - 1 - \frac{2p+5}{6} \right\} \ln |\mathbf{R}| \quad (1)$$

dimana

n : jumlah observasi

p : jumlah variabel independen

\mathbf{R} : matrik korelasi dari masing-masing variabel independen

Jika sudah ditentukan nilai α maka H_0 ditolak jika nilai

$$\chi_{hitung}^2 > \chi_{(\alpha, \frac{1}{2}p(p-1))}^2$$

Sehingga disimpulkan bahwa terdapat korelasi antar variabel independen.

Identifikasi Pencilan

Menurut Herdiani. *et al* (2019), Pencilan adalah pengamatan yang tidak mengikuti pola sebaran data dan dapat menyebabkan penyimpangan dari analisis data. Salah satu metode yang sering digunakan untuk mendeteksi pencilan adalah jarak Mahalanobis, yaitu dengan menghitung jarak dari setiap observasi pusat data. Adapun rumus perhitungan jarak Mahalanobis adalah

$$d_i^2 = (x_i - \mu)' \Sigma^{-1} (x_i - \mu) \quad i = 1, 2, \dots, n \quad (2)$$

dengan

x_i : Vektor random dari variabel independen

μ : Vektor rata-rata

Σ : Matriks kovarian.

Analisis Komponen Utama

Analisis komponen utama pertama kali dikemukakan oleh K Pearson sebagai alat *fitting* bidang dengan kuadrat terkecil ortogonal. Selanjutnya analisis kompoen utama dikembangkan oleh *T Hotelling* untuk tujuan menganalisa struktur korelasi (Johnson & Wichern, 2002).

Terdapat beberapa tujuan analisis komponen utama. Pertama, mereduksi variabel dengan menjaga sebanyak mungkin informasi dari data asal. Kedua, mendapatkan variabel baru yang memiliki bersifat *orthogonal*. Ketiga, mentransformasikan variabel independen asal yang saling berkorelasi menjadi variabel baru yang saling bebas, sehingga tidak terjadi korelasi antar variabel independennya.

Langkah-langkah penggunaan analisis komponen utama adalah:

1. Jika Σ matriks varian-kovarian dari variabel random x' . Pasangan nilai *eigen* dan vektor *eigen* Σ adalah $(\lambda_1, e_1), (\lambda_2, e_2), \dots, (\lambda_p, e_p)$ dimana $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$, maka komponen utama ke- i adalah:

$$Y_i = \underline{e}_i' \underline{x} = e_{i1}x_1 + e_{i2}x_2 + \dots + e_{ip}x_p, \quad i = 1, 2, \dots, p \quad (3)$$

$$\text{dengan } \text{Var}(Y_i) = \underline{e}_i' \Sigma \underline{e}_i = \lambda_i \quad (4)$$

$$\text{dan } \text{Cov}(Y_i, Y_k) = \underline{e}_i' \Sigma \underline{e}_k = 0. \quad (5)$$

2. Proporsi total varian yang diterangkan komponen utama ke- k adalah

$$\frac{\lambda_k}{\lambda_1 + \lambda_2 + \dots + \lambda_p}; \quad k = 1, 2, \dots, p \quad (6)$$

Korelasi komponen utama ke- i dengan variabel asal ke- k adalah

$$\rho_{Y_{ik}} = \frac{e_{ik}\sqrt{\lambda_i}}{\sigma_{ii}} ; i, k = 1, 2, \dots, p \quad (7)$$

3. Analisis komponen utama dari variabel yang distandarisasi adalah sebagai berikut $Z_1 = \left(\frac{x_1 - \mu_1}{\sqrt{\sigma_{11}}}\right); Z_2 = \left(\frac{x_2 - \mu_2}{\sqrt{\sigma_{22}}}\right); \dots; Z_p = \left(\frac{x_p - \mu_p}{\sqrt{\sigma_{pp}}}\right)$; atau dalam notasi matriks $Z = \left(v^{\frac{1}{2}}\right)^{-1} (\underline{x} - \mu)$ dengan $v^{\frac{1}{2}}$ matriks diagonal dengan elemen $\sqrt{\sigma_{ii}}$ maka kovariansnya $Z = \left(v^{\frac{1}{2}}\right)^{-1} \Sigma \left(v^{\frac{1}{2}}\right)^{-1} = \rho$. Komponen utama ke- i adalah $Y_i = \underline{e}'_i Z = \underline{e}'_i \left(v^{\frac{1}{2}}\right)^{-1} (\underline{x} - \mu)$. Proporsi varian yang diterangkan komponen ke- k adalah $\frac{\lambda_k}{q}$ dimana λ_k adalah nilai eigen dari ρ dan $q = \text{tr}(\rho)$.

ROBPCA (Robust PCA)

ROBPCA (ROBust PCA), atau juga disebut Hubert PCA diperkenalkan oleh Hubert et al (Hubert, Rousseeuw, & Branden, 2005). Metode ROBPCA ini adalah kombinasi dari konsep *Projection Pursuit* (PP) dengan penduga kovarian *Minimum Covariance Determinant* (MCD) yang pada tahun 1999 telah dimodifikasi bersama Van Driessen menjadi *Fast Minimum Covariance Determinant* (FASTMCD). Adapun kegunaan model ini untuk mendapatkan komponen utama yang tidak terpengaruh terlalu banyak dengan kehadiran data pencilan atau dengan kata lain robust terhadap data pencilan.

Adapun algoritma dalam Metode ROBPCA () adalah sebagai berikut:

1. Memilih $\frac{1}{2} < \alpha < 1$ untuk mendapatkan

$$h = \max\left(\alpha_n, \left\lfloor \frac{h+p+1}{2} \right\rfloor\right) \quad (8)$$

2. Menghitung outlyingness x_i setiap data dengan rumus Stanhel-Donoho

$$O(x_i) = \max\left(\alpha_n, \left\lfloor \frac{h+p+1}{2} \right\rfloor\right) \quad (9)$$

3. Matriks kovarian (S_o) dikomposisi sehingga diperoleh komponen utamanya.
4. Pada n, k dari algoritma ke-3 dihitung kembali penduga nilai tengah ($\hat{\mu}_2$) dan matriks kovarian MCD S_1 menggunakan FAST-MCD yang diadaptasi. Komponen utamaakhir adalah vector eigen dari matriks kovarian tersebut (S_1)

Kriteria Pemilihan Komponen Utama

Tujuan analisis komponen utama salah satunya adalah mereduksi dimensi data asal yang berjumlah p variabel independen menjadi k komponen utama ($k < p$). Menurut Sri Jumiati, Shantika Martha, Nurfitri Imro'ah (2018), terdapat 3 (tiga) cara dalam menentukan k komponen utama antara lain :

- a. Nilai *eigen* > 1
- b. Proporsi kumulatif varian dari p variabel independen asal yang dijelaskan oleh k komponen utama minimal 80 persen (Artigue & Smith, 2019).
- c. Berdasarkan *scree plot* (plot antara i dengan λ_i) yaitu melihat letak terjadinya belokan dengan menghapus komponen utama yang menghasilkan beberapa nilai *eigen* kecil membentuk pola garis lurus (Rencher, 1998). Bila kurva masih curam maka masih bisa dilakukan penambahan komponen, sebaliknya bila kurva sudah melandai maka penentuan jumlah komponen sudah selesai meskipun penentuan landai atau curamnya *scree plot* bersifat subjektif oleh setiap orang.

Jenis Data

Data yang digunakan dalam penelitian ini adalah data sekunder rata-rata pengeluaran perkapita sebulan menurut kelompok komoditas makanan pada Kabupaten atau Kota di Provinsi Sulawesi Selatan Tahun 2020. Variabel yang digunakan pada penelitian ini yang disebut sebagai variabel independen disajikan pada Tabel 1.

Tabel 1. Variabel penelitian

Variabel	Keterangan
X ₁	Padi-padian
X ₂	Umbi-umbian
X ₃	Ikan
X ₄	Daging
X ₅	Telur dan susu
X ₆	Sayur-sayuran
X ₇	Kacang-kacangan
X ₈	Buah-buahan
X ₉	Minyak dan kelapa
X ₁₀	Bahan minuman
X ₁₁	Bumbu-bumbuan
X ₁₂	Bahan makanan lainnya
X ₁₃	Makanan minuman jadi
X ₁₄	Rokok dan tembakau

Sumber Data

Sumber data dalam penelitian ini adalah data yang berasal dari Publikasi Statistik Kesejahteraan Rakyat 2020 pada masing-masing Badan Pusat Statistik (BPS) Kabupaten/Kota di Provinsi Sulawesi Selatan. Data tersebut merupakan hasil pendataan Survei Sosial Ekonomi Nasional (Susenas) Maret 2020.

Tahapan Prosedur Analisis Komponen Utama Robust

Tahapan prosedur dalam melakukan Analisis Komponen Utama Robust adalah sebagai berikut (Susilowati & Sihombing, 2020) :

1. Uji Korelasi antara variabel asal (X).
2. Identifikasi keberadaan Pencilan
3. Melakukan Analisis Komponen Utama
 - a. Penentuan nilai *eigen*
 - b. Penentuan proporsi kumulatif
 - c. Penentuan komponen utama
4. Melakukan Analisis Komponen Utama Robust menggunakan ROBPCA
 - a. Penentuan nilai *eigen*
 - b. Penentuan proporsi kumulatif
 - c. Penentuan komponen utama *robust*
5. Perbandingan analisis komponen utama dengan analisis komponen utama robust
6. Interpretasi model

HASIL DAN PEMBAHASAN

Uji Korelasi Menggunakan Uji Statistik Bartlett Sphericity

Uji korelasi digunakan untuk mengetahui hubungan antara variabel independen dalam penelitian ini adalah variabel komoditas makanan. Hasil uji korelasi variabel komoditas makanan adalah

- Hipotesis :

$H_0 : \mathbf{R} = \mathbf{I}$ (Variabel komoditas makanan tidak berkorelasi)

$H_1 : \mathbf{R} \neq \mathbf{I}$ (Variabel komoditas makanan berkorelasi)

- Nilai uji *bartlett sphericity* (*chi-square* hitung)

$$\chi_{hitung}^2 = 338,22$$

- Nilai *chi-square* tabel (derajat bebas 91 dan $\alpha = 0,05$)

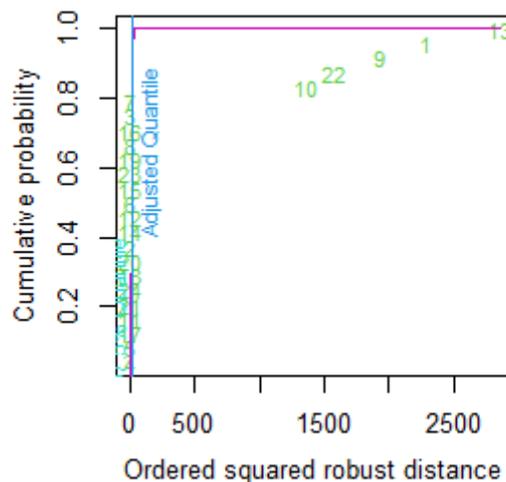
$$\chi_{(0,05,91)}^2 = 70$$

- Kesimpulan :

Karena nilai $\chi_{hitung}^2 = 338,22 > \chi_{(0,05,91)}^2 = 70$ maka dapat disimpulkan bahwa variabel komoditas makanan tidak berkorelasi secara simultan.

Identifikasi Pencilan

Tahap selanjutnya adalah mendeteksi keberadaan pencilan pada data secara multivariat. Proses pendeteksian pencilan pada data penelitian ini menggunakan program R 4.0.3. Hasil pendeteksian pencilan dapat dilihat pada gambar 1.



Gambar 1. *Chi-Square* Q-Q Plot Pendektian Pencilan

Pada Gambar 1 dapat dilihat bahwa terdapat 5 (lima) pengamatan yang teridentifikasi sebagai pencilan yaitu pengamatan ke-1, 9, 10, 13 dan 22. Kelima pengamatan tersebut berturut-turut adalah Kepulauan Selayar, Pangkep, Barru, Wajo dan Makassar.

Analisis Komponen Utama

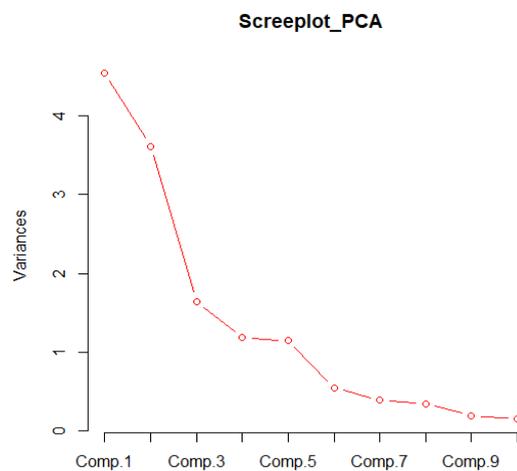
Salah satu tujuan dari analisis komponen utama adalah mereduksi variabel independen asal (komoditas makanan) yang berjumlah p (14 variabel) menjadi k komponen utama. Komponen utama yang terbentuk ditentukan dengan melihat nilai *eigen* yang nilainya lebih dari 1 dan proporsi kumulatif varian terhadap total varian mencapai 80 persen. Nilai *eigen* dan proporsi kumulatif varian untuk masing-masing komponen utama disajikan pada Tabel 2.

Pada Tabel 2 terlihat bahwa terdapat nilai *eigen* lebih dari 1 yaitu pada komponen utama pertama sampai kelima. Selain itu, dilihat dari proporsi kumulatif varian komponen utama pertama sampai kelima mampu menjelaskan keragaman data komoditas makanan sebesar 86,70 persen. Dapat disimpulkan bahwa komponen utama yang terbentuk sebanyak lima komponen utama. Untuk lebih memperjelas pembentukan ketiga komponen tersebut, berikut disajikan gambar *scree plot*.

Tabel 2. Nilai *eigen* dan proporsi kumulatif varian masing-masing komponen utama

Komponen utama	Nilai <i>eigen</i>	Proporsi kumulatif varian (Persen)
1	4,54552	32,47
2	3,61775	58,31
3	1,64547	70,06
4	1,18604	78,53
5	1,14272	86,70
6	0,54671	90,60
7	0,38478	93,35
8	0,34532	95,82
9	0,18791	97,159
10	0,15387	98,26
11	0,10694	99,02
12	0,08992	99,66
13	0,02431	99,84
14	0,02273	100,00

Berdasarkan Gambar 2, terlihat titik-titik yang menunjukkan nilai *eigen* masing-masing komponen utama. Komponen utama yang memenuhi kriteria proporsi kumulatif varians lebih dari 80 persen dapat dilihat pada titik yang berada di atas 1 atau nilai *eigen* lebih dari 1 yaitu komponen utama ke-1 sampai komponen utama ke-5. Jika digunakan pendekatan titik *elbow*, diperoleh bahwa pada komponen ke-6 kurva sudah mulai landai (sebelah kiri curam dan sebelah kanan landai) sehingga dapat diambil lima komponen utama. Model baru mencapai kumulatif varian 100 persen pada komponen ke-14.



Gambar 2. Scree Plot Analisis Komponen Utama

ROBPCA

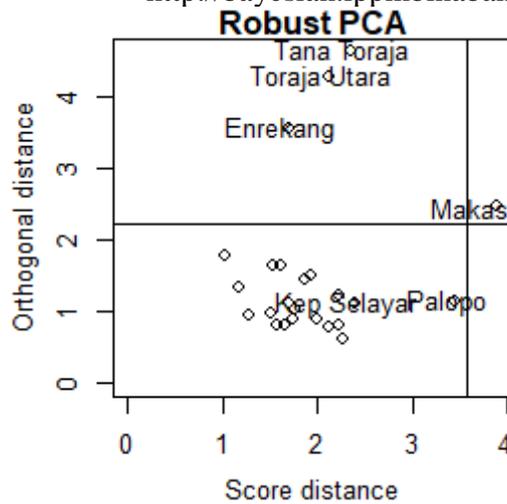
Hasil dari identifikasi pencilan menunjukkan keberadaan pencilan pada data yang dianalisis. Apabila data terdapat pencilan analisis komponen utama klasik bisa digunakan tetapi akurasiya berkurang karena matrik kovarian yang digunakan sensitive terhadap pencilan. Sehingga digunakanlah metode ROBPCA.

Tujuan dari ROBPCA adalah mereduksi variabel independen asal (komoditas makanan) yang berjumlah p yang mengandung pencilan (14 variabel) menjadi k komponen utama *robust*. Sama pada analisis komponen utama klasik pemilihan banyaknya komponen utama *robust* berdasarkan proporsi kumulatif varians lebih dari 80 persen. Nilai eigen dan proporsi kumulatif varian untuk masing-masing komponen utama *robust* disajikan pada Tabel 3.

Tabel 3. Nilai *eigen* dan proporsi kumulatif varian masing-masing komponen utama *robust*

Komponen utama	Nilai <i>eigen</i>	Proporsi kumulatif varian (Persen)
1	2,1936	37,24
2	2,0052	68,36
3	1,2624	80,69
4	1,1830	91,52
5	1,04658	100,00

Pada Tabel 3 terlihat bahwa terdapat nilai *eigen* lebih dari 1 yaitu pada komponen utama *robust* pertama sampai kelima. Proporsi kumulatif varians lebih dari 80 persen terbentuk dengan komponen utama *robust* pertama sampai komponen utama *robust* ketiga dengan nilai 80,69 persen. Sedangkan untuk komponen utama *robust* pertama sampai komponen utama *robust* kelima memiliki proporsi kumulatif varians 100 persen. Artinya bahwa informasi dari 14 variabel independen asal sudah sepenuhnya terwakilkan oleh 5 komponen utama *robust* yang terbentuk. Dari hasil ini dapat dilihat bahwa model ROBPCA lebih baik dari PCA kalsik karena hanya dengan jumlah komponen yang lebih sedikit mampu menjelaskan variasi yang leboh besar, bahkan mencapau 100 persen hanya dengan 5 komponen. Hal yang sama didapat oleh Yuliani, Sukarsa dan Srinadi (Yuliani, Sukarsa, & Srinadi, 2013) sehingga yang menyatakan bahwa ROBPCA lebih baik daripada metode PCA kalsik sehingga ROBPCA memiliki nilai estimasi yang lebih baik dibandingkan dengan metode klasik. Hal ini njuga didukung oleh Sunaryo dan Siagian (Sunaryo & Siagian, 2011) yang menyatakan bahwa hasil model ROBPCA memberikan performa yang lebih baik dari PCA klasik.



Gambar 3. Output RobPCA

Dari plot di atas terlihat bahwa observasi Enrekang, Tana Toraja dan Toraja Utara, merupakan observasi yang outlier orthogonal karena mempunyai OD (*Orthogonal Distance*) besar ($>cut-off\ point$), dan SD (*Score Distance*) kecil ($< cut-off\ point$). Daerah Kepulauan Selayar, Palopo, dan Makassar merupakan *bad leverage point* karena OD bernilai kecil namun SD besar.

Tabel 4. Nilai *Loading* Masing-Masing Variabel Pada Komponen RobPCA

Variabel/ loading	PC1	PC2	PC3
[1,]	0.255	-0.221	0.117
[2,]	0.454	-0.356	-0.107
[3,]	0.264	0.222	0.144
[4,]	0.120	-0.061	-0.010
[5,]	0.168	0.187	0.278
[6,]	0.402	-0.088	0.054
[7,]	0.238	0.021	-0.047
[8,]	0.260	-0.153	-0.494
[9,]	0.268	0.346	-0.129
[10,]	0.361	-0.247	0.270
[11,]	0.213	0.120	-0.483
[12,]	0.189	0.523	-0.025
[13,]	0.135	0.481	-0.127
[14,]	0.169	0.072	-0.537

Dari Tabel 4, nilai Y_1 dapat diterangkan paling baik oleh variabel 2 (umbi-umbian) dan variabel 6 (sayur-sayuran) dengan korelasi antara dengan kedua variabel tersebut adalah positif. Hal ini berarti apabila nilai kedua variabel tersebut meningkat maka nilai Y_1 juga akan meningkat. Nilai Y_2 dapat diterangkan paling baik oleh variabel 12 (bahan makanan lainnya) dan variabel 13 (makanan jadi) dengan korelasi antara dengan kedua variabel tersebut adalah positif. Hal ini berarti apabila nilai kedua variabel tersebut meningkat maka nilai Y_2 juga akan meningkat. Sedangkan Nilai Y_3 dapat diterangkan paling baik oleh variabel 8 (buah-buahan) dan variabel 11 (bumbu-bumbu) dengan korelasi antara dengan kedua variabel tersebut adalah negatif. Hal ini berarti apabila nilai kedua variabel tersebut meningkat maka nilai Y_3 juga akan menurun.

KESIMPULAN

Berdasarkan penelitian yang telah dilakukan dan diuraikan pada pembahasan sebelumnya, dapat ditarik kesimpulan bahwa pada penelitian terhadap rata-rata pengeluaran per kapita sebulan pada empat belas kelompok makanan, metode ROBPCA memberikan performa yang lebih baik dari pada PCA klasik, hal ini dapat dilihat model ROBPCA mampu menghasilkan jumlah komponen utama yang lebih sedikit daripada variabel asalnya, hanya dengan 3 komponen sudah mampu menjelaskan hingga 80,69 persen.

DAFTAR PUSTAKA

- Artigue, H., & Smith, G. (2019). The Principal Problem With Principal Components Regression. *Cogent Mathematics & Statistics*, 6(1).
- Badan Pusat Statistik Provinsi Sulawesi Selatan. (2020). *Produk Domestik Regional Bruto Provinsi Sulawesi Selatan Menurut Lapangan Usaha 2015-2019*. Makassar: Badan Pusat Statistik.
- Herdiani, E., Sari, P., & Sunusi, N. (2019). Detection of Outliers in Multivariate Data using Minimum Vector Variance Method. *Journal of Physics: Conference Series* 1341.
- Hubert, M., Rousseeuw, P., & Branden, K. (2005). ROBPCA: a New Approach to Robust Principal Component Analysis. *American Statistical and the American Society for Quality Technometric*, 7(1).
- Johnson, R., & Wichern, D. (2002). *Applied Multivariate Statistical Analysis, 5th edition*. Pearson Education International.
- Jumiati, S., Martha, S., & Imro'ah, N. (2018). Penerapan Analisis Komponen Utama untuk Mereduksi Variabel dalam Pengukuran Desain Helm . *Buletin Ilmiah Mat, Stat, dan Terapannya (Bimaster)*, 7(3), 225-230.
- Mishra, S. P., Sarkar, U., Taraphder, S., Datta, S., Swain, D. P., Sasmita, R. S., & Laishram, P. M. (2017). Mishra, S.P, et all. Multivariate Statistical Data Analysis-Principal Component Analysis (PCA). *International Journal of Livestock Research.*, 7(5), 60-78.
- Morrison, D. (1978). *Multivariate Statistical Methods, Second edition*. New York : McGraw-Hill.
- Rencher, A. (1998). *Multivariate Statistical Inference and Application*. Brigham: Wiley-Interscience Publication.
- Sunaryo, S., & Siagian, T. H. (2011). Mengatasi Masalah Multikolinearitas Dan Outlier Dengan Pendekatan ROBPCA. *Jurnal Matematika, Saint dan Teknologi*, 12(1), 1-10.
- Susilowati, B. E., & Sihombing, P. R. (2020). Metode ROBPCA (Robust Principal Component Analysis) dan Clara (Clustering Large Area) pada Data dengan Outlier. *Jurnal Ilmu Komputer*, 13(2).
- Yuliani, N. W., Sukarsa, K. G., & Srinadi, I. G. (2013). Perbandingan Regresi Komponen Utama dan RobPCA Dalam Mengatasi Multikolinearitas dan Pencilan Pada Regresi Linear Berganda. *E-Jurnal Matematika*, 2(4), 1-5.