JURNAL FOURIER | Oktober 2021, Vol. 10, No. 2, 89-98 ISSN 2252-763X
DOI: 10.14421/fourier.2021.102.89-98 E-ISSN 2541-5239

Perbandingan Analisis Faktor Klasik dan Robust secara Empiris dan Simulasi (Studi kasus: Mobilitas Aktivitas Masyarakat Setiap Provinsi Di Masa Covid-19)

Yogi Anggara[1*], Epha Diana Supandi[2]

[1,2]*Program Studi Matematika, Fakultas Sains dan Teknologi, UIN Sunan Kalijaga Yogyakarta.*
Jl. Marsda Adisucipto No. 1 Yogyakarta, 55281, Indonesia.

Korespondensi: Yogi Anggara, Email: yogianggaraaa@gmail.com

Abstrak

Analisis faktor adalah kajian pada statistika yang digunakan untuk menggambarkan varians antara variabel berkorelasi yang diamati dalam jumlah yang lebih kecil dari variabel yang diamati yang mana disebut sebagai faktor. Ketika dihadapkan dengan data yang berisikan pencilan didalamnya, kerap kali model faktor yang terbentuk melenceng dari hasil yang diharapkan. Alternatif yang dapat digunakan adalah dengan mengestimasi kovarian yang *robust* terhadap pencilan menggunakan Minimum Covariance Determinant (MCD). Data yang digunakan pada penelitian ini adalah data mobilitas provinsi di Indonesia yang diambil melalui Google pada tanggal 2 Maret 2020 sampai 2 Maret 2021. Dari hasil analisis, diperoleh faktor optimal pada data mobilitas sebanyak 3 faktor, dengan varians kumulatif sebesar 84,3%. Dengan membandingkan sensitivitas loading factor antara model analisis faktor *robust* menggunakan MCD dengan analisis faktor klasik, didapatkan bahwa loading factor analisis faktor *robust* menggunakan MCD lebih stabil dibandingkan analisis faktor klasik.

Kata Kunci: Analisis Faktor; MCD; *Robust*; COVID-19

Abstract

Factor analysis is the study of statistics used to describe the variance between correlated variables that are observed in a smaller number of observed variables which are referred to as factors. When faced with data that contains outliers in it, the factor model formed often deviates from the expected results. An alternative that can be used is to estimate a robust covariance for outliers using the Minimum Covariance Determinant (MCD). The data used in this study is mobility data for provinces in Indonesia taken via Google from March 2 2020 to March 2 2021. From the analysis results, the optimal factor for mobility data is 3 factors, with a cumulative variance of 84.3%. By comparing the sensitivity of factor loading between robust factor analysis models using MCD and classical factor analysis, it is found that the factor loading of robust factor analysis using MCD is more stable than classical factor analysis.

Keywords: Factor Analysis; MCD; Robust; COVID-19

Pendahuluan

Coronavirus Disease (COVID-19) adalah penyakit yang disebabkan oleh *Severe Acute Respiratory Syndrome Coronavirus-2* (SARS-CoV-2). Kasus pertama ditemukan di Wuhan pada Desember 2019. Kasus COVID-19 terus mengalami peningkatan hingga akhirnya menginfeksi banyak negara [1]. Pada tanggal 2 Maret 2020, Indonesia melaporkan dua kasus pertama COVID-19 di Depok [2]. Kasus ini menyebabkan terguncangnya stabilitas beberapa sektor, antara lain sektor bisnis, yaitu pada pergerakan harga saham [3]. Untuk menekan angka kenaikan COVID-19, tentunya harus memahami faktor apa yang menyebabkan peningkatan mobilitas di beberapa lokasi. Pemahaman terkait mobilitas aktivitas masyarakat tentunya dapat membantu memahami kebijakan-kebijakan apa yang perlu dilakukan di tiap daerah.

© 2021 JURNAL FOURIER

Pada ilmu statistika, terdapat analisis yang dapat digunakan untuk memahami faktor apa yang mempengaruhi suatu data, yaitu analisis faktor. Tujuan penting dari analisis faktor adalah untuk menggambarkan hubungan kovarians antara banyak variabel, bila memungkinkan, dalam hal variabel acak yang mendasari tetapi tidak dapat diamati disebut faktor. Pada dasarnya, model faktor dilatarbelakangi oleh pembahasan berikut: Misalkan variabel dapat dikelompokkan menurut korelasinya. Asumsikan bahwa semua variabel dalam kelompok tertentu berkorelasi tinggi satu sama lain, tetapi berkorelasi relatif buruk dengan variabel dalam kelompok lain. Dengan demikian, dapat ditentukan bahwa setiap set variabel mewakili konstruk atau faktor yang mendasari hasil dari nilai korelasi yang diamati [4]. Analisis faktor adalah suatu studi tentang saling ketergantungan antar variabel yang bertujuan untuk menemukan sekumpulan variabel baru yang lebih sedikit dari variabel aslinya dan menunjukkan variabel asal mana yang merupakan faktor persekutuan [5]. Analisis faktor merupakan bagian dari analisis multivariat, yaitu suatu metode analisis statistik yang memperlakukan sekumpulan variabel yang berkorelasi sebagai suatu sistem dan memperhitungkan korelasi antar variabel tersebut [5]. Analisis faktor mengelompokan variabel-variabel ke dalam suatu faktor yang memiliki karakteristik yang hampir sama, sehingga dapat mempermudah analisis data. Dalam prosesnya, analisis faktor memanfaatkan kovarian variabel asal. Dengan demikian, pencilan pada data sangat berpengaruh dalam proses pembentukan faktor. Untuk mengatasi hal tersebut, kovarian variabel asal dapat diestimasi dengan metode yang *robust*, seperti *Minimum Covariance Determinant* (MCD). Pada penelitian sebelumnya, dengan membandingkan sensitivitas loading factor antara analisis faktor *robust* menggunakan MCD ddengan analisis faktor klasik, didapatkan hasil bahwa *loading factor* yang dibentuk oleh metode MCD lebih stabil terhadap pencilan dibandingkan dengan metode klasik [6]. Beberapa penelitian yang menggunakan MCD dalam pembentukan analisis faktor *robust* di antaranya ialah terkait analisis faktor inflasi kelompok bahan makanan di Jawa Tengah oleh Erna dan Abdul Mukid [7]. Dari penelitian tersebut diperoleh hasil bahwa analisis faktor *robust* menghasilkan proporsi kumulatif varian total lebih besar dibandingkan dengan analisis faktor klasik.

Pada penelitian ini, akan dilakukan analisis faktor *robust* dengan MCD pada data mobilitas aktivitas masyarakat di beberapa lokasi yang telah disediakan oleh Google. Dari hasil tersebut, ditentukan faktor-faktor yang mempengaruhi setiap variabel. Selain itu dilakukan juga simulasi sensitivitas *loading factor* antara analisis faktor *robust* menggunakan MCD dengan analisis faktor klasik. Data simulasi menggukan data mobilitas yang telah ditambahkan pencilan. Analisis data dan proses perhitungan menggunakan bahasa pemrogran R versi 3.6.3 dan *package* robustfa versi 1.0-5 [8]. R adalah bahasa pemrograman yang sangat kuat yang dapat melakukan berbagai analisis, yang mana sangat cocok dalam melakukan berbagai proses perhitungan metode statistika [9]. Penelitian ini diharapkan dapat membantu dalam memahami faktor apa saja yang mempengaruhi mobilitas di masa pandemi.

Landasan Teori

Pembahasan pada bagian ini dibagi menjadi dua bagian, bagian pertama menjelaskan mengenai data yang digunakan dalam penelitian, sementara bagian berikutnya menjelaskan metode yang digunakan dalam penelitian ini.

Data

Data pada penelitian ini diambil dari data laporan mobilitas harian masyarakat setiap provinsi di beberapa titik lokasi yang direkap oleh Google melalui aktivitas masyarakat pada Google Maps. Pengukuran dilakukan dengan cara menghitung perubahan mobilitas setiap harinya dengan nilai acuannya, yaitu median data mobilitas di tanggal 3 Januari sampai 6 Februari 2020. Data ini menunjukan perubahan kunjungan masyarakat ke berbagai tempat, seperti ritel dan rekreasi, toko makanan dan apotek, taman, transportasi umum, tempat kerja, hingga pemukiman. Lokasi ritel dan rekreasi terdiri dari tempat-tempat

seperti restoran, kafe, pusat perbelanjaan, taman hiburan, museum, perpustakaan, dan bioskop. Toko makanan dan apotek terdiri dari tempat-tempat seperti supermarket, toko grosir makanan, pasar tradisional, toko makanan khusus, toko obat, dan apotek. Taman terdiri dari tempat-tempat seperti taman lokal, taman nasional, pantai umum, dermaga, taman hewan peliharaan, lapangan terbuka, dan taman umum. Pusat transportasi umum terdiri dari tempat-tempat seperti pusat transportasi umum misalnya stasiun KRL, terminal bus, dan stasiun kereta api. Pada penelitian ini, data diambil dati tanggal 2 Maret 2020 sampai 2 Maret 2021 yang mana kemudian dirata-ratakan.

Tabel 1. Data rata-rata mobilitas pada lima provinsi pertama

	Ritel dan rekreasi	Toko makanan dan apotek	Taman	Pusat transportasi umum	Tempat kerja	Pemukinan
Aceh	-11.128	1.163	-14.074	-27.713	-17.303	2.705
Bali	-36.730	-26.728	-44.782	-63.439	-35.970	13.681
Bangka Belitung	-22.005	-1.632	-10.943	-38.888	-19.703	6.071
Banten	-17.589	-5.872	-3.191	-46.087	-24.016	13.240
Bengkulu	-9.308	3.747	-11.963	-29.841	-19.989	5.807

Metode

Penjelasan mengenai metode yang digunakan dalam penelitian ini terbagi menjadi dua, pertama akan dijelaskan mengenai analisis faktor, dan pembahasan dilanjutkan dengan penejelasan mengenai *minimum covariant determinan.*

1. **Analisis Faktor**

Analisis faktor (*Factor Analysis*, FA) adalah teknik untuk mempelajari keterkaitan antar variabel. Misalkan kita memiliki satu set variable $X_1, X_2, \dots, X_p$. Tujuan FA adalah untuk mewakili masing-masing variabel ini sebagai kombinasi linier dari sekumpulan faktor umum yang lebih kecil ditambah faktor unik untuk masing-masing variabel respons. Kombinasi linier ini dapat diwakili oleh persamaan berikut:

$$X_i = l_{i1}F_1 + l_{i2}F_2 + \cdots + l_{ik}F_k + e_i; \;\; i = 1,2, \dots, p \tag{1}$$

Dimana

l_{ij} : Loading factor untuk variabel ke-i dari faktor ke-j.
$F_1, F_2, \dots, F_k$: Faktor persekutuan,
$e_1, e_2, \cdots, e_i$: Faktor unik atau spesifik
k : Banyaknya faktor persekutuan dimana $k \leq p$.

Persamaan (1) juga dapat ditulis sebagai:

$$X = LF + \epsilon \tag{2}$$

Dimana,

$X = \left(X_1, X_2, \dots, X_p\right)^T$: Set variabel
$L \in \mathbb{R}^{p \times k}$: Matriks *loading factor*
$F = \left(F_1, F_2, \dots, F_p\right)^T$: Matriks faktor persekutuan
$\epsilon = \left(e_1, e_2, \dots, e_p\right)^T$: Matriks faktor unik atau spesifik

Model analisis faktor mengasumsikan bahwa:

1. vektor acak L dan F saling bebas
2. $E(F) = E(\epsilon) = 0$
3. $cov(F) = 1$
4. $cov(\epsilon) = \varphi$ di mana $\varphi = \left(\varphi_1, \varphi_2, \dots, \varphi_p\right)^T$ adalah matriks diagonal.

Misalkan $cov(X) = \Sigma$, maka dengan menggunakan kondisi yang diberikan pada F, diperoleh

$$cov(X) = cov(LF + \epsilon)$$
$$cov(X) = Lcov(F)L^T + cov(\epsilon)$$

atau

$$\Sigma = LL^T \quad (3)$$

Setelah dekomposisi pada persamaan (2) dan (3) dilakukan maka diperoleh estimasi $\hat{L}$ untuk L dan estimasi $\hat{\varphi}$ untuk φ; sehingga menghasilkan perkiraan $\hat{F}$ untuk F. Jadi, model FA bergantung pada dekomposisi matriks kovarians dari X. Dalam analisis faktor klasik, matriks kovarians sampel digunakan untuk memperkirakan matriks kovarians dari X. Misalkan jumlah pengamatan adalah n maka matriks kovarians sampel dapat dihitung sebagai $\hat{\Sigma} = \frac{1}{n}(x_i - \hat{\mu})(x_i - \hat{\mu})^T$, di mana $\hat{\mu} = \frac{1}{n}\sum_{i=1}^{p} r_i$.

Pendekatan klasik ini memiliki kemampuan yang sangat baik jika datanya berdistribusi normal multivariat. Namun, jika distribusi data menyimpang dari distribusi ideal ini, kovarians yang diperkirakan dapat menjadi sangat bias [10].

2. *Minimum Covariance Determinant*

Untuk membangun model yang lebih kuat, dapat dilakukan dengan cara mengestimasi Σ pada persamaan (2.3) menggunakan *scatter matrix* yang *robust*. Estimasi parameter L dan φ yang belum diketahui kemudian dapat diperoleh dengan cara melakukan dekomposisi matriks kovarians yang diestimasi secara *robust*.

Salah satu cara yang dapat dilakukan adalah dengan menggunakan estimasi *Minimum Covariance Determinant* (MCD). Diberikan matriks data $n \times p$, tujuan MCD adalah menemukan h pengamatan di mana $\left\lfloor \frac{n+N+1}{2} \right\rfloor \leq h \leq n$ sehingga diperoleh matriks kovarian yang memiliki determinan terkecil. Estimasi lokasi MCD kemudian ditentukan dari rata-rata titik h, dan MCD dari *scatter* adalah matriks kovariansnya. Estimasi lokasi MCD dapat dihitung menggunakan rumus $\mu = \frac{1}{h}\sum_{i=1}^{h} r_i$ dan taksiran sebar didefinisikan sebagai $\hat{\Sigma}^r = \frac{1}{h}\sum_{i=1}^{p}(x_i - \hat{\mu}^r)(x_i - \hat{\mu}^r)^T$, (*superscript* r singkatan dari *robust*).

Pada tahun 1999, ditemukan metode untuk membangun algoritma yang sangat cepat dalam perhitungan estimator MCD yang disebut sebagai FAST-MCD [11]. Berikut adalah langkah-langkah dari algoritma FAST-MCD:

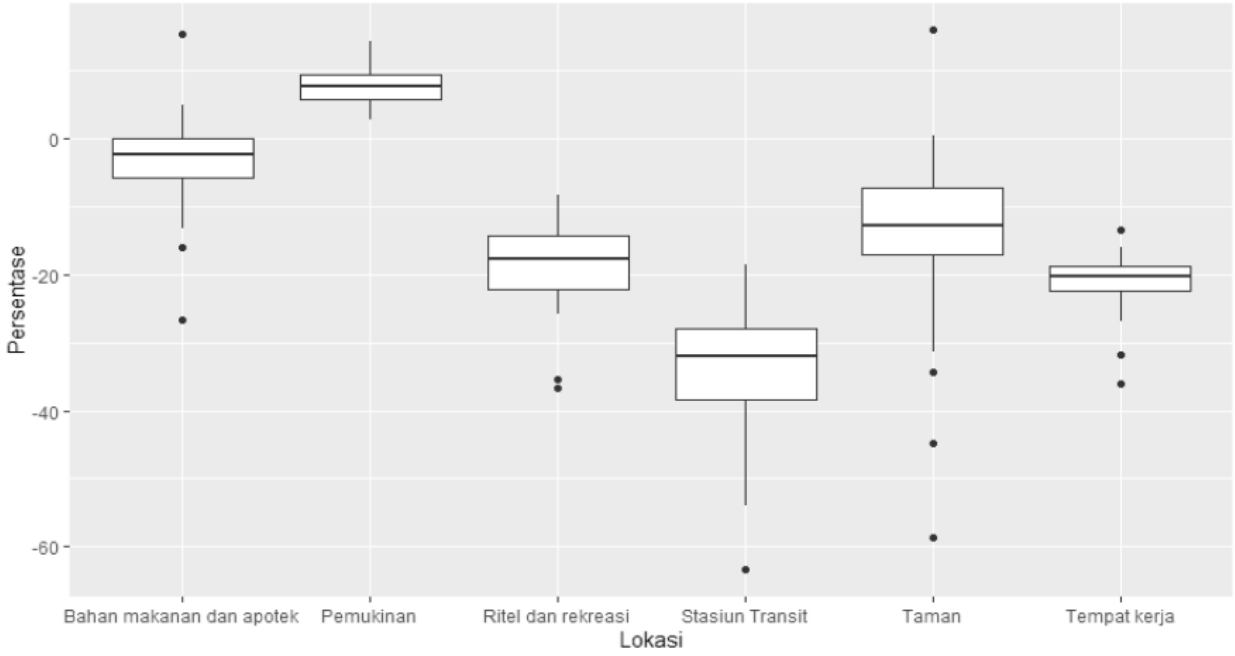
1. Tentukan *h*-subset awal H_1, dimulai dengan $(p + 1)$ subset acak J.
2. Hitung $\mu_0^r = \frac{1}{p+1}\sum_{i \in J} r_i$ dan $\hat{\Sigma}_0^r = \frac{1}{h}\sum_{i=1}^{p}(x_i - \hat{\mu}_0^r)(x_i - \hat{\mu}_0^r)^T$. Jika $\left|\hat{\Sigma}_0^r\right| = 0$, maka pengamatan acak ditambahkan ke J hingga $\left|\hat{\Sigma}_0^r\right| > 0$.
3. Terapkan *C-step* ke *h*-subset H_1 awal dan dapatkan $\left(\mu_1^r, \hat{\Sigma}_1^r\right)$. Jika $\left|\hat{\Sigma}_0^r\right| = 0$ atau $\left|\hat{\Sigma}_0^r\right| = \left|\hat{\Sigma}_1^r\right|$, berhenti; jika tidak, jalankan *C-step* lain untuk menghasilkan $\left(\mu_2^r, \hat{\Sigma}_2^r\right)$, dan seterusnya, hingga konvergensi tercapai.

Hasil dan Pembahasan

Pembahasan dimulai dengan deskripsi data, dilanjutkan dengan menguji kelayakan data dan diakhiri dengan analisis faktor dan sensitivitas.

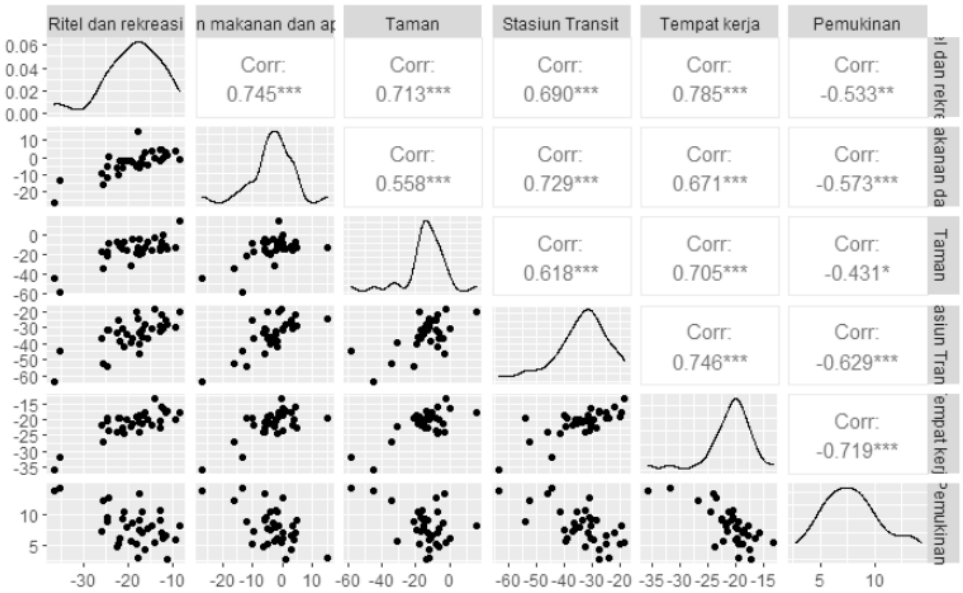
Deskripsi Data

Pada tabel 1, dapat dilihat bahwa pada kelima provinsi, mobilitas di setiap tempat hampir semuanya bernilai negatif. Hal ini dikarenakan pada pelaporan pertama kasus COVID-19 di Indonesia, pemerinta beberapa kali memberlakukan *social distancing* sehingga aktivitas di luar ruangan menjadi terbatas. Kemudian, nampak bahwa mobilitas pada lokasi bahan makanan dan apotek menjadi yang tertinggi di kelima provinsi. Hal ini dikarenakan pada *social distancing*, komponen yang paling penting untuk keberlangsungan hidup adalah makanan dan obat-obatan, sehingga banyak masyarakat yang berbonodong-bondong membeli makanan dan obat-obatan.



Gambar 1. Boxplot mobilitas di tiap lokasi

Pada Gambar 1, nampak bahwa lokasi pemukiman memiliki median yang paling besar dan *range* yang paling kecil. Ini menunjukan bahwa nilai mobilitas pada lokasi pemukiman di tiap provinsi sangat memusat pada satu titik. Lokasi stasiun transit memiliki median paling kecil. Kemudian lokasi taman memiliki *range* yang paling besar, yang mana berarti nilai mobilitas lokasi taman di setiap provinsi sangat variatif. Lokasi tempat kerja bernilai negatif di semua provinsi. Selain itu, lokasi tempat kerja memiliki range yang cukup kecil. Hal ini menandakan bahwa perusahaan di setiap provinsi mulai memberlakukan *work from home*. Selain lokasi pemukiman, semua lokasi memiliki pencilan.



Gambar 2. Korelasi dan persebaran data di tiap lokasi

Pada Gambar 2, dapat dilihat bahwa nilai mutlak dari korelasi tiap pasangan variabel lebih dari 0,4. Kemudian terlihat juga bahwa tiap pasangan variabel memiliki tanda bintang (*), yang mana menunjukan bahwa setiap pasangan variabel memiliki hubungan linear dengan tingkat kepercayaan 95% [12]. Pada *scatter plot*, terlihat beberapa titik terletak jauh dari titik-titik lainnya. Ini mengindikasikan bahwa titik-titik tersebut merupakan pencilan.

Uji Kelayakan Data

Tabel 2. Hasil uji KMO di tiap lokasi

Ritel dan rekreasi	Toko makanan dan apotek	Taman	Transportasi umum	Tempat kerja	Pemukinan	Total
0.846523	0.861239	0.890634	0.90654	0.829209	0.833813	0.86

Analisis faktor dibangun berdasarkan korelasi antar variabel pada data. Dengan demikian, variabel yang digunakan haruslah variabel yang saling berkorelasi. Oleh sebab itu, langkah berikutnya yang harus dilakukan adalah menguji kelayakan data apakah tiap variabel pada data memang memiliki korelasi satu sama lain atau tidak. Pengujian matriks korelasi dapat menggunakan uji *Bartlett Test of Sphericity* dan *Kaiser Meyer Olkin* (KMO).

Tabel 3. Hasil uji Bartlett

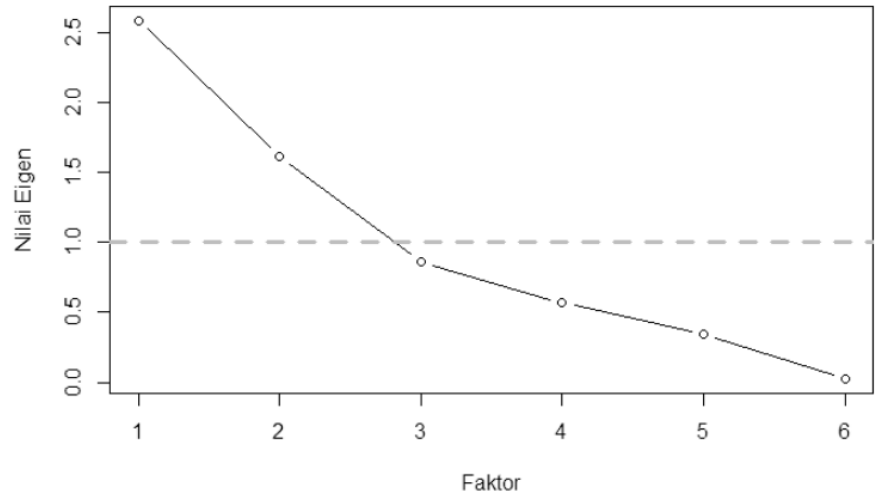
Chi-Square	P-value	df
137.351	7.66E-22	15

Berdasarkan Tabel 2, dapat dilihat bahwa semua variabel lokasi memilki nilai KMO di atas 80. Ini menunjukan bahwa data baik (meritorious) untuk analisis faktor. Kemudian pada tabel 3, diperoleh p-value pada uji Bartlett mendekati 0. Dengan demikian dapat disimpulkan bahwa matriks korelasi bukan merupakan matriks identitas, yang mana berarti dengan kata lain terdapat variabel pada data yang saling

berkorelasi. Berdasarkan kedua uji tersebut, dapat ditarik kesimpulan bahwa data sudah layak untuk dilakukan analisis faktor.

Analisis Faktor

Terdapat banyak cara dalam menentukan jumlah faktor. Di antaranya adalah dengan menggunakan kriteria apriori, nilai Eigen (kriteria akar laten), plot scree visual dan jumlah varians tertentu yang dijelaskan oleh faktor [13]. Pada penelitian ini, penentuan jumlah faktor didasari atas visualisasi scree plot dan besarnya nilai eigen yang lebih dari 1. Berdasarkan Gambar 4, dapat dilihat bahwa scree plot mulai melandai pada faktor 3. Dengan demikian, menurut kriteria pada scree plot, faktor yang dapat dipilih adalah faktor 1, 2, dan 3. Akan tetapi, faktor yang nilai eigennya lebih dari 1 hanya faktor 1 dengan nilai eigen sebesar 2.585 dan faktor 2 dengan nilai eigen sebesar 1.613. Namun karena nilai eigen dari faktor 3 cukup dekat dengan 1, yaitu sebesar 0,86, maka pada penelitian ini faktor 3 tetap dipilih. Ketiga faktor tersebut dapat menjelaskan 84,3% dari variansi data**.**



Gambar 3. Scree plot nilai eigen tiap faktor

Tabel 4. Hasil Analisis Faktor *Robust* menggunakan MCD

	Factor1	Factor2	Factor3
X1	0.915	-0.050	0.163
X2	0.899	-0.077	0.085
X3	0.146	0.028	0.986
X4	0.568	-0.578	-0.071
X5	0.124	-0.876	-0.008
X6	0.008	0.967	0.011
SS loadings	2.005	2.044	1.010
Proportion Var	0.334	0.341	0.168
Cumulative Var	0.334	0.675	0.843

Pada Tabel 4, diperoleh bentul model dari analisis faktor *robust* adalah sebagai berikut:

$$X_1 = 0{,}915F_1 - 0{,}050F_2 + 0{,}163F_3$$

$$X_2 = 0{,}146F_1 - 0{,}028F_2 + 0{,}085F_3$$

$$X_3 = 0{,}146F_1 - 0{,}028F_2 + 0{,}986F_3$$

$$X_4 = 0{,}568F_1 - 0{,}578F_2 + 0{,}071F_3$$

$$X_5 = 0{,}124F_1 - 0{,}876F_2 + 0{,}008F_3$$

$$X_6 = 0{,}008F_1 - 0{,}967F_2 + 0{,}011F_3$$

Elemen di setiap baris matriks *loading factor* yang nilai mutlaknya paling besar dan lebih dari 0,5 ditandai dengan warna merah. Elemen-elemen ini memiliki efek penjelas utama dari variabel yang bersesuaian. Faktor 1 sebagian besar menjelaskan variabel X_1 (ritel dan rekreasi) dan X_2 (toko makanan dan apotek). Dengan demikian, dapat dikatakan bahwa faktor 1 merupakan faktor aktivitas dalam memenuhi kesehatan fisik dan mental. Faktor 2 sebagian besar menjelaskan variabel X_4 (transportasi umum), X_5 (tempat kerja), dan X_6 (pemukiman). Dengan demikian, dapat dikatakan bahwa faktor 2 merupakan faktor aktivitas primer (bekerja dan sosialisasi). Faktor 3 sebagian besar menjelaskan variabel X_3 (taman). Dengan demikian, faktor 3 merupakan faktor aktivitas untuk bersantai.

Analisis Sensitivitas

Pada bagian ini dilakukan simulasi analisis sensitivitas loading factor data mobilitas menggunakan MCD, yang mana kemudian dibandingkan dengan metode klasik. Simulasi ini telah dilakukan pada penelitian sebelumnya dengan menghitung *Mean Square Error* (MSE) antara nilai *loading factor* data asli dengan data yang telah ditambahkan pencilan [6]. Misalkan data mobilitas yang telah diskalakan dengan metode *min-max scaler* adalah $X^{(0)}$, matirks *noise* $err^{(s)}$, dan matirks $xout^{(s)}$ yang mana memuat sebanyak n_{out} pencilan, maka data hasil pembangkitan yang mengandung pencilan adalah:

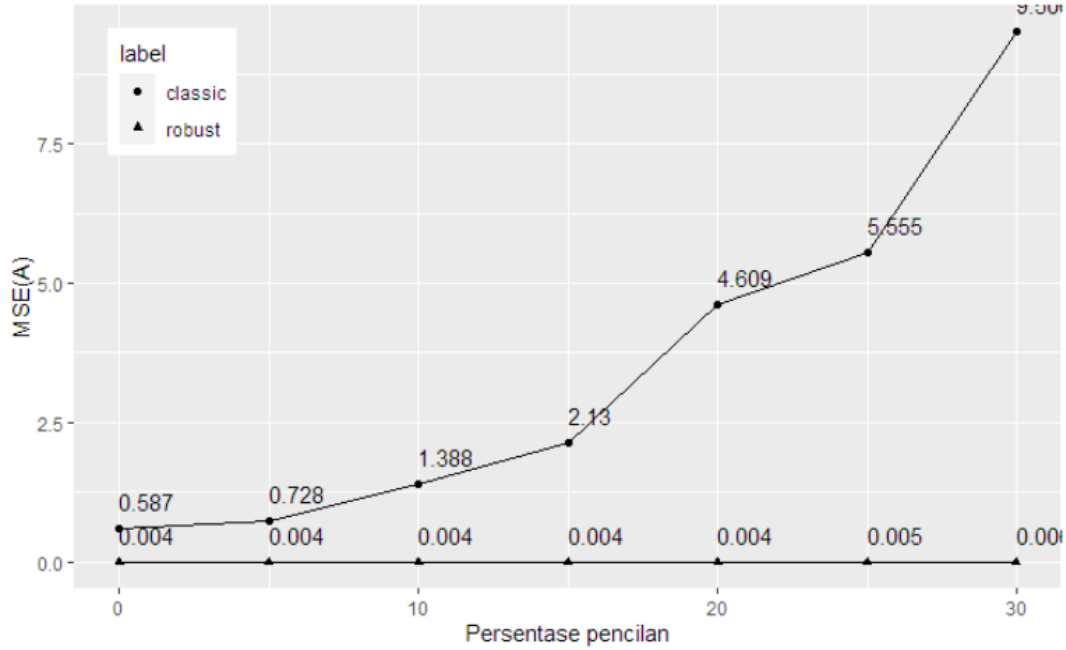
$$X^{(s)} = X^{(0)} + err + xout;\ \ s = 1, 2, 3, \dots, m$$

Dengan melakukan analisis faktor pada data $X^{(0)}$ dan $X^{(s)}$, maka akan didapatkan matriks *loading factor* $L^{(0)}, L^{(s)} \in \mathbb{R}^{m\times k}$, di mana k adalah banyaknya loading factor dan n adalah banyaknya data. Misalkan $A^{(0)} = L^{(0)}\left(L^{(0)}\right)^T$ dan $A^{(s)} = L^{(s)}\left(L^{(s)}\right)^T$, maka didapatkan formula MSE untuk menghitung sensitivitas *loading factor* terhadap data pencilan adalah:

$$MSE(A) = \frac{1}{m}\sum_{s=1}^{m}\left(a_{ij}^{(s)} - a_{ij}^{(0)}\right)$$

Pada analisis sensitivitas di penilitian ini, dipilih $k = 3$, $m = 50$, $err \sim N(0; 0{,}5^2)$, dan $xout \sim N(10; 0{,}5^2)$.

Pada Gambar 4, ditunjukan nilai dari $MSE(\boldsymbol{A})$ untuk analisis faktor *robust* dan klasik berdasarkan persentase pencilannya. dapat dilihat bahwa nilai MSE untuk analisis faktor klasik naik secara eksponensial seiring bertambahnya persentase pencilan. Berbeda dengan analisis faktor *robust* dengan MCD yang nilai MSE-nya cenderung konstan di sekitar 0,004 sampai 0,006. Dengan demikian, loading factor yang dibentuk oleh analisis faktor *robust* menggunakan MCD lebih stabil terhadap pencilan dibandingkan dengan analisis faktor klasik.



Gambar 4. Plot analisis sensitivitas loading factor dengan MSE

Kesimpulan

Pada penelitian ini, untuk data mobilitas, analisis faktor *robust* menggunakan MCD memiliki kemampuan yang lebih baik dalam membentuk *loading factor* yang stabil terhadap data pencilan dibandingkan dengan analisis faktor klasik. Faktor optimal yang terbentuk adalah sebanyak 3 faktor, yaitu faktor aktivitas dalam memenuhi kesehatan fisik dan mental, faktor aktivitas primer (bekerja dan sosialisasi), dan faktor aktivitas untuk bersantai. Ketiga faktor tersebut mampu menjelaskan 84,3% dari variansi data.

Data mobilitas pada penelitian yang merupakan data runtun waktu disederhanakan dengan menghitung rata-ratanya. Untuk penelitian berikutnya disarankan untuk membuat analisis faktor robust yang dinamis terhadap pergerakan waktu.

Referensi

[1] H. C *et al.*, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet (London, England)*, vol. 395, no. 10223, pp. 497–506, Feb. 2020, doi: 10.1016/S0140-6736(20)30183-5.

[2] A. Susilo *et al.*, "Coronavirus Disease 2019: Tinjauan Literatur Terkini," *J. Penyakit Dalam Indones.*, vol. 7, no. 1, p. 45, 2020, doi: 10.7454/jpdi.v7i1.415.

[3] D. Rakhmawati, R. Wahyudi, and C. G. Yuliawan, "Pemodelan Harga Saham Ihsg Selama Pandemi Covid-19 Menggunakan Arima Non Musiman," *J. Pro Bisnis*, vol. 13, no. 2, pp. 39–48, 2020.

[4] W. K. Härdle and L. Simar, *Applied Multivariate Statistical Analysis*. Pearson Prentice Hall, 2013.

[5] Suryanto, *Metode Statistika Multivariat*. Jakarta: Departemen Pendidikan dan Kebudayaan, 1988.

[6] G. Pison, P. J. Rousseeuw, P. Filzmoser, and C. Croux, "Robust Factor Analysis," *J. Multivar. Anal.*, vol. 84, no. 1, pp. 145–172, 2003, doi: 10.1016/S0047-259X(02)00007-6.

[7] E. Puspitasari and M. Abdul Mukid, "PERBANDINGAN ANALISIS FAKTOR KLASIK DAN ANALISIS FAKTOR ROBUST UNTUK DATA INFLASI KELOMPOK BAHAN MAKANAN DI JAWA TENGAH," *J. Gaussian*, vol. 3, no. 3, pp. 343–352, 2014, doi: 10.14710/J.GAUSS.V3I3.6445.

[8] Y.-Y. Zhang, "An Object-Oriented Solution for Robust Factor Analysis." 2013.

[9] A. Alexander Beaujean and A. Alexander, "Factor Analysis using R," *Pract. Assessment, Res. Eval.*, vol. 18, 2013, doi: 10.7275/z8wr-4j42.

[10] R. Kosfeld, "Robust Exploratory Factor Analysis," *Stat. Pap. 1996 372*, vol. 37, no. 2, pp. 105–122, 1996, doi: 10.1007/BF02926576.

[11] P. J. Rousseeuw and K. Van Driessen, "A Fast Algorithm for the Minimum Covariance Determinant Estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999, doi: 10.1080/00401706.1999.10485670.
[12] B. Schloerke *et al.*, "Package 'GGally.'" 2021.
[13] W. B. Joseph F. Hair, Ronald L. Tatham, Rolph E. Anderson, *Multivariate Data Analysis*, 5th ed. New Jersey: Prentice Hall, 1998.