

## PENGEMBANGAN REPRESENTASI DATA RUMAH SAKIT DI INDONESIA MENGGUNAKAN *EXPLORATORY FACTOR ANALYSIS*

Aisha Gemala Jondya<sup>(1)</sup>, Reny Fitri Yani<sup>2</sup>, Akbar Shiddiq<sup>3</sup>, Linna Oktaviana Sari<sup>4</sup>, Aditya Nugraha Pratama Saiya<sup>5</sup>

<sup>1</sup> Universitas Riau, Pekanbaru

<sup>2</sup> Universitas Riau, Pekanbaru

<sup>3</sup> Universitas Riau, Pekanbaru

<sup>4</sup> Universitas Riau, Pekanbaru

<sup>5</sup> Universitas Riau, Pekanbaru

### **Abstract**

*This study aims to develop a feature engineering approach based on Exploratory Factor Analysis (EFA) to construct a more compact and informative representation of hospital data in Indonesia. Data were collected from the Indonesian Hospital Information System (SIRS) and grouped into five domains: human resources, healthcare services, bed capacity, supporting facilities, and operational support. The suitability of each domain for factor analysis was evaluated using the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's Test of Sphericity. The results indicate that all domains satisfied the requirements for EFA, with KMO values ranging from 0.555 to 0.981 and significant Bartlett's Test results ( $p < 0.001$ ). A total of 488 variables were successfully reduced into 35 latent factors, demonstrating the effectiveness of EFA in reducing data complexity while preserving essential information. The resulting factors provide a more structured representation of hospital characteristics and can serve as a foundation for future hospital segmentation and healthcare data analytics.*

**Kata Kunci:** Eksplorasi, Dataset, Rumah Sakit, *Factor Analysis*, *Machine Learning*

### *Informasi Artikel:*

Dikirim : 04 Juni 2026

Ditelaah: 12 Juni 2026

Diterima: 19 Juni 2026

Publikasi: 29 Juni 2026

Januari – Juni 2026, Vol 7 (1) : hlm 31-42

©2026 Institut Teknologi dan Bisnis Ahmad Dahlan.

All rights reserved.

(\*) Korespondensi: [aisha.gemala@lecturer.unri.ac.id](mailto:aisha.gemala@lecturer.unri.ac.id) (Aisha Gemala Jondya)

## PENDAHULUAN

Rumah sakit merupakan komponen penting dalam sistem pelayanan kesehatan yang berperan dalam menyediakan layanan medis bagi masyarakat. Pemanfaatan teknologi untuk administrasi rumah sakit dan pelayanan telah banyak membantu meningkatkan kualitas layanan Rumah Sakit di seluruh dunia (Jondya, Rahma, & Sophia, 2024). Peraturan Menteri Kesehatan Republik Indonesia Nomor 3 Tahun 2020 tentang Klasifikasi dan Perizinan Rumah Sakit menetapkan bahwa klasifikasi rumah sakit didasarkan pada aspek pelayanan, sumber daya manusia, serta sarana dan prasarana (Kementerian Kesehatan Republik Indonesia, 2020). Meskipun telah mencakup berbagai aspek, pendekatan tersebut belum sepenuhnya didukung oleh analisis berbasis data yang mampu menggambarkan keseimbangan antara kapasitas layanan dan ketersediaan tenaga medis secara komprehensif (Dong, et al., 2023).

Dalam praktiknya, rumah sakit dengan tipe kelas yang sama dapat menunjukkan variasi dalam hal kapasitas, sarana, maupun sumber daya manusia (Jiarui Han, 2025). Variasi ini mengindikasikan bahwa klasifikasi administratif belum sepenuhnya merepresentasikan kondisi operasional rumah sakit secara rinci (Hammad, 2022), khususnya terkait keseimbangan antara kapasitas layanan dan sumber daya manusia (Mirna Raafiana, 2025). Namun demikian, pola distribusi serta hubungan antara kedua aspek tersebut belum banyak dianalisis secara sistematis menggunakan pendekatan berbasis data dan *machine learning*, sehingga pemahaman mengenai karakteristik rumah sakit masih cenderung terbatas pada aspek *administrative* (Marengo, Santamato, & Iacoviello, 2025).

Perkembangan digitalisasi sektor kesehatan telah menghasilkan ketersediaan data dalam jumlah besar melalui berbagai sistem informasi kesehatan. Salah satu sumber data rumah sakit di Indonesia yang dapat dimanfaatkan adalah Sistem Informasi Rumah Sakit (SIRS) (Direktorat Jenderal Kesehatan Lanjutan, 2025) yang dikelola oleh Kementerian Kesehatan Republik Indonesia. Data tersebut mencakup berbagai aspek rumah sakit, mulai dari sumber daya manusia, kapasitas tempat tidur, layanan kesehatan, fasilitas penunjang, hingga infrastruktur kesehatan. Ketersediaan data yang sangat kaya ini membuka peluang untuk melakukan analisis berbasis data dan *machine learning* guna memperoleh pemahaman yang lebih mendalam mengenai karakteristik rumah sakit di Indonesia.

Meskipun demikian, pemanfaatan data rumah sakit masih menghadapi tantangan berupa tingginya dimensi data dan banyaknya variabel yang saling berkaitan (Ray, Reddy, & Banerjee, 2021). Penggunaan seluruh variabel secara langsung dalam proses analisis berpotensi menghasilkan redundansi informasi (G. Manikandan, 2021), meningkatkan kompleksitas komputasi, serta menyulitkan interpretasi hasil (Cavina & et al., 2026). Oleh karena itu, diperlukan suatu pendekatan yang mampu membangun representasi data yang lebih ringkas namun tetap mempertahankan informasi penting yang terkandung di dalamnya (Shu, Huang, & et al., 2021).

Salah satu pendekatan yang dapat digunakan adalah *feature engineering* (Verdonck & et al., 2024), yaitu proses membangun, memilih, dan mentransformasikan fitur agar lebih representatif terhadap karakteristik objek yang dianalisis. Dalam penelitian ini, *feature engineering* dikombinasikan dengan *Exploratory Factor Analysis* (EFA) untuk mengidentifikasi faktor-faktor laten yang mendasari hubungan antar variabel rumah sakit. Melalui pendekatan tersebut, sejumlah besar variabel dapat direduksi (Howard, 2023) menjadi

beberapa faktor yang lebih mudah diinterpretasikan dan berpotensi merepresentasikan dimensi utama karakteristik rumah sakit di Indonesia.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan mengembangkan feature engineering berbasis *factor analysis* untuk membangun representasi data rumah sakit yang lebih informatif dan terstruktur. Hasil penelitian diharapkan menghasilkan faktor-faktor laten yang mampu menggambarkan karakteristik rumah sakit secara komprehensif serta dapat digunakan sebagai dasar dalam penelitian segmentasi rumah sakit berbasis machine learning pada tahap selanjutnya.

## **METODE PENELITIAN**

Penelitian ini menggunakan pendekatan kuantitatif berbasis data science untuk mengembangkan representasi data rumah sakit melalui proses feature engineering dan *Exploratory Factor Analysis* (EFA). Pendekatan ini dipilih karena data rumah sakit yang diperoleh dari Sistem Informasi Rumah Sakit (SIRS) memiliki karakteristik berdimensi tinggi dengan jumlah variabel yang besar dan saling berkaitan, sehingga berpotensi menimbulkan redundansi informasi serta menyulitkan proses analisis dan interpretasi. Oleh karena itu, diperlukan suatu metode yang mampu menyederhanakan struktur data tanpa menghilangkan informasi penting yang merepresentasikan karakteristik rumah sakit.

Dalam penelitian ini, feature engineering digunakan sebagai proses untuk membangun representasi data yang lebih informatif melalui transformasi dan reduksi variabel. Proses tersebut dikombinasikan dengan *Exploratory Factor Analysis* (EFA) untuk mengidentifikasi faktor-faktor laten yang mendasari hubungan antarvariabel (Finch, 2023). Melalui pendekatan ini, sejumlah besar variabel yang menggambarkan sumber daya manusia, kapasitas layanan, fasilitas kesehatan, dan infrastruktur rumah sakit dapat direduksi menjadi sejumlah faktor yang lebih ringkas, mudah diinterpretasikan, dan mampu merepresentasikan dimensi utama karakteristik rumah sakit di Indonesia.

Tahapan penelitian meliputi akuisisi data, *data understanding*, *preprocessing*, *feature engineering*, ekstraksi faktor laten menggunakan EFA, evaluasi hasil analisis faktor, serta pembentukan representasi data baru (dataset yang lebih ringkas dan optimal) berdasarkan *factor score* yang dihasilkan. Representasi data tersebut diharapkan dapat menjadi dasar yang lebih efektif untuk analisis lanjutan, khususnya dalam pengembangan model segmentasi rumah sakit berbasis machine learning pada penelitian berikutnya.

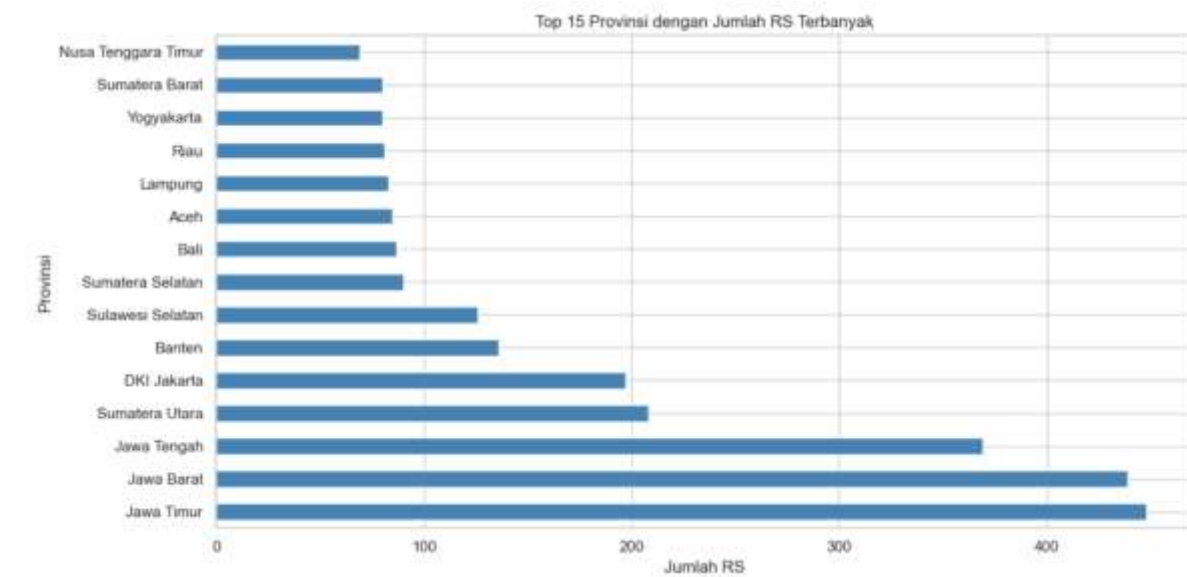


Hasil data yang didapatkan setelah proses scrapping yang telah dilakukan berjumlah 3.304 baris Rumah Sakit dengan jumlah kolom 802. Kondisi awal dataset ditemukan banyak data yang *missing* dan tidak lengkap sehingga perlu dilakukannya pra-pemrosesan data. Jumlah kolom yang sangat banyak juga mengidentifikasi perlu adanya efisiensi jumlah fitur sehingga dataset dapat lebih optimal. Hal ini memperlihatkan bahwa penelitian dengan metode *Exploratory Data Analysis* ini sangat perlu dilakukan untuk dapat melihat karakteristik sebenarnya dari data Rumah Sakit di seluruh Indonesia.

### **Data Understanding**

Tahap ini bertujuan memahami karakteristik dataset sebelum dilakukan analisis lebih lanjut. Aktivitas yang dilakukan meliputi identifikasi struktur data, tipe variabel, distribusi data, serta pemeriksaan kualitas data. Selain itu, dilakukan analisis statistik deskriptif dan *Exploratory Data Analysis* (EDA) untuk mengidentifikasi pola awal, mendeteksi *Outlier*, dan mengevaluasi hubungan antarvariabel.

Berdasarkan hasil *Exploratory Data Analysis* (EDA) terlihat bahwa tipe Rumah Sakit di Indonesia didominasi Rumah Sakit Umum dengan jumlah 2.792 RS = 84.5%, sementara Rumah Sakit Ibu dan Anak menempati posisi ke dua terbanyak dengan jumlah 314 RS = 9,5%. Untuk sebaran lokasi terlihat tidak seimbang dimana Pulau Jawa memiliki lebih banyak jumlah Rumah Sakit dibandingkan pulau lainnya. Status kepemilikan Rumah Sakit paling banyak dimiliki oleh swasta sebanyak 934 Rumah Sakit dengan status BLU/BLUD mendominasi yaitu sebanyak 2.410 Rumah Sakit.



Gambar 3. Distribusi Rumah Sakit Indonesia

Berdasarkan hasil EDA dilakukan pula analisis statistic terhadap 3 fitur numerical yaitu *total\_tt*, *total\_layanan*, dan *total\_sdm*. Berdasarkan hasil analisis histogram dan box plot ditemukan bahwa ketiga fitur menunjukkan distribusi right skewed (condong kanan) dan banyak *Outlier* di sisi atas. Hal ini menunjukkan mayoritas Rumah Sakit mempunyai kapasitas kecil-menengah meskipun ada segelintir Rumah Sakit dengan kapasitas sangat besar yang jadi

*outlier*. *Outlier* tersebut bukan merupakan error karena memang ada Rumah Sakit yang sangat besar seperti RSUP, Rumah Sakit tipe A, dan Rumah Sakit vertikal Kemenkes.

### **Data Preprocessing**

Data *preprocessing* dilakukan untuk meningkatkan kualitas data sebelum proses analisis. Tahapan ini meliputi penanganan missing values, penghapusan duplikasi data, identifikasi dan penanganan *Outlier*. Pada tahap ini juga dilakukan normalisasi data numerik. Selain itu, dilakukan seleksi awal terhadap variabel yang memiliki variasi sangat rendah atau mengandung informasi yang tidak relevan sehingga dapat mengurangi kompleksitas analisis.

Pada tahap awal dilakukan data *cleaning*. Dari hasil data *cleaning* didapatkan 3.304 data rumah sakit dimana jumlah fitur atau kolom mengalami pengurangan sebanyak 10 kolom menjadi 792 fitur. Fitur tersebut terdiri dari 10 fitur identitas / kategorikal dan 782 fitur numerik. Selanjutnya tahapan *preprocessing* dilanjutkan dengan melakukan *scaling* atau normalisasi dan analisis data lebih lanjut. Dataset akhir yang dihasilkan dari proses ini adalah total rumah sakit sejumlah 3.292 dengan fitur sejumlah 785.

### **Exploratory Factor Analysis (EFA)**

EFA digunakan untuk mengidentifikasi faktor-faktor laten yang mendasari hubungan antarvariabel dalam dataset. Sebelum proses ekstraksi faktor dilakukan, terlebih dahulu dilakukan pengujian kelayakan data menggunakan Kaiser-Meyer-Olkin (KMO) *Measure of Sampling Adequacy* dan *Bartlett's Test of Sphericity* (Thao, Tan, & Mai, 2022). Selanjutnya dilakukan ekstraksi faktor menggunakan metode yang sesuai, diikuti proses rotasi faktor untuk memperoleh struktur faktor yang lebih mudah diinterpretasikan. Faktor yang terbentuk diharapkan mampu merepresentasikan dimensi utama karakteristik rumah sakit di Indonesia.

Pada penelitian ini, seluruh fitur yang didapatkan dari website SIRS dan kemudian melalui tahap *preprocessing* menghasilkan sejumlah 785 fitur. Sebagian fitur tidak diikuti dalam proses EFA karena tidak signifikan mempengaruhi segmentasi seperti identitas Rumah sakit yaitu, alamat, nomor telepon, nama direktur, dan sebagainya. Sebelum dilakukan EFA, fitur-fitur tersebut dikelompokkan berdasarkan domainnya. Pengelompokan domain dilakukan untuk mempermudah interpretasi hasil sekaligus menjaga kesesuaian konteks antarvariabel yang dianalisis. Domain fitur dataset pada penelitian ini ditampilkan pada tabel 1 berikut.

Tabel 1. Domain Fitur Dataset Rumah Sakit Indonesia

No.	Domain	Jumlah Fitur
1.	SDM	122
2.	Layanan	316
3.	Kapasitas	32
4.	Fasilitas	14
5.	Penunjang	4

Proses ekstraksi faktor dilakukan menggunakan metode *Principal Axis Factoring* (PAF), karena metode ini dirancang untuk mengidentifikasi faktor laten yang menjelaskan varians bersama (*common variance*) antarvariabel. Setelah faktor diekstraksi, dilakukan rotasi faktor menggunakan metode *Varimax Orthogonal Rotation* untuk menghasilkan struktur faktor

yang lebih sederhana dan mudah diinterpretasikan. Rotasi Varimax dipilih karena mampu memaksimalkan perbedaan loading antar faktor sehingga setiap variabel cenderung memiliki keterkaitan yang lebih kuat pada satu faktor tertentu.

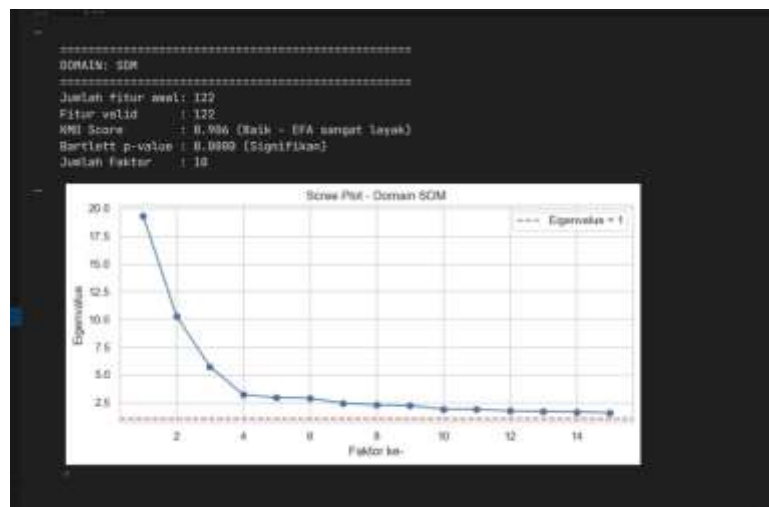
Penentuan jumlah faktor dilakukan menggunakan *Kaiser Criterion* dengan mempertahankan faktor yang memiliki nilai eigenvalue lebih besar dari 1. Selain itu, scree plot digunakan sebagai pendukung untuk mengidentifikasi titik *elbow* yang menunjukkan jumlah faktor yang paling representatif dalam menjelaskan struktur data. Kombinasi kedua pendekatan tersebut digunakan untuk memperoleh jumlah faktor yang optimal pada masing-masing domain analisis.

## HASIL DAN PEMBAHASAN

Sebelum proses ekstraksi faktor dilakukan, setiap domain dievaluasi menggunakan Kaiser-Meyer-Olkin (KMO) *Measure of Sampling Adequacy* dan *Bartlett's Test of Sphericity* untuk memastikan kelayakan penerapan EFA. Nilai KMO digunakan untuk mengukur kecukupan sampel dan tingkat keterkaitan antarvariabel, sedangkan Bartlett's Test digunakan untuk menguji apakah matriks korelasi yang terbentuk memenuhi syarat untuk dilakukan analisis faktor. Selanjutnya, jumlah faktor laten ditentukan berdasarkan kriteria eigenvalue lebih besar dari satu (Kaiser Criterion) yang didukung dengan analisis scree plot.

### 1. Hasil EFA Domain SDM

Domain SDM terdiri atas 122 variabel yang menggambarkan berbagai jenis tenaga kesehatan dan tenaga pendukung yang tersedia pada rumah sakit. Hasil pengujian menunjukkan nilai KMO sebesar 0,906 yang termasuk kategori sangat baik, sedangkan Bartlett's Test menghasilkan nilai signifikansi  $p < 0,001$ . Hasil tersebut menunjukkan bahwa korelasi antarvariabel cukup kuat dan memenuhi syarat untuk dilakukan analisis faktor. Berdasarkan kriteria eigenvalue lebih besar dari satu, diperoleh 10 faktor laten yang mampu merepresentasikan struktur data SDM rumah sakit.

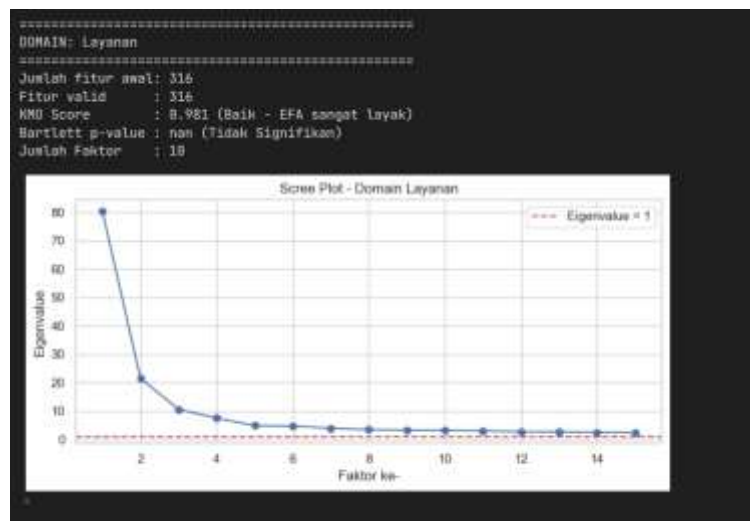


Gambar 4. Hasil EFA Domain SDM

Temuan ini menunjukkan bahwa meskipun jumlah variabel SDM sangat besar, sebagian besar informasi yang terkandung di dalamnya dapat diringkas ke dalam sejumlah dimensi utama yang lebih sederhana. Dengan demikian, EFA berhasil mengurangi kompleksitas data SDM sekaligus mempertahankan informasi penting yang diperlukan untuk menggambarkan karakteristik sumber daya manusia rumah sakit.

## 2. Hasil EFA Domain Layanan

Domain layanan kesehatan merupakan domain dengan jumlah variabel terbesar, yaitu sebanyak 316 variabel yang mencakup berbagai jenis layanan medis, spesialis, subspecialis, dan layanan penunjang kesehatan lainnya. Hasil analisis menghasilkan nilai KMO sebesar 0,981 yang menunjukkan tingkat kecukupan sampel yang sangat tinggi. Nilai tersebut mengindikasikan bahwa variabel-variabel layanan memiliki pola korelasi yang kuat sehingga sangat sesuai untuk dilakukan analisis faktor. Berdasarkan kriteria eigenvalue lebih besar dari satu diperoleh 10 faktor laten yang mampu merepresentasikan keseluruhan domain layanan.

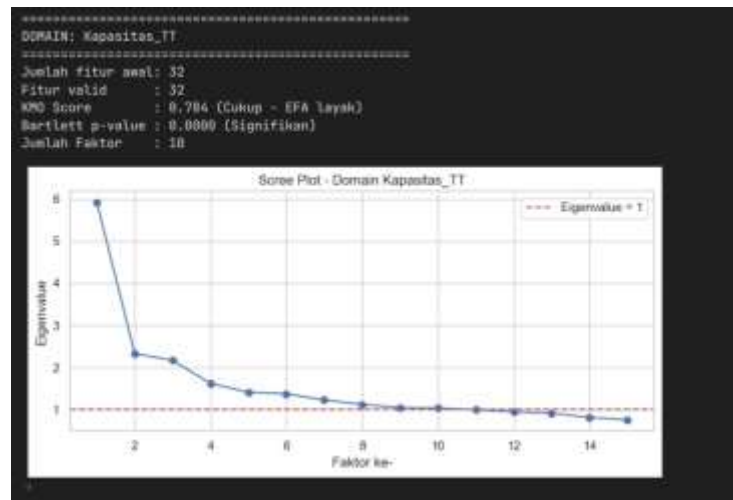


Gambar 5. Hasil EFA Domain Layanan

Hasil ini menunjukkan bahwa ratusan variabel layanan sebenarnya memiliki struktur laten yang relatif terorganisasi dan dapat diringkas ke dalam sejumlah dimensi utama. Temuan tersebut memperlihatkan efektivitas EFA dalam mengatasi tingginya dimensi data layanan rumah sakit sekaligus menghasilkan representasi data yang lebih ringkas dan mudah digunakan untuk analisis lanjutan.

## 3. Hasil EFA Domain Kapasitas

Domain kapasitas terdiri atas 32 variabel yang merepresentasikan berbagai jenis tempat tidur dan kapasitas pelayanan rawat inap yang tersedia di rumah sakit. Hasil analisis menunjukkan nilai KMO sebesar 0,784 yang mengindikasikan tingkat kecukupan sampel yang baik untuk penerapan EFA.

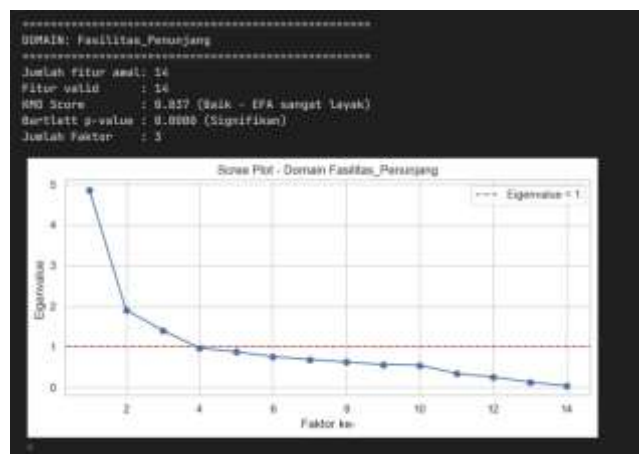


Gambar 6. Hasil EFA Domain Kapasitas

Selain itu, Bartlett's Test menunjukkan hasil yang signifikan ( $p < 0,001$ ), sehingga hubungan antarvariabel dinilai memadai untuk membentuk faktor laten. Berdasarkan hasil ekstraksi diperoleh 10 faktor yang memiliki eigenvalue lebih besar dari satu. Hasil ini menunjukkan bahwa variasi kapasitas tempat tidur rumah sakit tidak hanya dipengaruhi oleh satu atau dua aspek utama, tetapi terdiri atas beberapa dimensi yang mencerminkan keragaman layanan rawat inap yang tersedia. Meskipun demikian, jumlah faktor yang dihasilkan masih jauh lebih sedikit dibandingkan jumlah variabel awal, sehingga EFA tetap memberikan manfaat dalam mereduksi dimensi data dan meningkatkan kemudahan interpretasi.

#### 4. Hasil EFA Domain Fasilitas

Domain fasilitas penunjang terdiri atas 14 variabel yang menggambarkan keberadaan berbagai fasilitas pendukung pelayanan kesehatan di rumah sakit. Hasil pengujian menunjukkan nilai KMO sebesar 0,837 yang termasuk kategori baik, sementara Bartlett's Test menunjukkan hasil signifikan ( $p < 0,001$ ). Hasil tersebut menunjukkan bahwa hubungan antarvariabel cukup kuat untuk membentuk struktur faktor yang stabil.

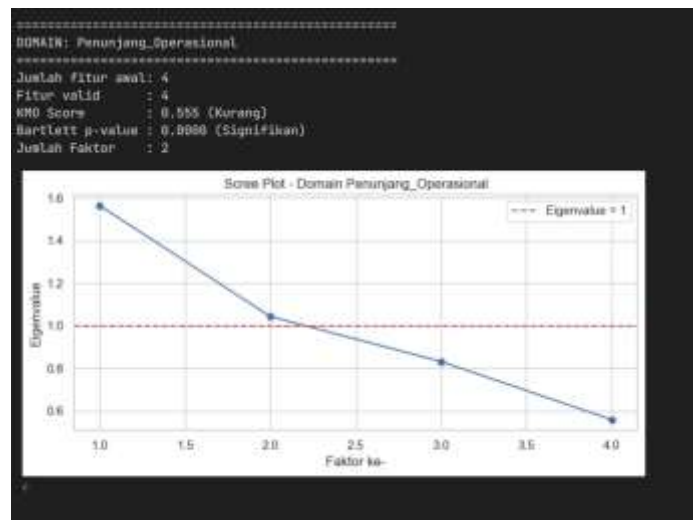


Gambar 7. Hasil EFA Domain Fasilitas

Berdasarkan proses ekstraksi diperoleh tiga faktor laten yang memiliki eigenvalue lebih besar dari satu. Jumlah faktor yang relatif sedikit dibandingkan jumlah variabel awal menunjukkan bahwa karakteristik fasilitas penunjang memiliki tingkat kemiripan yang cukup tinggi sehingga dapat direpresentasikan oleh beberapa dimensi utama saja.

#### 5. Hasil EFA Domain Penunjang

Domain penunjang operasional terdiri atas empat variabel yang menggambarkan fasilitas dan sumber daya pendukung operasional rumah sakit. Hasil analisis menunjukkan nilai KMO sebesar 0,555 yang masih berada di atas batas minimum kelayakan EFA, sedangkan Bartlett's Test menunjukkan hasil yang signifikan ( $p < 0,001$ ). Hasil ini menunjukkan bahwa meskipun tingkat korelasi antarvariabel tidak sekuat domain lainnya, struktur data masih memungkinkan untuk dianalisis menggunakan pendekatan faktor. Berdasarkan kriteria eigenvalue lebih besar dari satu diperoleh dua faktor laten yang mampu merepresentasikan domain penunjang operasional.



Gambar 8. Hasil EFA Domain Penunjang

Jumlah faktor yang terbentuk menunjukkan bahwa variasi karakteristik operasional rumah sakit dapat diringkas menjadi beberapa dimensi utama yang lebih sederhana. Meskipun kualitas struktur faktor pada domain ini relatif lebih rendah dibandingkan domain lainnya, hasil yang diperoleh tetap memberikan informasi yang berguna dalam proses pembentukan representasi data rumah sakit secara keseluruhan.

## KESIMPULAN

Penelitian ini berhasil menerapkan pendekatan feature engineering berbasis Exploratory Factor Analysis (EFA) untuk membangun representasi data rumah sakit yang lebih ringkas dan informatif. Berdasarkan data rumah sakit yang diperoleh dari Sistem Informasi Rumah Sakit (SIRS), variabel-variabel yang memiliki dimensi tinggi dikelompokkan ke dalam lima domain utama, yaitu sumber daya manusia (SDM), layanan kesehatan, kapasitas tempat

tidur, fasilitas penunjang, dan penunjang operasional. Hasil pengujian menunjukkan bahwa seluruh domain memenuhi syarat untuk dilakukan analisis faktor dengan nilai KMO berkisar antara 0,555 hingga 0,981 serta Bartlett's Test yang signifikan. Temuan ini mengindikasikan bahwa data rumah sakit memiliki struktur korelasi yang memadai untuk dieksplorasi menggunakan EFA.

Melalui proses ekstraksi faktor, sebanyak 488 variabel berhasil direduksi menjadi 35 faktor laten yang mampu merepresentasikan karakteristik utama rumah sakit di Indonesia. Hasil tersebut menunjukkan bahwa EFA efektif digunakan untuk mengurangi kompleksitas data sekaligus mempertahankan informasi penting yang terkandung di dalamnya. Representasi data yang dihasilkan tidak hanya mempermudah interpretasi karakteristik rumah sakit, tetapi juga menyediakan dasar yang lebih terstruktur untuk pengembangan analisis lanjutan, seperti segmentasi rumah sakit, pengelompokan layanan kesehatan, maupun penerapan metode machine learning lainnya. Dengan demikian, penelitian ini memberikan kontribusi dalam pengembangan pendekatan data-driven untuk analisis data rumah sakit serta mendukung pemanfaatan data kesehatan yang lebih optimal dalam proses pengambilan keputusan. Pada penelitian selanjutnya, faktor-faktor laten yang dihasilkan dapat dimanfaatkan sebagai fitur masukan dalam proses segmentasi rumah sakit menggunakan metode clustering untuk mengidentifikasi kelompok rumah sakit yang memiliki karakteristik serupa berdasarkan kapasitas layanan dan sumber daya yang dimiliki.

## DAFTAR PUSTAKA

- Abodayeh, A. H. (2023). **Web scraping for data analytics: A beautifulsoup implementation**. *sixth international conference of women in data science at prince Sultan University*. IEEE.
- Cavina, P., & et al. (2026). Beyond tabular data: **Feature extraction and selection for high-dimensional data**. *Array*.
- Direktorat Jenderal Kesehatan Lanjutan. (2025, December). *Dashboard RS Online*. Retrieved from Sistem Informasi Rumah Sakit: [https://sirs.kemkes.go.id/fo/home/dashboard\\_rs?id=0](https://sirs.kemkes.go.id/fo/home/dashboard_rs?id=0)
- Dong, E., Sun 4, X., Xu , T., Zhang, S., Wang, T., & et al. (2023). **Measuring the inequalities in healthcare resource in facility and workforce: A longitudinal study in China**. *National Library of Medicine*.
- Finch, W. (2023). **A Comparison of Methods for Determining the Number of Factors to Retain with Exploratory Factor Analysis of Dichotomous Data**. *Psych*.
- G. Manikandan, S. A. (2021). **An efficient feature selection framework based on information theory for high dimensional data**. *Applied Soft Computing*.
- Hammad, A. R. (2022). **PERKEMBANGAN RUMAH SAKIT DI INDONESIA DI ERA DISRUPTIF**. *Jurnal Keperawatan Merdeka*, 2(1), 1-6.
- Howard, M. C. (2023). **A systematic literature review of exploratory factor analyses in management**. *Journal of Business Research*.
- Jiarui Han, J. H. (2025). **Spatial accessibility evaluation and clustering analysis of primary healthcare within 15 min circles in a northern Chinese megacity**. *Scientific Report*.

- Jondya, A. G., Rahma, A. M., & Sophia, T. (2024). **Enhancing Public Health Records: Development of Death Certificate Management System**. *4th International Conference on Electrical Engineering and Informatics (ICon EEI)*. Pekanbaru.
- Kementerian Kesehatan Republik Indonesia. (2020). **Peraturan Menteri Kesehatan Republik Indonesia Nomor 3 Tahun 2020**. Retrieved April 16, 2026, from Peraturan Badan Pemeriksa Keuangan: <https://peraturan.bpk.go.id/Details/152506/permenkes-no-3-tahun-2020>
- Marengo, A., Santamato, V., & Iacoviello, M. (2025). **Machine Learning in Biomedical Informatics: Optimizing Resource Allocation and Energy Efficiency in Public Hospitals**. *IEEE Access*, 13.
- Mirna Raafiana, H. A. (2025). **KESIAPAN RUMAH SAKIT DALAM IMPLEMENTASI KEBIJAKAN KELAS RAWAT INAP STANDAR (KRIS) JKN : LITERATURE REVIEW**. *Syntax Literate: Jurnal Ilmiah Indonesia*, 10(1).
- Mumuni, A., & Mumuni, F. (2025). **Automated data processing and feature engineering for deep learning and big data applications: A survey**. *Journal of Information and Intelligence*, 113-153.
- Ray, P., Reddy, S. S., & Banerjee, T. (2021). **Various dimension reduction techniques for high dimensional data analysis: a review**. *Artificial Intelligence Review*, 3473–3515.
- Shu, L., Huang, K., & et al. (2021). **Feature selection using autoencoders with Bayesian methods to high-dimensional data**. *Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology*.
- Thao, N., Tan, N., & Mai, T. (2022). **KMO and Bartlett's Test for Components of Workers' Working Motivation and Loyalty at Enterprises in Dong Nai Province of Vietnam**. *International Transaction Journal of Engineering, Management, & Applied Sciences & Technologies*.
- Verdonck, T., Baesens, B., Óskarsdóttir, M., & vanden Broucke, S. (2024). **Special issue on feature engineering editorial**. *Machine Learning*, 3917-3928.