



PENDAHULUAN

Analisis faktor adalah metode statistik yang digunakan untuk mengidentifikasi dan memahami hubungan antara sekelompok variabel.

Tujuan utama dari analisis faktor adalah untuk mengidentifikasi faktor-faktor yang mendasari pola korelasi antara variabel-variabel tersebut.

Langkah-langkah umum dalam melakukan analisis faktor adalah sebagai berikut:

1. Menentukan tujuan penelitian: Langkah pertama dalam analisis faktor adalah menentukan tujuan penelitian. Kita perlu memahami apa yang ingin kita ketahui atau jelaskan melalui analisis faktor. Tujuan utama dari analisis faktor adalah untuk menyederhanakan data kompleks dengan mengidentifikasi faktor-faktor yang mendasari variabilitas dalam data tersebut. Dengan demikian, analisis faktor dapat membantu mengungkap struktur yang tersembunyi dalam data dan memberikan wawasan yang berharga bagi pengambilan keputusan.



- 2. Mengumpulkan data: Selanjutnya, kita perlu mengumpulkan data yang relevan untuk analisis faktor. Data ini dapat berupa data survei, data observasi, atau data dari sumber lainnya yang sesuai dengan tujuan penelitian kita.
- 3. Menentukan jumlah faktor: Setelah mengumpulkan data, kita perlu menentukan jumlah faktor yang akan diekstraksi. Hal ini dapat dilakukan dengan menggunakan metode-metode seperti Kaiser's Rule atau scree plot.
- 4. Mengevaluasi kecocokan model: Selanjutnya, kita perlu mengevaluasi kecocokan model faktor yang kita ekstraksi. Metode yang umum digunakan untuk ini adalah analisis komponen utama, di mana kita menghitung komunalitas, eigenvalue, dan variance explain d oleh setiap faktor.
- 5. Menafsirkan faktor: Setelah mengkonfirmasi model faktor yang baik, kita perlu menafsirkan faktor-faktor yang diekstraksi. Halini melibatkan melihat pada pola beban faktor (factor loadings) dan memahami arti di balik setiap faktor.
- 6. Membuat interpretasi dan laporan: Langkah terakhir adalah membuat interpretasi dari hasil analisis faktor dan menyusun laporan penelitian yang merangkum temuan kita.



IMPORT DATA

Kode di samping digunakan untuk menginisiasi sebuah variabel fifa_results yang akan digunakan untuk menyimpan data dari file CSV yang berlokasi di 'C:\Users\TOSHIBA\Downloads\S2 FMIPA MATEMATIKA UI\Analisis Statistika Lanjut\Project UAS\FIFA 2018 Statistics.csv'. Setelah membaca file CSV, kolom-kolom dalam data tersebut akan diubah namanya agar lebih mudah dalam pemrograman.

TAMPILAN DATA

Date ÷	Team	Opponent	Goal.Scored *	Ball.Possession	Attempts	On.Target
14-06-2018	Russia	Saudi Arabia	5	40	13	7
14-06-2018	Saudi Arabia	Russia	0	60	6	0
15-06-2018	Egypt	Uruguay	0	43	8	3
15-06-2018	Uruguay	Egypt	1	57	14	4
15-06-2018	Morocco	Iran	0	64	13	3
15-06-2018	Iran	Morocco	1	36	8	2
15-06-2018	Portugal	Spain	3	39	8	3
15-06-2018	Spain	Portugal	3	61	12	5
16-06-2018	France	Australia	2	51	12	5
16-06-2018	Australia	France	1	49	4	1
16-06-2018	Argentina	Iceland	1	72	26	7
16-06-2018	Iceland	Argentina	1	28	9	3
16-06-2018	Peru	Denmark	0	52	18	6
16-06-2018	Denmark	Peru	1	48	10	3
17-06-2018	Croatia	Nigeria	2	54	11	2
17-06-2018	Nigeria	Croatia	0	46	14	2
17-06-2018	Costa Rica	Serbia	0	50	10	3
17-06-2018	Serbia	Costa Rica	1	50	10	3
17-06-2018	Germany	Mexico	0	60	25	9
17-06-2018	Mexico	Germany	1	40	12	4
17-06-2018	Brazil	Switzerland	1	52	20	4

VARIABEL	DESKRIPSI	VARIABEL	DESKRIPSI					
Date	Tanggal pertandingan	Passes	Jumlah umpan yang dilakukan oleh tim tersebut					
Team	Tim yang bermain	Distance.Covered.K	Jarak yang ditempuh oleh tim tersebut dalam kilometer					
Opponent	Tim lawan	Fouls.Committed	Jumlah pelanggaran yang dilakukan oleh tim tersebut					
Goal.Scored	Jumlah gol yang dicetak oleh tim tersebut	Yellow.Card	Jumlah kartu kuning yang diterima oleh pemain tim tersebut					
Ball.Possessio	Persentase penguasaan bola oleh tim tersebut	Yellow.Red	Jumlah kartu kuning kedua yang mengakibatkan kartu merah oleh pemain tim tersebut					
Attempts	Jumlah percobaan tembakan	Red	Jumlah kartu merah yang diterima oleh pemain tim tersebut					
On.Target	Jumlah tembakan yang mengarah ke gawang	Man.of.the.Match	Pemain terbaik dalam pertandingan tersebut					
Off.Targe	Jumlah tembakan yang tidak mengarah ke gawang	first_goal	Waktu terjadinya gol pertama dalam pertandingan tersebut					
Blocked	Jumlah tembakan yang diblokir oleh pemain lawan	round	Babak atau putaran dari pertandingan tersebut					
Corners	Jumlah tendangan pojok yang diberikan oleh tim tersebut	PSO	Apakah pertandingan tersebut dilanjutkan dengan adu penalti (Penalty Shootout) atau tidak					
Offsides	Jumlah pelanggaran offside oleh tim tersebut	Goals.in.PSO	Jumlah gol yang dicetak dalam adu penalti					
Free.Kicks	Jumlah tendangan bebas yang diberikan oleh tim tersebut	Own.Goals	Jumlah gol bunuh diri yang terjadi dalam pertandingan tersebut					
Saves	Jumlah penyelamatan yang dilakukan oleh kiper tim tersebut	Own.Goals.Time	Waktu terjadinya gol bunuh diri					
Pass.Accurac	Persentase akurasi umpan oleh tim tersebut							

VARIABEL

SELEKSI VARIABEL

```
# Taking only the numeric columns

df <- fifa_results[,c(4:20,22)]

# Check the number of rows and columns

print("no. of rows & columns")
```

Pada kode di atas, terdapat beberapa langkah yang dilakukan: Membuat variabel df yang akan menyimpan data dari fifa_results. Dalam variabel df, hanya kolom-kolom numerik yang dipilih untuk diproses. Kolom-kolom numerik tersebut memiliki indeks dari 4 hingga 20 dan juga indeks 22. Menggunakan fungsi dim() untuk mencetak jumlah baris dan kolom dari dataframe df yang mana dalam dataframe df ini terdapat 128 baris (team) dan 18 kolom (variabel). Pernyataan print("no. of rows & columns") digunakan untuk memberikan label pada output agar lebih deskriptif.

Pemilihan variabel hanya dari jenis numerik dalam analisis faktor didasarkan pada beberapa alasan yang mendasar. Berikut adalah beberapa alasan mengapa variabel numerik lebih umum digunakan dalam analisis faktor:

- 1. Skala pengukuran yang jelas: Variabel numerik memiliki skala pengukuran yang jelas dan terukur. Dalam analisis faktor, perhitungan hubungan antara variabel-variabel menggunakan metode kovariansi atau korelasi. Oleh karena itu, penting untuk memiliki variabel yang dapat diukur secara kuantitatif dan memiliki skala pengukuran yang jelas.
- 2. Kemampuan perhitungan statistik: Variabel numerik memungkinkan perhitungan statistik yang lebih lanjut, seperti menghitung rata-rata, varians, korelasi, dan lainnya. Ini memungkinkan analisis faktor untuk menggali lebih dalam hubungan antara variabel-variabel numerik dan faktor-faktor laten yang mendasarinya.
- 3. Representasi informasi yang lebih kaya: Variabel numerik dapat menggambarkan informasi yang lebih kaya daripada variabel kategorikal. Dalam analisis faktor, variasi dalam data dianalisis untuk mengungkap struktur faktor laten. Variabel numerik dapat memberikan variasi yang lebih besar dan beragam, sehingga memberikan gambaran yang lebih lengkap tentang struktur yang mendasarinya.

Meskipun variabel numerik lebih umum dalam analisis faktor, tidak berarti variabel kategorikal tidak dapat digunakan sama sekali. Dalam beberapa kasus, variabel kategorikal dapat diubah menjadi bentuk yang sesuai dengan analisis faktor, seperti menggunakan teknik pengkodean kategori atau menggantikan variabel dengan indikator biner. Namun, dalam prakteknya, analisis faktor pada umumnya lebih mudah dan lebih tepat diterapkan pada variabel numerik yang langsung mencerminkan tingkat atau jumlah suatu atribut.

MENGATASI MISSING VALUE

```
# Replacing NA values with 0

df[is.na(df)] <- 0

head(df)
```

Berdasarkan kode di atas, terlihat bahwa penting untuk mengganti nilai-nilai yang hilang (NA) (missing value) dalam dataframe df dengan angka 0 menggunakan pernyataan df[is.na(df)] <- 0. Hal ini dilakukan untuk memungkinkan analisis statistika lanjut atau pemrosesan data lainnya tanpa terpengaruh oleh nilai yang hilang.

```
Goal.Scored Ball.Possession Attempts On.Target Off.Target Blocked Corners
## 1
                                      13
                                       6
## 2
                                       8
                             43
## 3
                                      14
                             57
## 4
## 5
                             64
                                      13
                                       8
                             36
## 6
    Offsides Free.Kicks Saves Pass.Accuracy Passes Distance.Covered.Kms
                           0
2
3
## 1
                     11
                                         78
                                               306
                                                                   118
## 2
                     25
                                         86
                                               511
                                                                   105
## 3
                                         78
                                              395
                                                                   112
           1
                     13
## 4
                                         86
                                               589
                                                                   111
           0
                     14
## 5
                                               433
                                         86
                                                                   101
                     22
                                               194
## 6
                                                                   100
    Fouls.Committed Yellow.Card Yellow.Red Red first_goal
## 1
                 22
                                                       12
                 10
## 2
                                                        0
## 3
                12
## 4
                 6
                                                       89
                 22
## 5
## 6
                 14
                                                       90
```

MENENTUKAN JUMLAH FAKTOR

TES KAISER-MEYER OLKIN (KMO)

Selanjutnya, temukan variabel mana yang dapat dianggap untuk Analisis Faktor: Variabel mana yang relevan?

Tes Kaiser-Meyer-Olkin (KMO) dapat dilakukan untuk menjawab pertanyaan ini. Langkah-langkah melakukan tes KMO pada program R.

- 1. Mengimpor paket psych menggunakan pernyataan library(psych). Paket ini digunakan untuk melakukan analisis psikologis dan termasuk fungsi yang relevan untuk analisis faktor.
- 2. Membuat matriks korelasi dengan menggunakan fungsi cor() pada dataframe df. Matriks korelasi ini akan digunakan untuk menghitung KMO.
- 3. Menggunakan fungsi KMO() dari paket psych untuk menghitung faktor kecukupan KMO. Fungsi ini akan mengevaluasi apakah variabel-variabel dalam matriks korelasi tersebut cocok untuk dilakukan analisis faktor. Output dari fungsi ini akan memberikan skor KMO dan juga interpretasinya.

Berdasarkan output KMO di samping, terdapat skor MSA (Measure of Sampling Adequacy) adalah ukuran untuk pengecualian variabel.

Jika MSA < 0,5, maka variabel tersebut sebaiknya dihilangkan. Variabel dengan MSA > 0,6 cocok. Variabel dengan MSA > 0,8 sangat cocok untuk analisis faktor.

Hasil tersebut memberitahu kita untuk menghilangkan variabel "On.Target", "Off.Target", "Blocked", "Offsides", "Distance.Covered.Kms", "Yellow.Red", "Red", dan "first_goal".

library(psych)

```
## Warning: package 'psych' was built under R version 4.1.3
```

df_corr <- cor(df) # Create a correlation matrix
KMO(df_corr) # Kaiser-Meyer-Olkin factor adequacy</pre>

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = df corr)
## Overall MSA = 0.56
## MSA for each item =
                             Ball.Possession
           Goal.Scored
                                                         Attempts
##
##
                  0.62
                                       0.74
                                                             0.52
                                                          Blocked
             On.Target
                                 Off.Target
##
                   0.37
                                        0.35
                                                             0.43
                                   Offsides
                                                       Free Kicks
##
               Corners
                  0.93
                                        0.49
                                                             0.53
##
                               Pass.Accuracy
                  Saves
                                                           Passes
##
                  0.78
                                        0.83
                                                             0.70
                             Fouls.Committed
## Distance Covered Kms
                                                      Yellow.Card
##
                  0.22
                                                             0.60
                                        0.68
            Yellow.Red
##
                                         Red
                                                       first goal
##
                   0.49
                                                             0.45
                                        0.52
```

```
df1 <- df[,c("Goal.Scored","Ball.Possession","Attempts","Corners","Free.Kicks","Saves","Pass.Accuracy","Passes","Fouls.Committe
d","Yellow.Card","Red")]
head(df1)</pre>
```

```
Goal.Scored Ball.Possession Attempts Corners Free.Kicks Saves Pass.Accuracy
## 1
                                      13
                                                        11
                                                               0
                                                                            78
                             60
## 2
                                                                            86
                                      8
14
## 3
                             43
                                                                            78
                             57
## 4
                                                                            86
                                                        14
## 5
                             64
                                      13
                                                                            86
                             36
                                                        22
                                                               2
## 6
                                                                            86
    Passes Fouls.Committed Yellow.Card Red
## 1
       306
                        22
                                         0
## 2
       511
                        10
## 3
       395
                        12
## 4
       589
## 5
       433
                        22
                                     3
                        14
## 6
       194
```

MATRIKS KORELASI

Selanjutnya kita akan membuat matriks korelasi dari dataframe yang telah diekstraksi. Pada kode yang diberikan di bawah, kita akan menggunakan package corrplot untuk memvisualisasikan matriks korelasi.

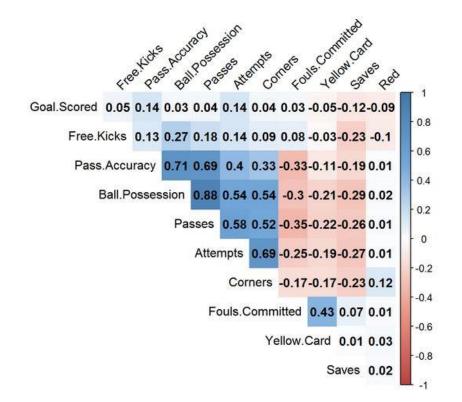
Lakukan perhitungan matriks korelasi menggunakan fungsi cor() yang hasilnya disimpan dalam variabel df1_corr.

Selanjutnya, lakukan pemetaan warna pada matriks korelasi menggunakan fungsi colorRampPalette() dengan menggunakan skala warna yang ditentukan dalam vektor warna c("#BB4444", "#EE9988", "#FFFFFF", "#77AADD", "#4477AA").

Gunakan fungsi corrplot() untuk membuat plot matriks korelasi. Beberapa argumen yang digunakan adalah sebagai berikut:

- round(df1_corr, 2): Matriks korelasi yang dibulatkan menjadi dua desimal, akan digunakan sebagai input untuk plot;
- method="color": Metode visualisasi menggunakan warna;
- col=col(200): Skema warna yang diberikan sebagai argumen col dari fungsi corrplot;
- type="upper": Hanya memperlihatkan bagian atas matriks korelasi;
- order="hclust": Mengurutkan variabel menggunakan metode hclust;
- addCoef.col="black": Menambahkan koefisien korelasi dengan warna hitam;
- tl.col="black": Warna teks label;
- tl.srt=45: Rotasi teks label sebesar 45 derajat;
- diag=FALSE: Menyembunyikan koefisien korelasi di diagonal utama matriks korelasi.

Terakhir, lakukan pencetakan matriks korelasi yang telah dibulatkan menjadi dua desimal menggunakan fungsi round(df1_corr, 2). Ini akan mencetak matriks korelasi dalam bentuk tabel.



##	Goal.Scored Ba	all.Pos	session	Attempts	Corners Fre	e.Kicks	Saves
## Goal.Scored	1.00		0.03	0.14	0.04	0.05	-0.12
## Ball.Possession	0.03		1.00	0.54	0.54	0.27	-0.29
## Attempts	0.14		0.54	1.00	0.69	0.14	-0.27
## Corners	0.04		0.54	0.69	1.00	0.09	-0.23
## Free.Kicks	0.05		0.27	0.14	0.09	1.00	-0.23
## Saves	-0.12		-0.29	-0.27	-0.23	-0.23	1.00
## Pass.Accuracy	0.14		0.71	0.40	0.33	0.13	-0.19
## Passes	0.04		0.88	0.58	0.52	0.18	-0.26
## Fouls.Committed	0.03		-0.30	-0.25	-0.17	0.08	0.07
## Yellow.Card	-0.05		-0.21	-0.19	-0.17	-0.03	0.01
## Red	-0.09		0.02	0.01	0.12	-0.10	0.02
##	Pass.Accuracy	Passes	Fouls.C	ommitted	Yellow.Card	Red	
## Goal.Scored	0.14	0.04		0.03	-0.05	-0.09	
## Ball.Possession	0.71	0.88		-0.30	-0.21	0.02	
## Attempts	0.40	0.58		-0.25	-0.19	0.01	
## Corners	0.33	0.52		-0.17	-0.17	0.12	
## Free.Kicks	0.13	0.18		0.08	-0.03	-0.10	
## Saves	-0.19	-0.26		0.07	0.01	0.02	
## Pass.Accuracy	1.00	0.69		-0.33	-0.11	0.01	
## Passes	0.69	1.00		-0.35	-0.22	0.01	
## Fouls.Committed	-0.33	-0.35		1.00	0.43	0.01	
## Yellow.Card	-0.11	-0.22		0.43	1.00	0.03	
## Red	0.01	0.01		0.01	0.03	1.00	

Berdasarkan output di atas, matriks korelasi yang diberikan adalah representasi dari hubungan korelasi antara variabel-variabel yang terkait dengan analisis faktor. Setiap angka dalam matriks menunjukkan koefisien korelasi antara dua variabel yang berdekatan.

INTERPRETASI MATRIKS KORELASI

Berikut adalah beberapa interpretasi dari matriks korelasi tersebut:

- Variabel "Yellow.Card" memiliki korelasi negatif dengan variabel "Goal.Scored", "Ball.Possession", "Attempts", "Corners", "Free.Kicks", dan "Pass.Accuracy". Artinya, semakin tinggi jumlah kartu kuning yang diterima, semakin rendah skor gol, kepemilikan bola, upaya menyerang, tendangan sudut, tendangan bebas, dan akurasi umpan.
- 2. Variabel "Red" memiliki korelasi positif dengan variabel "Corners" dan korelasi negatif dengan variabel "Ball.Possession" dan "Free.Kicks". Artinya, semakin tinggi jumlah kartu merah yang diterima, semakin tinggi jumlah tendangan sudut, dan semakin rendah kepemilikan bola dan jumlah tendangan bebas.
- Variabel "Saves" memiliki korelasi negatif dengan variabel "Ball.Possession", "Attempts", "Corners", "Pass.Accuracy", dan "Passes". Artinya, semakin tinggi jumlah penyelamatan, semakin rendah kepemilikan bola, upaya menyerang, tendangan sudut, akurasi umpan, dan jumlah umpan yang dilakukan.
- 4. Variabel "Pass.Accuracy" dan "Passes" memiliki korelasi positif dengan variabel "Ball.Possession" dan korelasi negatif dengan variabel "Fouls.Committed" dan "Yellow.Card". Artinya, semakin tinggi akurasi umpan dan jumlah umpan yang dilakukan, semakin tinggi kepemilikan bola, dan semakin rendah jumlah pelanggaran dan kartu kuning.
- 5. Variabel "Fouls.Committed" dan "Yellow.Card" memiliki korelasi positif. Artinya, semakin tinggi jumlah pelanggaran yang dilakukan, semakin tinggi kemungkinan untuk menerima kartu kuning.

METODE MINRES DAN ROTASI VARIMAX

```
nfactors <- 4 # Based on the screeplot suggestion
nvars <- dim(dfi_corr)[1]
factors <- fa(r = dfi_corr, nfactors = nfactors, rotate = "Varimax")
```

factors

```
## Factor Analysis using method = minres
## Call: fa(r = df1_corr, nfactors = nfactors, rotate = "Varimax")
## Standardized loadings (pattern matrix) based upon correlation matrix
                    MR1 MR3 MR2 MR4 h2 u2 com
## Goal Scored
                  0.05 0.06 0.01 0.25 0.067 0.933 1.2
## Ball.Possession 0.93 0.25 -0.10 0.06 0.932 0.068 1.2
## Attempts
                  0.41 0.62 -0.14 0.19 0.614 0.386 2.1
## Corners
                  0.33 0.89 -0.05 -0.03 0.897 0.103 1.3
## Free.Kicks
                  0.24 0.02 0.12 0.38 0.214 0.786 1.9
## Saves
                  -0.25 -0.18 -0.01 -0.34 0.207 0.793 2.4
## Pass.Accuracy 0.72 0.10 -0.17 0.03 0.563 0.437 1.2
                  0.86 0.28 -0.18 0.02 0.856 0.144 1.3
## Fouls.Committed -0.20 -0.06 0.98 0.01 0.996 0.004 1.1
## Yellow.Card
                 -0.11 -0.12 0.41 -0.06 0.200 0.300 1.4
                  0.01 0.12 0.03 -0.26 0.083 0.917 1.4
                        MR1 MR3 MR2 MR4
## SS loadings
                       2.58 1.39 1.23 0.43
## Proportion Var
                       0.23 0.13 0.11 0.04
## Cumulative Var
                       0.23 0.36 0.47 0.51
## Proportion Explained 0.46 0.25 0.22 0.08
## Cumulative Proportion 0.46 0.70 0.92 1.00
## Mean item complexity = 1.5
## Test of the hypothesis that 4 factors are sufficient.
## df null model = 55 with the objective function = 4.21
## df of the model are 17 and the objective function was 0.17
## The root mean square of the residuals (RMSR) is 0.03
## The df corrected root mean square of the residuals is 0.05
## Fit based upon off diagonal values = 0.99
## Measures of factor score adequacy
                                                   MR1 MR3 MR2 MR4
## Correlation of (regression) scores with factors 0.96 0.93 1.00 0.60
## Multiple R square of scores with factors
                                                  0.93 0.87 1.00 0.36
## Minimum correlation of possible factor scores 0.86 0.75 0.99 -0.28
```

Berikut adalah interpretasi dari hasil menggunakan metode "minres" dan rotasi "Varimax":

Loadings Standar (matriks pola) berdasarkan matriks korelasi:

- 1. MR1 (Faktor 1): Memuat 2.58 variabel;
- 2. MR3 (Faktor 3): Memuat 1.39 variabel;
- 3. MR2 (Faktor 2): Memuat 1.23 variabel;
- 4. MR4 (Faktor 4): Memuat 0.43 variabel;

SS Loadings: Jumlah kuadrat standar (SS) untuk setiap faktor. Faktor 1 memiliki SS Loadings tertinggi yaitu 2.58, diikuti oleh Faktor 3 dengan 1.39, Faktor 2 dengan 1.23, dan Faktor 4 dengan 0.43.

Proportion Var: Proporsi varians yang dijelaskan oleh setiap faktor. Faktor 1 menjelaskan 23% varians, Faktor 3 menjelaskan 13% varians, Faktor 2 menjelaskan 11% varians, dan Faktor 4 menjelaskan 4% varians.

Cumulative Var: Jumlah kumulatif proporsi varians yang dijelaskan oleh faktor-faktor sejauh ini. Pada Faktor 1, jumlah kumulatif proporsi varians adalah 23%, pada Faktor 3 adalah 36%, pada Faktor 2 adalah 47%, dan pada Faktor 4 adalah 51%.

Proportion Explained: Proporsi varians yang dijelaskan oleh setiap faktor secara relatif terhadap total varians. Faktor 1 menjelaskan 46% varians, Faktor 3 menjelaskan 25% varians, Faktor 2 menjelaskan 22% varians, dan Faktor 4 menjelaskan 8% varians.

Cumulative Proportion: Jumlah kumulatif proporsi varians yang dijelaskan oleh faktor-faktor secara relatif terhadap total varians. Pada Faktor 1, jumlah kumulatif proporsi varians adalah 46%, pada Faktor 3 adalah 70%, pada Faktor 2 adalah 92%, dan pada Faktor 4 adalah 100%.

Mean item complexity: Kompleksitas rata-rata item dalam analisis faktor, yaitu 1.5.

Test of the hypothesis that 4 factors are sufficient: Pengujian hipotesis bahwa 4 faktor sudah cukup untuk menjelaskan data. Dalam kasus ini, df null model adalah 55 dengan nilai fungsi tujuan 4.21, sedangkan df model yang digunakan adalah 17 dengan nilai fungsi tujuan 0.17.

Root mean square of the residuals (RMSR): Nilai akar rata-rata kuadrat residual, yaitu 0.03.

df corrected root mean square of the residuals: Nilai akar rata-rata kuadrat residual yang dikoreksi berdasarkan derajat kebebasan, yaitu 0.05.

Fit based upon off diagonal values: Nilai kesesuaian berdasarkan nilai di luar diagonal, yaitu 0.99.

Measures of factor score adequacy: Ukuran kecukupan skor faktor.

Korelasi skor regresi dengan faktor-faktor: Korelasi antara skor regresi dengan faktor-faktor adalah 0.96 untuk Faktor 1, 0.93 untuk Faktor 3, 1.00 untuk Faktor 2, dan 0.60 untuk Faktor 4.

Multiple R square of scores with factors: Nilai R kuadrat berganda antara skor dengan faktor-faktor adalah 0.93 untuk Faktor 1, 0.87 untuk Faktor 3, 1.00 untuk Faktor 2, dan 0.36 untuk Faktor 4.

Minimum correlation of possible factor scores: Korelasi minimum dari skor faktor yang mungkin adalah 0.86 untuk Faktor 1, 0.75 untuk Faktor 3, 0.99 untuk Faktor 2, dan -0.28 untuk Faktor 4.

Selain itu, kita juga bisa menilai analisis faktor berdasarkan:

Nilai h2 (merupakan koefisien determinasi (R squared) yang mengindikasikan seberapa baik variabel dijelaskan oleh faktor-faktor).

Nilai u2 (merupakan koefisien unik yang menggambarkan seberapa besar variasi yang tidak dijelaskan oleh faktor-faktor).

Nilai com (merupakan kompleksitas rata-rata item yang menunjukkan seberapa kompleks item-item tersebut dalam menjelaskan faktorfaktor).

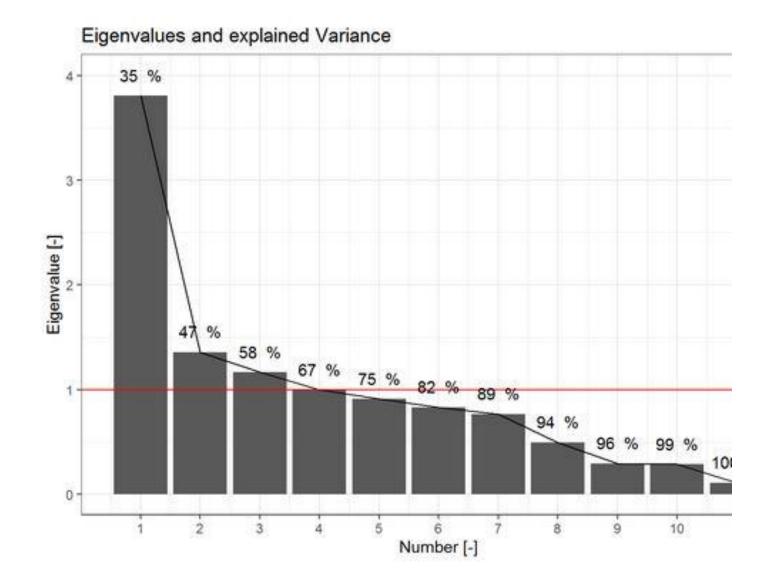
Sehingga berdasarkan nilai h2, terlihat bahwa variabel "ball possession", "corners", passes", dan "fouls committed" sangat baik dijelaskan oleh faktor-faktor.



```
library(ggplot2)
## Warning: package 'ggplot2' was built under R version 4.1.3
## Attaching package: 'ggplot2'
## The following objects are masked from 'package:psych':
      %+%, alpha
# Plot Eigenvalues / Represented Variance
eigenvalues <- data.frame(factors$e.values)
colnames(eigenvalues) <- c("Values")
eigenvalues$Number <- 1:nrow(df1_corr)
eigenvalues$RepresentedVariance <- NA
for (i in 1:nrow(df1_corr)) {
   eigenvalues$RepresentedVariance[i] <- sum(eigenvalues$Values[1:i])/sum(eigenvalues$Values) *
eigenvalues$RepresentedVariance_text <- paste(round(eigenvalues$RepresentedVariance,
   0), "%")
e1 <- ggplot(eigenvalues, aes(Number, y = Values), group = 1)
e1 <- e1 + geom_bar(stat = "identity")
e1 <- e1 + geom_line(aes(y = Values), group = 2)
e1 <- e1 + xlab("Number [-]")
e1 <- e1 + ylab("Eigenvalue [-]")
e1 <- e1 + geom_hline(aes(yintercept = 1), col = "red")
e1 <- e1 + geom_text(aes(label = RepresentedVariance_text), nudge_y = 0.2)
e1 <- e1 + ggtitle("Eigenvalues and explained Variance")
e1 <- e1 + theme_bw()
e1 <- e1 + scale_x_continuous(breaks = seq(1, 10, 1))
e1
```

SCREE-PLOT

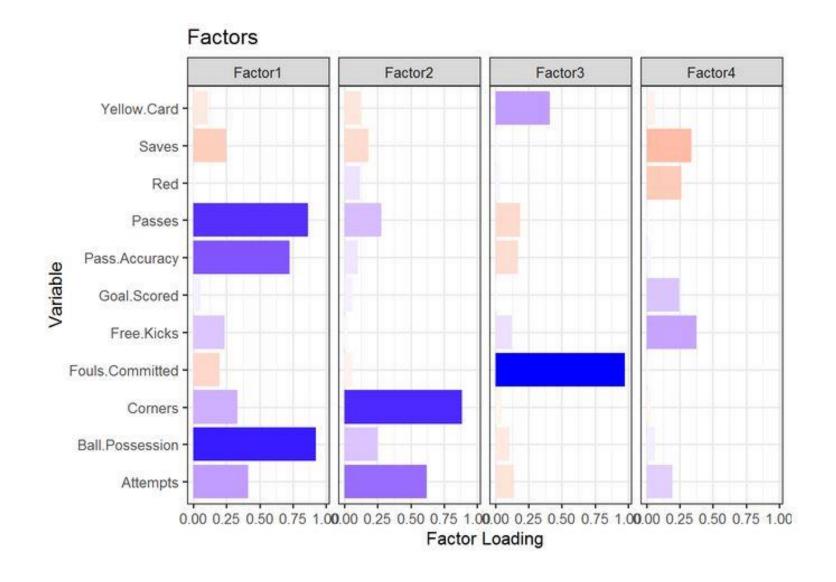
- •Berdasakran grafik di atas, menunjukan grafik Screeplot yang menampilkan Eigenvalues dan varian yang dijelaskan secara kumulatif.
- Grafik ini menggunakan paket ggplot2 untuk visualisasi.
- •Pada sumbu x, terdapat nomor faktor yang digunakan, sedangkan pada sumbu y, terdapat nilai Eigenvalues.
- Setiap batang pada grafik menunjukkan nilai Eigenvalues untuk faktor tersebut.
- Garis yang menghubungkan batang-batang tersebut merupakan representasi nilai Eigenvalues.
- Garis merah pada grafik Scree-plot menunjukkan nilai eigenvalue yang sama dengan satu.
- Eigenvalues yang lebih besar dari satu menandakan bahwa faktor tersebut mampu menjelaskan lebih banyak varian daripada varian yang dapat dijelaskan oleh variabel asli.
- •Dalam analisis faktor, kita cenderung memilih jumlah faktor sebelum eigenvalues mulai mendekati atau turun di bawah nilai satu.
- •Pemilihan jumlah faktor ini didasarkan pada asumsi bahwa faktor-faktor dengan eigenvalues lebih besar dari satu mampu menjelaskan varian yang signifikan dalam data.
- •Sehingga, kesimpulannya kita dapat memilih jumlah faktor sebanyak 1, 2, 3, atau 4.



MENAFSIRKAN FAKTOR

```
loadings_mat <- as.data.frame(matrix(nrow = nvars, ncol =nfactors))
loadings_mat$variable <- colnames(df1)
for (i in 1:nfactors) {
    for (j in 1:nvars) {
        loadings_mat[j, i] <- factors$loadings[j, i]
    }
}
colnames(loadings_mat) <- c("Factor1", "Factor2", "Factor3", "Factor4", "Variable")
loadings_mat_gather <- loadings_mat %% gather("Factor", "Value", 1:nfactors)</pre>
```

FACTOR LOADINGS



INTERPRETASI

Berdasarkan output factor loadings di atas, akhirnya kita dapat memberi nama kepada faktor-faktor berdasarkan variabel yang dimuat.

- Faktor 1: Kontrol Permainan ('Ball Possession', 'Passes', dan 'Pass Accuracy');
- Faktor 2: Serangan ('Corners' dan 'Attempts');
- Faktor 3: Disiplin ('Fouls Committed' dan 'Yellow Card');
- Faktor 4: Lainnya? ('Free Kicks' dan 'Goal scored').

Variabel yang tersisa ('Saves' dan 'Red') mungkin telah dijelaskan dalam faktor-faktor lain (pertahanan, dll...) karena kita hanya memilih 4 faktor, kita hanya dapat melihat faktor-faktor ini.

Kita juga dapat melihat bahwa variabel-variabel ini memiliki beban yang negatif pada beberapa faktor.

Kita masih bisa mencoba dengan jumlah faktor yang berbeda untuk melihat apakah hal tersebut lebih masuk akal untuk data ini.



TERIMAKASIH