



## **Laporan Tugas 3**

# **Analisis Faktor Eksploratori & Analisis Faktor Konfirmatori**

Analisis Multivariat Kelas B

Kelompok L

Evan Haryowidyatna	2006485011
--------------------	------------

Muhammad Jauhar Hakim	2006463982
-----------------------	------------

Siskawati Simandalahi	2006572970
-----------------------	------------

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Indonesia

Depok

Maret 2022

## 1. Penjelasan Data

Data yang kami gunakan merupakan data risiko perilaku kanker serviks. Sumber data yang kami gunakan yaitu: <https://archive.ics.uci.edu/ml/datasets/Cervical+Cancer+Behavior+Risk> Data risiko perilaku kanker serviks terdiri dari 19 atribut data numerik yaitu :

- 1) behavior\_eating (perilaku makan)
- 2) behavior\_personalHygiene (perilaku kebersihan pribadi)
- 3) intention\_aggregation (niat agregasi)
- 4) intention\_commitment (niat komitmen)
- 5) attitude\_consistency (sikap konsisten)
- 6) attitude\_spontaneity (sikap spontanitas)
- 7) norm\_significantPerson (norma kepada orang penting)
- 8) norm\_fulfillment (norma pemenuhan)
- 9) perception\_vulnerability (persepsi kerentanan)
- 10) perception\_severity (persepsi keparahan)
- 11) motivation\_strength (motivasi kekuatan)
- 12) motivation\_willingness (motivasi kemauan)
- 13) socialSupport\_emotionality (dukungan sosial emosionalitas)
- 14) socialSupport\_appreciation (dukungan sosial apresiasi)
- 15) socialSupport\_instrumental (dukungan sosial instrumental)
- 16) empowerment\_knowledge (pemberdayaan pengetahuan)
- 17) empowerment\_abilities (pemberdayaan kemampuan)
- 18) empowerment\_desires (pemberdayaan keinginan)
- 19) ca\_cervix (merupakan atribut kategorik, 1 = memiliki kanker serviks, 0 = tidak memiliki kanker serviks)

## II. Tujuan

Analisis faktor adalah analisis statistika yang bertujuan untuk mereduksi dimensi data dengan cara menyatakan variabel asal sebagai kombinasi linear sejumlah faktor, sedemikian hingga sejumlah faktor tersebut mampu menjelaskan sebesar mungkin keragaman data yang terkandung dalam variabel asal.

Analisis faktor terbagi dua, yaitu analisis faktor eksploratori dan analisis faktor konfirmatori. Analisis faktor eksploratori merupakan suatu teknik analisis faktor dimana beberapa faktor yang akan terbentuk berupa variabel laten yang belum dapat ditentukan sebelum analisis dilakukan, sedangkan analisis faktor konfirmatori yaitu suatu teknik analisis faktor dimana faktor yang dibentuk berdasarkan teori dan konsep yang sudah diketahui atau ditentukan sebelumnya beserta variabel apa saja yang dapat mengukur masing-masing faktor yang dibentuk. Dengan kata lain analisis faktor konfirmatori dapat di ciri kan sebagai analisis yang sangat mengarah pada dunia statistika, di mana kita melakukan pengujian hipotesis atas faktor yang kita sudah tentukan sebelumnya.

## III. Dasar Teori

Analisis faktor adalah sebuah teknik yang digunakan untuk mencari faktor-faktor yang mampu menjelaskan hubungan atau korelasi antara berbagai indikator independen yang diobservasi. Analisis faktor merupakan perluasan dari analisis komponen utama. Digunakan juga untuk mengidentifikasi sejumlah faktor yang relatif kecil yang dapat digunakan untuk menjelaskan sebagian besar variabel yang saling berhubungan.

Sehingga variabel-variabel dalam satu faktor mempunyai korelasi yang tinggi, sedangkan korelasi dengan variabel-variabel pada faktor lain relatif rendah. Tiap-tiap kelompok dari variabel mewakili suatu konstruksi dasar yang disebut faktor. Untuk meningkatkan daya interpretasi faktor, harus dilakukan transformasi pada matriks *loadings*. Transformasi dilakukan dengan merotasi matriks tersebut.

Hasil rotasi ini akan mengakibatkan setiap variabel asal mempunyai korelasi tinggi dengan faktor tertentu saja dan dengan faktor yang lain korelasi relatif rendah sehingga setiap faktor akan lebih mudah untuk diinterpretasikan.

Analisis faktor terbagi menjadi dua macam yaitu:

### 1. Analisis Faktor Eksploratori

Analisis faktor eksploratori adalah teknik analisis faktor di mana beberapa faktor yang akan terbentuk berupa variabel laten yang belum dapat ditentukan sebelum analisis dilakukan. Pada prinsipnya analisis faktor eksploratori adalah suatu teknik yang bertujuan untuk mereduksi data dari variabel asal atau variabel awal menjadi variabel baru atau faktor yang jumlahnya lebih kecil dari pada variabel awal. Proses analisis tersebut mencoba untuk menemukan hubungan antar variabel baru atau faktor yang terbentuk yang saling independen, sehingga bisa dibuat satu atau beberapa kumpulan variabel laten atau faktor yang lebih sedikit dari jumlah variabel awal yang bebas atau tidak berkorelasi. Jadi antar faktor yang terbentuk tidak berkorelasi.

### 2. Analisis Faktor Konfirmatori

Analisis faktor konfirmatori yaitu suatu teknik analisis faktor berdasarkan teori dan konsep yang sudah diketahui atau ditentukan sebelumnya, maka dibuat sejumlah faktor yang akan dibentuk serta variabel apa saja yang termasuk ke dalam masing-masing faktor yang dibentuk dan sudah pasti tujuannya. Pembentukan faktor konfirmatori (CFA) secara sengaja berdasarkan teori dan konsep, dalam upaya untuk mendapatkan variabel baru atau faktor yang mewakili beberapa item atau sub-variabel yang merupakan variabel teramati (*observerb variable*).

### KMO dan *Bartlett Test of Sphericity*

Sebelum melakukan analisis faktor terlebih dahulu dilakukan pengujian untuk memeriksa kelayakan perlu tidaknya analisis faktor dilakukan melalui pengujian Kaiser Mayer Olkin (KMO) *Measure of Sampling Adequacy* dan *Bartlett Test of Sphericity*. Kaiser Mayer Olkin (KMO) *Measure of Sampling Adequacy* adalah mengukur kecukupan sampling (sampling adequacy). Nilai ini membandingkan besarnya nilai koefisien korelasi yang diamati terhadap koefisien korelasi parsial.

Nilai KMO yang kecil atau  $< 0.5$  menunjukkan bahwa bahwa korelasi antar pasangan variabel tidak bisa diterangkan oleh variabel lainnya dan analisis faktor tidak dapat dilakukan. Sedangkan *Bartlett Test of Sphericity* dilakukan untuk menguji apakah matriks korelasi adalah matriks identitas atau bukan.

Berikut ini merupakan kriteria nilai uji KMO dari matriks antara variabel di antaranya adalah:

Nilai uji  $0,9 < KMO \leq 1,00$  = sangat memuaskan,

Nilai uji  $0,8 < KMO \leq 0,9$  = sangat baik,

Nilai uji  $0,7 < KMO \leq 0,8$  = baik,

Nilai uji  $0,6 < KMO \leq 0,7$  = cukup baik,

Nilai uji  $0,5 < KMO \leq 0,6$  = jelek,

Nilai uji  $\leq 0,5$  = ditolak.

Dan uji hipotesis untuk *Bartlett Test of Sphericity* adalah

$H_0$  : Matriks korelasi data merupakan matriks identitas

$H_1$  : Matriks korelasi data bukan merupakan matriks identitas

Untuk mengestimasi parameter *loadings* dan *communalities* dengan metode *maximum likelihood* dibutuhkan asumsi multivariat normal, sehingga berikut ini merupakan uji dari asumsi multivariat normal yaitu

$H_0$  : Data berdistribusi multivariat normal

$H_0$  : Data tidak berdistribusi multivariat normal

#### IV. Pemodelan

Langkah kerja analisis faktor eksploratori :

1. Menguji apakah data bisa dilakukan analisis faktor dengan uji KMO (Kaiser Meyer Olkin) dan *Bartlett Test of Sphericity*
2. Uji asumsi multivariat normal
3. Penentuan berapa banyak faktor yang bisa diekstrak.
4. Ekstraksi dan rotasi faktor
5. Penamaan faktor jika memungkinkan
6. Menghitung skor faktor
7. Membuat *path* diagram dari data beserta faktor
8. Interpretasi hasilnya

Langkah kerja analisis faktor konfirmatori :

1. Definisikan hipotesis modelnya (persamaan model dan *path* diagram)
2. Estimasi model parameter
3. Uji Hipotesis goodness of fit (chi-square, CFI, RMSEA, SRMR)
4. Buat *path* diagram beserta skor faktor
5. Interpretasi hasilnya

## V. Proses Komputasi

*Install package dan apply library*

```
packages <- c("Hmisc", "matlib", "Matrix", "expm", "matrixcalc", "ellipsis",
              "Hotelling", "dplyr", "psych", "RcmdrMisc", "Rcsdp", "mvnrmtest", "lavaan",
              "Ssemplot")

if ( length(missing_pkgs <- setdiff(packages, rownames(installed.packages()))) > 0)
{
  message("Installing missing package(s): ", paste(missing_pkgs, collapse = ", "))
  install.packages(missing_pkgs)
}

lapply(packages, library, character.only = TRUE)
```

### Proses Komputasi EFA

#### 1. Impor File

```
data <- read.table("sobar72.csv", header=TRUE, sep=",")
data <- as.matrix(data)
```

#### 2. Membuat matriks korelasi dari data

```
cor(data)
```

#### 3. Mencari KMO dan *Bartlett Test*

```
KMO(cor(data))
cortest.bartlett(cor(data), n = nrow(data))
```

#### 4. Uji asumsi multivariat normal

```
mshapiro.test(t(data))
```

## 5. Menentukan banyak faktor yang akan diekstrak

-Membuat vektor eigen dan nilai eigen

```
eig <- eigen(cordata)
```

```
eig
```

-Mencari Variansi Kumulatif

```
cumpro <- cumsum(eig$values^2 / sum(eig$values^2))
```

```
Cumpro
```

-Membuat *Parallel Analysis Scree Plot*

```
fa.parallel(data, fa="fa")
```

## 6. Menjalankan EFA dengan rotasi faktor dan distribusi variabel

```
EFA <- fa(data, nfactors = 2, rotate = "Promax", fm = "pa")
```

```
EFA$loadings
```

## 7. Membuat diagram *loadings* dan menghapus faktor yang tidak berpengaruh besar

```
loads <- EFA$loadings
```

```
fa.diagram(loads)
```

```
#drop kolom attention aggregation (4)
```

```
#drop kolom behavior sexual risk (1)
```

```
#drop kolom intention commitment (5)
```

```
#drop kolom attitude consistency (6)
```

```
#drop kolom attitude spontaneity (7)
```

```
#drop kolom behavior eating (2)
```

```
loads[-c(1,2,4,5,6,7), ]
```

```
fa.diagram(loads[-c(1,2,4,5,6,7), ])
```



## Proses Komputasi CFA

### 1. Membuat model dari faktor yang sudah kami tentukan

```
hs.mod <- '
    factor1 =~ empowerment_abilities + socialSupport_emotionality +
empowerment_knowledge + empowerment_desires + socialSupport_appreciation +
socialSupport_instrumental + motivation_willingness + ca_cervix

    factor2 =~ norm_fulfillment + perception_severity +
perception_vulnerability + norm_significantPerson + ca_cervix
```

### 2. Menjalankan proses CFA dan mengeluarkan hasil dengan modul Lavaan

```
hs.fit <- cfa(hs.mod, data = data)

hs.fit

summary(hs.fit, fit.measures = TRUE, standardized = TRUE)
```

### 3. Membuat path diagram beserta factor score nya

```
semPaths(hs.fit, "std", layout = "tree", intercepts = F, residuals = T,
nDigits = 3, label.cex = 1, edge.label.cex=.95, fade = F)
```

## VI. Hasil Komputasi

### Hasil Komputasi EFA

#### 1. Mengimpor Data

behavior_sexualRisk	behavior_eating	behavior_personalHygiene	intention_aggregation	intention_commitment
10	13	12	4	7
10	11	11	10	14
10	15	3	2	14
10	11	10	10	15
8	11	7	8	10
10	14	8	6	15
10	15	4	6	14
8	12	9	10	10
10	15	7	2	15
7	15	7	6	11
7	15	7	10	14
10	15	8	9	15
10	15	12	10	15
9	12	14	9	15

#### 2. Membuat matriks korelasi dari data

	behavior_sexualRisk	behavior_eating	behavior_personalHygiene	intention_aggreg
behavior_sexualRisk	1.000000000	-0.165857419	0.003911805	-0.005778998
behavior_eating	-0.165857419	1.000000000	0.224623078	0.116634406
behavior_personalHygiene	0.003911805	0.224623078	1.000000000	0.441812169
intention_aggregation	-0.005778998	0.116634406	0.441812169	1.000000000
intention_commitment	0.126616118	0.116074795	0.009612683	0.265216361
attitude_consistency	-0.067541012	0.124196640	0.152173227	-0.039641814
attitude_spontaneity	-0.057419533	0.307610235	-0.118432846	-0.175529145
norm_significantPerson	0.057869139	0.038375946	0.237062616	0.116700620
norm_fulfillment	0.158799335	-0.048262113	0.252654885	0.059117769
perception_vulnerability	0.175791278	-0.000406886	0.139975249	-0.044996908
perception_severity	0.067469233	-0.077466622	0.252095300	0.061594841
motivation_strength	-0.041937374	-0.141731699	0.386605083	0.337716869
motivation_willingness	0.309356605	-0.077380016	0.432541729	0.277540427

### 3. Mencari KMO dan Bartlett Test

```
Kaiser-Meyer-Olkin factor adequacy
Call: KMO(r = cor(data))
Overall MSA = 0.73
MSA for each item =
```

	behavior_sexualRisk	behavior_eating
	0.47	0.39
behavior_personalHygiene		
	0.69	0.58
intention_commitment		
	0.56	0.36
attitude_spontaneity		
	0.43	0.68
norm_fulfillment		
	0.79	0.78
perception_severity		
	0.72	0.80
motivation_willingness		
	0.86	0.81
socialSupport_appreciation		
	0.66	0.70
empowerment_knowledge		
	0.73	0.83
empowerment_desires		
	0.89	0.77

```
cortest.bartlett(cor(data), n = nrow(data))
```

```
$chisq
906.298897052485

$p.value
8.9262312299281e-94

$df
190
```

### 4. Uji asumsi multivariat normal

```
Shapiro-Wilk normality test
```

```
data: Z
W = 0.52999, p-value = 7.979e-14
```

## 5. Menentukan banyak faktor yang akan diekstrak

-Membuat vektor eigen dan nilai eigen

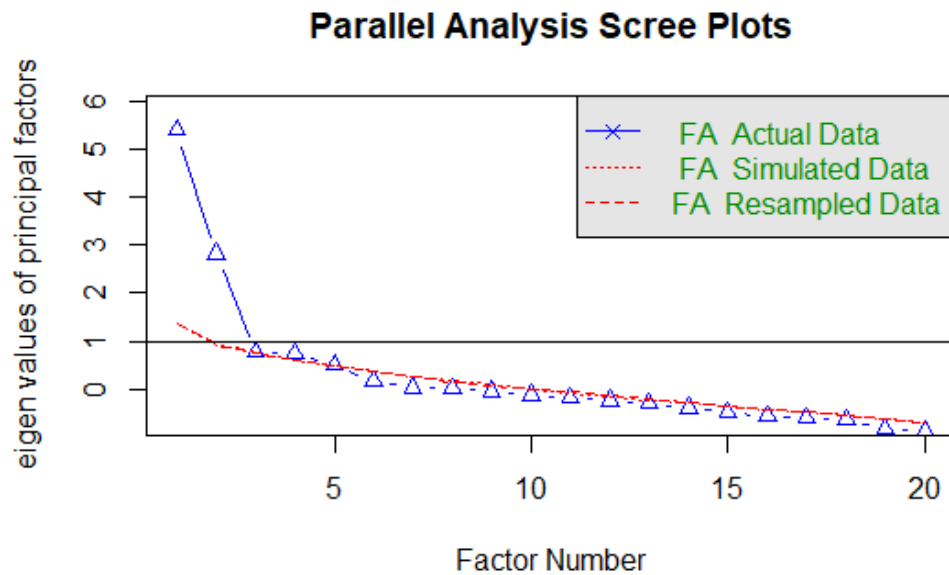
```
eigen() decomposition
$values
[1] 5.82760996 3.72985113 1.68856998 1.58194773 1.35086050 0.96858654
[7] 0.85529088 0.67690944 0.57373428 0.47785367 0.46743266 0.39707537
[13] 0.34970762 0.25270530 0.21596743 0.18918000 0.14216132 0.11137150
[19] 0.08486557 0.05831912

$vectors
      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
[1,] -0.11301219  0.057509606  0.257760730 -0.13339608 -0.44822351 -0.32590197
[2,]  0.01141402 -0.033673352 -0.539827421 -0.17192684  0.15545345 -0.54148578
[3,] -0.22302084  0.129804055 -0.281221411  0.27392194  0.28472330 -0.15435596
[4,] -0.13312680  0.058639287 -0.291610106  0.51868282 -0.09530748 -0.20967219
[5,] -0.06593157  0.001985812 -0.324239717  0.07548856 -0.61021941  0.01212646
[6,]  0.02472924  0.140135506 -0.333141031 -0.23724896  0.23952477  0.37510308
[7,]  0.02208661 -0.079844587 -0.422088869 -0.42910813 -0.23873703  0.21747280
[8,] -0.02365649  0.411213693 -0.032071945  0.01350735  0.07556632 -0.14472810
[9,] -0.04679466  0.474547874  0.044646287 -0.09428475  0.02144053 -0.05555994
[10,] -0.09908459  0.416654592  0.081640195 -0.27335972  0.01685735 -0.08155021
[11,] -0.10458922  0.452027138  0.045351515 -0.13019426  0.07567065  0.06958465
[12,] -0.20510677  0.099751248 -0.144588997  0.32786641 -0.05717464  0.53189852
[13,] -0.33170184 -0.021040424  0.066014420  0.10478336  0.01555761 -0.07094390
[14,] -0.34041327 -0.117606650  0.096275323  0.03894020  0.16743205 -0.02736905
[15,] -0.31426968 -0.168064983  0.090662599 -0.03163284  0.20669985 -0.01241567
[16,] -0.28130587 -0.233696130  0.039997698 -0.23069522  0.11651473 -0.03292500
[17,] -0.35139044 -0.056451607 -0.138229434 -0.10567511 -0.02425945  0.08909436
[18,] -0.36691481 -0.049931923  0.041110566 -0.14128191  0.03246184  0.03565452
[19,] -0.33272687 -0.121927735  0.005574776 -0.21526115 -0.13221487 -0.05317001
```

-Mencari Variansi Kumulatif

```
0.581277301131497 · 0.819391562201695 · 0.868193890171126 · 0.911027696819254 · 0.942261395376884 · 0.958318944922884 · 0.970839687577457 · 0.978682350525409 ·
0.984316443406947 · 0.988224782781653 · 0.991964515112753 · 0.994663176015989 · 0.99675638502516 · 0.997849412624233 · 0.998647736158499 · 0.999260302082989 ·
0.99960621384133 · 0.999818514127001 · 0.999941786414589 · 1
```

-Membuat Parallel Analysis Scree Plot



## 6. Menjalankan EFA dengan rotasi faktor dan distribusi variabel

Factor analysis with Call: `fa(r = data, nfactors = 2, rotate = "Promax", fm = "pa")`

Test of the hypothesis that 2 factors are sufficient.

The degrees of freedom for the model is 151 and the objective function was 4.61

The number of observations was 72 with Chi Square = 286.43 with prob < 1.9e-10

The root mean square of the residuals (RMSA) is 0.09

The df corrected root mean square of the residuals is 0.1

Tucker Lewis Index of factoring reliability = 0.756

RMSEA index = 0.111 and the 10 % confidence intervals are 0.092 0.132

BIC = -359.35

With factor correlations of

PA1 PA2

PA1 1.00 0.15

PA2 0.15 1.00

socialSupport_emotionality	0.84	-0.10	0.6975	0.30	1.0
socialSupport_appreciation	0.80	-0.20	0.6342	0.37	1.1
socialSupport_instrumental	0.76	-0.32	0.6064	0.39	1.3
empowerment_knowledge	0.84	0.02	0.7053	0.29	1.0
empowerment_abilities	0.88	0.04	0.7935	0.21	1.0
empowerment_desires	0.82	-0.11	0.6630	0.34	1.0
ca_cervix	-0.50	-0.49	0.5618	0.44	2.0

	PA1	PA2
SS loadings	5.37	3.50
Proportion Var	0.27	0.17
Cumulative Var	0.27	0.44
Proportion Explained	0.61	0.39
Cumulative Proportion	0.61	1.00

## EFA\$loadings

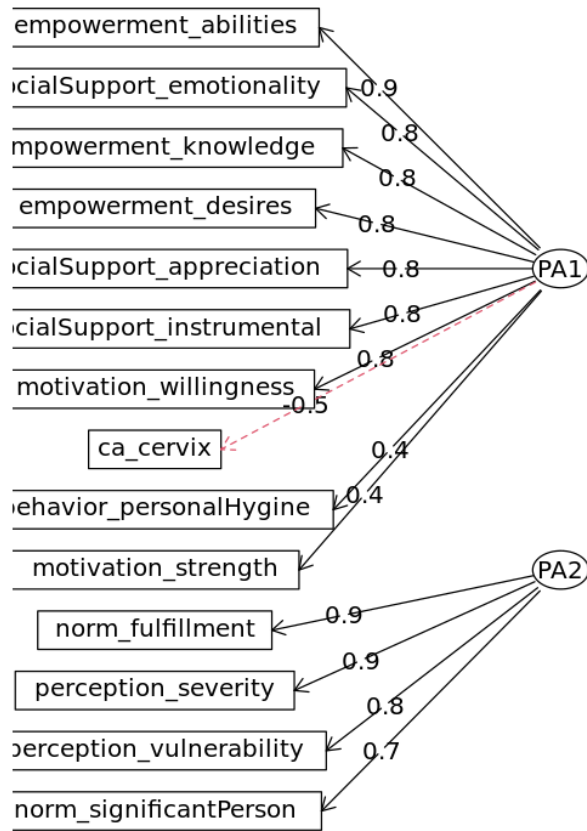
Loadings:

	PA1	PA2
behavior_sexualRisk	0.205	0.125
behavior_eating		
behavior_personalHygiene	0.411	0.277
intention_aggregation	0.242	0.122
intention_commitment	0.130	
attitude_consistency	-0.112	0.207
attitude_spontaneity		-0.124
norm_significantPerson	-0.158	0.733
norm_fulfillment	-0.157	0.944
perception_vulnerability		0.805
perception_severity		0.912
motivation_strength	0.382	0.218
motivation_willingness	0.755	
socialSupport_emotionality	0.844	-0.100
socialSupport_appreciation	0.802	-0.198
socialSupport_instrumental	0.759	-0.322
empowerment_knowledge	0.837	
empowerment_abilities	0.884	
empowerment_desires	0.823	-0.106
ca_cervix	-0.499	-0.489

	PA1	PA2
SS loadings	5.407	3.535
Proportion Var	0.270	0.177
Cumulative Var	0.270	0.447

7. Membuat diagram loading dan menghapus faktor yang tidak berpengaruh besar

### Factor Analysis



## Hasil Komputasi CFA

### 1. Menjalankan proses CFA dan mengeluarkan hasil dengan modul Lavaan

lavaan 0.6-10 ended normally after 64 iterations

Estimator	ML
Optimization method	NLMINB
Number of model parameters	26
Number of observations	72

Model Test User Model:

Test statistic	142.095
Degrees of freedom	52
P-value (Chi-square)	0.000

lavaan 0.6-10 ended normally after 64 iterations

Estimator	ML
Optimization method	NLMINB
Number of model parameters	26
Number of observations	72

Model Test User Model:

Test statistic	142.095
Degrees of freedom	52
P-value (Chi-square)	0.000

Model Test Baseline Model:

Test statistic	768.756
Degrees of freedom	66
P-value	0.000

User Model versus Baseline Model:

Comparative Fit Index (CFI)	0.872
Tucker-Lewis Index (TLI)	0.837

Loglikelihood and Information Criteria:

Loglikelihood user model (H0)	-1906.952
Loglikelihood unrestricted model (H1)	-1835.904



Akaike (AIC)	3865.903
Bayesian (BIC)	3925.096
Sample-size adjusted Bayesian (BIC)	3843.179

Root Mean Square Error of Approximation:

RMSEA	0.155
90 Percent confidence interval - lower	0.125
90 Percent confidence interval - upper	0.186
P-value RMSEA <= 0.05	0.000

Standardized Root Mean Square Residual:

SRMR	0.093
------	-------

Parameter Estimates:

Standard errors	Standard
Information	Expected
Information saturated (h1) model	Structured

Latent Variables:

	Estimate	Std.Err	z-value	P(> z )	Std.lv	Std.all
factor1 =~						
empowrmnt_blt	1.000				3.779	0.910
sclSpprt_mtnlt	0.933	0.092	10.104	0.000	3.525	0.836
empwrmnt_knwld	0.969	0.094	10.307	0.000	3.661	0.844
empowrmnt_dsrs	0.974	0.099	9.866	0.000	3.681	0.827
sclSpprt_pprct	0.579	0.069	8.404	0.000	2.189	0.761
sclSpprt_nstrm	0.850	0.104	8.177	0.000	3.212	0.749
mtvtn_wllngnss	0.780	0.102	7.616	0.000	2.947	0.719
ca_cervix	-0.061	0.011	-5.507	0.000	-0.229	-0.505
factor2 =~						
norm_fulflmnt	1.000				4.477	0.919
perceptn_svrty	0.703	0.053	13.194	0.000	3.148	0.932
prcptn_vlnrbt	0.818	0.074	11.023	0.000	3.664	0.863
nrm_sgnfcntPrs	0.284	0.039	7.240	0.000	1.270	0.693
ca_cervix	-0.050	0.009	-5.371	0.000	-0.222	-0.489

variances:

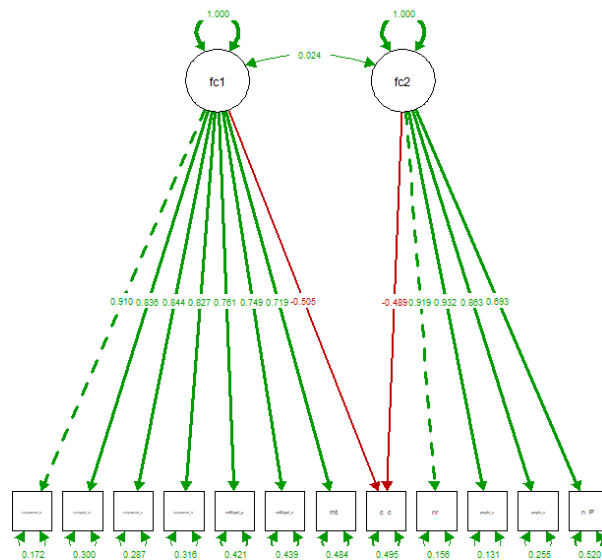
	Estimate	Std.Err	z-value	P(> z )	Std.lv	Std.all
factor1 ~~						
factor2	0.413	2.121	0.195	0.846	0.024	0.024

variances:

	Estimate	Std.Err	z-value	P(> z )	Std.lv	Std.all
.empowrmnt_blt	2.964	0.713	4.156	0.000	2.964	0.172
.sclSpprt_mtnlt	5.332	1.043	5.110	0.000	5.332	0.300
.empwrmnt_knwld	5.401	1.069	5.051	0.000	5.401	0.287
.empowrmnt_dsrs	6.260	1.210	5.174	0.000	6.260	0.316
.sclSpprt_pprct	3.485	0.636	5.479	0.000	3.485	0.421
.sclSpprt_nstrm	8.059	1.461	5.515	0.000	8.059	0.439
.mtvtn_wllngnss	8.136	1.454	5.597	0.000	8.136	0.484
.ca_cervix	0.102	0.018	5.639	0.000	0.102	0.495
.norm_fulflmnt	3.711	0.983	3.775	0.000	3.711	0.156
.perceptn_svrty	1.493	0.449	3.329	0.001	1.493	0.131

.prcptn_vlnrb1t	4.604	0.946	4.866	0.000	4.604	0.255
.nrm_sgnfcntPrs	1.747	0.309	5.655	0.000	1.747	0.520
factor1	14.282	2.878	4.963	0.000	1.000	1.000
factor2	20.039	3.984	5.030	0.000	1.000	1.000

2. Menampilkan path diagram beserta faktor skornya



## VII. Pembahasan

### EFA

Berdasarkan proses komputasi yang telah dilakukan sebelumnya, dapat dilihat bahwa hasil dari KMO test yang bernilai 0.73 mengindikasikan bahwa data yang dimiliki cukup baik untuk dilakukan analisis faktor. Selain itu hasil p-value dari Bartlett Test juga memberikan nilai yang kurang dari 0.05 juga mendukung pernyataan sebelumnya bahwa analisis faktor dapat memberikan hasil yang baik dari data yang dimiliki.

Estimasi parameter loadings dan communalities (komunitas) dapat dilakukan dengan metode likelihood maksimum dengan syarat terpenuhinya asumsi data yang berdistribusi multivariat normal dengan  $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ , oleh karena itu kami melakukan uji asumsi multivariat normal. Dari uji hipotesis didapatkan p-value yang kurang dari 0.05 yang membuat hipotesis ditolak. Artinya datanya tidak berdistribusi normal, sehingga kami menggunakan metode faktor utama (principal factor method atau principal axis method) untuk mengestimasi parameter loadings dan communalities (komunitas).

Saat kami melakukan proses EFA untuk melihat berapa faktor yang dapat diambil, kami memulainya dengan membuat parallel analysis scree plot. Dari sini, hasil yang diberikan adalah untuk mengambil dua faktor dari dua puluh variabel yang dimiliki pada data yang kami punya. Hal ini bisa dilihat dari adanya dua titik yang terletak di atas batas pada parallel analysis scree plot yang kami buat. Selanjutnya kami melakukan proses EFA dengan faktor dua dan rotate = “varimax”, ini kami lakukan karena kami ingin melihat bagaimana variabel yang dimiliki di data dibuat menjadi dua variabel.

Selain itu, rotasi kami pilih berdasarkan yang dapat memberikan hasil terbaik untuk loadings dari setiap variabel pada faktor yang ditentukan dari keseluruhan data, karena dengan menggunakan rotasi varimax yang termasuk dalam rotasi ortogonal untuk mendapatkan struktur faktor yang terdiri dari kombinasi variabel-variabel yang memiliki nilai faktor loading yang sangat tinggi hanya pada satu faktor. Dengan demikian jika terbentuk lebih dari satu faktor maka nilai variabel loading suatu variabel pada satu faktor akan sangat tinggi dan pada faktor lainnya mendekati nol. Oleh karenanya, hasil struktur faktor dengan menggunakan pendekatan rotasi varimax dapat dengan mudah dibedakan dengan struktur faktor yang lainnya (antar faktor terbentuk berdasarkan kombinasi variabel-variabel pembentuknya).

Kami membuat diagram dari loadings per variabel pada setiap faktor untuk melihat lebih jelas lagi variabel mana yang sebetulnya tidak memiliki pengaruh besar pada data kami. Setelah membuat diagram kami menemukan bahwa ternyata terdapat enam faktor yang tidak memiliki pengaruh besar pada analisis faktor, oleh karena itu kami menghapusnya dari dua faktor yang kami akan tentukan. Setelah menentukan faktor dan variabel pembentuknya maka proses EFA dapat dikatakan sudah selesai.

## CFA

Berdasarkan proses komputasi yang telah dilakukan sebelumnya, kami memulai proses CFA dengan mengambil model pemfaktoran yang telah kami tentukan di proses EFA di atas. Yang pertama dilakukan tentunya adalah membuat permodelan dua faktor dengan variabel pembentuknya yang sudah ditetapkan dari proses EFA dan diagram loadings sebelumnya.

Kami melanjutkannya dengan menjalankan proses CFA pada R guna melihat ringkasan dan hasil dari proses CFA tersebut pada permodelan data dua faktor dari data yang kita miliki. Kami membuat hipotesis awal bahwa permodelan dua faktor dapat memberikan hasil yang baik pada analisis faktor dari data yang kami miliki. Setelah menjalankannya, kami menemukan bahwa nilai p-value pada uji goodness of fit chisquare bernilai lebih kecil dari 0.05. Hal ini mengindikasikan bahwa terdapat cukup bukti untuk mengatakan bahwa model yang kami bentuk kurang tepat.

Selain itu bila kita melihat nilai CFI dan RMSEA dari proses CFA yang dilakukan, nilai CFI yang kami dapat adalah 0.872 sedangkan permodelan dapat dikatakan baik bila nilai CFI di atas 0.9, lalu nilai RMSEA yang kami dapat adalah 0.155 sedangkan di sisi lain permodelan dapat dikatakan baik jika nilai RMSEA di bawah 0.06. Kedua hal ini mendukung hasil dari uji goodness of fit chisquare sebelumnya, walaupun pada tes SRMR nilai yang kami dapat yaitu 0.093 sudah dapat dikatakan baik oleh beberapa sumber yang mengatakan bahwa permodelan benar jika nilai SRMR di bawah 0.1.

Oleh karena itu, kami dapat menyimpulkan bahwa proses CFA mengindikasikan bahwa permodelan dua faktor dengan variabel yang telah ditentukan kurang tepat untuk melakukan analisis faktor pada data yang kami miliki. Penyebab dari kesalahan permodelan ini dapat disebabkan oleh banyak hal, seperti proses EFA yang kurang sempurna atau adanya hasil dari proses EFA yang lebih baik dari yang kami lakukan di atas.

14. 3 Consider the 1-factor confirmatory factor analysis model in (14.2) with

$$\Lambda = \begin{bmatrix} 1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix}, \quad \text{var}(f) = \phi, \quad \text{cov}(\epsilon) = \begin{bmatrix} \psi_{11} & 0 & 0 \\ 0 & \psi_{22} & 0 \\ 0 & 0 & \psi_{33} \end{bmatrix}$$

(a) Use the rule presented in Section 14.2.2 to determine whether or not model is identified. Explain your reasoning.

- Diketahui terdapat 3 variabel sehingga  $p = 3$
- Diketahui terdapat 1 faktor sehingga  $m = 1$

Berdasarkan order condition, model akan identified ketika setidaknya jumlah statistik unik di matriks kovariansi sampel sebesar jumlah parameter yang perlu dibutuhkan untuk menentukan matriks kovariansi populasi di bawah hipotesis.

Model pada soal ini seperti pada model pada persamaan (14.7) yaitu :

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} + \begin{pmatrix} 1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} f + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{pmatrix}$$

Sehingga jumlah statistik uniknya yaitu,

$$\frac{p(p+1)}{2} = \frac{3(3+1)}{2} = 6.$$

Parameter estimasi yang dibutuhkan yaitu,

- Estimasi factor loadings  $(p-m)m$
- Estimasi variansi dan kovariansi faktor  $m(m+1)/2$
- Estimasi variansi error  $p$

Sehingga sebesar

$$(p-m)m + \frac{m(m+1)}{2} + p = (3-1)1 + \frac{1(1+1)}{2} + 3 = 2 + 1 + 3 = 6$$

Sehingga didapatkan

$$\frac{p(p+1)}{2} = (p-m)m + \frac{m(m+1)}{2} + p = 6$$

Sehingga dapat diketahui model identified atau lebih tepatnya model just-identified karena statistik unik dari matriks kovariansi sampel sama dengan parameter yang dibutuhkan untuk mengestimasi matriks kovariansi populasi.

(b) How many degrees of freedom will be associated with the  $\chi^2$  goodness-of-fit statistic described in (14.13)? What does this imply about the nature of the estimate  $\hat{\boldsymbol{\theta}}$ —that is, how close will  $\Sigma(\hat{\boldsymbol{\theta}})$  be to the sample covariance matrix for the data  $\mathbf{S}$ ?

Misalkan  $p$  adalah jumlah variabel sehingga  $p = 3$

Misalkan  $q$  adalah jumlah parameter yang dibutuhkan untuk mengestimasi matriks kovariansi populasi sehingga dari soal sebelumnya didapatkan  $q = 6$

Diketahui derajat bebas yang dibutuhkan sebesar

$$\frac{p(p+1)}{2} - q = \frac{3(3+1)}{2} - 6 = 6 - 6 = 0$$

Sehingga derajat bebas yang digunakan  $\chi^2$  goodness-of-fit statistic sebesar 0.

Dengan derajat sebesar 0 (model just-identified) akan dihasilkan satu solusi unik, namun tidak dapat ditentukan apakah estimasi  $\hat{\boldsymbol{\theta}}$  akan dekat dengan  $\Sigma(\hat{\boldsymbol{\theta}})$ . Ketika derajat bebas-nya sebesar 0, maka uji hipotesis tidak terlalu berguna dikarenakan parameternya akan tetap (fixed) sehingga mau bagaimanapun bentuk datanya, akan dihasilkan hasil yang sempurna.

C. Find the model covariance matrix  $\Sigma(\theta)$ , with each element written as a function of  $\lambda_2, \lambda_3, \phi, \psi_{11}, \psi_{22}$ , dan  $\psi_{33}$

Jawab:

↳ Matriks kovariansi untuk populasi yaitu

$$\Sigma(\theta) = \Lambda \Phi \Lambda' + \Psi$$

maka,

$$\Sigma(\theta) = \begin{bmatrix} 1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} \phi \begin{bmatrix} 1 & \lambda_2 & \lambda_3 \end{bmatrix} + \text{cov}(\epsilon)$$

$$= \begin{bmatrix} \phi \\ \phi \lambda_2 \\ \phi \lambda_3 \end{bmatrix} \begin{bmatrix} 1 & \lambda_2 & \lambda_3 \end{bmatrix} + \begin{bmatrix} \psi_{11} & 0 & 0 \\ 0 & \psi_{22} & 0 \\ 0 & 0 & \psi_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \phi & \phi \lambda_2 & \phi \lambda_3 \\ \phi \lambda_2 & \phi \lambda_2^2 & \phi \lambda_2 \lambda_3 \\ \phi \lambda_3 & \phi \lambda_2 \lambda_3 & \phi \lambda_3^2 \end{bmatrix} + \begin{bmatrix} \psi_{11} & 0 & 0 \\ 0 & \psi_{22} & 0 \\ 0 & 0 & \psi_{33} \end{bmatrix}$$

$$= \begin{bmatrix} \phi + \psi_{11} & \phi \lambda_2 & \phi \lambda_3 \\ \phi \lambda_2 & \phi \lambda_2^2 + \psi_{22} & \phi \lambda_2 \lambda_3 \\ \phi \lambda_3 & \phi \lambda_2 \lambda_3 & \phi \lambda_3^2 + \psi_{33} \end{bmatrix}$$

maka, tiap elemen  $\Sigma(\theta)$  dalam fungsi  $\lambda_2, \lambda_3, \phi, \psi_{11}, \psi_{22}$ , dan  $\psi_{33}$  adalah

$$\begin{aligned} \rightarrow \text{Var}(y_1) &= \phi + \psi_{11} & \rightarrow \text{cov}(y_1, y_2) &= \phi \lambda_2 \\ \rightarrow \text{Var}(y_2) &= \phi \lambda_2^2 + \psi_{22} & \rightarrow \text{cov}(y_1, y_3) &= \phi \lambda_3 \\ \rightarrow \text{Var}(y_3) &= \phi \lambda_3^2 + \psi_{33} & \rightarrow \text{cov}(y_2, y_3) &= \phi \lambda_2 \lambda_3 \end{aligned}$$

d. Let the sample covariance matrix for the data be denoted

$$S = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix}$$

Set the unique elements of  $S$  equal to the corresponding elements of  $\Sigma(\theta)$ .

With six equations and six parameters to estimate, find estimates of  $\lambda_2, \lambda_3, \phi, \psi_{11}, \psi_{22}$ , and  $\psi_{33}$  as functions of sample variances and covariances ( $s_{11}, s_{12}, s_{13}, s_{22}, s_{23}$  and  $s_{33}$ )



Jawab :

$$L_0 S = \Sigma(\theta)$$

$$\begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} = \begin{bmatrix} \phi + \psi_{11} & \phi \lambda_2 & \phi \lambda_3 \\ \phi \lambda_2 & \phi \lambda_2^2 + \psi_{22} & \phi \lambda_2 \lambda_3 \\ \phi \lambda_3 & \phi \lambda_2 \lambda_3 & \phi \lambda_3^2 + \psi_{33} \end{bmatrix}$$

maka, didapat

$$\begin{aligned} \rightarrow S_{11} &= \phi + \psi_{11} & \rightarrow S_{12} &= \phi \lambda_2 & \rightarrow S_{13} &= \phi \lambda_3 \\ \rightarrow S_{21} &= \phi \lambda_2 & \rightarrow S_{22} &= \phi \lambda_2^2 + \psi_{22} & \rightarrow S_{23} &= \phi \lambda_2 \lambda_3 \\ \rightarrow S_{31} &= \phi \lambda_3 & \rightarrow S_{32} &= \phi \lambda_2 \lambda_3 & \rightarrow S_{33} &= \phi \lambda_3^2 + \psi_{33} \end{aligned}$$

Perhatikan bahwa :

$$\begin{aligned} \rightarrow \frac{S_{23}}{S_{12}} &= \frac{\phi \lambda_2 \lambda_3}{\phi \lambda_2} & \rightarrow \frac{S_{23}}{S_{13}} &= \frac{\phi \lambda_2 \lambda_3}{\phi \lambda_3} & \rightarrow \frac{S_{12} \cdot S_{13}}{S_{23}} &= \frac{\phi \lambda_2 \cdot \phi \lambda_3}{\phi \lambda_2 \lambda_3} \\ &= \lambda_3 & &= \lambda_2 & &= \phi \end{aligned}$$

$$\begin{aligned} \rightarrow S_{11} &= \phi + \psi_{11} & \rightarrow S_{22} &= \phi \lambda_2^2 + \psi_{22} & \rightarrow S_{33} &= \phi \lambda_3^2 + \psi_{33} \\ \psi_{11} &= S_{11} - \phi & \psi_{22} &= S_{22} - \phi \lambda_2^2 & \psi_{33} &= S_{33} - \phi \lambda_3^2 \\ &= S_{11} - \frac{S_{12} \cdot S_{13}}{S_{23}} & &= S_{22} - \frac{S_{12} \cdot S_{13}}{S_{23}} \cdot \left( \frac{S_{23}}{S_{13}} \right)^2 & &= S_{33} - \frac{S_{12} \cdot S_{13}}{S_{23}} \cdot \left( \frac{S_{23}}{S_{12}} \right)^2 \\ & & &= S_{22} - \frac{S_{12} \cdot S_{23}}{S_{13}} & &= S_{33} - \frac{S_{13} \cdot S_{23}}{S_{12}} \end{aligned}$$

Jadi, diperoleh bahwa

$$\begin{aligned} \rightarrow \lambda_2 &= \frac{S_{23}}{S_{13}} & \rightarrow \psi_{11} &= S_{11} - \frac{S_{12} \cdot S_{13}}{S_{23}} \\ \rightarrow \lambda_3 &= \frac{S_{23}}{S_{12}} & \rightarrow \psi_{22} &= S_{22} - \frac{S_{12} \cdot S_{23}}{S_{13}} \\ \rightarrow \phi &= \frac{S_{12} \cdot S_{13}}{S_{23}} & \rightarrow \psi_{33} &= S_{33} - \frac{S_{13} \cdot S_{23}}{S_{12}} \end{aligned}$$

**14.5** Use the Seishu data of Table 7.1. Conduct a confirmatory factor analysis of the covariance matrix for all 10 variables. Use maximum likelihood to fit the model and test the hypothesis that the observations are driven by four factors:

$f_1$  = “flavor”

$f_2$  = “acidity”

$f_3$  = “sweetness”

$f_4$  = “alcohol content”

To fit an identifiable model, define the observed variable “Taste” to be equal to  $f_1$  plus error, define the observed variable “pH” to be equal to  $f_2$  plus error, define the observed variable “Total Sugar” to be equal to  $f_3$  plus error, and define the observed variable “Alcohol” to be equal to  $f_4$  plus error. In your initial model, allow the other 6 variables to be functions of all 4 factors, for a total of 24 factor loadings to be estimated.

(a) Assess goodness of fit with the criteria discussed in Section 14.3.3.

(b) Using goodness-of-fit statistics and hypothesis tests on each of the factor loadings, conduct an exploratory analysis using the iterative model-selection process discussed at the end of Section 14.5 to identify a simpler model. Specifically, determine whether a model can be specified with fewer factor loadings (i.e., simpler structure) and still exhibit a good fit. Propose a simpler model and appropriately justify your proposal using goodness-of-fit statistics and hypothesis tests.

## JAWABAN

**Tabel**

$y_1$	$y_2$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$	$x_8$
1.0	.8	4.05	1.68	.85	3.0	3.97	5.00	16.90	122.0
.1	.2	3.81	1.39	.30	.6	3.62	4.52	15.80	62.0
.5	.0	4.20	1.63	.92	-2.3	3.48	4.46	15.80	139.0
.7	.7	4.35	1.43	.97	-1.6	3.45	3.98	15.40	150.0
-.1	-1.1	4.35	1.53	.87	-2.0	3.67	4.22	15.40	138.0
.4	.5	4.05	1.84	.95	-2.5	3.61	5.00	16.78	123.0
.2	-.3	4.20	1.61	1.09	-1.7	3.25	4.15	15.81	172.0
.3	-.1	4.32	1.43	.93	-5.0	4.16	5.45	16.78	144.0
.7	.4	4.21	1.74	.95	-1.5	3.40	4.25	16.62	153.0
.5	-.1	4.17	1.72	.92	-1.2	3.62	4.31	16.70	121.0
-.1	.1	4.45	1.78	1.19	-2.0	3.09	3.92	16.50	176.0
.5	-.5	4.45	1.48	.86	-2.0	3.32	4.09	15.40	128.0
.5	.8	4.25	1.53	.83	-3.0	3.48	4.54	15.55	126.0
.6	.2	4.25	1.49	.86	2.0	3.13	3.45	15.60	128.0
.0	-.5	4.05	1.48	.30	.0	3.67	4.52	15.38	99.0
-.2	-.2	4.22	1.64	.90	-2.2	3.59	4.49	16.37	122.8
.0	-.2	4.10	1.55	.85	1.8	3.02	3.62	15.31	114.0
.2	.2	4.28	1.52	.75	-4.8	3.64	4.93	15.77	125.0
-.1	-.2	4.32	1.54	.83	-2.0	3.17	4.62	16.60	119.0
.6	.1	4.12	1.68	.84	-2.1	3.72	4.83	16.93	111.0
.8	.5	4.30	1.50	.92	-1.5	2.98	3.92	15.10	68.0
.5	.2	4.55	1.50	1.14	.9	2.60	3.45	15.70	197.0
.4	.7	4.15	1.62	.78	-7.0	4.11	5.55	15.50	106.0
.6	-.3	4.15	1.32	.31	.8	3.56	4.42	15.40	49.5
-.7	-.3	4.25	1.77	1.12	.5	2.84	4.15	16.65	164.0
-.2	.0	3.95	1.36	.25	1.0	3.67	4.52	15.98	29.5
.3	-.1	4.35	1.42	.96	-2.5	3.40	4.12	15.30	131.0
.1	.4	4.15	1.17	1.06	-4.5	3.89	5.00	16.79	168.2
.4	.5	4.16	1.61	.91	-2.1	3.93	4.35	15.70	118.0
-.6	-.3	3.85	1.32	.30	.7	3.61	4.29	15.71	48.0

a.

Akan dilakukan proses CFA dengan menggunakan kriteria yang sudah ditentukan di soal

**Hipotesis,**

$H_0 =$  Observasi dapat dibentuk dengan menggunakan empat faktor

$H_1 =$  Observasi tidak dapat dibentuk dengan menggunakan empat faktor

**Model yang dibentuk,**

$$factor\ 1 = Y_1 + Y_2 + X_2 + X_3 + X_4 + X_5 + X_8$$

$$factor\ 2 = X_1 + Y_2 + X_2 + X_3 + X_4 + X_5 + X_8$$

$$factor\ 3 = X_6 + Y_2 + X_2 + X_3 + X_4 + X_5 + X_8$$

$$factor\ 4 = X_7 + Y_2 + X_2 + X_3 + X_4 + X_5 + X_8$$

Dengan metode pada 14.3.3 akan dihitung CFI, RMSEA, dan SRMR

**Uji goodness of fit chisquare (diterima p-val>0,05)**

$$\chi^2_{df} = 0,178$$

Karena p-value > 0,05 maka terdapat cukup bukti untuk menerima  $H_0$

**CFI (model baik jika CFI>0,9)**

$$CFI = 1 - \frac{\max(\chi^2_M - df_M, 0)}{\max(\chi^2_B - df_B, \chi^2_M - df_M, 0)}$$

$$CFI = 0,976$$

Nilai CFI > 0,9 di mana ini merupakan batas bawah untuk menentukan sebuah model baik atau tidak sehingga dengan CFI dapat dikatakan model baik.

**RMSEA (model baik jika RMSEA<0,06)**

$$RMSEA = \sqrt{\frac{\max(\chi^2_M - df_M, 0)}{(n - 1)df_M}}$$

$$RMSEA = 0,112$$

Nilai RMSEA yang didapat > 0,06 sehingga dapat dikatakan bahwa model tidak cukup baik.

**SRMR (model baik jika SRMR<0,1)**

$$SRMR = \sqrt{\frac{2}{p(p+1)} \sum_{i=1}^p \sum_{j=1}^i \frac{(s_{ij} - \sigma_{ij})^2}{S_{ii}S_{jj}}}$$

$$SRMR = 0,072$$

Nilai SRMR yang didapat < 0,1 sehingga dapat dikatakan bahwa model sudah cukup baik.

### **Coding R pengerjaan CFA**

```
packages <- c("Hmisc", "matlib",
"Matrix", "expm", "matrixcalc", "ellipsis", "Hotelling", "dplyr", "psych", "RcmdrMisc", "Rcsdp", "mvnrmtest", "lavaan")

if ( length(missing_pkgs <- setdiff(packages,
rownames(installed.packages()))) > 0) {

  message("Installing missing package(s): ", paste(missing_pkgs, collapse =
", "))

  install.packages(missing_pkgs)

}

lapply(packages, library, character.only = TRUE)

hs.mod <- '

factor1 =~ y1 + y2 + x2 + x3 + x4 + x5 + x8
factor2 =~ x1 + y2 + x2 + x3 + x4 + x5 + x8
factor3 =~ x6 + y2 + x2 + x3 + x4 + x5 + x8
factor4 =~ x7 + y2 + x2 + x3 + x4 + x5 + x8
'

hs.fit <- cfa(hs.mod, data = data)

hs.fit

summary(hs.fit, fit.measures = TRUE, standardized = TRUE)
```

### **Kesimpulan**

Dari hasil perhitungan goodness of fit 14.3.3 di atas, dapat dikatakan bahwa model yang dibentuk sudah cukup baik dan terdapat cukup bukti untuk menerima  $H_0$  yang mengatakan

bahwa observasi dapat dibentuk menggunakan empat faktor yaitu “flavor”, “acidity”, “sweetness”, dan “alcohol content”.

b.

Untuk mencari model yang lebih sederhana, proses EFA bisa digunakan untuk melihat berapa faktor yang terbaik untuk dibuat dalam permodelan.

- Melakukan uji KMO dan Bartlett test guna melihat apakah analisis faktor dari data dapat memberikan hasil yang tepat dengan menggunakan R

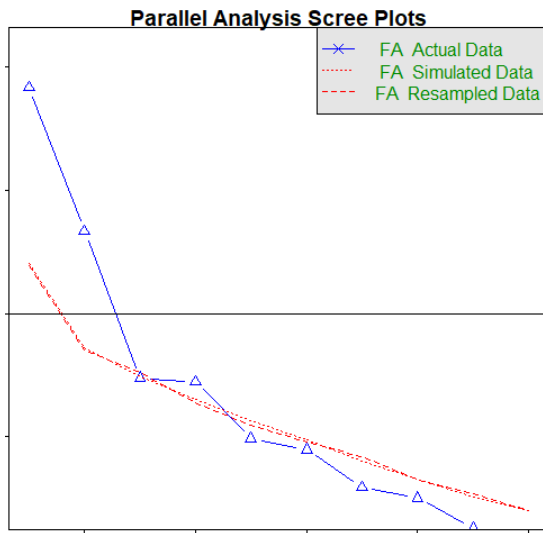
```
KMO(data)
cortest.bartlett(data)
```

didapatkan hasil dari uji KMO adalah 0,55 dan p-value dari bartlett test adalah  $1.157322e-114$ , ini menandakan bahwa proses analisis faktor masih dapat memberikan hasil yang cukup baik sebagaimana nilai KMO  $>0,5$  dan p-value dari uji Bartlett masih  $< 0,05$

- Dilanjutkan dengan melakukan proses EFA dan juga membuat screeplot untuk melihat berapa jumlah faktor yang terbaik untuk dilakukan permodelan pada data ini dengan menggunakan R

```
hs.efa <- fa(data, nfactors = 10,
             rotate = "none")
hs.efa

plot(hs.efa$e.values)
fa.parallel(data, fa = "fa", fm = "pa")
```



Proses ini mengindikasikan bahwa permodelan yang paling baik adalah dengan membuatnya menjadi dua faktor

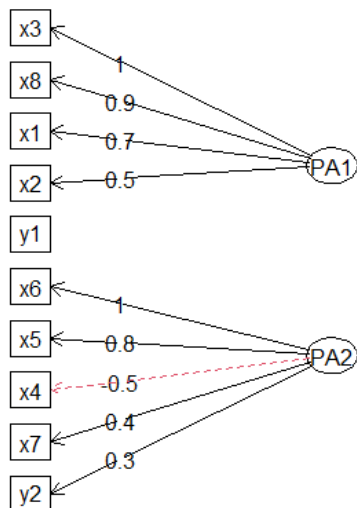
- Setelah mengetahui banyaknya faktor, kita harus mengetahui variabel mana yang akan dimasukan ke setiap faktornya dengan cara mencari loadings dari setiap variabelnya dan membuat diagram dari loads setelah memeriksa proses EFA dengan dua faktor

```
hs.efa2 <- fa(data, nfactors = 2, rotate = 'oblimin', fm = 'pa')
hs.efa2
```

```
hs.efa2$loadings
loads <- hs.efa2$loadings
```

```
fa.diagram(loads)
```

### Factor Analysis



Setelah menentukan jumlah faktor dan menentukan variabel yang akan digunakan, selanjutnya akan dilakukan pengujian hipotesis dengan goodness of fit test menggunakan CFA pada permodelan yang baru.

### **Hipotesis,**

$H_0 = \text{Observasi dapat dibentuk dengan menggunakan dua faktor}$

$H_1 = \text{Observasi tidak dapat dibentuk dengan menggunakan dua faktor}$

### **Model yang dibentuk,**

$$\text{factor 1} = X_3 + X_8 + X_1 + X_2$$

$$\text{factor 2} = X_6 + X_5$$

Ini dilakukan dengan cara lihat diagram loads dan memilih variabel yang memiliki load di atas 0,4 dan memiliki garis lurus.

Dengan metode pada 14.3.3 akan dihitung CFI, RMSEA, dan SRMR

### **Uji goodness of fit chisquare (diterima p-val>0,05)**

$$\chi^2_{df} = 0,408$$

Karena p-value > 0,05 maka terdapat cukup bukti untuk menerima  $H_0$

### **CFI (model baik jika CFI>0,9)**

$$CFI = 1 - \frac{\max(\chi^2_M - df_M, 0)}{\max(\chi^2_B - df_B, \chi^2_M - df_M, 0)}$$

$$CFI = 0,997$$

Nilai CFI > 0,9 di mana ini merupakan batas bawah untuk menentukan sebuah model baik atau tidak sehingga dengan CFI dapat dikatakan model baik.

### **RMSEA (model baik jika RMSEA<0,06)**

$$RMSEA = \sqrt{\frac{\max(\chi_M^2 - df_M, 0)}{(n-1)df_M}}$$

$$RMSEA = 0,03$$

Nilai RMSEA yang didapat < 0,06 sehingga dapat dikatakan bahwa model cukup baik.

**SRMR (model baik jika SRMR<0,1)**

$$SRMR = \sqrt{\frac{2}{p(p+1)} \sum_{i=1}^p \sum_{j=1}^i \frac{(s_{ij} - \sigma_{ij})^2}{S_{ii}S_{jj}}}$$

$$SRMR = 0,075$$

Nilai SRMR yang didapat < 0,1 sehingga dapat dikatakan bahwa model sudah cukup baik.

## Coding R pengerjaan CFA

```
packages <- c("Hmisc", "matlib",
"Matrix", "expm", "matrixcalc", "ellipsis", "Hotelling", "dplyr", "psych", "RcmdrMisc", "Rcsdp",
"mvnrmtest", "lavaan")

if ( length(missing_pkgs <- setdiff(packages, rownames(installed.packages()))) > 0) {
  message("Installing missing package(s): ", paste(missing_pkgs, collapse = ", "))
  install.packages(missing_pkgs)
}

lapply(packages, library, character.only = TRUE)

hs.mod <- '
factor1 =~ x3 + x8 + x1 +x2
factor2 =~ x6 + x5
'

hs.fit <- cfa(hs.mod, data = data)

hs.fit

summary(hs.fit, fit.measures = TRUE, standardized = TRUE)
```

## Kesimpulan



Dengan menggunakan EFA untuk menentukan jumlah faktor dan jenis variabel yang digunakan dalam permodelan, jika kita melihat hasil dari ketiga metode pada 14.3.3 di atas, maka terdapat cukup bukti untuk mengatakan  $H_0$  benar dan observasi bisa dilakukan dengan menggunakan dua faktor