

SEN4016

MULTIVARIATE DATA ANALYSIS

Week # 5

EXPLORATORY FACTOR ANALYSIS

What is Exploratory Factor Analysis (EFA)?

- Exploratory Factor Analysis is a **statistical technique** used to **identify hidden (latent) factors** that explain the **patterns of correlations** among a set of observed variables.
- The idea is that many measured variables may actually reflect a **smaller number of underlying traits**.

Intuitive Example

Orange, Motorcycle, Bus,

Apple, Banana, Car

Anything in Common?

Intuitive Example

- Let's group them:

Group 1: Orange, Apple, Banana

Group 2: Motorcycle, Bus, Car

Intuitive Example

- Correlation Matrix

Items	1	2	3	4	5	6
1. Orange	1.00					
2. Apple	.67	1.00				
3. Banana	.70	.81	1.00			
4. Motorcycle	.11	.08	.05	1.00		
5. Bus	.08	.12	.09	.75	1.00	
6. Car	.18	.12	.22	.89	.83	1.00

Intuitive Example

- Name the groups

Fruits	Vehicles
Orange	Motorcycle
Apple	Bus
Banana	Car

- EFA is about **identifying hidden (latent) factors** that explain the **patterns of correlations** among a set of observed variables.
- The aim of the EFA is to divide the variables into groups to separate those variables that **correlate highly** from those that correlate less strongly.

How EFA works?

1. Determine the Suitability of Data for Factor Analysis
2. Factor Extraction
3. Factor Rotation
4. Interpret and Label the Factors

1. Determine the Suitability of Data for Factor Analysis

- **Kaiser Meyer Olkin (KMO) Measure:**

Verify the sampling adequacy. A value greater than 0.6 is generally considered acceptable.

- **Bartlett's Test:** Check the significance level to determine if the correlation matrix is suitable for factor analysis.

Kaiser–Meyer–Olkin (KMO) Measure

- Factor analysis works only if some variables are correlated with each other — but not too highly (to avoid multicollinearity).
- The KMO measure compares simple and partial correlations
 - If partial correlations are small, it means variables share common factors — good for EFA.
 - If partial correlations are large, it means variables are mostly unrelated — bad for EFA.

Interpreting Kaiser–Meyer–Olkin (KMO) Values

KMO Value	Interpretation
0.90 – 1.00	Excellent
0.80 – 0.89	Very good
0.70 – 0.79	Good
0.60 – 0.69	Acceptable
0.50 – 0.59	Poor
< 0.50	Inadequate for factor analysis

Bartlett's Test

- **Bartlett's Test** checks whether your correlation matrix is significantly different from an identity matrix.
- If the correlation matrix is close to an **identity matrix**, there is **no common variance** to explain — so factor analysis is not possible.

Bartlett's Test

- Bartlett Test is useful to assess the hypothesis that the sample came from a population in which the variables are uncorrelated:

H_0 : The variables are uncorrelated in the population.

H_1 : The variables are correlated in the population.

P-value < 0.05 indicates presence of correlations.

2. Factor Extraction

Factor extraction is the **process of identifying the underlying factors** that explain the **patterns of correlations** among a set of observed variables.

Factor extraction finds

- The **number** of factors
- The **factor loadings** (how strongly each variable relates to each factor)
- The **amount of variance** explained by each factor

The Factor Model

The basic model for EFA is:

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \dots + a_{im}F_m + \epsilon_i$$

Where:

- X_i : observed variable
- F_m : extracted common factors
- a_{im} : loading of variable i on factor m
- ϵ_i : unique variance (specific + error variance)

2. Factor Extraction

Steps of Factor Extraction

- Choosing Factor Extraction Method
- Determine the Number of Factors

Choosing Extraction Method

- Principal Component Analysis, Principal Axis Factoring, Maximum likelihood are the most common methods.
- Determines how communalities are estimated and what variance is modeled.

Partitioning the Variance

Each variable's variance can be divided into:

- **Common variance (communality):** the proportion of variance in each observed variable that can be explained by the factors.
- **Unique variance:** specific to that variable
- **Error variance:** random noise

Extraction focuses on explaining as much *common variance* as possible.

- A good factor solution is one that explains the great share of the variance with the fewest factors
- Researchers are happy with 50 to 75% of the variance explained

What is Communality?

- Each variable has a **communality** – which indicates the proportion of the variable's variance explained by the extracted factors
- Communalities can range between
 - 0 (no variance explained)
 - 1 (all variance explained)

What is Communality?

- High communalities ($> .5$): Extracted factors explain most of the variance in the variable
- Low communalities ($< .5$): A variable has considerable variance unexplained by the extracted factors.

Factor Extraction Methods

Principal Component Analysis (PCA): Used when the main goal is data reduction.

- Analysis total variance
- Initial communalities = 1.0 (because total variance = 1).

Principal Axis Factoring (PAF): Used when the main goal is to identify underlying factors.

- Analysis shared variance
- Initial communality is often the Squared Multiple Correlation (SMC) of each variable with all others.

Determine the Number of Factors

- **Kaiser rule**

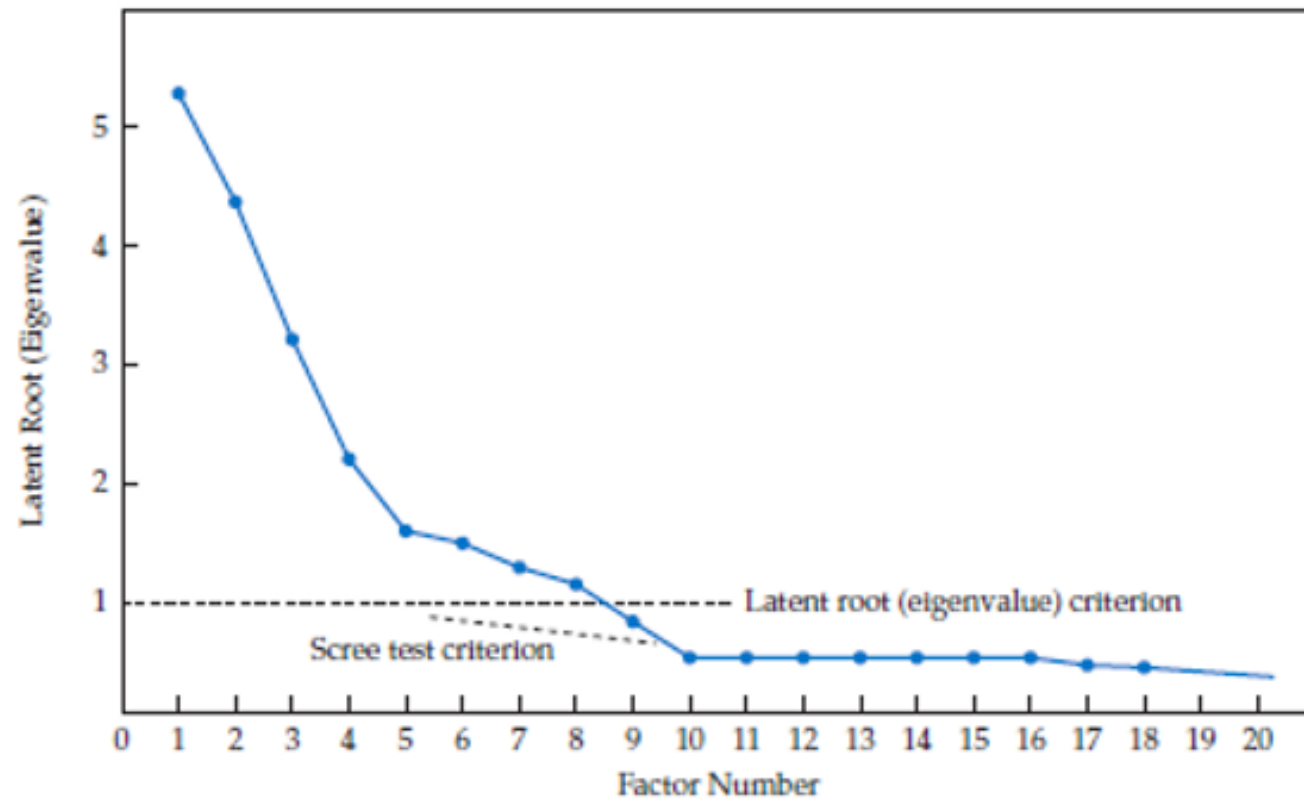
- most commonly used technique.
- factors having eigenvalues greater than 1 are considered significant.
- most applicable to PCA where the diagonal value representing the amount of variance for each variable is 1.0.
- less accurate with a small number of variables or lower communalities.

- **Percentage of Variance**

- Extract enough components to achieve a specified cumulative percentage of total variance extracted.

Determine the Number of Factors

- **Scree Plot**



Factor Extraction

Once the diagonal of your correlation matrix is set according to your chosen method:

- Perform **eigen-decomposition** (PCA or PAF).
- Extract the top m factors (based on eigenvalues > 1 , scree, etc.).
- Compute **factor loadings** and **updated communalities** (sum of squared loadings).

The Factor Model

The basic model for EFA is:

$$X_i = a_{i1}F_1 + a_{i2}F_2 + \dots + a_{im}F_m + \epsilon_i$$

Where:

- X_i : observed variable
- F_m : extracted common factors
- a_{im} : loading of variable i on factor m
- ϵ_i : unique variance (specific + error variance)

3. Factor Rotation

The reference axes of the factors are turned about the origin until some other position has been reached. Loadings of each variable remain fixed relative to other loadings.

- **Impact**

- The ultimate effect of rotating the factor matrix is to redistribute the variance from earlier factors to later ones to achieve a simpler, theoretically more meaningful factor pattern.

- **Alternative Methods**

- Orthogonal rotation – simplest approach which maintains orthogonality of factors.
 - Most common method – VARIMAX.
- Oblique rotation – allows for correlation among rotated factors.

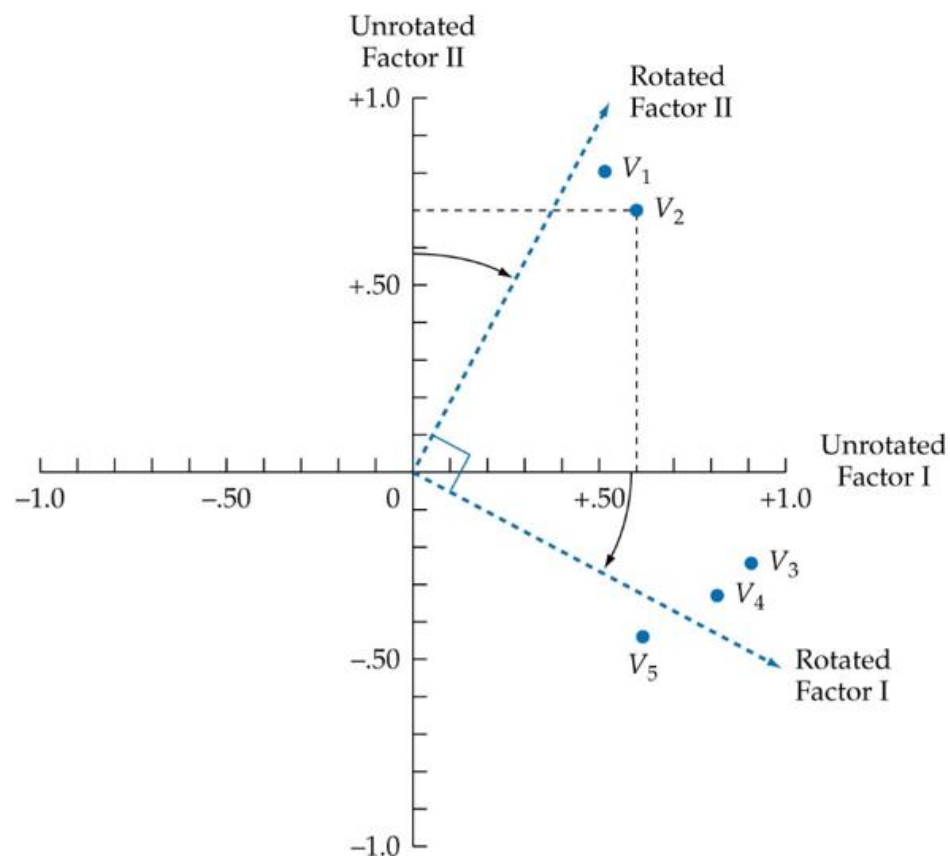
3. Factor Rotation

In this step, factors are rotated.

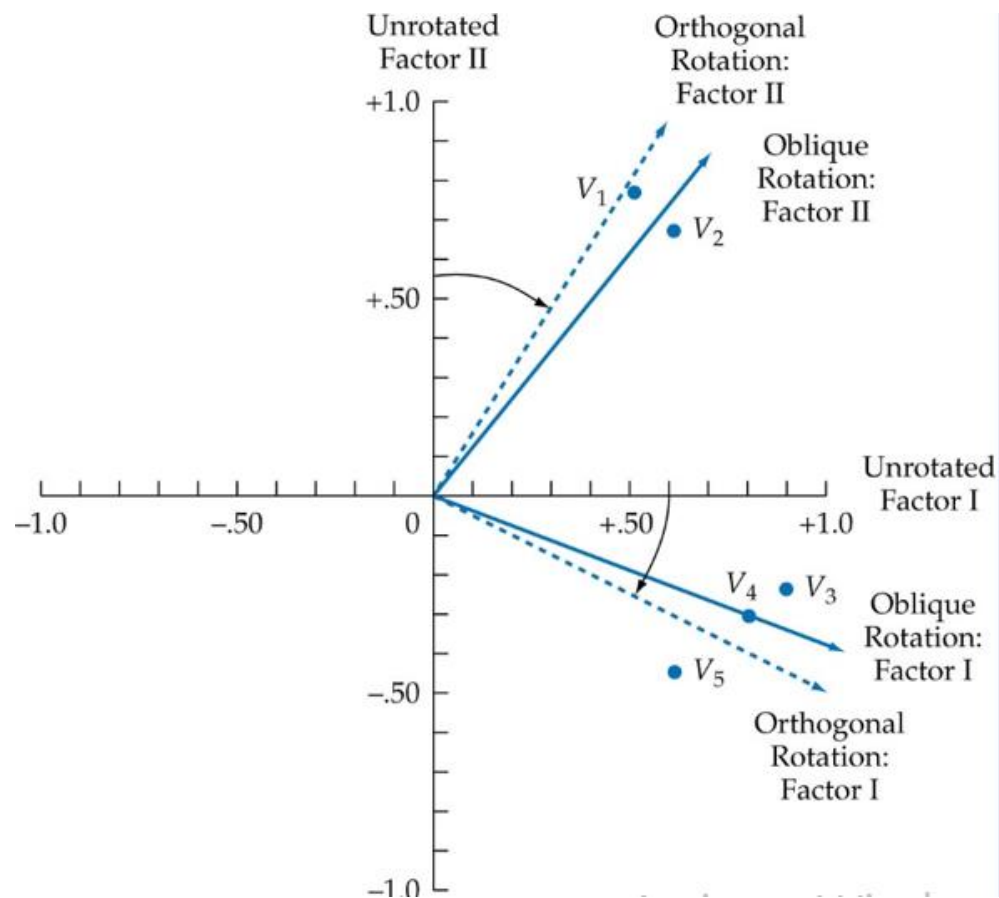
- Un-rotated factors are typically not very interpretable (most factors are correlated with many variables).
- Factors are rotated to make them more meaningful and easier to interpret (each variable is associated with a minimal number of factors).
- Different rotation methods may result in the identification of somewhat different factors.

Factor Rotation

- Orthogonal Rotation



- Oblique Rotation



4. Interpret and Label Factors

Identify Strong Loadings

- Look for loadings ≥ 0.50 → strong relationship between variable and factor
- Loadings between **0.30–0.49** → moderate; interpret cautiously
- Each variable should **load highly on only one factor**
- Avoid variables that load on **multiple factors** (cross-loadings)
- **Use Meaningful Names**
- Group variables with high loadings → interpret their common theme
- Label each factor with a **concise, conceptual name**

Numerical Example

20 two-wheeler users were surveyed about their perception and image attributes of vehicles they owned.

- 1. I use a two-wheeler because it is affordable.
- 2. It gives me sense of freedom to own a two-wheeler.
- 3. Low maintenance cost makes a two-wheeler very economical in the long run.
- 4. A two-wheeler is essentially a man's vehicle.
- 5. I feel very powerful when I am on my two-wheeler.
- 6. Some of my friends who don't have their own vehicle are jealous of me.
- 7. I feel good whenever I see the ad for my two-wheeler on TV, in a magazine or on a hording.
- 8. My vehicle gives me a comfortable ride.
- 9. I think two-wheelers are safe way to travel.
- 10. Three people should be legally allowed to travel on a two-wheeler

Data

S. NO.	QUESTION NO.									
	1	2	3	4	5	6	7	8	9	10
1	1	4	1	6	5	6	5	2	3	2
2	2	3	2	4	3	3	3	5	5	2
3	2	2	2	1	2	1	1	7	6	2
4	5	1	4	2	2	2	2	3	2	3
5	1	2	2	5	4	4	4	1	1	2
6	3	2	3	3	3	3	3	6	5	3
7	2	2	5	1	2	1	2	4	4	5
8	4	4	3	4	4	5	3	2	3	3
9	2	3	2	6	5	6	5	1	4	1
10	1	4	2	2	1	2	1	4	4	1
11	1	5	1	3	2	3	2	2	2	1
12	1	6	1	1	1	1	1	1	2	2
13	3	1	4	4	4	3	3	6	5	3
14	2	2	2	2	2	2	2	1	3	2
15	2	5	1	3	2	3	2	2	1	6
16	5	6	3	2	1	3	2	5	5	4
17	1	4	2	2	1	2	1	1	1	3
18	2	3	1	1	2	2	2	3	2	2
19	3	3	2	3	4	3	4	3	3	3
20	4	3	2	7	6	6	6	2	3	6

Suitability of Data

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.618
Bartlett's Test of Sphericity	Approx. Chi-Square	164.098
	df	45
	Sig.	.000

The KMO Measure of Sampling.

- High values (close to 1.0) generally indicate that a factor analysis may be useful with your data.
- If the value is less than 0.50, the results of the factor analysis probably won't be very useful.

Bartlett's test of sphericity

- Small values (less than 0.05) of the significance level indicate that a factor analysis may be useful with your data.

Results

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.883	38.828	38.828	3.883	38.828	38.828	3.841	38.409	38.409
2	2.777	27.770	66.598	2.777	27.770	66.598	2.429	24.294	62.703
3	1.375	13.747	80.346	1.375	13.747	80.346	1.764	17.643	80.346
4	.945	9.449	89.795						
5	.479	4.793	94.588						
6	.292	2.923	97.511						
7	.117	1.166	98.677						
8	.068	.680	99.356						
9	.037	.374	99.730						
10	.027	.270	100.000						

Extraction Method: Principal Component Analysis.

Results

Component Matrix^a

	Component		
	1	2	3
VAR00001	.176	.670	.493
VAR00002	-.136	-.608	.254
VAR00003	-.107	.820	.218
VAR00004	.966	-.036	-.097
VAR00005	.951	.166	-.136
VAR00006	.952	-.084	-.025
VAR00007	.971	.096	-.046
VAR00008	-.322	.775	-.308
VAR00009	-.069	.735	-.482
VAR00010	.161	.319	.814

Extraction Method: Principal Component Analysis.

a. 3 components extracted.

Rotated Component Matrix^a

	Component		
	1	2	3
VAR00001	.126	.313	.780
VAR00002	-.181	-.639	-.107
VAR00003	-.116	.604	.594
VAR00004	.970	-.064	-.006
VAR00005	.964	.131	.063
VAR00006	.945	-.140	.030
VAR00007	.971	.024	.106
VAR00008	-.262	.848	.101
VAR00009	.010	.881	-.044
VAR00010	.063	-.149	.874

Extraction Method: Principal Component Analysis.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Results

Communalities		
	Initial	Extraction
VAR00001	1.000	.722
VAR00002	1.000	.452
VAR00003	1.000	.731
VAR00004	1.000	.945
VAR00005	1.000	.950
VAR00006	1.000	.914
VAR00007	1.000	.955
VAR00008	1.000	.799
VAR00009	1.000	.777
VAR00010	1.000	.789
Extraction Method: Principal		

Initial = 1.000:

Because PCA starts from the *correlation matrix*, all variables are standardized (variance = 1).

So, each variable's total variance is 1.0 at the beginning.

Extraction:

After extracting the principal components (factors), each value shows the **proportion of variance explained by the retained factors**.

- e.g. VAR00004 = 0.945 means 94.5% of its variance is explained by the extracted components.
- VAR00002 = 0.452 means only 45.2% of its variance is explained — suggesting this variable doesn't fit the factor model well.

Results

Each **extraction value** equals the **sum of squared loadings** of that variable across the retained factors.

$$h_i^2 = (l_{i1})^2 + (l_{i2})^2 + (l_{i3})^2$$

If VAR00002 had loadings:

−0.136, −0.608, 0.254

Then its communality:

$$h_2^2 = (-0.136)^2 + (-0.608)^2 + (0.254)^2 = 0.018 + 0.370 + 0.065 = 0.453$$

Interpretation for VAR00002 (as an example)

Step	Observation	Interpretation
Communality = 0.452	Only 45.2% of its variance is captured by the extracted components	VAR00002 is not well represented by the factor solution — it may not relate strongly to the common underlying constructs.
Component Matrix loading = -0.608 on Factor 2	Largest absolute loading	VAR00002 contributes mainly (negatively) to Factor 2 , but still weak overall compared to others (since large variance unexplained).
Rotation	Simplifies pattern, aligns VAR00002 cleanly with Factor 2	Negative loading still means it measures the opposite side of the construct captured by Factor 2.

Interpretation for VAR2 (as an example)

Communality Range	Interpretation
> 0.80	Excellent representation by the extracted factors (e.g., VAR00004, VAR00005, VAR00007)
0.60–0.79	Acceptable representation
0.40–0.59	Weak — variable may not belong conceptually (e.g., VAR00002)
< 0.40	Very poor — candidate for removal

Labeling the Factors

Factors	Variables	Name
Factor 1	Var4, Var5, Var6,Var7	'MALE EGO' or 'PRIDE OF OWNERSHIP'
Factor 2	Var3, Var8, Var9	'COMFORT' or 'SAFETY'
Factor 3	Var1, Var10	'AFFORDABILITY'

Communality of variable 2 is 45.2%.

- It implies that the only 45.2% of variation in variable 2 is captured by our extracted factors.
- This may also partially explain why variable 2 is not appearing in our final interpretation of the table.

References

- <https://www.grandacademicportal.education/assets/images/documents/20180623113120.pdf>