

Exploratory Factor Analysis.

What is it?

At its core Factor Analysis is nothing more than combining factors in a dataset that will inherently lower the number of variables during an analysis.

What comes next?

Once Factor Analysis is being approached, we must first identify whether the data is Confirmatory or Exploratory. Confirmatory Analysis is going to be further discussed in Chapter 13. Exploratory Analysis, on the other hand, is the examination of multiple variables and find a common factor or something that ties them together.

How do we approach Exploratory Analysis?

Well it all depends whether the analysis is going to have a specific unit of value, i.e time variables, type of material, services, or products in a brand. For those types of situations, there are two types of factor analysis. Correlation Matrix and Cluster Analysis. The Correlation Matrix and more specifically R Factor Analysis is the most common type of factor analysis given that checks for how variables are correlating with one another. This method is also useful to identify correlating items that are not easily observed. Then we have Cluster Analysis, in this method Q Factor Analysis is often used given that it checks for groups that may appear in a dataset.

Assumptions in Factor Analysis

During Factor Analysis there will be a plethora of issues we may encounter. Conceptual issues arrive when an analyst fails to identify the character the variables have with respect to the factor being constructed. For example, adding apples as a variable when researching vegetables as a factor. Statistical issues, arrives when the mathematical model is overly or under restrictive, meaning the significance of the correlation may not affect the variables needed to identify our factor.

Important Note: Mathematics will provide us with potential matches and items that we think are correlating with one another. It now the job of us humans, the ones who created Mathematics, to begin with, and decide whether the information provided has meaning. Therefore, we must ask the following questions. What variables are included? How are the variables measured? Will

the results given be clear representation fo the sample size we wanted? Additionally, A statistical significant Bartlett's test of sphericity (sig. <0.5) indicates that a sufficient correlation exists among the variables to proceed. Lastly, Measure of sampling adequacy (MSA) values must exceed .50 should be omitted from the factor analysis one at a time, with the smallest being omitted each time.

Selecting the Factors

As previously mentioned factors are created from variables and factor analysis is trying to find those variables. However, there is a Common Factor Analysis and Component Analysis. Common Factor Analysis looks at the common or shared variances. Component Analysis looks at the total variance and derives factors that contain small proportions of unique variance. That is to say, a factor is created when a certain event occurs in a small portion of variables.

Interpreting the Factors

While there is no clear process or guideline to interpreting the factors and while there is no better way than understand the underlying variables in a data set. However, a good rule of thumb include the following:

- An optimal structure exists when all variables have a high loading only on a single factor.
- Variables shows generally have commonalities of greater than .5
- Deleting variables that may affect the correlation
- changing rotation methods
- Increase or decrease the number of factors

Validation of the Factors

After interpreting our factor and concluding that they are indeed there to represent the variables we have grouped together. We must validate their existence. In this part of the factor analysis, we must stop and see what our factors are and check whether those factors are going to represent the data and the population the data derived from. We wouldn't want to consider customer satisfaction with the price of an item. While there could be a correlation with one another, it doesn't really affect the overall goal of the analysis (predict the cost of the next generation of a product.) Additionally, the analyst must also be wary of variables that may influence the observations, i.e outliers, missing data, etc.

Additional Uses of Factor Analysis Results

Finally, we have created, interpret, and validated my factors, what more could we possibly want? Well, let us remind ourselves that while combining variables into a factor sounds neat, there could be the possibility of overusing our simplification method. There could be a subset of variables that can be further investigated and an analyst that could potentially lead to a statistical discovery. Additionally, as scientists, we must also ask ourselves, will the replacing of the original variables with a smaller dataset impact the presence of a factor. In theory, this sounds time-consuming and a tedious process. However, factor analysis if done right could lead an incredible analytical review and representation/discovery of variables affecting the dataset.