



# DATA ANALYTICS CARD 1: FACTOR ANALYSIS

Source: Palmatier, R.W. & Sridhar, S. (2017): Marketing Strategy - Based on First Principles and Data Analytics, Palgrave, London, p. 46.

## DESCRIPTION

Factor analysis is a data reduction technique that can be used to identify a small number of latent "factors" that explain the variation in a large number of observed variables.

## WHEN TO USE IT

- ▶ To condense a large pool of potential customer needs, wants, and preferences into a short set of similar characteristics.
- ▶ To reduce high correlation among predictors.

## HOW IT WORKS

We begin with a large number of measured variables (e.g., 30) of customer survey measures. The factor analysis algorithm synthesizes the large number of measured variables into smaller sets (e.g., 3-4) of latent "factors" that capture the essence of the meaning in the larger number of measures. To choose the total number of factors to retain, we observe how many factors have an Eigenvalue greater than 1. The strength of the association between a measure variable and its factor is called the "factor loading." When a measured variable has a factor loading greater than 0.3, it is generally associated with a factor. We categorize the measured variable with a factor where it has the highest loading (e.g., if a measured variable has factor loadings of 0.01 and 0.8 with Factors 1 and 2, we would associate the measured variable with Factor 2). Finally, we interpret what each latent factor represents, by surmising the conceptual commonality underlying the measured variables' loading on the factor.

## EXAMPLE

The manager of an online website collected customer satisfaction data from a survey of 1,000 customers on eight aspects of the company's focal product. The table shows the factor loadings of a few variables after conducting a factor analysis with three factors. Factor 1 is highly associated with product diversity, specialty, and price; thus, it can be interpreted as the "product" factor. Factor 2 is associated with cash back and discounts, and is thus labeled the "promotion" factor. For Factor 3, the "service" factor, delivery service and customer service have the highest factor loadings. The factors can be used as data input for segmentation analyses. The figure shows the focal attributes associated with each factor.

Attribute	Factor 1	Factor 2	Factor 3	Focal attributes associated with factors
Product diversity	<b>0.665</b>	-0.016	0.017	Diversity
Product specialty	<b>0.681</b>	-0.056	0.006	Specialty
Product price	<b>0.638</b>	0.284	0.173	Price
Cash back	-0.042	<b>0.712</b>	0.051	Cash back
Discount	0.216	<b>0.781</b>	0.103	Discount
Delivery service	-0.007	0.178	<b>0.752</b>	Delivery service
Customer service	0.155	0.199	<b>0.739</b>	Customer service