Data Analysis Types

Data analysis is the process of producing useful information by inspecting, cleansing, transforming, and modeling raw data.

There are various methodologies for analyzing data. Each has its own benefits and limitations. There is also some debate about how to categorize data analysis methods and their specific definitions.

In general, the following types of analysis are the most commonly used.

1. **Descriptive Analysis**

   Descriptive Analysis looks at data to find trends, highlight whether KPIs (Key Performance Indicators) were met, etc. This is the most common type of analysis and is used frequently for building "dashboards" of important information. The following measures (centers and spread) are frequently used:

   - Mean
   - Median
   - Mode
   - Standard deviation
   - Variance
   - Interquartile Range (IQR)

2. **Diagnostic (Inferential) Analysis**

   Diagnostic Analysis is often used for finding a "why". For example, why did a machine fail after a short period of time? Why did a movie not do as well as predicted? As with all analysis, the key is collecting enough data to make a good diagnosis. For example, you cannot make a diagnosis about a movie failure without capturing viewer data. Similarly, machine failures require data not just on machines that failed, but also for machines that did not fail. The key is deriving conclusions about the whole from analyzing part of the data.

3. **Predictive Analysis**

   Predictive Analysis is the art of making educated guesses about the future. The "educated" part of guess comes from the data. Forecasting can be as imprecise as next week's weather forecast, but businesses need to make decisions based on their best guess about the future. Waiting is not an option for goods and services that are provided in advance. For example, clothes manufacturers need to have their goods in stores before the season in which they sell. Grocery stores need food

restocked every day, but can't wait for weekly or even monthly reports to tell them what to stock. The analysis here is to use historical data and patterns to predict what may happen in the future.

4. **Prescriptive Analysis**

   Sometimes grouped with Prescriptive Analysis, this form of analysis is for optimizing choices found in Predictive Analysis. For example, if your Predictive Analysis shows a machine failure before an optimal service interval, you can choose to shorten the service interval or improve the quality of the parts, depending on what your Prescriptive Analysis shows. Risk management often plays a role in this type of analysis.

5. **Exploratory Analysis**

   Statistical Analysis is analyzing data sets to discover the relationship (if any) between variables. Specifically, the goal is to find:

   - The variables and their relationships.
   - Find outliers and/or the most influential variables.
   - Usually published in the form of reports or dashboards.

   While exploratory data analysis can discover relationships between variables, answering the "why/how are they related" questions is usually done through diagnostic analysis.

Resources

1. Accounting is Analytics: https://accountingisanalytics.com/
2. NIH Methods of Data Collection, Representation, and Analysis: https://www.ncbi.nlm.nih.gov/books/NBK546485/
3. Analyzing and using data: https://www.education.vic.gov.au/school/teachers/teachingresources/practice/Pages/insight-data.aspx
4. Data Analysis and School Administration–opinion paper on the (mis)use of data analysis: https://eric.ed.gov/?id=ED453596
5. Root Cause Analysis for Beginners (PDF): https://ldh.la.gov/assets/medicaid/hss/docs/NH/RootCauseForBeginners.pdf

This material is based upon work supported by the National Science Foundation under Grant DUE 2055411.Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.