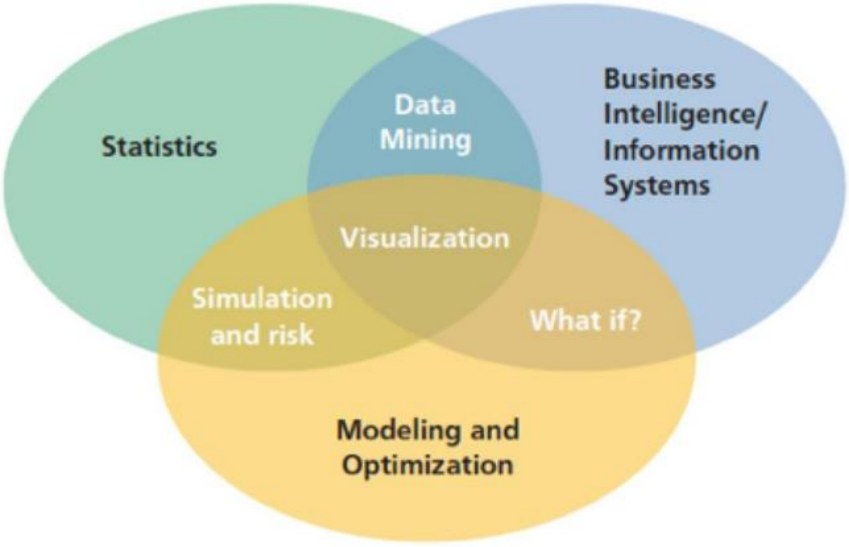
1 BUSINESS ANALYTICS AND DESCRIPTIVE STATISTICAL MEASURES

Business Analytics

Analytics is the use of data, information technology, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight into their business operations and make better, fact-based decisions.

- Pricing
  - setting prices for consumer and industrial goods, government contracts, and maintenance contracts
- Customer segmentation
  - identifying and targeting key customer groups in the retail, insurance, and credit card industries
- Merchandising
  - determining brands to buy, quantities, and allocations
- Location
  - finding the best location for bank branches and ATMs, or where to service industrial equipment
- Social Media
  - understand trends and customer perceptions; assist marketing managers and product designers

A Visual Perspective of Business Analytics



Impacts and Challenges

- Benefits
    - …reduced costs, better risk management, faster decisions, better productivity, and enhanced bottom-line performance such as profitability and customer satisfaction.
- Challenges
    - …lack of understanding of how to use analytics, competing business priorities, insufficient analytical skills, difficulty in getting good data and sharing information, and not understanding the benefits versus perceived costs of analytics studies.

Scope of Business Analytics

- **Descriptive analytics**: the use of data to understand past and current business performance and make informed decisions

- **Predictive analytics**: predict the future by examining historical data, detecting patterns or relationships in these data, and then extrapolating these relationships forward in time.

- **Prescriptive analytics**: identify the best alternatives to minimize or maximize some objective

Example 1.1: Retail Markdown Decisions

- Most department stores clear seasonal inventory by reducing prices.

- Key question: When to reduce the price and by how much to maximize revenue?

- Potential applications of analytics:
    - Descriptive analytics: examine historical data for similar products (prices, units sold, advertising, …)
    - Predictive analytics: predict sales based on price
    - Prescriptive analytics: find the best sets of pricing and advertising to maximize sales revenue

Descriptive and Inferential Statistics

When analyzing data, such as the marks achieved by 100 students for a piece of coursework, it is possible to use both descriptive and inferential statistics in your analysis of their marks. Typically, in most research conducted on groups of people, you will use both descriptive and inferential statistics to analyze your results and draw conclusions. So what are descriptive and inferential statistics? And what are their differences?

Descriptive Statistics

Descriptive statistics is the term given to the analysis of data that helps describe, show or summarize data in a meaningful way such that, for example, patterns might emerge from the data. Descriptive statistics do not, however, allow us to make conclusions beyond the data we have analyzed or reach conclusions regarding any hypotheses we might have made. They are simply a way to describe our data.

Descriptive statistics are very important because if we simply presented our raw data it would be hard to visualize what the data was showing, especially if there was a lot of it. Descriptive statistics, therefore, enables us to present the data in a more meaningful way, which allows a simpler interpretation of the data. For example, if we had the results of 100 pieces of students' coursework, we may be interested in the overall performance of those students. We would also be interested in the distribution or spread of the marks. Descriptive statistics allow us to do this. How to properly describe data through statistics and graphs is an important topic and discussed in other Laerd Statistics guides. Typically, there are two general types of statistics that are used to describe data:

- **Measures of central tendency**: these are ways of describing the central position of a frequency distribution for a group of data. In this case, the frequency distribution is simply the distribution and pattern of marks scored by the 100 students from the lowest to the highest. We can describe this central position using several statistics, including the mode, median, and mean. You can learn more in our guide: Measures of Central Tendency.
- **Measures of spread**: these are ways of summarizing a group of data by describing how to spread out the scores are. For example, the mean score of our 100 students maybe 65 out of 100. However, not all students will have scored 65 marks. Rather, their scores will be spread out. Some will be lower and others higher. Measures of spread help us to summarize how spread out these scores are. To describe this spread, several statistics are available to us, including the range, quartiles, absolute deviation, variance, and standard deviation.

When we use descriptive statistics it is useful to summarize our group of data using a combination of tabulated description (i.e., tables), graphical description (i.e., graphs and charts), and statistical commentary (i.e., a discussion of the results).

Inferential Statistics

We have seen that descriptive statistics provide information about our immediate group of data. For example, we could calculate the mean and standard deviation of the exam marks for the 100 students and this could provide valuable information about this group of 100 students. Any group of data like this, which includes all the data you are interested in, is called a population. A population can be small or large, as long as it includes all the data you are interested in. For example, if you were only interested in the exam marks of 100 students, the 100 students would represent your population. Descriptive statistics are applied to populations, and the properties of populations, like the mean or standard deviation, are called parameters as they represent the whole population (i.e., everybody you are interested in).

Often, however, you do not have access to the whole population you are interested in investigating, but only a limited number of data instead. For example, you might be interested in the exam marks of all students in the UK. It is not feasible to measure all exam marks of all students in the whole of the UK so you have to measure a smaller sample of students (e.g., 100 students), which are used to represent the larger population of all UK students. Properties of samples, such as the mean or standard deviation, are not called parameters, but statistics. Inferential statistics are techniques that allow us to use these samples to make generalizations about the populations from which the samples were drawn. It is, therefore, important that the sample accurately represents the population. The process of achieving this is called sampling (sampling strategies are discussed in detail in the section, Sampling Strategy, on our sister site). Inferential statistics arise out of the fact that sampling naturally incurs sampling error and thus a sample is not expected to perfectly represent the population. The methods of inferential statistics are (1) the estimation of the parameter(s) and (2) the testing of statistical hypotheses.

What are the similarities between descriptive and inferential statistics?

Both descriptive and inferential statistics rely on the same set of data. Descriptive statistics rely solely on this set of data, whilst inferential statistics also rely on this data to make generalizations about a larger population.

What are the strengths of using descriptive statistics to examine the distribution of scores?

Other than the clarity with which descriptive statistics can clarify large volumes of data, there are no uncertainties about the values you get (other than only measurement error, etc.).

What are the limitations of descriptive statistics?

Descriptive statistics are limited so much that they only allow you to make summations about the people or objects that you have measured. You cannot use the data you have collected to generalize to other people or objects (i.e., using data from a sample to infer the properties/parameters of a population). For example, if you tested a drug to beat cancer and it worked in your patients, you cannot claim that it would work in other cancer patients only relying on descriptive statistics (but inferential statistics would give you this opportunity).

What are the limitations of inferential statistics?

There are two main limitations to the use of inferential statistics. The first, and most important limitation, which is present in all inferential statistics, is that you are providing data about a population that you have not fully measured, and therefore, cannot ever be completely sure that the values/statistics you calculate are correct. Remember, inferential statistics are based on the concept of using the values measured in a sample to estimate/infer the values that would be measured in a population; there will always be a degree of uncertainty in doing this. The second limitation is connected with the first limitation. Some, but not all, inferential tests require the user (i.e., you) to make educated guesses (based on theory) to run the inferential tests. Again, there will be some uncertainty in this process, which will have repercussions on the certainty of the results of some inferential statistics.