

Data Analysis: Introduction and Exploration

Course Outline



Overview

In the 21st century, good decision making is driven by data. Data needs to be analyzed and interpreted and translated into the appropriate policies. It is important that one develops an understanding that they will need to make informed decisions using data, and to communicate the results effectively. The course offered by e-Biome is an introduction to the essential concepts, tools and methods of statistics in science and technology, business, economics and similar disciplines, although it may have wider interest.

The focus of this course is on concepts, reasoning, interpretation and thinking rather than computation, formulae and theory. It therefore pulls on participants' knowledge to write effectively and communicate their ideas with clarity.

The Data Analysis course is the best way to get quality training in statistics in a convenient manner. Our course will equip you to work with vast amounts of data to make predictions regarding customer behavior, anticipate market trends, and perform a host of other tasks that will help any business or non-profit succeed.

Course Description

Data analysis and analytics are evolving disciplines. Advanced techniques rest on fundamentals which can be applied in many job roles. This course quickly equips you with the necessary foundation.

The course covers two main branches of statistics: descriptive statistics and inferential statistics. Descriptive statistics includes collecting data and summarizing and interpreting them through numerical and graphical techniques. Inferential statistics includes selecting and applying the correct statistical technique in order to make estimates or test claims about a population based on a sample. Topics covered may include descriptive statistics, correlation and simple regression, probability, and hypothesis testing.

By the end of this course, participants should understand and know how to use statistics. Participants will also develop some understanding of the limitations of statistical inference and of the ethics of data analysis and statistics.

Objectives

- Apply correctly a variety of statistical techniques, both descriptive and inferential.
- Interpret, in plain language, the application and outcomes of statistical techniques.
- Interpret computer output (Excel and SPSS) and use it to solve problems.

- Recognize inappropriate use or interpretation of statistics and comment critically on the appropriateness of this use of statistics.

Learning Outcomes

At the end of this course, students will be able to understand the following:

1. Data and Information
 - a. Data in the Real World
 - b. Data vs. Information
 - c. Structured Data and Unstructured Data
 - d. Types of Data
2. Data Analysis Defined
 - a. Why do we analyze data?
 - b. Data Analysis Steps
 - c. Data Analysis Defined
 - d. Descriptive Statistics vs Inferential Statistics
3. Types of Variables
 - a. Categorical vs Numerical
 - b. Nominal Variables
 - c. Ordinal Variables
 - d. Interval Variables
 - e. Ratio Variables
4. Central Tendency of Data
 - a. (Arithmetic) Mean
 - b. Median
 - c. Mode
5. Distributions, Variance, and Standard Deviation
 - a. Data Collection and Sampling
 - b. Discrete Distributions
 - c. Continuous Distributions
 - d. Range
 - e. Quartiles
 - f. Variance
 - g. Standard Deviation
 - h. Population vs. Sample
 - i. Distributions in Excel

- j. Statistical Inference and Sampling Distribution
6. Fitting Data
- a. Bivariate Data (Two Variables)
 - b. Covariance and Correlation
 - c. Simple Linear Regression
 - d. Linear Regression
 - e. Graphical Descriptive Techniques - Nominal Data
 - f. Graphical Descriptive Techniques - Numerical Data

Course Content

UNIT	CONTENT	ASSESSMENT
Unit 1: Data and Information	<ul style="list-style-type: none"> • Data in the Real World • Data vs. Information • Structured Data and Unstructured Data • Types of Data 	Final MCQ 30 Marks
Unit 2: Data Analysis Defined	<ul style="list-style-type: none"> • Why do we analyze data? • Data Analysis Steps • Data Analysis Defined • Descriptive Statistics vs Inferential Statistics 	
Unit 3: Types of Variables	<ul style="list-style-type: none"> • Categorical vs Numerical • Nominal Variables • Ordinal Variables • Interval Variables • Ratio Variables 	
Unit 4: Central Tendency of Data	<ul style="list-style-type: none"> • (Arithmetic) Mean • Median • Mode 	
Unit 5: Distributions, Variance, and Standard Deviation	<ul style="list-style-type: none"> • Data Collection and Sampling • Discrete Distributions • Continuous Distributions • Range • Quartiles • Variance • Standard Deviation • Population vs. Sample • Distributions in Excel • Statistical Inference and Sampling Distribution 	
Unit 6: Fitting Data	<ul style="list-style-type: none"> • Bivariate Data (Two Variables) 	

	<ul style="list-style-type: none"> • Covariance and Correlation • Simple Linear Regression • Linear Regression • Graphical Descriptive Techniques - Nominal Data • Graphical Descriptive Techniques - Numerical Data 	
--	---	--

Resources:

1. <https://www.math.arizona.edu/~jwatkins/statbook.pdf>
2. <https://d3bxy9euw4e147.cloudfront.net/oscms-prodcms/media/documents/IntroductoryBusinessStatistics-OP.pdf>
3. <http://www.utstat.toronto.edu/mikevans/jeffrosenthal/book.pdf>
4. <https://files.eric.ed.gov/fulltext/ED536788.pdf>
5. http://www.ru.ac.bd/wp-content/uploads/sites/25/2019/03/102_10_Longnecker_An-Introduction-to-Statistical-Methods-and-Data-Analysis-6th-Ed.pdf
6. <https://www.statsref.com/StatsRefSample.pdf>