

DATA ANALYSIS

Course description

Data analysis is progressing fast in almost all areas of business as more voluminous data sets are generated and decision making with these is more appropriate. On the other hand, securing, cleaning, processing, analyzing and presenting data are becoming more sophisticated and complicated. This course intends to introduce students the basic tools and material of data analysis towards promoting them to grasp all aspects of data analysis.

Course objective

More frequent use of social media and computerization of almost all economic and industrial activity lead to generation of huge unstructured data sets in forms of emails, tweets, photos, videos, logs, and speeches. On the other hand, ultra powerful processors in highly capable computers enable making use of such data for better management of companies. Currently data analysts have gained the ability to store huge volumes of data, analyze them, and extract information to run business more efficiently. The key objective of this course is to familiarize the students with main tools of information technology used in processing big data. From the statistical point of view, this course reviews and expands upon core topics in probability and statistics through theory and practice of data analysis. These topics include but are not limited to numerical and graphical summaries of data, hypothesis testing, confidence intervals, analysis of variance, regression, and correlation. Having completed the course, students will be able to think critically about data and apply standard statistical inference procedures for more appropriate decision making in management.

Learning outcomes

At the end of the course students will be able to:

1. Contemplate the importance of decision making in business, its typical difficulties and pitfalls, and the importance of essential data analysis in decision making.
2. Decide on exact tools and methods of statistics and probability theory to process data.
3. Apply common quantitative and visual techniques to aid in management decision making.
4. Develop an appreciation of how data and data analytics can be used by managers to make better decisions.
5. Be familiar to a set of commonly used terms and techniques in the area of “big data” and data analytics that are in use today.

Course content

This course is divided into five units to be covered:

1. Introduction
 - a What is data analysis/science/engineering?

- b Application areas
- c The major steps for solving a data analysis problem
 - i Determining the business problem
 - ii Collecting data
 - iii Reading, manipulating and cleaning data
 - iv Analyzing data
 - v Building a model
 - vi Evaluation and interpreting the results
 - vii Reporting and visualization

2. Tools for data analysis

- a Excel
- b Python
- c Databases(SQL)

3. Statistical foundations for data analysis: Descriptive statistics, statistical inference and outlier treatment

- a Population, sample, parameter, statistic
- b Types of data: qualitative, quantitative, discrete, continuous
- c Levels of measurement
- d Organization and presentation of data
- e Descriptive statistics for central tendency and dispersion
- f Frequency distributions
- g Probability and probability distributions
- h Statistical inference: hypothesis testing, confidence interval construction
- i Correlation and regression
- j Residual analysis and outlier treatment

4. Data analysis practices/case studies

5. Comprehensive Case Study

Course Outline*:

<u>Week 1</u>	<u>Unit 1</u>
<u>Week 2</u>	<u>Unit 2</u>
<u>Week 3</u>	<u>Unit 2</u>
<u>Week 4</u>	<u>Unit 2</u>
<u>Week 5</u>	<u>Unit 3</u>
<u>Week 6</u>	<u>Unit 3</u>
<u>Week 7</u>	<u>Unit 3/4</u>
<u>Week 8</u>	<u>Unit 4</u>

*There may be changes due to unexpected conditions

Mode of delivery

Online.

Assessment

The assessment will be on pass/fail decision based on the student performance in the project scheduled to the end of the course. Each student is supposed to complete the comprehensive project and submit it on time. Successful students will be granted the certificate.

Student responsibility

Regular attendance at lectures is expected. The end-of-course project shall be completed and submitted by each student.