

Course Profile: Introduction to statistics and data analysis

Course Number: INF 6490

Credits: 3

Prerequisite(s): None (see below)

Course Description

Statistical methods of data analysis are necessary for understanding important aspects of our world and are essential for a variety of research and professional activities. This course offers an applied introduction to statistics and data analysis. The course covers the process of research question formulation to data organization, analysis and communication of results. It covers the elements of descriptive and exploratory statistics, including probability concepts and data visualization. The course covers inferential statistics, including techniques of correlation and regression, as well as confidence intervals and hypothesis testing. The course uses practical examples to guide the student through the various concepts, issues and analyses, and provides tutorials on how to use statistical software (e.g. spreadsheet, Python or R). Course activities include readings, videos, quizzes, data analysis labs, and a final data analysis project. Helpful pre-requisite knowledge: introductory algebra; basic command of computer technology (e.g. Web browsing, downloading/installing programs, opening/saving files).

Learning Outcomes:

Students who successfully complete this course will be able to:

1. Articulate the role and value of statistics and data analysis across distinct settings.
2. Define concepts of probability, data types and distributions.
3. Explain the role of random sampling, causality, and inference.
4. Perform exploratory data analysis and data visualization with statistical software.
5. Apply and assess results of correlation and regression analysis.
6. Conduct hypothesis testing and explain p-values and confidence intervals.
7. Assess adequacies and fallacies with approaches to data analysis.

Topics:

1. Research question formulation and study design.
2. Exploratory data analysis (categorical data, quantitative data, intro to R).
3. Probability and distributions (sampling, random variables, continuous variables).
4. Multivariable models (variable interactions, multiple regression).
5. Inference for categorical data (normal distribution, proportions, errors, power).
6. Inference for quantitative data (one mean, paired means, two means, t-tests).
7. Inference in regression (linear regression, model assumptions).
8. Writing and communication of data and statistics.

Course Activities:

1. Readings
2. Videos
3. Quizzes
4. Discussions
5. Data analysis labs
6. Data analysis project

Evaluation of Student Performance:

The student's performance will be evaluated using:

1. Quizzes
2. Data analysis labs
3. Data analysis project
4. Discussion posts

Last Updated: May 11 2022