האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

The Hebrew University of Jerusalem

Syllabus

INTRODUCTION TO DATA SCIENCE - 71253

Last update 16-03-2021

HU Credits: 3

Degree/Cycle: 1st degree (Bachelor)

Responsible Department: agro informatics

Academic year: 0

Semester: 2nd Semester

Teaching Languages: Hebrew

Campus: Rehovot

Course/Module Coordinator: Dr. Jonathan Friedman

Coordinator Email: yonatan.friedman@mail.huji.ac.il

Coordinator Office Hours: Appointments should be coordinated by email

Teaching Staff:

Dr. Yonatan Friedman,
Mr. Isaac Kramer

Course/Module description:

In recent years, there has been a dramatic increase in the amount of data generated in the world, and data science has become an major discipline with numerous applications across science and industry.
The course introduces the mathematical foundations and computational tools used in data science by combining theoretical lectures with hands-on experience using Python for the analysis and visualization of datasets related to agriculture, food, and the environment.

Course/Module aims:

The course aims to provide the theoretical background and practical skills required for analyzing and visualizing data in order to gain novel insights.

Learning outcomes - On successful completion of this module, students should be able to:

- Collect, clean, and manipulate data
- Formulate precise, quantitative questions
- Analyze data and detect differences between groups and dependencies between variables.
- Choose and create data visualizations according to the nature of the data and the analysis.

Attendance requirements(%):

Teaching arrangement and method of instruction: Lectures + students' presentations of their projects

Course/Module Content:

Foundation skills
1. Introduction to data science
2. Data types; reading and cleaning data
3. Organizing and manipulating data

4. Principles of data visualization

Analyzing single variables
5. Distributions and summary statistics
6. Smoothing and interpolation

Analyzing pairs of variables
7. Differences between groups
8. Dependencies between variables.
9. Nonparametric statistics and bootstrapping

Analyzing multiple variables
10. Dimension reduction
11. Clustering analysis
12. Comparing multiple groups

Required Reading:
Data Science from Scratch: First Principles with Python 2nd Edition

Additional Reading Material:
Storytelling with Data: A Data Visualization Guide for Business Professionals

Course/Module evaluation:
End of year written/oral examination 0 %
Presentation 0 %
Participation in Tutorials 0 %
Project work 50 %
Assignments 50 %
Reports 0 %
Research project 0 %
Quizzes 0 %
Other 0 %

Additional information: