**ADA 440**

**COURSE SYLLABUS**

**Python for Data Science**

**(3 - 0) 3**

|  |
| --- |
| **Instructors:** Assoc. Prof. Dr.Yıldırım Akbal |
|  |
| **Course Description:** |
| This course will introduce the fundamental concepts of Python programming language, its syntax, functions and packages to enable them to implement algorithms to manipulate, analyze and retrieve data. The course will also cover main packages such as Numpy, Pandas and Matplotlib. |
|  |
| Text Book: |
| Python for Data Analysis: Data Wrangling with pandas, NumPy, and Jupyter 3rd Edition (Wes McKinney) |
|  |
| **Course Objectives:** |
| The main objective of the course is to provide students with the basic concepts of Python, its syntax, functions and packages to enable them to write scripts for data manipulation and analysis. The course develops skills of writing and running a code using Python. The course covers various variables types and their features, basic operators and statements, loops, as well as the main packages for data science: NumPy, Pandas, Matplotlib. At the end of the course, students should be able to write short scripts to import, prepare and analyze data. |
|  |
| **Course outcomes:** Upon successful completion of this course students |
| 1. Understand the programming principles such as control structures, data types, conditional statements; |
| 1. **Understand the principles of object-oriented programming,** 2. **Create own tools, rather than using ready-to-go ones,** |
| 1. Analyze data using Python programming language, 2. Infer information from data by visualizing it, 3. Design simple tests, |
| **Course policy**: For the first 5 – 6 weeks of the lecture we will be looking at the. This will form a solid background for the following weeks. We will then be looking at how to use Python and its packages efficiently. |
|  |
| **Grading Policy:**   1. ALE: 6 x (6%) (The number of ALE’s may change – may be given as an HW) 2. Midterm: 25% 3. Final: 40 % |
|  |

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **COURSE CHART**   |  |  |  |  | | --- | --- | --- | --- | | **WEEK** | **Dates** | **Course Topics** |  | | 1 |  | Basic data types | ALE1 | | 2 |  | Functions, Modules |  | | 3 |  | Conditional Statements | ALE2 | | 4 |  | Loops – Exception Handling |  | | 5 |  | Object Oriented Programming |  | | 6 |  | Object Oriented Programming Cont’d  -- Capstone Project (Text generation) | ALE4 | | 7 |  | Data visualization (Matplotlib – Seaborn) |  | | 8 |  | Operation with Arrays (Numpy)  --- Stochastic Gradient Descent with Linear Regression !!!!  --- KNN from scratch | ALE5 | | 9 |  | Midterm | MT1 | | 10 |  | Numpy cont’d – Pandas (data cleaning – some normalization, flling out missing values). |  | | 11 |  | Pandas cont’d |  | | 12 |  | Case study | ALE6 | | 13 |  | Introduction to SQL (Lite) |  | | 14 |  | SQL (Lite) cont’d |  | |  |