CLARUSWAY

Lesson Plan

Data Analysis with Python

This course will give you comprehensive and valuable information about Numpy and Pandas libraries and by doing hands-on exercises you will learn how to use Python to analyze data. At the end of the course, you will have the intuition to prepare the data for any Decision Support Purposes or any Machine learning algorithm.

Custodian : Matthew Connor-Instructor (matthew_c@clarusway.com)

In-class Sessions : 14 Sessions (42 hours)

Lab Sessions : 3/4 Labs (3/4 hours)

Certification Requirements:

1. % 70 attendance to in-class lessons (at least 10/14 for DAwPython Course)
2. Successfully completing and submitting assignments & projects (at least 1 assignment & 1 project for DAwPy Course)

Prerequisites

- Basic Math
- Basic Statistics
- Python Experience

Course Outline

1. NumPy
   - 1.1. NumPy Arrays and Operations
     - 1.1.1 Introduction to NumPy
     - 1.1.2 NumPy Arrays
     - 1.1.3 NumPy Array Indexing
     - 1.1.4 NumPy Operations
   - 1.2. Numpy Exercise

2. Pandas
    - 2.1. Pandas Part-1
        - 2.1.1 Introduction to Pandas
        - 2.1.2. Pandas Series
        - 2.1.3 Data Frame Basics-1 (Attributes, Indexing, Column Creating)
        - 2.1.4 Data Frame Basics-2 (Properties)
        - 2.1.5 Data Frame Basics-3 (Dropping,Selecting)
        - 2.1.6 Data Frame Basics-4 (Slicing Methods, Filtering)
    - 2.2. Pandas Part-2
        - 2.2.1. Groupby
        - 2.2.2. Pivot Table & Stack
        - 2.2.3 Useful Methods
        - 2.2.4 Missing Data
        - 2.2.5 Outliers
        - 2.2.6 Merging-Joining-Concatenating
        - 2.2.7 Text and Time Methods
        - 2.2.8 Regular Expressions (REGEX)
        - 2.2.9 Data Input and Output
    - 2.3 Pandas Exercise

Materials & Resources

- Clarusway Learning Management System (LMS)
- Numpy Official Documentation
- Pandas Official Documentation
- Jake VanderPlas (2016). Python Data Science Handbook: Essential Tools for Working with Data, O'Reilly.

Tools and Software

- Zoom, Slack, Kahoot, Peardeck Applications
- Jupyter Notebook / Google Colab
- SciPy, Numpy, Pandas, Matplotlib, Seaborn

Assignments & Projects

Assignments

1. Assignment-1 (Numpy Arrays & Operations, Pandas DataFrame Basics & Operations, and DataFrame Indexing & Selection)
2. Assignment-2 (Groupby, Combining DataFrames, Handling Missing Data, and Text & Time Methods)

Projects

1. Exploratory Data Analysis Project (Analyze US Citizens).