The first step through the world of data Analysis is learning a program language. For this Course we use Python which is the most widely used data science programming language in the world today.

So to start coding in python you should learn it's syntax and basic concepts (Variables, list, tuple, dictionaries, array….) so here are some tutorials for beginners (if you are familiar with these basics you can skip):

- <u>Python for Beginners (full Course)</u> (Youtube - Step by Step Instruction)

- <u>Python tutorial - Tutorialspoint </u>(A comprehensive guide for beginners)

- <u>A Primer on Scientific Programming with Python </u>by H. P. Langtangen

- <u>W3schools Python Tutorial </u>(educational website for learning coding online step by step)

- Notes from Dr. Alaei:
    - Installation of Python in Windows and Linux
    - Introduction to Python
    - Matplotlib Tutorial
    - Input data and Error handling
    - File Strings and Dictionaries
    - Array Computing and Curve Plotting
    - Introduction to Classes

Once you get familiar with the elementary concepts, the next step is to learn how to manipulate data using Python. As you may know, Python was the most popular data science programming language of 2021, and the reasons why are endless. It is easy to use, and easy to learn. Python provides all the necessary tools for the 4 steps of problem solving — data collection & cleaning, data exploration, data modeling and data visualization. So here are some References you can use to get familiar with These Concepts:

- <u>Python Data Science Handbook</u> by Jake VanderPlas
- <u>Python for Data Analysis</u>   by Wes McKinney
- <u>Introduction to Machine Learning with Python</u> by Andreas C. Müller and Sarah Guido

These three textbooks are well known and standard Refs that make your learning enjoyable with clear explanations of different Data Analysis libraries in Python (Numpy, Pandas, matplot).