Data Visualization in Python

Introduction

In this course, we will learn how to manipulate and plot data that is useful to us. For that, we will use Python and its libraries. There are a few libraries that we can use for visualization and manipulation. In this course, we will use Matplotlib to plot data and Pandas to manipulate it.

There are several libraries for plotting data in Python. In this course, as we mentioned, we will use Matplotlib. Still, you can use libraries such as Seaborn, Plotly, etc... Matplotlib is a low-level graph plotting library in Python that serves as a visualization utility. It is open source and we can use it freely. An interesting fact about Matplotlib is that the creator of Matplotlib, John D. Hunter, originally developed it to visualize brain signals while working in a neurobiology lab. It has since evolved into the most widely used Python plotting library. You can find all the functionalities of the Matplotlib library here: https://matplotlib.org.

The name "Pandas" has a reference to both "Panel Data", and "Python Data Analysis". Pandas is used to analyze data. It is used for working with data sets. It's widely used in data science. Pandas allow us to analyze big data and make conclusions based on statistical theories. Pandas can clean messy data sets, and make them readable and relevant. It is highly optimized for performance and can handle millions of rows of data in memory efficiently. This makes it a go-to tool for data cleaning and analysis. Relevant data is very important in data science. More details about this library can be found here: https://pandas.pydata.org/

Installation

Installation is very easy. You can install library Matplotlib by using pip or conda. Commands are shown below.

Pip installation:

```
pip install matplotlib
```

Conda installation:

```
conda install -c conda-forge matplotlib
```

Installation for Pandas is also very easy. Command for installation is shown below.

```
pip install pandas
```