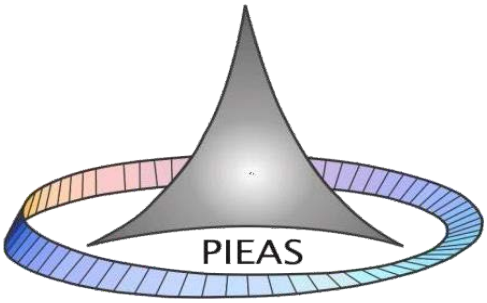
Week-1

Stock Price Prediction



Submitted to:

Dr Aneela Zameer

Submitted by:

Group:

Aniq Javed

Usman Asad

Ahmed Raza

Pakistan Institute of Engineering and Applied Sciences (PIEAS)

- ## Data Extraction:

We were able to get the data for S&P 500 from following link:

https://www.kaggle.com/datasets/andrewmvd/sp-500-stocks

The data contains more than **2 million examples** which are continuously **updating** every day.

There are three datasets in this repository:

1. S&P companies Data

```
----------Information about S&P Companies Data--------------------

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 494 entries, 0 to 493
Data columns (total 16 columns):
 #   Column               Non-Null Count  Dtype
---  ------               --------------  -----
 0   Exchange             494 non-null    object
 1   Symbol               494 non-null    object
 2   Shortname            494 non-null    object
 3   Longname             494 non-null    object
 4   Sector               494 non-null    object
 5   Industry             494 non-null    object
 6   Currentprice         494 non-null    float64
 7   Marketcap            494 non-null    int64
 8   Ebitda               451 non-null    float64
 9   Revenuegrowth        481 non-null    float64
 10  City                 494 non-null    object
 11  State                474 non-null    object
 12  Country              494 non-null    object
 13  Fulltimeemployees    490 non-null    float64
 14  Longbusinesssummary  494 non-null    object
 15  Weight               494 non-null    float64
dtypes: float64(5), int64(1), object(10)
```

2.S&P Index Data

```
----------Information about S&P Index Data-------------------

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2519 entries, 0 to 2518
Data columns (total 2 columns):
 #   Column  Non-Null Count  Dtype
---  ------  --------------  -----
 0   Date    2519 non-null   object
 1   S&P500  2519 non-null   float64
dtypes: float64(1), object(1)
```

3. S&P stock Data

```
----------Information about S&P Stocks Data----------

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1644307 entries, 0 to 1644306
Data columns (total 8 columns):
 #   Column     Non-Null Count    Dtype
---  ------     --------------    -----
 0   Date       1644307 non-null  object
 1   Symbol     1644307 non-null  object
 2   Adj Close  1575810 non-null  float64
 3   Close      1575810 non-null  float64
 4   High       1575810 non-null  float64
 5   Low        1575810 non-null  float64
 6   Open       1575810 non-null  float64
 7   Volume     1575810 non-null  float64
dtypes: float64(6), object(2)
```

- ## Heat Map Generation

We were able to generate heatmap by using following fields as shown below:

```
#heatmap
sns.heatmap(sp_500_comp[['Currentprice','Marketcap','Ebitda','Revenuegrowth','Fulltimeemployees','Weight']].corr(), annot=True)
plt.show()
```