University of New Haven

Connecticut

Spring 2022

CSCI-6674 (Data Mining)

Phase 4 – Data Exploration

Submitted by   **Submitted to**

Team Data Diggers   Dr. Shivanjali Khare

Department of Computer Science

Date – 03/13/2022

1) Team Name – **Team Data Diggers**

Team Members –

1. Kunal Dhanaitkar (kdhan1@unh.newhaven.edu)

2. Anil Kumar Reddy Allikepalli (aalli2@unh.newhaven.edu)

3. Koushik Akkenapally (kakk1@unh.newhaven.edu)

2) The datasets we have selected are from the **past few years** and are from an authenticated website**.** We'll be working with datasets from the Facebook, Apple, Amazon, Netflix, and Google (FAANG) big-tech companies. Since these five companies are technological epitome and being technocrats ourselves, we would like to study their stocks and understand how they have exponentially grown over the years and made an impact on the world. Also, we would like to study on which company we can invest in the near future.

3) **Exploration Techniques Used:**

In our project we used have used bivariate analysis for the data exploration. The dataset which we have contains only numerical attributes. So, we have used the following data exploration techniques:

1. Line plot
2. Scatter plot
3. Box plot

4) **Description and Findings of Data Exploration Techniques Used:**

The exploration techniques which we have used in our project is bivariate analysis. The attributes which we have considered for this analysis are close price and date. We have used both the line and scatter plots for this analysis.

According to the exploration techniques used, the box plot exhibits that the Apple stock has performed consistently well except with few ups and downs. During the years 2018-2019 it has seen a downside but picked up thereafter which thereby suggests the stock is a good one to be added in the portfolio.

For the line graph plotted for Google, we can observe the graph shows that the stock performed very well over the years, with a very little downsides, it managed to thrive, thereby indicating this can be bought to have profits.

From the line graph plotted for Netflix, there is certain rise and downfall thereafter due to covid restrictions. So, Netflix is not a proper stock to be invested in.

From the graph plotted for Amazon, there is a drastic rise is the stock price due to people opting for online shopping because of social distancing. This indicates that the stock is a good one to be invested in.

5) **<u>Github Repository Link:</u>**

https://github.com/aanilkumarreddy/DataDiggers