# **Project Guidelines**

## Documentation Components:

### 1.Timeline

**Project Topic Discussion:** September 20th & September 21st, decided between exploring finance data, and using data from the Govt. of Canada website to assess the nature of an increase in Food Prices over time.

**Project Topic Finalization:** September 24th, decided to analyze stock data prices from the yahoo finance website.

**Data Extraction Guide:** September 25th, explored how the data would need to be sourced, webscrapping, manual download, using a specific python library to extract data from a compatible website.

**Data Extraction:** The data was Extracted from the Yahoo Finance Website on September 26th, 2023, using a compatible python library.

**Data Cleaning:** The data was cleaned to be easily sourced and accessed on September 27th.

**Data Manipulation:** Data was converted into a JSON format, so that it could be utilized for data analysis on September 27th.

**Database Storage:** Utilized XML instead to store the data in order for it be accessed on September 27th.

**API Creation:** Developed and created an API to be sourced from with the help of flask on September 28th.

**Website Creation:** Utilized javascript, html & css to a create working website on September 28th.

**Website** **Modification:** Modified the data visualizations to be more interactive and stable on September 30th & October 1st.

**Project Analysis:** Discussed the structure of the project, main contents and analysis required for completion on October 2nd.

**PPT Creation:** Created a structured slide deck to present our data on stock analysis findings on October 3rd.

### 2.Data Source

We conducted a data collection on the Yahoo Financial website, gathering information on 100 companies and cross-referencing their acronyms to ensure the accuracy. We used the Yahoo Finance Library, which allowed us to search and collect the historical data spanning the full year of 2023.

Links:

<https://finance.yahoo.com/most-active/?count=100&offset=0>

<https://pypi.org/project/yfinance/#files>

### 3.Steps

* Utilized the Py.Pi library through Python in order for the data to be extracted and stored in the form of a csv file.
* Converted the csv file into a json format, so it could be stored in a database.
* Utilized XML as the preferred method to store and access the data.
* Developed an API with the help of Flask in order for the data to be sourced.
* Utilized Javascript libraries to plot data as a line chart, column chart and Pie Chart.
* Developed a presentation in order for the data to be shared, and to discuss key findings

### 4.Data Components

Date: Choose Dates from Sept 27th,2022 to Sept 28th,2023 for 100 stocks.

Open: Opening Price is the value of a stock when the market opens for the day.

High: The highest price of a stock when traded throughout the day.

Low: The lowest price of a stock when traded throughout the day.

Close: Closing Price is the value of a stock when the market closes for the day.

Volume: The number of shares traded daily as per the Data Set.

Dividends: Profit Sharing payment made in the form of cash or reinvested shares

Company: The name of the company

Core Data Visualizations:

1. Line Graph: Utilized the company name and closing price of each stock daily to highlight a particular trend in the data.
2. Column Chart: Aggregated the total volume of all the stock data daily and highlighted the top 5 stocks by trading volume out of 100 companies.
3. Pie Chart: Highlighted the Top 5 stocks that offer the highest dividends annually out of 100 companies.

### 5.Roles and Responsibilities Presentation

Ana – Data Collection

Hikmet: Frontend and Backend Development of website with visualizations

Tyler: API Creation & Design

Abish: Project Design & Analysis