**Term Project**

**Foreign exchange Rate**

Akila Selvaraj

Masters of Data Science,

Bellevue University

DSC 540 T301

Aug 06,2022

**Topic**

This project is to analyze how the forex rates of the countries have changed over the years and the impact of factors that influence currency rates in the countries. To achieve that, I had to collect exchange rate information from three different sources and in three different formats. First format of the data is in csv format which has history data, and the second in html table which gives recent values as per last working day the and third from API which gives real time rates.

**Data Handling**

For this project, first step is to search and find the data from three different sources which can be related to each other. I got the first csv file from Kaggle and the second html data from currency calculator website which contains the table of exchange rates, and the third from currency layer API link.

After finding all the sources, next step is to extract the data and clean it so that data would be in usable format and can be combined with other data. One of the main things to consider while cleaning the data is format of the data. Cleaning the data requires removal of duplications, removing or replacing missing entries, correcting misfielded values, ensuring consistent formatting and a host of other tasks which take a considerable amount of time. We need to make sure that data should be in correct format. In this project, as this involves various time period, I need to make sure that the data and time format are in sync. Handling missing values is the next step as missing values would lead to incorrect results in our analysis. Transform the data is also necessary to make it usable. If the csv data is huge, I have filtered only the required data for my analysis.

To use the data, I have analyzed each dataset, looked for duplicates, got rid of the outliers to create appropriate results. Columns from different sources have been renamed to make the join easy. Some of the packages used are API requests, html\_parser to parse and handle html, pprint to print the data, BeautifulSoup for web scraping.

Ethical implications should always be taken in consideration when we work on any data science project. The data I have collected is public data and it is not a personal data belonging to any particular individual or an organization. In this forex exchange data, I used the collected information ethically for analyzing the price variation trends without changing the meaning of the data. Unethical data handling can result in loss of reputation and would lead to loss of trust. In this project, none of the information or content has been amplified or reduced, and the original information has been used as is, ethically.