

Mehmet Egehan Pala

31089

CS 210 Term Project Phase II Project Report

Goal of This Project:

The main aim of this project is to find out whether there is a correlation between the Consumer Price Index (CPI) of USA- Microsoft Stock Prices and between the CPI of USA - Gold Prices between the period 2019-2023. The CPI is one of the primary indicators of inflation, which reflects the change in the general consumer goods prices over a constant time. By understanding the effects of the CPI fluctuations over other datasets, valuable predictions can be obtained over stock prices.

Expected Outcome of This Project:

If there is a significant correlation found after Pearson correlation test, a linear regression model will be built to predict to outcomes of the inflation over the datasets.

Building the linear regression model and analysing it will provide some valuable insights about stock fluctuations which will significantly help investors to hedge market behaviour accordingly in the inflationary periods.

Expected Insights of This Project:

- **Inflation Sensitivity:** By analysing the correlation between CPI and other datasets, we can understand the reason behind the change in stock prices on inflationary pressures. Understanding how fluctuations in CPI affect investor behaviour and stock valuations is an important issue in predicting market reactions to macroeconomic trends.
- **Investor Behaviour:** By analysing the bond between CPI and stock prices investors can arrange their behaviour accordingly during inflationary times. Investors may adjust their portfolio allocations, risk preferences, and expectations based on inflation forecasts, influencing market dynamics.
- **Macroeconomic Implications:** The outcome of this project can teach people more about the effects of inflation in business side and the entire stock market. The information that is derived from this project could be used by the people who makes decisions in economy, such as economists and government officials. They can use the acquired information to keep the economy steady and decide what policies might or might not be needed to avoid economic fluctuations.

Data Description:

The dataset that is used in this project has 1 numerical column that is used which is the cpi column. The data set consists of 918 rows containing monthly CPI values from January 1947 to June 2023. Since the dates did not match, the given data set was changed in monthly order and the average value of that month was assigned so that they become paired. The dates were then matched with each other.

The CPI dataset has:

Mean: 117.794828

Standard Deviation: 83.736360

Min Value: 21.480000

Max Value: 303.841000

The CPI dataset is in numerical form (Float64), so since both the datasets are in numerical for Pearson Correlation test have been used in hypothesis testing.

References:

<https://www.kaggle.com/datasets/saketk511/2019-2024-us-stock-market-data?resource=download>

<https://www.kaggle.com/datasets/pavankrishnanarne/us-inflation-dataset-1947-present>