**July-Dec 2024**

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**

A close-up of a logo

Description automatically generated

**SOCIAL NETWORK ANALYSIS ASSIGNMENT**

Time Series Forecasting and prediction of US CPI

Submitted to:- Submitted by:-

Mr. Tej Bahadur Riddhi Gupta(E22BCAU0057)

Kamal Verma(E22BCAU0116)

**Objective:**

This project aims to develop an accurate time series forecasting model to predict the future values of the US Consumer Price Index (CPI). Specifically, it will:

1. To understand what is CPI and how it is related to inflation of an area, city, state, country or even globally.
2. Analyse historical CPI data to identify patterns and trends.
3. Preprocess the data for accuracy.
4. Evaluate and select the best forecasting model .
5. Result

The goal is to provide valuable insights into future inflation trends to aid in economic decision-making.

**1.To understand what is CPI and how it is related to inflation of an area, city, state, country or even globally.**

**Consumer Price Index (CPI)** measures the average change in prices paid by consumers for goods and services over time, reflecting the cost of living.

**Relation to Inflation:**

* **Inflation Rate**: Calculated using CPI to indicate how fast prices are rising.
* **Regional Analysis**: CPI can be analyzed for cities, states, or countries to reflect local economic conditions.
* **Global Perspective**: Comparing CPI across countries provides insights into global economic stability.

**Applications:**

* **Adjusting Income**: Governments use CPI to adjust wages and benefits.
* **Policy Decisions**: Central banks use CPI for monetary policy.
* **Economic Analysis**: Economists gauge economic health using CPI.

In summary, CPI helps understand and manage inflation, informing policy decisions and economic strategies.

### **2.Analysis of Historical US CPI Data (1910-2020)**

We analysed the historical data of the United States Consumer Price Index (CPI) from 1910 to 2020 to identify patterns and trends.

#### **Period of 1910 to 1940**

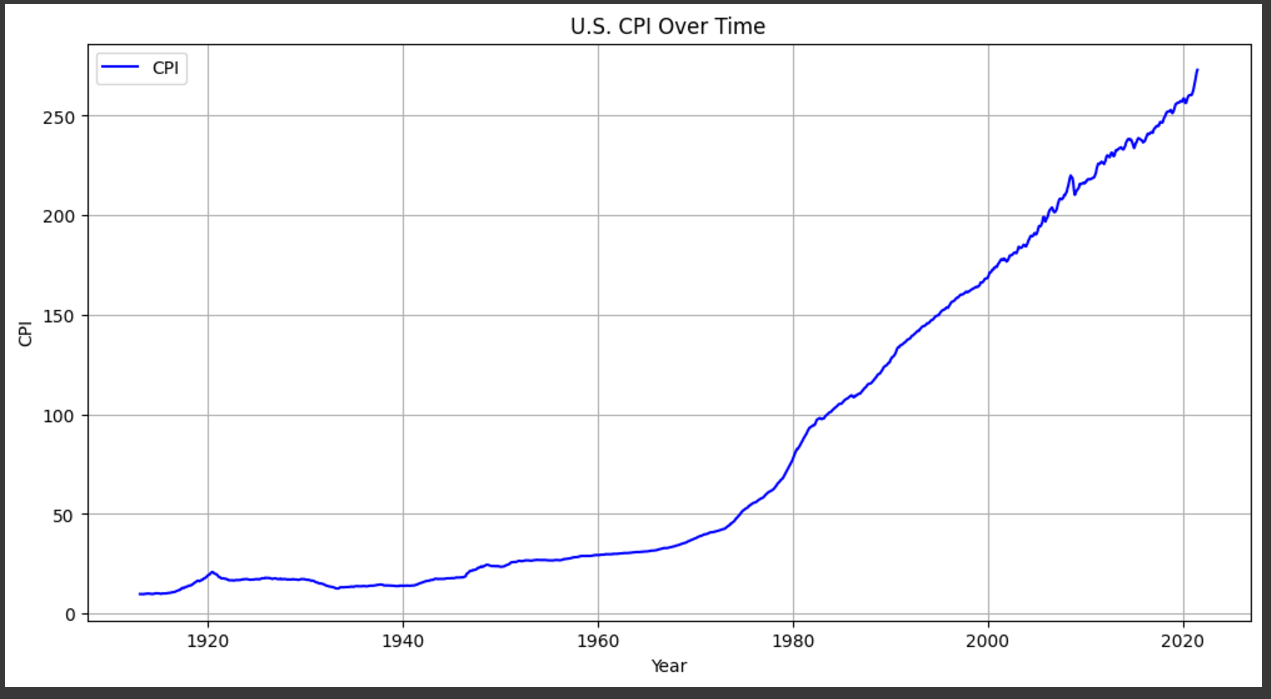
During this period, CPI showed minimal change. Stability can be attributed to economic policies and market conditions of the time, including the impact of the Great Depression.

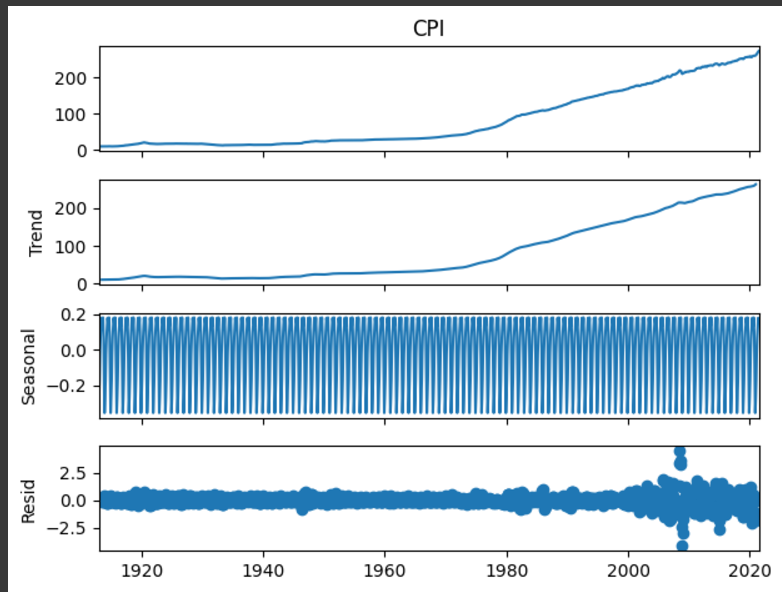
#### **Post-1940 Trend**

From 1940 onwards, CPI increased significantly, reflecting economic growth, post-war reconstruction, and inflationary pressures due to increased consumer demand.

***Seasonality***

The analysis also revealed distinct seasonal trends influenced by factors like weather, holidays, and agricultural cycles.



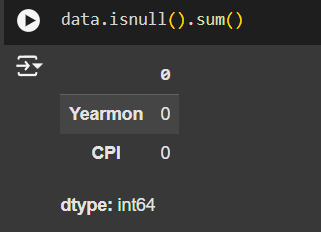


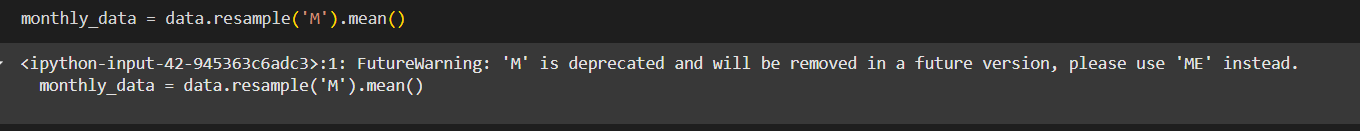
**3.Data Preprocessing for Accuracy**

In the preprocessing phase, we ensured the data's accuracy by checking for missing values and assessing the need for resampling:

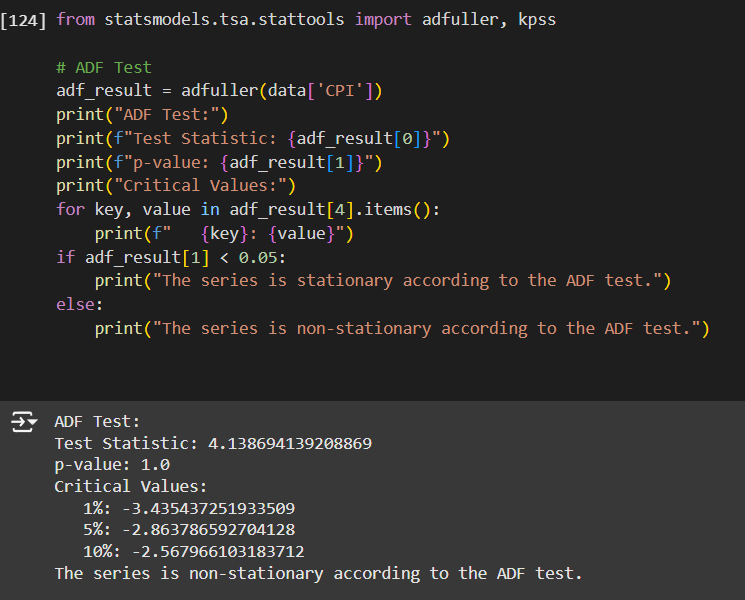
* **Missing Values**: We conducted an analysis to detect any missing values within the dataset. The results confirmed that there were no missing values, ensuring the integrity of the data.
* **Data Resampling**: Although resampling is a common preprocessing step to standardize time intervals, our dataset was already recorded at monthly intervals. Therefore, resampling did not have any impact on the data.

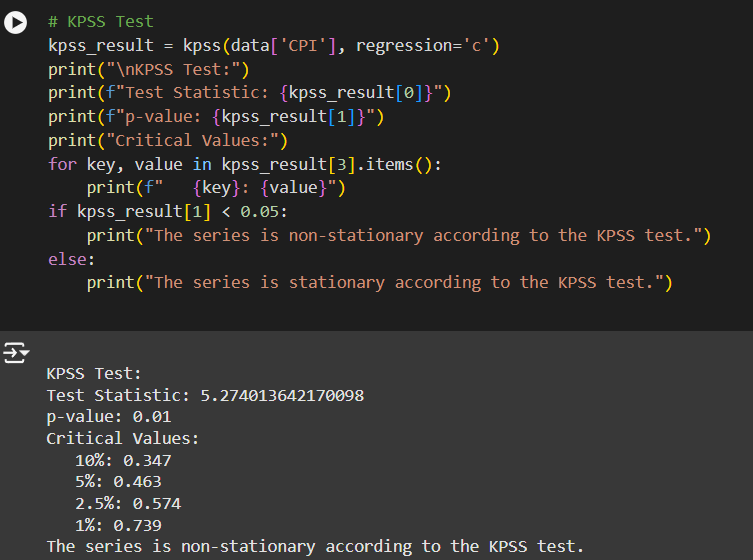
These preprocessing steps assured us that the data was clean and consistent, ready for effective time series forecasting.





We also performed the stationarity test such as ADF (Augmented-Dicky Fuller) test and KPSS (Kwiatkowski-Phillips-Schmidt-Shin) test.   
And by performing we got to know that both the tests deduced that the series is non-stationary.





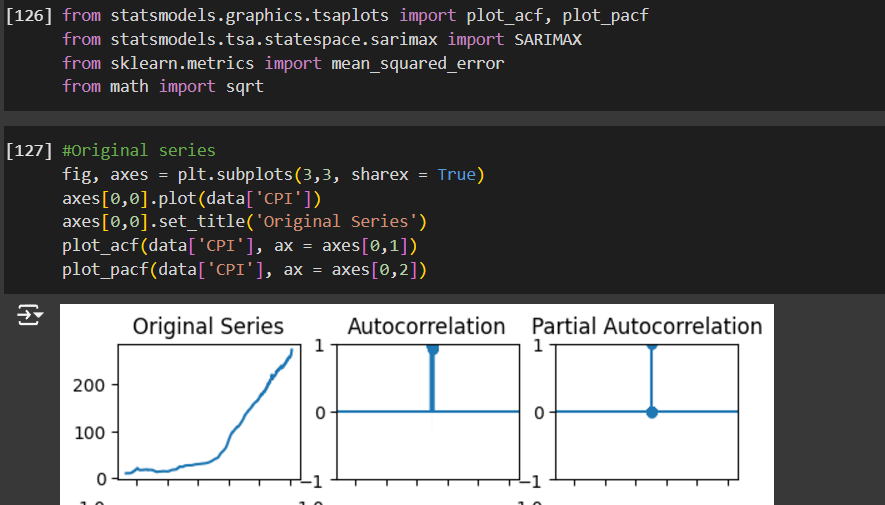
**Evaluation and Selection of the Best Forecasting Model**

In this section, we evaluated three different forecasting models: SARIMA, AutoARIMA, and Holt-Winters Exponential Smoothing, to determine the most effective model for predicting the US Consumer Price Index (CPI).

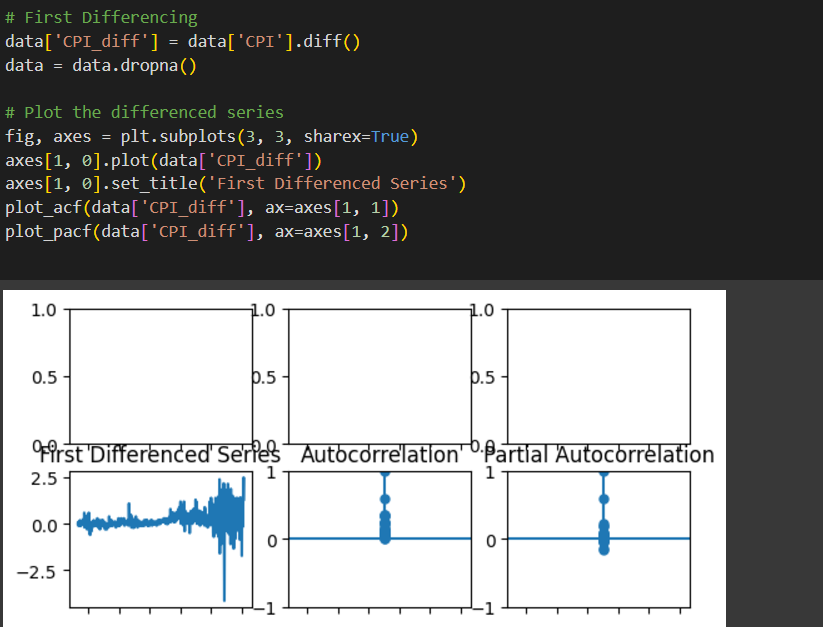
**Models Evaluated:**

* 1. **SARIMA** (Seasonal AutoRegressive Integrated Moving Average) – we also plotted the acf and pacf plots to determine the value of p,d and q and P,D,Q and S. We also used evaluation metrices such as RMSE and MSE to assess the performance of the forecasting model.

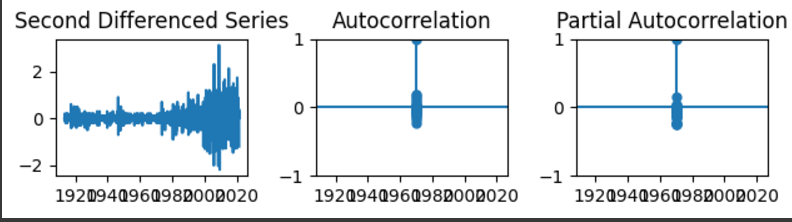
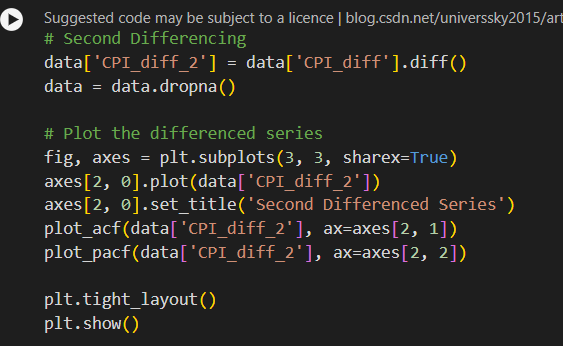
- original series



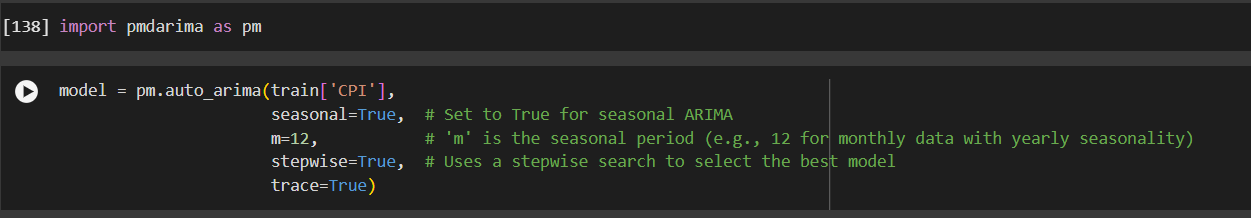
-First Order Differencing

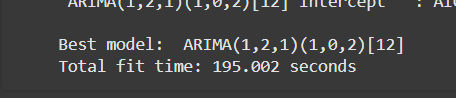


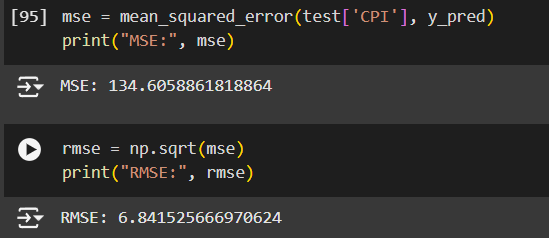
-Second Order Differencing



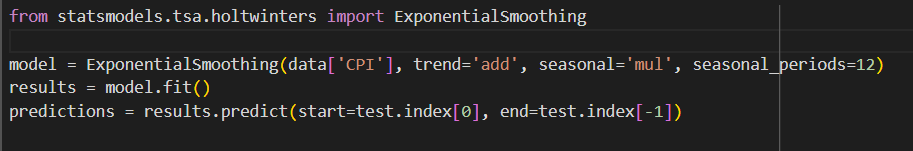
* 1. **AutoARIMA (Automatic ARIMA)** - **AutoARIMA** is a method that automatically identifies the best ARIMA model by testing different combinations of parameters (p, d, q) and selecting the optimal set based on a chosen criterion.

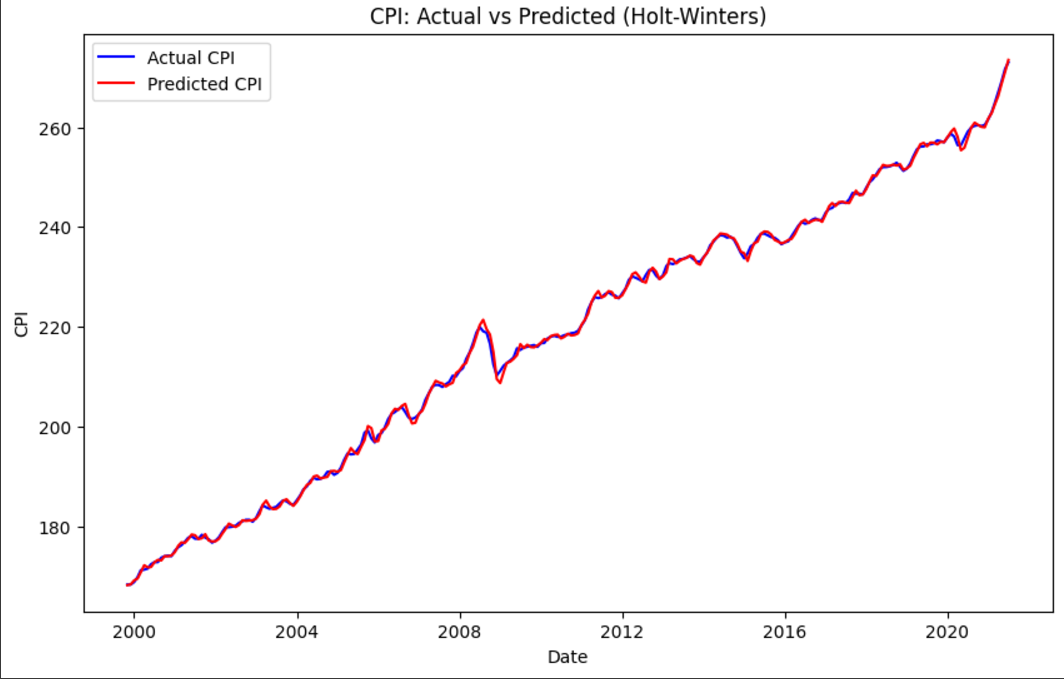


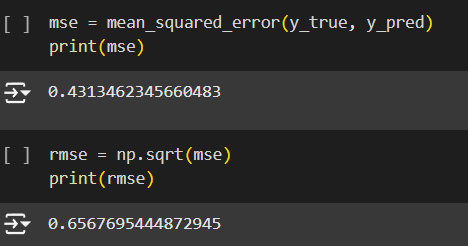


****

* 1. **Holt-Winters Exponential Smoothing –** Holt-Winter's Exponential Smoothing is a time series forecasting method that accounts for trend and seasonality in the data. It focuses on making immediate or near-future predictions based on recent historical data.







**Result:-**

In this study, we implemented three forecasting models—**SARIMA**, **AutoARIMA**, and **Holt-Winter's Exponential Smoothing**—to predict the future values of the US Consumer Price Index (CPI). After evaluating the models using performance metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), and analyzing the prediction plots, we observed that Holt-Winter's Exponential Smoothing provided the most accurate results, with the lowest MSE and RMSE among the models. Consequently, this model was chosen to forecast the US CPI for the next 12 months, ensuring reliable and precise predictions for economic analysis.\

