**Predicting & Combating Inflation Report**

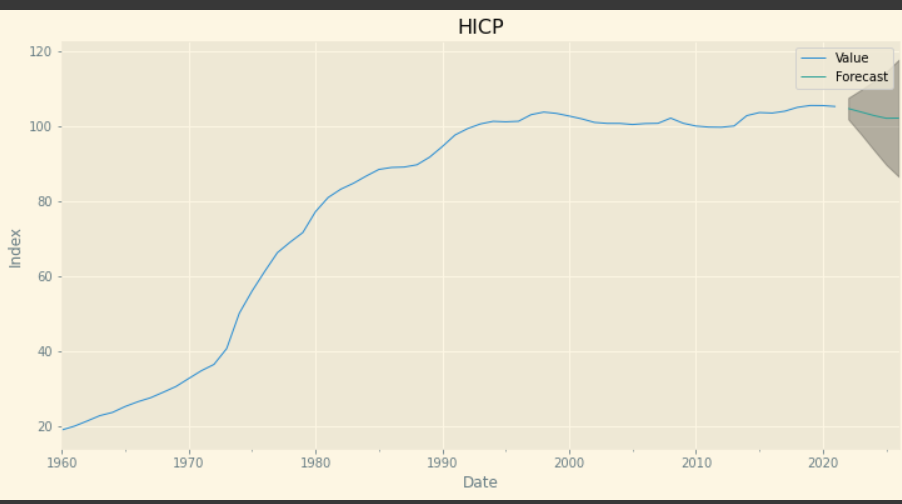
| **Name** | Alfaiz Shaikh |
| --- | --- |
| **Country** | JAPAN |
| **Colab Notebook Link** | [Japan data.ipynb - Colaboratory (google.com)](https://colab.research.google.com/drive/1TJ5MUqVNBs_f5CdmtWZwpcJpHmuYOyXM#scrollTo=Ny0hrlKJ0nmg) |
| **Dataset Link** | [Consumer price index (2010 = 100) | Data (worldbank.org)](https://data.worldbank.org/indicator/FP.CPI.TOTL) |

**Description:** By the help ofMachine learning models we predicted inflation rates for the next 5 years in Japan. Taking into account these inflation predictions, the portfolios were made , spanning high risk, medium risk and low risk portfolios.

**HICP/CPI:** Consumer price index is referred to as that index that is used in calculating the retail inflation in the economy by tracking the changes in prices of most commonly used goods and services. It changes yearly.

CPI can be used to calculate the cost of living of the people of a country and also the changes in the purchasing power of the currency of a nation.

**CPI = (Cost of market basket in a given year / Cost of market basket in base year) x 100**

****

**Time-series Components:** Time series data analysis is the way to predict time series based on past behaviour. Prediction is made by analysing underlying patterns in the time-series data.

Given the discrete nature of time series, the data is often taken from discrete time periods. Here are the 4 major components:

**Trend component**: This is useful in predicting future movements. Over a long period of time, the trend shows whether the data tends to increase or decrease.

**Seasonal component**: The seasonal component of a time series is the variation in some variable due to some predetermined patterns in its behaviour.  The pattern is appearing at a regular interval, E.g., Quarterly GDP of India.

**Cyclical component**: The cyclical component in a time series is the part of the movement in the variable which can be explained by other cyclical movements in the economy. This also depends on regular patterns but frequency is more than a year. The data shows fluctuations at any time of the year. E.g., Union Budget.

**Residual:** The data remaining after removing the above three is called Residual. It is also known as random or irregular component.

By using Autocorrelation and Partial Autocorrelation we find out the best p and q values for our SARIMA model and find out that values of CPI generally increase from year 1986 to 2021 thus showing an increasing trend.

**Diagnostics:**

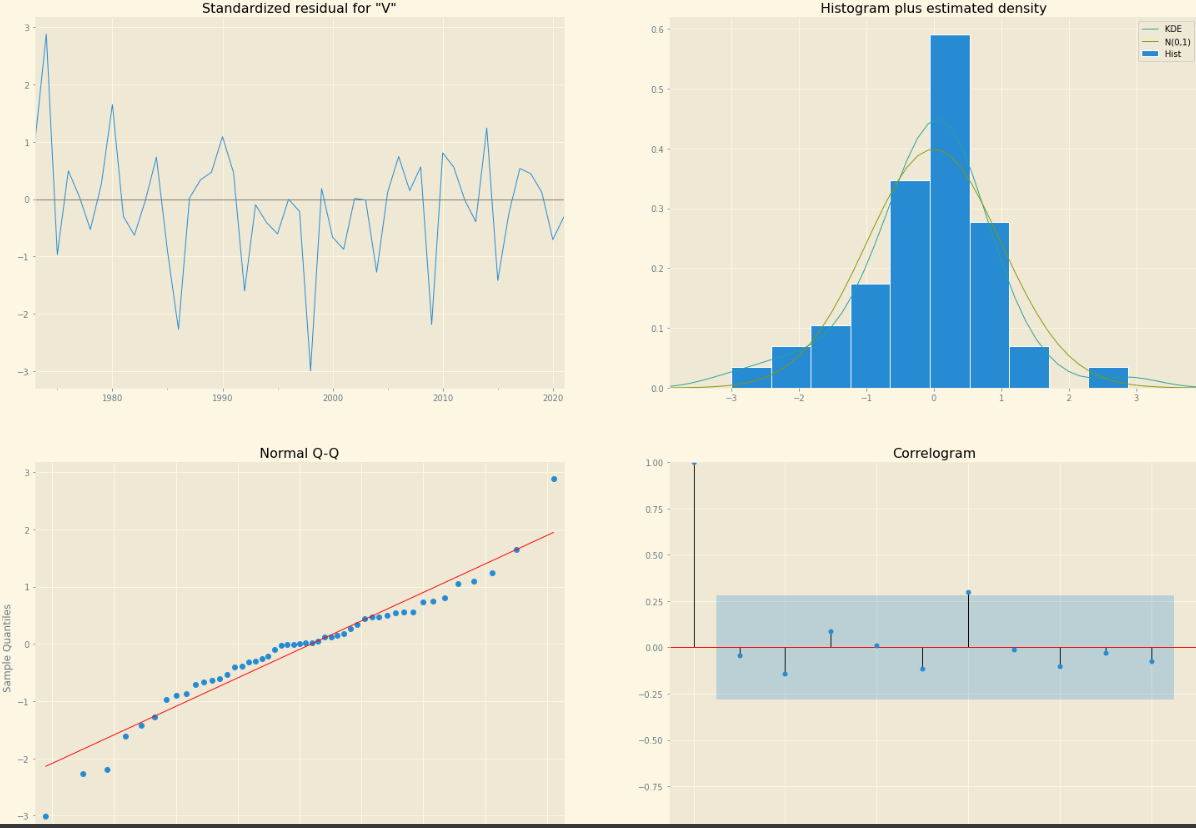
**Q-Q plots:** “Quantile-Quantile” plot, is used to assess whether or not a set of data potentially came from some theoretical distribution.

If the data is normally distributed, the points in a Q-Q plot will lie on a straight diagonal line. Thus, as seen in the figure below the data is normally distributed as most points lie on a straight diagonal line.

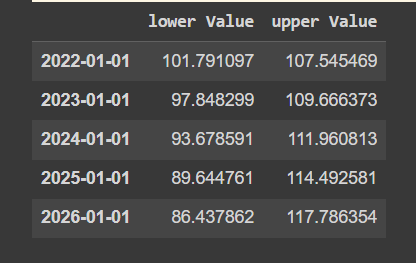
**Correlogram:** The correlogram is a commonly used tool for checking [randomness](https://en.wikipedia.org/wiki/Randomness) in a [data set](https://en.wikipedia.org/wiki/Data_set). If random, autocorrelations should be near zero for any and all time-lag separations. If non-random, then one or more of the autocorrelations will be significantly non-zero.

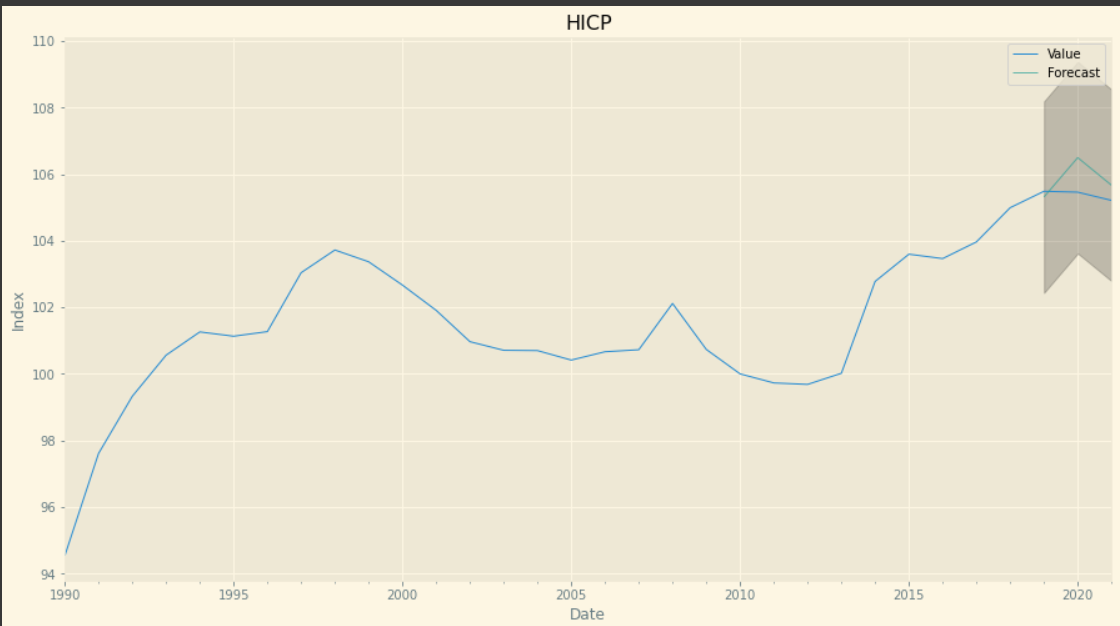
As in the given figure many points are above 1 thus the dataset is not random!

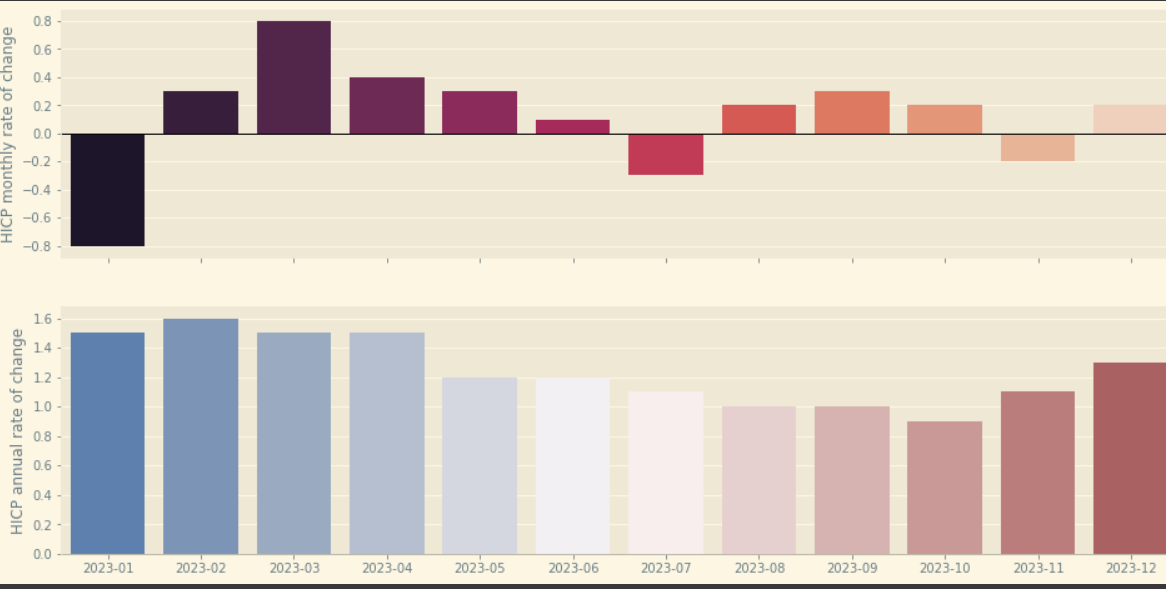
**Standardized residual:** A residual is the difference between an observed value and a predicted value in a [regression model](https://www.statology.org/linear-regression/). One type of residual we often use to identify outliers in a regression model is known as a **standardized residual**. From the plot we can see that none of the points exceed 3 or -3 thus none of the observations appear to be outliers.



**HICP Predictions:** In Japan we can see that the inflation rate is constant so its easy to beat inflation but for better earnings we should invest more percentage of our capital in international market or real esatate.



****

****