**PREDICTIVE ANALYTICS LABORATORY**



**Submitted to**

**Ms.K.Vani**

**CURRENCY SWAP PREDICTION**

**REPORT**

**Submitted by**

**DUSHYANTH.M**

### COIMBATORE INSTITUTE OF TECHNOLOGY

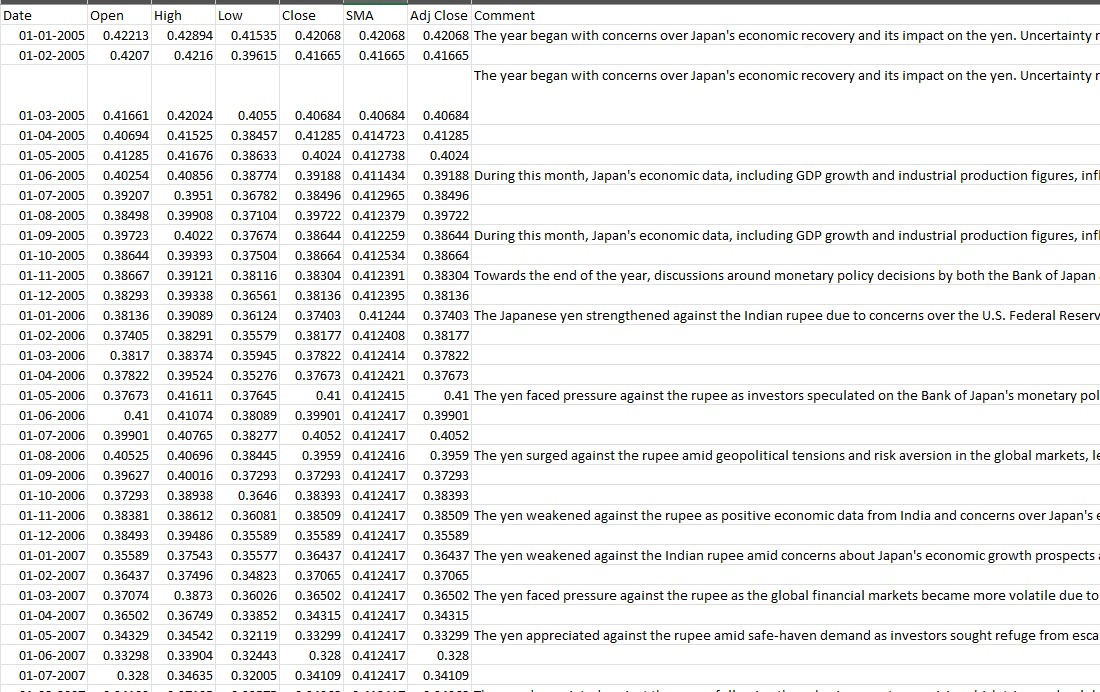
### (A Government Aided Autonomous Institution Affiliated to Anna University)

### Coimbatore – 641 014,Tamil Nadu

**ABSTRACT:**

Our Dataset is about the prediction of the currency swaps using the Multilinear regression,Polynomial regression and testing such as the Ftest and Ttest. And the dataset we used in this was the currency exchange rate dataset from the yahoo finance The currency exchange rate dataset provides a comprehensive collection of historical exchange rates between various global currencies. This dataset is essential for financial analysts, economists, and researchers interested in studying currency trends, foreign exchange market behaviors, and macroeconomic indicators. It includes daily, monthly, and annual exchange rates, capturing fluctuations influenced by geopolitical events, economic policies, and market dynamics. The data spans multiple decades, offering a longitudinal perspective.

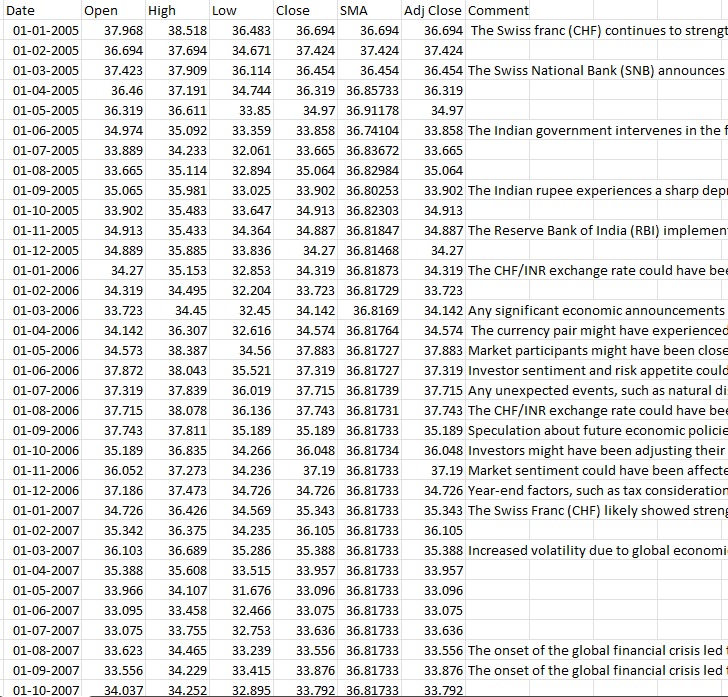
**JAPAN DATASET:**



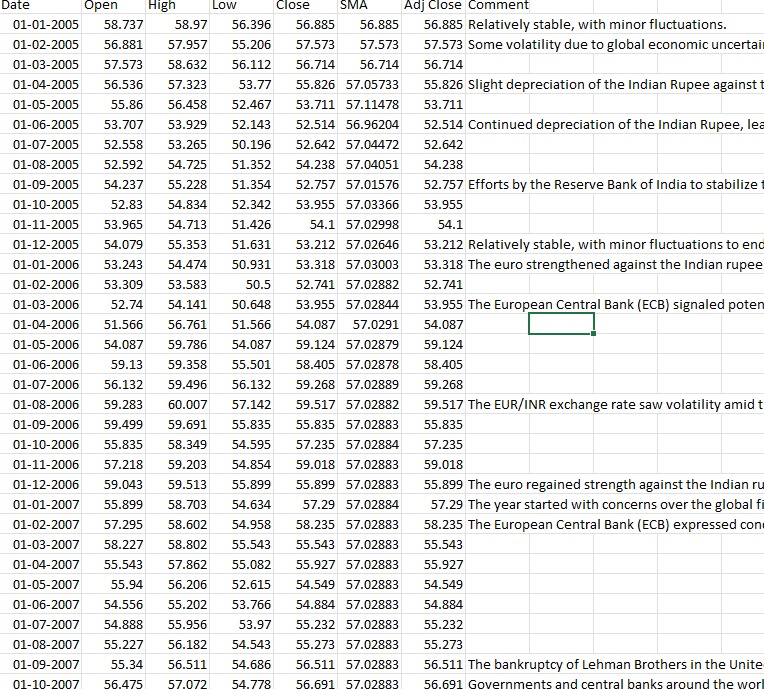
AUSTRALIA DATASET:



SWITZERLAND:



EUROPE:



**REGRESSION:**

**1)Multi linear regression:**

Multiple linear regression is a statistical technique used to model the relationship between a dependent variable and two or more independent variables. The primary goal is to understand how the dependent variable changes when any one of the independent variables is varied, while the other independent variables are held constant. In multiple linear regression, the coefficients (*𝛽1,𝛽**2,…,𝛽𝑛β*1 ,*β*2 ,…,*β**n* ) are estimated using methods such as ordinary least squares (OLS), which minimizes the sum of the squared differences between the observed values and the values predicted by the model. This technique allows for the assessment of the relative importance of each independent variable and helps in making predictions or inferences about the relationships within the data.

When performing multiple linear regression in Python for currency exchange rates, the typical steps involve data collection, preprocessing, model building, evaluation, and interpretation. Here's a detailed outline of what we typically do:

**MODEL OF MULTI LINEAR REGRESSION:**

1. **Data Collection**:
   * Gather historical currency exchange rate data from reliable sources (e.g., financial APIs, online databases).
   * Ensure the data includes the dependent variable (e.g., exchange rate) and multiple independent variables (e.g., economic indicators, interest rates, inflation rates).
2. **Data Preprocessing**:
   * **Cleaning**: Handle missing values, outliers, and any inconsistencies in the data.
   * **Feature Selection**: Choose relevant independent variables that potentially influence the exchange rates.
   * **Feature Engineering**: Create new features if necessary (e.g., lagged variables, moving averages).
   * **Normalization/Standardization**: Scale the data to ensure all variables are on a comparable scale.
3. **Exploratory Data Analysis (EDA)**:
   * Visualize the data to understand relationships between variables (e.g., scatter plots, correlation matrix).
   * Perform statistical tests to confirm the significance of relationships.
4. **Model Building**:
   * Import necessary libraries (e.g., **pandas**, **NumPy**, **scikit-learn**).
   * Split the data into training and testing sets.
   * Fit a multiple linear regression model using the training data

5. **Model Evaluation**:

* Make predictions on the testing set.
* Evaluate the model using metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared (R²).

6. **Interpretation**:

* Analyze the coefficients to understand the impact of each independent variable on the exchange rate.
* Determine which variables are statistically significant

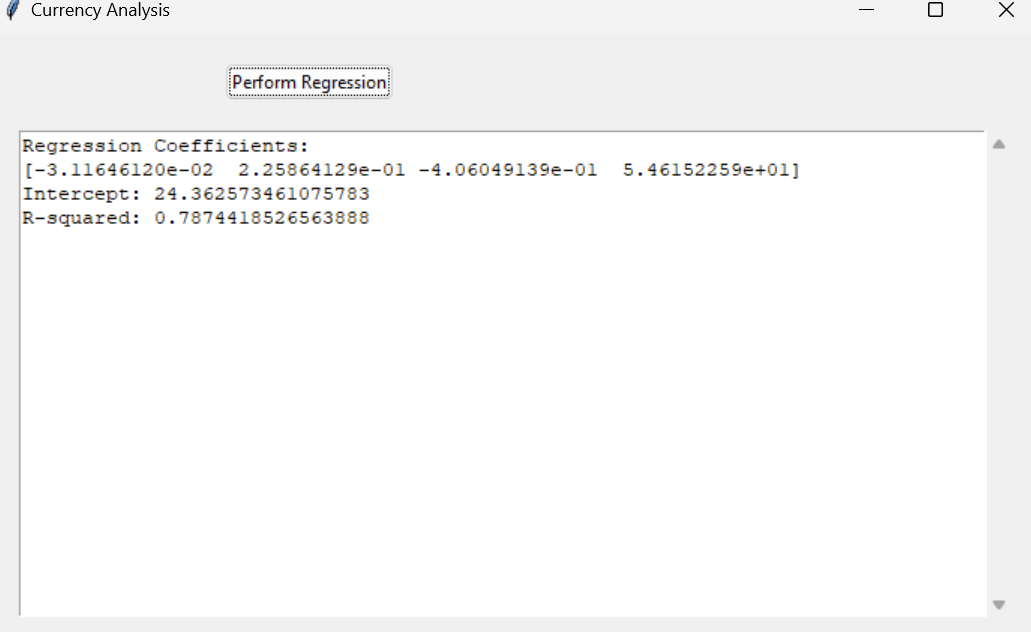
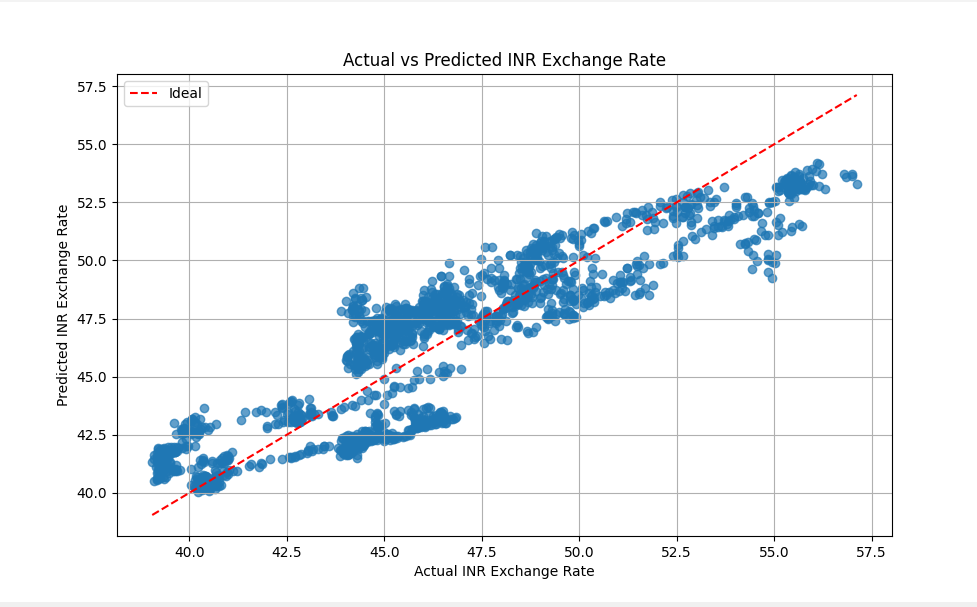
7. **Model Refinement**:

* 1. If necessary, refine the model by removing insignificant variables, adding new variables, or using regularization techniques to improve performance.

8**.Deployment**:

* 1. Once the model is satisfactory, it can be deployed to make real-time predictions or integrated into a larger financial analysis system.

By following these steps, we can effectively apply multiple linear regression to analyze and predict currency exchange rates, gaining insights into the factors that influence these rates and making informed decisions based on the model's predictions.



**INFERENCE:**

In summary, multiple linear regression analysis of currency exchange rates allows us to:

* Determine how economic factors like interest rates, inflation, and GDP affect exchange rates.
* Identify statistically significant variables influencing exchange rate movements.
* Assess model performance using metrics like R-squared, Mean Squared Error, and Root Mean Squared Error.
* Make informed decisions based on the model's insights, while acknowledging limitations like historical bias and multicollinearity.
* Continuously refine the model with additional variables or techniques to improve accuracy and relevance over time.

**2)Polynomial regression:**

**Definition:**

Polynomial regression is a type of regression analysis used to model the relationship between a dependent variable and one or more independent variables by fitting a polynomial equation to the data. Unlike simple linear regression, which assumes a linear relationship between the variables, polynomial regression allows for a curved relationship. Polynomial regression can capture complex relationships between variables that cannot be adequately modeled by linear regression. By increasing the degree of the polynomial (i.e., the highest power of *𝑋X*), the model can accommodate more intricate patterns in the data. However, higher-degree polynomials can lead to overfitting, so careful model selection and evaluation are necessary.

Polynomial regression is widely used in various fields, including economics, engineering, and natural sciences, where nonlinear relationships are common. It provides a flexible framework for analyzing data and making predictions beyond the constraints of linear models.

**MODEL OF POLYNOMIAL REGRESSION:**

In Python, you can use libraries like NumPy and scikit-learn to perform polynomial regression for currency exchange rate analysis. Below is an example of how to implement polynomial regression using scikit-learn

1)Load the currency exchange rate data.

2)Define the independent and dependent variables.

3)Split the data into training and testing sets.

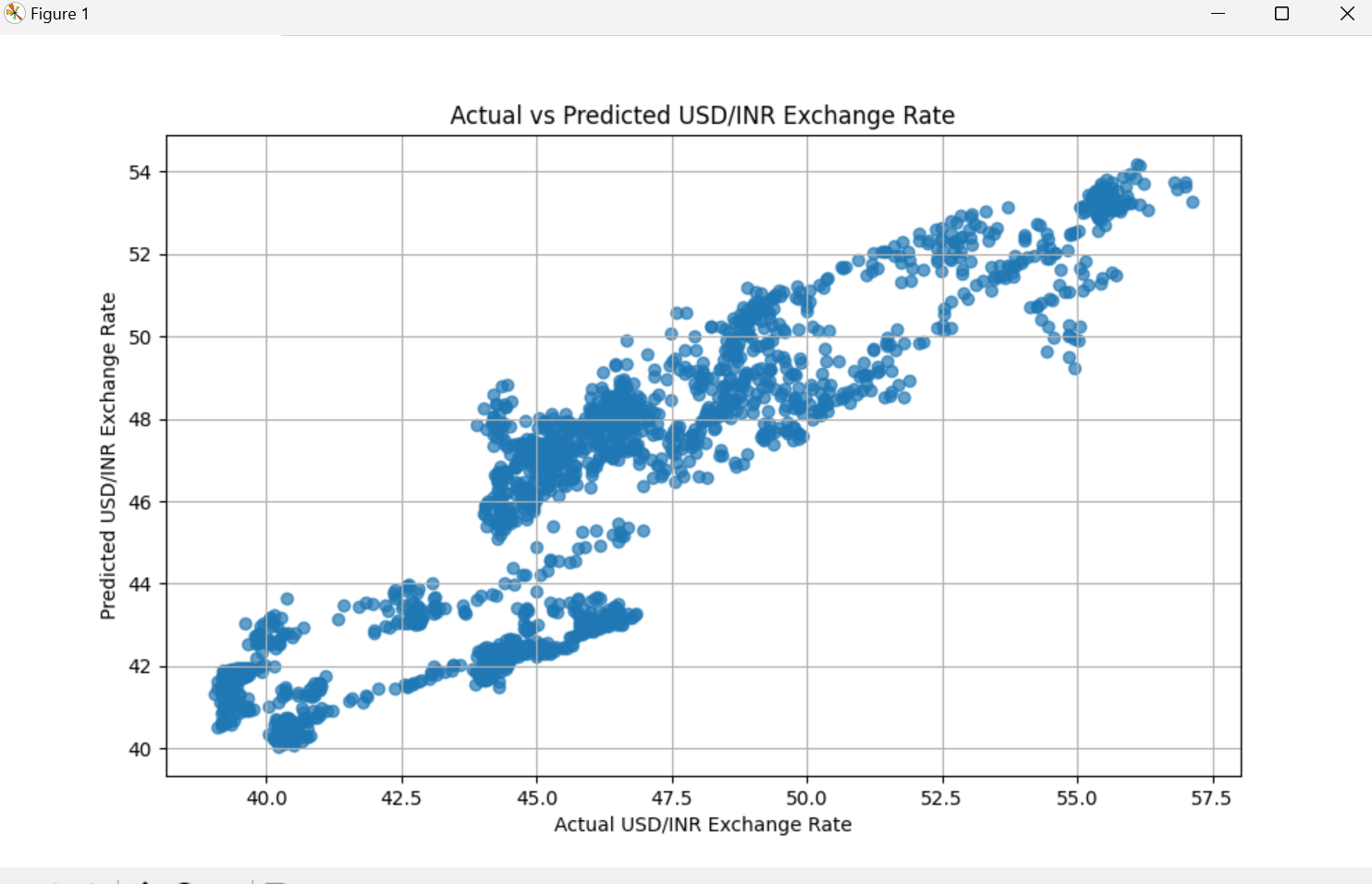
4)Use **PolynomialFeatures** from scikit-learn to transform the independent variable into polynomial features of a chosen degree.

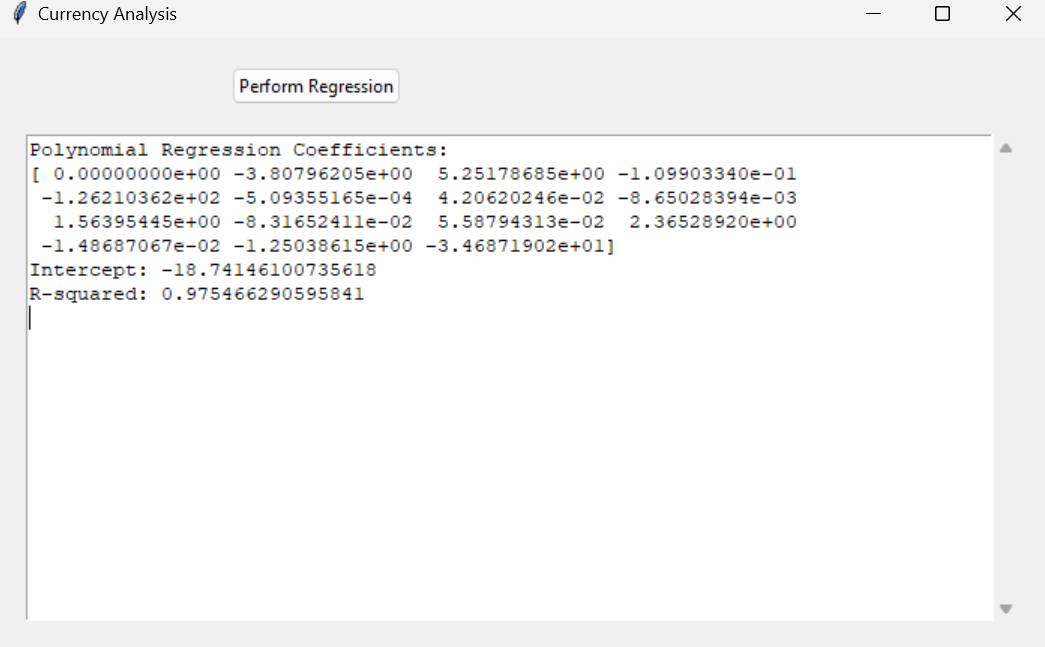
5)Fit a linear regression model to the polynomial features.

6)Make predictions on the test set.

7)Evaluate the model using metrics like Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R-squared.

8)Adjust the degree parameter in **PolynomialFeatures(degree=3)** to specify the degree of the polynomial you want to fit to your data. This controls the complexity of the polynomial model.





**INFERENCE:**

Polynomial regression for currency exchange rate analysis in Python allows us to:

* Capture nonlinear relationships between variables.
* Transform input features into polynomial terms of a chosen degree.
* Fit a linear regression model to the transformed features.
* Predict exchange rates and evaluate model performance using standard metrics like MSE, RMSE, and R-squared.
* Choose the polynomial degree carefully to balance model complexity and overfitting.

**TESTING:**

**1)F-TEST:**

The F-test is a statistical test used to determine whether there are significant differences between the variances of two or more groups. It is commonly used in the context of regression analysis and analysis of variance (ANOVA).

In the context of regression analysis, the F-test evaluates the overall significance of the regression model. Specifically, it tests the null hypothesis that all regression coefficients (except the intercept) are equal to zero, meaning that the independent variables do not explain the variability in the dependent variable. The alternative hypothesis is that at least one of the coefficients is not zero, indicating that the independent variables as a group do explain some of the variability in the dependent variable.

**MODEL USED IN F-TEST:**

**1) Import Libraries**: Import necessary libraries such as **pandas** for data handling and **statsmodels** for statistical modeling.

**2)Load and Prepare Data**: Load the currency exchange rate dataset and prepare the independent and dependent variables.

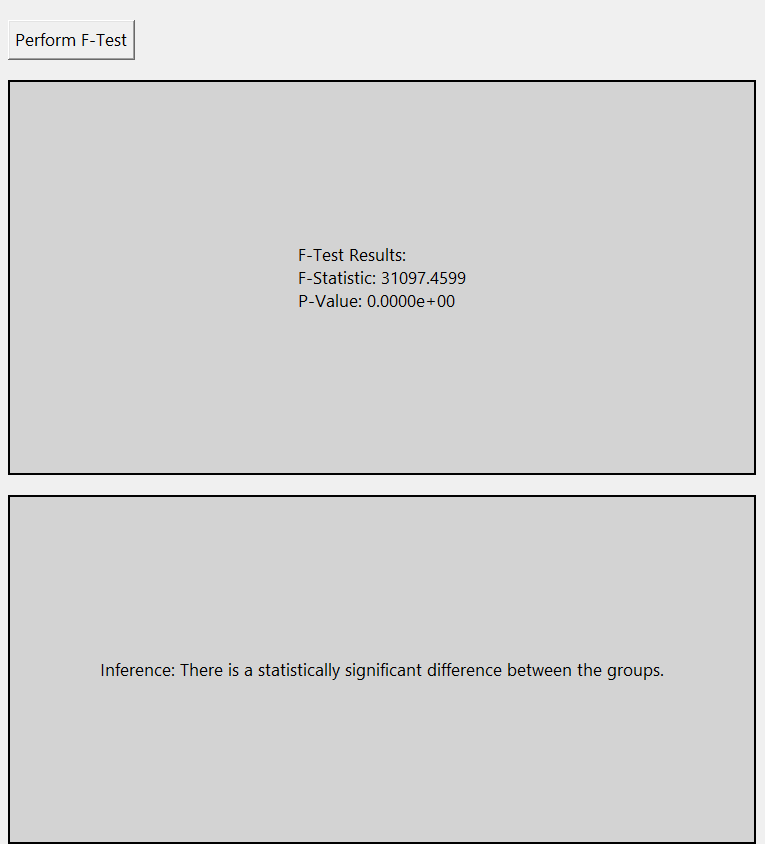
**3)Fit the Regression Model**: Use **statsmodels** to fit an Ordinary Least Squares (OLS) regression model.

**4)Perform the F-Test**: Extract the F-statistic and its associated p-value from the fitted model to evaluate the overall significance.

* **INTERPRETATION:**

**F-Statistic**:

* This value measures how well the independent variables explain the variability in the dependent variable (exchange rate). A higher F-statistic indicates that the model explains a significant portion of the variance.
* **P-Value**: The p-value associated with the F-statistic helps determine the statistical significance. If the p-value is less than a chosen significance level (e.g., 0.05), you reject the null hypothesis, concluding that at least one of the independent variables significantly affects the exchange rate.



**INFERENCE:**

Based on the F-test results:

* If the **F-statistic** is high and the **p-value** is low (typically below 0.05), it indicates that the regression model is statistically significant. This means that the set of independent variables (economic factors) collectively have a significant impact on the currency exchange rate.
* If the **p-value** is high (above 0.05), it suggests that the independent variables do not collectively explain the variation in the exchange rate, and the model may not be useful for predicting the exchange rate.

This process helps determine the overall effectiveness of the regression model in explaining currency exchange rate fluctuations and aids in identifying key economic indicators that influence exchange rates.

**2)T-TEST:**

**DEFINITION:**

The t-test is a statistical test used to determine whether there is a significant difference between the means of two groups. It helps to ascertain if the differences observed between sample data and a population parameter, or between two sample groups, are statistically significant or if they could have occurred by chance.

**MODEL USED IN T-test:**

In the context of analyzing currency exchange rates, a t-test can be used to compare the means of exchange rates under different conditions or time periods. Here, I'll demonstrate how to use a two-sample t-test to compare the average exchange rates between two different years.

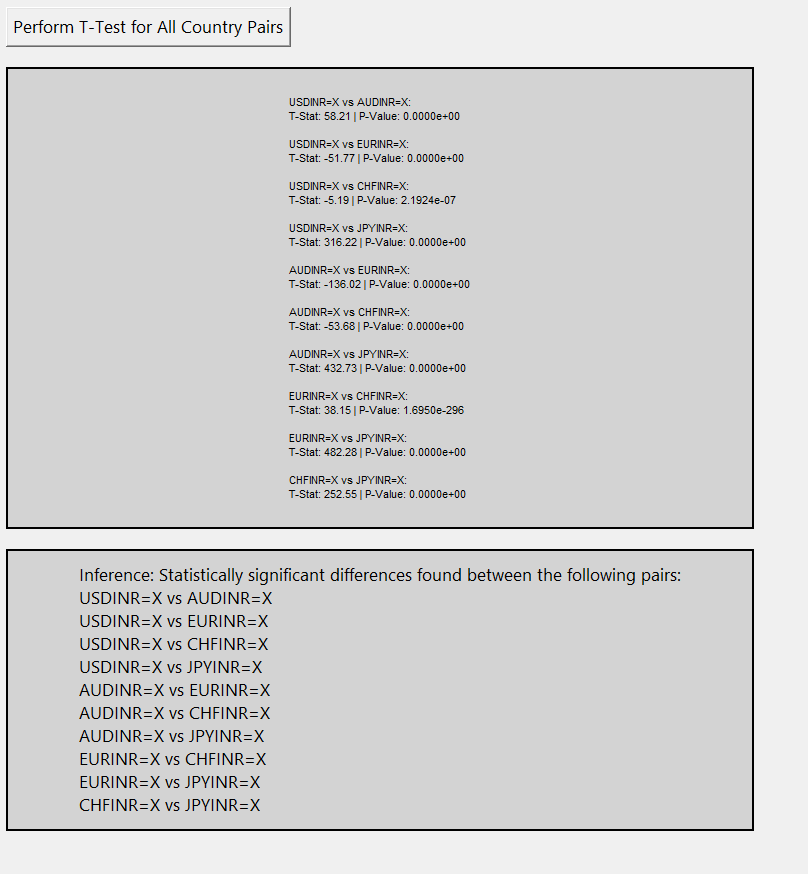
**1)Import Libraries**: Use libraries like **pandas** for data handling and **scipy.stats** for statistical testing.

**2)Load and Prepare Data**: Load the exchange rate data and prepare the data sets for the two groups you want to compare.

**3)Perform the T-Test**: Use the **ttest\_ind** function from **scipy.stats** to conduct the two-sample t-test.

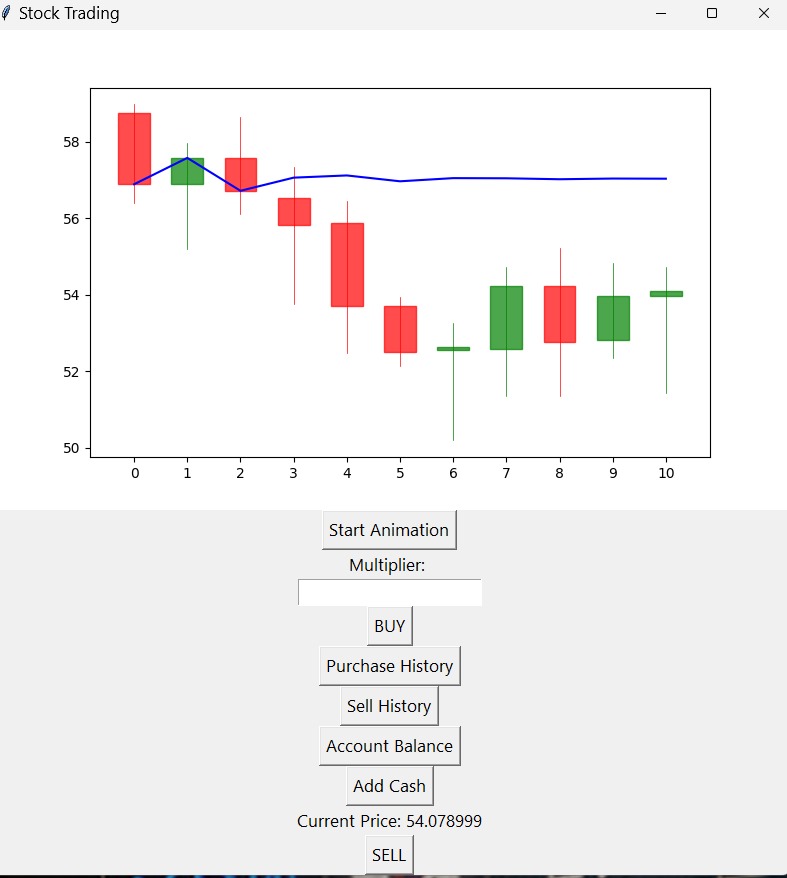
**INTERPRETATION:**

* **t-Statistic**: This value tells us how many standard deviations the sample mean of the exchange rate for the first year is from the sample mean of the second year. A larger absolute value indicates a larger difference between the groups.
* **p-Value**: This value helps determine the statistical significance of the observed difference. A p-value less than a chosen significance level (e.g., 0.05) indicates that the difference in means is statistically significant.

****

**INFERENCE:**

.

* if the **p-value** is high (above 0.05), it indicates that the observed difference in average exchange rates is not statistically significant, and any differences FOREX:

**GUI SCREESHOTS:**



