MDS531A: Econometrics

LAB 2: Data Collection and Analysis Task

Name: Neelanjan Dutta

Register number: 2448040

Tool used:

EViews (Econometric Views) is a powerful statistical and econometric software tool widely used for time series, cross-section, and panel data analysis. It offers an intuitive graphical user interface along with advanced features for statistical modeling, forecasting, and hypothesis testing. EViews is especially popular in economics, finance, and social sciences due to its ability to handle large datasets, perform regression analysis, and generate high-quality graphs.

The software allows users to import data from various sources, perform descriptive and inferential statistical analyses, compute correlation and covariance matrices, estimate econometric models, and interpret relationships between variables with precision. With EViews, researchers and analysts can efficiently explore datasets, test assumptions, and build models for forecasting and decision-making.

DATASET USED:

Dataset 1: Nifty 50

Link: https://www.nseindia.com/reports-indices-historical-index-data

Dataset 2: OIL

Link: https://in.investing.com/commodities/crude-oil

Dataset 3: USD_INR

Link: https://www.rbi.org.in/scripts/ReferenceRateArchive.aspx

Dataset 4: Any Individual Stock [INFOSYS]

Link: https://www.nseindia.com/get-quotes/equity?symbol=INFY

Objective:

The primary objective of this report is to empirically analyze the relationship between the Indian stock market and key global macroeconomic variables. Specifically, this study investigates the impact of daily returns of **WTI Crude Oil prices** and the **USD/INR** exchange rate on the daily returns of both the broader market, represented by the **Nifty 50 index**, and a key individual stock, of **Infosys**. The analysis covers the financial year from **April 1, 2024, to March 31, 2025**, and employs an Ordinary Least Squares (OLS) regression model to quantify the direction, magnitude, and statistical significance of these relationships after ensuring the stationarity of the data.

1. DATA PREPARATION

Q. Download and align the four series on common trading dates.

Ans. To align the three datasets, we had to use the **VLOOKUP function**. The formula is: **=VLOOKUP (A2, NIFTY50!A:E, 5, FALSE)** for the first column;

=VLOOKUP (A2, OIL!A:B, 2, FALSE) for the second column; **=VLOOKUP (A2, USD_INR!A:B, 2, FALSE)** for the third column; and **=VLOOKUP (A2, Infosys!A:H, 8, FALSE)** for the fourth column. There were some N/A values which was removed using the filter function.

Date	NIFTY50_CLOSE	OIL_PRICE	USD_INR_RATE	INFOSYS_CLOSE
02-Apr-24	22453.3	84.22	83.3585	1482.85
03-Apr-24	22434.65	84.61	83.4204	1480.65
04-Apr-24	22514.65	85.81	83.4475	1486.7
05-Apr-24	22513.7	86.1	83.4073	1479.1
08-Apr-24	22666.3	85.53	83.3176	1476.7
10-Apr-24	22753.8	85.44	83.2253	1506.8
12-Apr-24	22519.4	85.08	83.3931	1484.75
15-Apr-24	22272.5	84.86	83.4422	1468.15
16-Apr-24	22147.9	84.83	83.4982	1414.45
18-Apr-24	21995.85	82.1	83.517	1419.25

Dataset Uploading / Importing:

The dataset containing the observations was imported into **EViews** from an Excel file named Dataset.csv. The file was opened by selecting **File → Open → Foreign Data as Workfile**, and the CSV file was chosen from the system directory.

Descriptive Statistics:

Descriptive statistics were computed by selecting **Quick→ Group Statistics → Descriptive Statistics → Individual Samples** after opening the dataset in EViews. The analysis provides summary measures such as the mean, standard deviation, minimum, and maximum values.

Table 1: Descriptive statistics for prices

Group: UNTITLED Workfile: DATASET::Dataset\

View | Proc | Object | Print | Name | Freeze | Sample | Sheet | Stats | Spec

	INFOSYS_...	NIFTY50_C...	OIL_PRICE	USD_INR_...
Mean	1748.977	23796.49	73.87371	84.54765
Median	1829.750	23750.20	72.98000	83.96710
Maximum	1999.700	26216.05	86.10000	87.58580
Minimum	1393.650	21884.50	65.14000	83.06590
Std. Dev.	187.6745	1018.054	5.173782	1.288119
Skewness	-0.609358	0.067782	0.473721	1.002946
Kurtosis	1.845257	2.112761	2.183787	2.556947
Jarque-Bera	27.83468	7.955015	15.44304	41.67148
Probability	0.000001	0.018732	0.000443	0.000000
Sum	414507.6	5639769.	17508.07	20037.79
Sum Sq. Dev.	8312328.	2.45E+08	6317.253	391.5829
Observations	237	237	237	237

Table 2: Descriptive statistics for returns

Group: UNTITLED Workfile: DATASET::Dataset\

View | Proc | Object | Print | Name | Freeze | Sample | Sheet | Stats | Spec

	R_INFOSYS	R_NIFTY	R_OIL	R_USDINR
Mean	0.000244	0.000197	-0.000823	0.000112
Median	0.000154	4.79E-05	-0.000279	8.14E-05
Maximum	0.040826	0.033071	0.053043	0.006281
Minimum	-0.060383	-0.061124	-0.063203	-0.008021
Std. Dev.	0.015144	0.009026	0.017653	0.001394
Skewness	-0.394955	-0.968508	-0.337312	-0.219524
Kurtosis	4.587691	12.49312	3.725223	10.55744
Jarque-Bera	30.92311	923.0679	9.647148	563.5254
Probability	0.000000	0.000000	0.008038	0.000000
Sum	0.057524	0.046386	-0.194122	0.026317
Sum Sq. Dev.	0.053898	0.019145	0.073231	0.000456
Observations	236	236	236	236

Q. Compute daily logarithmic returns for each series.

Ans. To compute logarithmic returns for each series, we use the **dlog** function

The dlog() function in EViews is a shortcut that stands for the **difference of the natural logarithm**. It's a powerful tool specifically designed for financial and economic time series analysis.

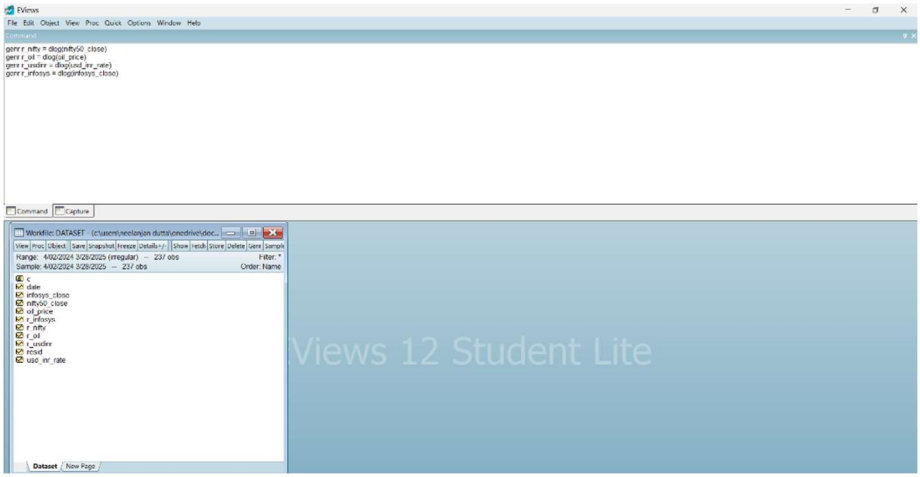


Fig 1: Logarithmic calculation for returns

It performs two mathematical operations in a single command:

1. **log**: It first calculates the natural logarithm of your series.
2. **d**: It then takes the first difference of that new log series (today's value minus yesterday's value).

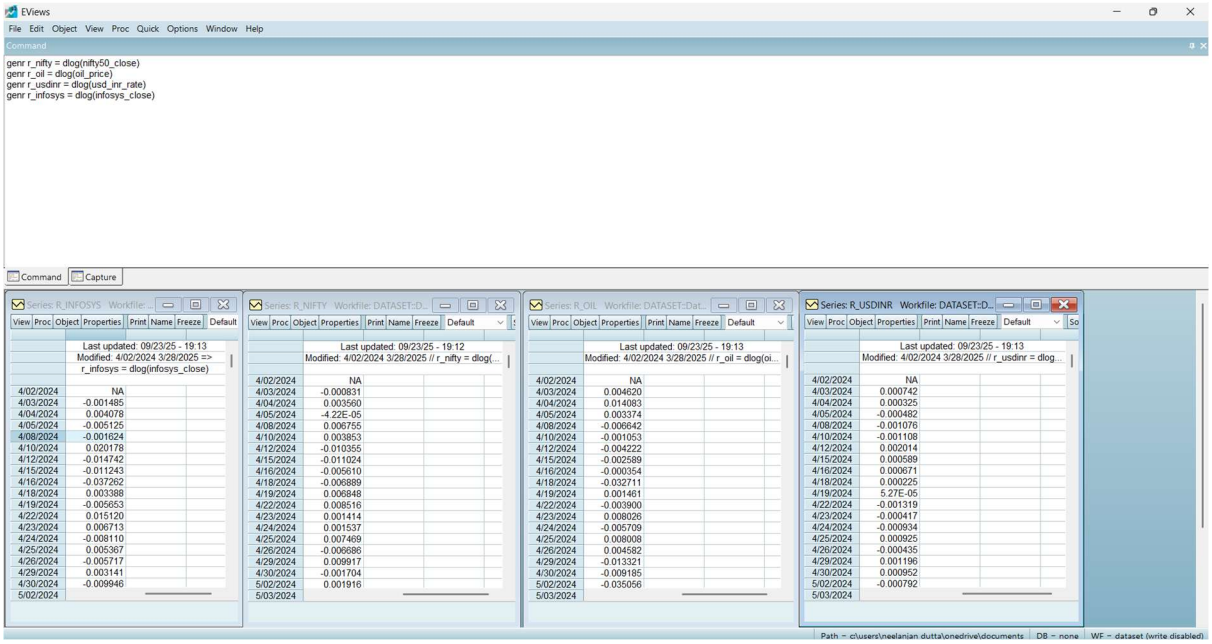


Fig 2: Visualizing the dataset for the returns

2. STATIONARY TESTING

Visual inspection:

Q. Plot the time series of both prices and returns to assess stationarity patterns.

Ans. <u>Plotting the price series</u>:

In our **EViews workfile window**, hold down the **Ctrl** key and click to select all four price series (nifty50_close, oil_price, usd_inr_rate, and Infosys_close). From the context menu, choose **Quick > Graph.**

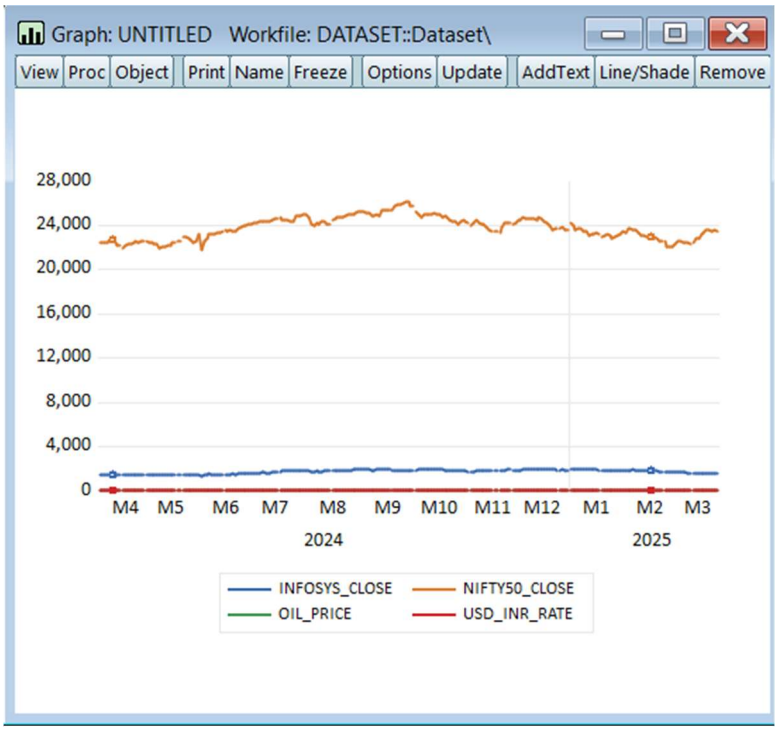


Fig 3: Time series plot for prices

<u>Interpretation:</u>

- **<u>Presence of Trends</u>**: As seen in the graph, none of the four series (NIFTY50_CLOSE, INFOSYS_CLOSE, OIL_PRICE, and USD_INR_RATE) fluctuates around a stable, constant average. Instead, they exhibit clear drifts and **"random walk"** behavior, meandering over time without returning to a central point. This is a classic visual sign of a non-stationary series.
- **<u>Changing Mean</u>**: The average value of each series is clearly different at the beginning of the period compared to the end.
- **<u>Unsuitability for Regression</u>**: The most important conclusion is that, in this raw price form, the data is not suitable for a reliable regression analysis. Using non-stationary data would likely lead to finding misleading or **"spurious"** relationships.

This plot perfectly demonstrates ***why*** we needed to calculate logarithmic returns before proceeding with your analysis.

<u>Plotting the return series</u>:

In our **EViews workfile window**, hold down the **Ctrl** key and click to select all three of your price series (r_nifty, r_oil, r_usdinr, r_infosys). From the context menu, choose **Quick > Graph.**

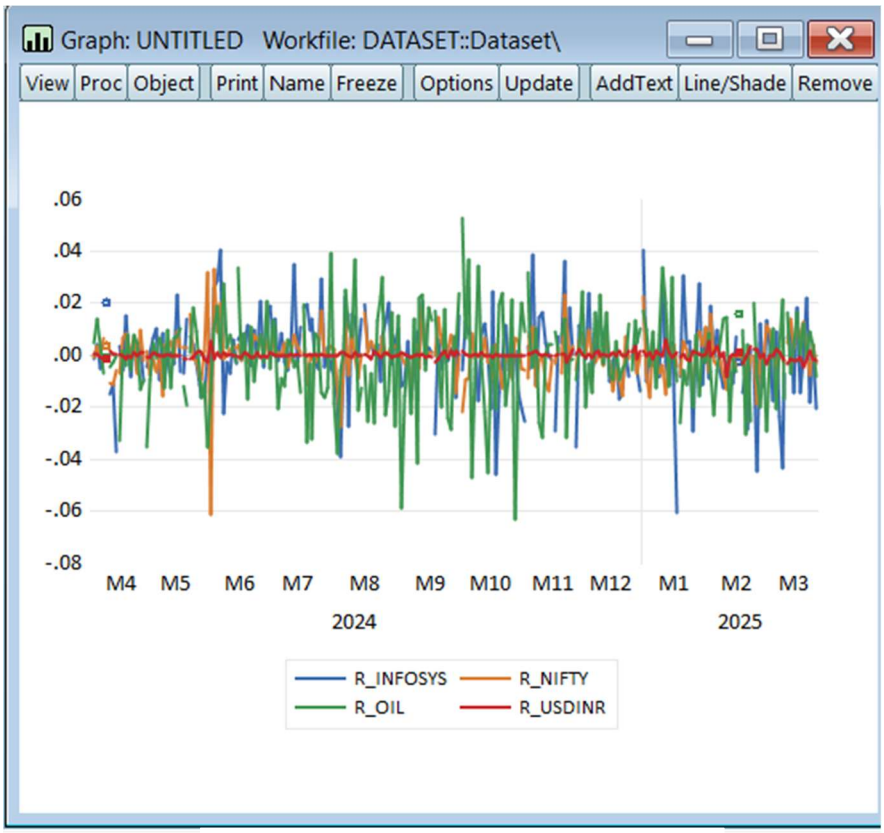


Fig 4: Time series plot for returns

Interpretation:

- **Mean Reversion**: Unlike the price series that drifted over time, all four of these return series (R_NIFTY50, R_INFOSYS, R_OIL, R_USDINR) fluctuate tightly around a constant central value of zero. There is no long-term upward or downward trend, which is the classic visual sign of a **stationary** series.
- **Constant Variance**: For the most part, the size of the fluctuations (the volatility) remains within a consistent band throughout the period, indicating that the variance is stable.
- **Suitability for Regression**: The most important conclusion from this plot is that our return series are visually confirmed to be stationary. This gives us the **"green light"** to use them in our OLS regression model, as we are now likely to find true economic relationships rather than spurious ones.

This visual check is a critical part of your analysis, and our graph clearly shows the desired outcome.

Statistical testing:

Q. Apply the Augmented Dickey–Fuller (ADF) unit root test to prices and returns separately. State the null and alternative hypotheses and conclude for each variable at the 5 % significance level.

Ans. The **Augmented Dickey-Fuller (ADF) test** is a formal statistical procedure used to test for the presence of a **"unit root"** in a time series. A unit root is the statistical property of a non-stationary series. In essence, the test helps us decide, with statistical confidence, whether our data is stationary or not.

For prices:

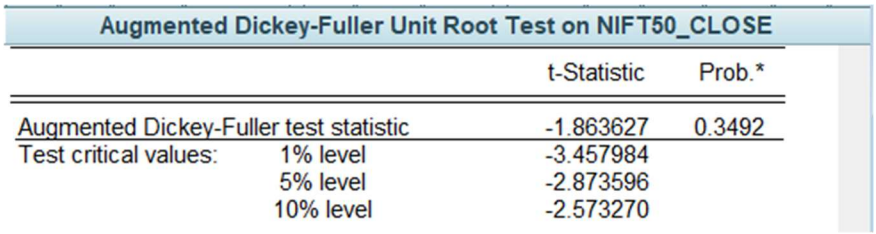
Let us set up the null hypothesis

H_0: The series has a unit root and is non-stationary, i.e., $\rho = 1$

Against the alternative hypothesis

H_1: The series does not have a unit root and is stationary, i.e., $\rho < 1$.

Result:



Augmented Dickey-Fuller Unit Root Test on NIFT50_CLOSE

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.863627	0.3492
Test critical values:	1% level	-3.457984	
	5% level	-2.873596	
	10% level	-2.573270	

Fig 5: ADF test value for prices [NIFTY50]

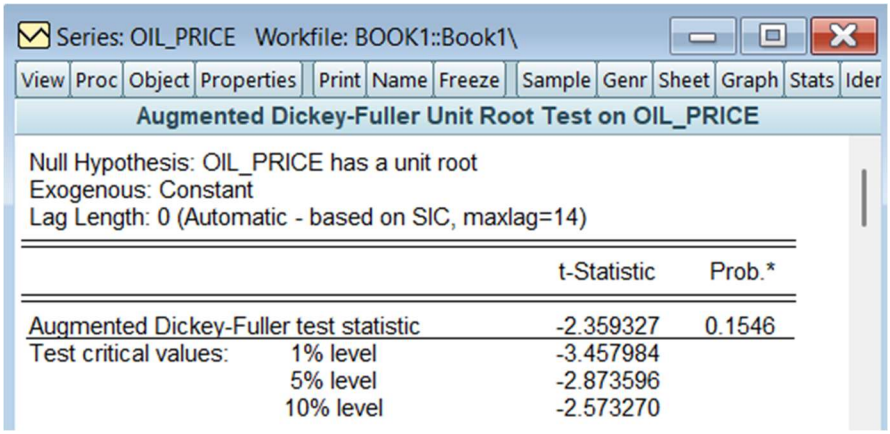
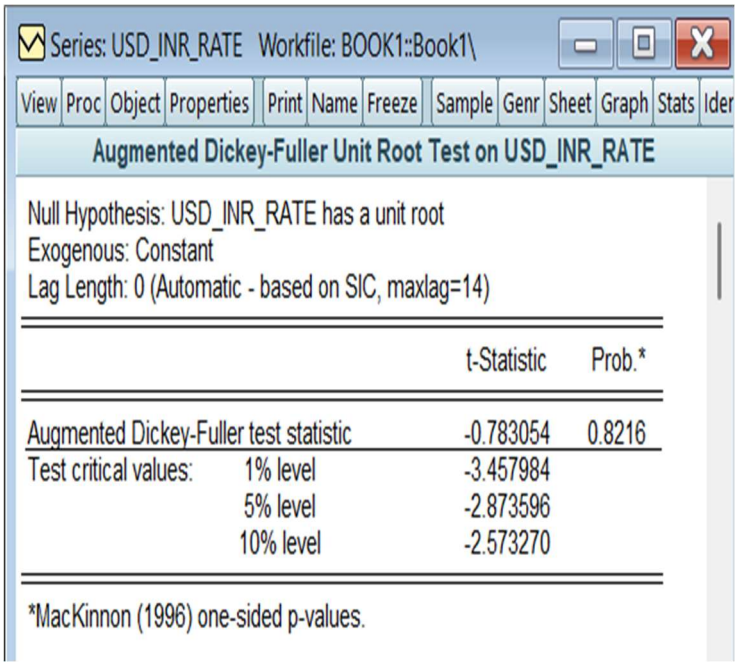


Fig 6: ADF test value for prices [OIL_PRICE]



Series: USD_INR_RATE Workfile: BOOK1::Book1\

View | Proc | Object | Properties | Print | Name | Freeze | Sample | Genr | Sheet | Graph | Stats | Ider

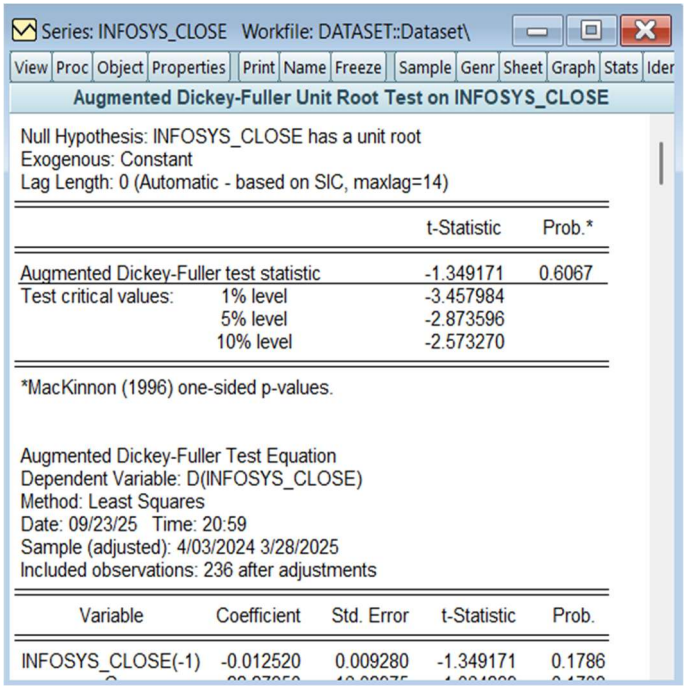
Augmented Dickey-Fuller Unit Root Test on USD_INR_RATE

Null Hypothesis: USD_INR_RATE has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-0.783054	0.8216
Test critical values:	1% level	-3.457984	
	5% level	-2.873596	
	10% level	-2.573270	

*MacKinnon (1996) one-sided p-values.

Fig 7: ADF test value for prices [USD_INR_RATE]



Series: INFOSYS_CLOSE Workfile: DATASET::Dataset\

View | Proc | Object | Properties | Print | Name | Freeze | Sample | Genr | Sheet | Graph | Stats | Ider

Augmented Dickey-Fuller Unit Root Test on INFOSYS_CLOSE

Null Hypothesis: INFOSYS_CLOSE has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-1.349171	0.6067
Test critical values:	1% level	-3.457984	
	5% level	-2.873596	
	10% level	-2.573270	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(INFOSYS_CLOSE)
Method: Least Squares
Date: 09/23/25 Time: 20:59
Sample (adjusted): 4/03/2024 3/28/2025
Included observations: 236 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
INFOSYS_CLOSE(-1)	-0.012520	0.009280	-1.349171	0.1786

Fig 8: ADF test value for prices [INFOSYS]

Interpretation:

We **fail to reject the null hypothesis** since the p-value is greater than 0.05, and conclude that all four NIFTY50_CLOSE, OIL_PRICE, USD/INR RATE, and INFOSYS_CLOSE have a **unit root** and are **non-stationary**.

For returns:

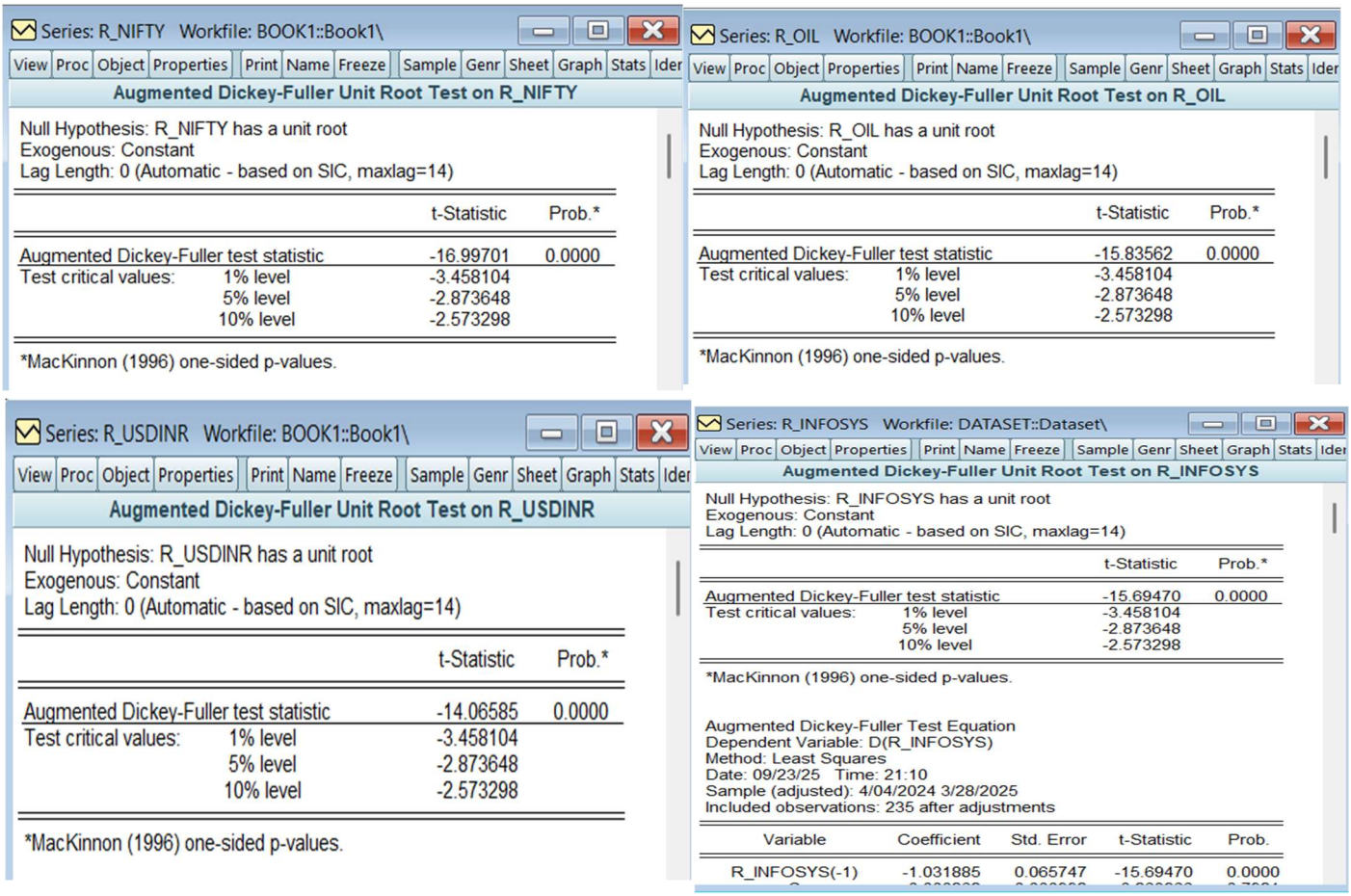
Let us set up the null hypothesis

H_0: The series has a unit root, i.e., $\rho = 1$

Against the alternative hypothesis

H_1: The series does not have a unit root, i.e., $\rho < 1$

Result:



Series: R_NIFTY Workfile: BOOK1::Book1\

Augmented Dickey-Fuller Unit Root Test on R_NIFTY

Null Hypothesis: R_NIFTY has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-16.99701	0.0000
Test critical values:	1% level	-3.458104	
	5% level	-2.873648	
	10% level	-2.573298	

*MacKinnon (1996) one-sided p-values.

Series: R_OIL Workfile: BOOK1::Book1\

Augmented Dickey-Fuller Unit Root Test on R_OIL

Null Hypothesis: R_OIL has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-15.83562	0.0000
Test critical values:	1% level	-3.458104	
	5% level	-2.873648	
	10% level	-2.573298	

*MacKinnon (1996) one-sided p-values.

Series: R_USDINR Workfile: BOOK1::Book1\

Augmented Dickey-Fuller Unit Root Test on R_USDINR

Null Hypothesis: R_USDINR has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-14.06585	0.0000
Test critical values:	1% level	-3.458104	
	5% level	-2.873648	
	10% level	-2.573298	

*MacKinnon (1996) one-sided p-values.

Series: R_INFOSYS Workfile: DATASET::Dataset\

Augmented Dickey-Fuller Unit Root Test on R_INFOSYS

Null Hypothesis: R_INFOSYS has a unit root
Exogenous: Constant
Lag Length: 0 (Automatic - based on SIC, maxlag=14)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-15.69470	0.0000
Test critical values:	1% level	-3.458104	
	5% level	-2.873648	
	10% level	-2.573298	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(R_INFOSYS)
Method: Least Squares
Date: 09/23/25 Time: 21:10
Sample (adjusted): 4/04/2024 3/28/2025
Included observations: 235 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
R_INFOSYS(-1)	-1.031885	0.065747	-15.69470	0.0000

Fig 9: ADF test value for returns [R_NIFTY50, R_OIL, R_USDINR, R_INFOSYS]

Interpretation:

We **reject the null hypothesis** since the p-value (0.0000) is less than 0.05 and conclude that all three return series, R_NIFTY, R_OIL, R_USDINR, and R_INFOSYS, do not have a **unit root** and are **stationary**.

3. Regression Analysis

Q. Estimate an ordinary least squares (OLS) model with Nifty 50 returns as the dependent variable and Crude Oil returns and USD/INR returns as independent variables.

Report coefficient estimates, standard errors, t-statistics, p-values, R^2, and diagnostic tests (e.g., Durbin–Watson, VIF).

Ans.

For the regression equation:

$R\_NIFTY = \beta_0 + \beta_1 R\_OIL + \beta_2 R\_USDINR + \beta_3 R\_INFOSYS + \epsilon$

Let us set up the null hypothesis

H_0: The independent variables Crude Oil returns, USD/INR returns, and Infosys returns have no significant impact on Nifty returns, i.e., $\beta_1 = \beta_2 = \beta_3 = 0$

Against the alternative hypothesis

H_1: At least one independent variable significantly impacts Nifty returns, i.e., at least one $\beta_i \neq 0$

A simple linear regression was performed using **Quick → Estimate Equation** in EViews. In the dialogue box that appears, we specified the regression equation by typing:

r_nifty c r_oil r_usdinr r_infosys

After entering this, we clicked **OK**, and EViews estimated the regression model using the **Ordinary Least Squares (OLS)** method.

For the Variance Inflation Factor, we click on **View → Coefficient Diagnostics→ Variance Inflation Factor**

Equation: UNTITLED Workfile: DATASET::Dataset\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Dependent Variable: R_NIFTY
Method: Least Squares
Date: 09/23/25 Time: 21:13
Sample (adjusted): 4/03/2024 3/28/2025
Included observations: 236 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.000330	0.000503	0.655317	0.5129
R_OIL	-0.009062	0.028458	-0.318418	0.7505
R_USDINR	-1.829813	0.360766	-5.072020	0.0000
R_INFOSYS	0.260021	0.033200	7.831924	0.0000

R-squared	0.282459	Mean dependent var	0.000197
Adjusted R-squared	0.273180	S.D. dependent var	0.009026
S.E. of regression	0.007695	Akaike info criterion	-6.879695
Sum squared resid	0.013737	Schwarz criterion	-6.820986
Log likelihood	815.8040	Hannan-Quinn criter.	-6.856028
F-statistic	30.44214	Durbin-Watson stat	2.270758
Prob(F-statistic)	0.000000		

Fig 10: Regression equation for R_Nifty50

Equation: UNTITLED Workfile: DATASET::Dataset\

View | Proc | Object | Print | Name | Freeze | Estimate | Forecast | Stats | Resids

Variance Inflation Factors
Date: 09/23/25 Time: 21:15
Sample: 4/02/2024 3/28/2025
Included observations: 236

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
C	2.53E-07	1.009283	NA
R_OIL	0.000810	1.003784	1.001600
R_USDINR	0.130152	1.009688	1.003238
R_INFOSYS	0.001102	1.003575	1.003314

Fig 11: VIF for R_Nifty50

Interpretation:

- **Crude Oil Returns (R_OIL)**
  - **Coefficient (-0.0091):** The relationship between crude oil returns and Nifty returns is negative but very close to zero. A **1% increase** in crude oil returns is associated with an expected **decrease of about 0.009% in Nifty returns**, holding other factors constant.
  - **Significance (p-value = 0.7505):** Since this p-value is greater than 0.05, the relationship is **not statistically significant**. We cannot conclude that crude oil price movements have a real impact on the returns of Infosys.
- **USD/INR Returns (R_USDINR)**
  - **Coefficient (- 1.8298):** There is a strong negative relationship between the USD/INR return and Nifty returns. A **1% increase** in USD/INR (i.e., rupee depreciation) is associated with an expected **1.83% decrease in Nifty returns**.
  - **Significance (p-value = 0.0000):** This result is **statistically significant** at the 1% level, indicating that currency fluctuations have a strong and reliable impact on Nifty returns.
- **INFOSYS Returns (R_INFOSYS)**
  - **Coefficient (0.2600):** There is a positive relationship between Infosys returns and Nifty returns. A **1% increase** in Infosys returns is associated with an expected **0.26% increase in Nifty returns**.
  - **Significance (p-value = 0.0000):** This relationship is **highly significant**, suggesting Infosys stock movements are a strong driver of overall Nifty performance.

Model Fit and Diagnostic Tests

- **R-squared (0.2825)**

- About 28.25% of the variation in Nifty returns is explained by crude oil, USD/INR, and Infosys returns. While not very high, this indicates the model captures some meaningful variation.

- **F Statistic (30.44, p-value = 0.0000)**
  - The model as a whole is statistically significant, confirming that at least one independent variable explains variations in Nifty returns.
- **Durbin-Watson Statistic (2.27)**
  - Close to 2, indicating no **serious autocorrelation** in residuals.
- **VIF (Variance Inflation Factors) [all <1.01]**
  - Variance Inflation Factors are very low, showing **no multicollinearity** issues, meaning the variables can be interpreted reliably.

Conclusion:

This report analysed the impact of daily WTI crude oil and USD/INR returns on both the Nifty 50 index and the individual stock, Infosys. After ensuring data stationarity through logarithmic returns, OLS regression models were estimated for both.

The analysis revealed that crude oil returns do not have a statistically significant effect on either the Nifty 50 or Infosys, suggesting that short-term oil price movements are not a meaningful driver of daily stock market fluctuations in this context. On the other hand, daily USD/INR returns were found to exert a **strong and statistically significant negative effect on the Nifty 50 index**, indicating that currency depreciation adversely impacts the broader market. However, the USD/INR effect was not significant for Infosys. Infosys returns, in contrast, displayed a **positive and highly significant influence on the Nifty 50**, highlighting the company's role as a key contributor to overall index performance.

Furthermore, the explanatory power of the models differed: the Nifty 50 regression achieved an R-squared of ~28%, indicating a moderate level of explanation, while the Infosys model explained only about 1.3% of the variation in its returns. In summary, while the broader Indian market shows significant sensitivity to currency fluctuations and the performance of large-cap stocks like

Infosys, crude oil price changes do not appear to be primary drivers of daily returns for either the market as a whole or for Infosys individually.

****