FORECAST EXCHANGE RATES

Lakshaya Maleti
QR2212

BUSINESS OBJECTIVE

- Data provided is related to USD/INR Exchange rates. The objective is to understand the underlying structure in the given dataset and come up with a suitable forecasting model which can effectively forecast USD/INR exchange rate for the next 30 days.

- This forecasting model will be used by exporting and importing companies to understand the currency movements and accordingly set their revenue expectations.

DATA

- The data was converted to .csv format.
- It had observation_date and Close columns.
- The observation date was converted to datetime format using “prase_date”

```
1 df.head()
```

	Close
Date	
2003-12-01	45.709999
2003-12-02	45.629002
2003-12-03	45.549999
2003-12-04	45.548000
2003-12-05	45.449001

```
1 df.tail()
```

	Close
Date	
2023-06-06	82.511200
2023-06-07	82.517197
2023-06-08	82.618103
2023-06-09	82.488297
2023-06-12	82.370003

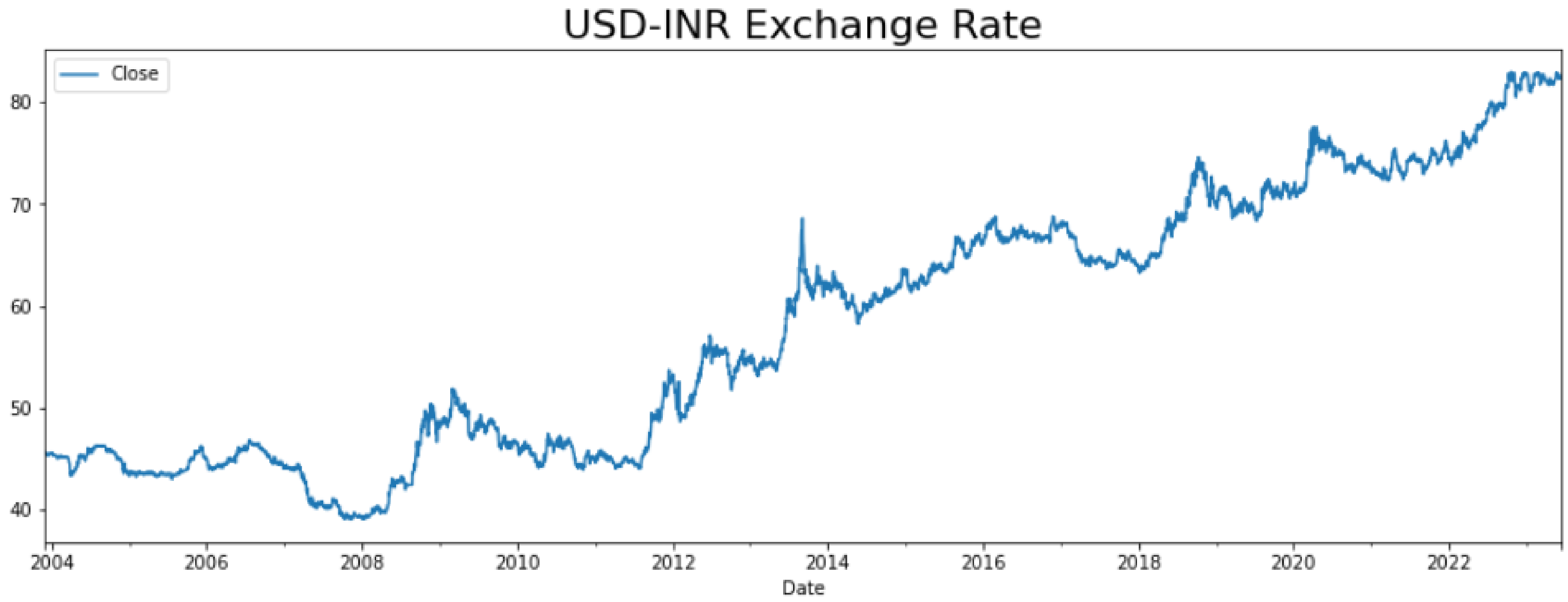
DATA PREPROCESSING (EDA)

- There were 5096 values in the data set.

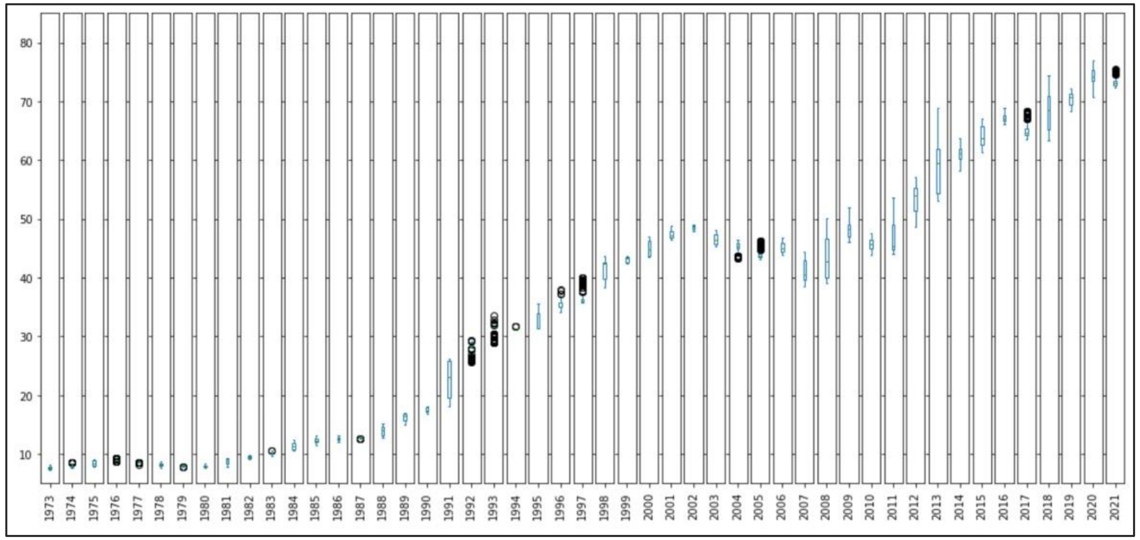
- The data also had multiple missing null values, which were 198, it's nearly 4% of the data, so we could not ignore it by dropping them. Instead We used Forward fill method to filling the missing values and then converted the rate into float.

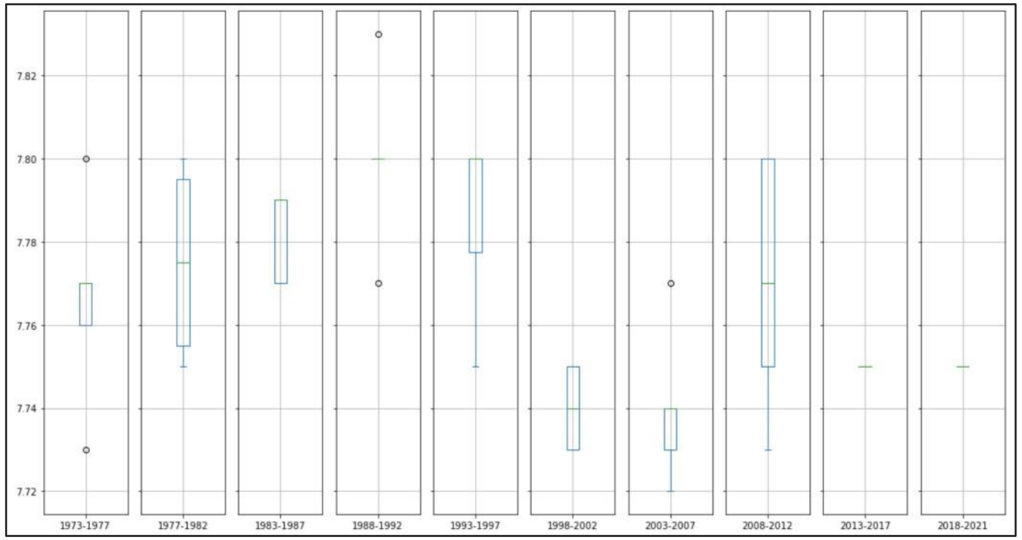
DATA VISUALIZATION

- Line plot of entire data



- Box plot of 5 years intervals

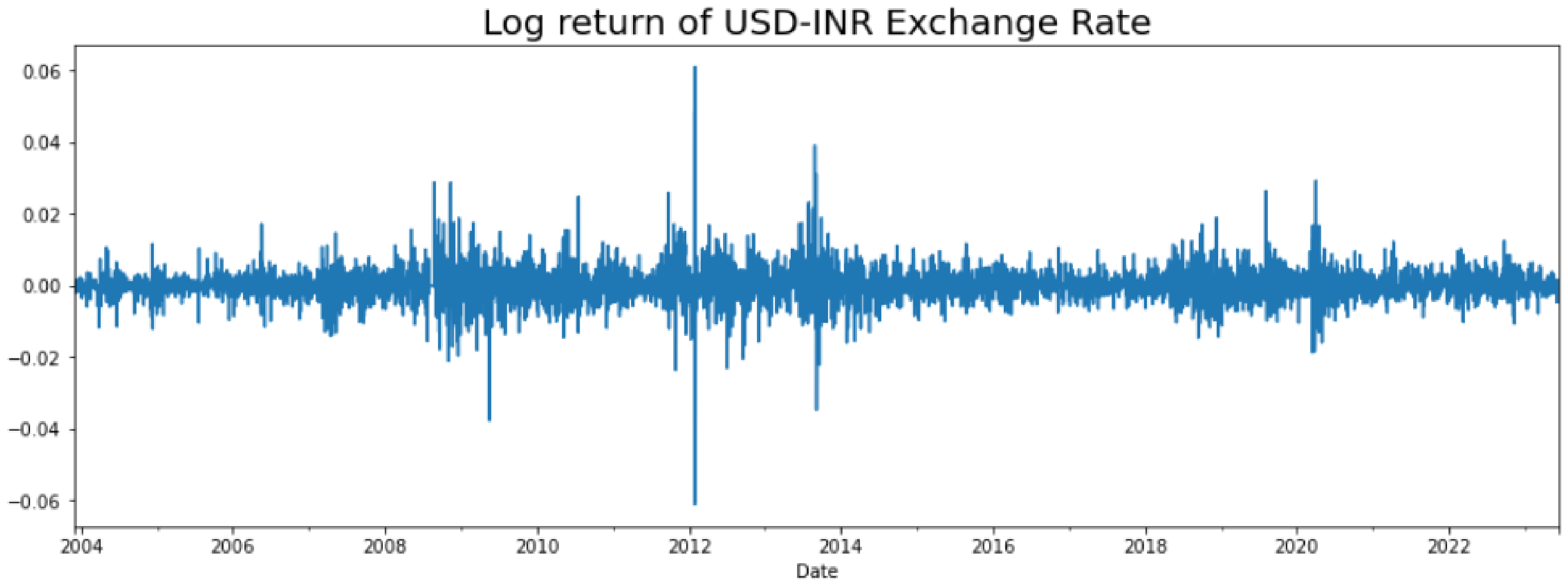




STATIONARITY

- Stationarity is an important concept in the field of time series analysis with tremendous influence on how the data is perceived and predicted.
- Since the data is not stationary we had to perform transformation use Difference and Logarithmic transformations.
- The stationarity test is conducted for confirming whether the data has unit root or not.

- H0 = The data has a unit root (In other words, the statistical properties of the time series remain constant over time.)



MODEL BUILDING

The Preprocessing and Stationarity test is completed and now the Time Series model building is the next step for forecasting results.

(AR) Autoregressive Process

These forecasts are based on a linear combination of past values of the variable to be forecasted. Given that εt is a purely random process with mean 0 and constant variance, and ERt is the nominal exchange rate at time'‘t’, the forecast of the exchange rate (ER), following an AR(p) process, at time ‘t+1’ is given by:

$ER(t) = c + \varphi 1*ER(t-1) + \varphi 2*ER(t-2) +\cdots+ \varphi p*ER(t-p) + \varepsilon(t)$, where εt is white noise.

(MA) Moving Average Process

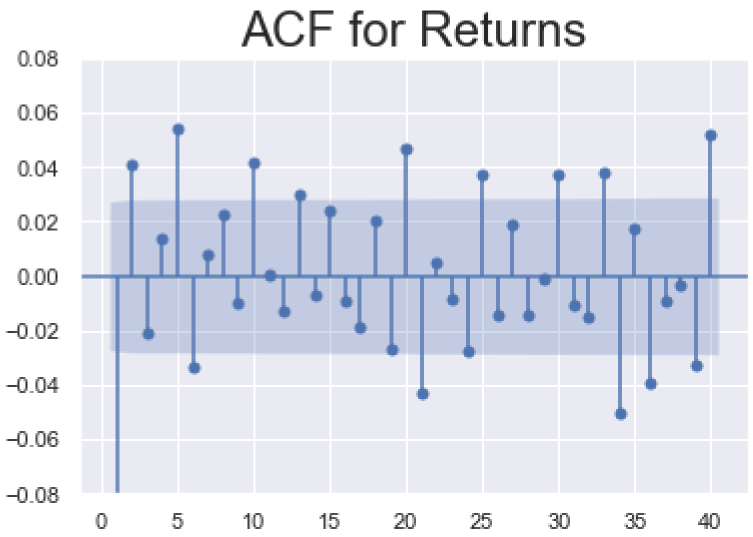
These models are based on the idea that current and past shocks systematically feed into the current value of the time series. An MA model is a regression model with nothing but current and lagged disturbances on the right-hand side. Given that εt is a purely random process with mean 0 and a constant variance, and ERt is the nominal exchange rate at time ‘t’, the forecast of the exchange rate (ER), following an MA(q) process, at time ‘t+1’ is given by:

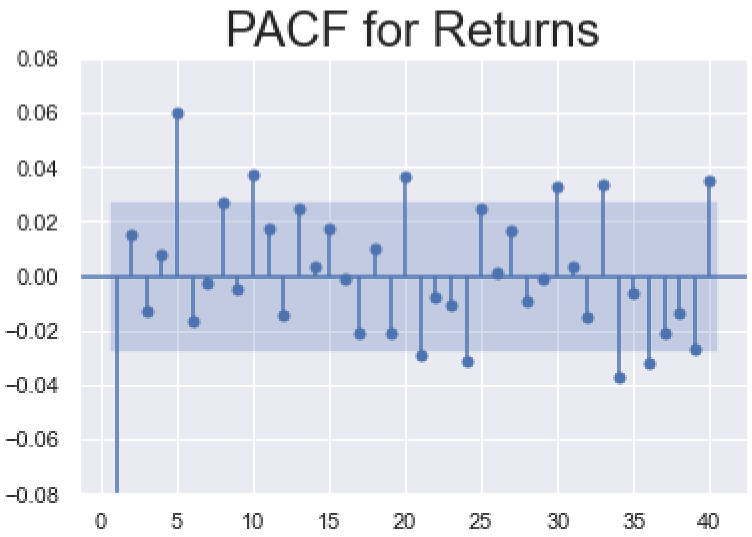
$ERt = c + \varepsilon t + \theta 1*\varepsilon(t-1) + \theta 2*\varepsilon(t-2) +\cdots+ \theta q*\varepsilon(t-q)$

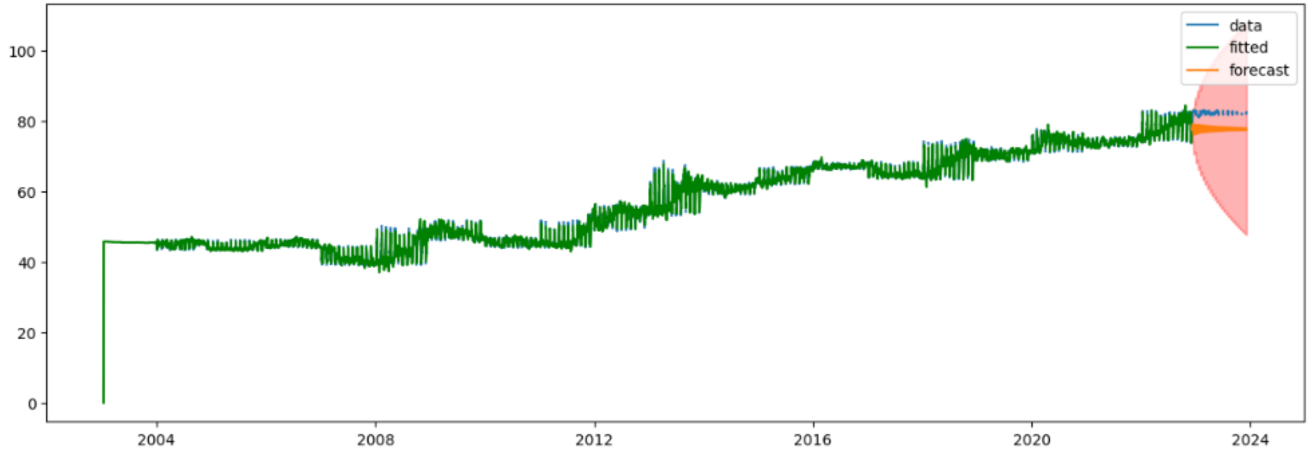
ARIMA MODEL

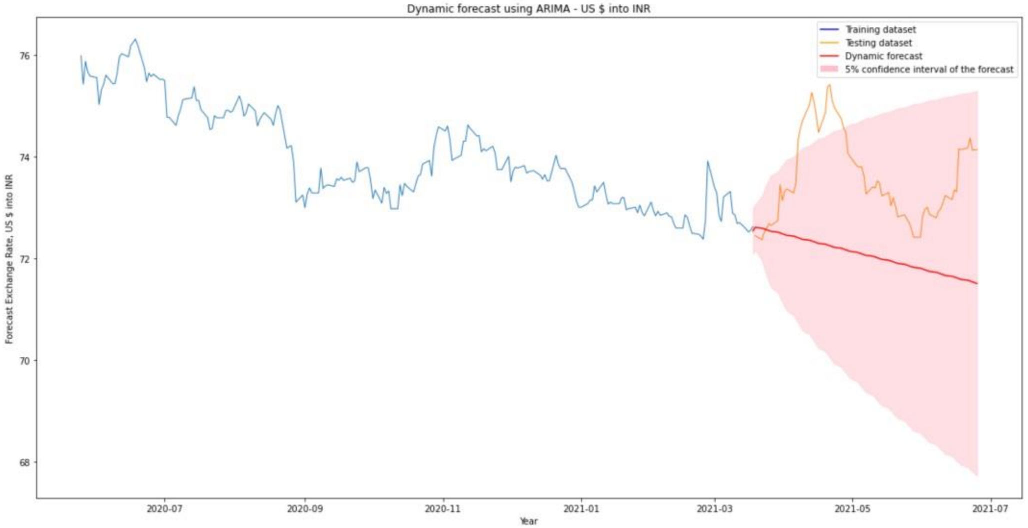
ARIMA models are a set of models that describe the process as a function of its own lags and white noise process (Box & Jenkins, 1974). A stochastic process (ERt) is called an ARIMA (p, d, q) process if it is integrated of order ‘d’ and the d times differenced process (ER’t) has an ARMA(p,q) representation, i.e.,

$ER'_t = c + \varphi_1 ER'_{t-1} + \cdots + \varphi_p ER'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$, where ER′ is the d-times differenced time series.









Further, Out of sample forecasting results were not as expected.

Information Criteria:

Utilize information criteria, such as the Akaike Information Criterion (AIC) or Bayesian Information Criterion (BIC), to compare different ARIMA models with varying values of p and q. Lower AIC or BIC values indicate a better fit for the model. Try several combinations of p and q and choose the model with the lowest AIC or BIC.

AIC = -2 * log-likelihood + 2 * k

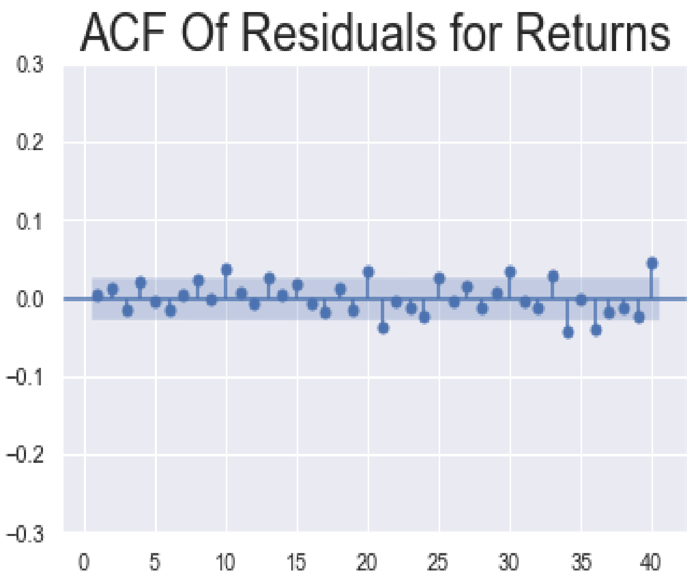
where:

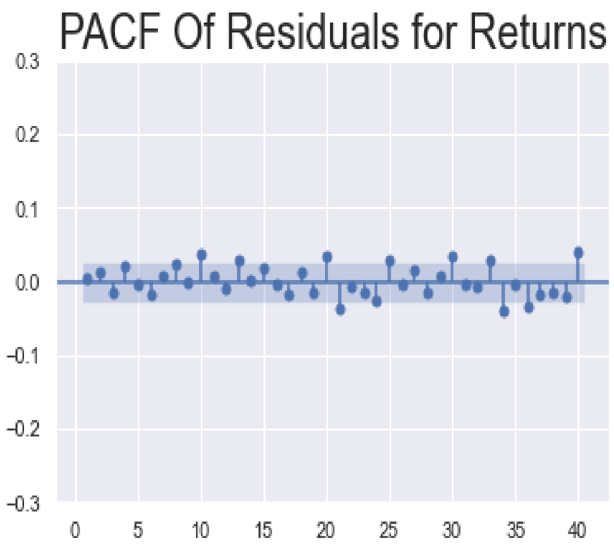
log-likelihood: The maximized log-likelihood of the model. It measures how well the model fits the data.

k: The number of parameters in the model. It represents the model's complexity, including both the number of regression coefficients and any estimated variance parameters.
The AIC penalizes models with more parameters, aiming to strike a balance between model fit and parsimony. Lower AIC values indicate better-fitting models with relatively fewer parameters, making them more favourable.

Bayesian Information Criterion (BIC):
The formula for BIC is as follows:
BIC = -2 * log-likelihood + k * log(n)
where:
Log-likelihood: The maximized log-likelihood of the model, same as in AIC.
n: The number of observations in the data.
The BIC is similar to AIC but penalizes models with more parameters more heavily, particularly for small sample sizes. It aims to prevent overfitting by providing a stronger penalty for additional parameters.

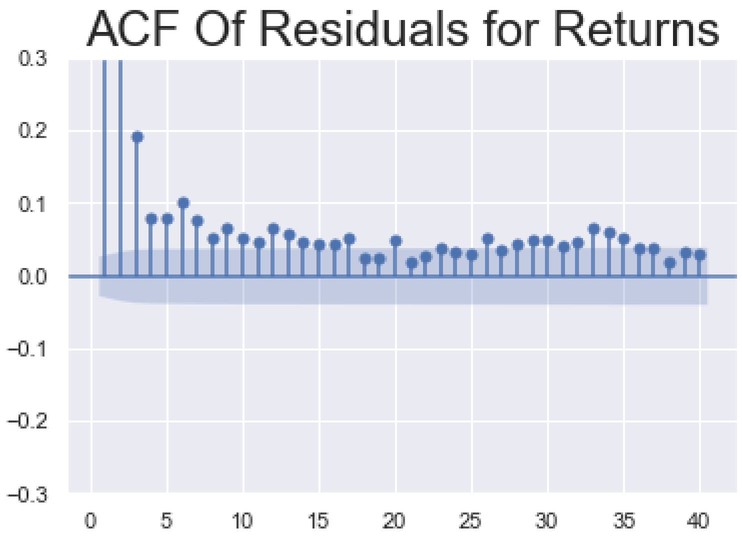
Residual analysis:

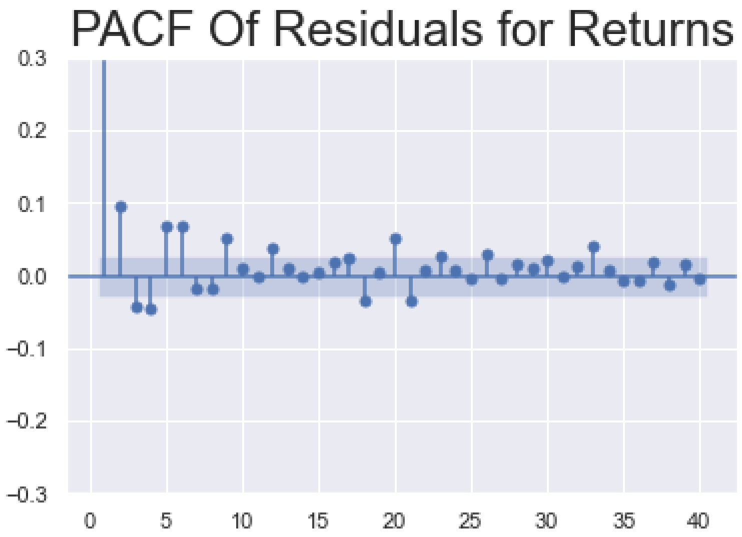




Therefore, It not having any sigificant lags upto 10^{th} lag, so is approximately follows white noice, So now there is no such pattern that is identifyable .

Now plotting Square of Residual:





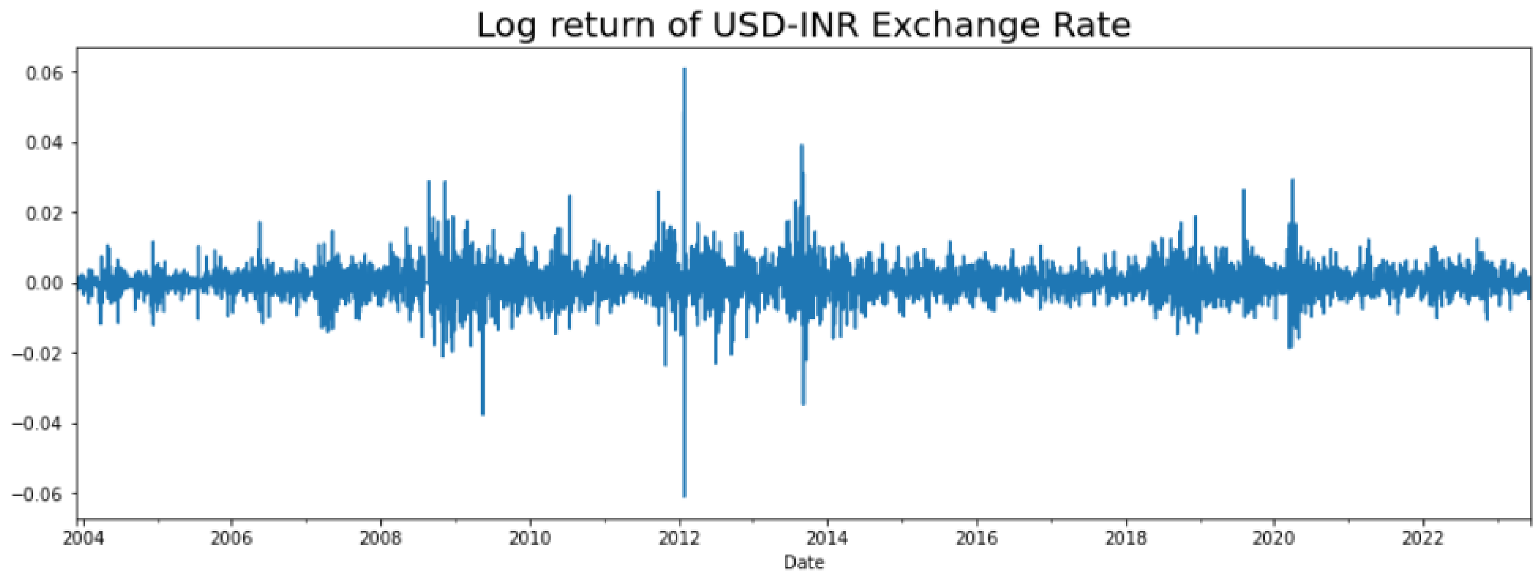
As, many significant lags are their so use ARCH/GARCH.

GARCH/ARCH :

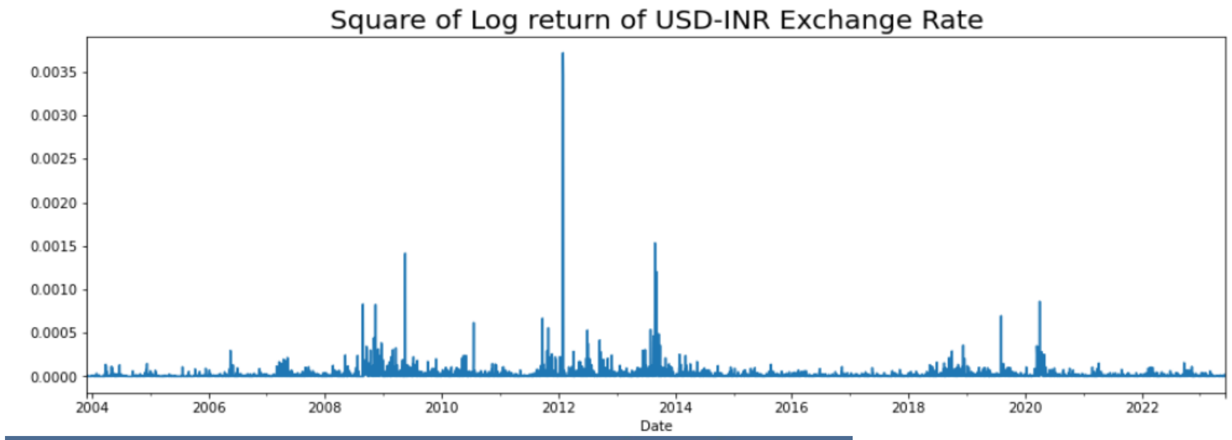
Volatility of returns refers to the degree of variation or fluctuation in the price or value over a specific period of time. It is a measure of the uncertainty or risk associated with the Exchange rate performance. Volatility is often quantified using statistical measures such as standard deviation or variance.

It is important for investors and traders to understand the volatility of an Exchange rate because it can affect investment decisions, risk management strategies, and the potential for both gains and losses.

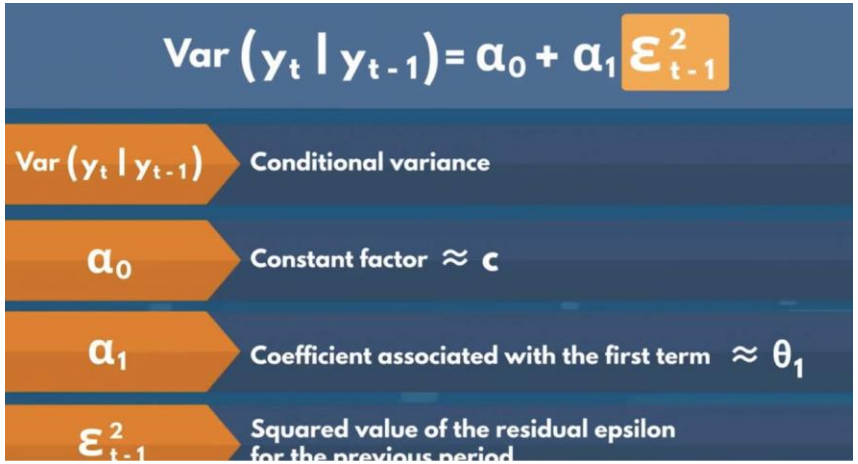
Volatility Modelling:



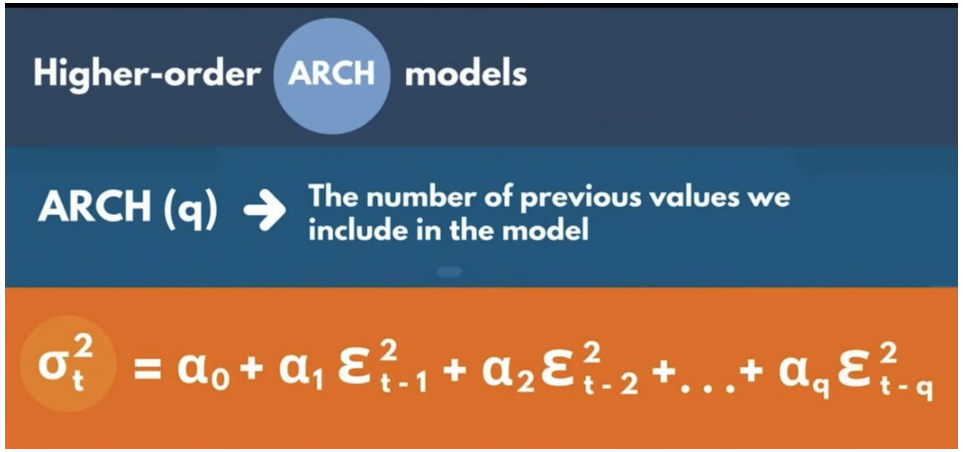
We see a lot of volatility clustering i.e. every time the return has a large magnitude, it surrounded by other return of large magnitude. Therefore we can model this on the basis of above observation.



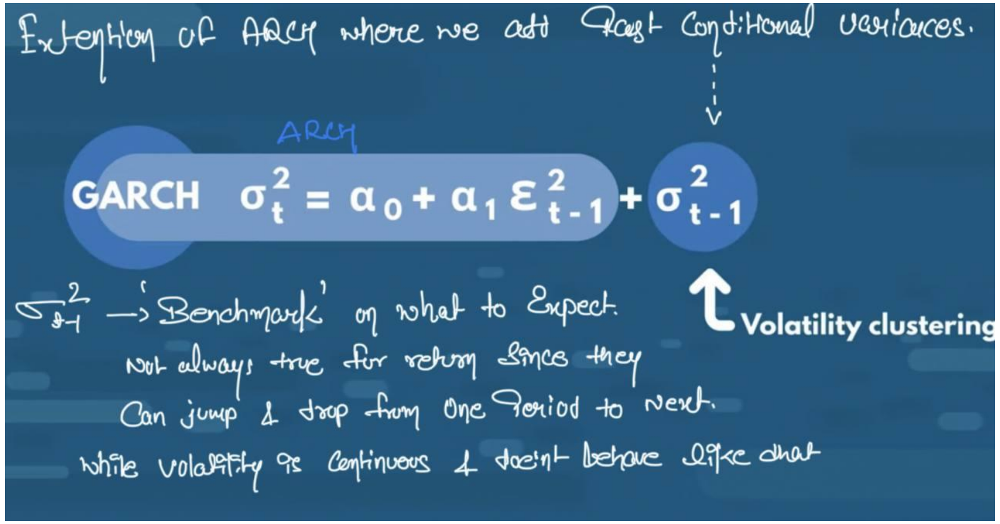
Square of Log return of USD-INR Exchange Rate
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022
Date

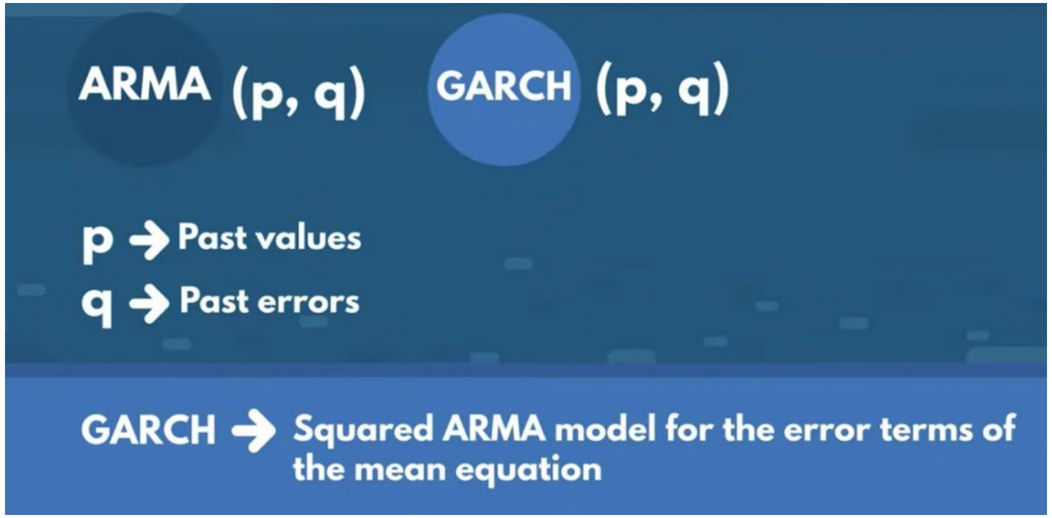


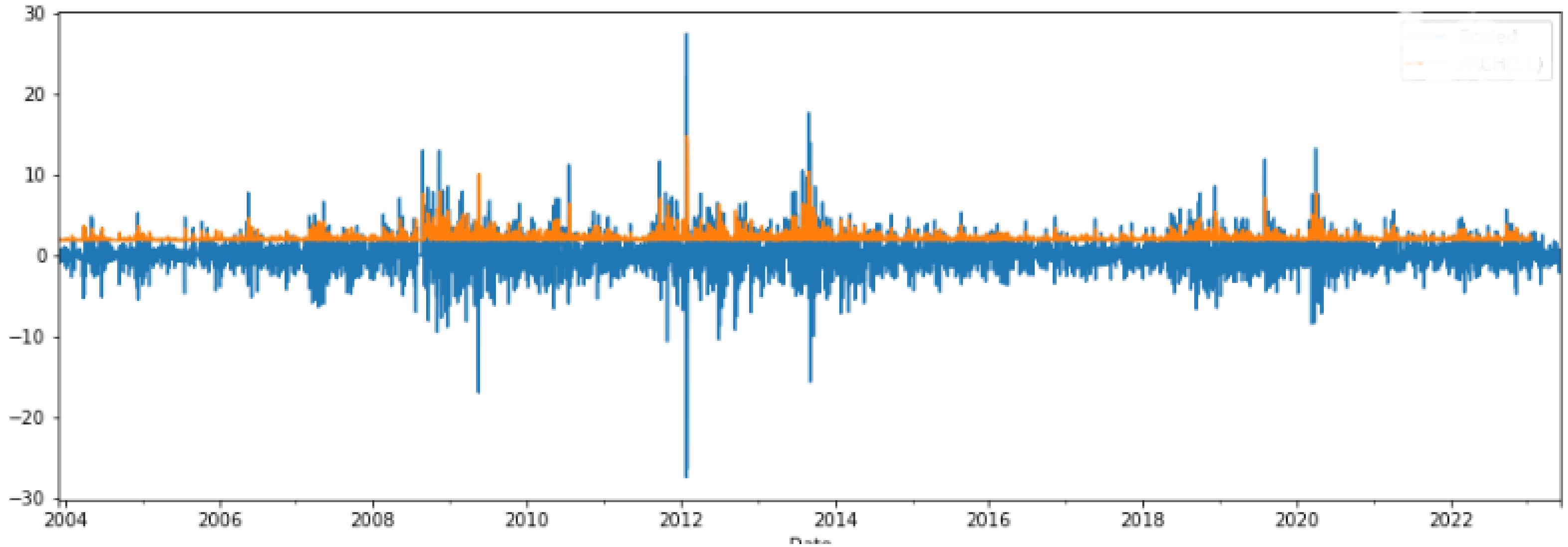
$Var(y_t \mid y_{t-1}) = \alpha_0 + \alpha_1 \varepsilon^2_{t-1}$
$Var(y_t \mid y_{t-1})$ Conditional variance
α_0 Constant factor $\approx c$
α_1 Coefficient associated with the first term $\approx \theta_1$
ε^2_{t-1} Squared value of the residual epsilon for the previous period

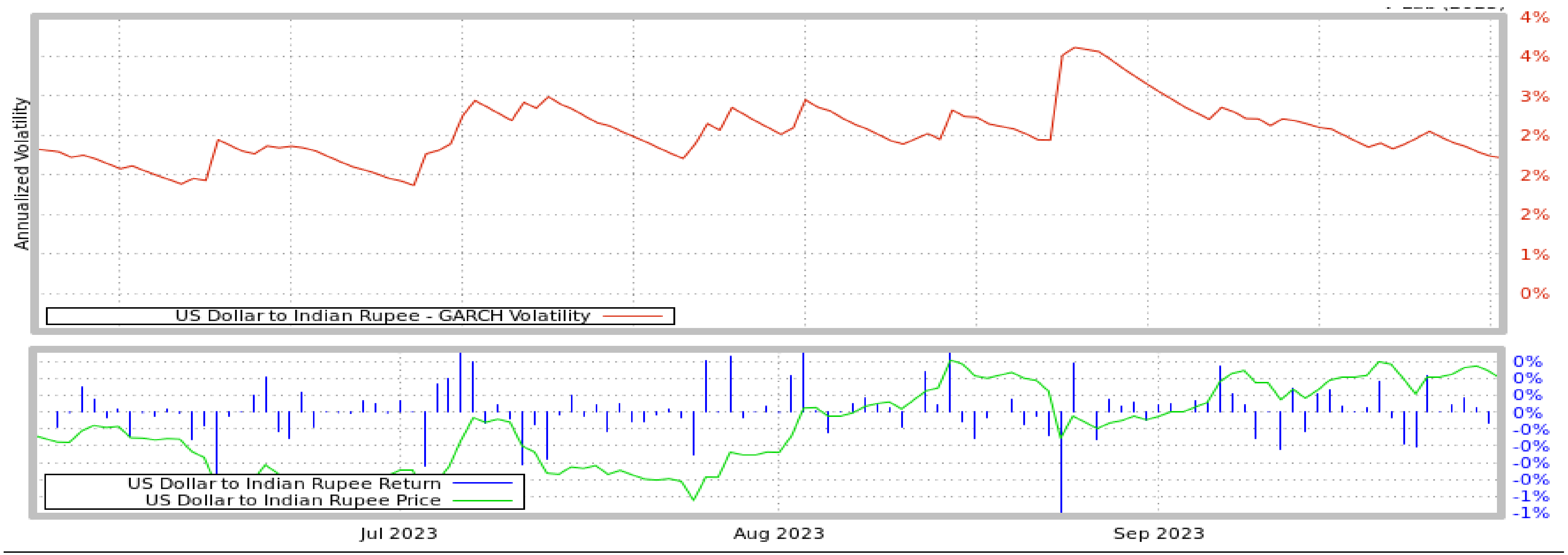


Higher-order ARCH models
ARCH (q) → The number of previous values we include in the model
$\sigma^2_t = \alpha_0 + \alpha_1 \varepsilon^2_{t-1} + \alpha_2 \varepsilon^2_{t-2} + \ldots + \alpha_q \varepsilon^2_{t-q}$


Extention of ARCH where we add Past Conditional variances.
ARCH
GARCH $\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \sigma_{t-1}^2$
σ_{t-1}^2 → 'Benchmark' on what to Expect.
Not always true for return since they
Can jump & drop from One Period to Next.
while volatility is Continuous & doesn't behave like that
Volatility clustering


ARMA (p, q)
GARCH (p, q)
p → Past values
q → Past errors
GARCH → Squared ARMA model for the error terms of the mean equation


30
20
10
0
−10
−20
−30
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022



Conclusion each of these models has its strengths and weaknesses, and their effectiveness depends on the specific characteristics of the time series data being analyzed. ARIMA is a good choice for linear and stationary data, while ARCH and GARCH models are valuable for capturing volatility patterns, particularly in financial time series. However, for more complex and nonlinear data, other advanced time series modeling approaches like state-space models, neural networks, or ensemble methods may be more appropriate. Additionally, proper model selection and parameter estimation are crucial for obtaining reliable and accurate forecasts with any of these models.

Thank you