# TIME SERIES FORECASTING

## ICICI Bank: Daily Stock Price & Returns


### Abstract

This project is based on the time series analysis of the stock under watchlist to forecast the price and returns as per the historical data.
For predicting and forecasting ARMA model is used for predicting and returns and ARIMA model for price forecasting for future prices.



SIDDHANT VISHWAKARMA
Third Year Undergraduate-Department of Humanities & Social Sciences, IIT Kharagpur

# Contents

# About Company: ICICI Bank

## Introduction

ICICI Bank is a leading private sector bank in India. The Bank's consolidated total assets stood at Rs. 14.76 trillion on September 30, 2020. ICICI Bank currently has a network of 5,288 branches and 15,158 ATMs across India.

ICICI was formed in 1955 at the initiative of the World Bank, the Government of India, and representatives of Indian industry. The principal objective was to create a development financial institution for providing medium-term and long-term project financing to Indian businesses. Until the late 1980s, ICICI primarily focused its activities on project finance, providing long-term funds to a variety of industrial projects. With the liberalization of the financial sector in India in the 1990s, ICICI transformed its business from a development financial institution offering only project finance to a diversified financial services provider that, along with its subsidiaries and other group companies, offered a wide variety of products and services. As India's economy became more market-oriented and integrated with the world economy, ICICI capitalized on the new opportunities to provide a wider range of financial products and services to a broader spectrum of clients. ICICI Bank was incorporated in 1994 as a part of the ICICI group. In 1999, ICICI became the first Indian company and the first bank or financial institution from non-Japan Asia to be listed on the New York Stock Exchange.

# Financial Performance

## FINANCIAL HIGHLIGHTS

The financial performance for fiscal 2020 is summarised in the following table:

| ₹ in billion, except percentages | Fiscal 2019 | Fiscal 2020 | % change |
|---|---|---|---|
| Net interest income and other income | 415.27 | **497.16** | 19.7% |
| Operating expenses | 180.89 | **216.14** | 19.5% |
| Core operating profit | 220.72 | **268.08** | 21.5% |
| Treasury income | 13.66 | **12.93** | (5.3)% |
| Operating profit | 234.38 | **281.01** | 19.9% |
| Provisions & contingencies (excluding tax) | 196.61 | **140.53** | (28.5)% |
| Profit before tax | 37.77 | **140.48** | 271.9% |
| Profit after tax | 33.63 | **79.31** | 135.8% |

| ₹ in billion, except percentages | Fiscal 2019 | Fiscal 2020 | % change |
|---|---|---|---|
| Consolidated profit before tax and minority interest | 74.08 | **185.89** | 150.9% |
| Consolidated profit after tax and minority interest | 42.54 | **95.66** | 124.9% |

***Shareholders of ICICI Bank with more than one percent holding at March 31, 2020***

| Name of the Shareholder | No. of Shares | % holding |
|---|---|---|
| Deutsche Bank Trust Company Americas* | 1,239,237,331 | 19.15 |
| Life Insurance Corporation of India | 521,902,188 | 8.06 |
| SBI Mutual Fund | 237,754,343 | 3.67 |
| HDFC Mutual Fund | 224,605,465 | 3.47 |
| Dodge & Cox International Stock Fund | 196,131,976 | 3.03 |
| ICICI Prudential Mutual Fund | 163,148,918 | 2.52 |
| Europacific Growth Fund | 135,615,372 | 2.10 |
| Reliance Mutual Fund | 112,506,655 | 1.74 |
| Government of Singapore | 108,965,533 | 1.68 |
| Aditya Birla Sun Life Mutual Fund | 107,202,608 | 1.66 |
| Kotak Mahindra Mutual Fund | 94,184,312 | 1.46 |
| NPS Trust | 88,627,973 | 1.37 |
| UTI Mutual Fund | 83,008,110 | 1.28 |
| Axis Mutual Fund | 73,705,550 | 1.14 |
| Abu Dhabi Investment Authority | 71,243,514 | 1.10 |

* *Deutsche Bank Trust Company Americas holds equity shares of ICICI Bank as depositary for ADS holders.*

## KEY RATIOS

The following table sets forth, for the periods indicated, the key financial ratios.

| Particulars | Fiscal 2019 | Fiscal 2020 |
|---|---|---|
| Return on average equity (%)[1] | 3.16 | **7.07** |
| Return on average assets (%)[2] | 0.39 | **0.81** |
| Net interest margin (%) | 3.42 | **3.73** |
| Cost to income (%)[3] | 43.56 | **43.50** |
| Earnings per share (₹) | 5.23 | **12.28** |
| Book value per share (₹) | 168.11 | **179.99** |

*1. Return on average equity is the ratio of the net profit after tax to the quarterly average equity share capital and reserves.*

*2. Return on average assets is the ratio of net profit after tax to average assets.*

*3. Cost represents operating expense. Income represents net interest income and non-interest income.*

The return on average equity, return on average assets and earnings per share increased primarily due to an increase in profit after tax.

## Assets

The following table sets forth, at the dates indicated, the principal components of assets.

₹ in billion, except percentages

| Assets | At March 31, 2019 | At March 31, 2020 | % change |
|---|---|---|---|
| Cash and bank balances | ₹ 802.96 | **₹ 1,191.56** | 48.4% |
| Investments | 2,077.33 | **2,495.31** | 20.1 |
| - Government and other approved investments[1] | 1,479.09 | **1,883.20** | 27.3 |
| - Equity investment in subsidiaries | 98.03 | **98.03** | 0.0 |
| - Other investments | 500.21 | **514.08** | 2.8 |
| Advances | 5,866.47 | **6,452.90** | 10.0 |
| - Domestic | 5,236.15 | **5,913.23** | 12.9 |
| - Overseas branches | 630.32 | **539.67** | (14.4) |
| Fixed assets (including leased assets) | 79.31 | **84.10** | 6.0 |
| Other assets | 818.52 | **759.78** | (7.2) |
| - RIDF and other related deposits[2] | 292.55 | **287.57** | (1.7) |
| **Total assets** | **₹ 9,644.59** | **₹ 10,983.65** | **13.9%** |

*1. Banks in India are required to maintain a specified percentage, currently 18.00% (at March 31, 2020), of their net demand and time liabilities by way of liquid assets like cash, gold or approved unencumbered securities.*

*2. Deposits made in Rural Infrastructure Development Fund and other related deposits pursuant to shortfall in the amount required to be lent to certain specified sectors called priority sector as per RBI guidelines.*

*3. All amounts have been rounded off to the nearest ₹ 10.0 million.*

Total assets of the Bank increased by 13.9% from ₹ 9,644.59 billion at March 31, 2019 to ₹ 10,983.65 billion at March 31, 2020, primarily due to a 10.0% increase in advances and a 20.1% increase in investments, offset, in part, by a 7.2% decrease in other assets.

## LIABILITIES

The following table sets forth, at the dates indicated, the principal components of liabilities (including capital and reserves).

₹ in billion, except percentages

| Liabilities | At March 31, 2019 | At March 31, 2020 | % change |
|---|---|---|---|
| Equity share capital | ₹ 12.94 | **₹ 12.98** | 0.3% |
| Reserves | 1,070.74 | **1,152.06** | 7.6 |
| Deposits | 6,529.20 | **7,709.69** | 18.1 |
| - Savings deposits | 2,276.71 | **2,455.91** | 7.9 |
| - Current deposits | 962.69 | **1,022.28** | 6.2 |
| - Term deposits | 3,289.80 | **4,231.51** | 28.6 |
| Borrowings (excluding subordinated debt) | 1,382.85 | **1,410.79** | 2.0 |
| - Domestic | 635.07 | **811.26** | 27.7 |
| - Overseas branches | 747.78 | **599.53** | (19.8) |
| Subordinated debt (included in Tier-1 and Tier-2 capital) | 270.35 | **218.17** | (19.3) |
| Domestic | 270.35 | **218.17** | (19.3) |
| Other liabilities | 378.51 | **479.95** | 26.8 |
| **Total liabilities** | **₹ 9,644.59** | **₹ 10,983.65** | **13.9%** |

*1. All amounts have been rounded off to the nearest ₹ 10.0 million.*

Total liabilities (including capital and reserves) increased by 13.9% from ₹ 9,644.59 billion at March 31, 2019 to ₹ 10,983.65 billion at March 31, 2020 primarily due to a 18.1% increase in deposits.

# Introduction to Time-Series Analysis

## Definition

Time series analysis is a statistical method to analyse the past data within a given duration of time to forecast the future. It comprises of ordered sequence of data at equally spaced interval.

## Importance of Time Series

Ample of time series data is being generated from a variety of fields. And hence the study time series analysis holds a lot of applications. Let us try to understand the importance of time series analysis in different areas.

- **Field of Economics:** Budget studies, census Analysis, etc.
- **Field of Finance**: Widely used in the field of finance such as to understand the stock market fluctuations, yield management, understand the market volatility, etc.
- **Healthcare**: An epidemiologist might be interested in knowing the number of people infected over the past years. Like in the current situation the researchers might be interested in knowing the people affected by the coronavirus over a period. Blood pressure traced over a period can be used in evaluating a drug.
- **Environmental Science:** Environmental time series data can help us explain the rise in temperature over the past few years. Plot shows the temperature data over a period.
- **Sales forecasting**: Understanding the sales number forecasting, the company's performance. The plot shows an earning with respect to time.

## Components of Time-Series

Time series can be divided into three-major components, which are:

- **Trend**: Linear Fluctuations in the variable.
- **Seasonality**: Repeating / Cyclical patterns observes during a time.
- **Residual**: Volatility in the variable under observation.

$$Time\ Series\ = Trend + Seasonality + Residual$$

All the time-series are having **trend** and **residual** to explain the model, but the seasonality is additional behaviour.

## Properties of Time-Series: Stationarity

Time series whose statistical properties such as mean, variance, autocorrelation, etc. are all constant over time are known as stationary time-series.

*Assumptions for Stationarity:*

$$\mu = Constant$$

$$\sigma^2 = Constant$$

$$Cov(X_n, X_{n+k}) = Cov(X_m, X_{m+k})$$

The time series which are following these three main properties can be used in forecasting. If the time-series is non-stationary, then we cannot predict or forecast it.

## Augmented Dickey-Fuller Test

Augmented Dickey Fuller test (ADF Test) is a common statistical test used to test whether a given Time series is stationary or not. It is one of the most used statistical tests when it comes to analysing the stationary of a series.

$$H_o: Non - stationary\ time\ series, unit\ root\ .$$

$$H_A: No\ root\ and\ time\ series\ is\ stationary.$$

If null hypothesis gets rejected at given significance level then, we can say that the time-series under observation is stationary.

## Autocorrelation & Partial Autocorrelation

Autocorrelation is a mathematical representation of the degree of similarity between a given time series and a lagged version of itself over successive time intervals.

Calculation of autocorrelation has two kind of impact associated in it:

- Indirect Impact
- Direct Impact

Partial Autocorrelation is used to find only direct impact on the current variables of the lagged variables, it omits out the indirect impact.

With the help of autocorrelation and partial autocorrelation, we can find out the lags which has significant degrees of similarity and using that lag we can predict the future variable.

## Log-Likelihood Ratio Test

The log likelihood test is the statistical test to compares the goodness of fit of two different nested models to test whether the alternative nested model is better than the simpler model or not.

$H_o$: All the parameters of both the models are same.

$H_A$: $Atleast\ one\ of\ the\ parameters\ of\ alternative\ model\ is\ better.$

On testing if we get the p-value under the significance level then the choose the alternative model otherwise we proceed with the simpler model.

# Forecasting Models

## ARMA / ARMAX / SARMAX

For forecasting the stationary data, we'll use Auto-Regressive-Moving-Average model (popularly known as ARMA) or its derivative models which are ARMAX and SARMAX. (X is for exogeneous variables in the model and S represent the Seasonality.)

These models are great in forecasting the stationary data but not performs well on non-stationary dataset. To use above mentioned models, we need to convert our datasets into stationary time-series.

ARMA (p, q)

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

Here, p is the no. of lag in past values and q is the no of lag the past error. These variables together work as unit to predict the present value.

## ARIMA / ARIMAX / SARIMAX

For forecasting of non-stationary data, we'll use Auto-Regressive-Integrated-Moving-Average model (popularly known as ARIMA) or its derivatives which are ARIMAX or SARIMAX. (X is for exogeneous variables in the model and S represent the Seasonality.)

These models are using integrated values or differences in the values to predict the future values. The level of integration is defined at which the non-stationary data converts into stationary data.

ARIMA (p, d, q)

$$\Delta^d(y_t) = c + \phi_1 \Delta^d(y_{t-1}) + \cdots + \phi_p \Delta^d(y_{t-p}) + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

Here, $\Delta$ is the level of integration in the model, which is represent as the no. of layers we used for the differencing purpose.

## ARCH / GARCH

These models are used for the volatility prediction and forecasting of given time-series. In this model we'll treat mean and variance as different equations and using lagged values in the variance we can predict the volatility using ARCH or GARCH model.

# Data Analysis for ICICI.NS

## Data Collection & Processing

For analysis we've collected the 5-year historical data of adjusted price of ICICI.NS from 1st January 2015 to 26th February 2021. The data is extracted from Yahoo. The analysis is done in two sections.

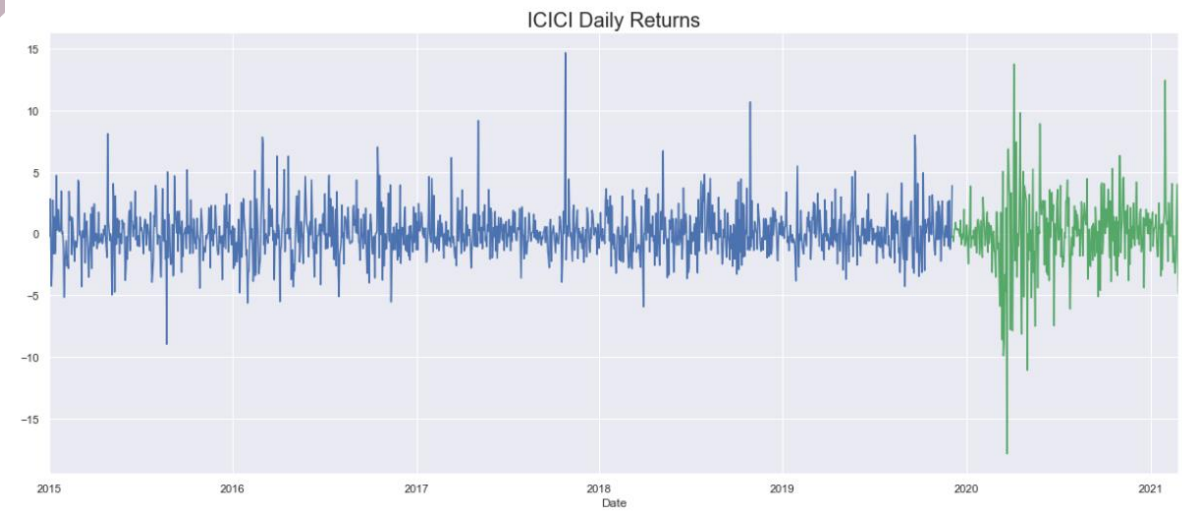
- Sub-section 1: ARIMA on prices (without using exogeneous variable)
- Sub-section 2: ARMAX on returns (with using exogeneous variables)

In Sub-section 2; the historical data of adjusted price of SENSEX, NIFTY, HDFC.NS is treated as exogenous variable and time frame is same as that of ICICI.NS.

## Train-Validation-Test Split

The given dataset is further categorised into two sets:

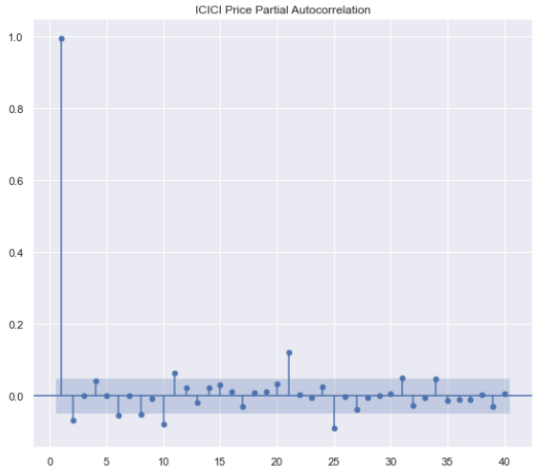
- Training Set (1st January 2015 – 5th December 2019)
- Test Set (6th December 2019 – 26th February 2021)

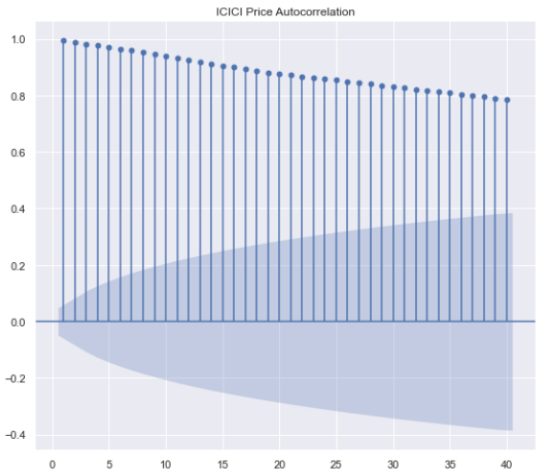



## Testing for stationarity

The price data is turned out be non-stationary, thus we convert prices into returns it after testing it with ADF Test. It came out to statistically significant.

## ACF & PACF for Prices

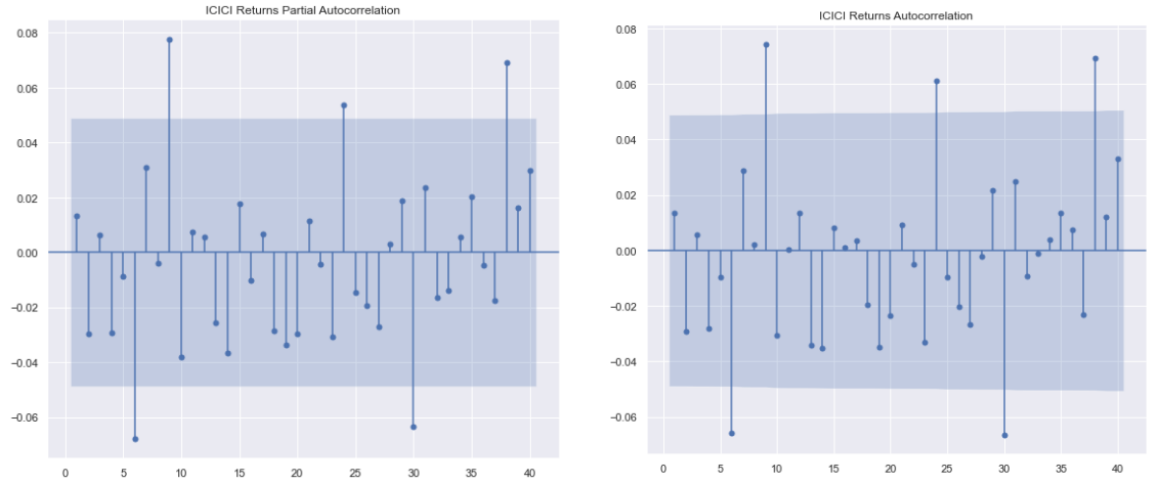



The ACF and PACF graphs for prices show that almost all the lags are significant. Hence, we can say that the prices are having high degrees of similarity with lagged coefficients.

But the time-series will not perform in forecasting.

## ACF & PACF for Returns

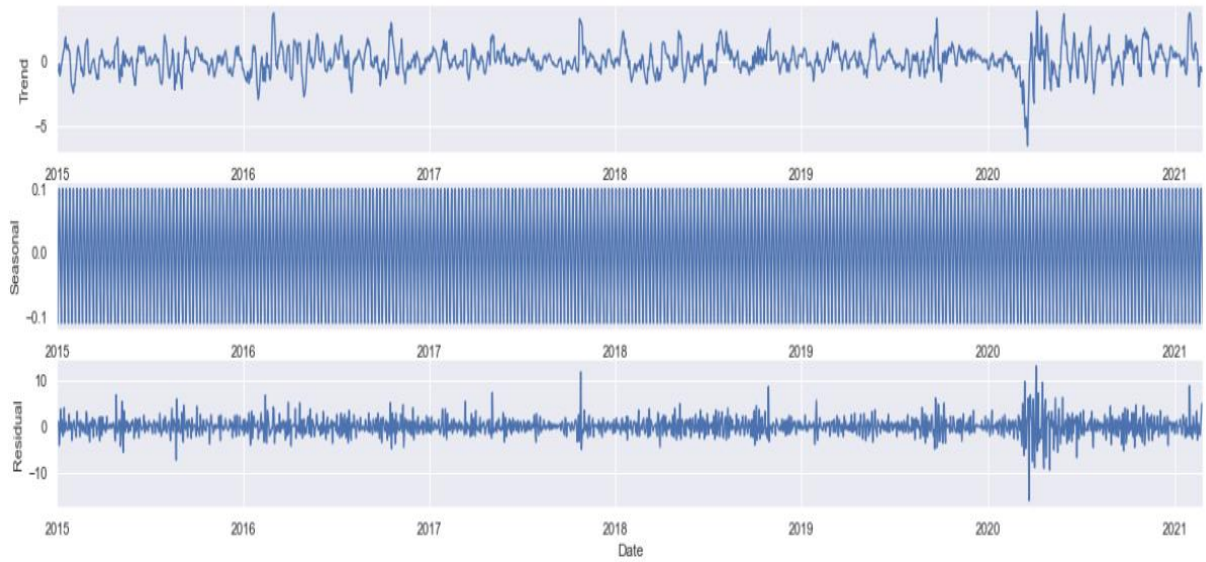



On the other hand, we have ACF & PACF for returns in which we can clearly see that only some the lags are significant, and rest of the lags are following normal distribution as they are not significant.

Since 6$^{th}$ lag in PACF and 6$^{th}$ lag in ACF is significant, we can say that the model can be predicted using the lags between 1$^{st}$ to 6$^{th}$ lag.

## Testing for Seasonality

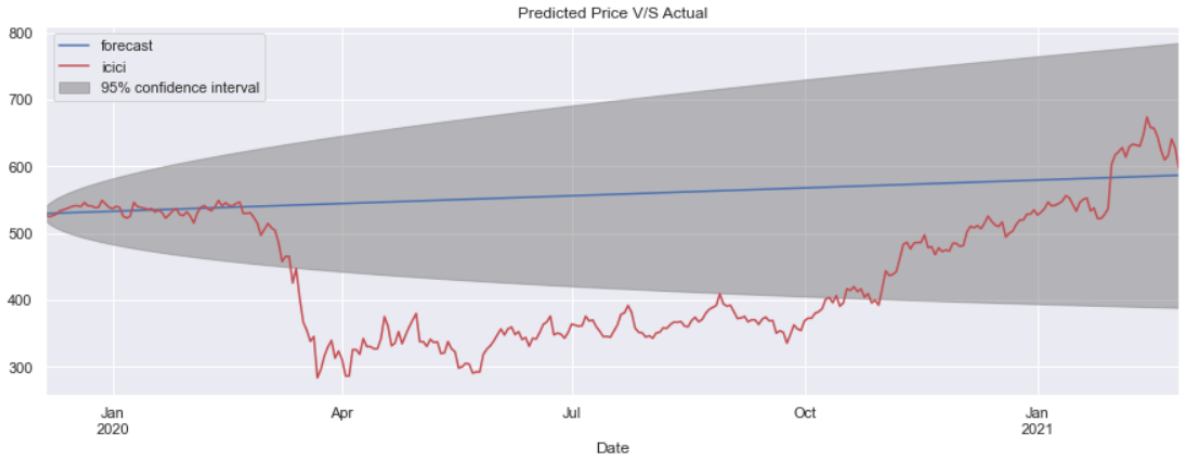
On testing the seasonality, we came across the conclusion that model is very frequently repeat itself in a year. Thus, there is no significant seasonality in our model as the data is constant every other day.



# Forecasting

## Sub-section 1: ARIMA

On training the model, we find that ARIMA (1,1,1) is best among them all. Then we have plotted the graph for the predicted and actual price.




Since, exogenous is not considered previously then it always gives a straight line. And we cannot talk about the fluctuations.

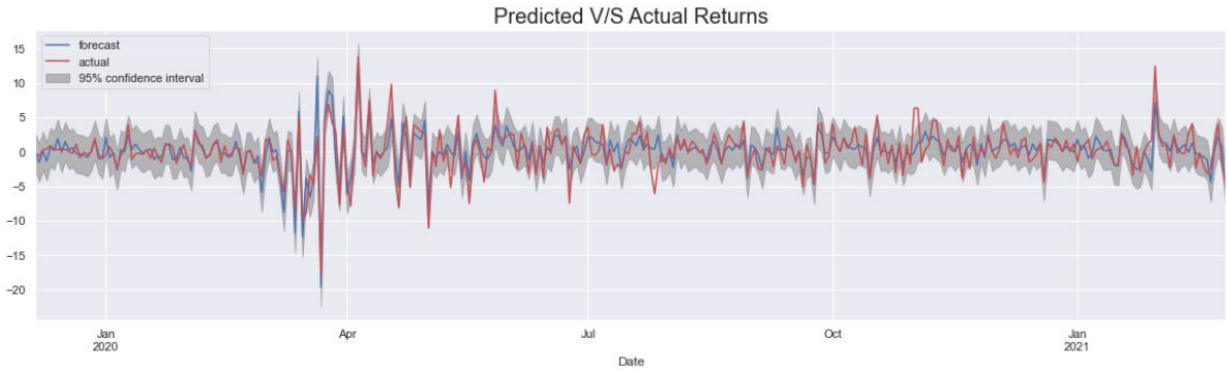
### *Forecasting Results*

The forecasting using ARIMA (1,1,1) gives us the results as follows:


We've found out that the prices of the stock ICICI.NS would be around INR 700.

## Sub-section 2: ARMAX

On training different models, we find that ARMAX (1,3) is the best model to predict the unknown timeseries. Considering the exogenous variables trained our model in accurate and more precise way.

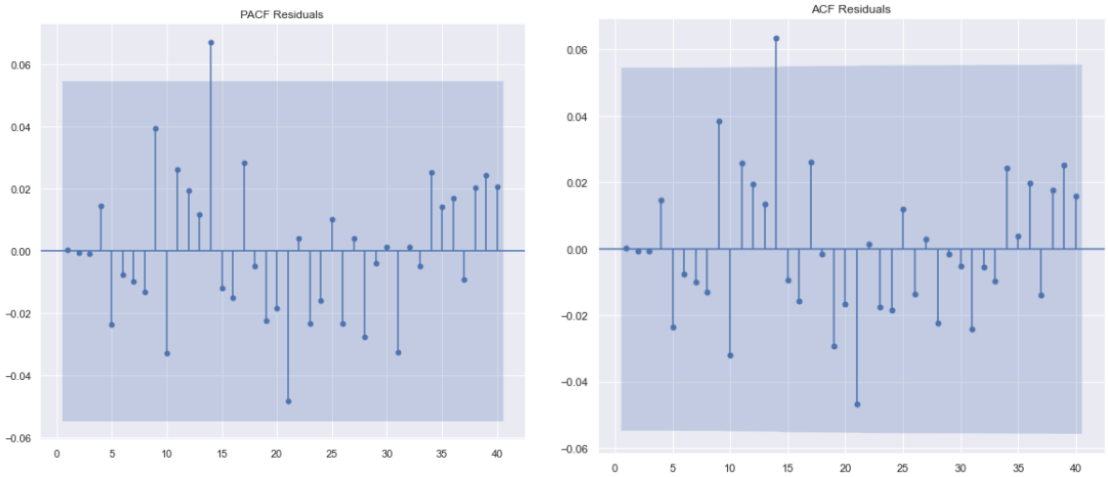



The prediction is closely matched with the unknown data of validation set. This model is the optimum model which will be used to predict the future returns. The goodness of fit of this model is 0.657.

## *Model Evaluation*

The optimum model is always prone to the overfitting case, where some of the residuals became significant in our model due to overfitting on training set and then failure in generalisation occurred.

Now, we need to check whether the residuals of the predicted model are following a normal distribution or not.

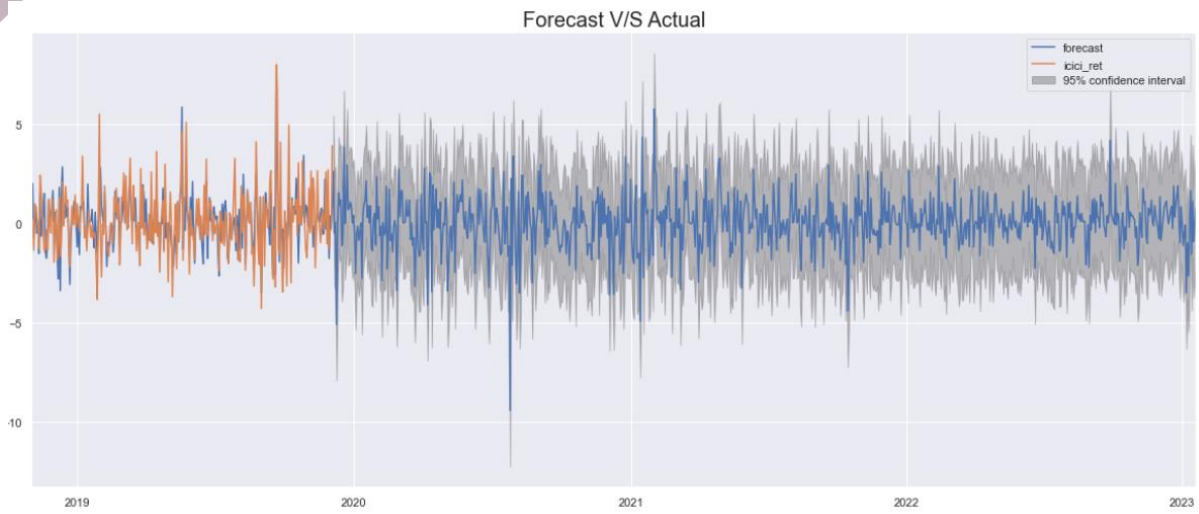



We can check from the ACF & PACF plots of the residuals that it is following a normal distribution. Hence, there is no overfitting.

Hence, we've arrived at the conclusion that the ARMA (1,3) is the best model to predict returns that too with $R^2$-score of 0.6432.

## *Forecasting Results*

With the help of the trained model the forecasted returns are as follows:




And from the forecasted returns the price movement till 2023 is as follows:


The expected price of ICICI.NS is around INR 700 by the end of 2023.

# Comparison of ARIMA & ARMAX

On comparing the models, we can say that both are equally good in forecasting as both gives the same results by the end of 2023. ARIMA model is failed to explain the short-term fluctuations. On the other hand, ARMA is good at forecasting the fluctuations as well.


# Current Market Scenario

As per the forecasting which was done using the training set. We're able to forecast the prices till the end of the year. And the price movement is within the 95% CI of the forecast, except the downfall occurred during the COVID outbreak.

# Conclusion

Time Series analysis is a tool which helps as to forecast the values using past errors and past values. The forecasting of the price is often helping an investor to think about the long-term opportunity and decide whether to buy the stock or not. Expected Price plays a key role in every trade in the market. Every investor before trading in the market has some expectation regarding the price. Investor forecast the expected price and if the situation seems profitable then only, he invests in that particular stock.

Forecasting can help the investor to earns profit. ARIMA and ARMA are some of the important techniques to forecast the prices. They are frequently used in the market to forecast. But it lags behind in accurately forecasting in the market. Many traders in the market are short-term players and hence using ARIMA and ARMA models. There is a possibility of making a huge loss. Thus, ARIMA and ARMA technique are not for scalper, intra-day traders.

There are more advanced deep learning models such as LSTM are there for accurately predicting the price and short-term fluctuations.

But ARIMA & ARMA are indeed good techniques to forecast and invest in long-term.