



# GOLD BOND or GENERAL BONDS

---

## Investment

An investment is an asset or item acquired with the goal of generating income or appreciation. Appreciation refers to an increase in the value of an asset over time.

**"Investing" we mean buying an asset for making a profit by selling it in the future, after it appreciates in value.**

GOLD BOND

GENERAL BONDS

# Project Goal

Finding the best investment between Gold Bond and General Bonds by building a forecasting model.



## Objectives

---

- Maximize the Investment returns

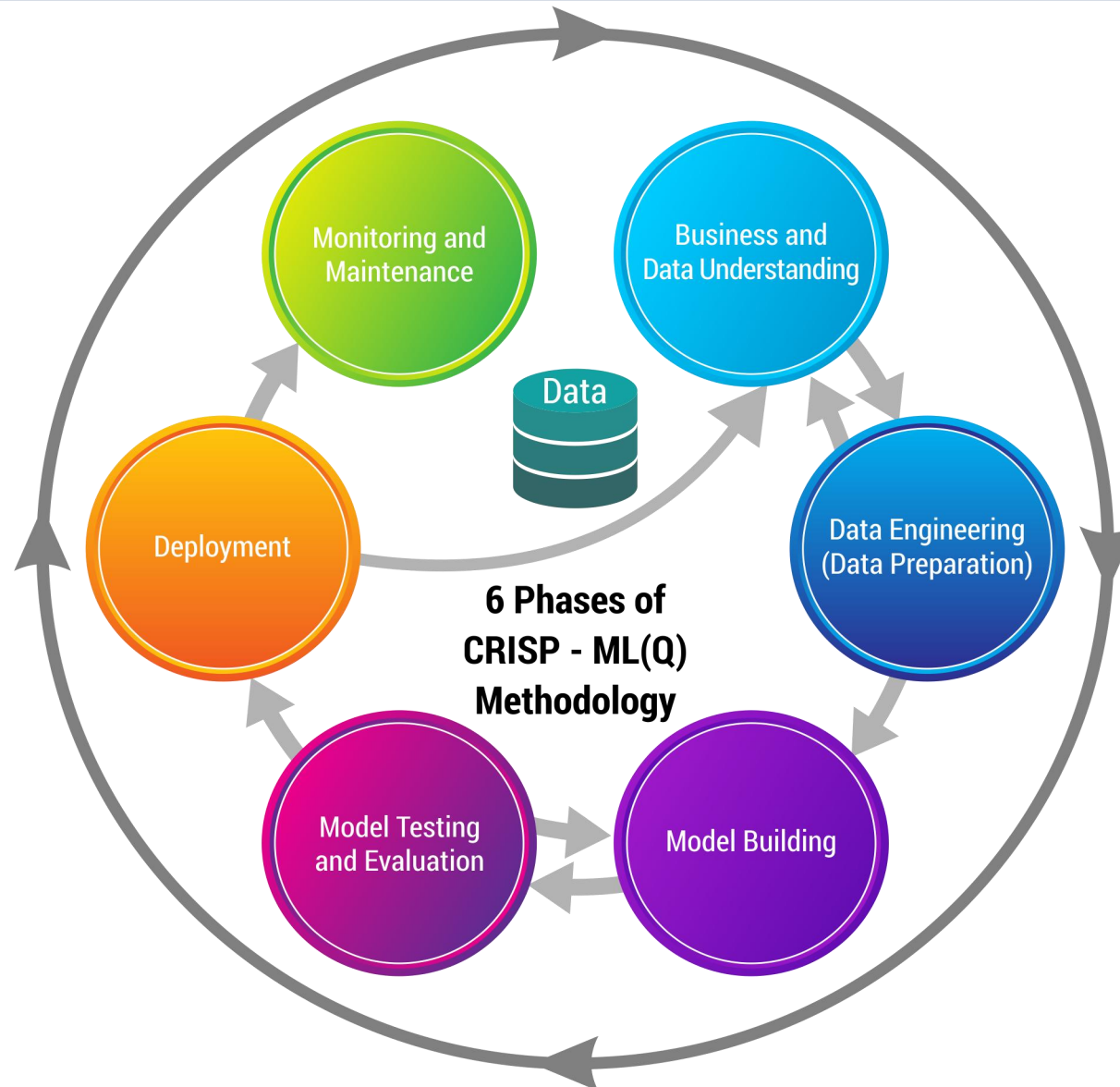


## Constraints

---

- Minimize the Investment risk

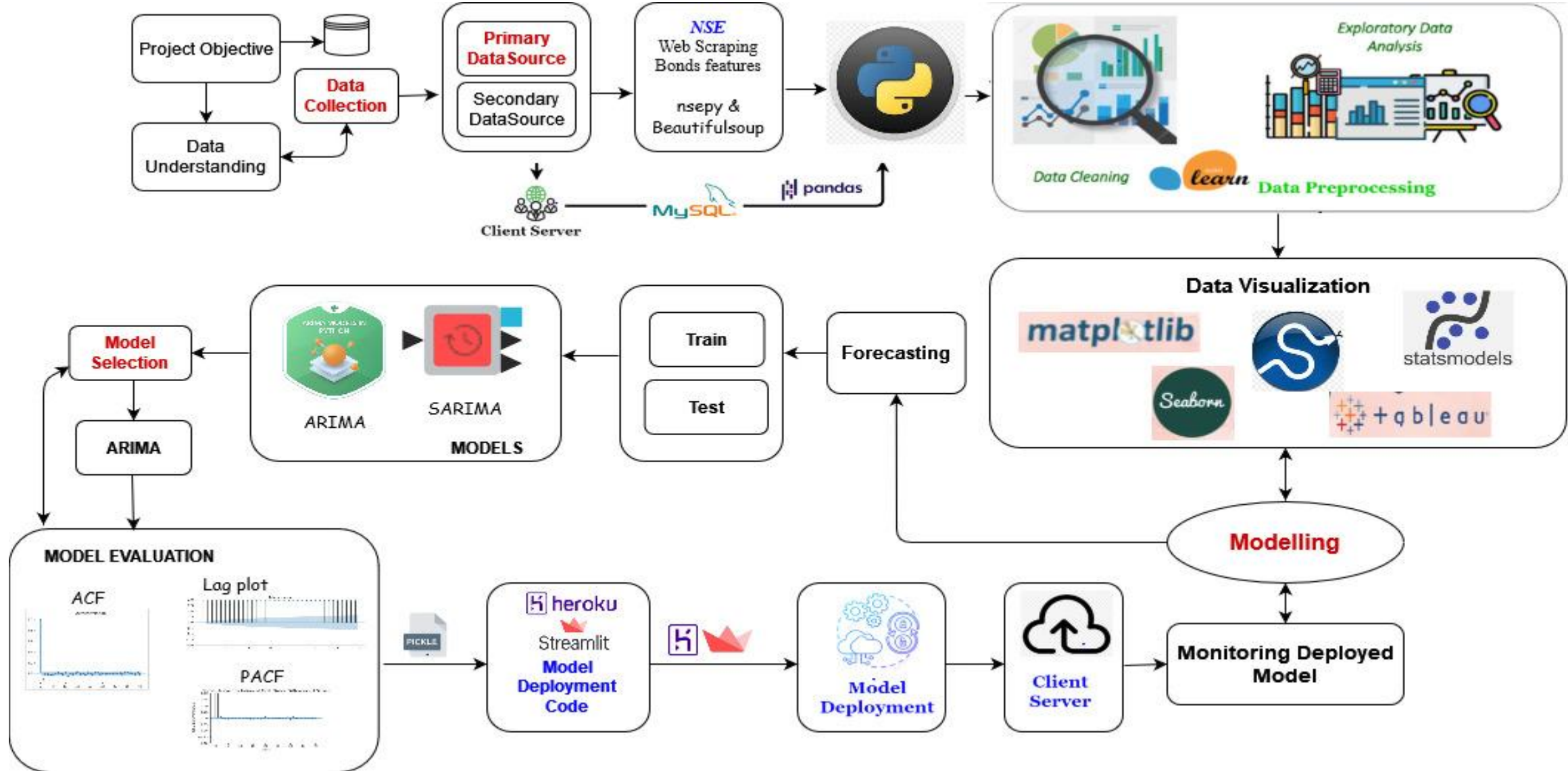
# CRISP-ML(Q) Methodology

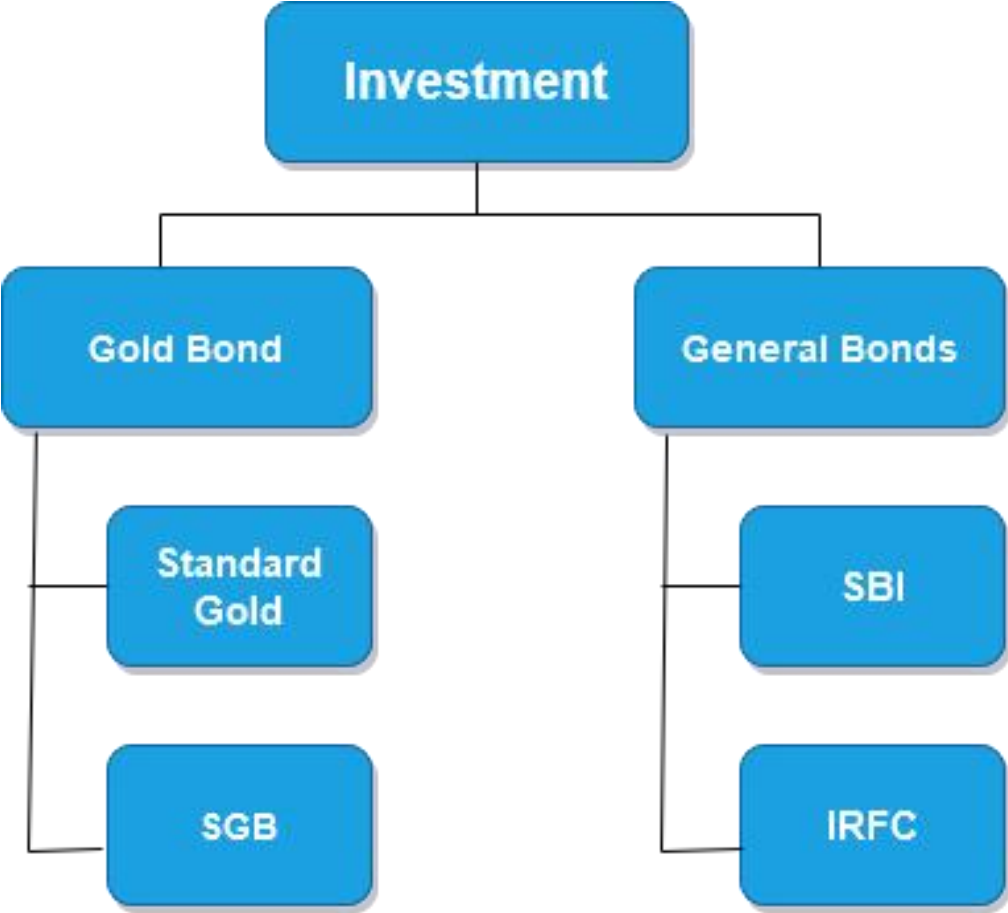


# Technical Stacks



# Project Architecture / Data Pipeline





	Data Sources
Standard Gold	<a href="https://www.livechennai.com/get_goldrate_history.asp">https://www.livechennai.com/get_goldrate_history.asp</a>
SGB	<a href="https://www.topsharebrokers.com/report/sovereign-gold-bond-return-calculator/317/">https://www.topsharebrokers.com/report/sovereign-gold-bond-return-calculator/317/</a>
SBI	<a href="https://in.investing.com/funds/sbi-life-bond-fund-historical-data?end_date=1632162600&amp;st_date=1313951400">https://in.investing.com/funds/sbi-life-bond-fund-historical-data?end_date=1632162600&amp;st_date=1313951400</a>
IRFC	<a href="https://www1.nseindia.com/products/content/equities/equities/eq_security.htm">https://www1.nseindia.com/products/content/equities/equities/eq_security.htm</a>



# Sovereign GOLD Bond

They are substitutes for holding physical gold. Investors have to pay the issue price in cash and the bonds will be redeemed in cash on maturity. The Bond is issued by Reserve Bank on behalf of Government of India.

## Benefits

- No GST when purchase
- No Making Charges
- Assured Returns of 2.5% P.A
- No Worry for Storage
- No Capital Gain Tax on Redemption
- Minimum 1 gram investment

## Who can Invest in ?

- Individuals
- HUFs
- Education Institutions- Universities
- Charitable organizations and trusts

Minimum investment in the Bond shall be one gram with a maximum limit of subscription of 4 kg for individuals.



# Data Understanding for SGB

- Data Source: Primary Data
- Dataset shape: 2474 observations, 3 features

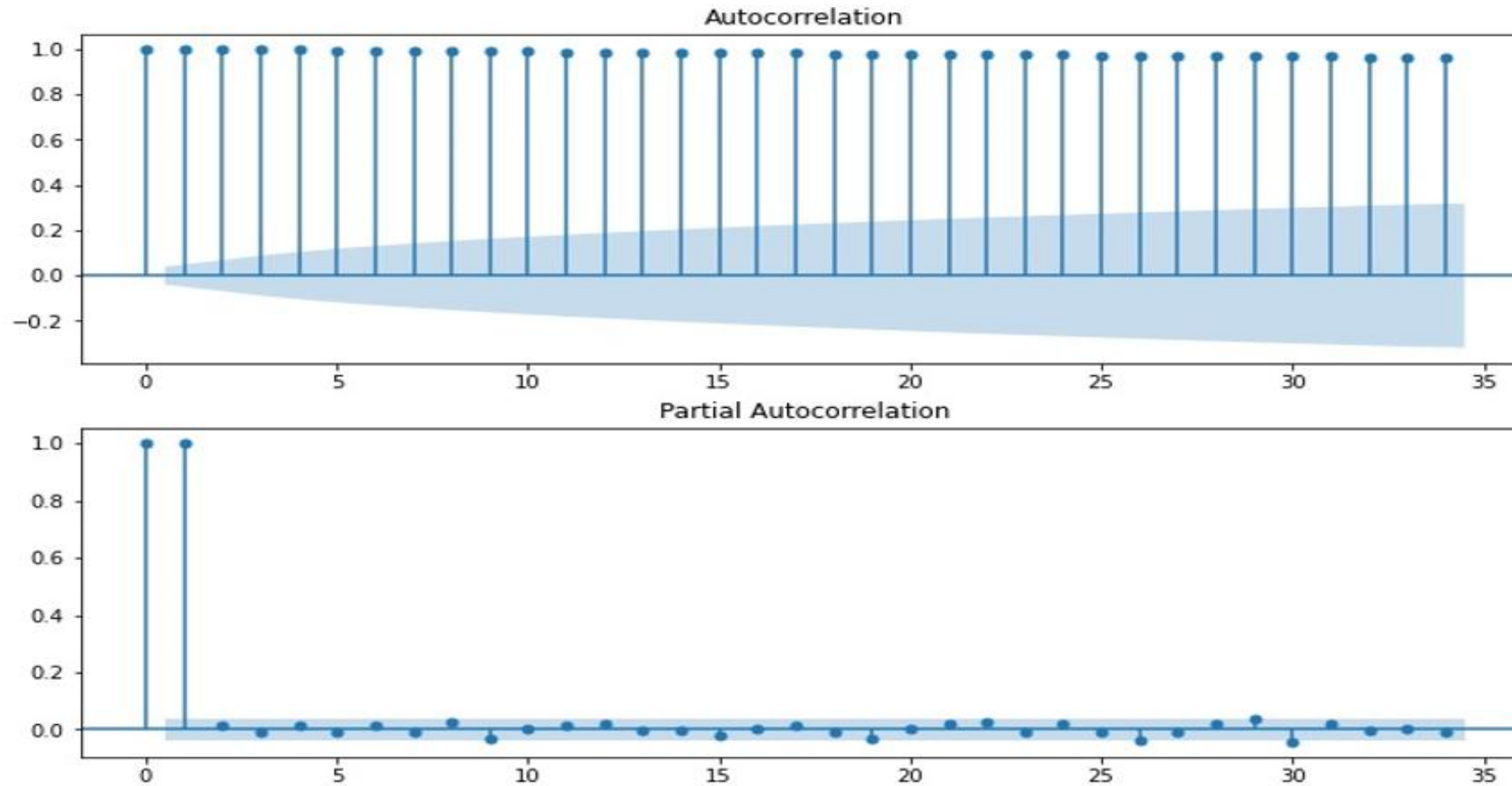
Name of Feature		Description
Date		Date format(YYYY/MM/DD)
Price		24 karat physical gold price
Price		22 Carat Standard gold

# Data Processing for Standard Gold/SGB

- Initial dataset was having rows 2474 and 3 columns.
- Following steps were taken to perform Exploratory Data Analysis using Python:
  - 🔍 Checked existence of NULL values, resulted in no “NA” value present.
  - 🔍 Selected 2 features (Date and Price) which would help to accomplish our purpose.
  - 🔍 Filled missing “Dates” in data set and “Pure Gold Prices” filled by Filled the missing dates in the data set and price was filled by using forward fill imputation.
  - 🔍 Final data set has 2496 rows and 2 columns.
  - 🔍 Checked Correlation using pair-plot and also checked seasonality.

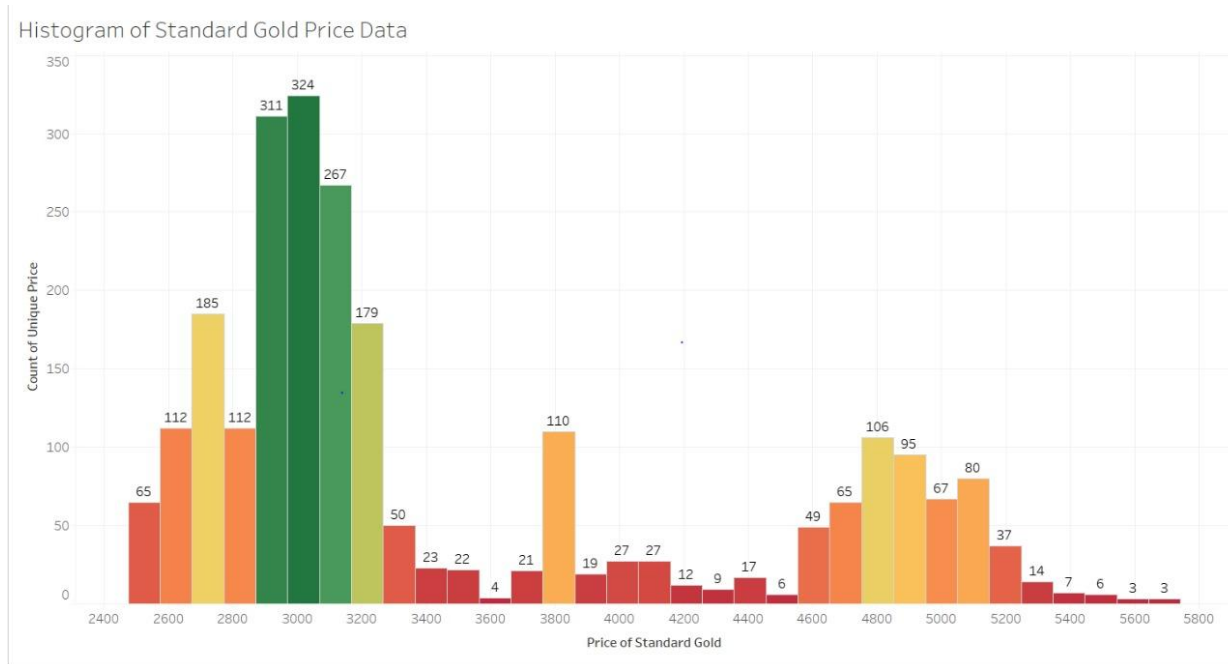


# EDA - Standard Gold/SGB



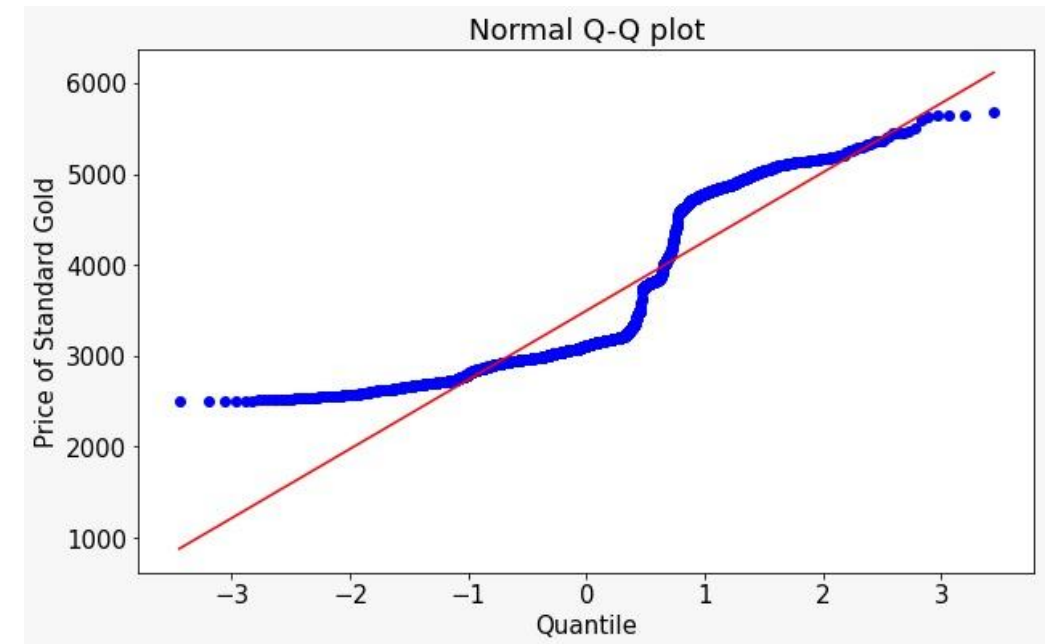
Here this plots are of Autocorrelation and Partial Autocorrelation which gives us the values for the hyperparameter which is P & Q.

# EDA – Standard Gold



**Fig: Histogram representation**

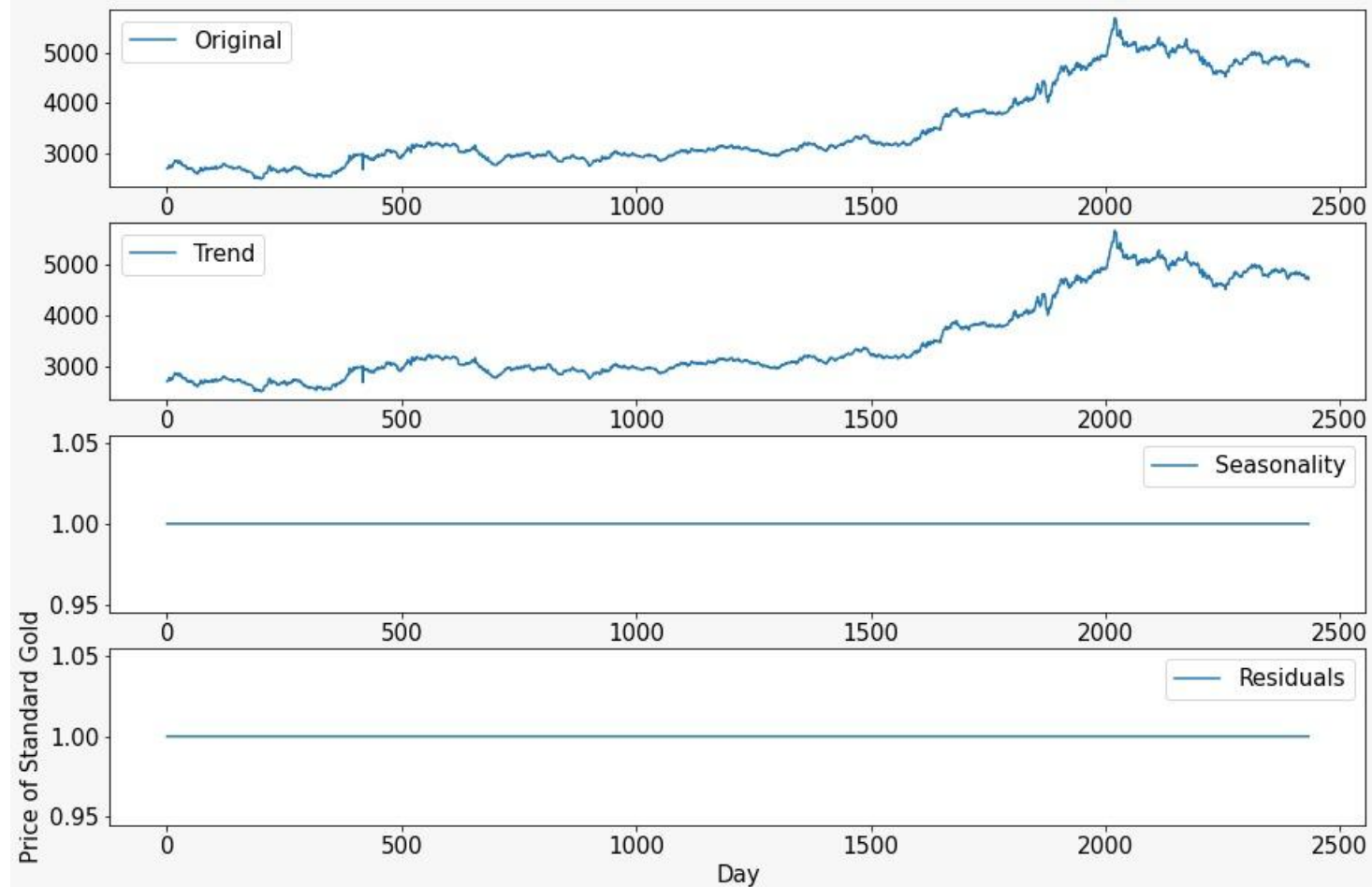
Data is not normally distributed and observed most data lies between 2400 to 3600.



**Fig: Q-Q plot representation**

It show the positive uptrend of the asset class with sustainable growth through pertinent quarters.

# EDA – Standard Gold



**Fig: Seasonality Graph**

In this graph, the data is showing upward trend and there is no seasonality in the data.

# GENERAL BONDS

A Bond is **loan from an investor to a borrower such as a company or government**. The borrower uses the money to fund its operations, and the investor receives interest on the investment. The market value of a bond can change over time.

Key Facts

Average returns  
Risks

Pros	Cons
Relatively safe	Low interest rates
Create a balancing force in portfolio	Difficulty in selling the bond you own
A form of fixed-income	Inflation can also reduce your purchasing power over time
Pay interest at regular, predictable rates and intervals.	
Solid asset	

# Data Understanding for SBI

- Data Source: Primary Data
- Dataset shape: 797 observations, 6 features

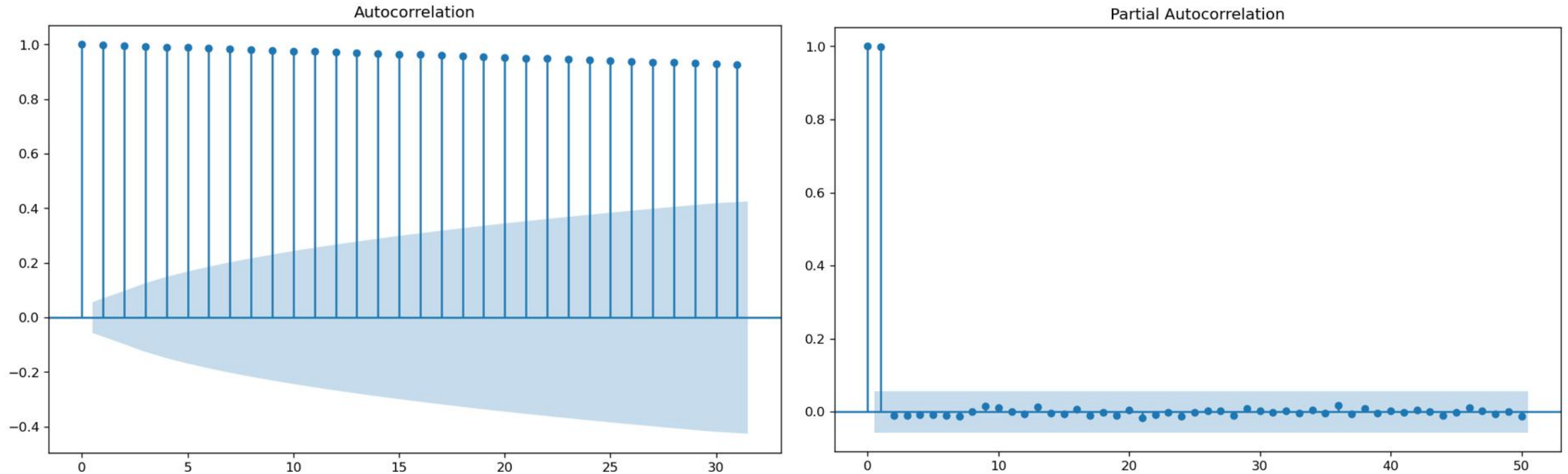
Name of Feature	Description
Date(Index)	Date format(YYYY/MM/DD)
Open / Price	Opening price of the bond
High	Highest Price of the bond
Low	Low Price of the bond
Close	Closing price of the bond
Change Pct	Change in the Open Price of Present-Day Price and Close Price of Previous day Price



# Data Processing for SBI

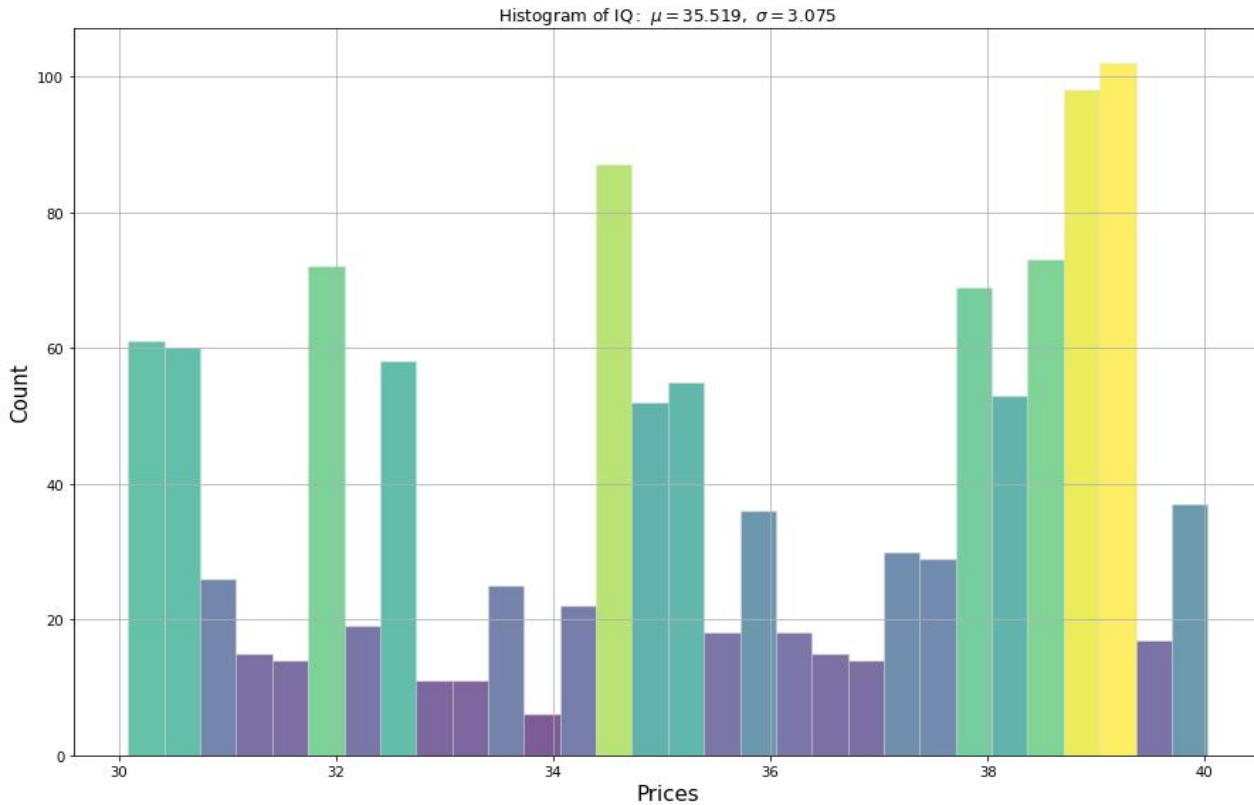
- Initial dataset was having 797 rows and 6 columns.
- Following steps were taken to perform Exploratory Data Analysis using Python:
  - 🔍 Checked existence of NULL values, resulted in no “NA” value present.
  - 🔍 Selected 2 features( Date, Price: Close) which would help to accomplish our purpose.
  - 🔍 Filled the missing dates in the data set and price was filled by using forward fill imputation.
  - 🔍 Final data set has 1204 rows and 2 columns.
  - 🔍 Checked Correlation using pair-plot and also checked seasonality.





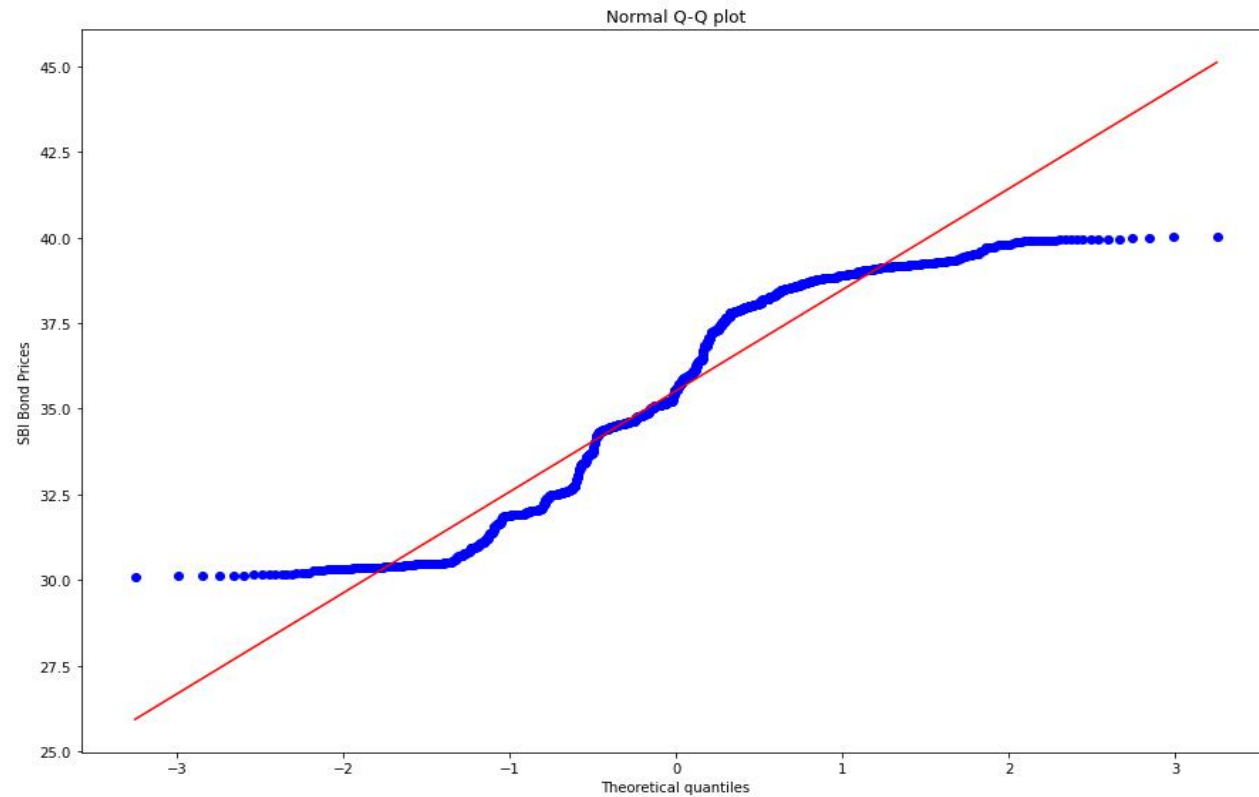
The above plot shows the Autocorrelation and Partial Autocorrelation which clearly displays the sudden steep from  $x=0$  to 1. Based on these changes, we have chosen the hyper parameter as  $P = 1$ ,  $Q = 0$ .

# EDA - SBI



**Fig: Histogram representation**

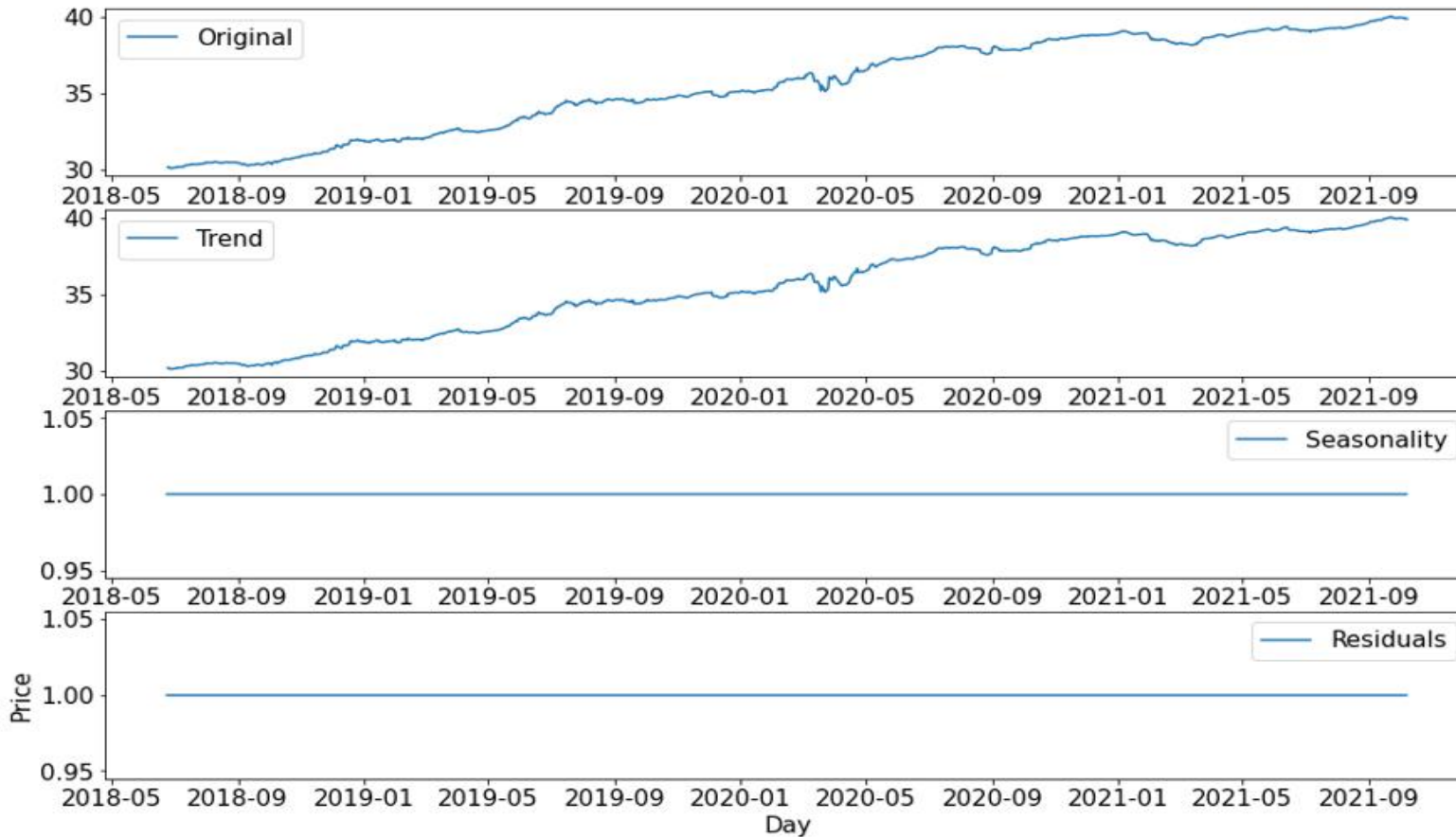
The price action is observed to be sinusoidal in nature with the maximum price observed between 30 to 40.



**Fig: Q-Q plot representation**

The Q-Q plot explains the normality of the data and also displays few outliers which will not be affecting the data.

# EDA - SBI



**Fig: Seasonality Graph**

The variation in price action seems to be in an uptrend, without an effect of seasonality. This shows stability and positive market.

# Data Understanding for IRFC

- Data Source: Secondary Data
- Dataset shape: 1596 observations, 8 features

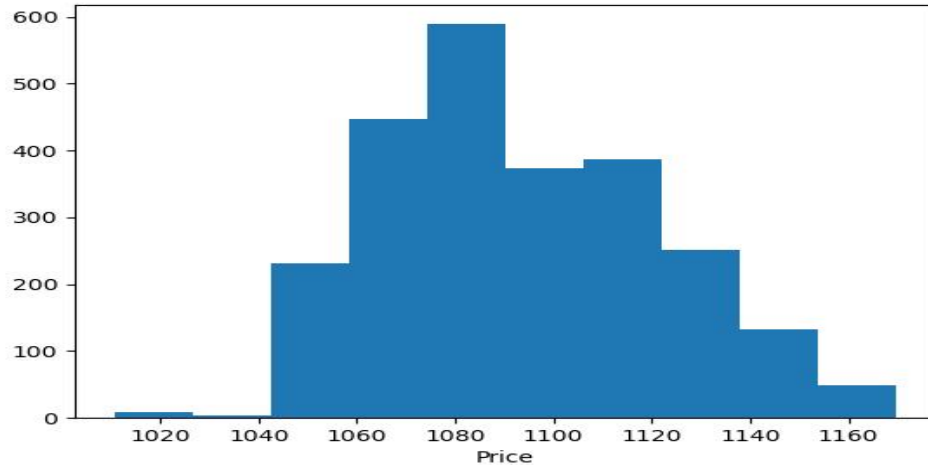
Name of Feature	Description
Date (Index)	Date format(YYYY/MM/DD)
Prev. Close(Price)	Closing price of the bond
Open	Opening price of bond
High	Highest price of bond reached for particular date
Low	Lowest price of bond reached for particular date
Last	Last price for the day
Close	Closed price for the day
VWAP(Avg)	Average of above features

# Data Processing for IRFC

➤ Following steps were taken to perform Exploratory Data Analysis in Python:

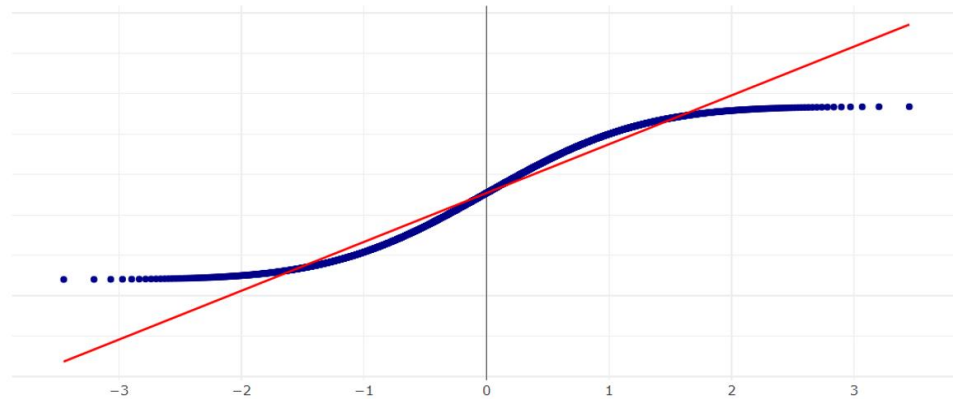
- 🔍 Checked existence of NULL values, resulted no “NA” value present.
- 🔍 Selected 2 features(Date, Price: Close) with the help of Correlation plot to accomplish our purpose.
- 🔍 Filled missing “Dates” in dataset and “Price” was filled by using forward imputation.
- 🔍 Final data set has 2473 rows and 2 columns.
- 🔍 Checked Correlation using pair-plot and also checked seasonality.





**Fig: Histogram representation**

Histogram representation of Price over last 5 years. Right skewness is observed.



**Fig: Q-Q plot representation**

QQ-Plot displays a normal distribution of the Price variable with few outliers which are not affecting our dataset.



Multiplicative Seasonality

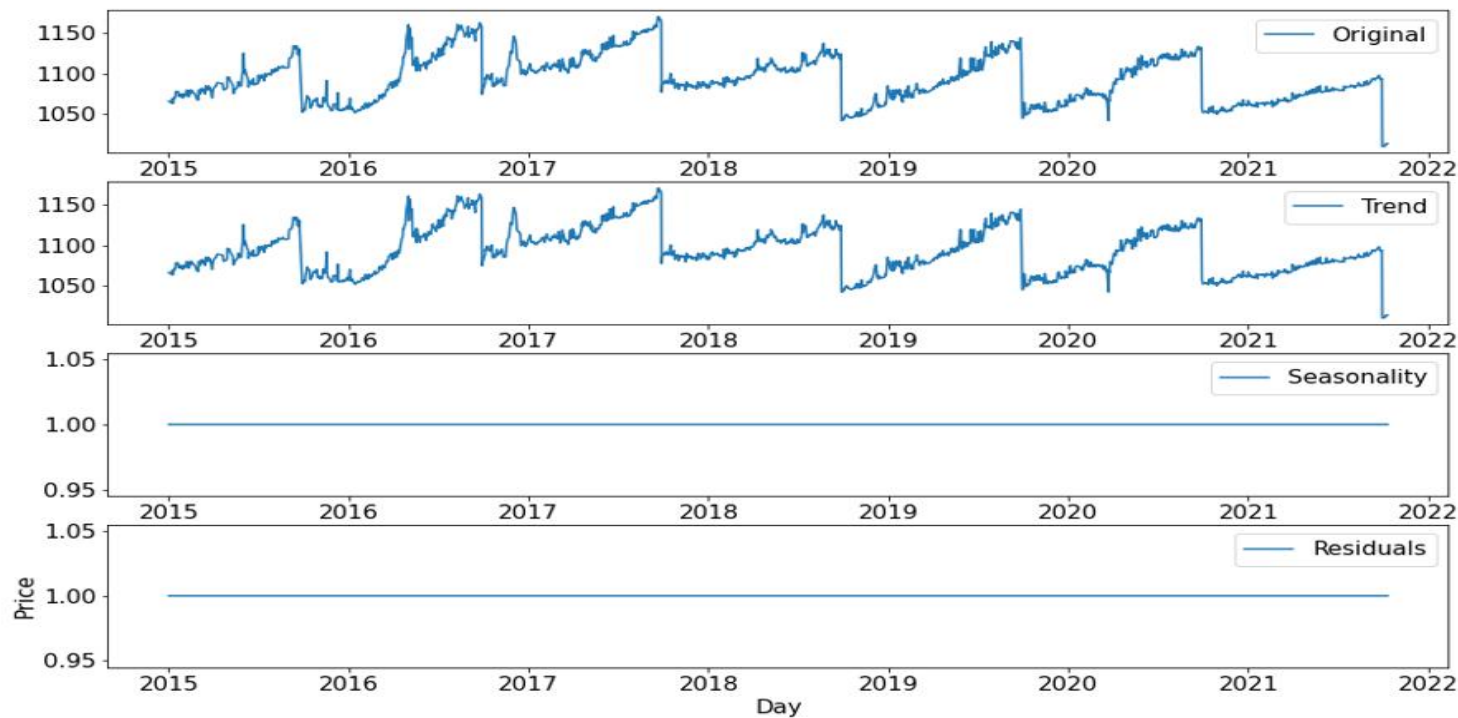
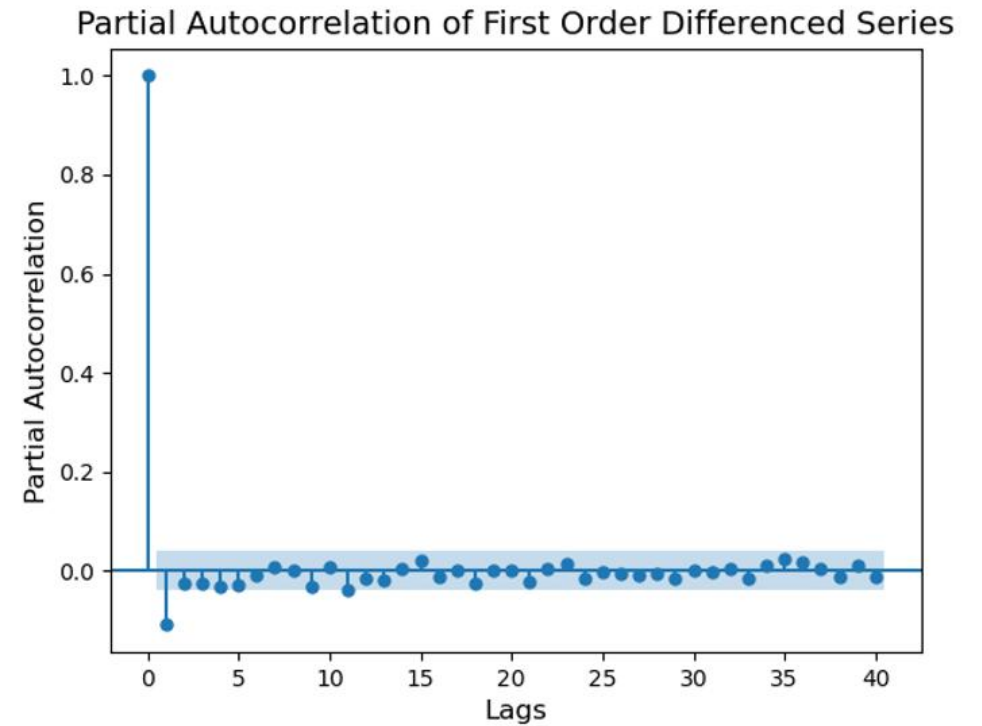
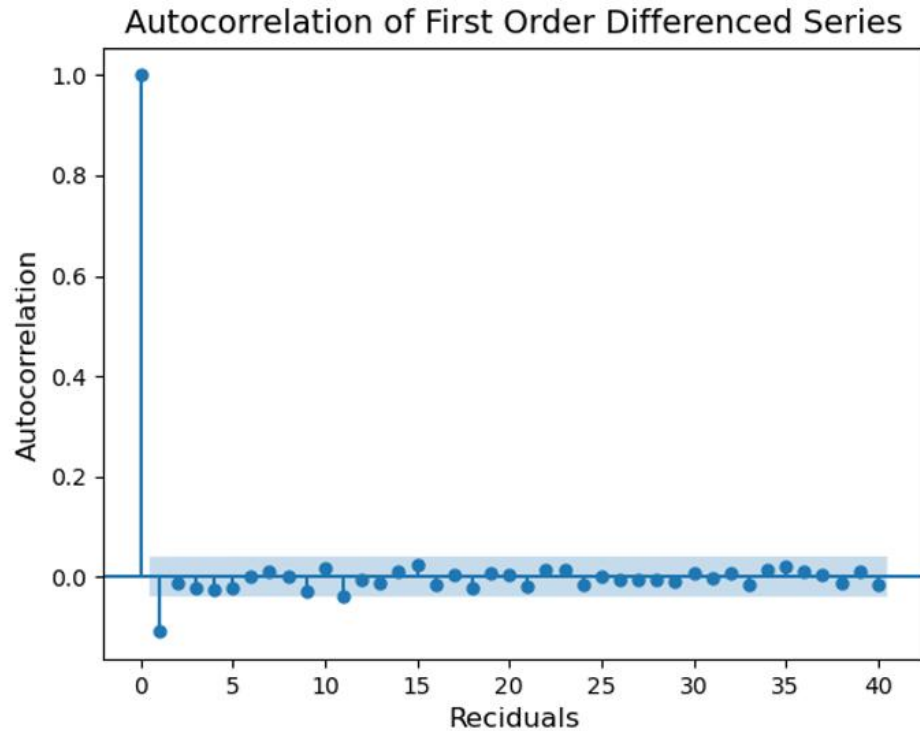


Fig: Seasonality Graph

The above plot shows the negative trend with no seasonality.

# EDA - IRFC



The above plot shows the Autocorrelation and Partial Autocorrelation which clearly displays the sudden steep from  $x=0$  to 1. Based on these changes, we have chosen the hyper parameter as  $P \ \& \ Q = 1$ .

# Model Building



# Model Building

## Autoregressive Integrated Moving Average Model:

This acronym is descriptive, capturing the key aspects of the model itself. Briefly, they are:

- **AR:** Autoregression. A model that uses the dependent relationship between an observation and some number of lagged observations.
- **I:** Integrated. The use of differencing of raw observations (e.g. subtracting an observation from an observation at the previous time step) in order to make the time series stationary.
- **MA:** Moving Average. A model that uses the dependency between an observation and a residual error from a moving average model applied to lagged observations.

The parameters of the ARIMA model are defined as follows:

- **p:** The number of lag observations included in the model, also called the lag order.
- **d:** The number of times that the raw observations are differenced, also called the degree of differencing.
- **q:** The size of the moving average window, also called the order of moving average.

# Model Description

Data	SGB (Pure Gold)	SBI	IRFC
(p, d, q)	(2, 1, 2)	(1, 1, 0)	(1,0,1)
Selection Process	auto_arima using pmdarima library	auto_arima using pmdarima library	Stationary check was conducted with Adfuller test using statsmodels library. Data is stationary, hence d = 0 p value was confirmed with autocorrelation plot. q value was confirmed with partial autocorrelation plot.

# Outputs of the Model in Codes

## SGB

```
=====
Dep. Variable:          D.pure_gold      No. Observations:          2474
Model:                  ARIMA(2, 1, 2)    Log Likelihood             -11437.527
Method:                  css-mle          S.D. of innovations        24.635
Date:                   Sun, 10 Oct 2021  AIC                           22887.053
Time:                   19:23:14          BIC                           22921.935
Sample:                 01-02-2015        HQIC                          22899.723
                   - 10-10-2021
=====

               coef      std err          z      P>|z|      [0.025      0.975]
-----
const           0.8514      0.507        1.680      0.093      -0.142      1.845
ar.L1.D.pure_gold -0.0150      0.205       -0.073      0.942      -0.416      0.386
ar.L2.D.pure_gold  0.7843      0.186        4.224      0.000       0.420      1.148
ma.L1.D.pure_gold -0.0036      0.209       -0.017      0.986      -0.413      0.406
ma.L2.D.pure_gold -0.7603      0.186       -4.078      0.000     -1.126     -0.395

                        Roots
=====
               Real      Imaginary      Modulus      Frequency
-----
AR.1          -1.1196      +0.0000j      1.1196      0.5000
AR.2           1.1388      +0.0000j      1.1388      0.0000
MA.1           1.1445      +0.0000j      1.1445      0.0000
MA.2          -1.1493      +0.0000j      1.1493      0.5000
-----
```

```
In [129]: forecast_test(fitted_SGB1, df_SGB_train, df_SGB_test)
MSE: 215394.55520053962
MAE: 433.9921869407817
RMSE: 464.1061895736143
```



# Outputs of the Model in Codes

## SBI

```
In [147]: fitted = model.fit(dis=-1)
```

```
In [148]: print(fitted.summary())
```

```
ARIMA Model Results
=====
Dep. Variable:          D.Close      No. Observations:          979
Model:                 ARIMA(1, 1, 0)  Log Likelihood             1341.784
Method:                css-mle        S.D. of innovations         0.061
Date:                  Sun, 10 Oct 2021  AIC                          -2677.568
Time:                  20:16:59         BIC                         -2662.908
Sample:                06-26-2018      HQIC                       -2671.991
                  - 02-28-2021

=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
const          0.0082      0.002      3.965      0.000      0.004      0.012
ar.L1.D.Close  0.0494      0.032      1.549      0.121     -0.013      0.112

Roots
=====
              Real      Imaginary      Modulus      Frequency
-----
AR.1          20.2330      +0.0000j      20.2330      0.0000
-----
```

```
In [149]: forecast_test(fitted, train_data, test_data)
MSE: 0.03231212962843829
MAE: 0.1576871822323348
RMSE: 0.1797557499175987
```



# Outputs of the Model in Codes

## IRFC

```
=====
                        ARMA Model Results
=====
Dep. Variable:          Close    No. Observations:          2192
Model:                 ARMA(1, 1)  Log Likelihood          -7004.607
Method:                css-mle    S.D. of innovations          5.905
Date:                  Sun, 10 Oct 2021    AIC              14017.214
Time:                  20:19:26    BIC              14039.985
Sample:                01-01-2015    HQIC             14025.536
                        - 12-31-2020
=====

              coef    std err          z      P>|z|      [0.025    0.975]
-----
const         1094.5550      6.396    171.135     0.000    1082.019    1107.091
ar.L1.Close      0.9829      0.004    246.325     0.000      0.975      0.991
ma.L1.Close     -0.1121      0.022    -5.080     0.000     -0.155     -0.069

                        Roots
=====
              Real      Imaginary      Modulus      Frequency
-----
AR.1           1.0174      +0.0000j      1.0174      0.0000
MA.1           8.9239      +0.0000j      8.9239      0.0000
-----

In [169]: forecast_test(fitted_IRFC, df_IRFC_train, df_IRFC_test)
MSE: 378.34917191475387
MAE: 13.979127605813128
RMSE: 19.451199755150167
```

# Comparison of Model Performance

Models	Model 1- SGB	Model 2-SBI	Model 3-IRFC
Accuracy	Medium	High	High
Returns (8 Years)	82.64%	111.51%	11.09%

# Deployment Strategy

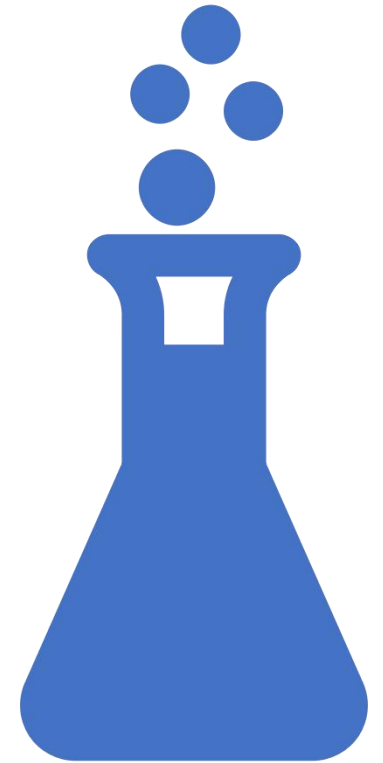
## Streamlit Features:

- Streamlit has been used in this project for building the web application
- No front-end programming experience required
- Easy to deploy
- Compatibility with all major frameworks/libraries



## Heroku Features:

- Support for modern open source languages
- Trusted application operations
- Built for continuous integration and delivery
- Simple horizontal and vertical scalability



# Let's see how it works

## Investment on Bonds

**Comparing Profit of Sovereign Gold Bond  
with General Bonds (Indian Railway Finance  
Corporation & SBI Life (Unit II Regular) -  
Bond Fund)**

how many years you plan for investment:

5

Predict

<https://compare-bonds.herokuapp.com/>

# Conclusion

- SBI Life (Unit II Regular) - Bond Fund is giving almost 111.51% of profit percentage for next 8 years (2021-29), this is the highest amongst all the 3 investments.
- As per SBI Life Insurance risk factor for SBI Life (Unit II Regular) - Bond Fund is Low to Medium.
- Indian Railway Finance Corporation is giving almost 11.09% of profit percentage for next 8 years (2021-29). This is the lowest amongst all the 3 investments, and risk factor is High.
- Sovereign Gold Bond is giving almost 82.64% of profit percentage for next 8 years (2021-29)., and risk factor is Low to Medium.
- As per our observations, we recommend SBI Life (Unit II Regular) - Bond Fund for investment, it's giving highest profit with low risk factor.

# Future Scope

- Similar implementation techniques could be applied to multiple company bonds (eg: Reliance, Tata).
- Data if provided by the client would help in understanding the requirements of the client and enhance the model performance Model improvisation will result in better accuracy.
- Deep diving into feature which may add value to the project and improve accuracy.
- Tailor made options as per customer's needs could result in best investment and good returns.

# Thank you

