Time Series Forecasting for Gold Price

Supervisor: Prof. Monika Bhattacharjee, Department of Mathematics, IIT Bombay

Name: Pritam Naskar
Roll No: 24N0071

Name: Shweta Verma
Roll No: 24N0058

Background

Gold prices are influenced by multiple factors, including:

- Global economic conditions
- Currency exchange rates
- Inflation trends
- Geopolitical events

Time series forecasting allows identification of trends, seasonality, and volatility.

Project Goals

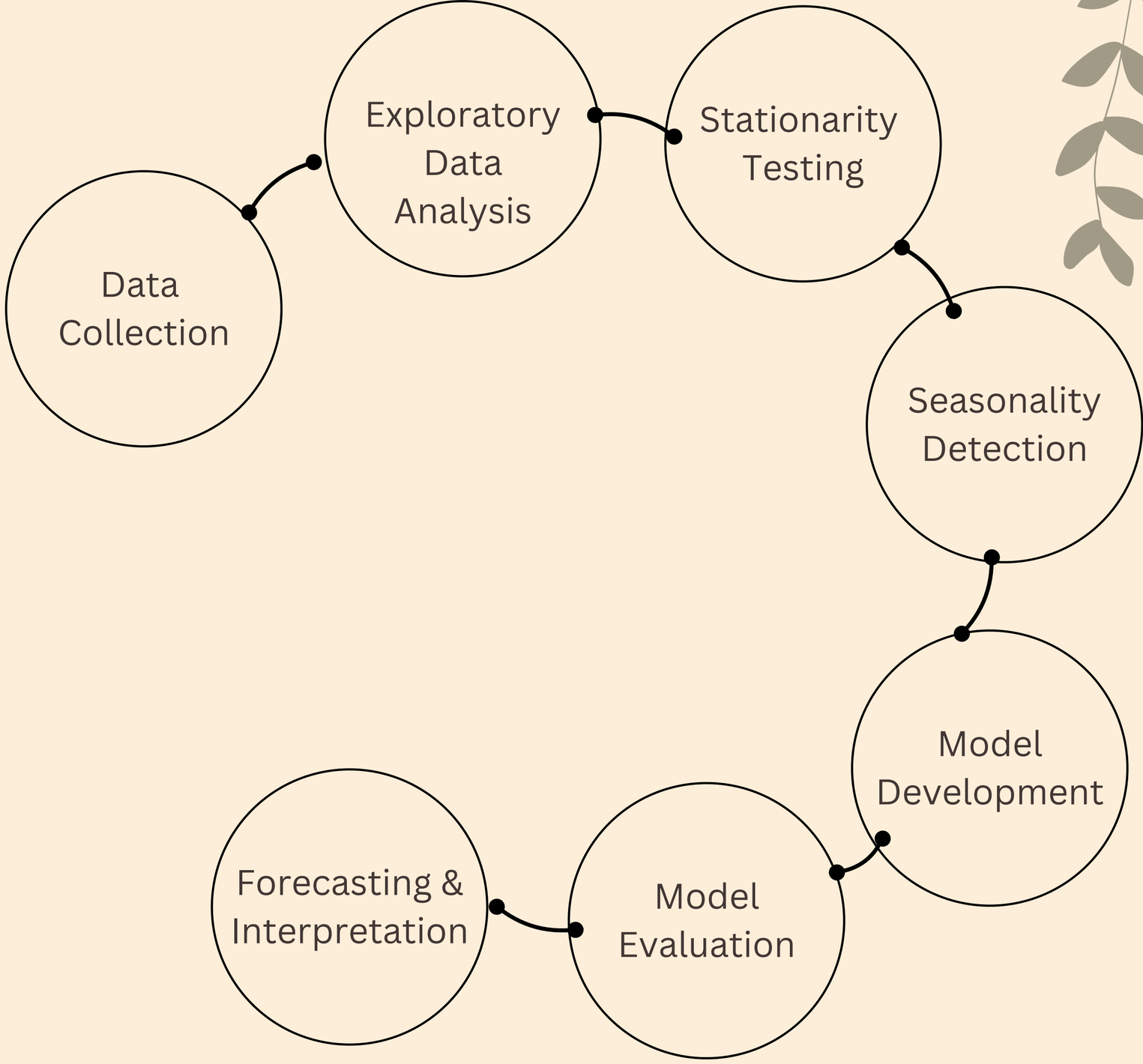
Primary Goal:

Forecast gold prices to assist decision-making in investment and trading.

Specific Objectives:

- Identify trends, seasonality, and noise in gold price data.
- Build and compare statistical forecasting models.
- Evaluate model performance using error metrics.

Methodology



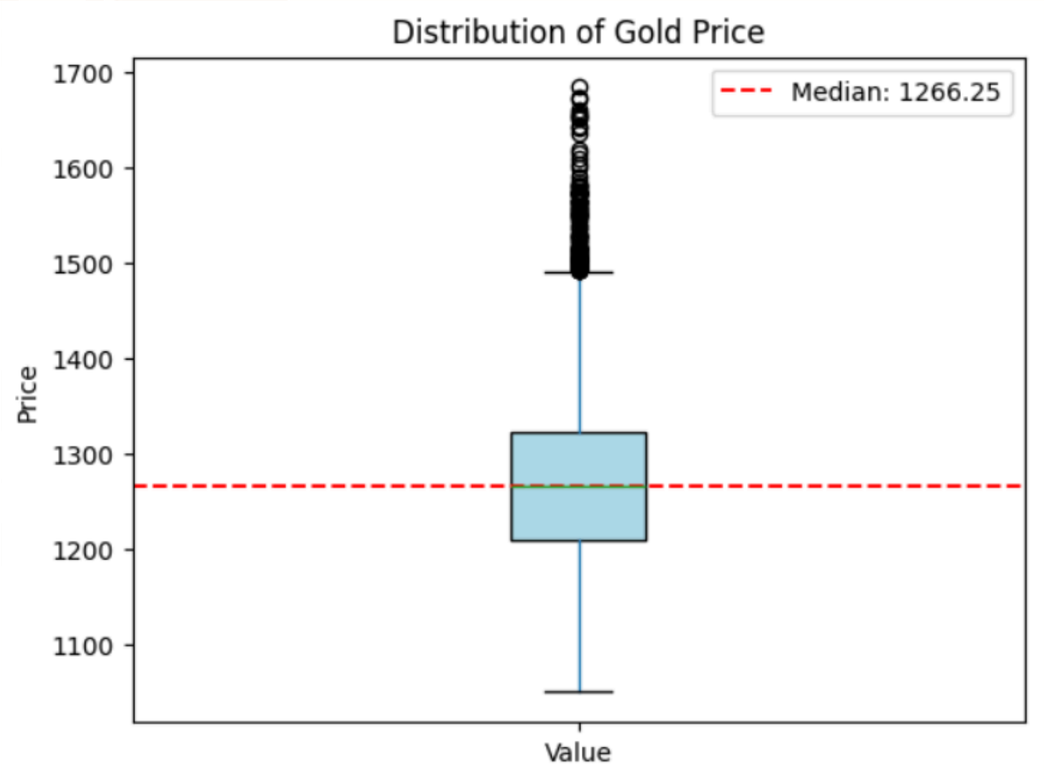
Data Description

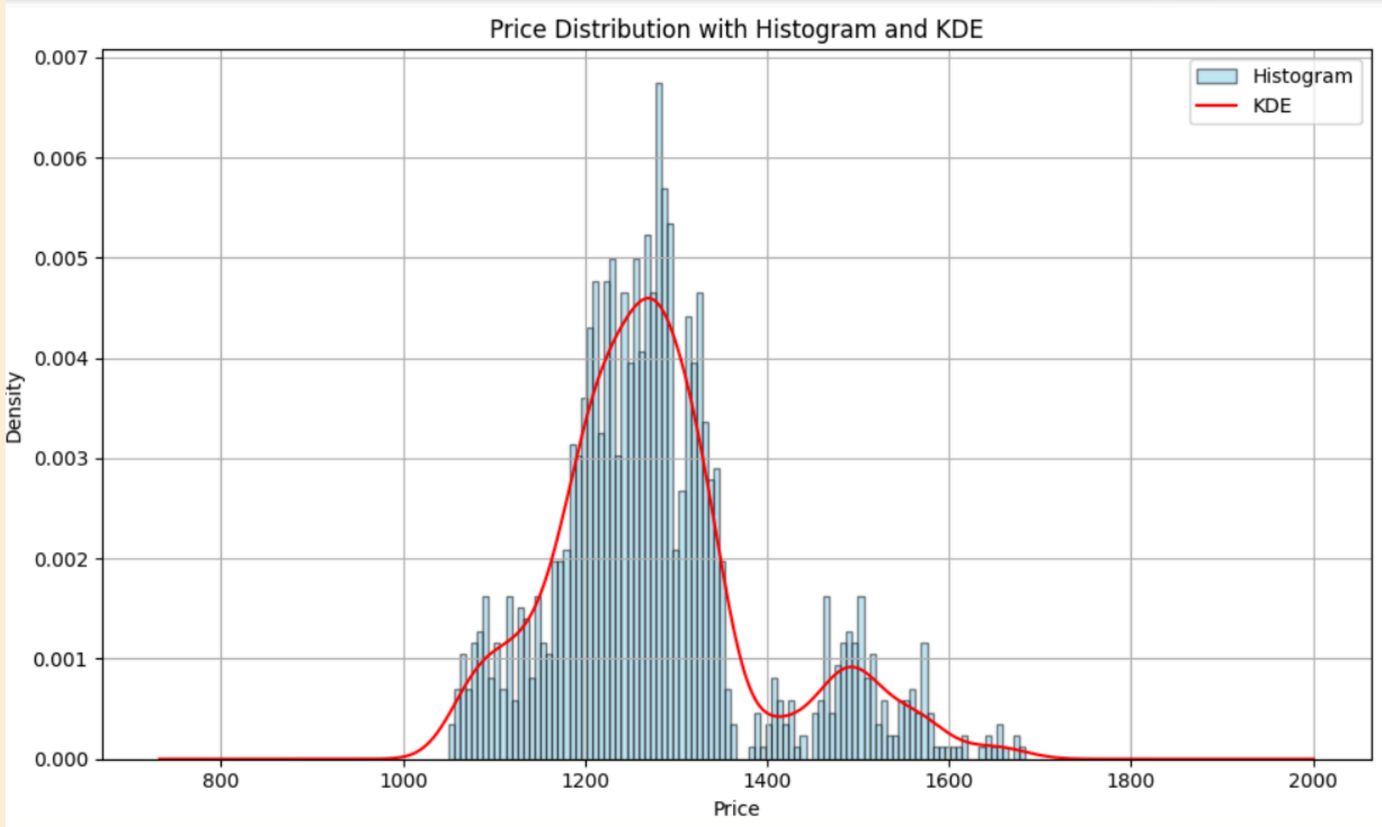
- **Source:** Kaggle Repository
- **Frequency:** Daily
- **Period:** Last 5 years (2015-2020)
- **Variables:** Date and Gold Price(per USD)

- **Initial Observations:**
  - Clear upward trends
  - Seasonal fluctuations present

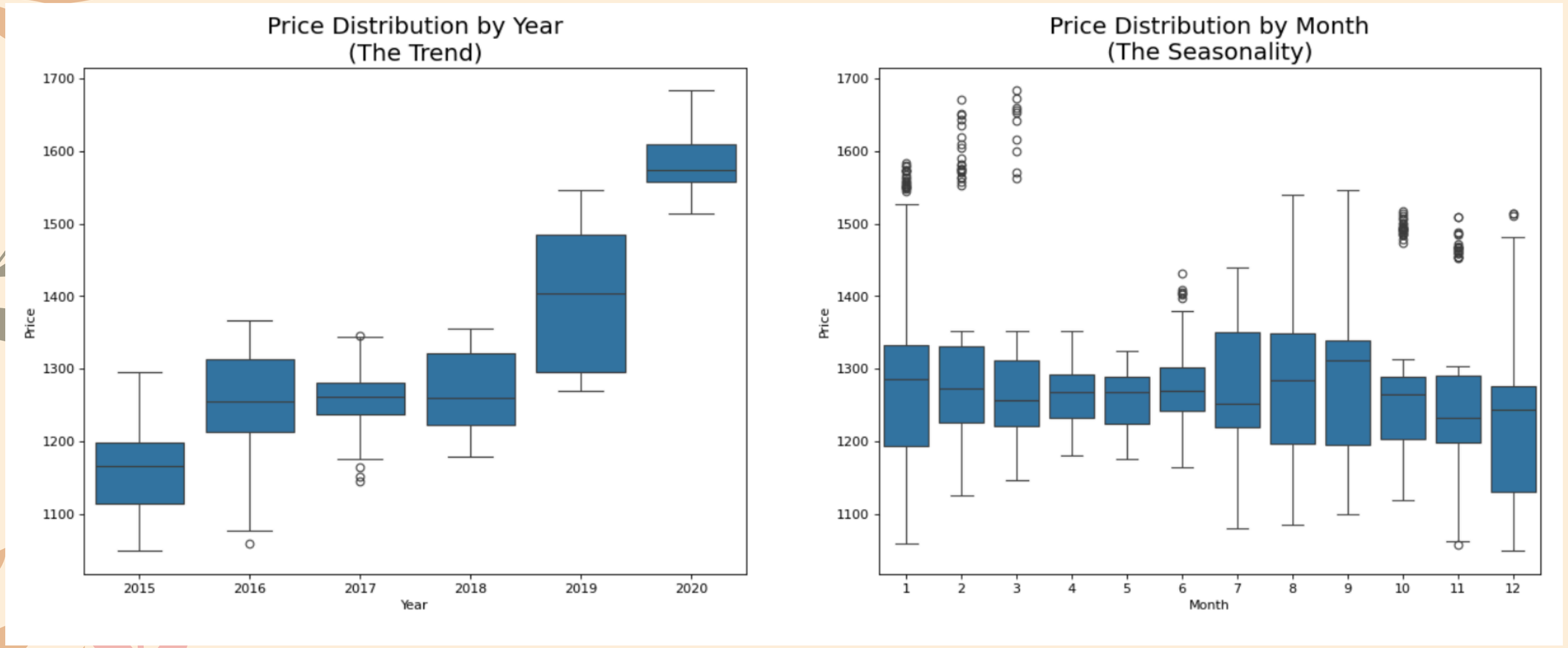
Exploratory Analysis

Exploratory Analysis

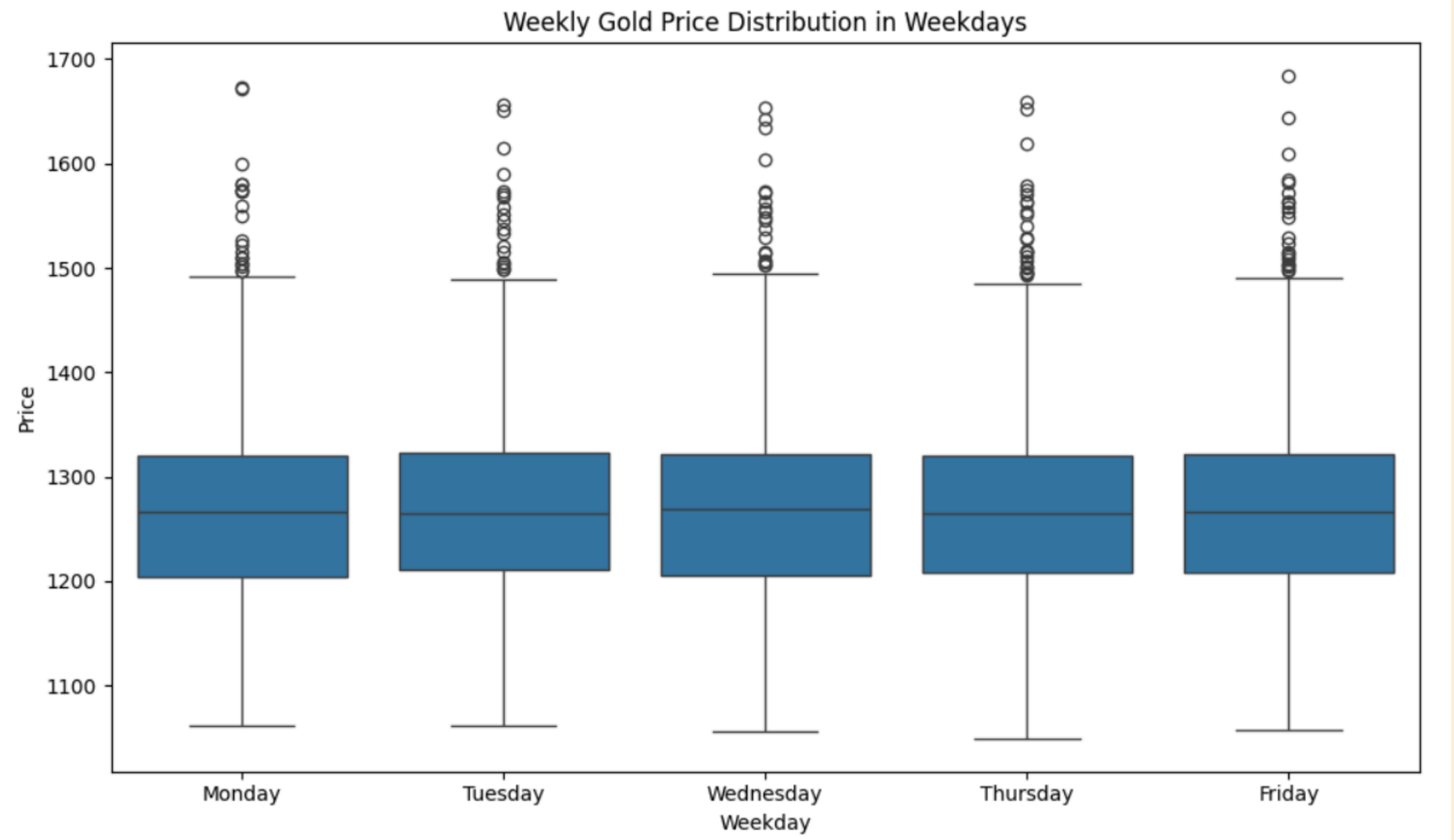




Exploratory Analysis

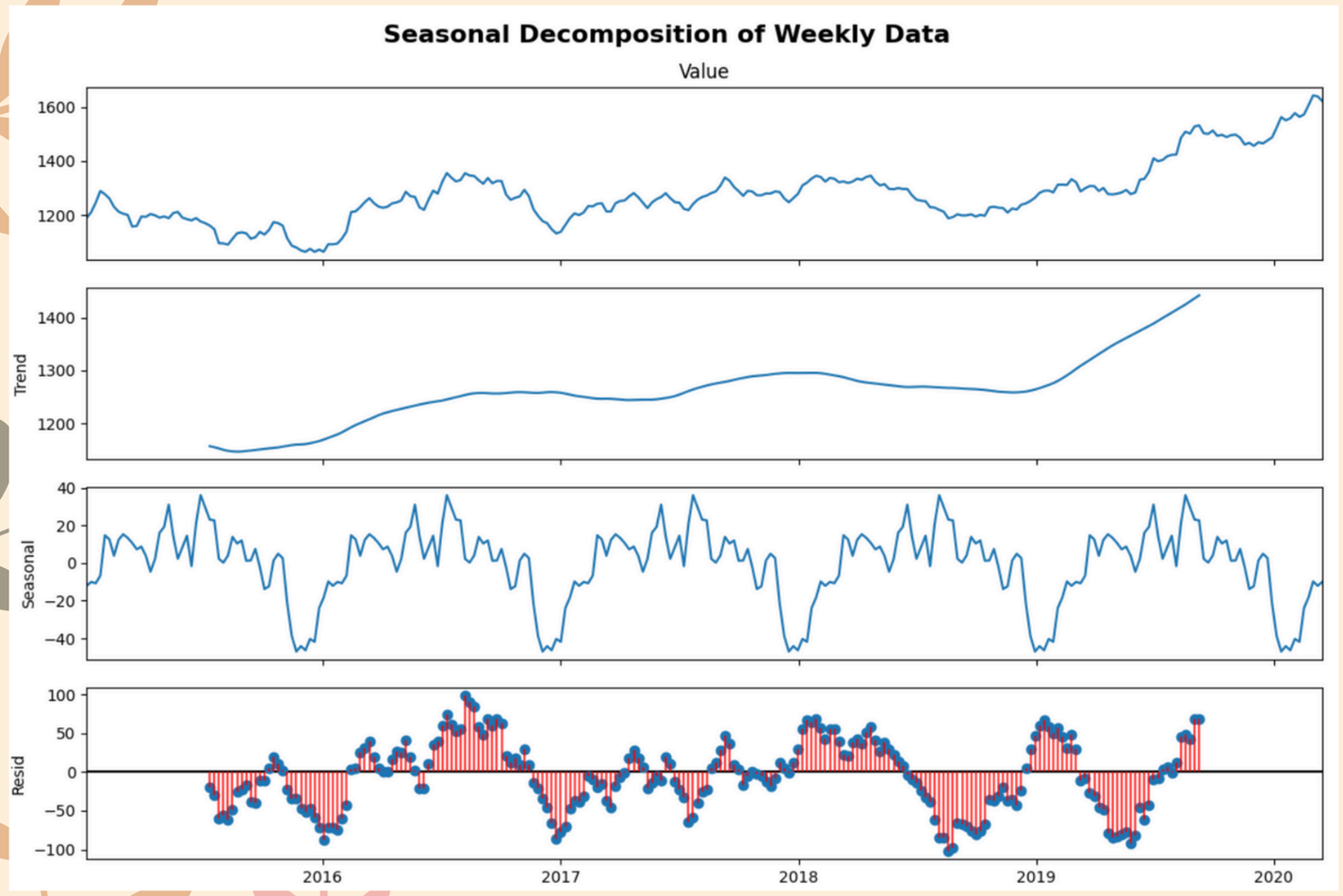


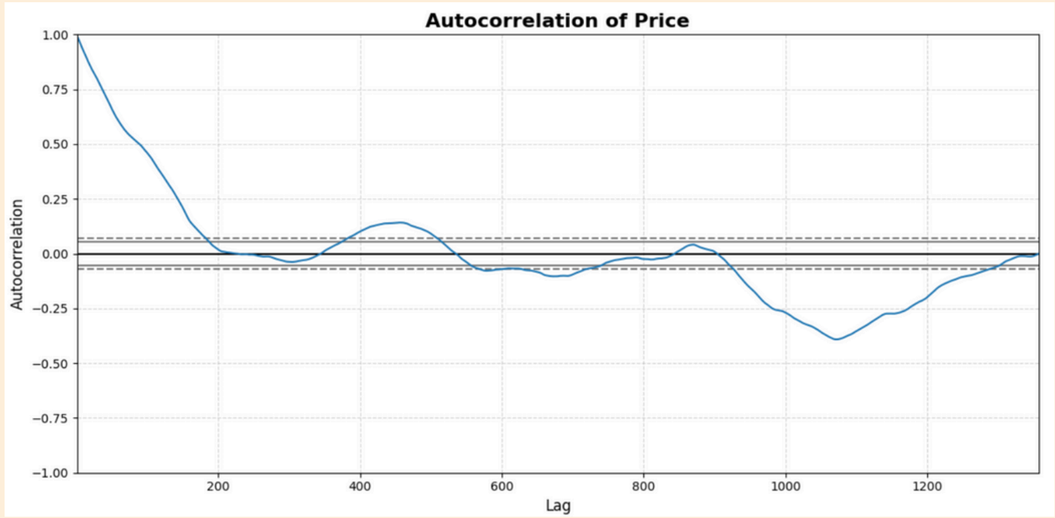
Exploratory Analysis



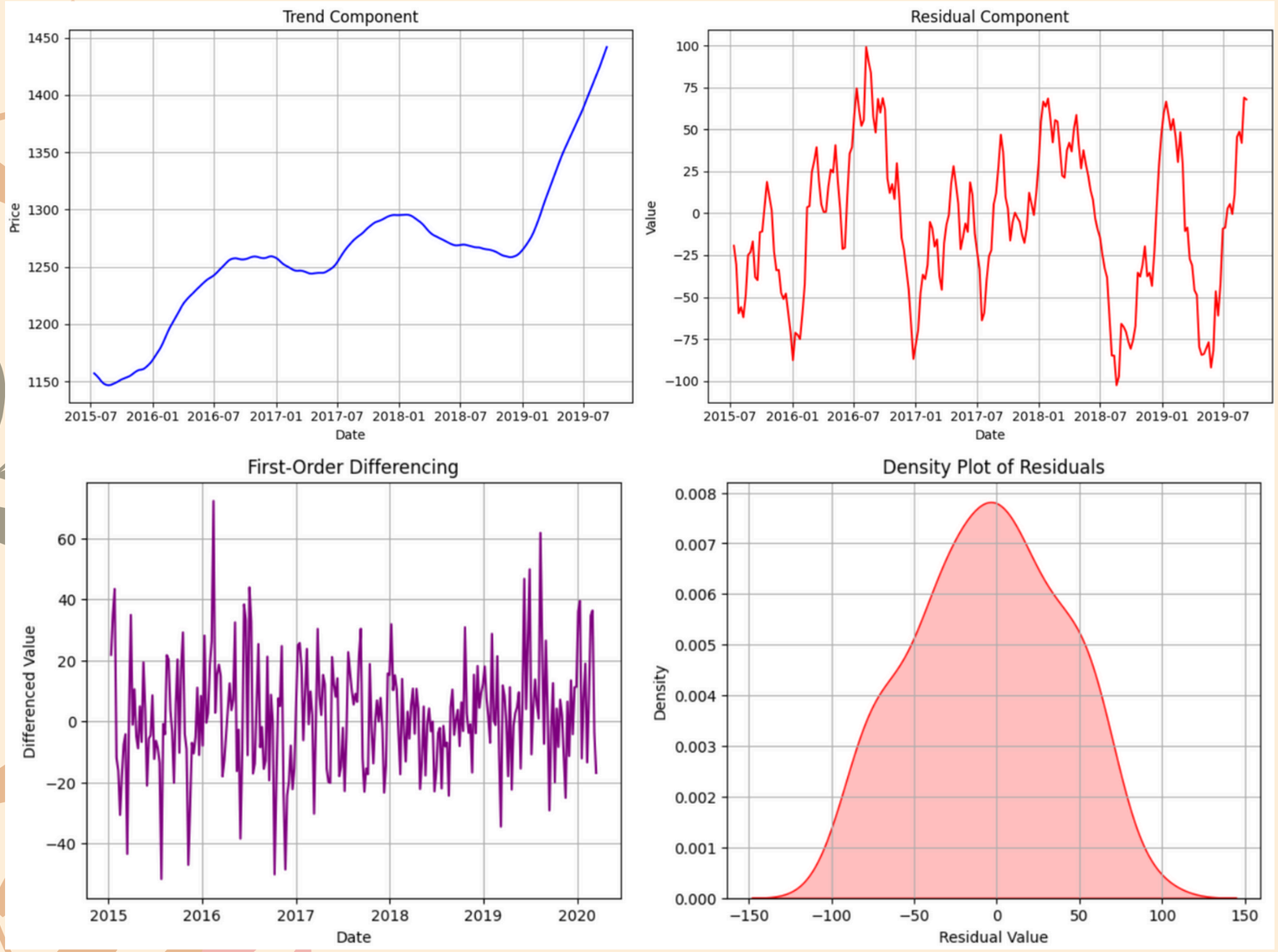
Exploratory Analysis

Exploratory Analysis





Exploratory Analysis



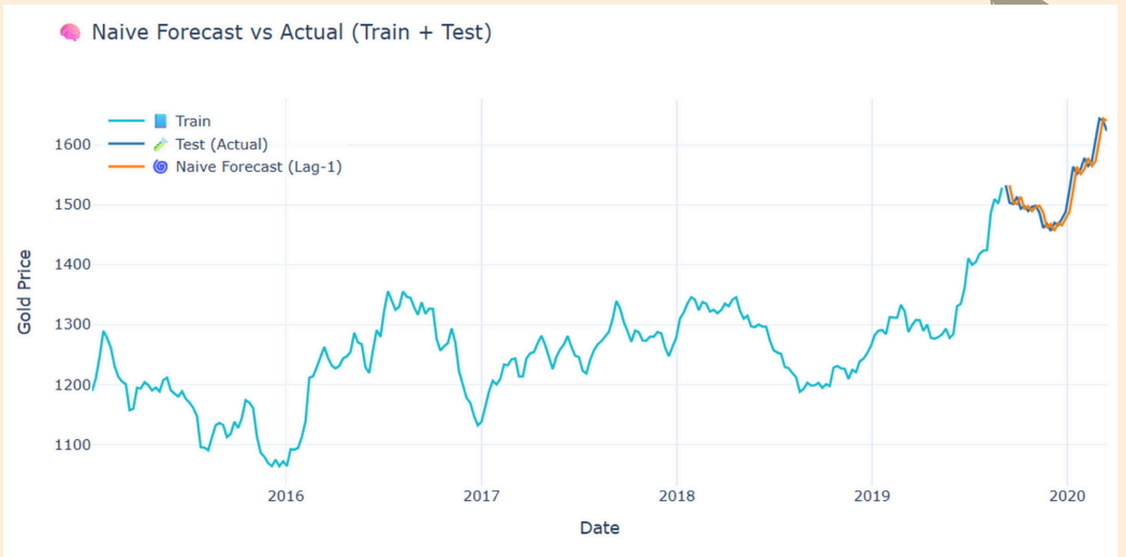
ADF Test:

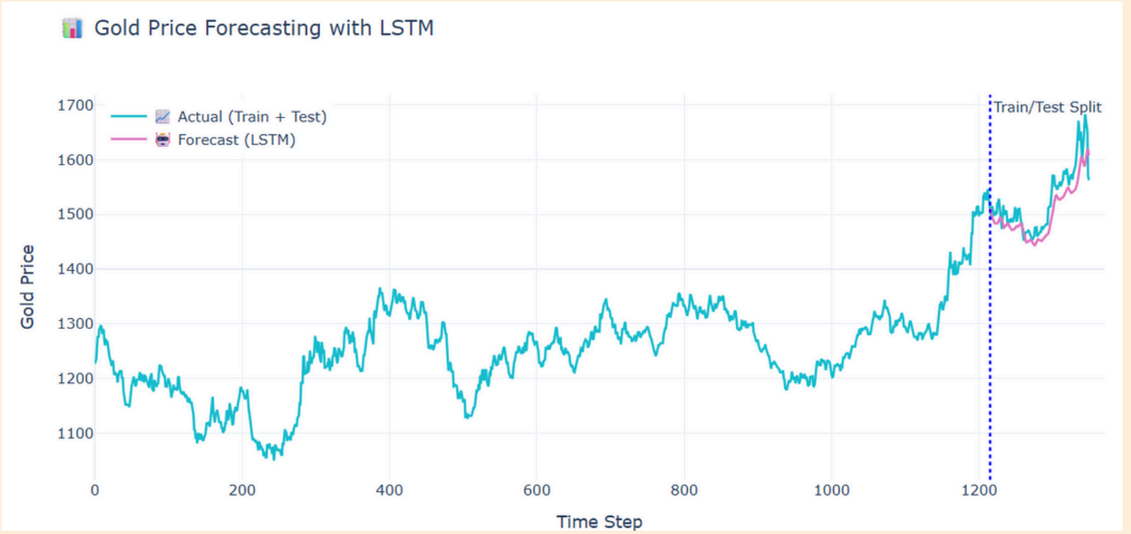
```
Before Differencing:
-----------------------------
ADF Statistic: -0.15778424031160418
p-value: 0.9433700996880114

After Differencing:
-----------------------------
ADF Statistic: -8.0551140788005
p-value: 1.7007407036919953e-12
```

Ljung-Box Test:

```
Ljung-Box Test Results:
      lb_stat  lb_pvalue
10  20.673555   0.023488
20  35.824932   0.016126
52  70.321536   0.046068
```

Exploratory Analysis





Exploratory Analysis

Forecasting Models: Performance & Statistical Significance

Model	MAE	RMSE	MAPE (%)
Naive	15.11	18.7	0.98
LSTM	26.25	32.09	1.69
Holt-Winters	28.33	38.49	1.81
SARIMA	29.67	37.44	1.91

Result

- The naïve model achieved the lowest errors, demonstrating strong short-term predictability.
- LSTM produced higher errors, likely due to limited tuning or data.
- Holt-Winters and SARIMA performed similarly, but with less accuracy, with SARIMA performing slightly worse.
- Simpler models outperformed more complex approaches in this case.

Conclusion

- **Key Findings:**
  - Differencing was required for stationarity
  - Naive Model yielded the best performance
  - Predictions achieved acceptable accuracy levels
- **Applications:**
  - Portfolio management, trading decisions, risk mitigation
- **Limitations:**
  - External shocks (e.g., war, pandemics) are not considered

Future Work

- Incorporate macroeconomic indicators (inflation rate, USD index)
- Test advanced ML models (FB Prophet)
- Explore intraday data for high-frequency forecasting

Thank You