

Case Study:

Historical Stock Analysis



[LINK TO KAGGLE NOTEBOOK](#)

Objective

Examine the closing price trends of large-cap companies listed on the NSE to identify patterns, fluctuations, and long-term performance insights. Exploring how macroeconomic factors and sectoral dynamics influence the stock prices of high-market-cap companies. Pinpoint stocks with consistent growth or volatility to help investors and analysts understand market leaders and laggards.

Dataset Overview

This dataset contains the historical closing price data for all stocks listed on the National Stock Exchange (NSE) of India with a market capitalization exceeding 500 crore INR. This dataset contains 1418 attributes and 8469 entries with data types as float64 for 14147 attributes and object for 1 attribute.

Features	Description
Stock Ticker:	Unique symbol representing each stock.
Date:	The specific trading date.
Closing Price	The price at which the stock closed on a given day.

Methodology

1. Handling Missing Values

Filling null values
using fillna()
feature using
method bfill.

2. Datetime Conversion

The date column
was converted to
datetime dtype.

3. Dropping null values

Dropping null values
which cannot be
imputed by setting
a threshold
value(0.5) of
missing entries

4. Outliers handling

Capping outliers
with upper and
lower bound.

5. Time Series decomposition

Understanding trend,
seasonality and
residual for ICICI
BANK stocks

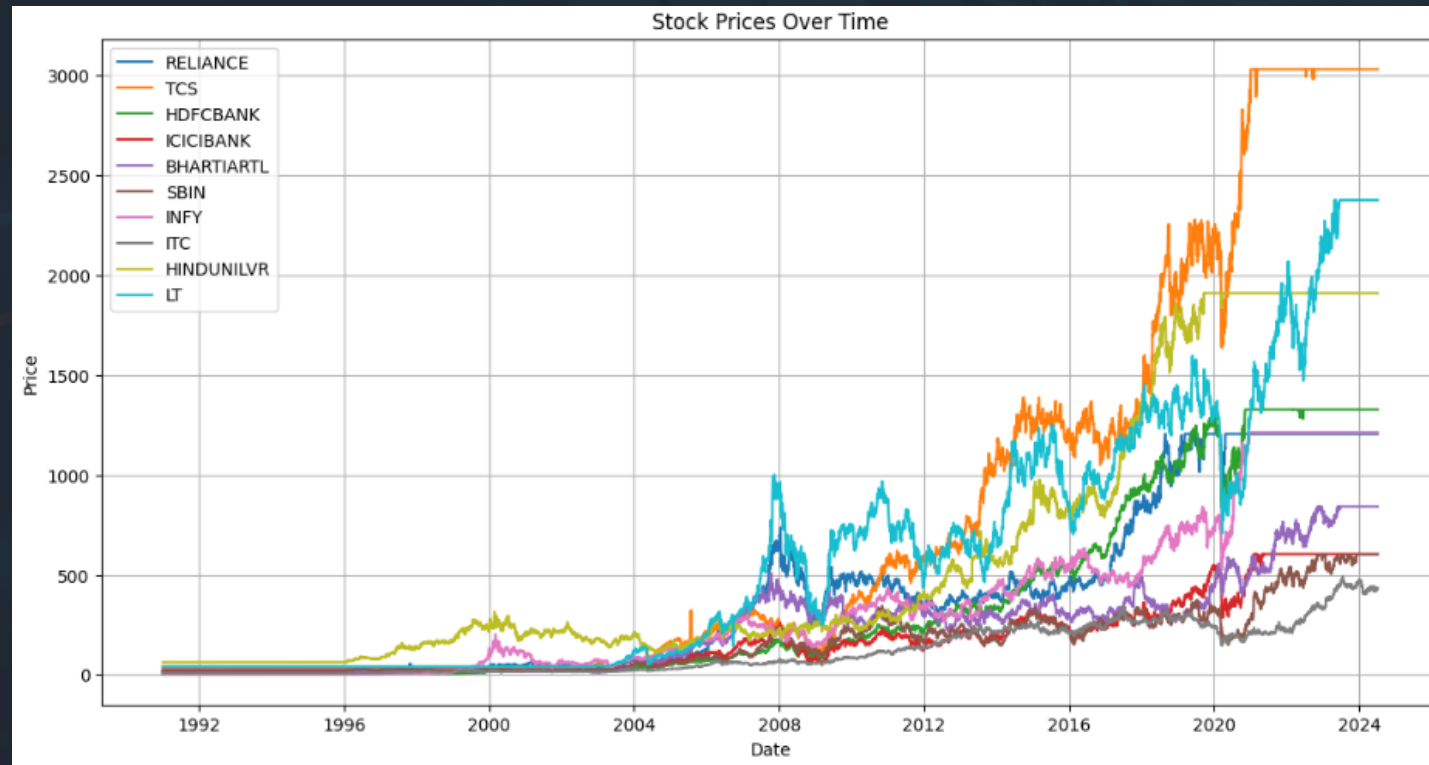
6. Anomaly detection

Detecting anomalies
using Isolation forest
and understanding
the reason behind it.

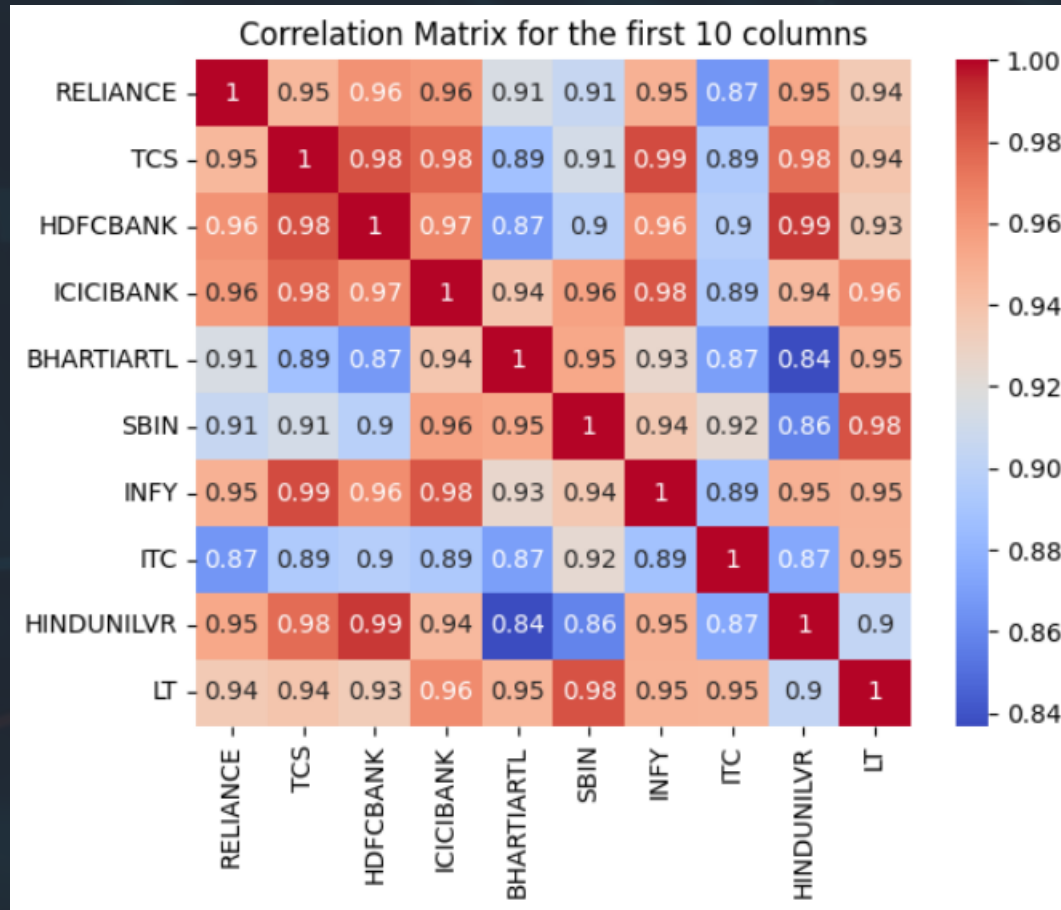
Statistical Analysis

- The dataset spans from 1991-01-02 to 2024-07-05, covering 33+ years of stock performance data.
- **Top 3 Average Stock Prices:**
 - **TCS:** ₹ 754.71
 - **LT:** ₹ 589.21
 - **HINDUNILVR:** ₹ 585.61
- These figures indicate that on average, TCS had the highest stock price among the listed companies.
- **Lowest Price:**
 - Infosys recorded the lowest stock price at ₹0.76, likely reflecting its early-stage valuation.
 - Other companies such as HDFC Bank and SBI also had very low starting prices, signaling substantial long-term growth.
- Companies like **TCS (₹988.47)** and **L&T (₹589.21)** show high standard deviation, indicating greater volatility in stock prices over time.

- On the contrary, companies such as **Bharti Airtel (₹184.24)** and **SBI (₹159.25)** exhibit relatively lower price volatility.
- The gap between the median and the mean indicates **skewness**, likely caused by significant price increases in later years.
- Over the 33 years, all companies experienced significant growth:
 - Reliance grew from a minimum of ₹10.98 to ₹1,205.24.
 - TCS rose from ₹35.47 to ₹3,031.64.
 - HDFC Bank started at ₹2.43 and reached ₹1,327.41.



Correlation Analysis



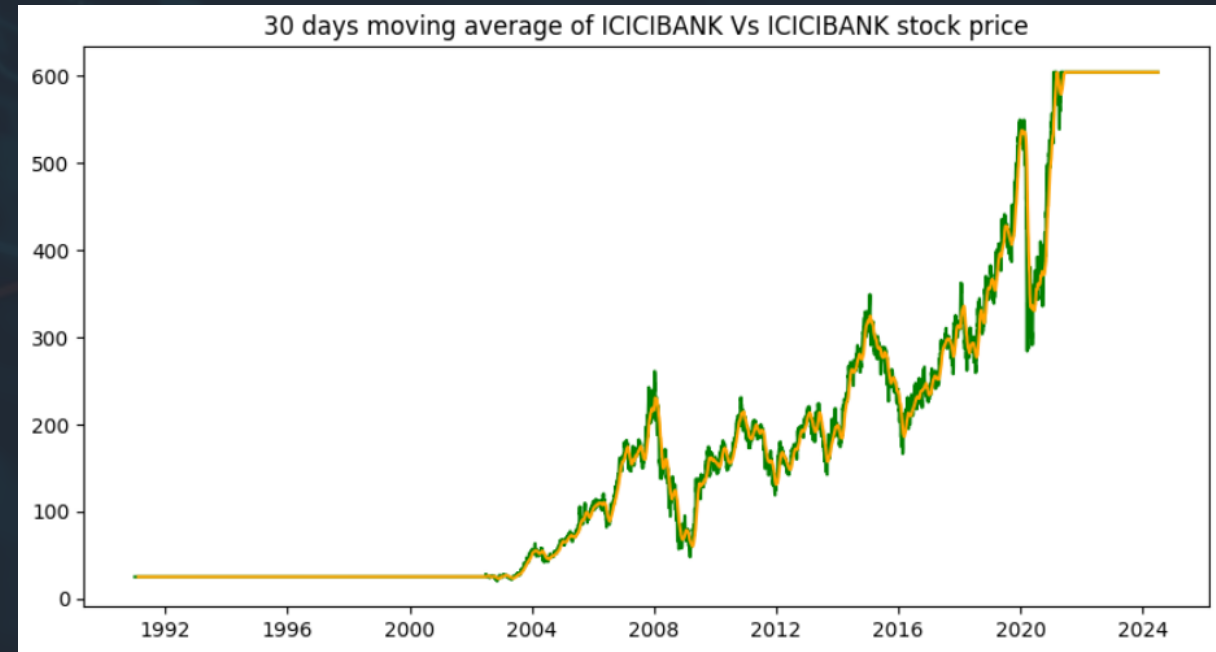
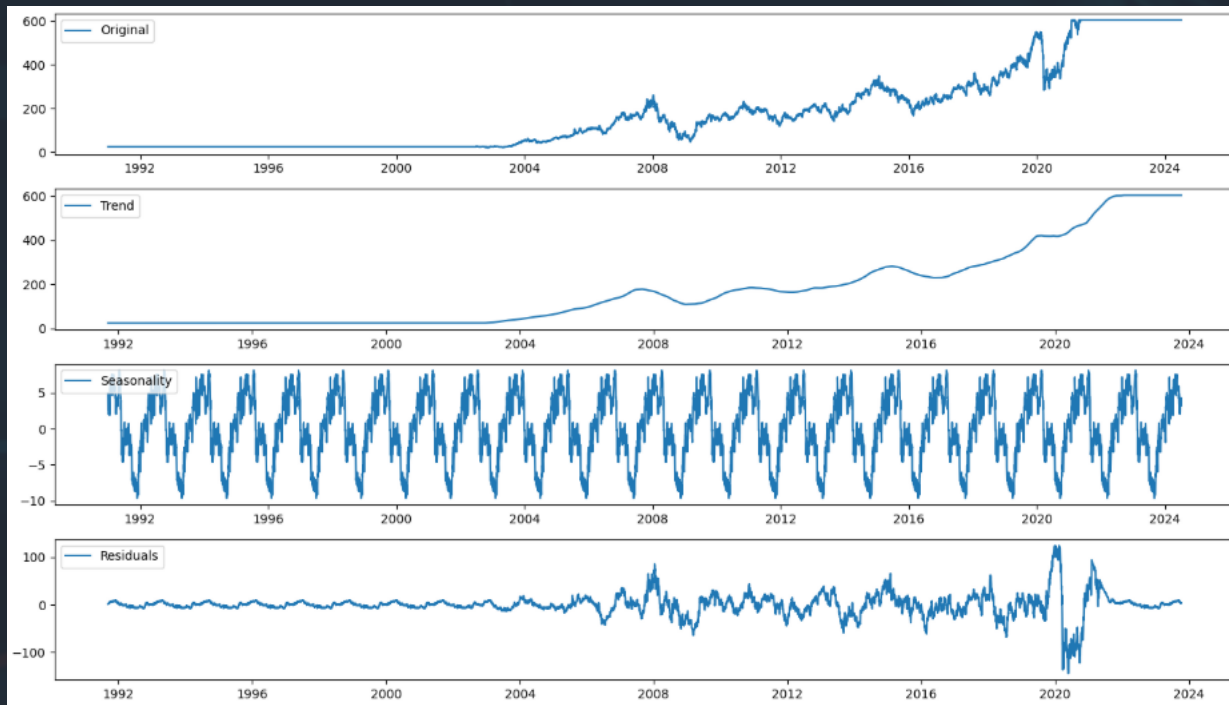
Bharti Airtel exhibits slightly lower correlation with other stocks, reflecting its distinct sectoral influences (telecom vs. IT or banking).

Banking Stocks (HDFC Bank, ICICI Bank, SBI): Strong inter-correlation, reflecting their similar responses to macroeconomic factors.

IT Stocks (TCS, Infosys): High correlation (0.99), likely due to their similar industry dynamics.

Time series decomposition

Stock- ICICI Bank



Trend:

- This component represents the upward movement in the stock price.
- There is a clear **upward trend** after a half-way, suggests that ICICIBANK stock value has been rising over time indicating a **long-term positive growth**.

Seasonality :

- This component captures a short-term repeated pattern in the data, which are consistent over the year.
- The cyclical pattern shows regular troughs and peaks in the data indicating that the data exhibits **seasonal effects**, which may be due to factors like the **macroeconomic conditions**.

Residual :

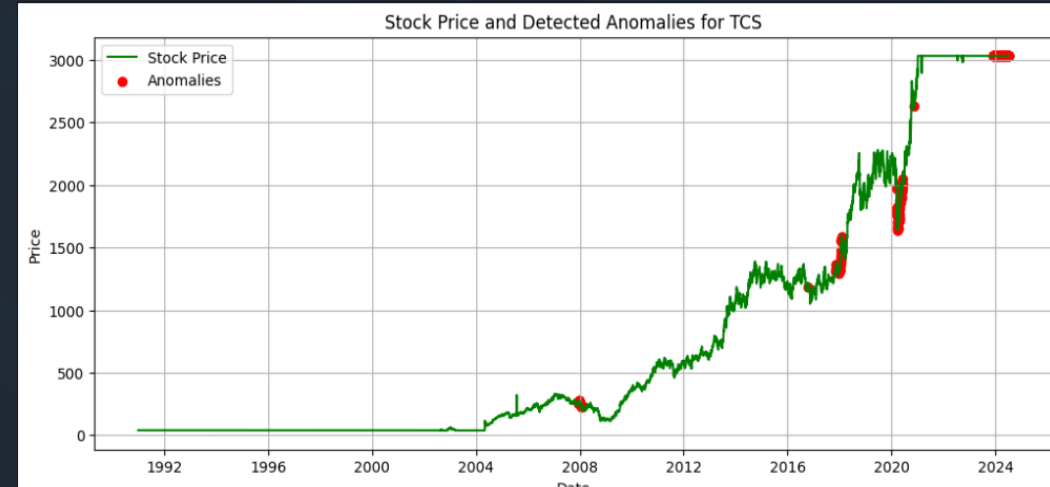
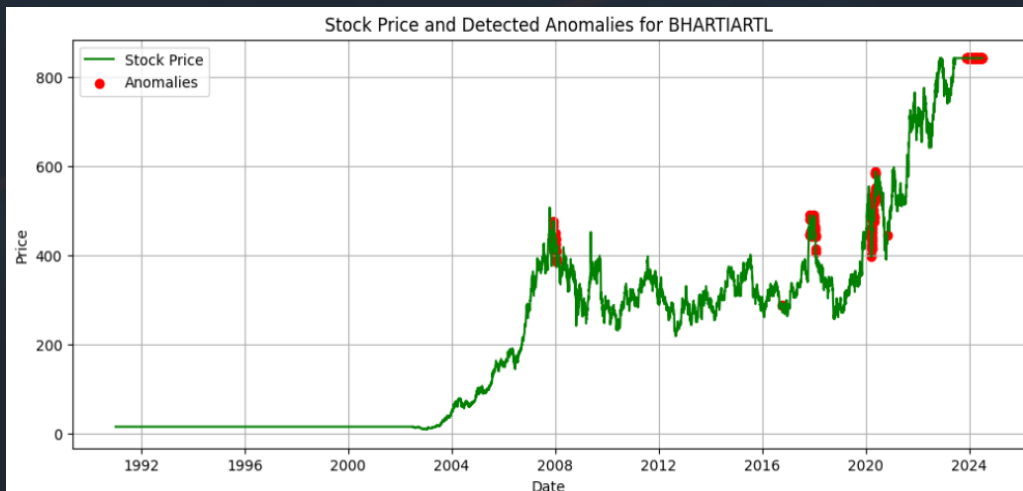
- This component captures the irregular fluctuations and noise in the data.
- The residual is somewhat constant at the beginning and later shows **higher volatility** which may be due to market events.

The upward trend suggests that the "**ICICIBANK**" stock has been **growing** over the long term. Traders may consider employing trend-following strategies, such as buying on dips during an **uptrend** and holding positions for the long term.

Anomaly Detection

Used **Isolation Forest** to detect anomalies in each stock of the data.

Here are the first 4 stocks anomaly-:



Reasons for anomaly

Most of the anomalies can be seen in year 2020 and 2008.

Few reasons for 2020 anomalies are:

- **Covid-19** outbreak can one of the major reasons for **anomalies(2020)** in the dataset.
- In the initial timeline of covid-19, rapid spread and **lockdown** caused panic among the investors, leading to sharp decline in stock prices.
- In the latter phase of covid-19(**recovery-period**), the reopening of market caused **volatility and rapid change** in the stock prices.

Few reasons for the 2008 anomalies are:

- The **financial crisis** of 2008 adversely impacted businesses, economies, and stock markets.
- This day is referred to as **Black Monday**.
- The widespread fear that the economy of the **USA** might go into **recession**
- A drop in the **interest rates** in the **US**
- Volatility in the commodity markets

Conclusion

The NSE historical stock market analysis revealed significant anomalies during March 2020, driven by global events like the COVID-19 pandemic. Key stocks such as RELIANCE, TCS, HDFCBANK, ICICIBANK, and BHARTIARTL contributed to these deviations, reflecting their impact on market movements. The analysis highlights the importance of monitoring blue-chip stocks and sectoral trends to understand market behavior, manage risks, and make informed investment decisions.

[LINK TO KAGGLE NOTEBOOK](#)