

TCS Stock Data Analysis and Prediction Project

Project Title : TCS Stock Data - Live and Latest

Domain : Data Analytics, Machine Learning

Tools & Technologies Used : Python, SQL, Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn

Project Objectives

- Analyze historical stock performance of Tata Consultancy Services (TCS).
- Identify trends, patterns, and correlations in stock price movements.
- Build predictive models to forecast closing prices using historical data.
- Evaluate model performance using appropriate metrics.
- Provide investment-related insights based on findings.

Dataset

Historical TCS stock data containing the following columns:

- Date
- Open, High, Low, Close prices
- Volume of shares traded
- Dividends
- Stock Splits

Data covers multiple years of trading activity, enabling both trend analysis and predictive modeling.

Data Cleaning & Preprocessing

- Handled missing values: Checked for NaN entries and filled using forward-fill method.
- Converted data types: Date column converted to datetime format.
- Sorted data by date to maintain chronological order.
- Converted price columns (Open, High, Low, Close) to numeric.
- Outlier detection: Checked volume and close price distributions.

Exploratory Data Analysis (EDA)

1. Price Trends

- Plotted Close Price over time to visualize historical movement.
- Calculated 30-day moving averages to smooth out fluctuations.
- Implemented moving average crossover strategy to generate buy/sell signals.

2. Volume & Corporate Actions

- Analyzed Volume trends over time.
- Compared Dividends and Stock Splits to price movement.
- Found weak correlation between dividends/stock splits and daily closing prices.

3. Correlation Analysis

- Open, High, Low, and Close showed near-perfect correlation (>0.99).

- Volume had a negative correlation with price (-0.15), indicating higher volume spikes during price drops.

4. Price Change Distribution

- Calculated Daily % Price Change and plotted histogram with KDE.
- Found normal-like distribution with occasional extreme spikes.

Feature Engineering

- Extracted Year, Month, Day, Day of Week from date.
- Created Lag Features: Previous day's close price (Prev_Close).
- Added Short-term & Long-term Moving Averages.
- Created Daily Price Change % column for volatility analysis.

Predictive Modeling

Model 1: Linear Regression

- **Features Used:** Open, High, Low, Volume, Prev_Close, Day_of_Week, Month
- **Target:** Close price
- **Performance Metrics:**
 - Mean Squared Error (MSE) : 39.05
 - R-squared (R^2) : 0.99994

Model 2: LSTM (Long Short-Term Memory)

- Designed a sequential model with:
 - LSTM layer (50 units)
 - Dense output layer (1 unit)

- **Performance:**

- Mean Absolute Error (MAE): 3.86

Observation: LSTM provided a more flexible modeling approach for sequential patterns but still showed room for improvement in predictive accuracy.

Visualizations

- Close Price over time
- Volume, Dividends, and Stock Splits over time
- Correlation heatmap
- Moving average crossover buy/sell signals
- Close Price vs Volume scatter plot
- Distribution of Daily % Price Change
- Actual vs Predicted Close Price (LSTM)

Insights & Key Findings

- TCS stock prices exhibit **strong positive correlation** among open, high, low, and close.
- **Volume tends to increase during price drops**, possibly indicating panic selling or profit booking.
- Moving averages can identify **potential buy/sell signals** but require backtesting for reliability.
- LSTM models can capture sequential dependencies but need **more hyperparameter tuning** for better accuracy.

- Historical data trends suggest **steady long-term growth** despite short-term volatility.

Conclusion

The TCS stock analysis showed steady long-term growth with short-term volatility. High correlation among price variables indicates consistent market movement, while moving averages offer potential trading signals. The predictive LSTM model demonstrated high accuracy with a low Mean Absolute Error of 3.86, indicating strong capability in capturing time-series patterns. Overall, the project highlights how data analytics and machine learning can provide valuable and reliable insights for stock forecasting.