# TCS Stock Data Analysis – Project Report

## 1. Introduction

This project focuses on the analysis of **Tata Consultancy Services (TCS)** stock data using Python. The aim is to understand historical stock price behavior, identify meaningful patterns, and apply basic analytical techniques that are commonly expected from a **Data Analytics Intern**.

The project is designed to reflect a **real-world analytical approach**, keeping the methodology practical, the explanations simple, and the insights business-oriented. Advanced or overly complex techniques have been intentionally avoided to maintain clarity and authenticity.

---

## 2. Business Context

Tata Consultancy Services (TCS) is one of India's largest IT services companies and a key constituent of Indian stock market indices. Stock price analysis of such a company provides insights into: - Long-term growth trends - Market volatility - Trading activity and investor behavior

Understanding these aspects is valuable for analysts, investors, and financial institutions.

---

## 3. Objective of the Project

The primary objectives of this project are: - To analyze historical stock price data of TCS - To perform data cleaning and validation - To conduct exploratory data analysis (EDA) - To identify trends using visualizations - To engineer simple analytical features - To build a basic predictive model for closing price estimation

This project is **analytical in nature**, not intended for real-time trading or financial advice.

---

## 4. Dataset Description

The analysis uses historical stock price data of TCS. The dataset contains daily trading information with the following key columns:

- **Date**: Trading date
- **Open**: Opening price
- **High**: Highest price of the day
- **Low**: Lowest price of the day
- **Close**: Closing price
- **Volume**: Number of shares traded
- **Dividends**: Dividend value (if any)

• **Stock Splits**: Stock split information

The dataset spans multiple years, making it suitable for long-term trend analysis.

---

# 5. Tools and Technologies Used

- **Programming Language**: Python
- **Libraries**:
- Pandas – data manipulation
- NumPy – numerical operations
- Matplotlib & Seaborn – visualization
- Scikit-learn – machine learning
- **Environment**: Jupyter Notebook / VS Code

These tools are widely used in the data analytics industry and suitable for academic as well as internship-level projects.

---

# 6. Data Cleaning and Preprocessing

Before analysis, the dataset was prepared carefully:

- Converted the `Date` column to datetime format
- Sorted the data chronologically
- Verified data types for numerical columns
- Checked for missing values (none were found)
- Ensured consistency in financial figures

This step ensured that the dataset was reliable and ready for analysis.

---

# 7. Exploratory Data Analysis (EDA)

## 7.1 Stock Price Trend Analysis

The closing price of TCS stock was plotted over time. The visualization revealed: - A strong upward trend over the long term - Periods of volatility during major market events - Consistent recovery after downturns

This indicates long-term value creation by the company.

---

**7.2 Trading Volume Analysis**

Trading volume showed frequent spikes, especially during high-volatility periods. However: - High volume did not always correspond to sharp price changes - This suggests institutional trading and long-term investor participation

---

**7.3 Correlation Analysis**

A correlation matrix was generated to understand relationships between variables:

Key observations: - Open, High, Low, and Close prices are highly correlated - Volume shows weak correlation with price - Dividends and stock splits have minimal short-term impact on price

This aligns with real-world financial behavior.

---

# 8. Feature Engineering

To enhance analysis, additional features were created:

## 8.1 Moving Averages

- 30-day moving average
- 100-day moving average

These were used to smooth price fluctuations and identify trend direction.

## 8.2 Daily Percentage Return

Daily percentage change in closing price was calculated to understand volatility and return distribution. The results showed: - Returns centered around zero - Presence of extreme values (market shocks)

---

# 9. Predictive Modeling

A **Linear Regression model** was implemented to estimate the closing price.

## 9.1 Model Inputs

Features used: - Open price - High price - Low price - Trading volume

Target variable: - Close price

**9.2 Model Evaluation**

     • Mean Squared Error (MSE)
     • R-squared score ($R^2$)

The model performed well due to strong correlation among price variables. However, it is important to note:
- The model is suitable for learning purposes - It does not account for external market factors

---

# 10. Results and Insights

     • TCS stock demonstrates strong long-term growth
     • Price-related features are strong predictors of closing price
     • Volume alone is not a reliable indicator of price movement
     • Simple models can explain historical patterns but not future uncertainty

---

# 11. Limitations of the Project

     • External factors such as news, global markets, and macroeconomic indicators were not included
     • Time-series forecasting models were not explored
     • The predictive model is not suitable for real trading decisions

These limitations were consciously accepted to keep the project focused and realistic.

---

# 12. Conclusion

This project successfully demonstrates a complete **data analytics workflow**, starting from data understanding to insights and basic prediction. It reflects the practical skills expected from a Data Analytics Intern, including:

     • Data cleaning
     • Visualization
     • Analytical thinking
     • Basic modeling

The project maintains a balance between technical depth and business understanding, making it suitable for academic submission and internship evaluation.

---

# 13. Future Scope

Possible extensions of this project include: - Time-series forecasting (ARIMA, LSTM) - Inclusion of external financial indicators - Dashboard development using Power BI or Tableau - Risk and volatility analysis

---

**End of Report**