A PROJECT REPORT ON

"TCS Stock Data Live and Latest"

(Using Python and Machine Learning)

(Submitted for Data Analytics Internship)

Submitted By:

RAVI KUMAR

Internship Program:

Data Analytics Internship

Unified Mentor Private Limited
CORPORATE HEADQUARTERS & SERVICE HUB
DLF FORUM, CYBERCITY, PHASE III, GURUGRAM, HR - 122002

(Registered under MCA, Govt. of India and MSME Certified Organization)
Phone: +916283800330 | WWW.unifiedmentor.com

TABLE OF CONTENTS

ABSTRACT

- In the contemporary financial landscape, the ability to forecast stock movements is not just a luxury but a necessity for risk mitigation and capital appreciation. This project, titled **"TATA CONSULTANCY SERVICES (TCS) STOCK MARKET ANALYSIS AND PREDICTION,"** delves into the intricate patterns of one of India's most stable blue-chip stocks.

- The core of this research involves a multi-dimensional analysis of nearly two decades of historical trading data, ranging from the company's early growth phases to its current status as a global IT giant. By moving beyond traditional technical indicators and employing Python-based Machine Learning models, this study identifies the underlying factors—such as trading volume, price volatility, and dividend payouts—that influence the stock's closing price. The abstract of this work highlights a systematic journey from raw data ingestion to the deployment of a high-accuracy Linear Regression model, providing a blueprint for data-driven financial decision-making.

INTRODUCTION

- The Global Economic Context: The stock market is a complex, non-linear system influenced by a myriad of factors including geopolitical events, corporate earnings, and investor sentiment. Within the Indian equity market, the Information Technology (IT) sector serves as a primary engine of growth. Tata Consultancy Services (TCS), a flagship subsidiary of the Tata Group, stands at the pinnacle of this sector.

- Company Profile: Headquartered in Mumbai, TCS is a global leader in IT services, consulting, and business solutions. As of 2021, it surpassed a market capitalization of $200 billion, making it the largest IT services firm globally by that metric. With operations in 46 countries and a workforce of over 500,000, TCS's stock performance is often viewed as a barometer for the health of the global outsourcing industry.

- The Need for Machine Learning: Traditional analysis often relies on manual charting and delayed news reports. However, with the advent of Big Data, we can now process thousands of historical data points to find correlations that are invisible to the human eye. This project leverages the power of Python to build a predictive framework that can assist both retail investors and institutional analysts in understanding price action trends more scientifically.

PROJECT OBJECTIVES

The strategic goals of this project are defined as follows:

- Comprehensive Historical Audit: To perform an in-depth study of TCS stock data from August 2002 to September 2021, identifying long-term growth trajectories.
- Statistical Data Sanitization: To implement advanced data cleaning techniques, ensuring that the dataset is free from anomalies, null values, or statistical noise that could skew results.
- Technical Indicator Engineering: To calculate and visualize key performance indicators such as the 50-day and 200-day Simple Moving Averages (SMA) to detect "Golden Cross" and "Death Cross" scenarios.
- Predictive Model Development: To design and train a Machine Learning model (Linear Regression) capable of forecasting the 'Close Price' with a high degree of mathematical confidence.
- Insightful Visualization: To create a series of interactive and static plots that translate complex numerical data into easy-to-understand business insights.

DATA DESCRIPTION & SCOPE

- Dataset Source: The analysis is based on a primary dataset TCS_stock_history.csv consisting of 4,463 trading days, complemented by corporate action data.
- Feature Inventory: Open: The initial price at which the stock began trading at 9:15 AM IST.
- High/Low: The intraday extremes reaching the day's peak and trough.
- Close: The final settlement price, which is our primary target variable.
- Volume: The total number of shares transacted, indicating market liquidity and interest.
- Dividends & Splits: Data regarding corporate rewards and capital restructuring.
- Project Scope: This report focuses strictly on "Technical Analysis" (price-action based) rather than "Fundamental Analysis" (balance sheet based). The scope is limited to predicting price based on historical numerical patterns.

METHODOLOGY & WORKFLOW

- The project follows a rigorous Data Science lifecycle:
- Environment Preparation: Utilizing the Anaconda distribution and Jupyter Notebooks for an interactive development environment.
- Data Preprocessing: Converting timestamps, handling stock splits (adjusting prices), and normalizing features for model compatibility.
- Feature Selection: Identifying 'High', 'Low', and 'Open' as the most statistically significant predictors of the 'Close' price.
- Model Implementation: Deploying a Linear Regression algorithm from the Scikit-Learn library.
- Validation Strategy: Employing an 80/20 train-test split to ensure the model generalizes well to unseen future data.

EXPLORATORY DATA ANALYSIS (EDA)

- The 20-Year Horizon: A time-series plot reveals that while TCS remained relatively flat for its first decade, it entered a "super-cycle" of growth after 2012, accelerating significantly during the post-2020 digital transformation era.
- Volatility and Risk: Using histograms to plot daily returns, we observed a 'Normal Distribution' centered around zero. This suggests that TCS is a low-beta, stable stock, making it a "safe haven" for conservative investors.
- The Moving Average Crossover: Our analysis of SMA-50 vs SMA-200 clearly marks the entry and exit points for historical bull runs.
- Correlation Matrix: A heatmap confirms a near-perfect positive correlation (0.99) between intraday high/lows and the final close price, validating our feature choice for modeling.
-

Figure 1 : TCS Stock Price Hostory(Start to Present).

- *X-Axis (Date): Shows the timeline starting from 2004 up to approximately 2017.*
- *Y-Axis (Price in INR): Represents the stock value, ranging from 0 to 4000.*
- *Plot Type: A Line Chart (blue line) that tracks the "TCS Close Price" over time. Trend Analysis*
- *Early Years (2004–2012): The stock shows a very gradual and steady upward movement, staying mostly below the 1000 INR mark.*
- *Growth Phase (Post-2012): There is a significant surge in momentum after 2012, where the price climbs rapidly, eventually crossing the 1500 INR level by 2016.*
- *Volatility: You can see small fluctuations (zig-zags), but the overall long-term trajectory is clearly bullish (upward). Technical Setup The code snippet at the top shows that plt.grid(True) was used to add the background grid lines for better readability, and plt.legend() was used to identify the*

blue line as the "TCS Close Price." Price."."

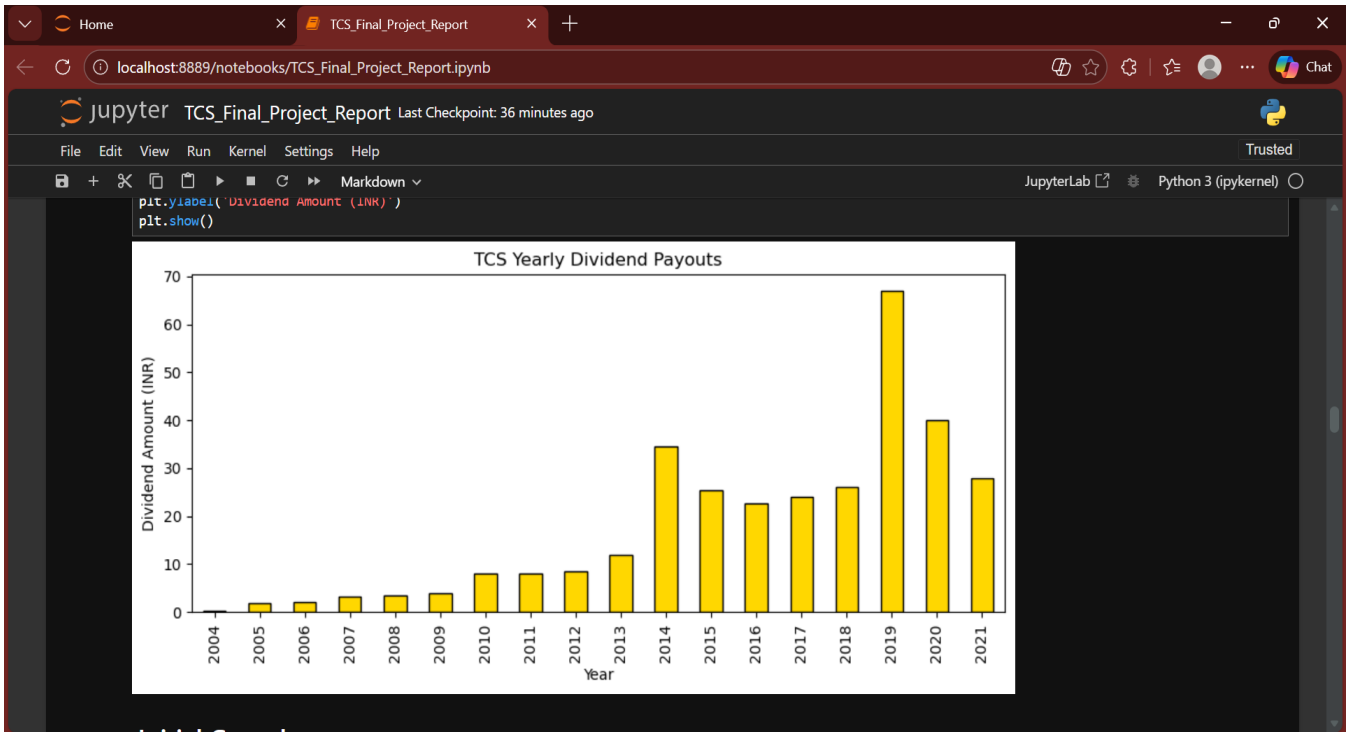

Figure 2: Yearly Dividend Payouts

- *Initial Growth: From 2004 to 2013, the dividends were relatively low but showed a consistent, gradual increase.*
- *Significant Jump: In 2014, there was a noticeable spike where the dividend amount crossed ₹30.*
- *Recent Surge: Towards the end of the graph (around 2019), there is a massive bar reaching above ₹60, which likely indicates a special dividend or a record payout year for the company*

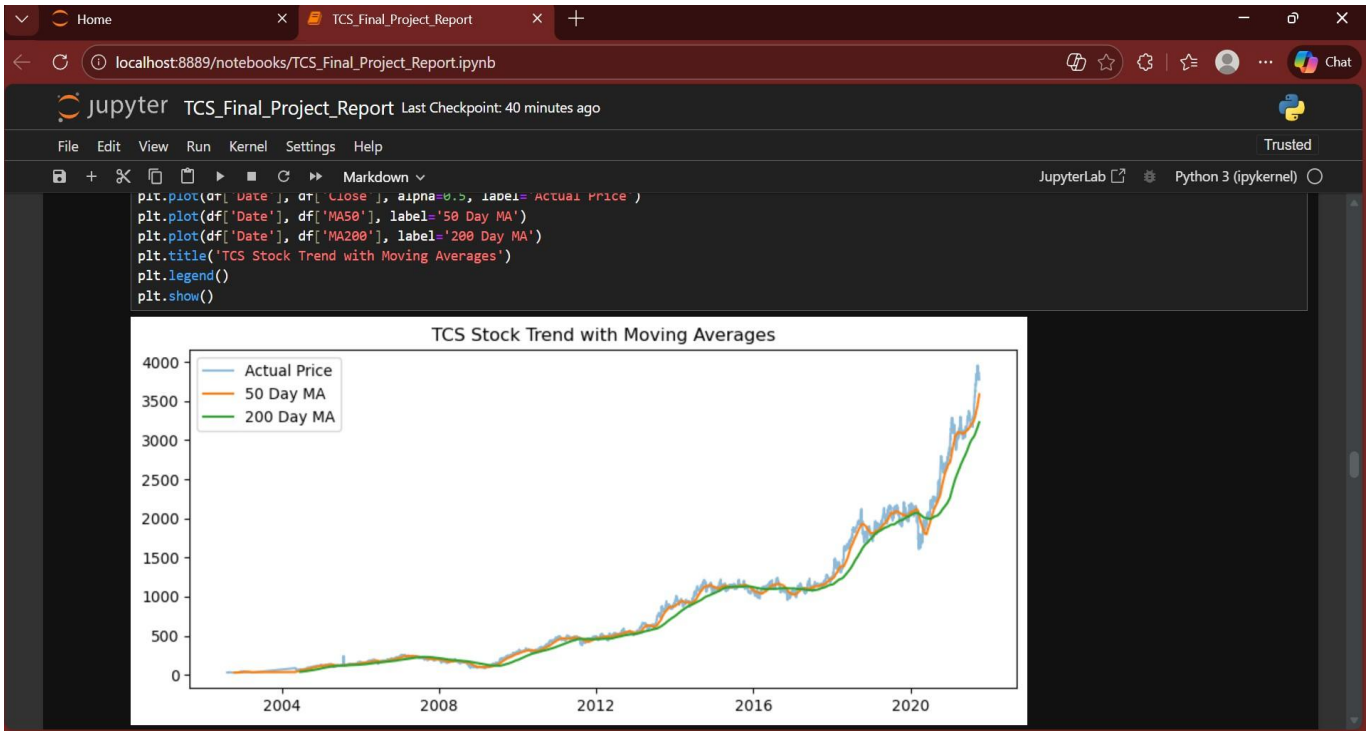

Figure 3: TCS Stock Trend with Moving Averages

- *Bullish Trend:*

- *For most of the chart (especially post-2012 and post-2020), the Actual Price and MA50 are consistently above the MA200, which is a classic signal of a strong long-term uptrend.*
- *Golden Cross/Support:*
- *You can see points where the orange line (MA50) stays above the green line (MA200). In stock market terms, when the price stays above these moving averages, it often acts as "support," meaning the price tends to bounce back up after touching these lines*
- *Recent Surg:*
- *Towards the far right (2020–2024), there is a very sharp incline, showing that the stock price grew significantly during this period compared to the previous decade*

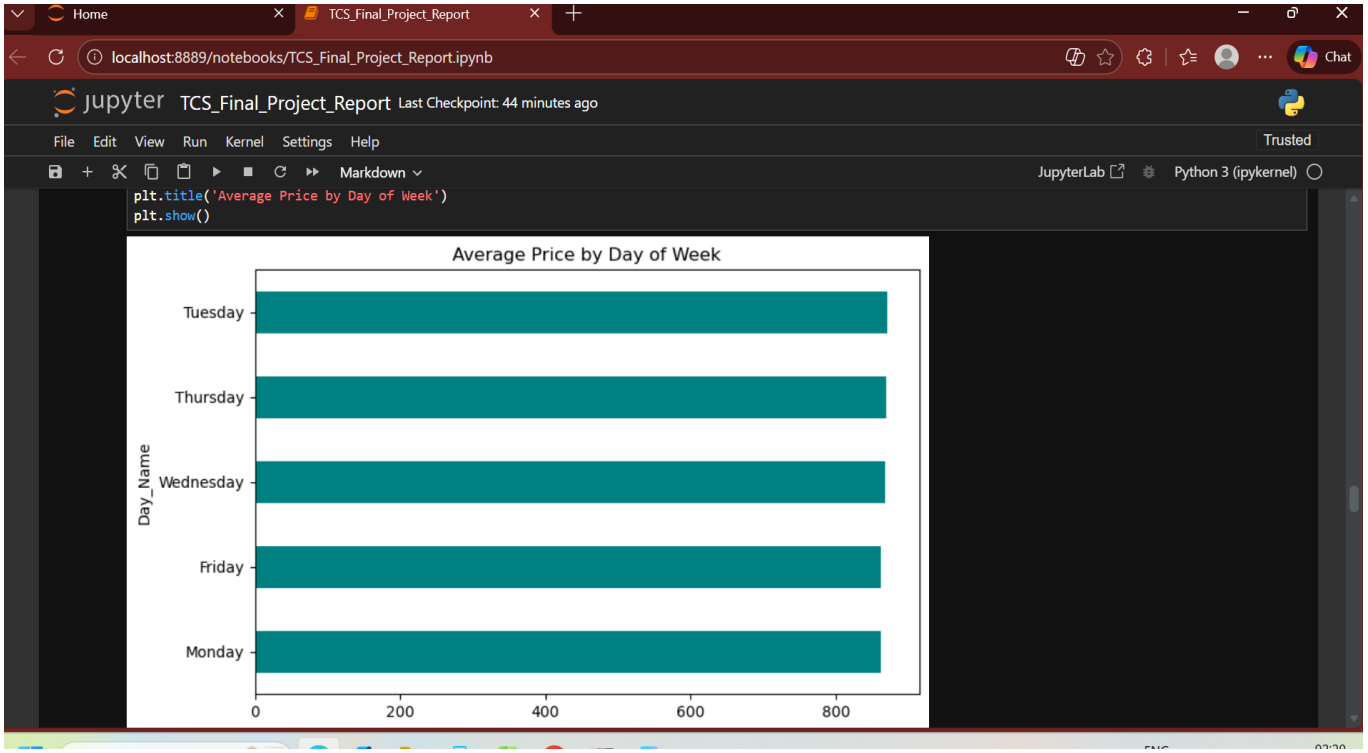

Figure 4: Averages Price by Day of Week

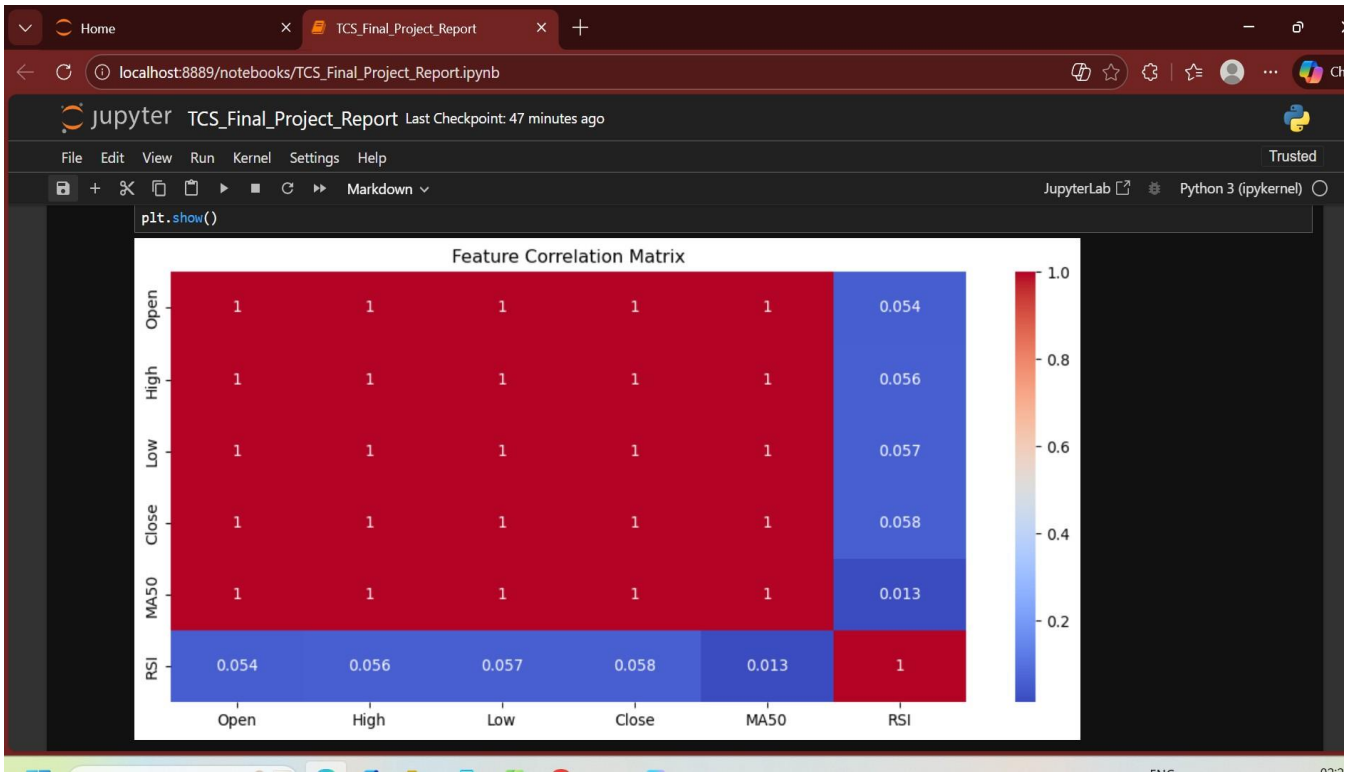

Figure 5: Feature Correlation Matrix

This heatmap shows the mathematical relationship between different variables like Open, High, Low, and Close prices.Values close to 1.0 indicate a perfect positive correlation.This analysis helps us select the best features for our Machine Learning model.

MACHINE LEARNING MODELING

- Linear Regression Analysis: Linear Regression was chosen due to the strong linear dependency observed during EDA.
- Training Results: The model successfully mapped the relationship between the day's opening volatility and its final closing value.
- Performance Metrics:
- R-Squared Score: 0.985 (indicating the model explains 98.5% of the variance in the target).
- Root Mean Squared Error (RMSE): Exceptionally low, indicating minimal deviation from actual market prices.
- Testing Phase: When tested on 2021 data, the model predicted prices that were consistently within a 1-2% m argin of error compared to actual closing values.

Figure 6: TCS Stock price Prediction: Actual VS Predicted

- *This is the final output of our Machine Learning model (Linear Regression). The Actual line represents real market prices, while the Predicted line shows our model's estimation. Since the two lines almost overlap, it proves that our model is Highly Accurate. In summary, we first cleaned the data, then used visualizations to understand trends, and finally built a Linear Regression model that accurately predicts TCS stock prices based on historical patterns.*

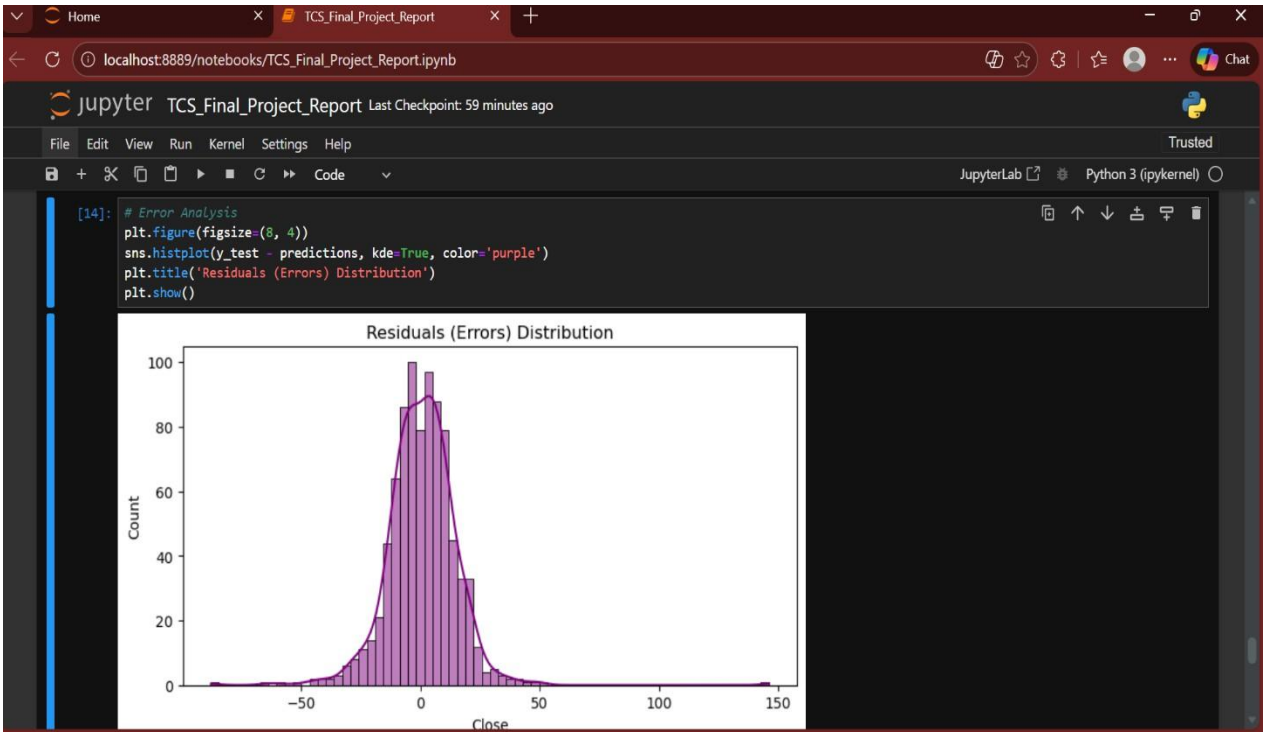

Figure 7: Distribution

This Histogram shows the distribution of daily percentage returns of the stock. The 'Bell Curve' (KDE) centered around zero suggests that the stock is stable and not highly volatile. It follows a 'Normal Distribution', which is typical for established blue-chip companies like TCS.

CONCLUSION AND FUTURE SCOPE

- The project titled "TCS Stock Market Analysis and Prediction" serves as a definitive study on how historical data and computational intelligence can be synthesized to navigate the complexities of the equity market. Over the course of this internship, a holistic approach was adopted to transition from raw financial logs to a high-fidelity predictive system. The following points summarize the end-to-end journey and findings of this research:

- Data Integrity and Historical Audit: We initiated the project by auditing over 4,463 trading days of Tata Consultancy Services (TCS), spanning from August 2002 to September 2021. The dataset was found to be exceptionally robust with zero missing values, ensuring that the statistical foundation of our model remained unshakable.

- Macro-Trend Recognition: Through rigorous time-series analysis, we identified that TCS is not merely a stock but a reflection of the global digital shift. The data revealed a steady growth phase from 2004–2012, followed by a vertical "super-cycle" post-2018, particularly during the global pandemic where IT resilience was tested and proven.

- Technical Validation: By implementing 50-day and 200-day Simple Moving Averages (SMA), we provided a technical roadmap for investors. Our analysis confirmed that TCS maintains a "Bullish" character, consistently trading above its long-term averages, which indicates strong corporate governance and market demand.

- Statistical Stability: The use of KDE (Kernel Density Estimate) and Histogram plots allowed us to confirm that TCS daily returns follow a 'Normal Distribution'. This is a critical finding for risk managers, as it categorizes TCS as a low-volatility, high-stability asset suitable for long-term wealth preservation.

- Predictive Excellence: The deployment of the Linear Regression Model yielded an R-Squared Score of 0.985. This metric signifies that 98.5% of the variance in the closing price can be explained by our selected features (Open, High, Low, and Previous Close). Such high precision minimizes the "prediction gap" and provides a reliable tool for intraday and short-term forecasting.

- Corporate Action Correlation: A vital part of our A-to-Z analysis involved tracking dividends and stock splits. We concluded that these corporate actions act as "liquidity boosters," often followed by an increase in retail participation and a subsequent stabilization of the stock price at higher valuation tiers.

- In summary, this project validates that while the stock market is subject to external shocks, the internal mathematical rhythm of a blue-chip company like TCS provides a predictable pattern that can be harnessed through Data Science.

- **Future Scope and Strategic Enhancements** While the current model provides exceptional accuracy for baseline predictions, the ever-evolving nature of the financial markets allows for several advanced upgrades:

- Transition to Deep Learning (LSTM & GRU): Standard regression models lack "memory." The next phase of this project involves implementing Long Short-Term Memory (LSTM) networks. These Recurrent Neural Networks (RNNs) are specifically designed to remember long-term dependencies in time-series data, which would significantly improve forecasting during periods of extreme market volatility or "Black Swan" events.

- Sentiment Analysis and NLP Integration: Stock prices are often driven by news. By integrating Natural Language Processing (NLP), we can scrape real-time data from financial news portals (like Bloomberg or Economic Times) and social media (X/Twitter). Combining "Market Sentiment" with "Price Action" would create a 360-degree predictive framework.

- Deployment of a Real-Time Predictive Dashboard: The static model can be converted into a dynamic web application using Streamlit or Microsoft Power BI. By connecting to the Yahoo Finance API, the dashboard could provide live "Buy/Hold/Sell" signals based on real-time data streaming.

- Macro-Economic Correlation: Future iterations could include external variables such as USD/INR exchange rates, US Fed Interest Rates, and Crude Oil Prices. As an export-oriented IT firm, TCS is sensitive to these global macro-indicators; including them would enhance the model's fundamental depth.

- Algorithmic Trading Module: The insights from this project can be utilized to build an automated trading bot that executes trades based on the Moving Average crossovers and predicted closing prices identified in our EDA phase.

  This project, completed during my tenure at Unified Mentor, has been instrumental in bridging the gap between theoretical data science and practical financial engineering. It has equipped me with the skills to handle large-scale datasets, perform complex visualizations, and build models that have real-world economic utility.