

# TCS Stock Data Analysis and Prediction Report

## 1. Project Overview

This project analyzes historical stock data of Tata Consultancy Services (TCS) to identify patterns, perform feature engineering, and build a linear regression model to predict the stock's closing price.

## 2. Dataset Description

- Source: TCS\_stock\_history.csv
- Features: Date, Open, High, Low, Close, Volume, Dividends, Stock Splits
- Target: Close price
- Total entries: 4,463
- Time span: From 2002 to recent

## 3. Data Preprocessing

- Checked for and filled missing values
- Converted columns to numeric where needed
- Forward filled NaNs using ffill

## 4. Exploratory Data Analysis

- Visualized Close price over time
- Computed 50-day and 200-day moving averages to reveal trends

## 5. Feature Engineering

- Extracted: Year, Month, Day, Day of Week from Date
- Created lag feature: Previous day's close price

## 6. Model Building and Evaluation

- Used Linear Regression to predict 'Close' price
- Features: Open, High, Low, Volume, Prev\_Close, Day\_of\_Week, Month
- Split: 80% training, 20% testing
- Evaluation:
  - Mean Squared Error: Measures average squared difference between predicted and actual values
  - R-squared Score: Indicates model fit

## 7. Result Summary

The model was able to predict TCS's stock close prices with reasonable accuracy. Predicted vs. actual price scatter plots show moderate linear correlation. Further improvements can be achieved with advanced models like Random Forest or LSTM.

## 8. Future Work

- Try more complex models (Random Forest, XGBoost, LSTM)
- Perform hyperparameter tuning
- Use time series forecasting techniques like ARIMA