Financial Forecasting through Predictive Analytics: Insights into Netflix's stock price evolution

TEAM 15

NETFLIX DATA ELEMENTS

ATTRIBUTES

Date

The date of the stock data entry (Temporal/

Categorical)

Open

The stock's opening price (Continuous/ Numerical)

High

The highest price of the stock during the trading day (Continuous/Numerical)

Low

The lowest price of the stock during the trading day (Continuous/Numerical)

Close

The stock's closing price

(Continuous/Numerical)

Adj Close

The stock's adjusted closing price, adjusted for splits and dividend distributions(Continuous/ Numerical).

KEY STATISTICS

Date Range: The data covers the period from January 2, 2003 to February 23, 2024.

Total Number of Days: **5,322 days**

DATA PRE-PROCESSING

- Conversion of Date: The 'Date' column was converted to a date type for accurate time series analysis.

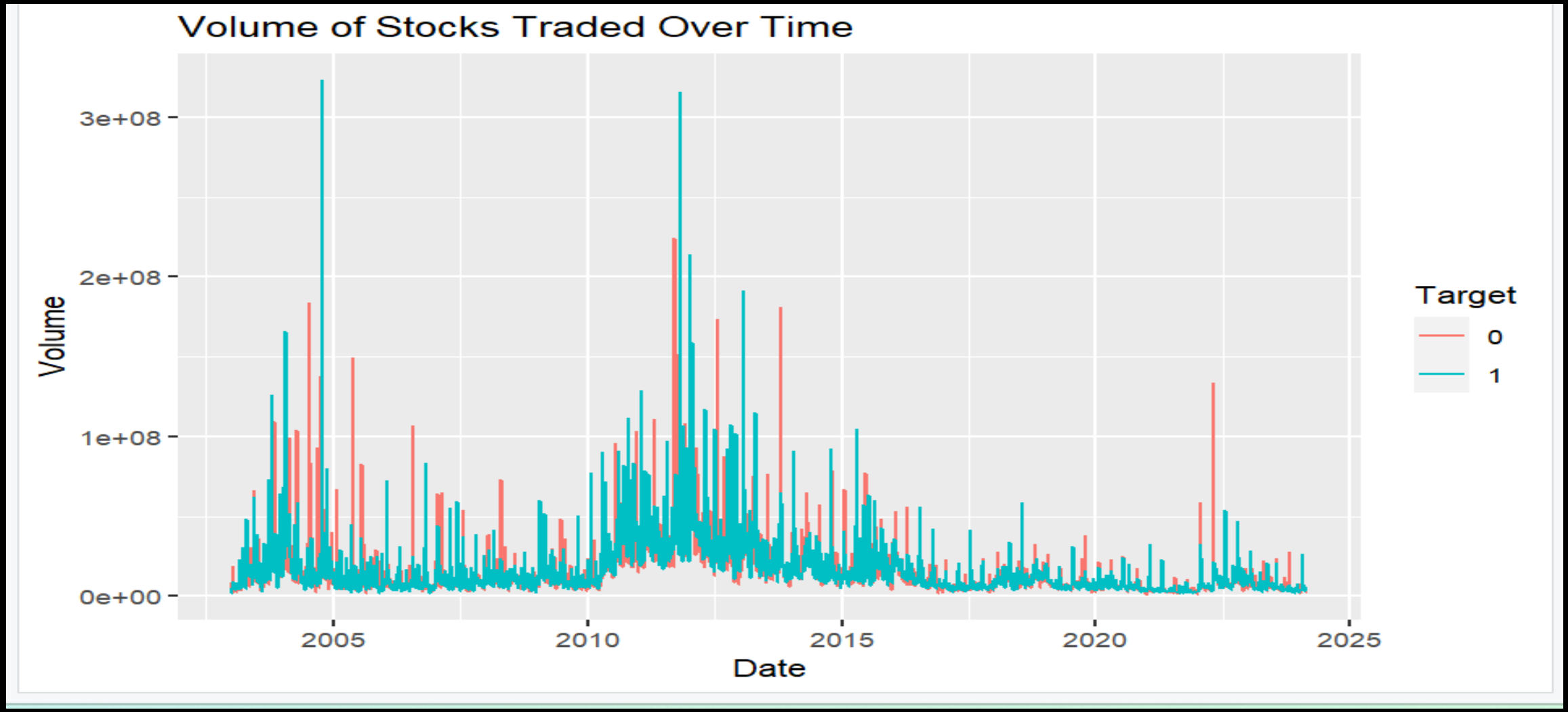
- Target Variable Creation: A binary target variable was created to indicate if the closing price was greater than the opening price (1 if Close > Open, 0 otherwise).

- Statistical Imputation Method: Imputation refers to the process of replacing missing data with substituted values. Specifically, missing values were filled using the median of the column.

Can we predict the direction of Netflix's stock closing price based on the attributes (or opening price, high, low, and volume of trading)?

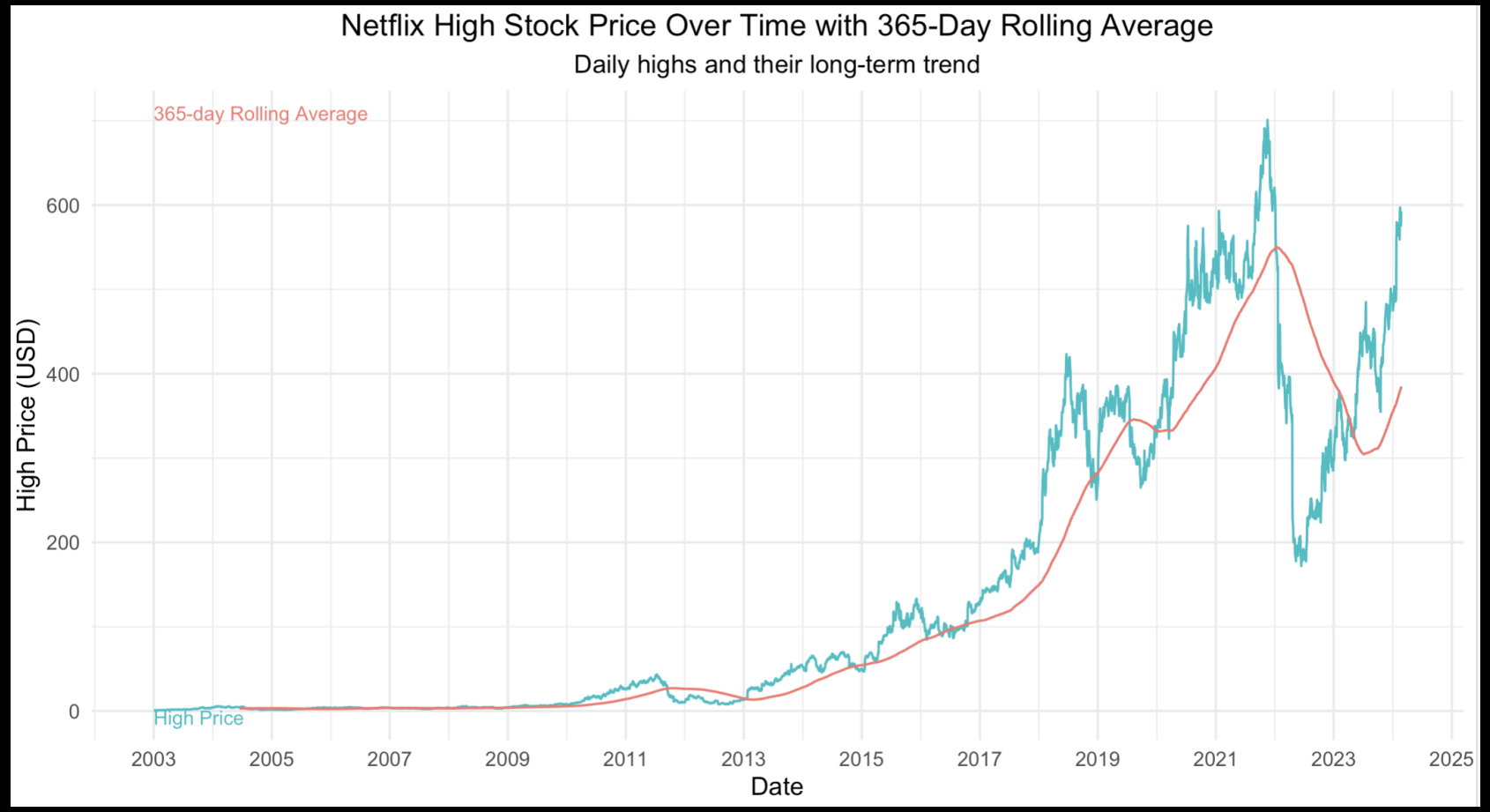
Closing Prices of Netflix's stock from 2003 to early 2024

The trend is upward which shows long-term growth in Netflix's stock value, where most growth occurred between 2013 - 2018 post when the price experienced more fluctuation



There doesn't appear to be a clear pattern where high-volume trade days consistently correlate with the stock closing higher (blue) or lower (red).

Analysis of 'High' attribute over time



The graph illustrates that Netflix's daily high prices, while volatile, follow an upward trajectory over time, as mirrored by the 365-day rolling average; this suggests a positive correlation with the closing prices, typically reflecting sustained investor confidence and market growth.

Random Forest Model

- The model above predicts which factors are the most critical markers for predicting the day's closing price:

- High and Low have highest (%Inc MSE) values which means they have the highest predicting power.

- Volume and Volume Change have lower importance indicating a lesser impact

- An RMSE of approximately 2.845 suggests that, on average, the model's predictions are within $2.85 of the actual closing price.

```
##                   %IncMSE IncNodePurity
## Prev_Close      13.513021     26654413.10
## Moving_Avg_5d   13.050397     24964656.67
## Volume_Change    7.141865        14206.31
## High            18.609626     39719557.25
## Low             18.001389     39502778.60
## Volume          13.022344      1955808.95
```

THANK YOU