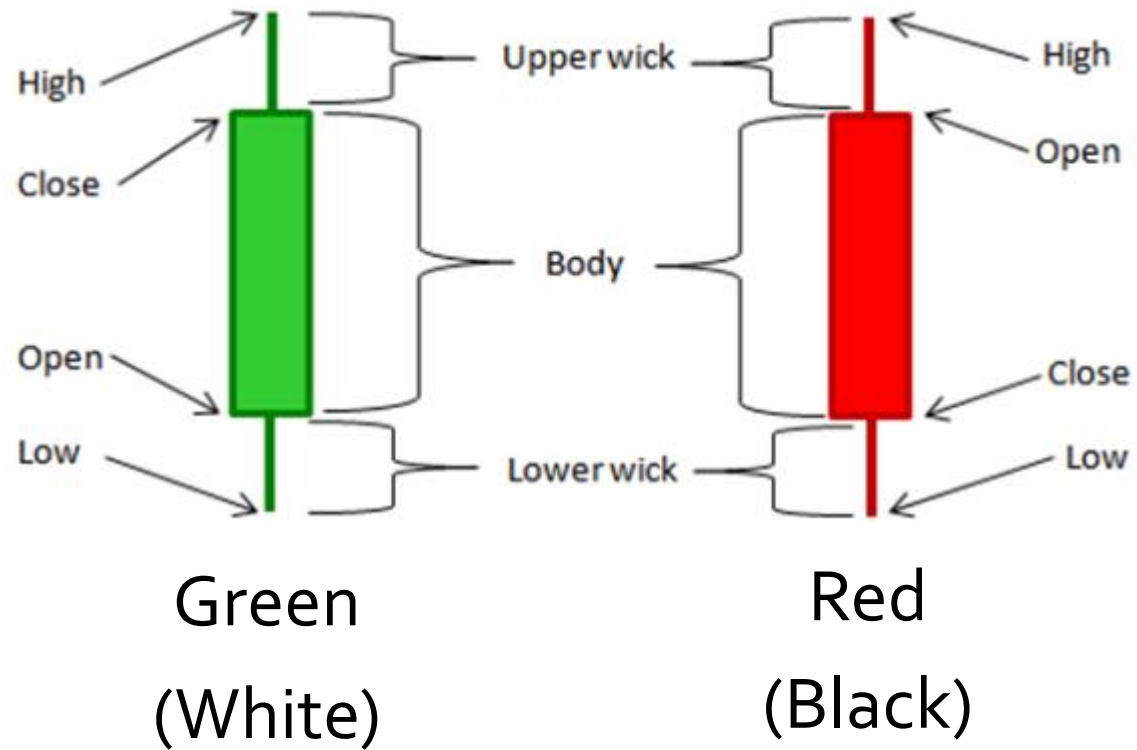


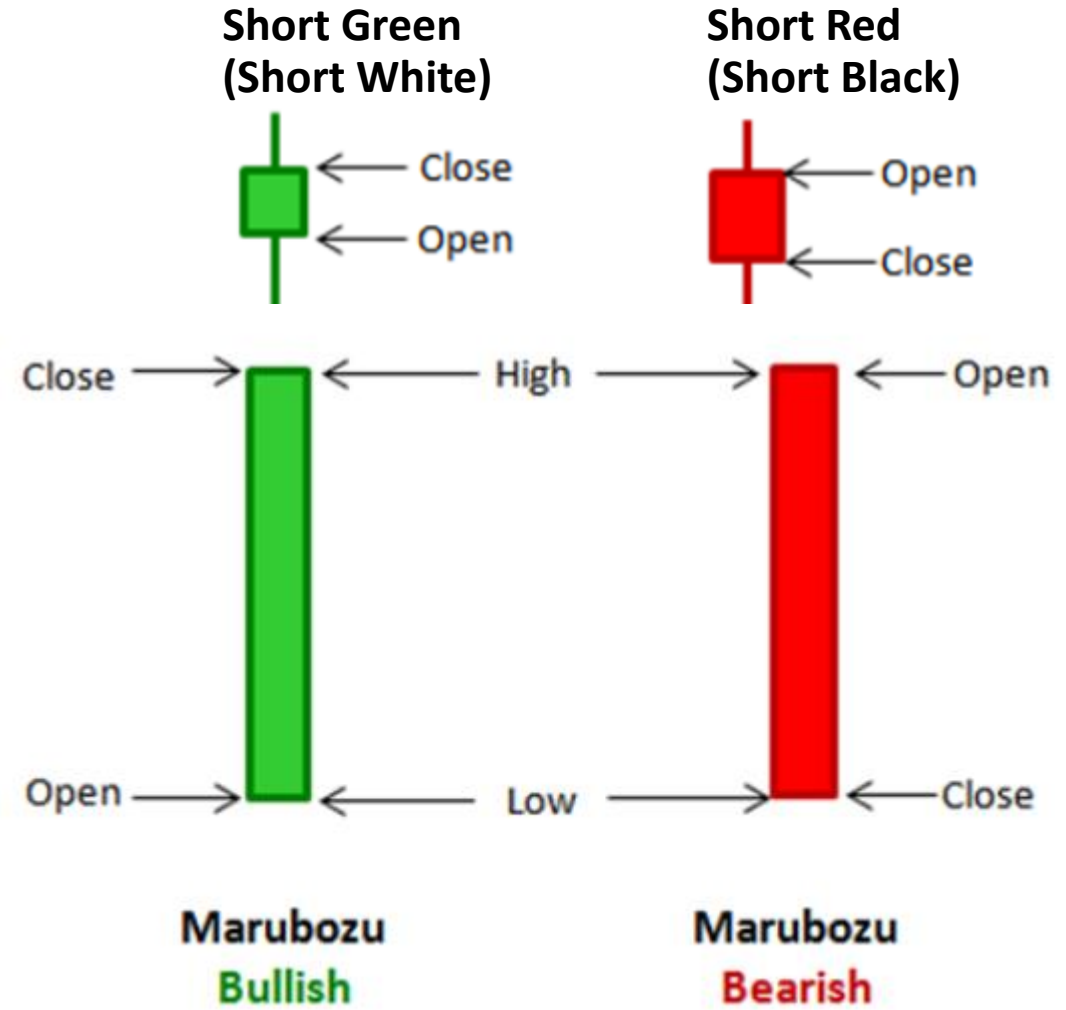
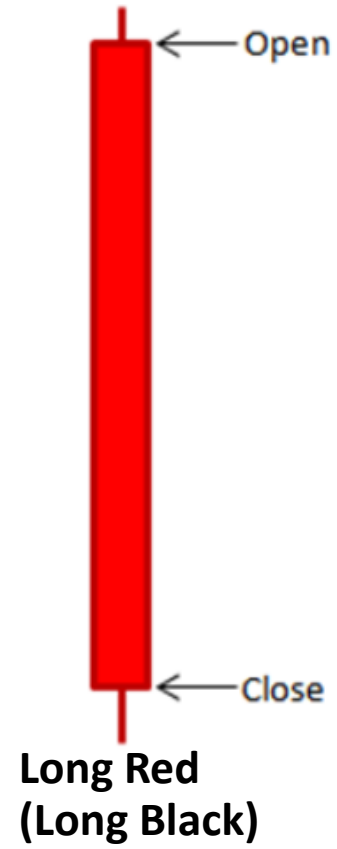
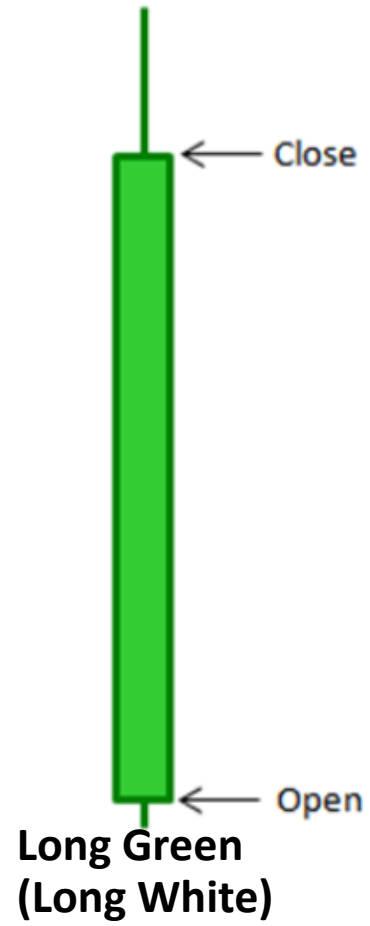
Using Candle Stick Patterns for Day-Trade

The Data Incubator - Capstone Project

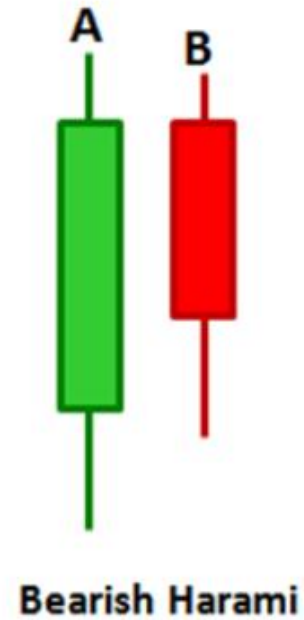
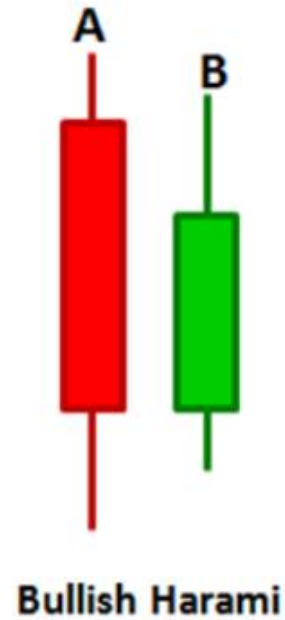
Suat Akbulut

Candle Sticks





Examples of Candle Stick Patterns



What I do?

Statistical significance
between Candlestick
patterns and Mean
Return of Day-Trade ?

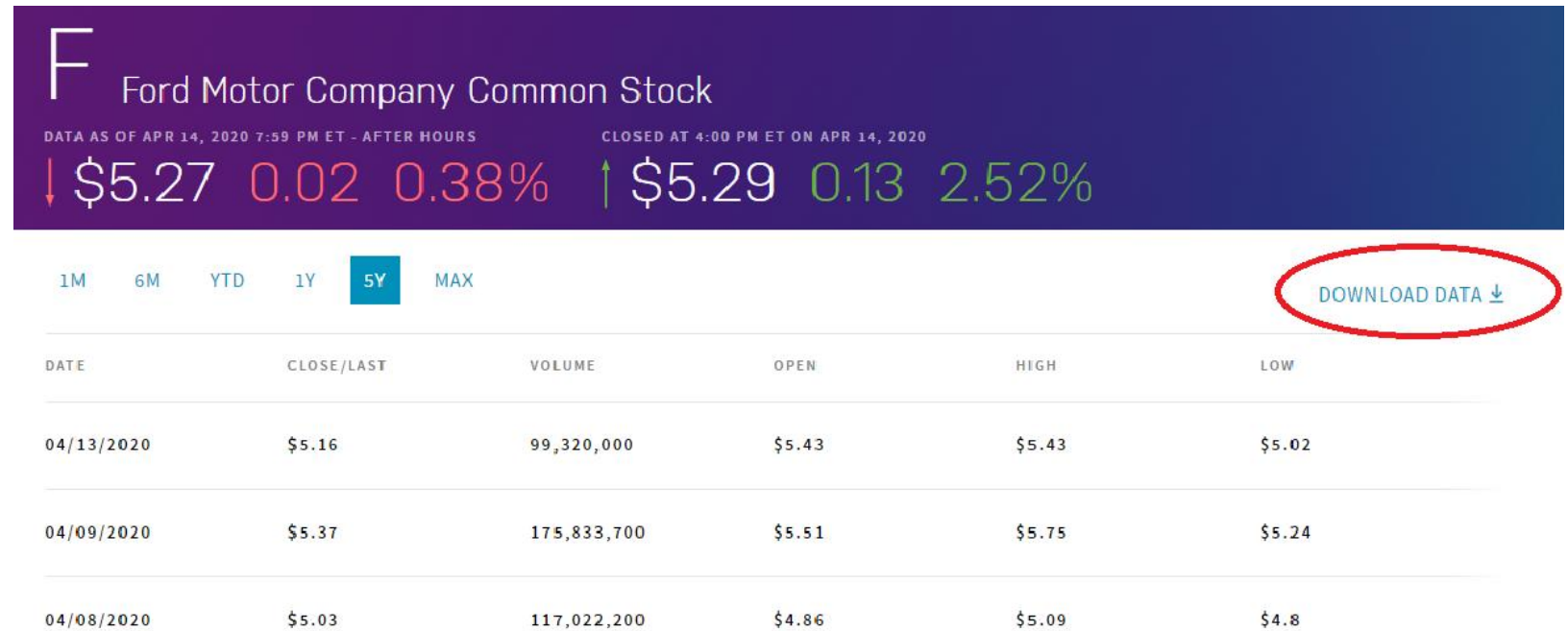
- Scrape 300,000 daily price data about the top 250 S&P500 stocks
- Create candlesticks, day-trade return
- Predict whether a day-trade is likely to be profitable with acceptable **precision**.

<u>Model</u>	<u>Accuracy</u>	<u>Presicion</u>	<u>Recall</u>
Logit (Non-Opt)	0.521	0.528	0.756
Random Forest (Non-Opt)	0.513	0.526	0.764
NN	0.516	0.530	0.674

Part 1: Scrape Data

www.nasdaq.com,
Top 250 S&P500 Stocks Last 5 year's daily
Open, High, Low, Close, Volume

Used requests
Saved as .csv files
ticks and Label the Patterns



The screenshot shows the stock page for Ford Motor Company Common Stock on NASDAQ.com. The header is dark blue with white text. Below the header, there are two rows of data: the current price and change, and the previous close and change. The current price is \$5.27, down 0.02 (0.38%) from the previous close of \$5.29 (up 0.13, 2.52%). Below this, there are tabs for different time periods: 1M, 6M, YTD, 1Y, 5Y (selected), and MAX. To the right of these tabs is a red circle around the 'DOWNLOAD DATA' link. Below the tabs is a table with columns: DATE, CLOSE/LAST, VOLUME, OPEN, HIGH, and LOW. The table contains three rows of data for the dates 04/13/2020, 04/09/2020, and 04/08/2020.

<div>F Ford Motor Company Common Stock</div> <div>DATA AS OF APR 14, 2020 7:59 PM ET - AFTER HOURS CLOSED AT 4:00 PM ET ON APR 14, 2020</div> <div>↓ \$5.27 0.02 0.38% ↑ \$5.29 0.13 2.52%</div> <div>1M 6M YTD 1Y 5Y MAX</div> <div>DOWNLOAD DATA ↓</div> <table><tr><th>DATE</th><th>CLOSE/LAST</th><th>VOLUME</th><th>OPEN</th><th>HIGH</th><th>LOW</th></tr><tr><td>04/13/2020</td><td>\$5.16</td><td>99,320,000</td><td>\$5.43</td><td>\$5.43</td><td>\$5.02</td></tr><tr><td>04/09/2020</td><td>\$5.37</td><td>175,833,700</td><td>\$5.51</td><td>\$5.75</td><td>\$5.24</td></tr><tr><td>04/08/2020</td><td>\$5.03</td><td>117,022,200</td><td>\$4.86</td><td>\$5.09</td><td>\$4.8</td></tr></table>						DATE	CLOSE/LAST	VOLUME	OPEN	HIGH	LOW	04/13/2020	\$5.16	99,320,000	\$5.43	\$5.43	\$5.02	04/09/2020	\$5.37	175,833,700	\$5.51	\$5.75	\$5.24	04/08/2020	\$5.03	117,022,200	\$4.86	\$5.09	\$4.8
DATE	CLOSE/LAST	VOLUME	OPEN	HIGH	LOW																								
04/13/2020	\$5.16	99,320,000	\$5.43	\$5.43	\$5.02																								
04/09/2020	\$5.37	175,833,700	\$5.51	\$5.75	\$5.24																								
04/08/2020	\$5.03	117,022,200	\$4.86	\$5.09	\$4.8																								
04/13/2020	\$5.16	99,320,000	\$5.43	\$5.43	\$5.02																								
04/09/2020	\$5.37	175,833,700	\$5.51	\$5.75	\$5.24																								
04/08/2020	\$5.03	117,022,200	\$4.86	\$5.09	\$4.8																								

Part 2: Create Candlesticks and Label the Patterns

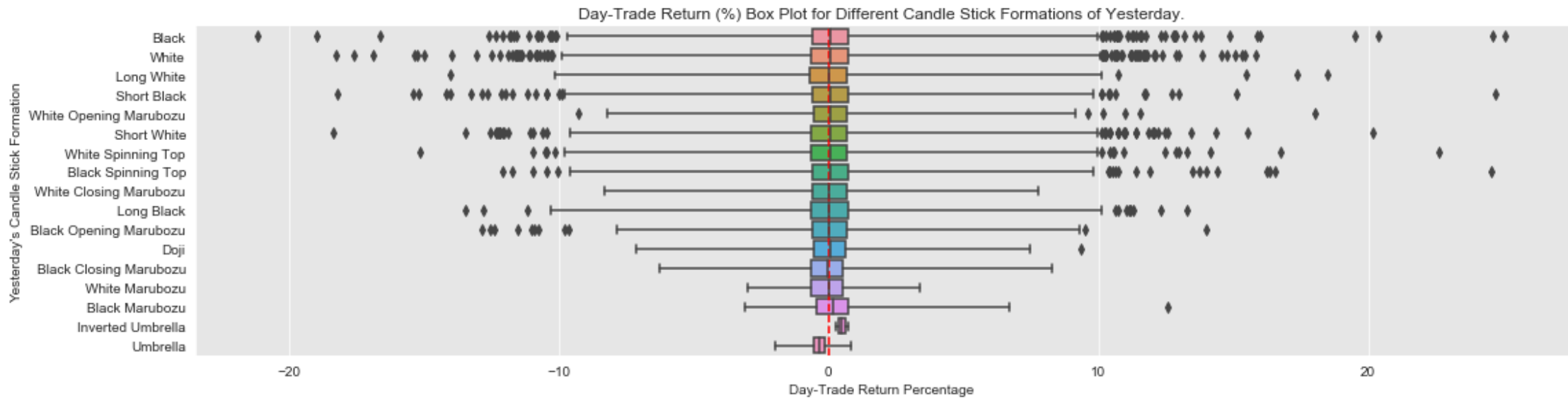
First 18 different candle sticks for the last 4 days

Volumes for the last 4 days

Day Trade Profit (%)

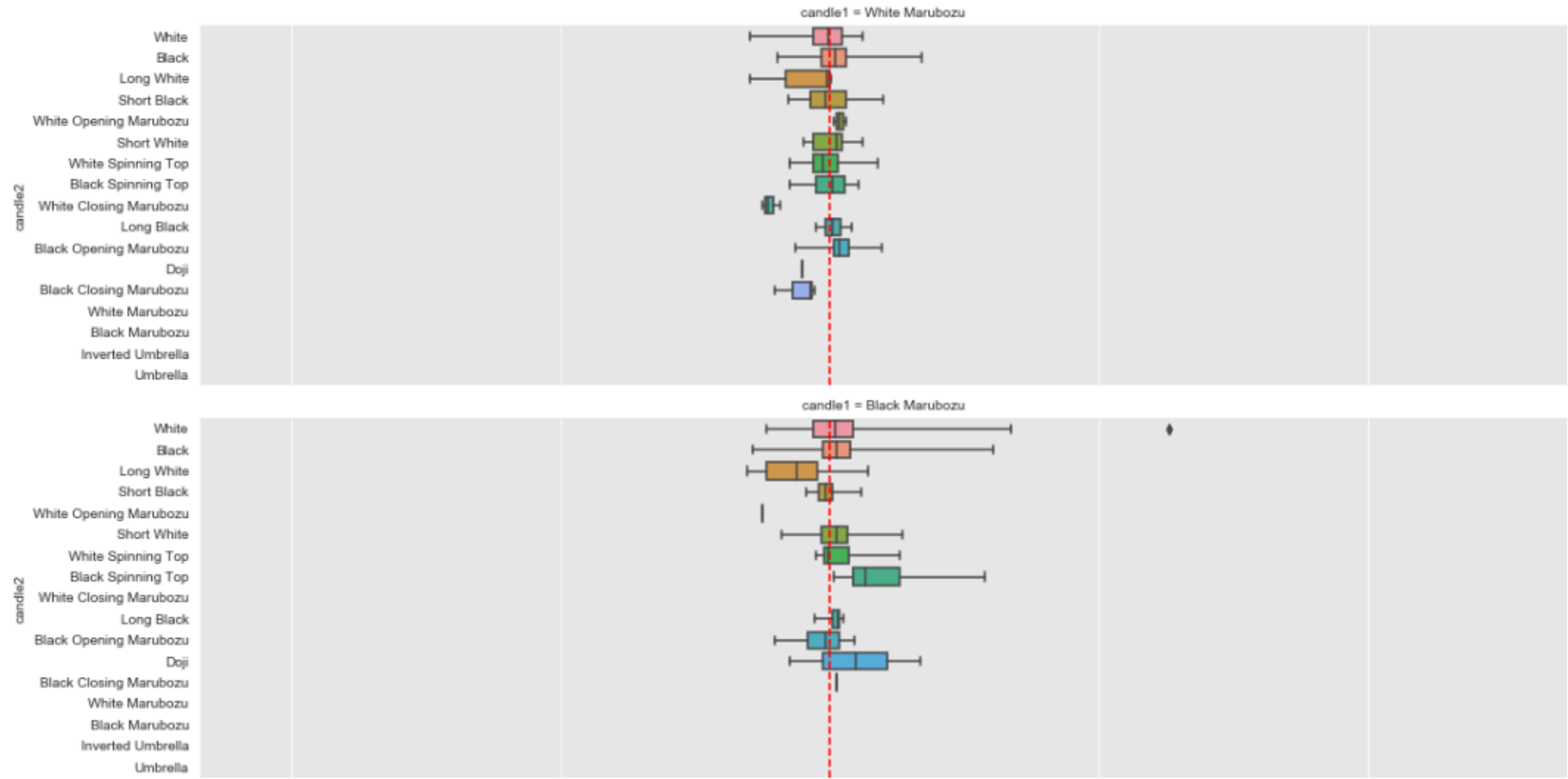
Day Trade Profit (Binary)

Labelling the patterns (NEXT)



Part 3: Exploratory Data Analysis

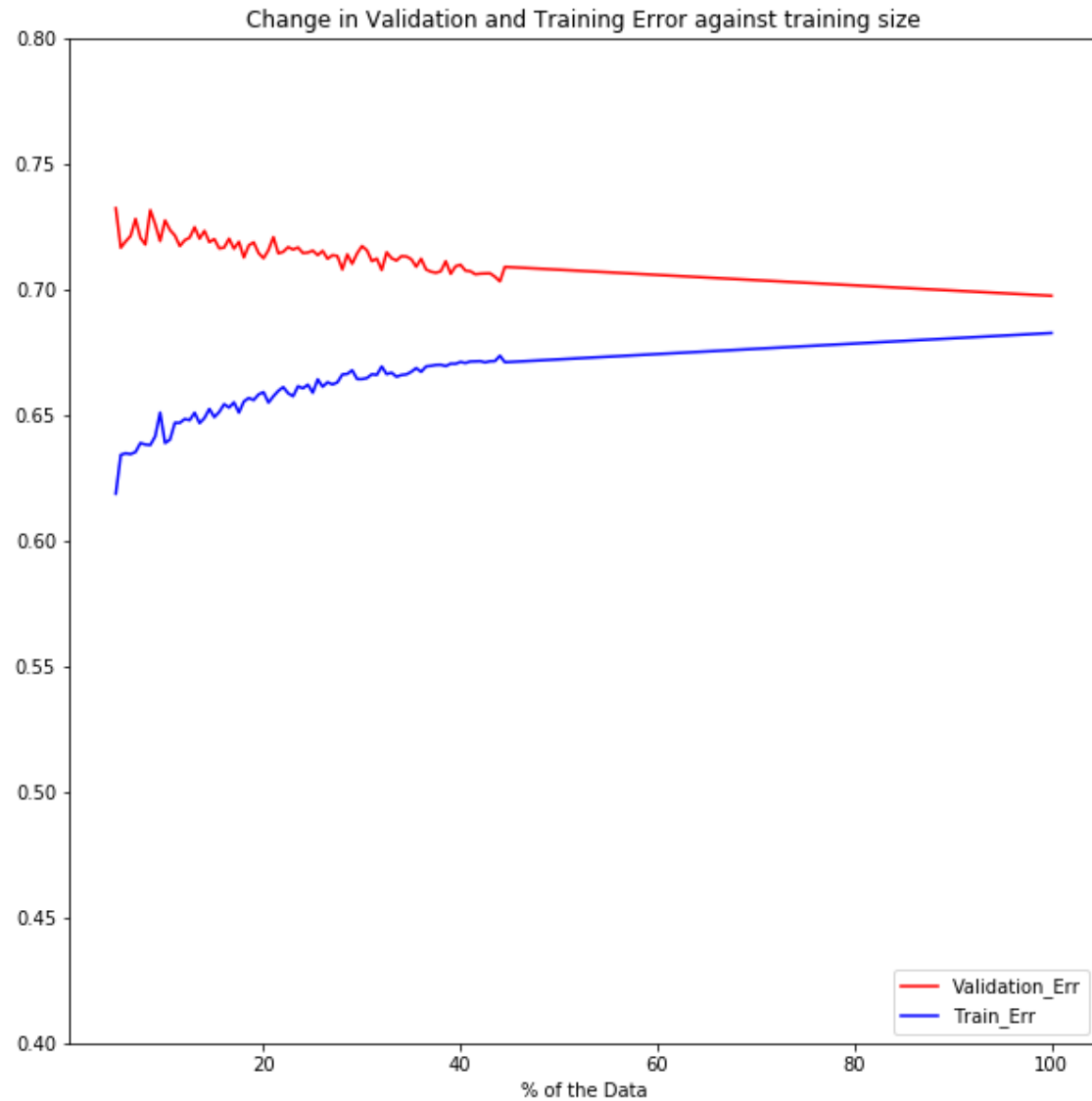
Last Two Day's Candle Stick Formation vs Percentage Return



Part 4: Preprocess and ML:

- Logistic Regression, Random Forest Classifier, and Neural Network
- Regressed IsProfit on Last 4 days' candles and volumes: Accuracy: Precision: Recall:
- Regressed IsProfit on Last 4 days' candles and volumes: Accuracy: Precision: Recall:

Part 5: More data or More Complication?



Next:

More
Complexity

- First 18 different candle sticks for the last 4 days
- Volumes for the last 4 days (std)
- Day Trade Profit (%)
- Day Trade Profit (Binary)
- (NEXT) Labelling the patterns, e.g. Bearish Harami, Bullish Engulfing, Bearish Evening Star, etc.
- Maybe candlestick pattern is more effective for smaller capitalized stocks