

Training Data Visualization

Kevin Lu and Jevin Maltais

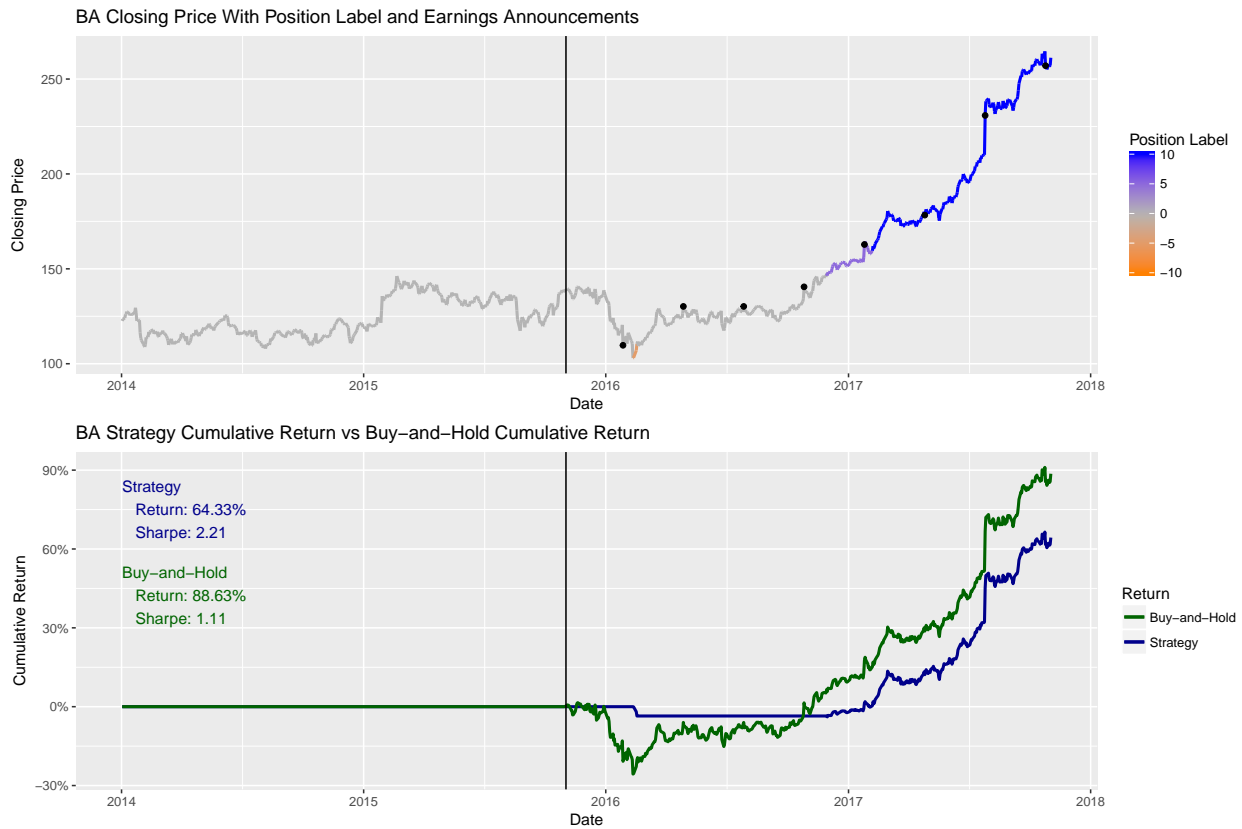
January 1, 2018

1. Introduction

This document visualizes and investigates the relationship between the position labels and the training data. The momentum-based strategy derived from the position labels is also compared to a buy-and-hold strategy. The investigation reveals that the strategy performs well on risk-adjusted terms and experiences extremely low drawdowns. The potential need for additional training examples for negative position labels is highlighted.

2. Interpreting Plots

An example of the plots contained within this document is included below and this section provides a brief description of how to interpret the plot.



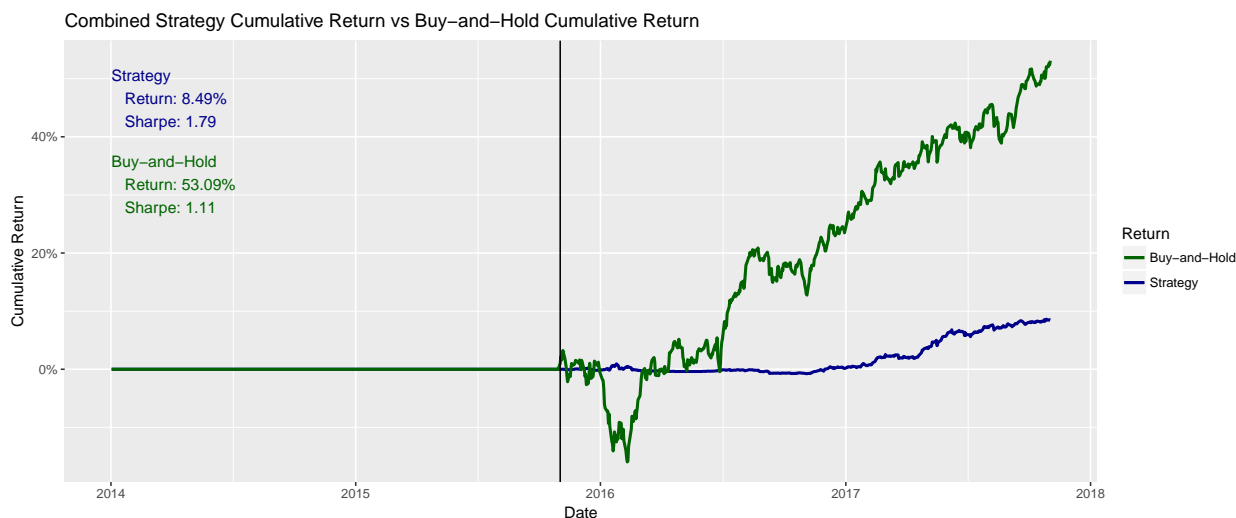
The plot consists of two subplots. The top subplot contains a visualization of the trading strategy over time. The subplot contains the adjusted closing price data starting approximately two years before the training period and continues to the end of the training period on November 2, 2017. The vertical black line in the middle of the plot indicates the beginning of the training period on November 2, 2015. Earnings announcement dates are indicated by small black points. The position label is displayed as a color gradient overlayed on top of the adjusted closing price – bright blue indicates a position label of +10, gray indicates a

position label of 0, and bright red indicates a position label of -10. Intermediate colors indicate intermediate position labels.

The bottom subplot contains a visualization of the performance of the trading strategy versus the buy-and-hold return of the underlying stock. The performance of the strategy is indicated by the dark blue line and the performance of the buy-and-hold strategy is indicated by the dark green line. The overall cumulative return of both strategies as well as the sharpe ratio are annotated in the upper left corner. The [sharpe ratio](#) is a measure of risk-adjusted returns with higher values associated with higher risk-adjusted performance.

3. Evaluation of Overall Strategy

The overall strategy is evaluated by assuming that the strategy would have an equal capital allocation to the 26 stocks contained in the training data. Under this hypothetical application of the strategy, the capital allocation to each stock is fully long when the position label is +10, half long when the position label is +5, half short when the position label is -5, and fully short when the position label is -10. The allocated capital is not invested when the position label is 0. While this does not reflect the actual application of the strategy which deploys capital to the most profitable opportunities at all times, it is a useful perspective to evaluate the strategy.



While the overall strategy constructed using equal capital allocation to each stock does not outperform the buy-and-hold return, the strategy sharpe ratio is significantly higher than the buy-and-hold sharpe ratio. Furthermore, the strategy experiences very shallow drawdowns and generally follows a steady upward trend which is very desirable from a risk management perspective. With the smart allocation of capital to the most profitable opportunities and perhaps with the careful application of leverage, the magnitude of the overall strategy return can outperform the buy-and-hold return. I believe the position labels represent a sound strategy that we can build a model on top of.

After reviewing the plots of individual stocks displayed in the following sections, a few patterns have emerged. The stocks contained within the training data can be grouped into distinct clusters based on the historical patterns of their stock prices, how much momentum-based behavior is exhibited, the volatility of the stock prices, proximity to earnings announcement dates, whether the stock follows patterns based on the Elliot wave principle, as well as the ability of the strategy to detect these patterns.

A careful review was conducted to ensure that the training data set has a diverse range of patterns so that a machine learning model trained on the data can generate good predictions across a wide range of stocks and market environments. The ideal training set contains sufficient training examples within each type or cluster of stock patterns and the distribution of position labels that is similar to the distribution in the real world.

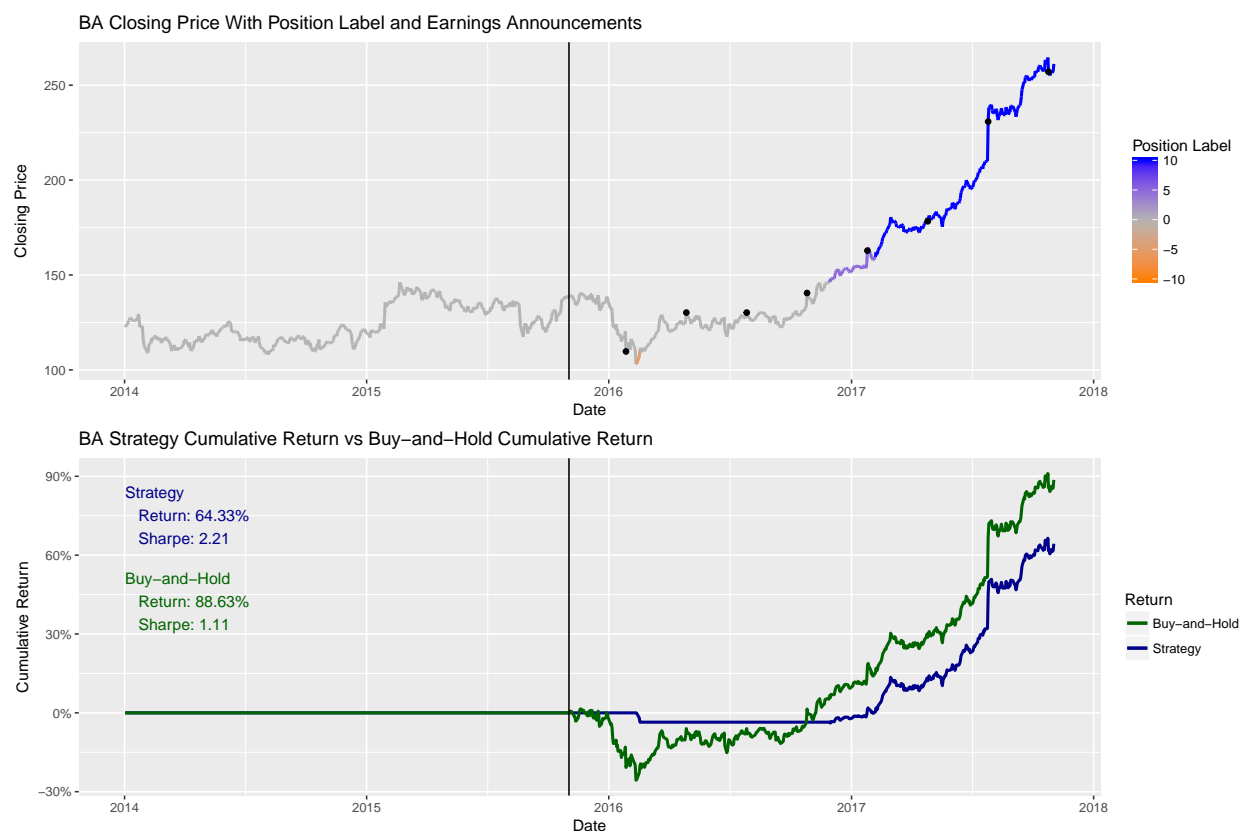
4. Detecting Desirable Stock Patterns

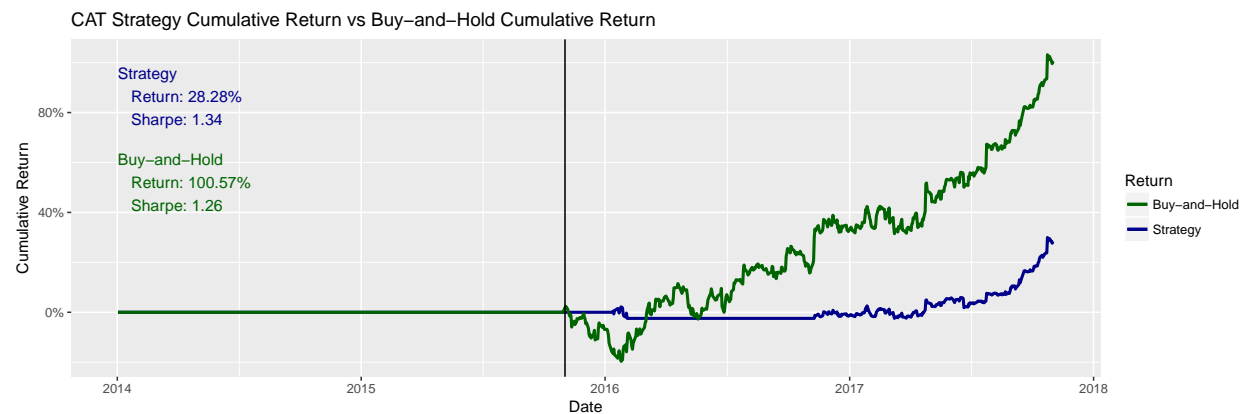
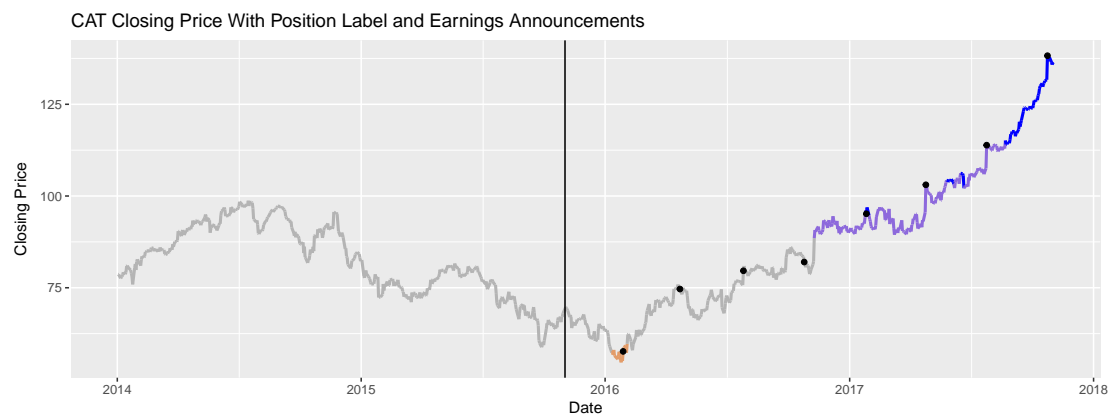
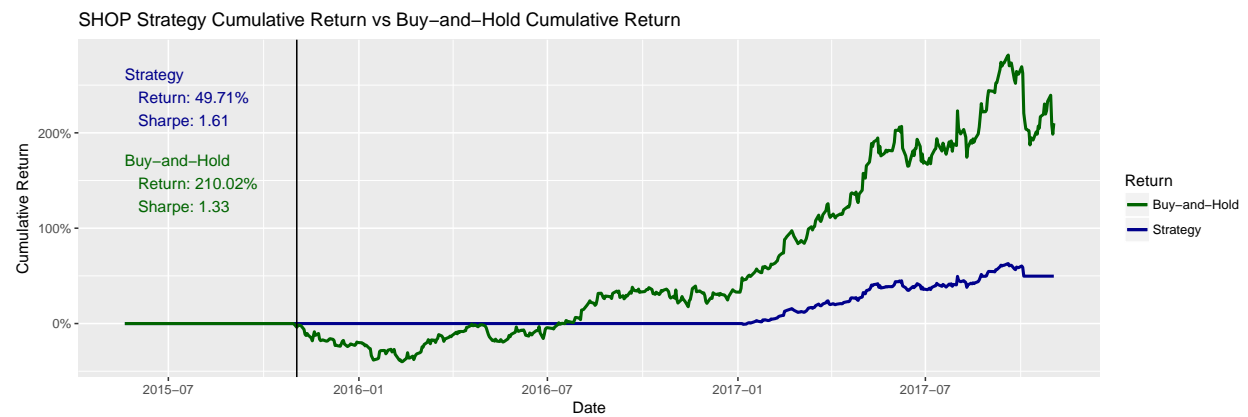
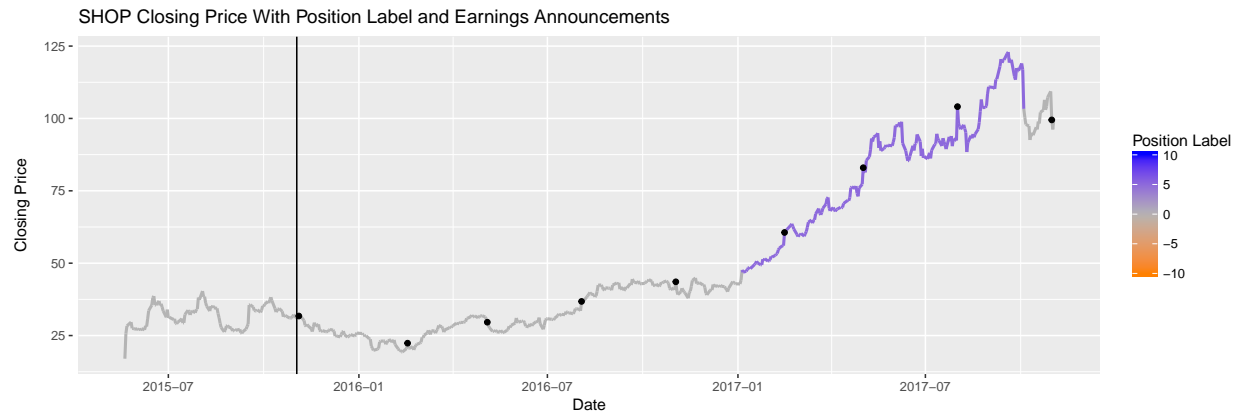
I identify two clusters that exhibit desirable stock patterns. Each stock within these two clusters contain a large number of training examples with a positive position label.

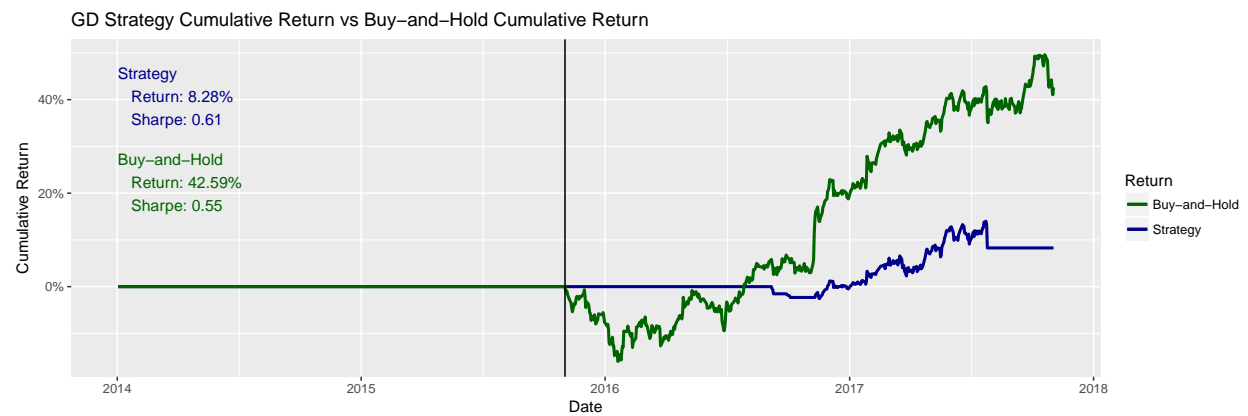
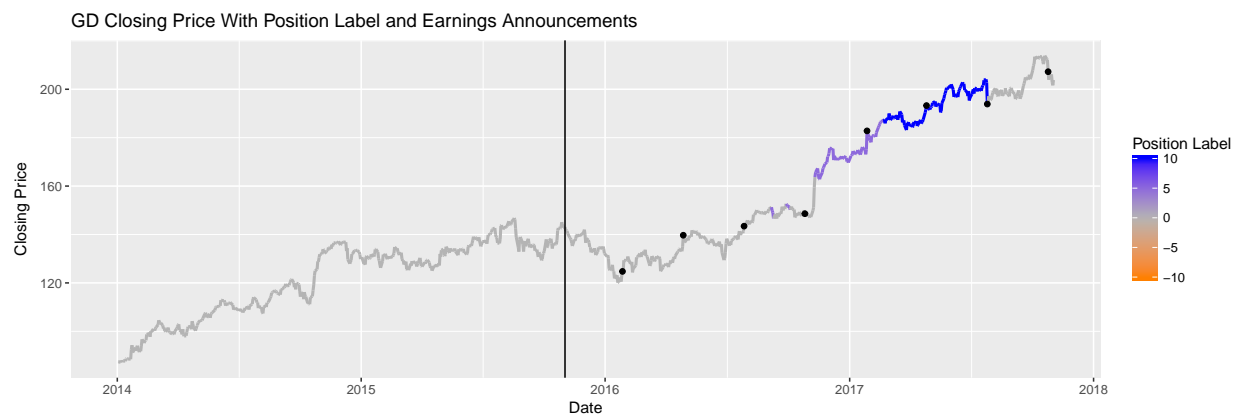
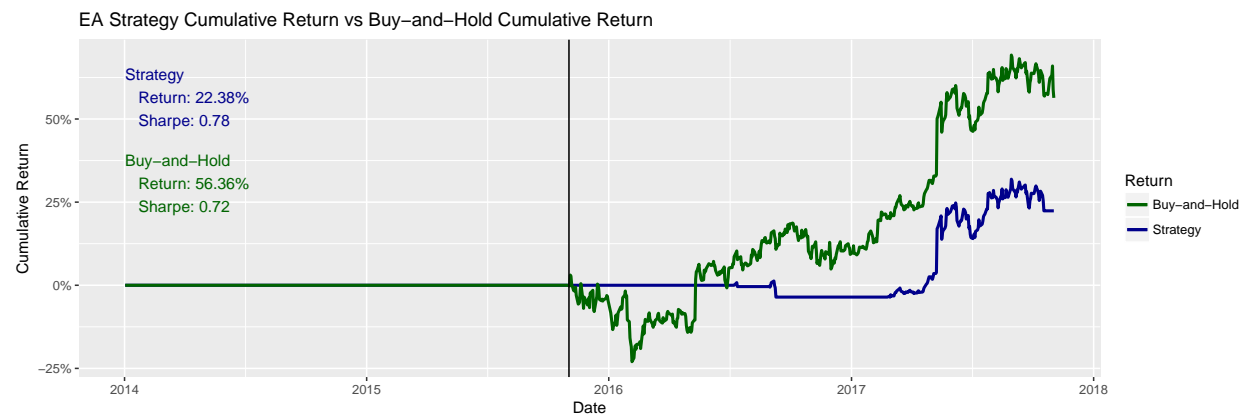
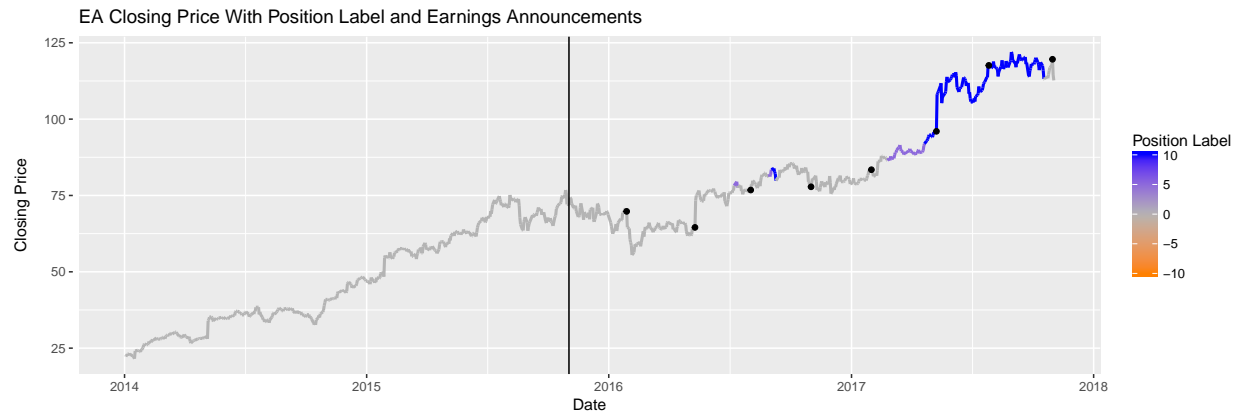
4.1 Strong Upward Trend With Low Volatility

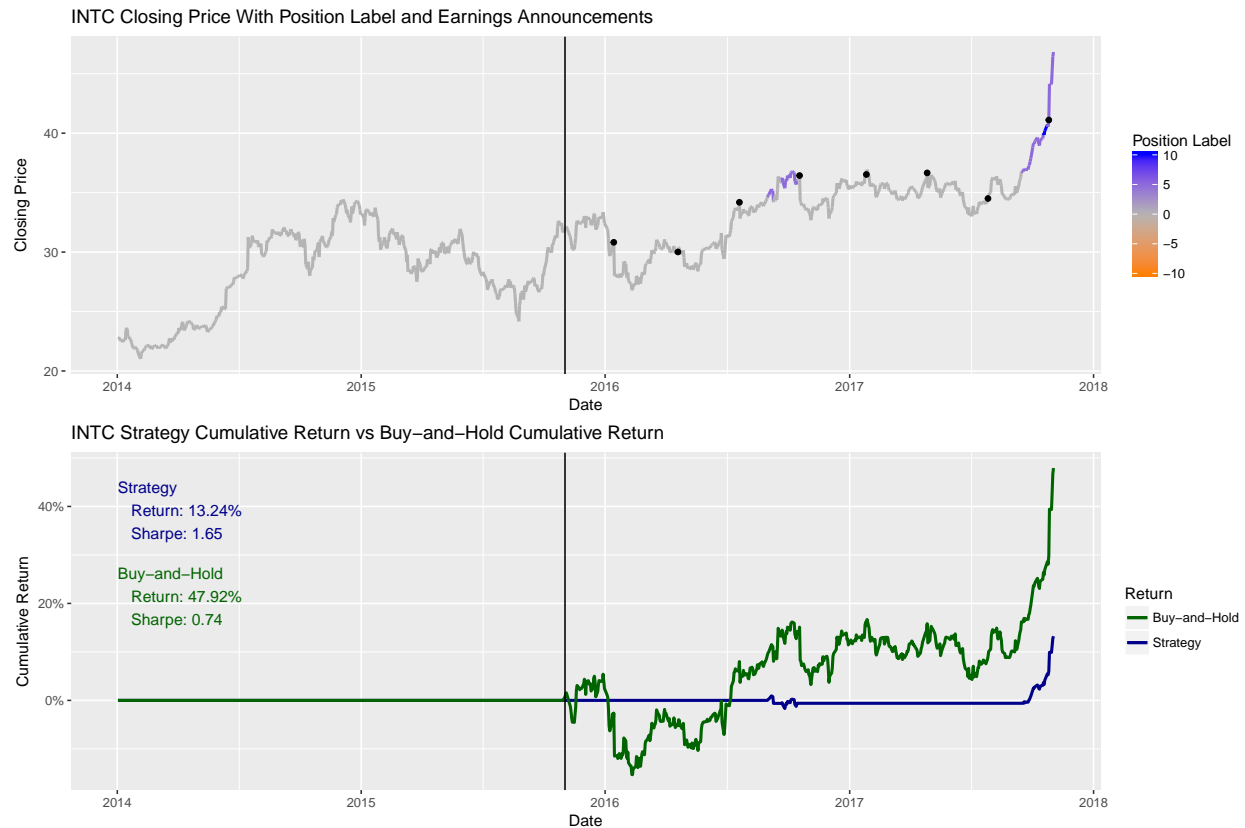
Stocks with a strong upward trend and with low volatility exhibit highly desirable stock patterns for a momentum-based trading strategy. Visually, each of these stocks follow a very similar pattern. The stock price first exhibits a long and steady consolidation period that trades in a narrow range for a period of several months to several years. Then, a breakout occurs on the upside with steady gains, low volatility, and without large single day changes. The positive trend persists for several months. The strategy as represented by the position labels is able to detect this pattern early on and each of these stocks contain long and uninterrupted periods where the position label is +5 or +10.

These patterns are highly profitable for the strategy and show excellent risk-adjusted return. These training examples in this cluster will allow the machine learning model would be to detect these patterns and there appear to be sufficient training examples for the model to learn from. The ultimate goal of the model would be to identify stocks that are exhibiting patterns similar to stocks within this cluster. Examples of stocks in this cluster include BA, SHOP, CAT, EA, GD, and INTC.







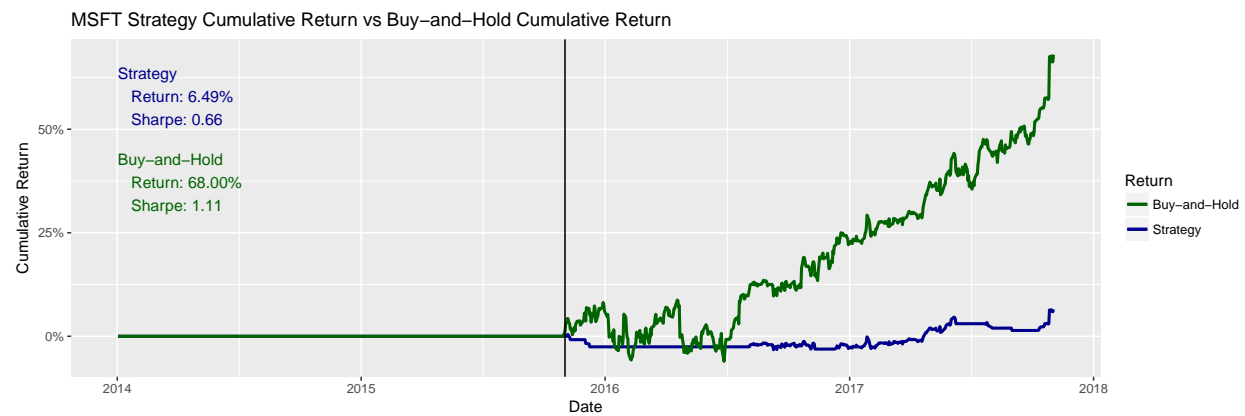
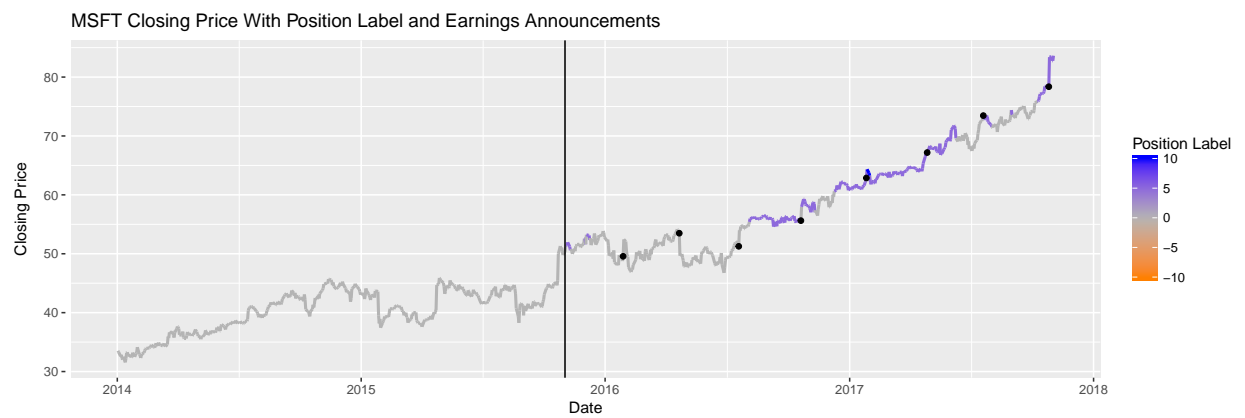
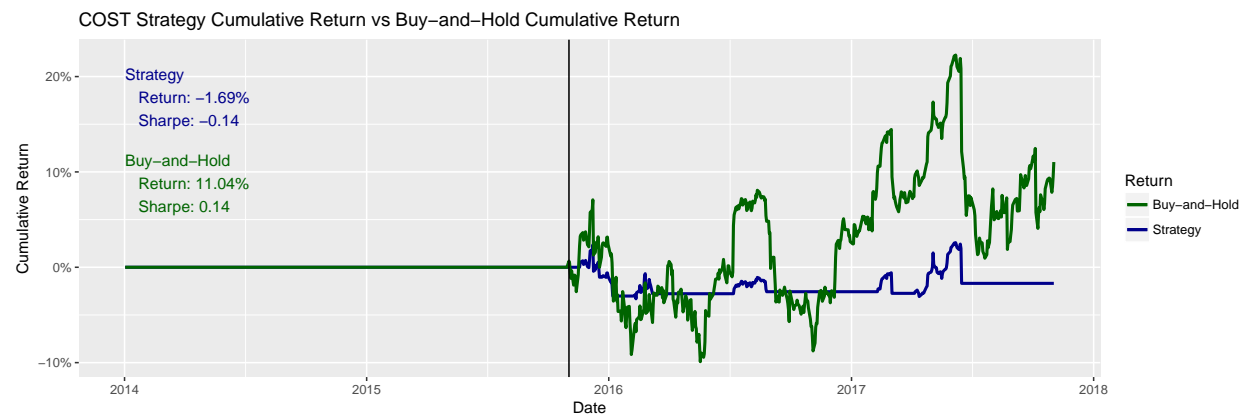
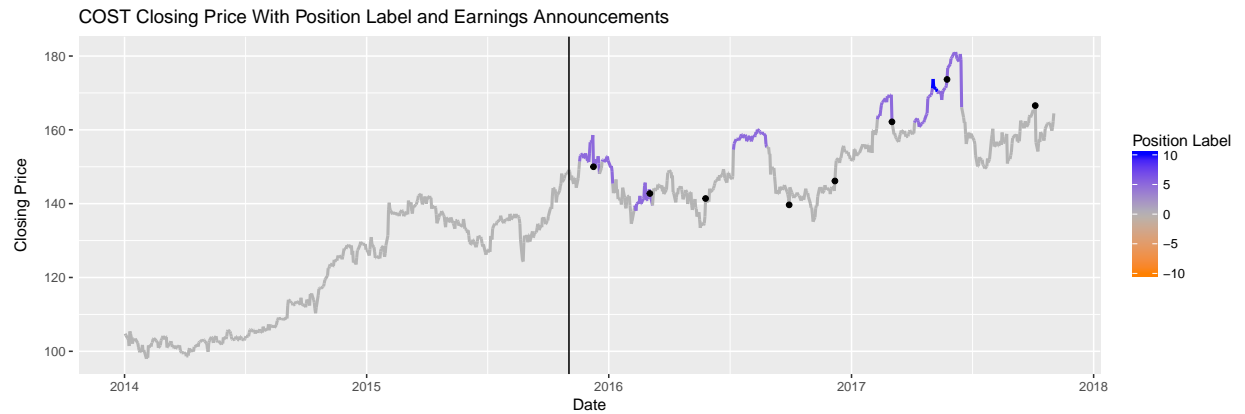


4.2 Upward Trend With High Volatility

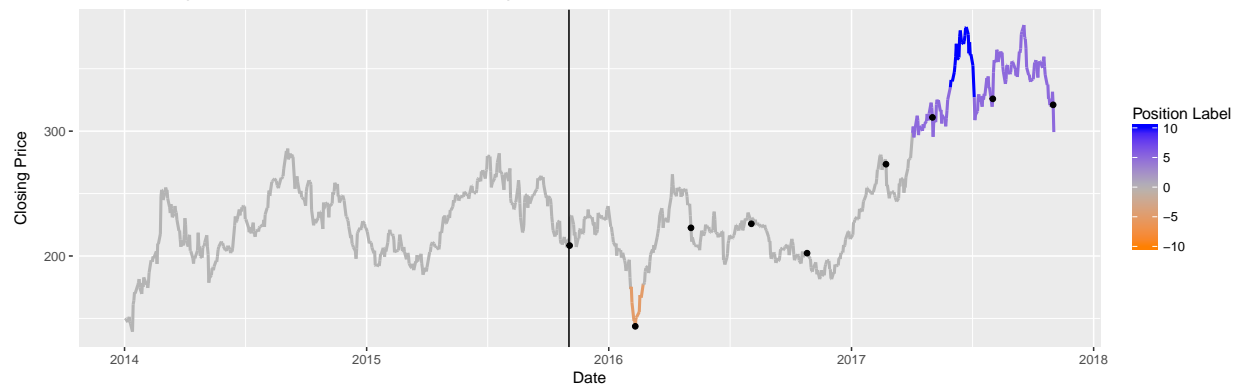
The stocks within this cluster are similar to the stocks discussed in section 4.1. Often times, the entry points show patterns that are remarkably similar to the stocks previously examined in that a breakout occurs on the upside following a period of time in which the stock trades within a narrow range. The stocks in this cluster differ in that the upward momentum is either too volatile in nature or not consistently sustained. The stock prices may also violate the patterns described by the Elliot wave principle. While the overall trend is up, it is far less well-defined compared to the stocks previously discussed.

While the strategy applied to many of the stocks in this cluster were not profitable, these stocks still contain valuable training examples that show attractive entry points. Experience and historical examples tell us that this kind of trade setup is profitable on average even if there are single examples that are not profitable. These stocks also contain valuable training examples that can teach the model strong risk management techniques and to exit unprofitable trades early.

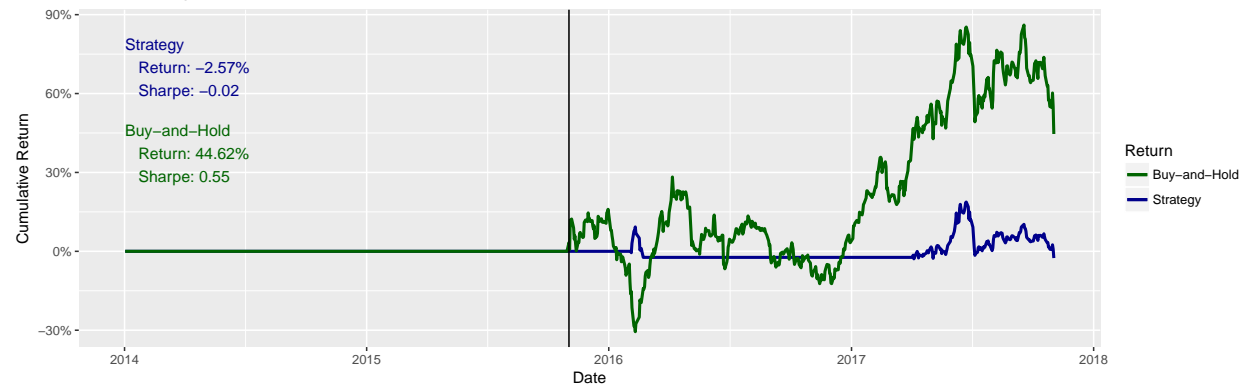
Examples of stocks in this cluster include COST, MSFT, TSLA, BAC, GDDY, and TRUE.



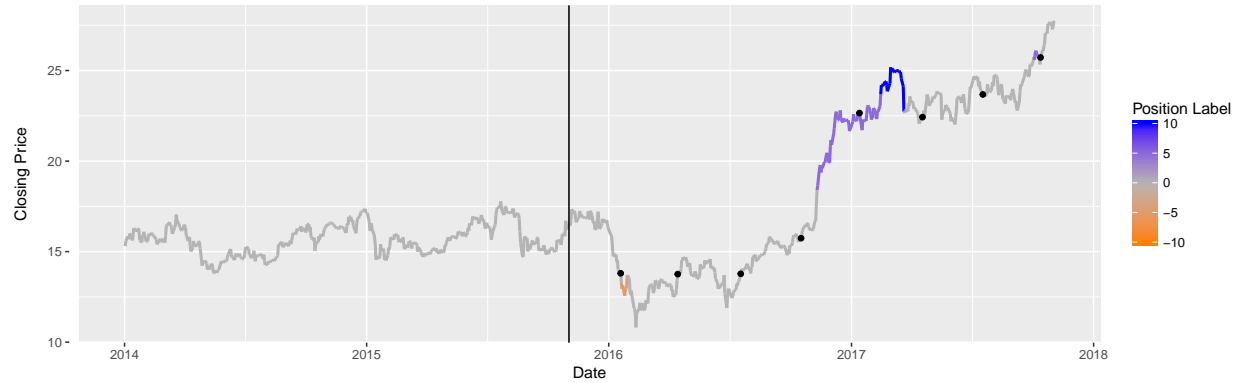
TSLA Closing Price With Position Label and Earnings Announcements



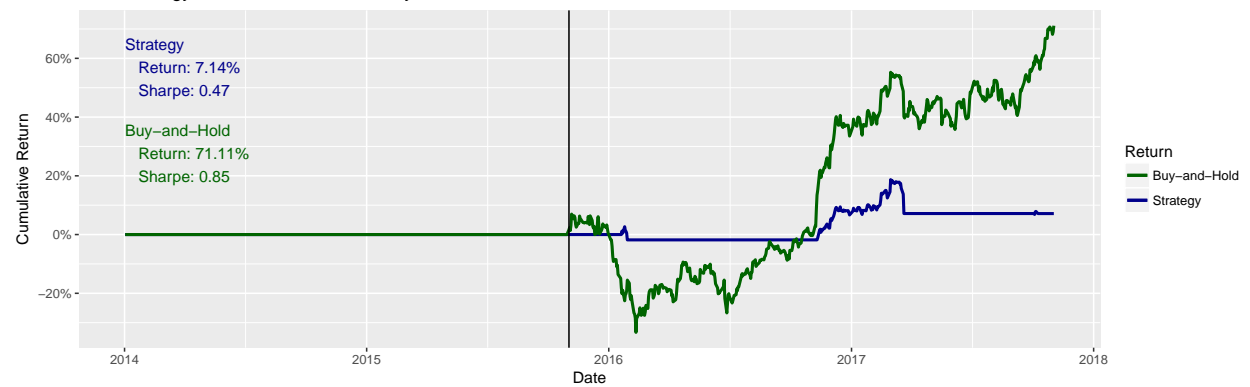
TSLA Strategy Cumulative Return vs Buy-and-Hold Cumulative Return



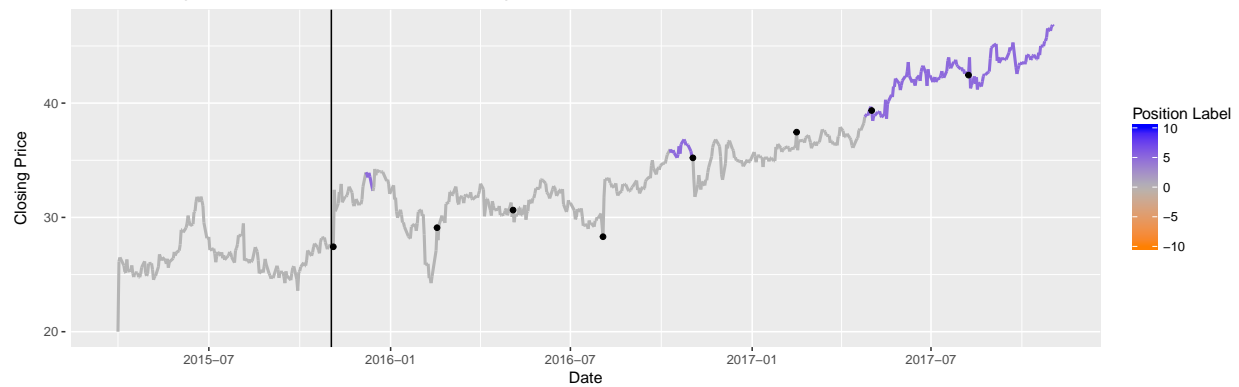
BAC Closing Price With Position Label and Earnings Announcements



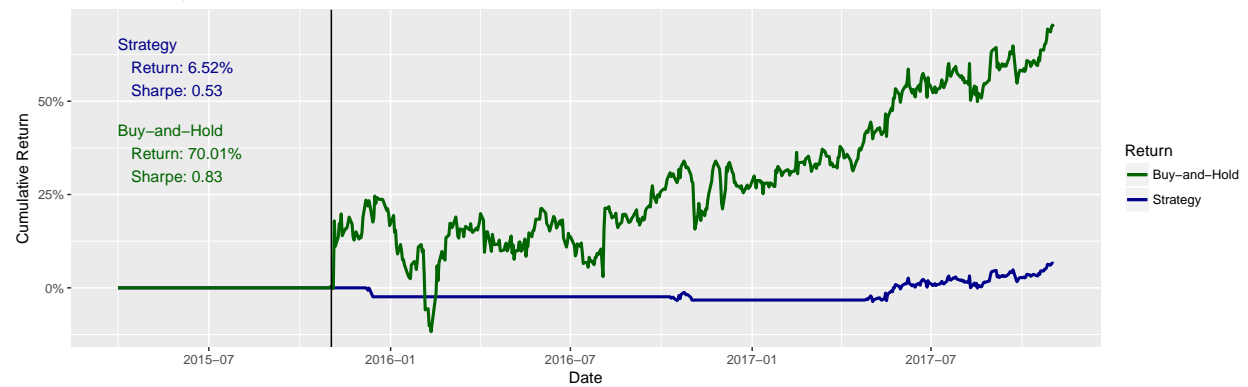
BAC Strategy Cumulative Return vs Buy-and-Hold Cumulative Return



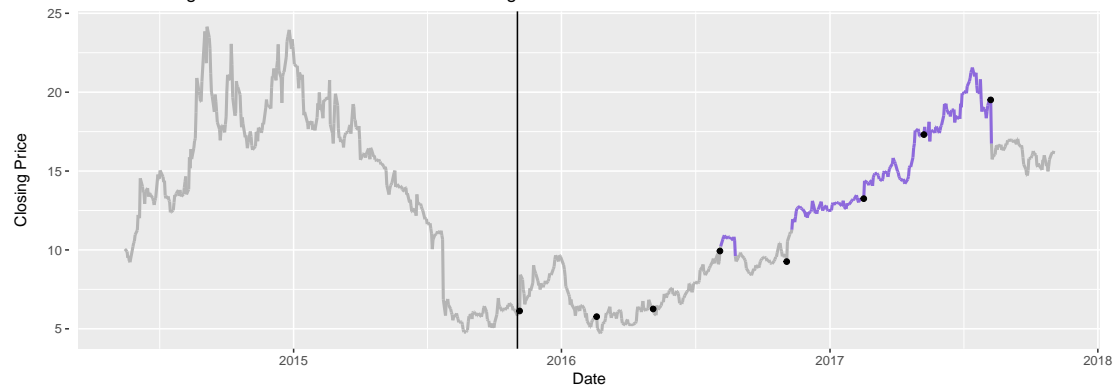
GDDY Closing Price With Position Label and Earnings Announcements



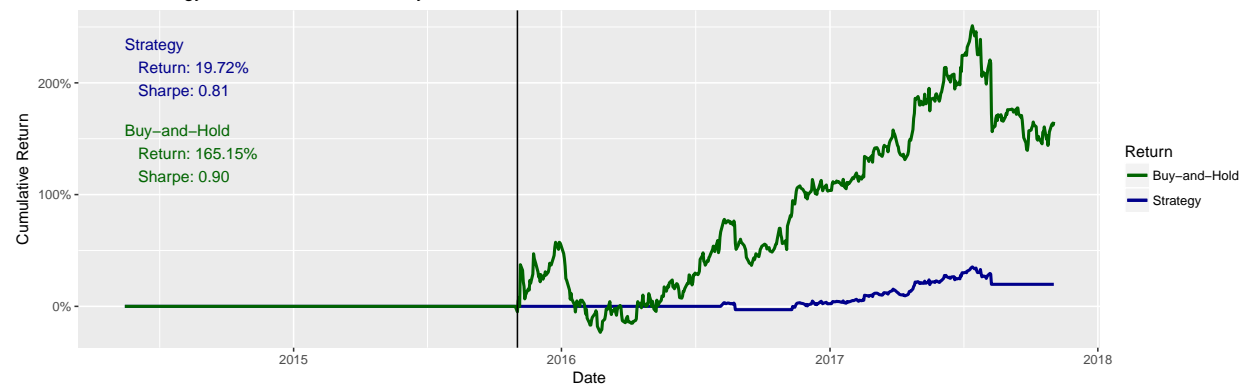
GDDY Strategy Cumulative Return vs Buy-and-Hold Cumulative Return



TRUE Closing Price With Position Label and Earnings Announcements



TRUE Strategy Cumulative Return vs Buy-and-Hold Cumulative Return



5. Detecting Undesireable Stock Patterns

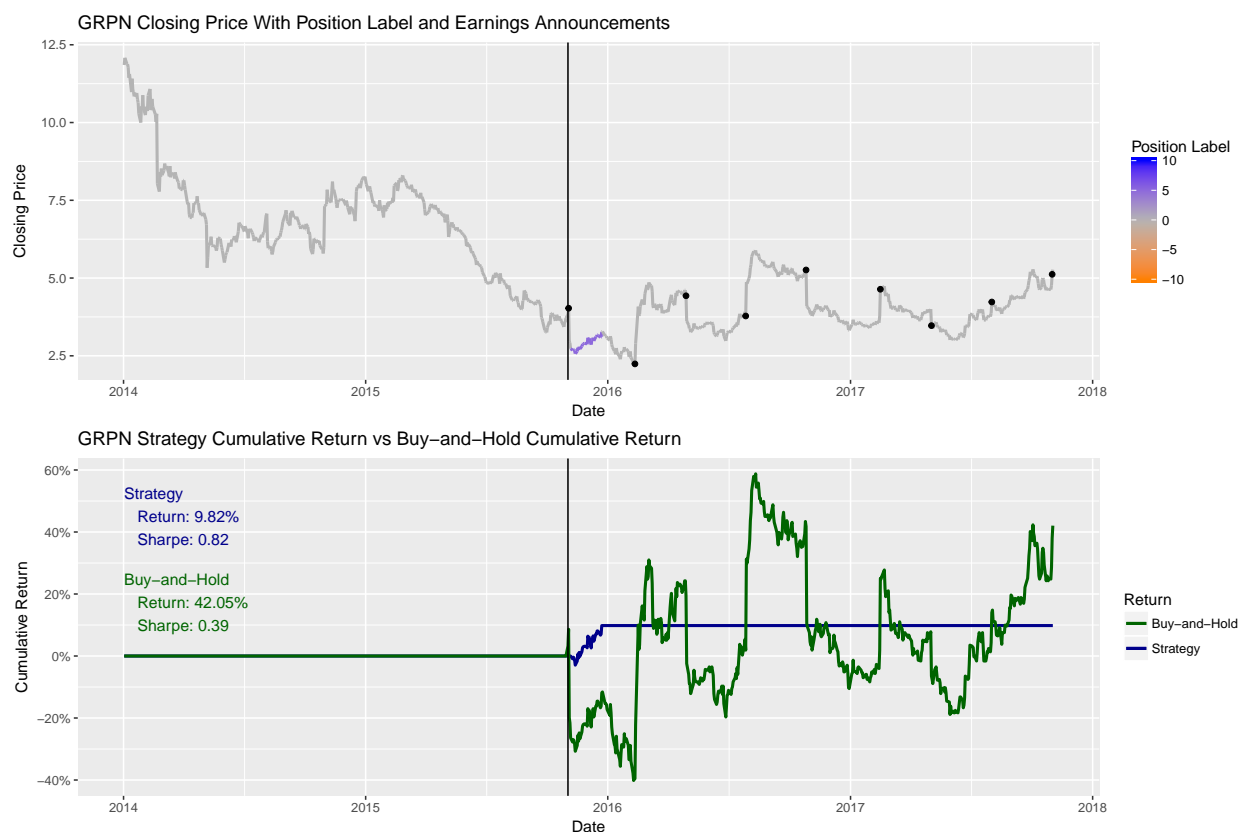
There are a number of clusters that contain undesirable stock patterns. These include clusters that exhibit no particular trend, are too volatile, or exhibit a negative trend.

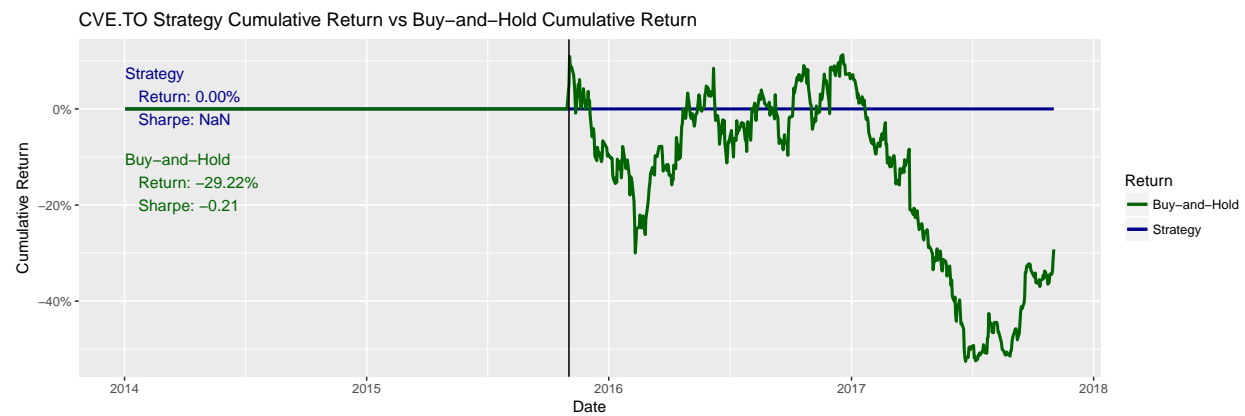
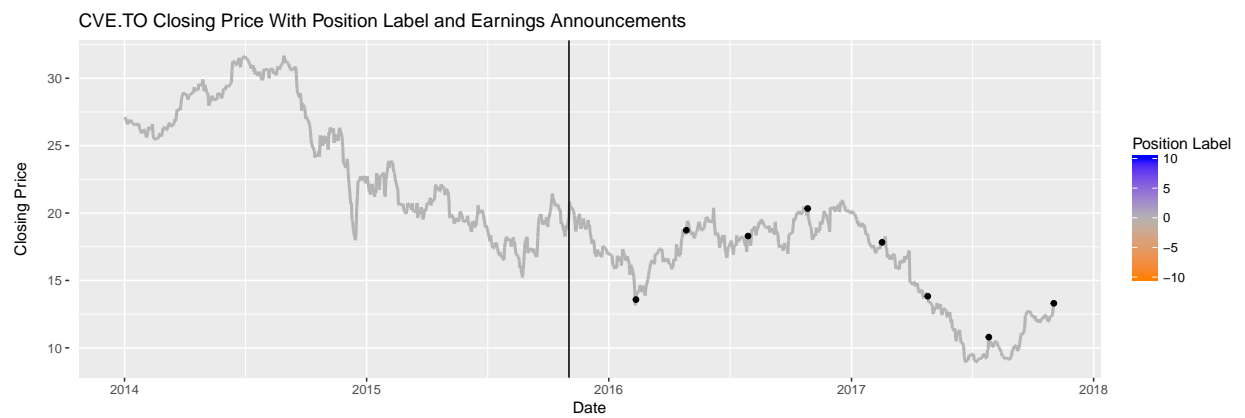
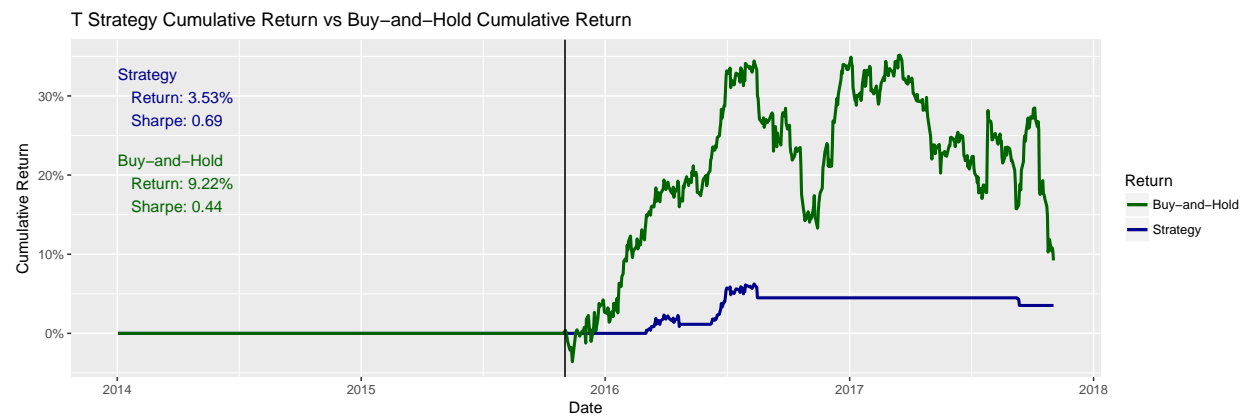
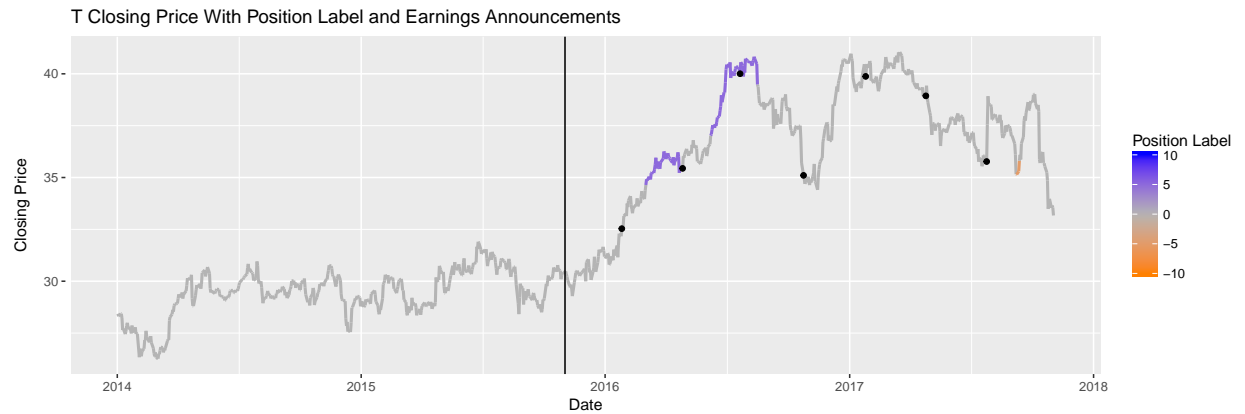
5.1 No Trend

Stocks within this cluster exhibit no trend. Some stocks trade within a narrow range and others trade within a large range. Stocks that trade within a narrow range are candidates for moving into the strong upward trend cluster but not until they break out of the established trading range.

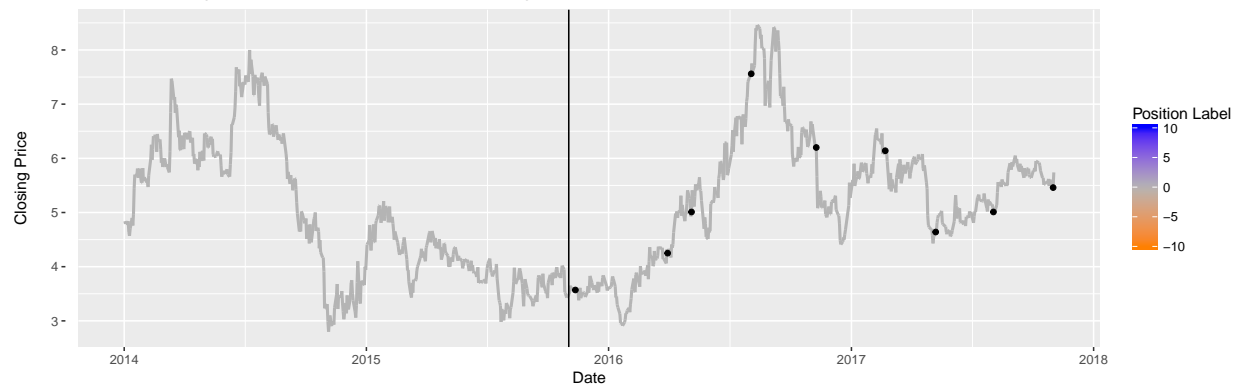
During typical market environments, there are large proportion of stocks that belong in this cluster and even stocks that do exhibit trends have large numbers of training examples where they are not trending. Most of the stocks in this cluster have a 0 position label for the entirety of their history. There are many training examples in this cluster for the model to learn from.

Examples of stocks in this cluster include GRPN, SHOP, BAC, T, CVE.TO, SSL.TO, RGC, NKE, IBM, VFC, and VET.TO.

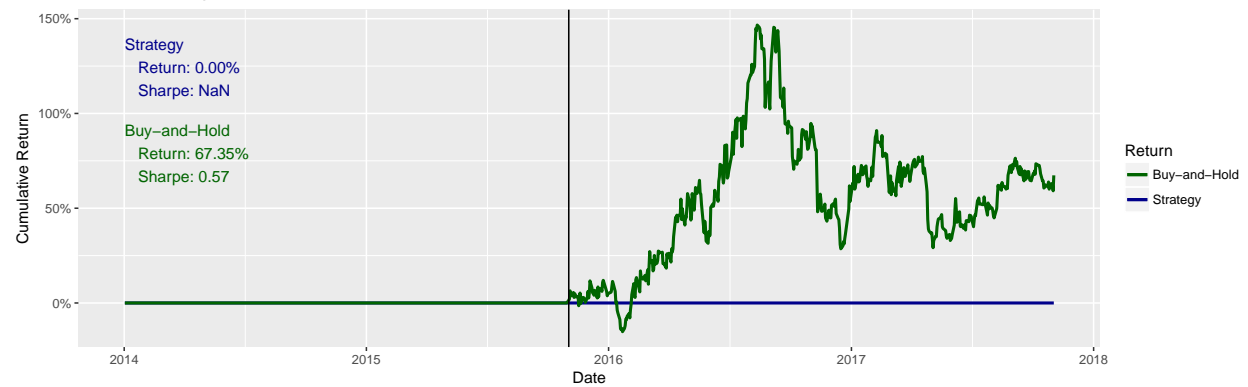




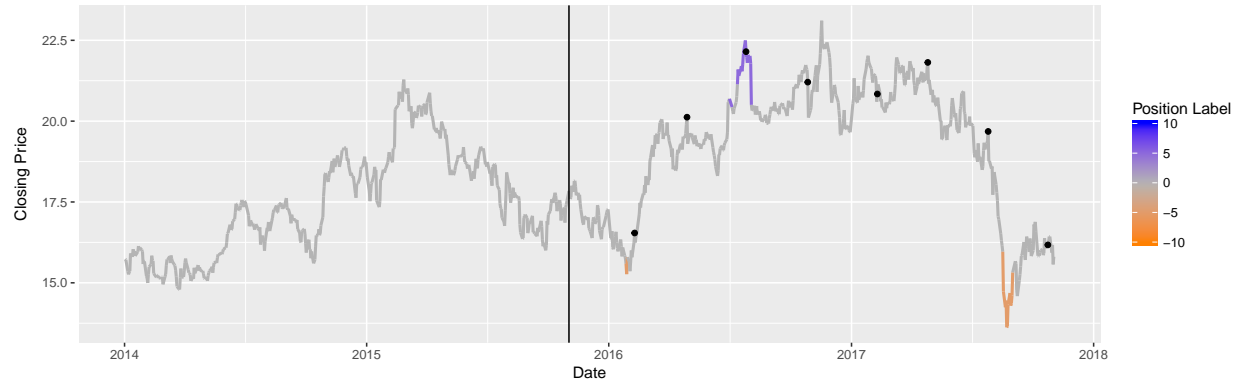
SSL.TO Closing Price With Position Label and Earnings Announcements



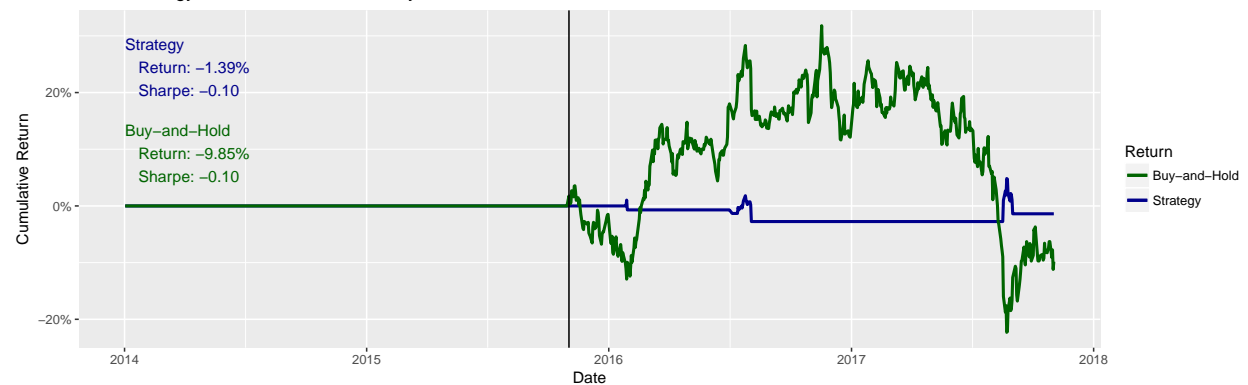
SSL.TO Strategy Cumulative Return vs Buy-and-Hold Cumulative Return

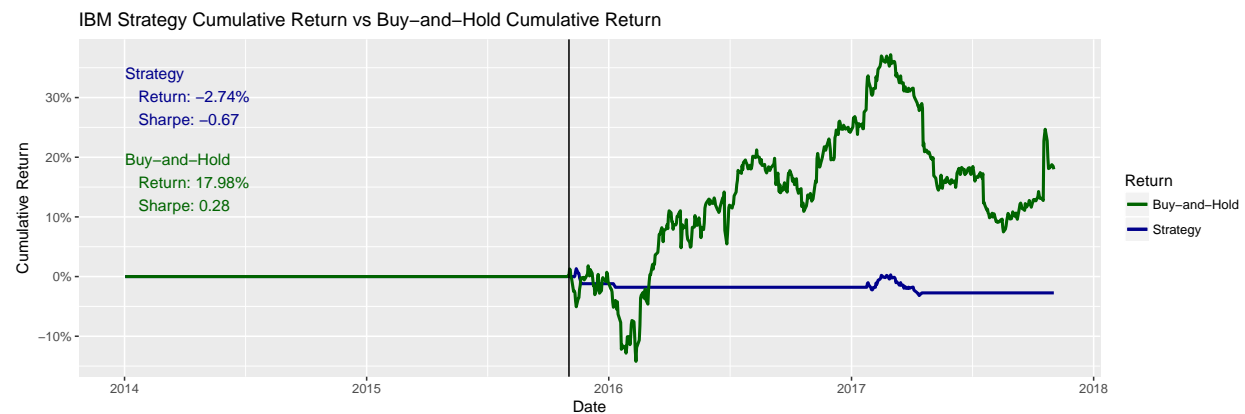
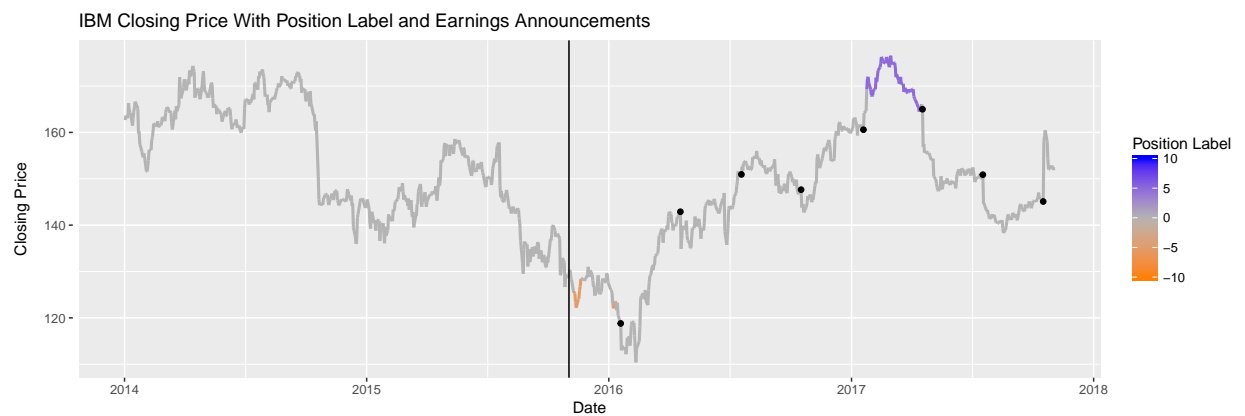
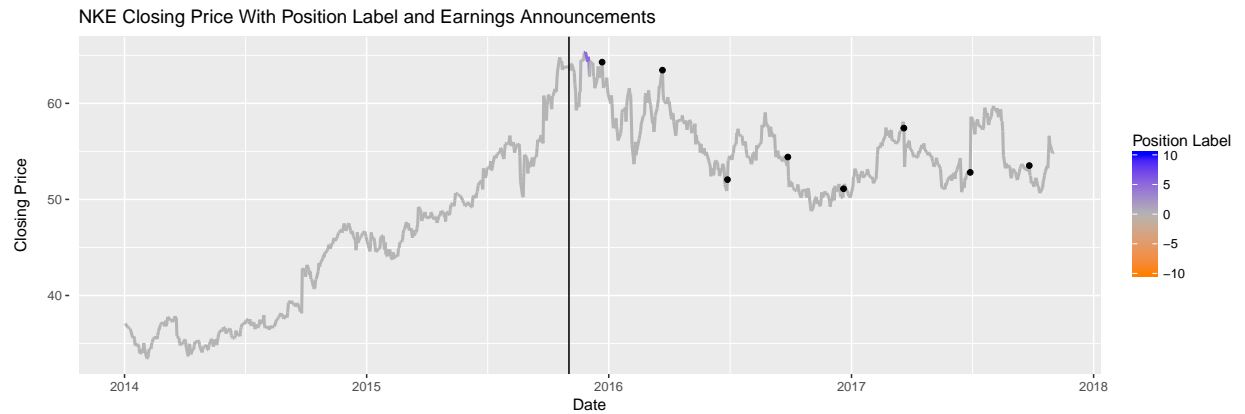


RGC Closing Price With Position Label and Earnings Announcements

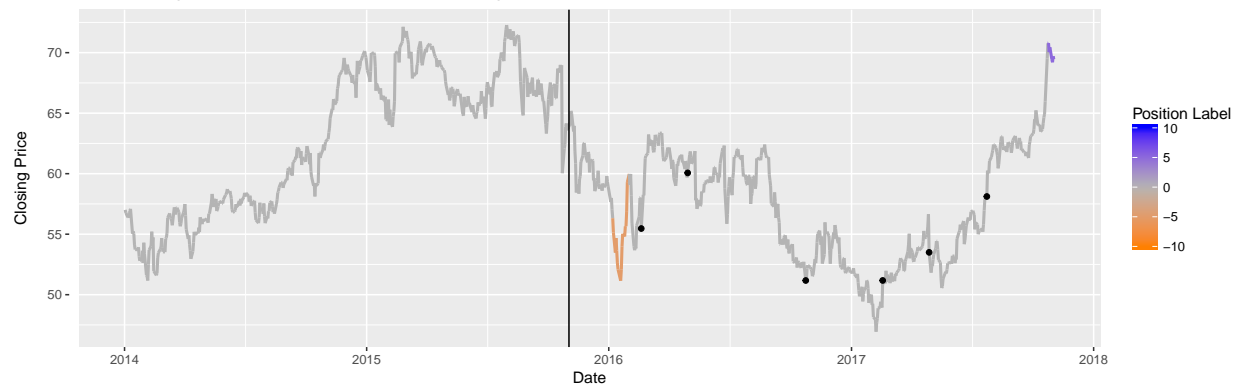


RGC Strategy Cumulative Return vs Buy-and-Hold Cumulative Return

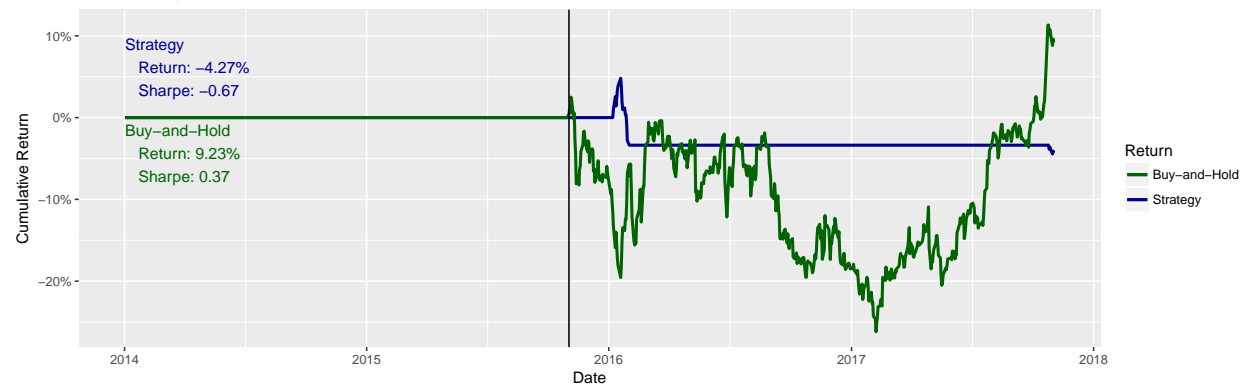




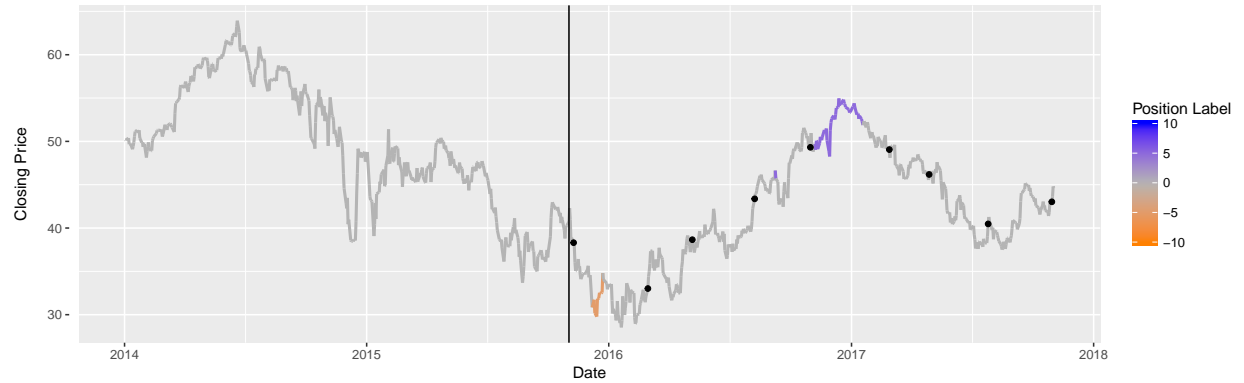
VFC Closing Price With Position Label and Earnings Announcements



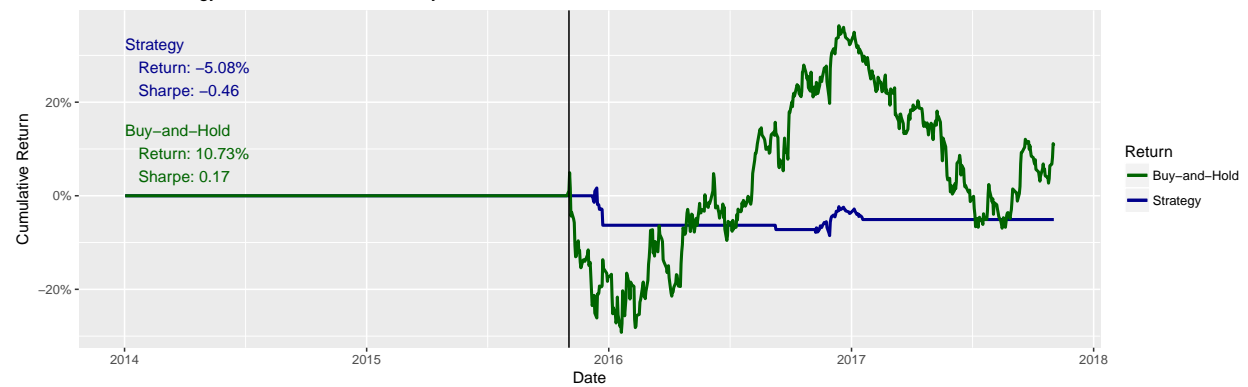
VFC Strategy Cumulative Return vs Buy-and-Hold Cumulative Return



VET.TO Closing Price With Position Label and Earnings Announcements



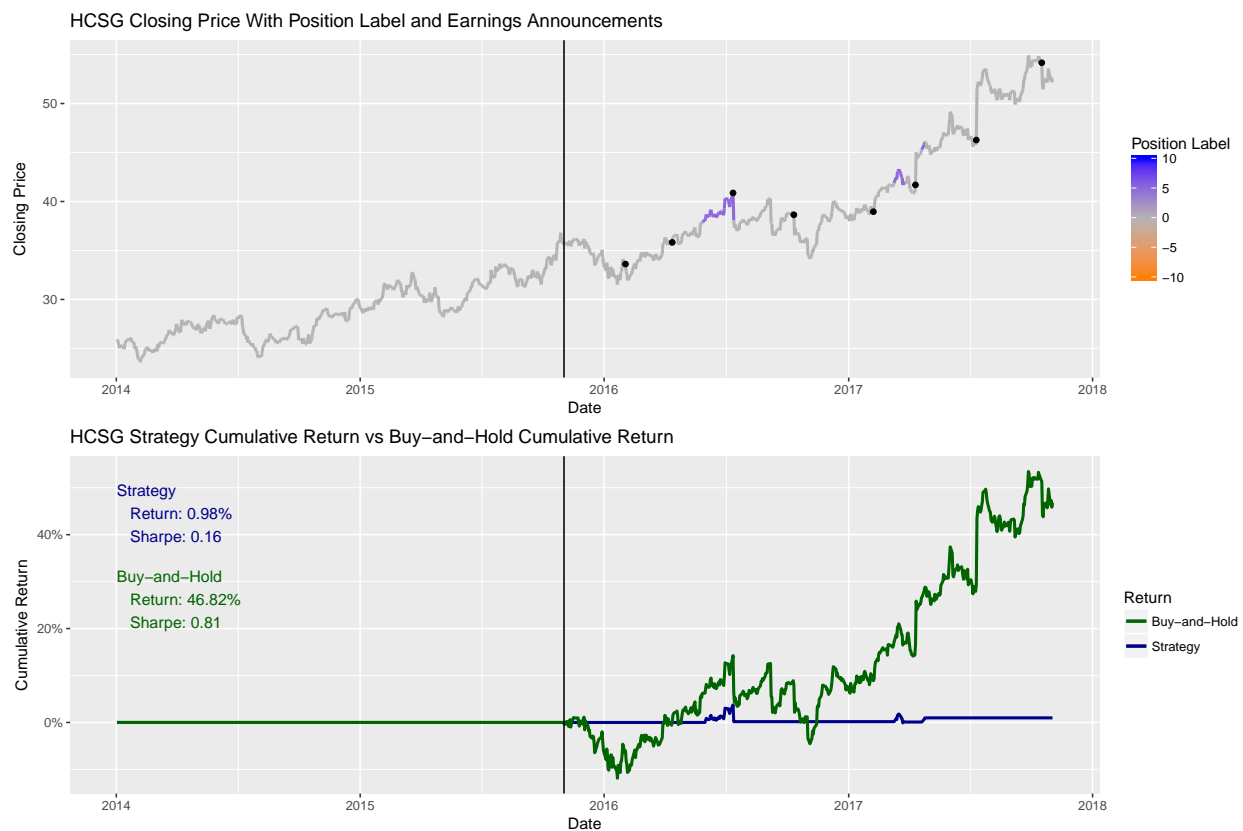
VET.TO Strategy Cumulative Return vs Buy-and-Hold Cumulative Return

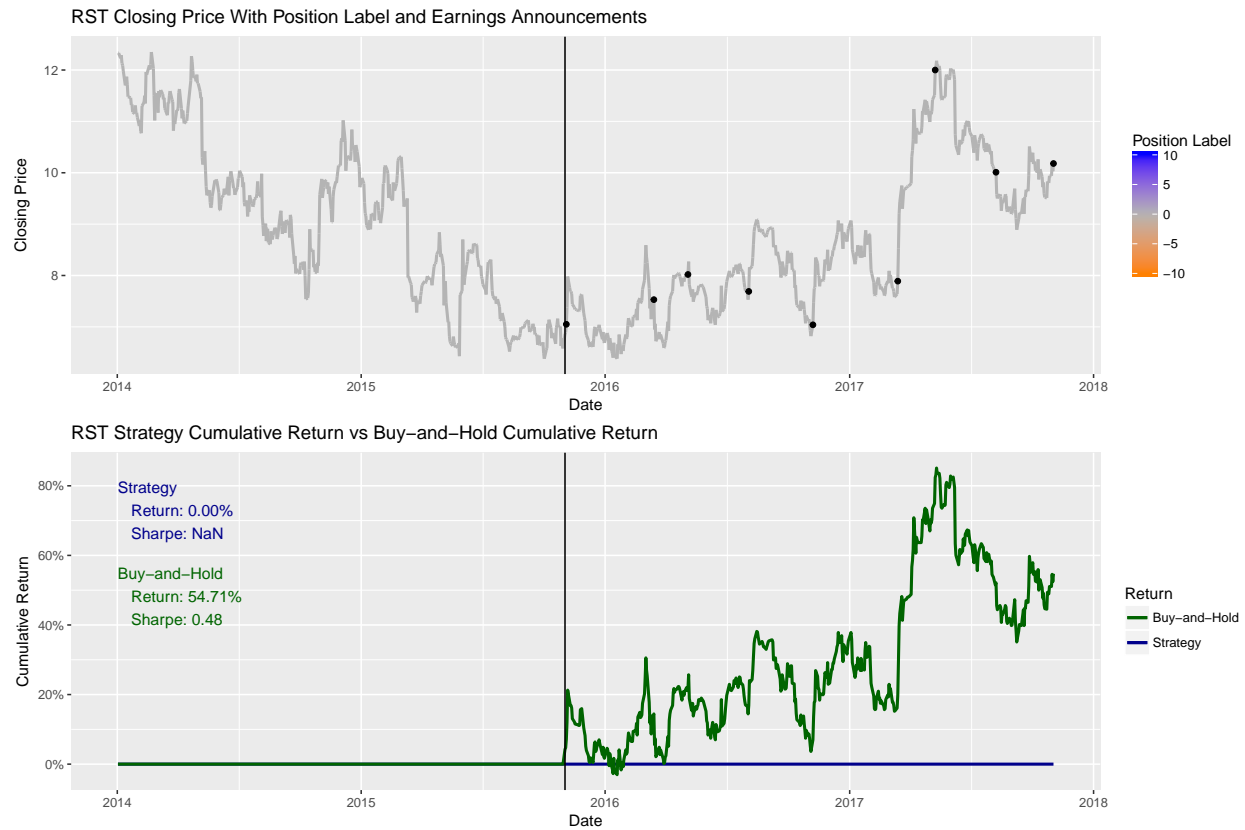


5.2 Too Volatile

The stocks in this cluster may exhibit an overall positive trend on a long time scale, but the nature of the gains are too volatile or inconsistent in nature. These stocks do not have a strong base to build on and are generally avoided by the strategy as indicated by the large number of days with a position label of 0 despite the overall positive trend.

Examples of stocks in this cluster include HCSG and RST.

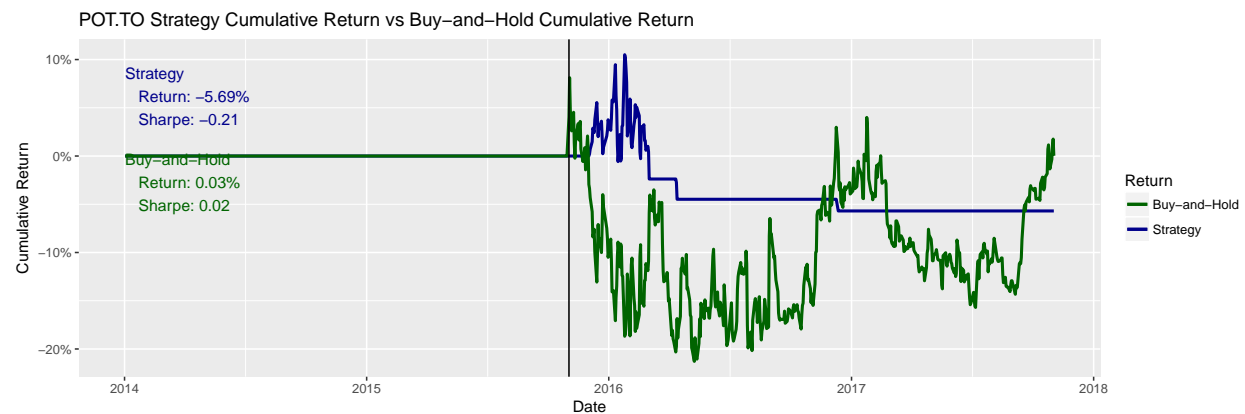
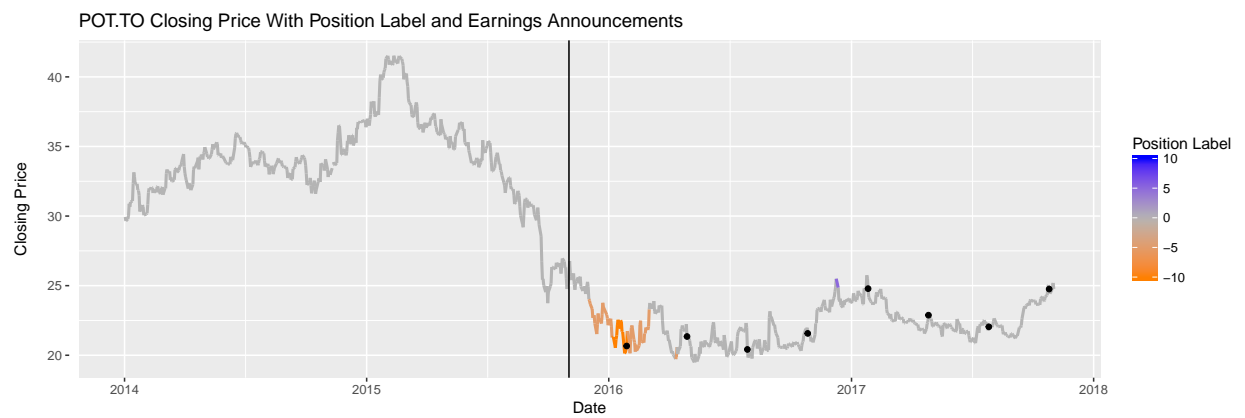
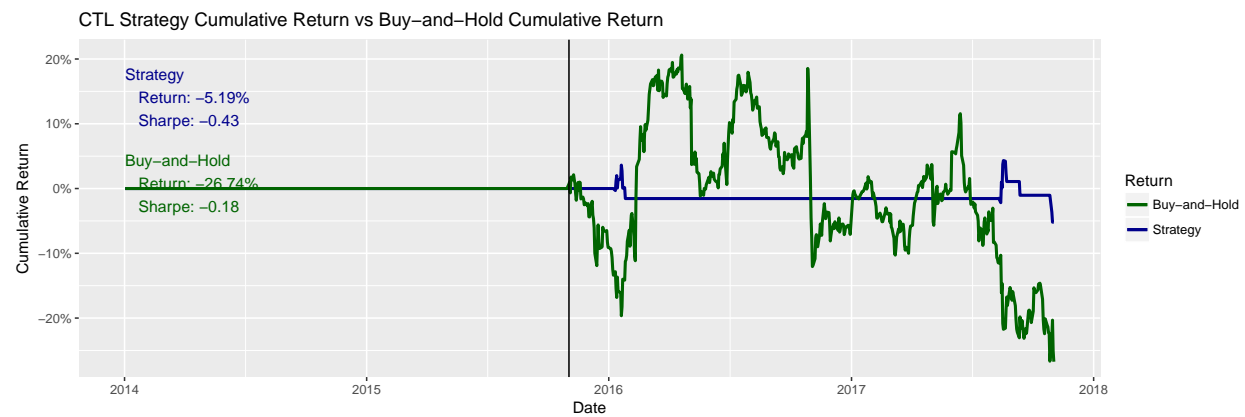
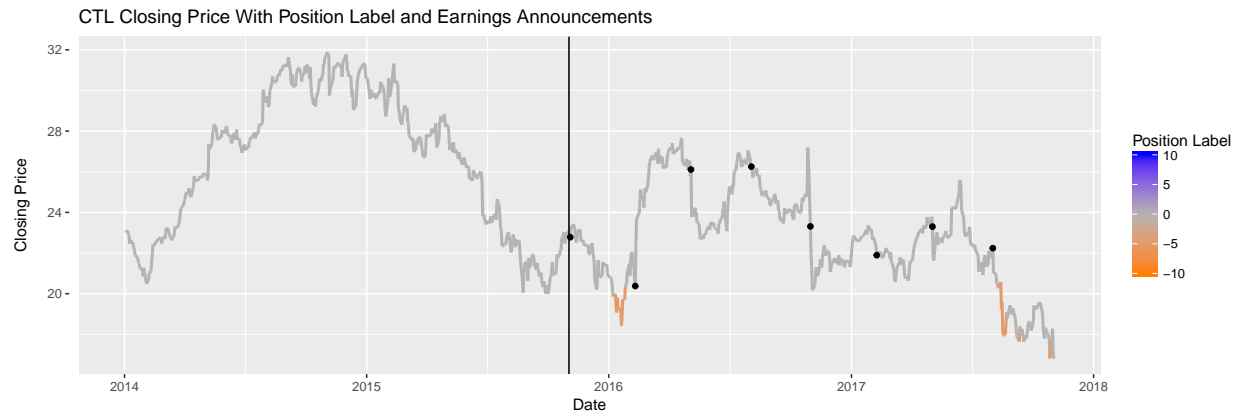




5.3 Downward Trend

The stocks in this cluster exhibit patterns that are the opposite of the clusters described in section 4.1 and 4.2. These stocks exhibit a downward trend and the strategy may enter into short positions if the trend persists with low volatility. If there is a downward trend with high volatility, the strategy remains on the sidelines.

Examples of stocks in this cluster include CTL and POT.TO.



6. Training Data Recommendations

Reviewing the training data and identifying these clusters was a helpful exercise for feature engineering and building the model. This exercise was also undertaken to determine whether the training data contains enough diverse examples to build a good predictive model. Based on this review, I recommend that additional training examples be provided that belong in the downward trend cluster described in section 5.3. While there are many stocks in the training data that have a lot of observations with either strong positive position labels and observations with a 0 position label, there are comparatively few stocks or training examples where the position label is negative.

A model built on a training set that contains more training examples that have negative position labels should perform better on unseen data and a diverse set of market environments. Positive momentum strategies have performed extremely well for the past several years in North American equity markets, but the market regime and environment may shift unpredictably in the future. A strategy that can go short with greater frequency may become more appropriate. Therefore, I recommend providing position labels for four more stocks that belong firmly to the downward trend cluster.