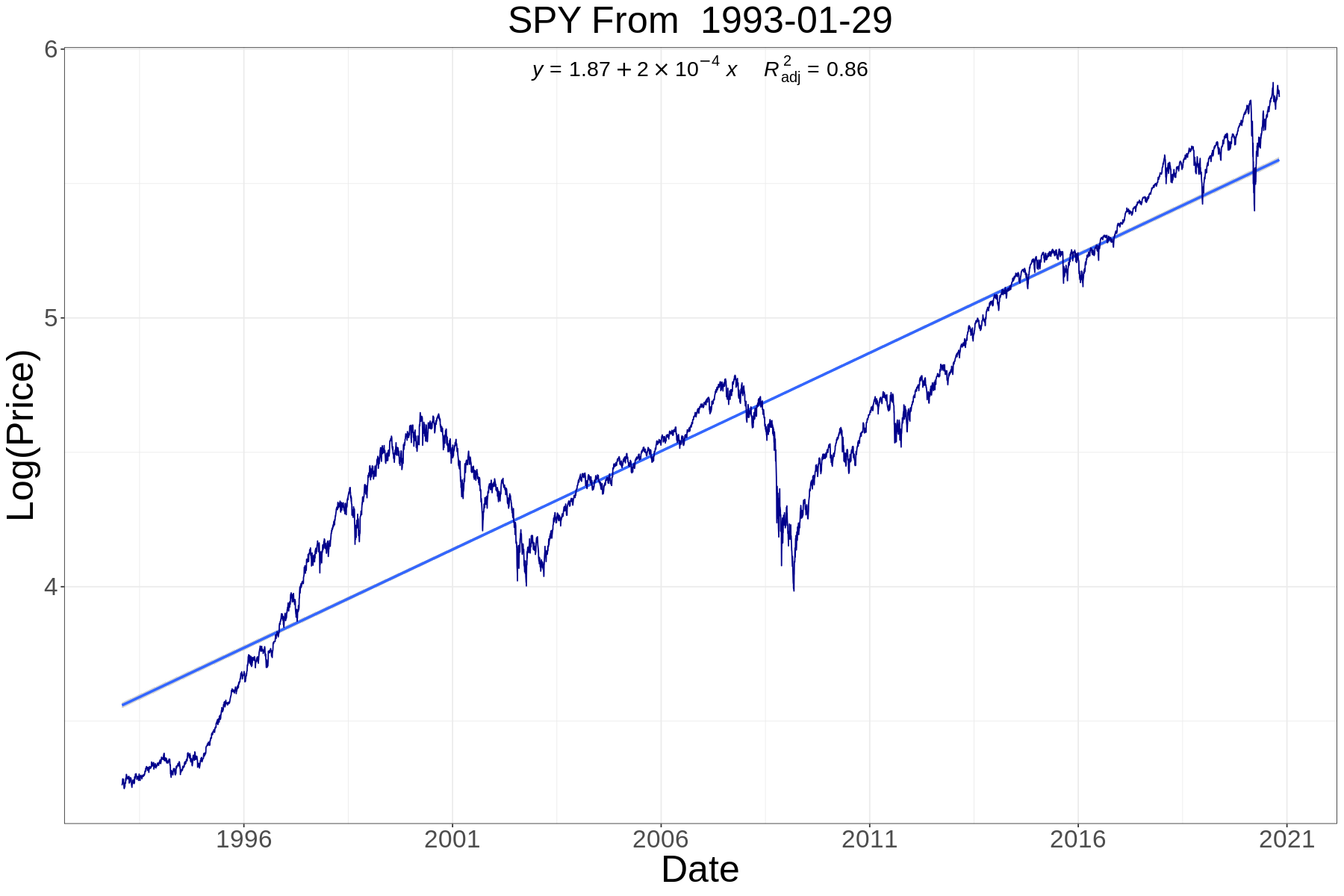
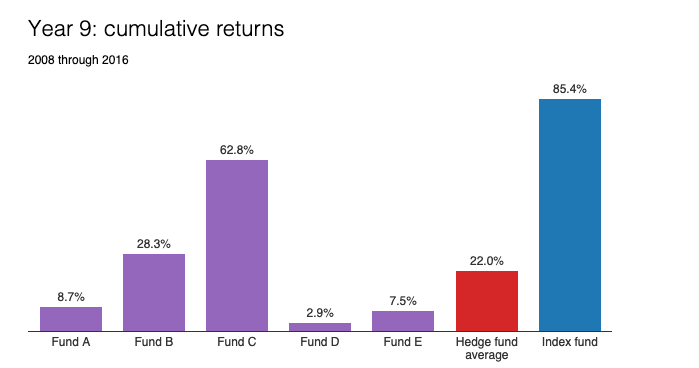
Titles: Smart Beta? What is the Return of Investing the newly added stocks in S&P 500?

Introduction

In stock markets, everyone wants to win the money. Actually, if invest on SPY, which is an ETF (exchange-traded fund) follow the S&P 500 index, from it was founded and stay in the market. You will get about 7.57% annualized rate of return (0.2 ‰ per day) (Figreu 1). However, some people may not satisfy with the market return, they tried to develop methods to beat market and earn more money. However, historical data showed that most of active investment methods will lose to market in the end. There is a famous bet. In 2008, Buffet issued a challenge to the hedge fund industry saying that no hedge funds can beat S&P 500 in the next 10 years. He wins in 2015 because his opponent conceded defeat ahead of the contest's scheduled wrap-up on December 31, 2017. (Figure 2) Since most of the people do not have the ability to select stock and time to invest and the active investment will finally lose to the market, smart beta strategies are popular in recent years. The smart beta strategies are passive intersegments that use alternative index construction rules. For example, use equal weighting instead of market capitalization weighting of S&P 500 is a smart beta. Here, we develop a new smart beta strategy that only invest on the newly added stocks in S&P500. We will compare the short-term and long-term return of our smart beta strategy with SPY. We also measure the volatility to evaluate the risk of our smart beta methods.



*Figure 1, Historical Price of SPY since it was founded.*



*Figure 2, 9 years cumulative return of 5 hedge funds and Vanguard's S&P 500 Admiral fund (Source:* [*BRK 2016 letter*](http://www.berkshirehathaway.com/letters/2016ltr.pdf)*)*

* Objective

Our object to propose a model for five and a ten-year return of investment of standard and poor's 500, by applying the analysis on the newest 500 hundred companies who had been added in S&P stocks since "20..-20..", in our model we will be subjected to the comparison between short term 5 years and long term 10 years with the most famous funds object to tracking the S&P 500 index SPY. We will test out the proposed smart beta strategy to determine how the strategy show the difference between short- and long-term investment and how can we help to minimize the investment risk by choosing the right term, in additional we will obtain how the trend behaves; and find the relation in the first five and ten years.

The proposal should include the below list.

The proposal should include:

(1) description of the dataset;

Five years and ten years return of investment for newly added companies. The dataset was constructed from R package quantmod and a list for companies joining date in the s&p500. Then the five- and ten-years investment return was calculated for each company from the first day of joining s&p500.

(2) the scientific goals, specific hypotheses;

Investigate a new investing strategy where the investor needs to invest in the newly added companies only. New investors usually do not know which company to invest in, and there are a lot of strategies for long term investment. One strategy is to invest in newly added companies

(3) preliminary plots,

(4) plan for analysis and modeling.

Our plan may use the test methods we covered in class. Firstly, we may compare these statistics—mean, weighted mean, median, quantile, variance, standard deviation, and covariance. Secondly, we will do the work related to hypothesis tests. For distribution function, we may use shapiro test and chi-square test. For mean and sd, we may use t test and var test. For test of independence, we may use chi-square test and fisher test. For test of correlation test, we may use correlation test. For analysis of variance, we may use binominal test, paired t test, shapiro test, bartlett test, Kruskal-Wallis test, and Friedman test. We need to determine whether there is a significant difference in the mean and comparing two populations mean.

As for the part of modelling, we may use linear regression model, regression tree, random forest, support vector machine, and analysis time series. We need to filter model independent variables, do regression diagnosis, and analyze outliers.

Topic: A New long term Investing Strategy in S&P 500.

Statistical analysis investment in Standard & Poor's “S&P” 500.

Materials

Table 1 contains the historical-components and changes of S&P500 back to 1957. The column date is the S&P 500 component date. The column name is the name of the company. The column value is the symbol of the ticket. The variable is the action, add tickets or remove tickets.

## Table1: [sp500-historical-components-and-changes](https://github.com/leosmigel/analyzingalpha/tree/master/sp500-historical-components-and-changes)

<https://github.com/leosmigel/analyzingalpha/blob/master/sp500-historical-components-and-changes/sp500_history.csv>

We used R library quantmod to get the historical data of stocks from source yahoo finance.

**Methods**

We used the oldest and one of the most popular funds SPDR S&P 500 ETF Trust (SPY) that aims to track the S&P 500 Index as benchmark. It was found on Jan 29, 1993. It means that we only focus the newly added components after Jan 29, 1993.

Strategies 1: One-time investment.

When S&P 500 added new components, we invest 1000 dollars to those companies and 1000 dollars to SPY and sells after 1,3,5, and 10 years. We used Anova or t-test (we will discuss on this) to test whether the 1, 3, 5, and 10 years return of our smart beta are different from the benchmark. We can also calculate the standard deviation of each ticket (you need not only the start point and the end point of the adjust price to calculate to standard deviation of each stock.) and then compare with benchmark

Strategies 2: Continue Cash Flow

When S&P 500 added new components, we invest 1000 dollars to those companies and 1000 dollars to SPY without withdrawal. We compare the return and standard deviation with benchmark.