Super Bowl Outcomes as a Predictor for Annual Stock Market Returns

The Fried Eggs (Katrina Francis, Kevin Le)

May 21, 2018

1 Introduction

For our project, we decided to investigate the old wives' tale that the league to which the winner of a given year's Super Bowl belongs can predict whether stocks will rise or fall overall. According to the tale, stocks will be "bearish" (will fall overall) if the winner of the Super Bowl belonged to the AFL before the merger with the NFL in the 1950's, or if they are a newer team that is a part of the AFC. On the other hand, if they belonged to the NFL before the merger (if they are newer and a part of the NFC), the tale claims that stocks will be "bullish" (will rise overall) [2].

We decided to pick some well-known indices of the stock market, along with indices of various industries and stocks of sports-related companies, to investigate whether or not the outcome of the Super Bowl can be linked to their annual returns.

2 Data Set

The dataset consists of three main categories: stock market indices, indices of a few different industries, and stocks of sports-related companies. It should be noted that many of the other stocks and ETFs (all of the industry indices except for the DJU) only date back to the mid-2000's, which greatly decreases the number of conclusions which can be drawn from them. Of course, the statistical tests will be able to account for this.

The dataset for each stock as we acquired it consists of standard, daily open, high, low, close, and volume values [1].

For the old wives' tale, the relevant returns are the annual returns. Thus, we selected the open price for the first day of that year for which we had data (as long as that day was in the first week of the year) as the annual open price. We used the close price for the last day of that year for which we had data (as long as it was in the last week of the year) as the annual close price. We used these to calculate the annual log returns, which we use throughout this report.

3 Project Goals

The goal of this project is ultimately to determine whether the outcome of the Super Bowl and the annual returns for our various stocks and indices are correlated. If such a correlation exists, the Super Bowl outcome could be an interesting predictor of the overall trends in the stock market, or perhaps even individual industries or stocks.

4 Analysis

4.1 Single Stock Analysis

4.1.1 Goodness of Fit Test

For the single stock analyses, we first wanted to check how well our stocks' respective yearly log returns fit a normal distribution. As our ultimate goal is to attempt a regression between the returns of a stock/index and the outcome of the Super Bowl, normality is a necessary pre-requisite for that type of analysis.

For our sports analysis, our calculations and tests are under the assumption that the yearly log returns are normal. The null hypothesis is that the data fits a normal distribution. The alternative hypothesis is that the data does not fit a normal distribution. We used a chi-squared goodness of fit test to see how well the data fit a normal distribution with a sample mean and sample standard deviation taken from the stock. Since parameters cannot be estimated in a Kolmogorov-Smirnov goodness of fit test, we decided to use a chi-squared goodness of fit test. To set up the data to run such a test, we wrote a binning function which groups ranges of yearly log returns and calculates the frequency of each range. For our significance level, the convention is to use 0.05. The SPX, IYW, and IYF all had P-values < 0.05. For these stocks, we reject the null hypothesis, indicating that there is sufficient evidence that the yearly log returns of these stocks do not follow a normal distribution. The remaining stocks had P-values > 0.05. For these stocks, we fail to reject the null hypothesis, indicating that there is not sufficient evidence that the yearly log returns of these stocks do not follow a normal distribution.

Table 1: Chi-Squared Goodness of Fit for Normal Distribution

Stock	P-Value
SPX	0.002584
DJI	0.06051
NDQ	0.13
IYW	0.04462
IYF	0.005293
MXI	0.9884
XLP	0.2356
XLU	0.7285
DJU	0.2629
NKE	0.1561
DKS	0.7584
FL	0.2578
LULU	0.7314
UAA	0.4549

A histogram of the yearly log returns appeared to present frequency distributions that were roughly normal, however as we saw in the P-values the SPX, IYW, and IYF are likely not normal. We decided to further check with normal probability plots for each stock's yearly log returns. In the cases of SPX, IYW, and IYF, the data plots do not appear linear and are skewed to the right or the left, further supporting that these stocks do not carry a normal distribution. However, the remaining stocks' normal probability plots are closer to being linear, further supporting that these stocks can be fit with a normal distribution.

4.1.2 Confidence Intervals for Mean and Standard Deviation

To better understand the data, we wanted to understand what the mean and variance of each stock's yearly log returns were. To do this, we looked at two-sided 95% confidence intervals for each respective mean and variance. Since the population μ and σ^2 of the data are unknown, we used the sample mean, sample variance, and sample deviation for our interval construction. For the confidence interval of each μ, we used a t-distribution with n-1 degrees of freedom. For the confidence interval of each σ^2, we used a chi-squared distribution with n-1 degrees of freedom.

Table 2: 95% Two-Sided Confidence Intervals for μ and σ^2

Stock	Confidence Interval for μ	Confidence interval for σ^2
SPX	[-0.318, 0.413]	[0.0272, 0.0442]
DJI	[-0.353, 0.462]	[0.0334, 0.0556]
NDQ	[-0.326, 0.5]	[0.0322, 0.0603]
IYW	[-0.444, 0.645]	[0.0308, 0.177]
IYF	[-0.596, 0.632]	[0.039, 0.224]
MXI	[-0.687, 0.74]	[0.05, 0.316]
XLP	[-0.117, 0.297]	[0.00443, 0.0254]
XLU	[-0.294, 0.419]	[0.0132, 0.0756]
DJU	[-0.376, 0.424]	[0.0308, 0.0557]
NKE	[-0.431, 0.806]	[0.058, 0.165]
DKS	[-0.856, 0.956]	[0.085, 0.488]
FL	[-0.762, 0.887]	[0.0704, 0.405]
LULU	[-1.73, 1.95]	[0.313, 2.2]
UAA	[-0.891, 1.07]	[0.0991, 0.569]

As seen, the 95% confidence intervals for μ were relatively small. Each mean confidence interval were around zero. The 95% confidence interval for variances were also all relatively small for all sets with the exception of FL, suggesting that the variability of the yearly log returns for each stock mostly possessed little variability. FL, on the other hand, seems to experience a fair amount of variability.

4.1.3 Simple Linear Regression

Table 3: Linear Regression of Yearly Log Returns Against Time

Stock	$\hat{\beta}_0$	$\hat{\beta}_1$	R^2
SPX	0.00918	0.00057	0.0141
DJI	0.0304	0.000399	0.00463
NDQ	0.0759	0.000278	0.00097
IYW	-0.0284	0.0198	0.0835
IYF	-0.205	0.0344	0.197
MXI	-0.0419	0.0114	0.014
XLP	0.0443	0.00699	0.072
XLU	-0.00662	0.0106	0.0556
DJU	-0.0311	0.00122	0.0245
NKE	0.308	-0.00774	0.0507
DKS	0.268	-0.0335	0.0861
FL	-0.0768	0.0214	0.0426
LULU	-0.0112	0.0216	0.00649
UAA	0.225	-0.0212	0.0295

Next, we performed a linear regression on each stock's yearly log returns against time. In all cases, the R^2 value is very small and close to zero, indicating little linearity or correlation. A look at the scatter plots with the linear regression line shows that many points are very disparate from the drawn line. The slope of each linear regression is also quite low, again indicating little linear correlation. Additionally, the residual plots do not show a symmetric distribution of points above and below zero in any pattern.

However, we wanted to take a closer look by observing the 95% confidence intervals on the slope ($\hat{\beta}_1$) and intercept ($\hat{\beta}_0$).

Table 4: 95% Two-sided confidence intervals for $\hat{\beta}_0$ and $\hat{\beta}_1$

Confidence interval for $\hat{\beta}_0$	Confidence interval for $\hat{\beta}_1$
[-0.0544, 0.0727]	[-0.000253, 0.00139]
[-0.0444, 0.105]	[-0.000664, 0.00146]
[-0.0179, 0.17]	[-0.00173, 0.00229]
[-0.369, 0.312]	[-0.0265, 0.0662]
[-0.565, 0.154]	[-0.0145, 0.0832]
[-0.532, 0.448]	[-0.0608, 0.0837]
[-0.0859, 0.174]	[-0.0107, 0.0247]
[-0.233, 0.22]	[-0.0202, 0.0413]
[-0.116, 0.0539]	[-0.00042, 0.00286]
[0.0777, 0.538]	[-0.0207, 0.00522]
[-0.298, 0.834]	[-0.11, 0.0434]
[-0.604, 0.45]	[-0.0502, 0.0931]
[-1.37, 1.34]	[-0.197, 0.24]
[-0.405, 0.855]	[-0.107, 0.0644]

Most revealing are the confidence intervals of the slope, or $\hat{\beta}_0$. In all cases, it is around zero, with the lower bound being negative and the upper bound being positive. This indicates little to no correlation.

Therefore, a linear regression is not a good predictor of future yearly log returns – perhaps the sports data will be better.

4.2 Two Stock Analysis

The two stock analyses will not be explored in depth, as they are not crucial to the main goal of this project, the examination of the old wives' tale.

4.2.1 Test for Independence

The app gives two separate options for tests for independence: the chi-squared contingency table test, and the distance correlation test for independence.

The contingency table test is intended for categorical data, so the application of it to two continuous variables (the log returns of two stocks), it is not ideal. Because the data must be 'binned', there is information lost about the data. For larger datasets, this has less of an impact, as the number of bins can be increased to compensate.

It should be noted that the app will automatically merge bins that don't meet a certain specified minimum value, as the assumptions of chi-squared test fail if bins are too empty.

In comparison, the distance correlation coefficient test for independence uses a more complex version of the correlation coefficient. While a finding of 0 for the linear correlation coefficient does not indicate independence between the two variables, the distance correlation coefficient is formulated such that a 0 can be taken to mean independence [3]. The app uses a predefined function which calculates a p-value through simulation.

4.3 Super Bowl Outcome Analysis

For the Super Bowl outcome analysis, we had two main aims. First of all, can some dependency be established between the outcome of the Super Bowl and the outcome of the stock market for that year? If so, how much predictive information can we extract? This written analysis will focus on the Dow Jones Industrial Average.

4.3.1 Test for Independence

To the end of answering that first question, we performed a chi-squared contingency table test, which can help ascertain whether there is any relationship between the two variables of interest.

The contingency table test requires two categorical data variables. The old wives' tale is very vague, so translates very easily into two binary variables. For a given year, the stock market will either go up or down (bullish or bearish, respectively), and either an AFL or a NFL team will win the Super Bowl. Thus, we can construct a contingency table, as shown below, for each stock or index in our dataset.

Table 5: Contingency table for Dow-Jones Industrial Average

	Bearish	Bullish
AFL	9	6
NFL	5	31

The chi-squared test assumes independence between the two variables as the null hypothesis, and results that deviate far from that assumption will cause the rejection of the null hypothesis in favor of the conclusion that the two variables are dependent. Thus, a lower p-value for this test indicates greater confidence in the existence of a dependency.

Table 6: P-values for chi-squared test for independence

Stock	P-value
SPX	0.059
DJI	0.003
NDQ	0.237
DJU	0.328
NKE	0.353

The S&P 500 and Dow-Jones Industrial Average both have acceptably low p-values to conclude with a reasonable degree of confidence that these indices are dependent on the outcome of the Super Bowl. Armed with this knowledge, we hoped to see if any other aspect of the Super Bowl might be able to make helpful predictions.

4.3.2 Multiple Linear Regression

To this end, we performed a multiple linear regression on two variables of interest: which league the winner belonged to, and what the point difference was. For this, we will look at the Dow-Jones Industrial Average, as it shows the most promise based off of the test for independence.

For the linear regression, it should be noted that since the NFL/AFL victory is a categorical variable, it was treated as having a value of 1 if the NFL won, or of 0 if the AFL won. Since this value only makes sense as a 1 or 0, this means that the multiple linear regression can be easily visualized in on a graph as two separate lines (Figure 1).

The equation below is the result of the linear regression model, where y is the annual log returns of the DJI, x_1 is the point difference in the Super Bowl, and x_2 is 0 if an AFL team wins, and 1 if an NFL team wins.

$$y = -0.000431x_1 + 0.136x_2 - 0.0228 \tag{1}$$

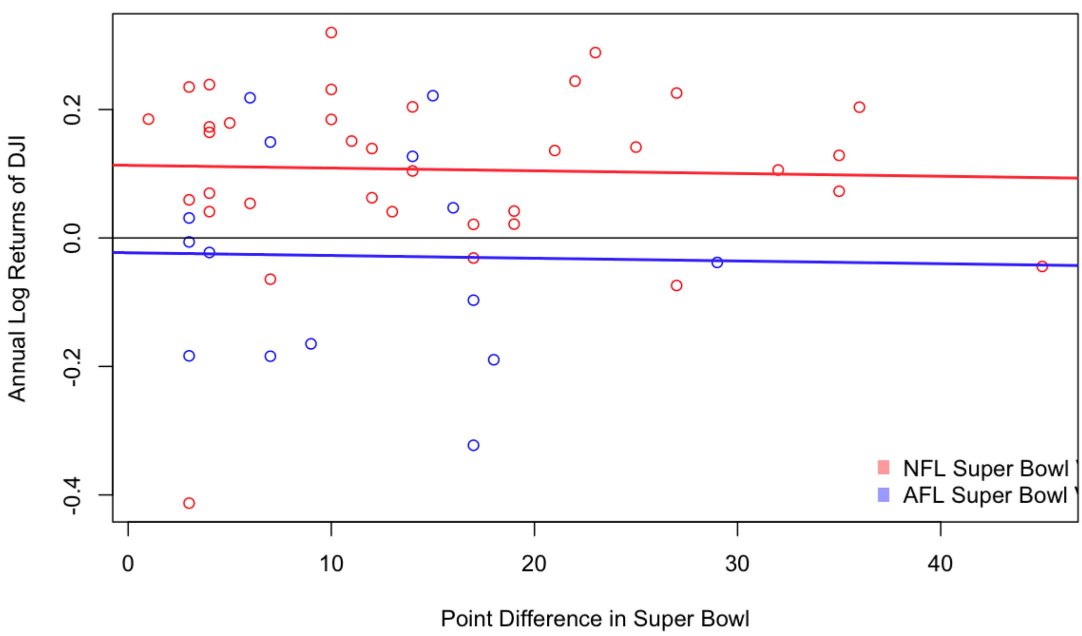
Table 7: 95% confidence intervals for the regression values

x_1 Slope	-0.00437	0.00351
x_2 Slope	0.0457	0.227
Intercept	-0.11	0.0639

Looking at the confidence intervals for the output of this multiple linear regression, we can conclude that in this model, point difference is not shown to be a useful component in the regression, as 0 lies within the 95% confidence interval. However, the dummy variable indicating AFL and NFL does not have 0 in the confidence interval, so this further supports the findings from the independence test that there is some correlation between the league of the Super Bowl and the DJI returns.

Examining the residual plot for the multiple linear regression (not pictured), it does not seem to bear any suspicious patterns. Thus, while we cannot make any claims about a linear correlation between the point difference and the stock returns, the regression has confirmed the dependency that the independence test introduced, and has further confirmed that it is a positive correlation, which is in line with the old wives' tale.

Figure 1: Regression Plot for Multiple Linear Regression on DJI



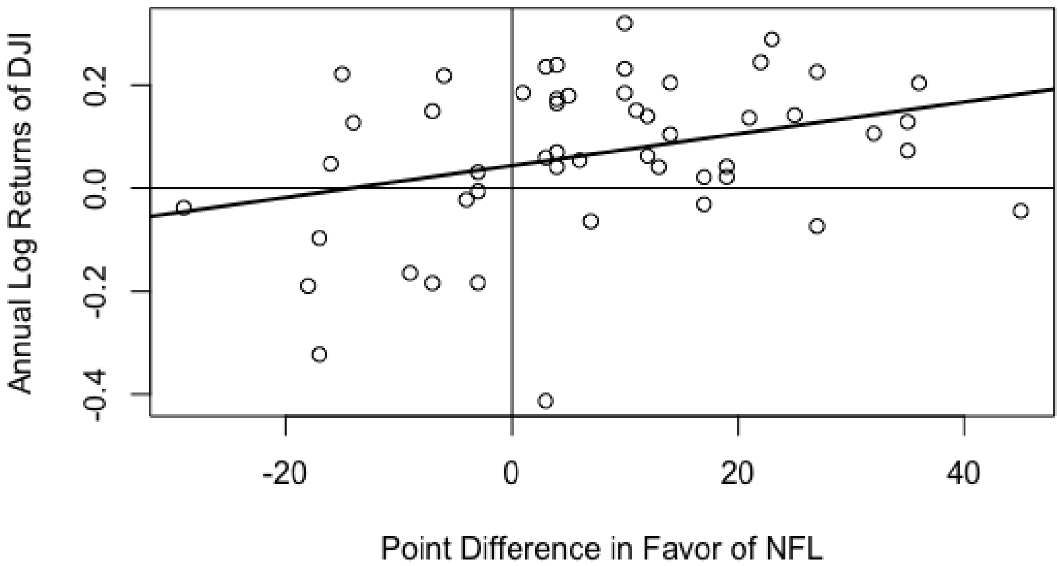
4.3.3 Simple Linear Regression

Not quite content, we also wanted to examine if perhaps the lack of finding in the multiple regression might be because of the point difference does not act independently of the league of the winner. Perhaps, the magnitude and direction should be considered jointly. That is, a bigger win by an NFL team might mean a very good year for the stock m market, whereas a very big win by an AFL team might mean a very bad year. The multiple linear regression would not allow for this type of relationship.

Thus, we performed simple linear regression with the point difference *in favor of the NFL* as the independent variable. The results of this are shown in Figure 2.

Examining the confidence intervals, the point difference variable now does not have 0 within its confidence interval, thus we can conclude that there is evidence of some correlation

Figure 2: Residual Plot for Simple Linear Regression on DJI



5 Conclusion

Having examined the correlation between the Super Bowl outcomes and the annual returns of the stock market, we conclude that there seems to be some evidence supporting the Super Bowl old wives' tale, but the Super Bowl old wives' tale does not yield any other predictions on the stock market besides if it goes up or down.

The chi-squared contingency table test provides evidence against Super Bowl results and stock market returns being independent. The highest P-value for independence, surprisingly, was for Nike – our selected sports stock. However, both the SPX and DJI stocks had low P-values suggesting some degree of dependence on the Super Bowl results. In the case of a general industry (DJU) and a specific stock (NKE), the annual returns appeared to be independent of the Super Bowl results.

However, the degree to which a team wins does not appear to have a strong correlation, based on our regression models for the relationship between the point difference in the Super Bowl and the annual returns of various stocks. In both the simple linear regression and multiple linear regression models, the correlation is less than 0.2. This indicates very little correlation between annual stock market returns and Super Bowl point differences.

Comparing the R^2 of the simple linear regression (0.104) with the R^2 of the multiple linear regression (0.162), it seems that it is an inferior model than the model which had point difference with an effective slope of 0. Looking at the two graphs again, the slopes are very low, and it is quite apparent neither of these models are particularly strong predictors. However, looking at the residuals, the data is not symmetrically distributed above and below 0. This suggests that the simple linear regression and multiple linear regression plots may not be good models of the relationship between annual stock market returns and the point difference in the Super Bowl.

Thus, we may perhaps conclude that the Super Bowl does not yield any more interesting predictions into the stock market than the old wives' tale suggests. There does appear to be some correlation between what team wins and whether or not the stock market goes up or down, which is consistent with the old wives' tale. However, the Super Bowl does not appear to be a good predictor of how much the stock market will go up or down, at least when looking at point differences. In other words, the Super Bowl old wives' tale only seems to suggest a binary association: either the stock market goes up by some degree, or it goes down by some degree.

Ultimately, we would not suggest investors base any significant investment on the results of the Super Bowl. However, if an investor does choose to place an investment based on the results of the Super Bowl, it should be concerning the SPX and DJI stocks.

References

[1] Stooq.com, apr 2018.

[2] William Power. Patriots bearish, eagles bullish in super bowl, feb 2018.

[3] Wikipedia. Distance correlation, jan 2018.