Significance of Date-Based Anomalies in the Stock Market

CMPT 353 @ Simon Fraser University

Spring 2024

Table of Contents

Project Members and Responsibilities

Alex Zhu:

- Formed research question
- Data gathering
- Wrote introduction, summary, and limitations and extensions
- Monday-Friday analysis and reporting
- Differences between months analysis and reporting

Divij Gupta:

- Beginning of the month analysis and reporting

Ege Candar

- Santa-Claus Rally analysis and reporting

Introduction

The stock market is fundamentally about transactions based on the perceived value of companies. However, value is a subjective, human concept and can be influenced by all kinds of disparate psychological factors. In this project, we examine various date-based phenomena that are claimed to influence stock prices to some extent.

There are two implications to this research:

1. Date-based phenomena challenge the efficient-market hypothesis (EMH), which states that asset prices only reflect the known information about the asset – in other words, it states that asset prices are based on the perceived profitability of the business
2. Exploiting date-based phenomena may lead to better or more consistent returns when making investments in practice

Monday-Friday Effect

A commonly cited date-based phenomena is one of the stock market performing better on Fridays than Mondays [1]. In more specific statements of this effect, it's suggested that Fridays may see higher stock prices than on the next Monday. (Note that trading typically doesn't occur on weekends.) A potential explanation involves traders feeling happier and more optimistic on Friday in anticipation of the weekend, and feeling less optimistic when returning to work on Monday. Another explanation considers that investor's optimism may be dampened when hearing about bad news on weekends, driving Monday prices down [1].

Big Data

To analyze this effect across the whole market, perhaps it's best to use a large dataset that reflects many different stocks. This dataset [2] is over 1 GB when unzipped, but it's actually smaller than that. It seems data has been duplicated. The Data folder contains the exact same ETFs and Stocks directories. We can also only look at the stocks rather than ETFs. Therefore, we delete the ETFs and Stocks folder in the root directory and keep the Data directory, leaving us with about 0.7GB of data.

Data	File folder
ETFs	File folder
Stocks	File folder

The CSVs inside each folder are titled with the stock name and have the following schema:

```
Date,Open,High,Low,Close,Volume,OpenInt
2015-11-11,18.5,25.9,18,24.5,1584600,0
2015-11-12,24.25,27.12,22.5,25,83000,0
2015-11-13,25.47,26.2,24.55,25.26,67300,0
```

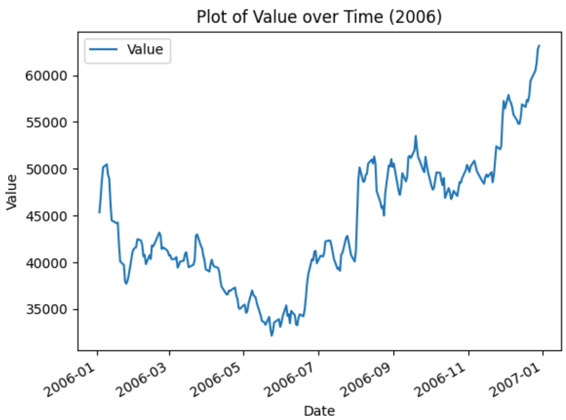
The open interest on the stock is not relevant to our analysis, and it would be much nicer to have a single stock price associated with each day rather than various prices. We'll perform an ETL with pyspark and for each stock, we'll average the low and high prices of each day, and then get a representative value for each day by averaging all the averages on each day for every available stock. Let's only look at two years for now, 2006 and 2016. The resulting CSVs should be small enough (about 10 years * 365 * 5/7 entries) to handle with pandas.

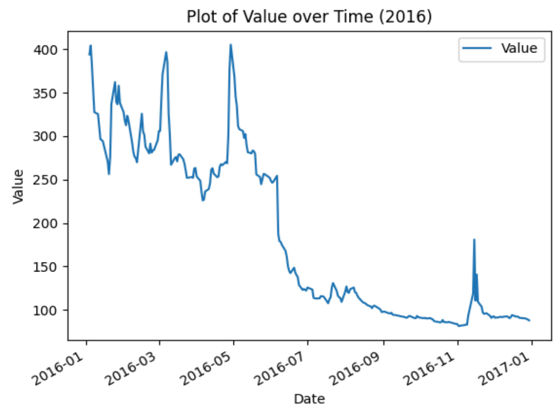
Once we have our data in a workable form, we'll pair Fridays with their following Monday. We'll then calculate the percent change from Friday to Monday, and perform two tests:

- A Mann-Whitney U-test between the values on Monday and on Friday to determine if stock values on Monday and Friday have the same median
- A Wilcoxon signed-rank test is like a MannWhitneyU test but on paired data instead of independent groups. In simple terms, it asks whether one half of the pair tends to be higher than the other. We'll perform this test on the Monday and Friday pairs.

Here are the results of the tests:
2006 - u-test p-value: 0.8654319153774157
2016 - u-test p-value: 0.8576045060966692
2006 - Wilcoxon p-value: 0.5382812738724851
2016 - Wilcoxon p-value: 0.11444247135693786

None of these tests meet the significance level of <0.05, thus we can't reject the null hypothesis, but more importantly, something concerning becomes clear when we look at a simple graph of the daily values:





Why are the values in 2006 so much higher than 2016, and why are stock values in 2016 declining so precipitously? The answer lies in the way we computed these values. The stocks available in the dataset around 2006 are mostly high-value stocks, while more and more low-value stocks start being introduced into the dataset later on, presumably since increased digitization makes it easier to track smaller companies. Mixing in these low-value stocks cause the average values to start dropping in later years. The analysis ends up reflecting the composition of the dataset more than the stock market itself. We will have to try something new….

S&P

The problem with considering a large set of stocks was that there was inconsistent availability of data when computing the average value for each day. We can resolve this issue by looking at a popular index called the S&P 500. An index essentially tracks the value across multiple stocks. The S&P in particular looks at the top 500 valued companies weighted by their market cap (their total value) and is widely used as an indicator of the overall market. Another advantage of looking at the S&P index is that we have much less data to deal with and can use pandas for our remaining analysis. For our dataset, we'll use this CSV dataset spanning 10 years [3] that has the following schema:

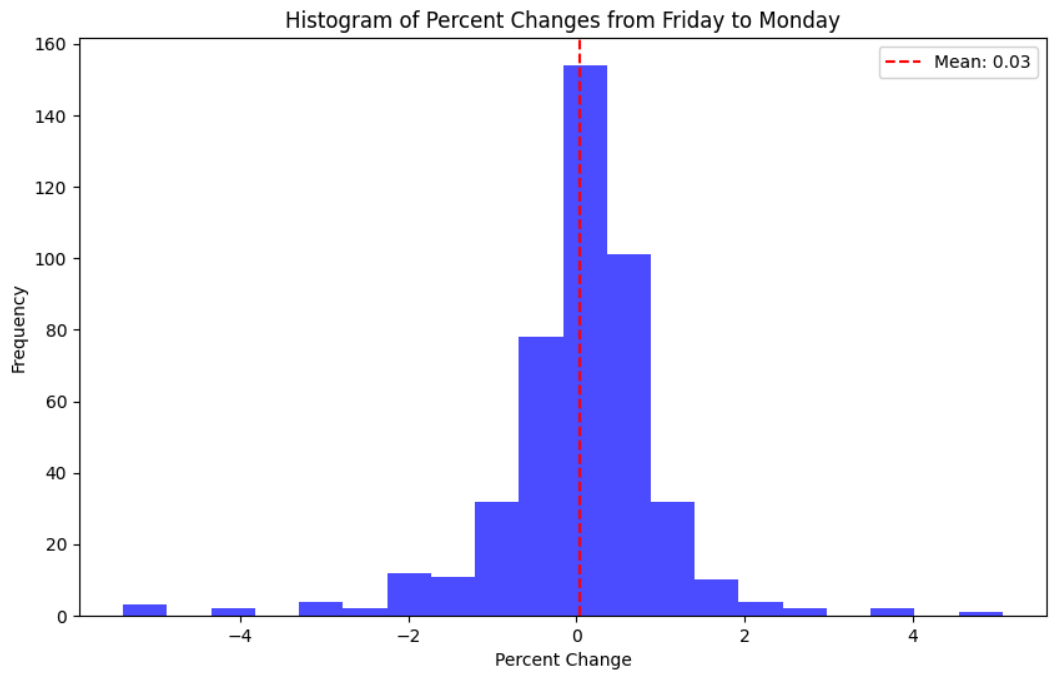
```
Date,Close/Last,Open,High,Low
03/28/2024,5254.35,5248.03,5264.85,5245.82
03/27/2024,5248.49,5226.31,5249.26,5213.92
03/26/2024,5203.58,5228.85,5235.16,5203.42
03/25/2024,5218.19,5219.52,5229.09,5216.09
03/22/2024,5234.18,5242.48,5246.09,5229.87
```

Again, we'll compute the value in the same way, averaging the high and low values, then dropping holidays which have a value of 0. We'll perform the same two tests, and also count the number of negative changes from Friday to Monday out of all the total changes.

Here's the result of the analysis:
u-test p-value: 0.9795398800077821
Wilcoxon p-value: 0.0015292442399880246
Number of Negative Changes: 175
Total Changes: 450
Average monday value: 3072.0334888888883
Average friday value: 3070.8543555555557

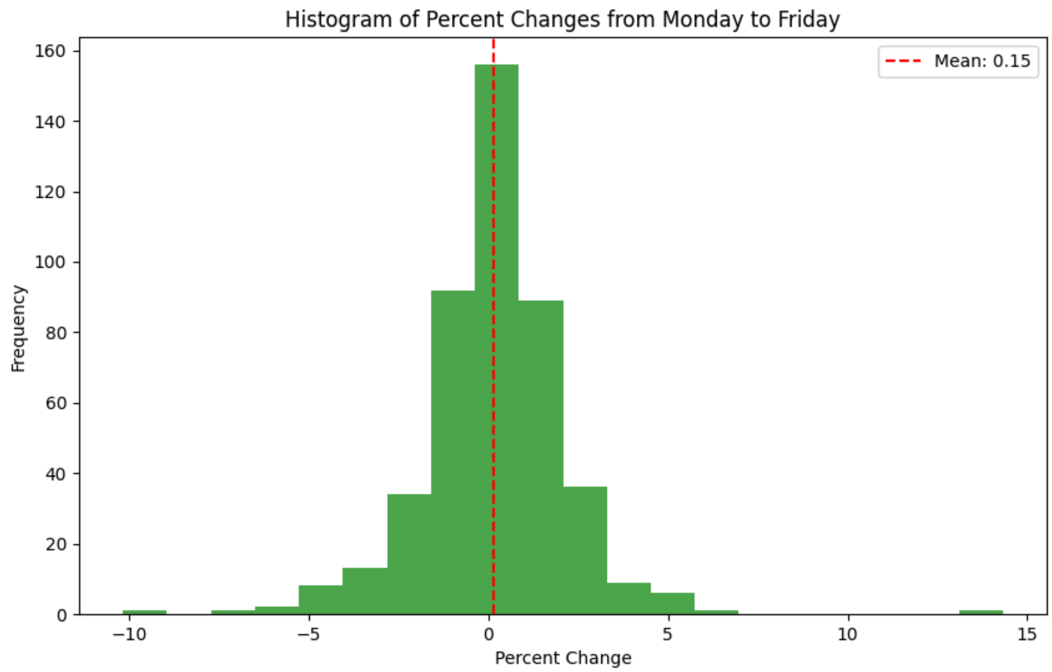
It looks like the Wilcoxon test gave a statistically significant result indicating one of the pairs is greater than the other. Looking at the changes and average values, it seems, surprisingly, that the stocks in the index have a higher price on Monday than their preceding Fridays.

A histogram of the percent changes also backs up this finding:

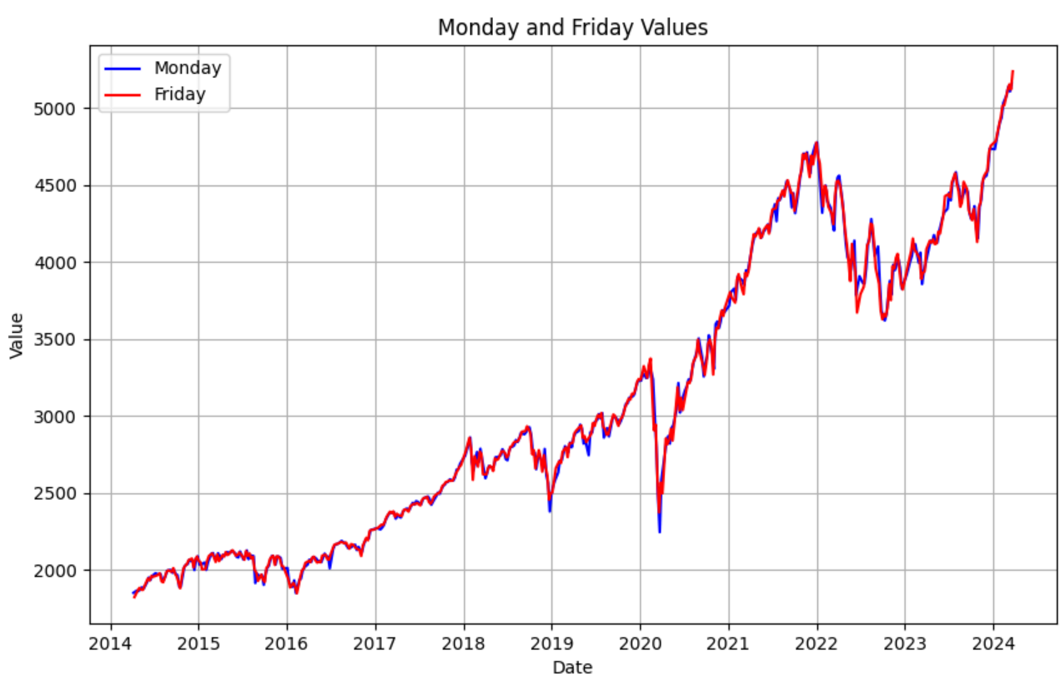


This could be simply due to the well-known fact that the overall market tends to grow in value over time. Since the Fridays are paired with a Monday that is one trading day ahead, the Mondays may be higher simply because they are later in time. It's also possible there is some kind of "reverse-weekend" effect at play [1].

In order to see if the overall market growth trend is primarily responsible for the differences, let's pair Mondays with the Friday that succeeds it, i.e. the Friday later in the week and see if it's Mondays or Fridays that tend to be higher. Here are the results:
u-test p-value: 0.9435875601615099
Wilcoxon p-value: 0.011557675144095662
Number of Negative Changes (from Monday to Friday): 194
Total Changes: 449
Average monday value: 3068.4192427616927
Average friday value: 3072.875345211581



The Wilcoxon test suggests another statistically significant result, and looking at the mean and number of negative changes, it seems that Fridays are now shown to be higher in price. The mean percent change in value from Friday to Monday was 0.03 (over a period of 1 trading day) and the mean percent change from Monday to Friday was 0.15 (over a period of 4 trading days), which suggests the overall change per day is very similar in either case. Looking at a graph of the Monday and Friday values further confirms that both values are similar:



If either Mondays or Fridays were intrinsically worse days for the market, we would expect to see one line consistently below the other, but that doesn't seem to be the case here. We can conclude that the overall trend of market growth over time far outweighs any possible price fluctuations caused by Fridays or Mondays.

S&P Daily Growth

Let's look at our initial question again. When we say "Fridays perform better than Mondays", perhaps we're not talking about the value change between Friday and Monday, but rather that Fridays experience more positive growth compared to the last trading day than Mondays. That means we'll calculate the percent growth in stock value from Thursday to Friday and the same for Friday to Monday and call that our daily growth for Friday and our daily growth for Monday. More generally, we use the equation:

$$Daily\_growth\_percent \;=\; \frac{(Value(Current\_Trading\_Day) - Value(Last\_Trading\_Day))}{Value(Last\_Trading\_Day)} * 100$$

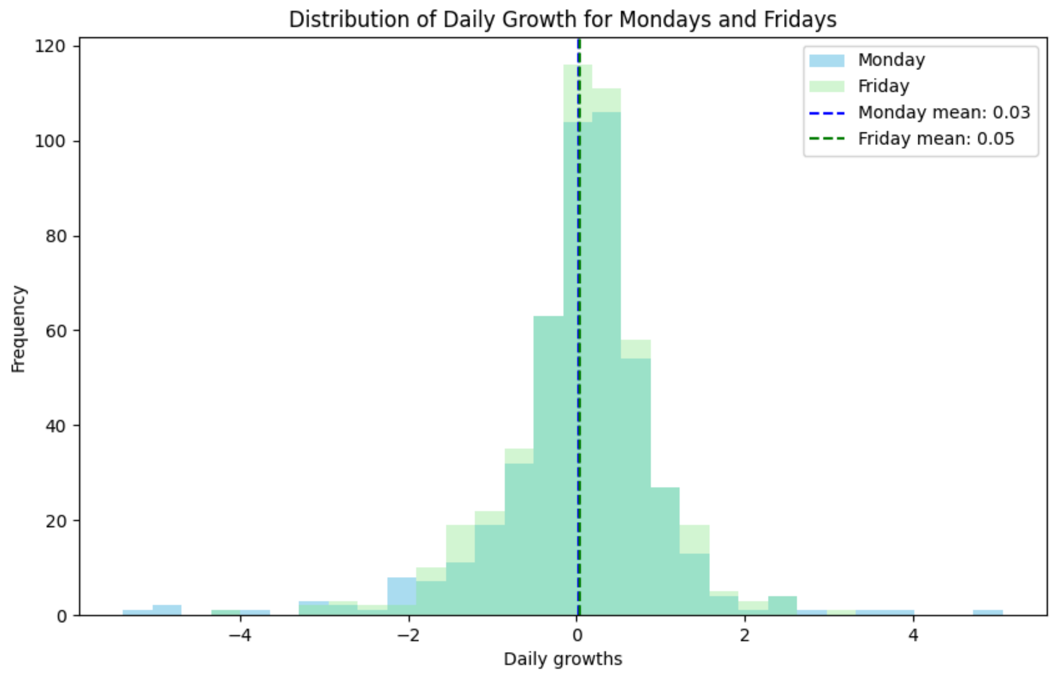
Assuming the stock market tends to grow at a relatively constant rate, this should reveal differences between growth on Mondays and Fridays without the ordering of the days in the pair affecting the analysis too much. We'll perform the same analysis with the Mann-Whitney U-test and the paired Wilcoxon test, but using the percent growths on each day instead of values.

Here's the results for when Mondays are paired with the previous Friday:
u-test p-value: 0.5983505973365202
Wilcoxon p-value: 0.40115220168040056
Number of decreased growths: 237
Total pairs: 450
Average monday growth percent: 0.0339777777777778
Average friday growth percent: 0.04051111111111112

And for when Mondays are paired with the next Friday:
u-test p-value: 0.6170947414278696
Wilcoxon p-value: 0.4542910521380904
Number of decreased growths: 229
Total pairs: 449
Average monday growth percent: 0.02463251670378619
Average friday growth percent: 0.030155902004454352

None of the tests gave a statistically significant result. Although it looks like Mondays did tend to grow a bit less than Fridays, they have very similar growths overall.

A histogram should illustrate the similarity between Mondays and Fridays:



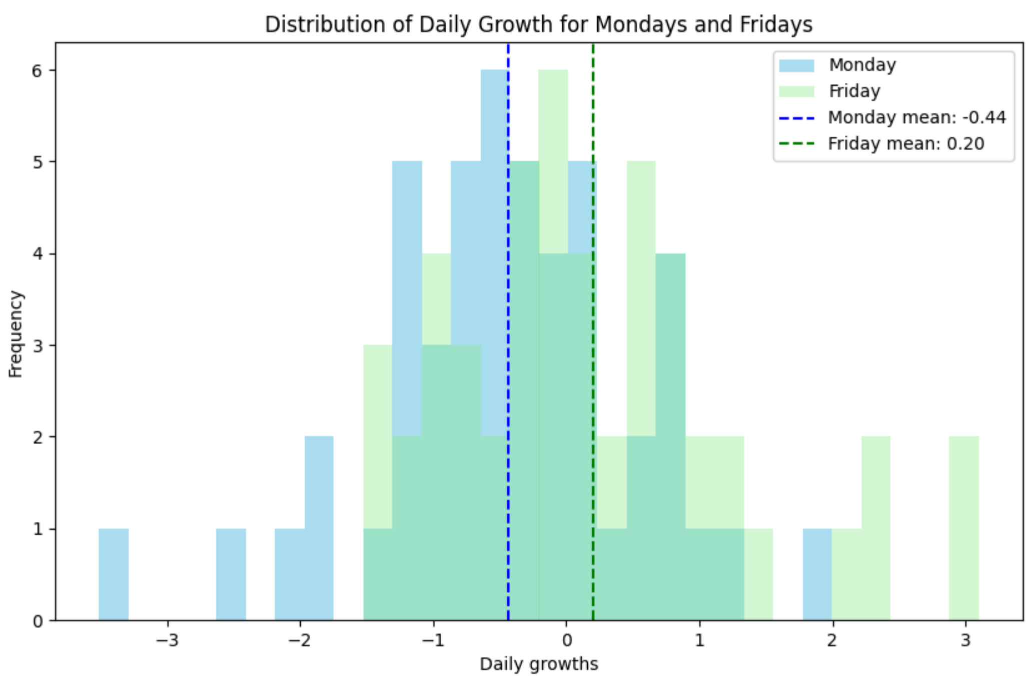
We can conclude that there's no statistically significant difference in growths on Monday and Friday even if the data seems slightly biased towards Fridays being the higher-growth day.

S&P in the Past

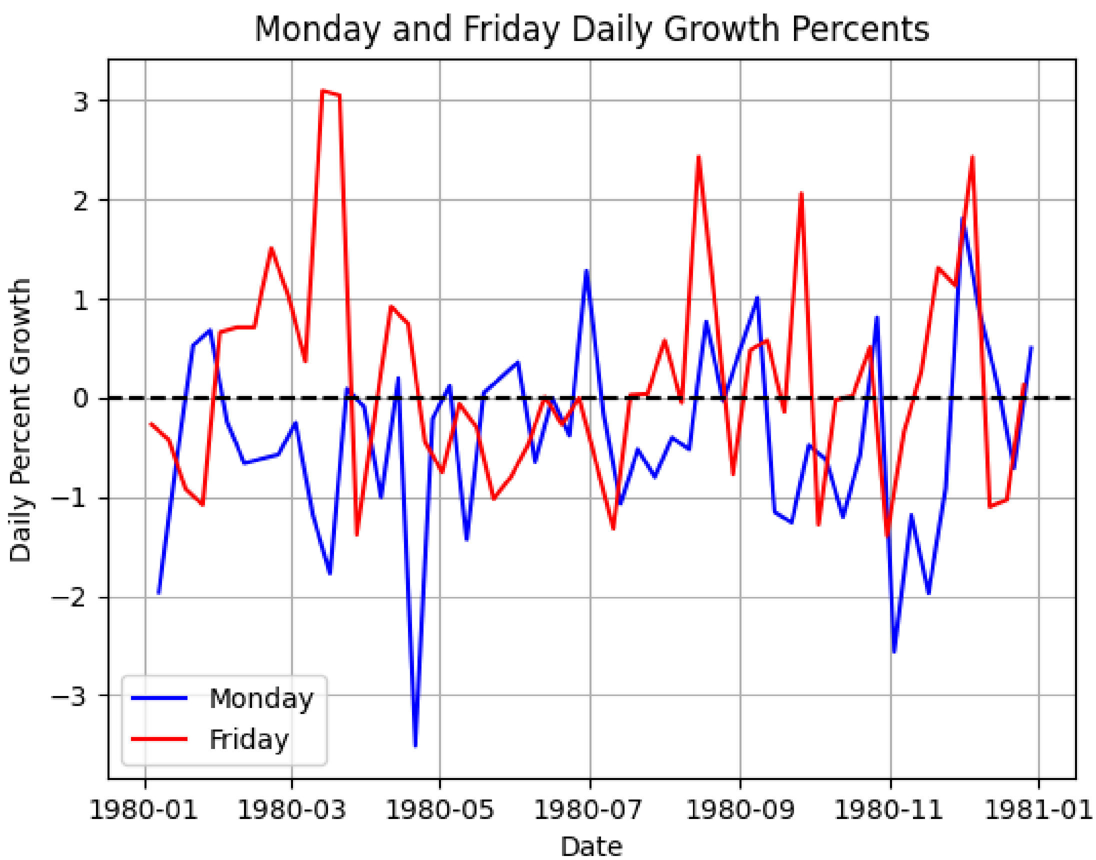
Perhaps there was a Monday-Friday effect that was more prominent in the past but has diminished over time as more people learned about it, causing the market to self-correct, or automated trading has caused the effect to decline [4]. We can use another S&P dataset that has data about years far in the past [5] (but only allows downloading datasets for a single year) and perform the previous analyses for various years to see if anything stands out. Considering that the percent changes in value analysis had problems dealing with the constant growth in the stock market, we'll only run the daily growth analysis. This kind of analysis was also not sensitive to whether it's the Friday or Monday that's first in the pair, so we'll just run it with Fridays before Mondays. We'll start with the years 1980, 1990, and 1997.
Of all these years, only 1980 showed statistically significant growth differences:

1980 with Fridays coming before Mondays in the pair:
u-test p-value: 0.009809503649952539
Wilcoxon p-value: 0.0035601492439525373
Number of decreased growths: 31
Total pairs: 47
Average monday growth percent: -0.4325531914893617
Average friday growth percent: 0.23638297872340427

In this case, Fridays showed statistically significant increased growth compared to Mondays.
Here's the corresponding histogram:



However, we should also look at the daily growths on Fridays and Mondays over the year to see the bigger picture:



Although Fridays do seem to have generally better growth than Mondays for this year, it seems like that may be a result of general market volatility and chance.We can also look at 1979 and 1981 and see if those years produce a statistically significant result. Running the analysis on both years didn't show any statistically significant result, so we can say with good confidence that 1980 is probably a statistical anomaly and that years in the past likely don't exhibit a clear bias towards increased growth on Fridays.

Monday-Friday Summary

Having run various analyses, there are a couple of conclusions we can make:

- The difference between stock values on Mondays and Fridays is primarily influenced by the constant growth in the market rather than any possible intrinsic difference between the days
- Although it's possible that stocks may seem to grow more on Fridays than Mondays (especially when looking at years farther in the past which had more volatility), this effect is typically **not statistically significant**
- Subsequently, trading strategies that seek to exploit any perceived gap between stock performance between Monday and Friday will not be effective

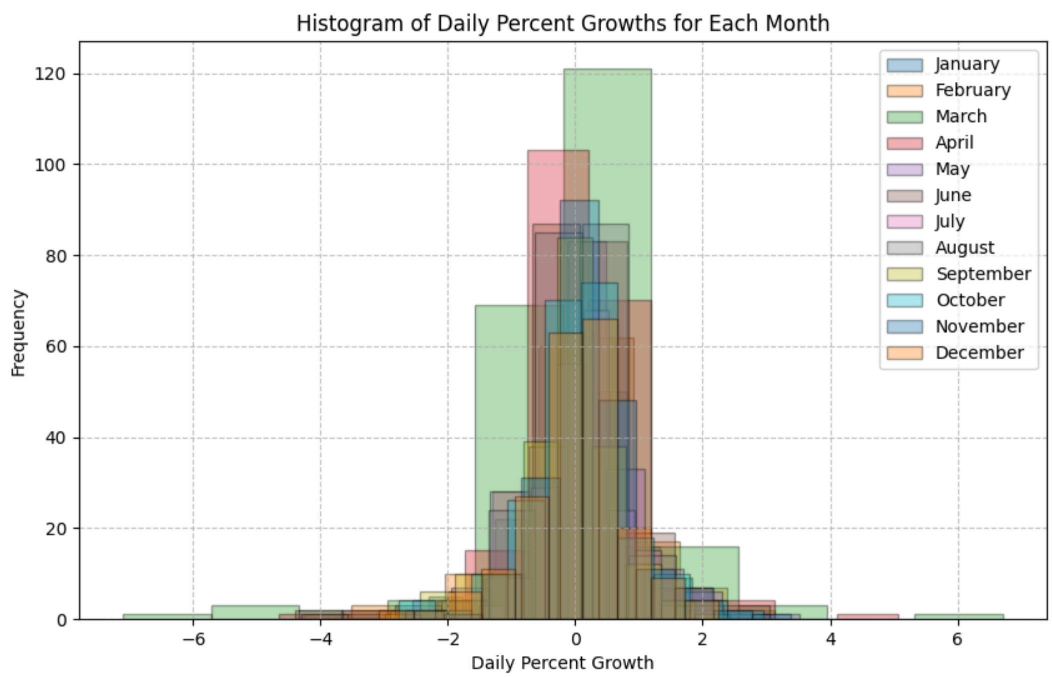
Differences Between Months

Certain months may experience more daily growth than others. In particular, September is known to be a particularly bad month for the market [4]. We can use the daily growth percentages that we generated from the Monday-Friday analysis as the basis for comparing months.

After grouping the months in ‘months.py’, we’ll run an ANOVA test to determine if there are multiple different means.

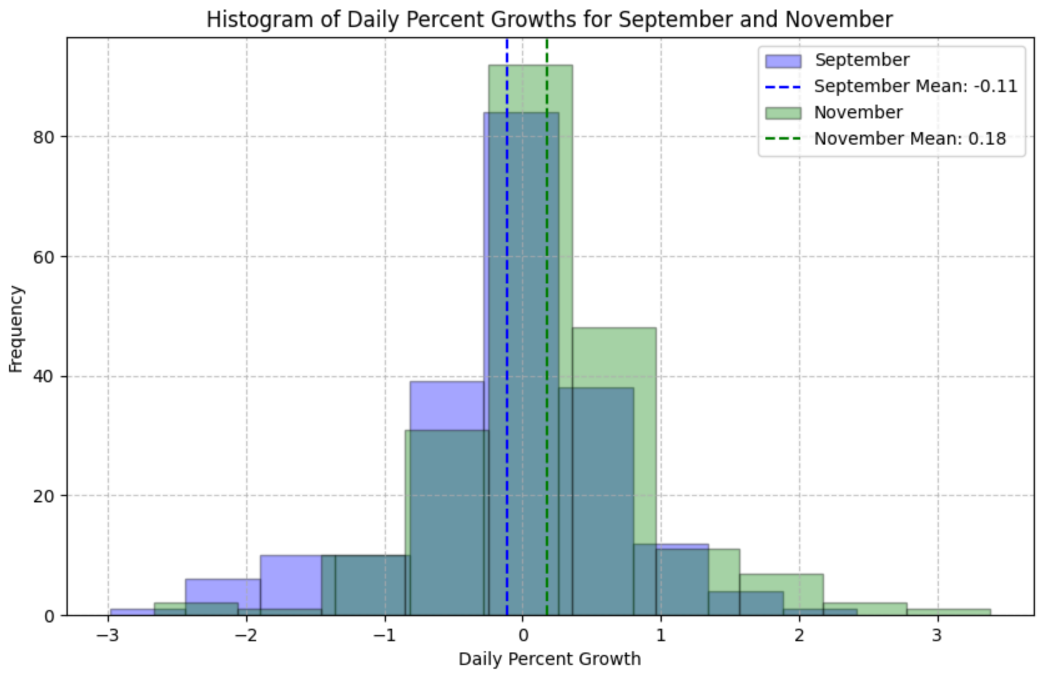
ANOVA p-value: 0.13882959156130908

We didn’t get a statistically significant result, and a histogram also seems to suggest that monthly daily-growth differences are not so pronounced:



However, let’s try running a post-hoc Tukey test just to see what we get. We don’t see a statistically significant difference between months except between November and September with a mean difference of -0.2897 and p-value of 0.0357. Looking into it, it just so happens that November is considered a particularly good month for stocks [6].

A histogram between these two months lets us see the difference more clearly:



It seems that the September growths are slightly left-skewed while the November growths are slightly right-skewed. This further confirms that the difference in daily growth between these months can be considered **statistically significant**. Therefore we can conclude that a reasonably good trading strategy would be to buy stocks during September when prices tend to fall and sell during November when prices tend to rise.

Beginning of the Month

The beginning of the month may see increased growth in stocks [4]. A possible cause of this date-base phenomena may be that mutual funds tend to receive fresh funding around the start of the month. When this funding is used to purchase stocks, this tends to increase demand, resulting in stock price growth. We will analyze this effect by comparing the daily percent growth on the first trading day of the month to the growths on the other trading days of the month.

ETL Process

For our initial dataset we will use the daily percent changes previously computed in the Monday-Friday analysis and stored in snpDataPercent.csv. We are primarily concerned with separating the first trading day of the month. This is the first day of the month excluding any weekends or holidays where trading doesn't occur. We follow this process for our ETL:

1. First we load our dataset and remove the unnecessary columns ("Day of Week" and "Value"). Additionally, data from April 2014 was also excluded due to the absence of growth data for the first trading day of the month.
2. Then, the first trading day of each month was identified and separated into its own DataFrame.
3. Similarly, the other days of the month were grouped together using the "***pd.dt.to_period('M')***" function to calculate the average daily growth percent for each month.
4. Finally, these two new DataFrames were merged to create the final DataFrame. The difference between the two metrics was calculated and stored in a new column.

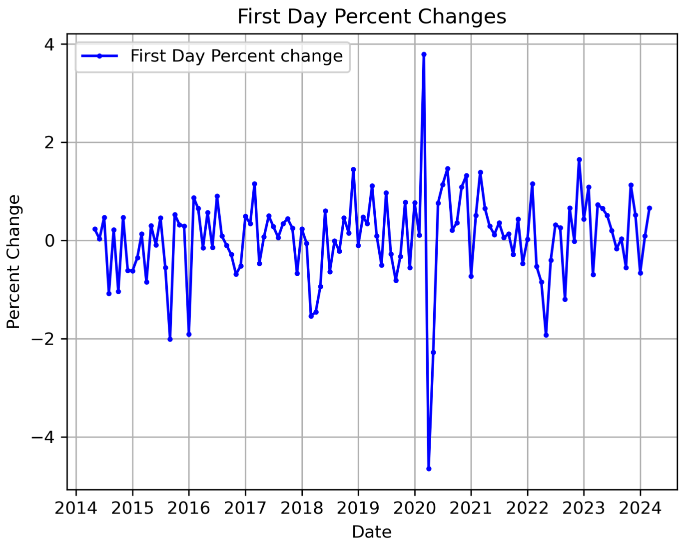
Our resulting CSV looks like this:

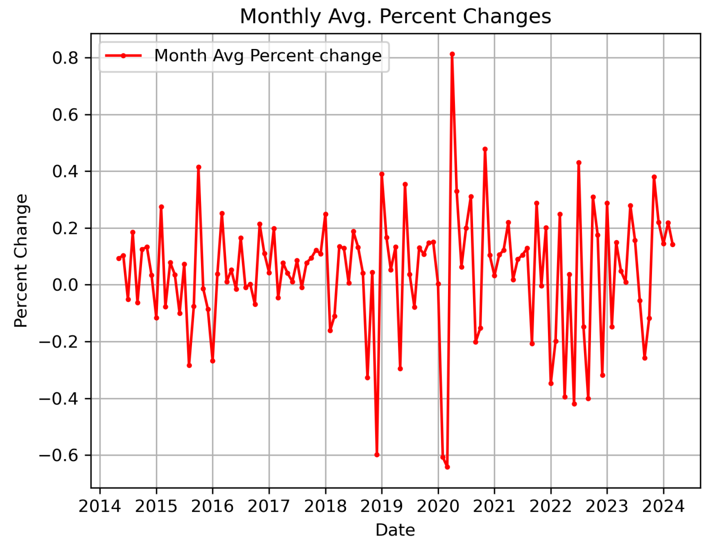
	Date	First Day Percent change	Month Avg Percent change	Difference
1	2014-05	0.23	0.09333333333333334	0.13666666666666666
2	2014-06	0.03	0.10350000000000001	-0.07350000000000001
3	2014-07	0.47	-0.05136363636363636	0.5213636363636364
4	2014-08	-1.08	0.185	-1.2650000000000001
5	2014-09	0.22	-0.06285714285714286	0.28285714285714286

Visualizing Data

Let's start by visualizing the data we produced to get a sense of any patterns that emerge. We'll plot both the percent growths on the first days of the month and the average for the other days of the month.

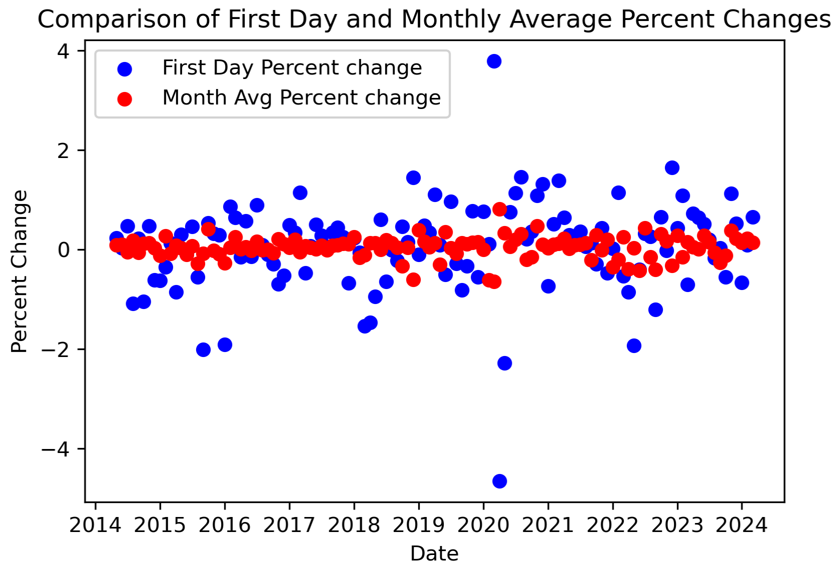
Below are the two graphs:





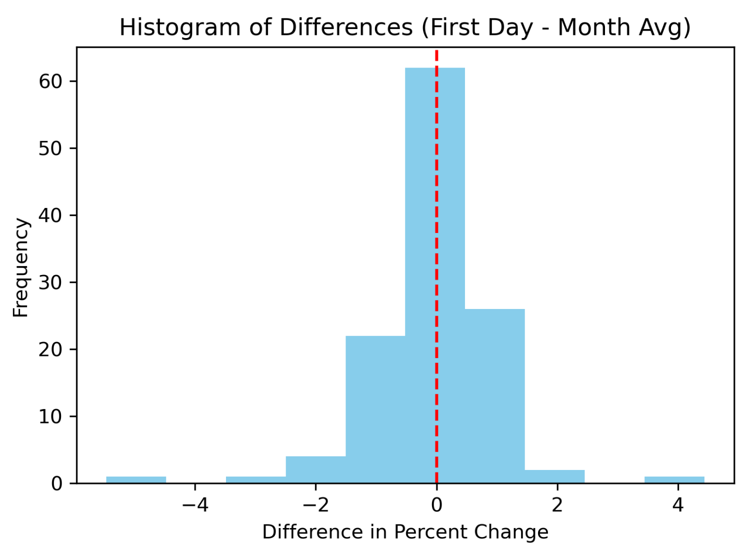
As we can see, the percent changes over the course of 10 years for the first day compared to the monthly average look fairly similar.

We also plot these two datasets on the same graph (in scatter plot format) to examine the central tendency of the data points. Note that the monthly averages tend to have less variance, since they are computed by taking an average.



From the graph above, we can observe that both datasets exhibit similar central tendencies.

Finally, we will also plot a graph of the differences between the growths (also referred to as percent changes) of the first day and the monthly average. This graph will help us visualize the differences between these two distributions and allow us to see if the central tendency of the difference deviates significantly from the null hypothesis of 0.



It appears the differences have a central tendency around 0, casting doubt on the idea that the first trading day of the month experiences additional growth.

Statistical Analysis

To conclude our analysis, we will perform several tests to verify what the graphs seem to be telling us – that the first trading day of the month does not experience a difference in growth from other days that is statistically significant.

Initially, we conducted normality and equal variance tests to determine if a t-test could be performed between these two datasets. It turns out that our data is neither normal nor does it have equal variance, as both tests resulted in a p-value < 0.05.

Given the non-normal distribution and unequal variance of our data, we proceeded with nonparametric tests for our analysis.

We performed two different non-parametric tests:

Mann-Whitney U-Test:
This test was used to assess whether the percent changes on the first trading day of the month are larger or smaller than the monthly averages. We obtained a p-value > 0.05, which prevents us from rejecting our null hypothesis that the two groups have the same median growth.

Wilcoxon Signed-Rank Test:
We can pair the growth on the first trading day of the month with the monthly average growths and apply a Wilcoxon test to see if one half of the pair experiences more growth than the other on average. This test also resulted in a p-value > 0.05, confirming that there is no significant difference between these two distributions.

These non-parametric tests lead us to conclude that there is no statistically significant difference in stock market behavior on the first trading day compared to the rest of the days of the month. This means that any trading strategy attempting to exploit any possible increased growth at the start of the month will likely be ineffective.

Here are all the results we obtained:

```
Mann-Whitney U-Test p-value:      0.08924407005728716
Wilcoxon Test p-value:            0.3478636904790874
Number of months of data:         119
Average First Day change percent: 0.06008403361344538
Average Other Days change percent:0.04397033263065991
```

Santa-Claus Rally

The "Santa-Claus Rally" phenomenon in the stock market refers to the supposed significant increase in stock prices during the last five trading days of December and the first two days of January likely caused by holiday optimism and increased consumerism[8]. We will test the statistical significance of this date-based phenomenon by considering periods that account for two weeks of trades. For Santa-Claus Rallies, we will construct two-week periods by taking the day one week before Christmas to be the start of the periods, accounting for 5 trading days, then taking the day 4 trading days after Christmas to be the end. This would comprise 10 trading days (2 weeks worth of trading days) if Christmas was a trading day. However, Christmas is not a trading day, but we can still consider this to be worth 2 weeks of trades because the underlying businesses tied to the stock values increase in value over time even if no trades are occuring. To model “normal” two-week periods we will randomly select two-week periods consisting of 10 trading days that are not during December. To analyze growth during our periods we will find the day within each period that has the maximal value and consider this to be our period peak. Finally we calculate the percent change in value from the start of the period to the peak day in the period and compare the start to peak changes for the two groups.

ETL Process

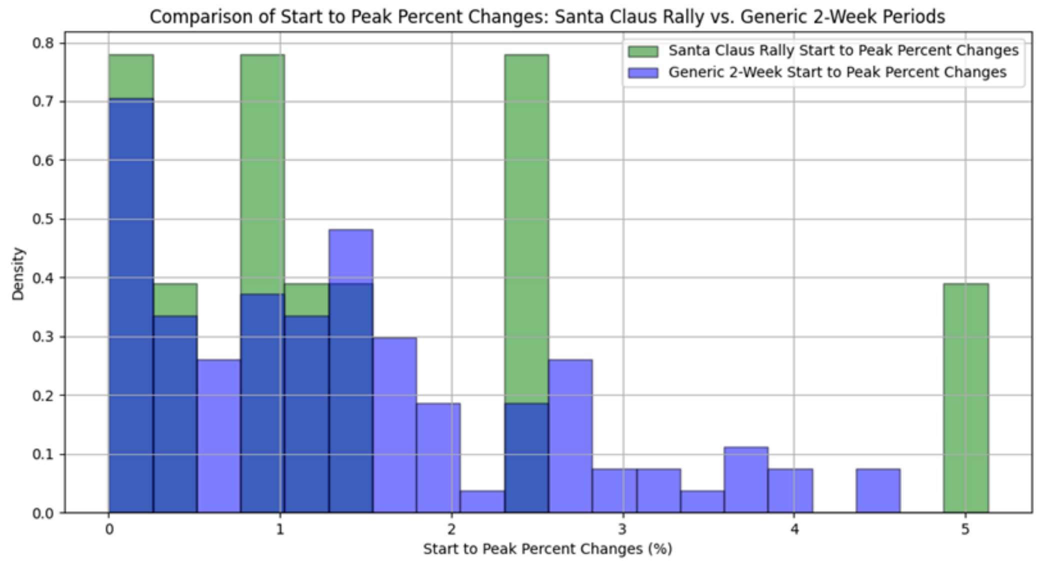
1. Using the same S&P dataset from the Monday-Friday analysis, we run “extract_santa_data.py” to isolate the relevant entries for the 2-week Santa-Claus Rally periods

2. We calculate the values for each day with “condensed_santa.py”

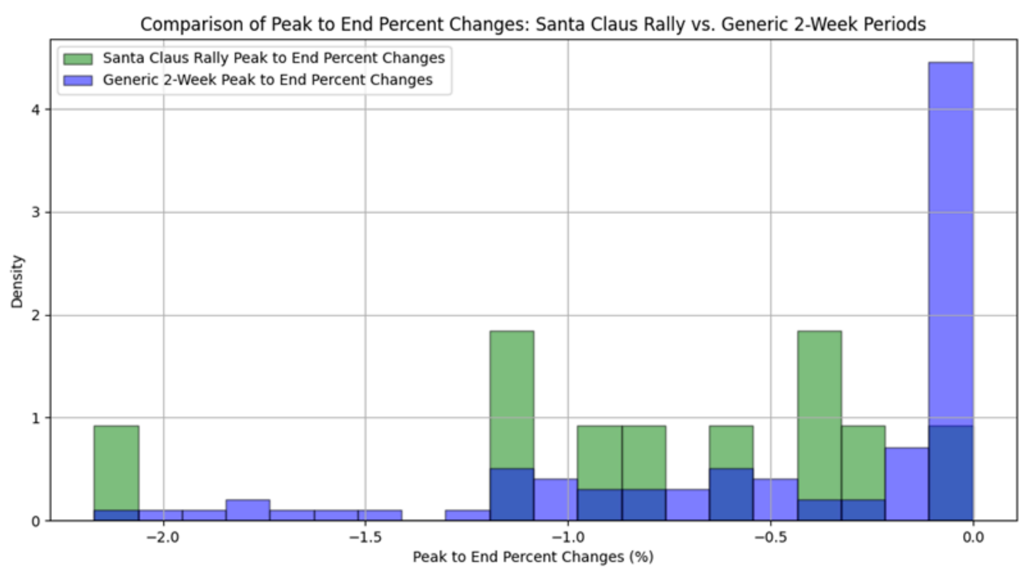
3. We use “compute_santa_peak.py” to group the condensed data into periods and calculate the start and end dates and values, and compute the start to peak change, along with the peak to end change.

4. “generic_two_week.py” performs identical calculations for generic two-week periods throughout the year, capturing the same set of data and metrics. This information is saved into another CSV file, "generic2week_snp_data_percent.csv".

5. “visualize.py” generates plots for ‘start to peak percent changes’ and ‘peak to end percent changes’ of both Santa Claus Rallies and other random 2-week periods.



As we can see from the plot, the Santa-Claus Rally period seems to tend towards higher start-to-peak change than generic 2-week periods, but not always. We can also say that it looks like rally periods may be more volatile, as evidenced by the greater spread of the peak to percent changes for Santa-Claus Rallies, although this could simply be because we can only take 10 Santa-Claus Rallies as samples while taking over a hundred generic 2-week samples.



Looking at the peak-to-end changes of the periods, again the rally periods seem to showcase a more pronounced peak (evidenced by a greater proportion of more extreme negative changes from peak to end) and greater variability.

Statistical Analysis

Considering the small sample size of the rally periods, we should run a statistical test to see if the rally periods are truly statistically significant. Using “statistical_comparison.py”, Mann-Whitney U-tests were conducted to compare percent changes from the start to peak and from peak to end of the rally versus the generic periods.

```
Mann-Whitney U-tests comparing the Santa Claus Rally data to the generic two-week periods:
Start to peak percent change: U-statistic = 510.0, P-value = 0.7072100036645202
Peak to end percent change: U-statistic = 472.0, P-value = 0.45454960787594223
```

The analysis shows high p-values in both cases, indicating **no statistically significant difference** between the peaks in the rally period and the peaks in any generic 2-week period. This suggests that the observed fluctuations during the Santa Claus Rally periods could be attributed to normal market variability rather than any specific seasonal effect.

Conclusions

We examined various date-based anomalies in the stock market with the results of our analysis summarized below. It should be noted that statistical significance does not guarantee success when using a strategy, as many other factors, primarily the perceived profitability of a business, play a much greater role in its stock price. If statistically significant, strategies should be interpreted more as a possible way to increase the odds of a successful trade by using them in conjunction with a well-proven strategy such as value investing rather than as an isolated strategy.

Effect	**Associated trading strategy**	**Statistically significant**
Monday-Friday effect	Buy on Mondays, sell on Fridays	No
September-November difference	Buy during September, sell during November	Yes
Beginning of the Month	Buy before the start of the month, sell shortly after the start of the month	No
Santa-Claus rally	Buy a week before Christmas, sell when prices increase around Christmas	No

Limitations and Extensions

Overall, our biggest challenge was determining the best way to do our analysis to effectively isolate the effects we were looking for. Swapping big data for the S&P index and switching from value-based analysis to growth-based analysis were crucial in this process and it would have saved a lot of time to have realized the limitations of our initial approaches earlier.
Possible extensions on our current analysis include:

- The Monday-Friday analysis could also consider differences in growth between other days of the week, potentially using an ANOVA test followed up by a Tukey test
- The Santa-Claus Rally analysis could also be applied to the period around Thanksgiving, which is also known to be a strong period for the stock market [7]
- To better account for smaller companies than simply the top 500, we could use different stock indexes or an average of various stock indexes as a measure of the overall market instead of the S&P 500
- Besides analyzing differences in stock growth, we could also analyze differences in stock price volatility. For example, is there a month that is significantly more volatile than others?

Additional related topics could include:

- Comparing how a given date-based trading strategy would fare against a standard trading strategy such as long-term investment in the S&P index over a period of 10 years by simulating and comparing the profit made from both investment strategies on an initial investment of $1000
- Catastrophic events like natural disasters, the COVID-19 pandemic, and terrorist attacks have been observed to have an effect on the stock market (It's even possible to spot increased volatility around 2020 due to the COVID-19 pandemic in some of our current graphs). It could be possible to statistically verify these effects by running a u-test between the daily growths during these periods and the daily growths in a "normal" period.
- Date-based anomalies may grow or shrink over time. It could be possible to calculate a correlation coefficient between time measured over decades and the strength of a phenomenon measured by its p-value.

Project Experience Summary

Alex:

- Identified various proposed date-related stock market anomalies and collected relevant datasets by engaging in informed research
- Actively coordinated, delegated, and explained research objectives and approaches to project collaborators through organizing frequent meetings and communication
- Aggregated and processed stock market data using a variety of approaches such as big-data, constructing paired data, and calculating differences in daily growth
- Identified statistical insignificance between the difference in Monday and Friday stock market growth by conducting nonparametric tests and performing data visualization
- Identified a statistically significant difference between stock market growth on September and November by applying a Tukey HSD test and performing data visualization
- Concisely communicated overall project objectives, scope, and conclusions in a written report

Divij:

- Conducted normality and variance tests to determine the suitability of statistical methods.
- Created visualizations including line graphs and histograms to illustrate and analyze trends.
- Executed non-parametric tests, such as Mann-Whitney U and Wilcoxon Signed-Rank, to analyze dataset differences.
- Identified statistical insignificance of stock growth on the first trading day of the month via applying nonparametric tests

Ege:

- Extracted and processed historical S&P 500 data for the designated Santa Claus Rally periods.
- Visualized data trends using histograms to compare the Santa Claus Rally to generic 2-week periods, focusing on start to peak and peak to end percent changes.
- Analyzed the Santa Claus Rally and generic two-week period data for statistical significance using Mann-Whitney U-tests

References (dataset sources highlighted):

[1] https://www.investopedia.com/terms/w/weekendeffect.asp
[2] https://www.kaggle.com/datasets/borismarjanovic/price-volume-data-for-all-us-stocks-etfs
[3] https://www.nasdaq.com/market-activity/index/spx/historical
[4] https://www.investopedia.com/day-trading/best-time-day-week-month-trade-stocks/
[5] https://www.marketwatch.com/investing/index/spx/download-data
[6] https://finance.yahoo.com/news/november-typically-best-month-stocks-213001829.html
[7]https://nasdaq.com/articles/feasting-on-profits:-does-thanksgiving-week-cook-up-stock-market-gains
[8] https://www.investopedia.com/terms/s/santaclauseffect.asp