

S&P 500 Analysis



Executive Summary

Summary of methodologies

- **Data collection** with Python yfinance module
- **Data wrangling** to transform data for better analysis
- **Exploratory data analysis (EDA)** using SQL and Pandas
- **Data visualization** with plots and charts

What I've found

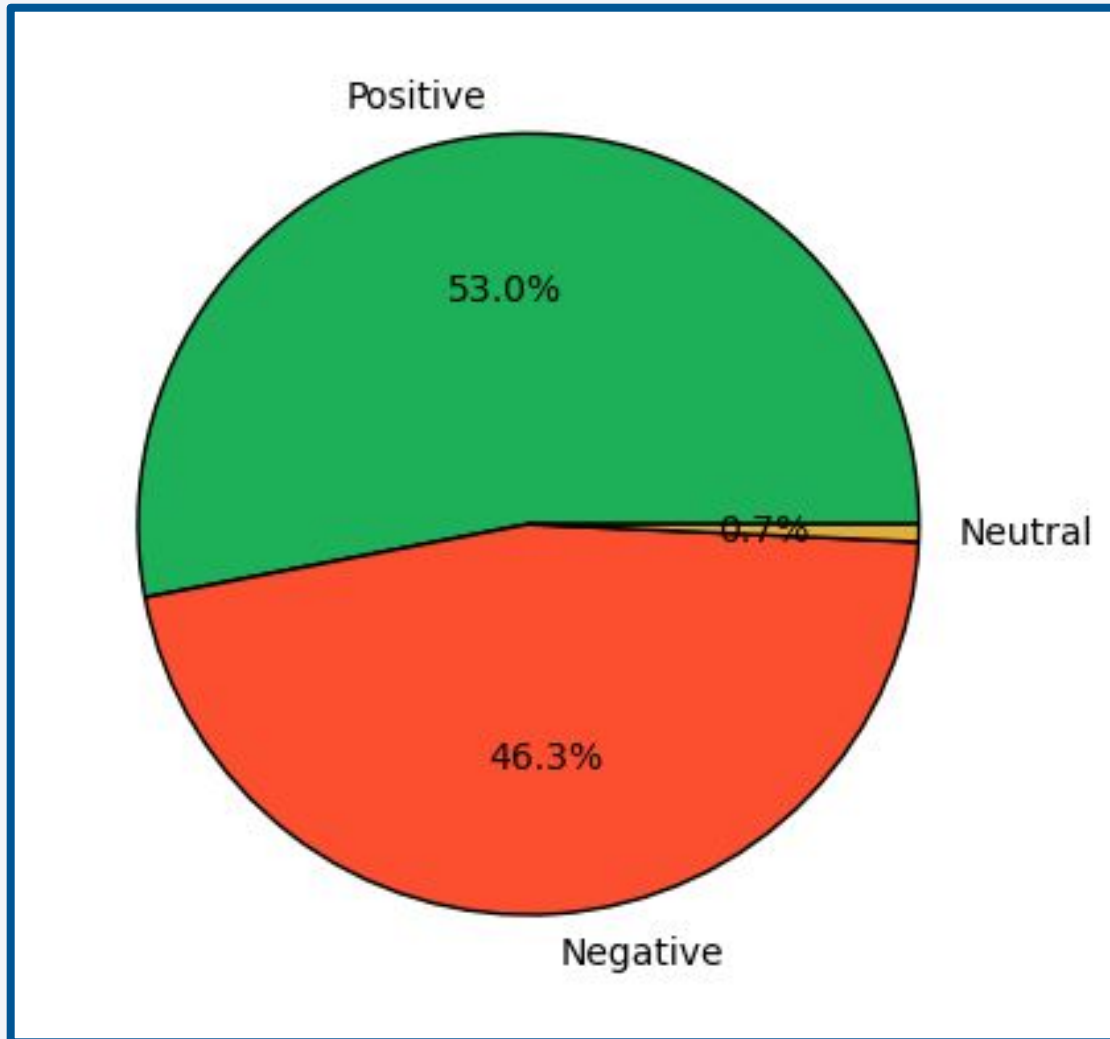
- Amount of positive and negative return days
- Day of the week that most times had a negative return
- Day of the week that most times had a positive return
- 20 worst days for the market since 2000
- Months that contained most negative return days
- Months that contained most positive return days

Introduction

Is monday really the worst day for the stock market? A lot of people say that it is. The goal of this project is to find that out, and if it really is the worse by how much.

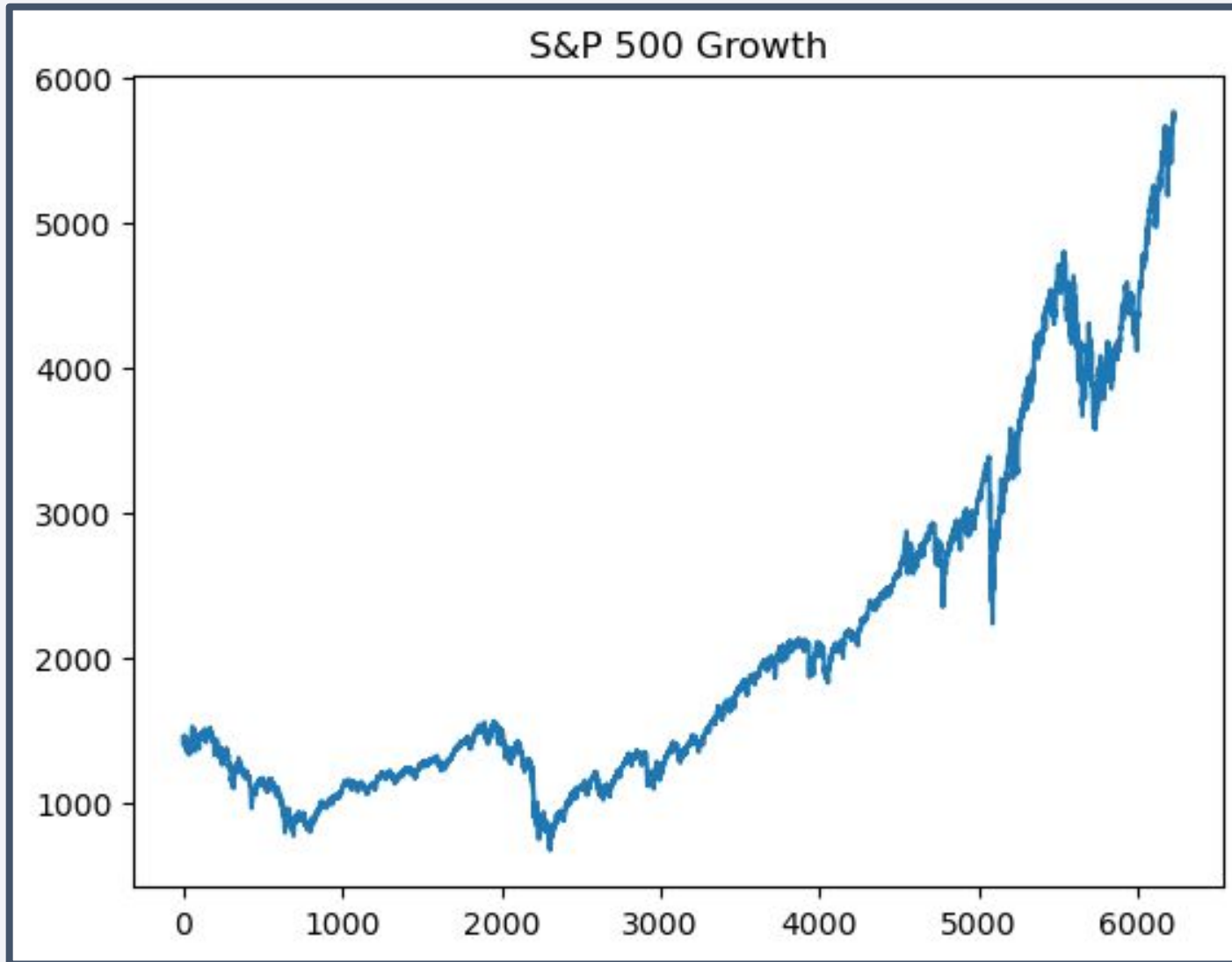
This research analyzed S&P 500 data, which is an index that tracks the 500 biggest companies in the US stock market. The data used in the project starts in the first trading day of the year 2000.

Positive, negative and neutral days



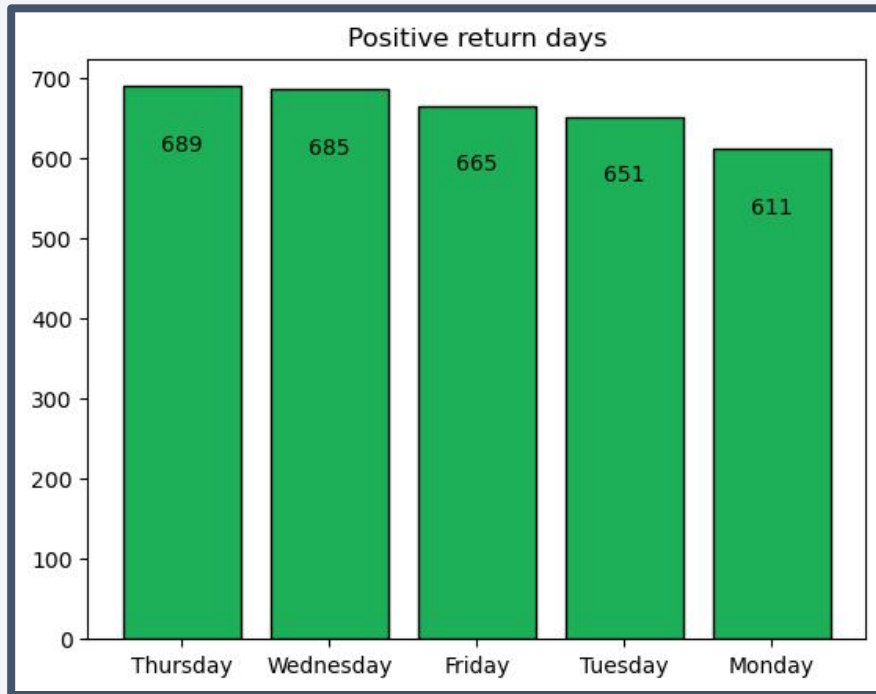
Out of the days observed 53% had positive returns, 46.3% had negative returns and 0.7% were neutral.

S&P 500 growth since 2000

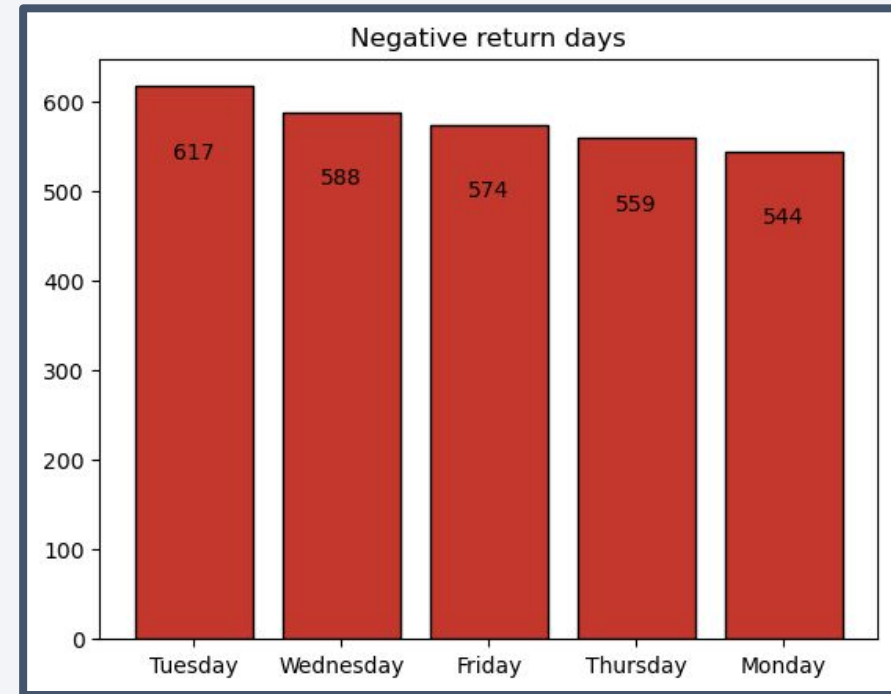


Even though 46.3% of the days had a negative return the index still performed really well, it grew almost 300%.

Count of positive and negative days (Day of the week)

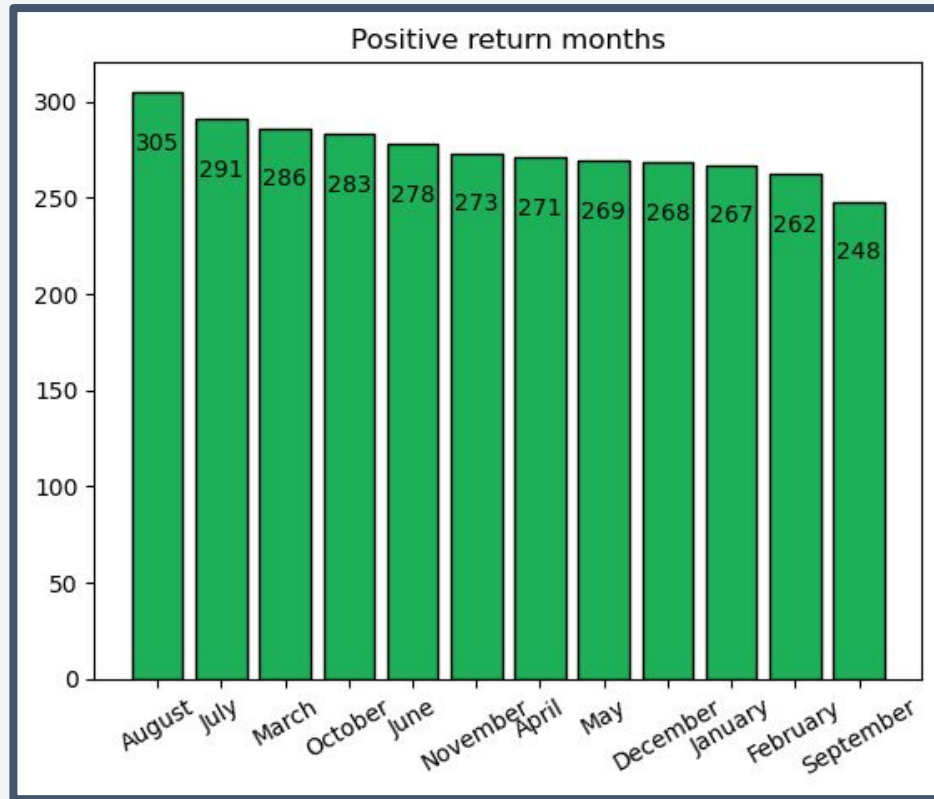


These are the days with more positive return occurrences, Thursday had most of the positive return days.

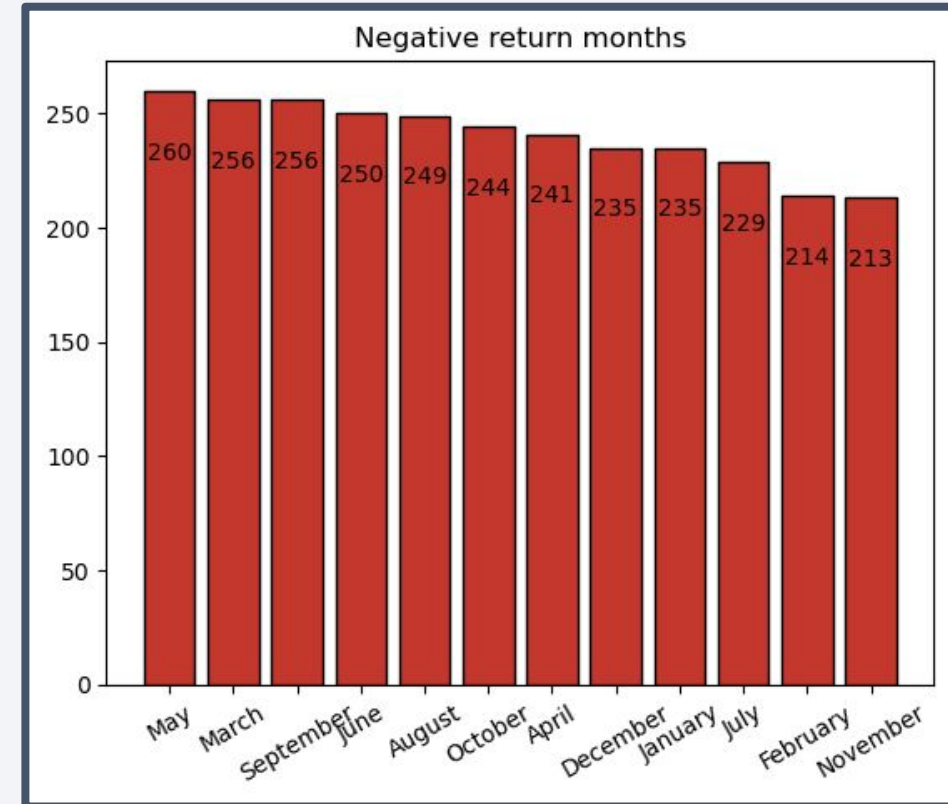


These are the days with more negative return occurrences, Tuesday had most of the negative return days. Monday actually had less occurrences than any other day.

Count of positive and negative days (Month)



These are the months with more positive return occurrences, August had most of the positive return days.



These are the months with more negative return occurrences, May had most of the negative return days.

Ranking the worst days for the market

Date	Return	Day	Month
2008-10-15 00:00:00-04:00	-8.72	Wednesday	October
2008-09-29 00:00:00-04:00	-8.49	Monday	September
2008-12-01 00:00:00-05:00	-8.15	Monday	December
2008-10-09 00:00:00-04:00	-7.94	Thursday	October
2008-11-20 00:00:00-05:00	-6.63	Thursday	November
2011-08-08 00:00:00-04:00	-6.59	Monday	August
2008-11-19 00:00:00-05:00	-6.11	Wednesday	November
2000-04-14 00:00:00-04:00	-5.83	Friday	April
2008-10-07 00:00:00-04:00	-5.80	Tuesday	October
2008-10-22 00:00:00-04:00	-5.77	Wednesday	October
2020-03-12 00:00:00-04:00	-5.71	Thursday	March
2009-01-20 00:00:00-05:00	-5.23	Tuesday	January
2020-03-20 00:00:00-04:00	-5.22	Friday	March
2008-11-06 00:00:00-05:00	-4.99	Thursday	November
2001-09-17 00:00:00-04:00	-4.92	Monday	September
2008-11-05 00:00:00-05:00	-4.90	Wednesday	November
2020-03-16 00:00:00-04:00	-4.88	Monday	March
2011-08-04 00:00:00-04:00	-4.77	Thursday	August
2008-09-15 00:00:00-04:00	-4.65	Monday	September
2008-11-12 00:00:00-05:00	-4.60	Wednesday	November

This is the list of the 20 days with the greatest percentage loss. Note that most of these days are from the 2008 financial crisis, the 2020 pandemic period is also present in the list. Between the greatest losses the day of the week with most occurrences is Monday.

Return of the worst days for the market

```
result = worst.groupby('Day')['Return'].sum()  
result.to_frame().sort_values(by='Return')
```

	Return
Day	
Monday	-37.68
Wednesday	-30.10
Thursday	-30.04
Friday	-11.05
Tuesday	-11.03

Adding the losses of the 20 worst days of the period we can see that in fact monday had greater losses.

Conclusion

- Although monday did not have the most occurrences of negative return days, among the 20 worst days it had the most occurrences and also the greatest percentage loss.
- Out of the days observed 53% had positive returns, 46.3% had negative returns and 0.7% were neutral.
- Thursday had most of the positive return days while Tuesday had most of the negative return days.
- August had most of the positive return days while May had most of the negative return days.
- The S&P 500 proved to be a great index for investments since it has grown since the year 2000.

Appendix

The notebook with all the project code can be found on Github:
https://github.com/awesomemegab/S-P500_Analysis

LinkedIn:

www.linkedin.com/in/gabriel-melo-72ba8524a

Gabriel Silva de Melo

Thank you!