S&P 500 Index prediction via sentiment analysis via Presidential tweets.

The Data

- The S&P 500 Index data encompassed a range of dates, starting Jan 1, 2017 (President Trump took office on the 20th) through Jan 1, 2019 on a daily frequency.

	High	Low	Open	Close	Volume	Adj Close
Date						
2018-12-31	2509.239990	2482.820068	2498.939941	2506.850098	3442870000	2506.850098
2018-12-28	2520.270020	2472.889893	2498.770020	2485.739990	3702620000	2485.739990
2018-12-27	2489.100098	2397.939941	2442.500000	2488.830078	4096610000	2488.830078
2018-12-26	2467.760010	2346.580078	2363.120117	2467.699951	4233990000	2467.699951

The Data

2018. There are exactly 252 **trading days in** 2018. February and September have the fewest (19), and August the most (23), with an average of 21 per month, or 63 per quarter. Out of a possible 365 **days**, 104 **days** are weekend **days** (Saturday and Sunday) when the stock exchanges are **closed**.

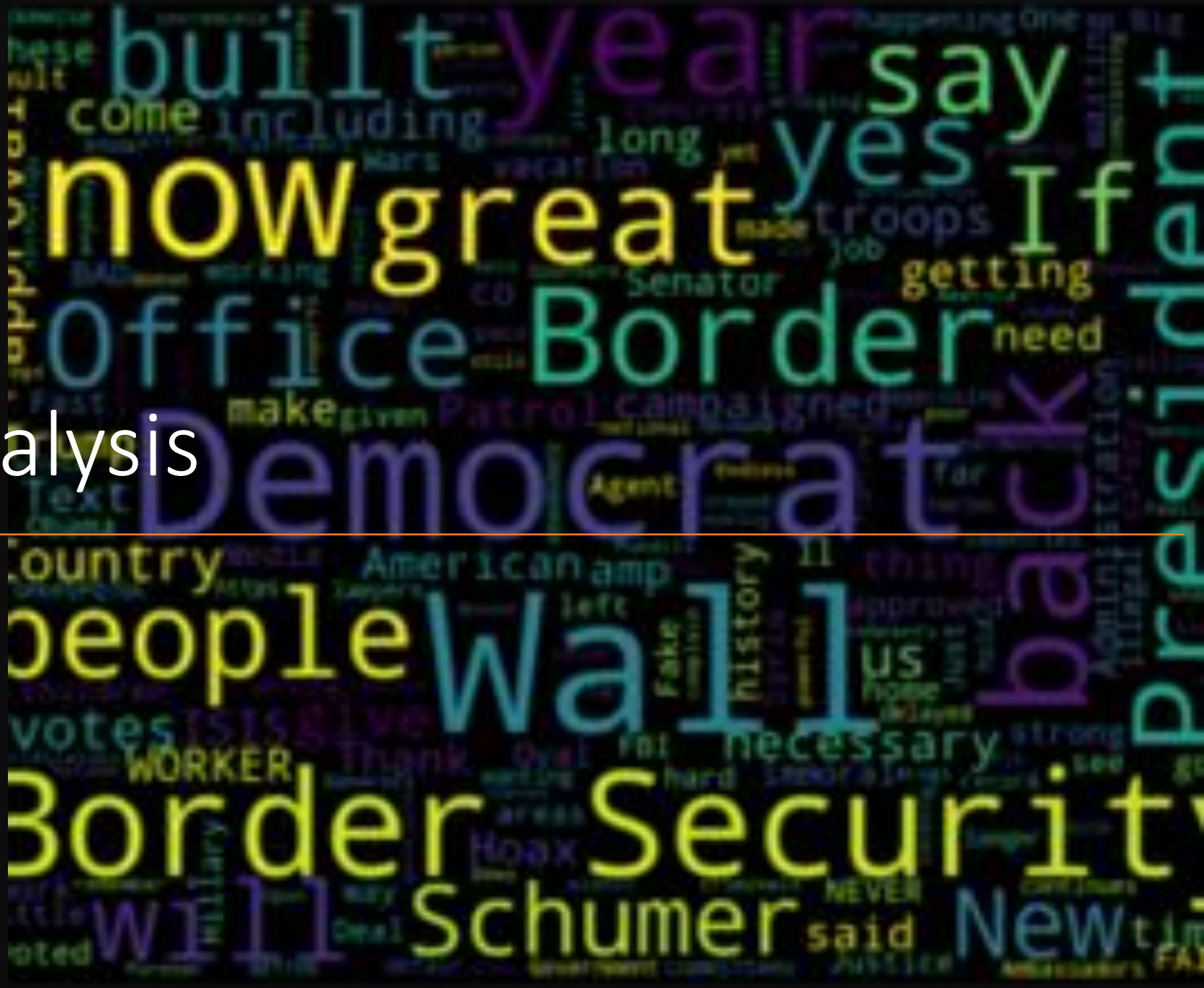
en.wikipedia.org › wiki › Trading_day

Trading day - Wikipedia

The Data

- Tweets were in the form of two json files.
- 2017 had 2605 tweets
- 2018 had 3510 tweets
- Retweets (RT) were removed and not being considered.
- Multiple tweet sentiment was averaged per day.

Sentiment analysis



How sentiment analysis was measured

- Polarity – measured how positive or negative a word is – 1 would be considered very negative. +1 would be considered very positive
    - ‘There is right now a full scale manhunt going on in California for an illegal immigrant accused of shooting and killing a police officer during a traffic stop. Time to get tough on Border Security. Build the Wall!’

Sentiment(polarity= -0.0875, subjectivity= 0.6047

How sentiment analysis was measured

- Subjectivity – measured how opinionated a word is a value of 0 would be considered a fact, whereas +1 considered very much opinionated
    - 'I hope everyone is having a great Christmas, then tomorrow it’s back to work in order to Make America Great Again (which is happening faster than anyone anticipated)!’

Sentiment (polarity= 0.6, subjectivity = 0.5)

Sentiment analysis data

```
1 Tweet_analysis.head()
```

	polarity	subjectivity
Date		
2017-01-01	0.812500	0.750000
2017-01-02	0.151407	0.602309
2017-01-03	-0.028750	0.487708
2017-01-04	-0.116617	0.499511
2017-01-05	-0.058434	0.437859

Bringing the data together

	polarity	subjectivity	Close
Date			
2018-12-20	0.140961	0.430085	0.005298
2018-12-19	0.154673	0.508193	0.021190
2018-12-18	0.223457	0.529954	0.033274
2018-12-17	0.268694	0.487194	0.029947
2018-12-14	0.287311	0.359091	0.043481

The future of stock forecasting and prediction isn't a new concept, with that said the S&P 500 has had unforeseen crashes within the last 50 years:

- 1973 | Conflict in the Middle East **-48.2 percent loss**, 70 months recovery
- 1980 | Stagflation **-27.1 percentage loss**, 3 months recovery
- 1987 | Black Monday **-33.5 percentage loss**, 20 months recovery
- 1990 | Gulf War **-19.9 percentage loss**, 4 month recovery
- 2000 | The Tech Bubble Bursts **-49.1 percentage loss**, 56 months recovery
- 2007 | Real Estate Goes Bust **-56.9 percentage loss**, 49 months recovery

Thank You!