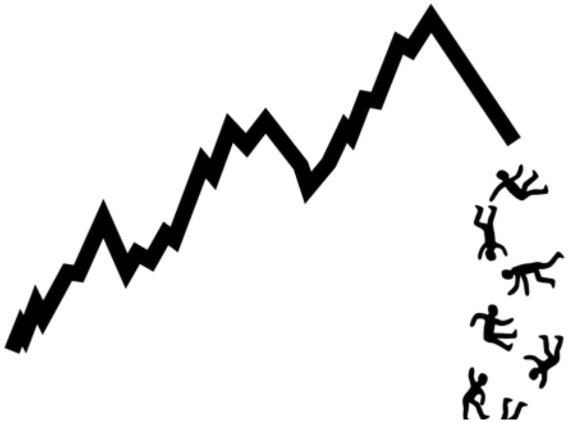
Bitcoin vs Average Stock Prices

John C, Will G, Jennie M, Jennifer P

Extract

- Using .csv files from Kaggle
- The first dataset for bitcoin was too large to work with
- Stock data (open, high, low, close, day, volume, and stock ticker name) from 2013-2018 for all companies found on the S&P 500 index
- Bitcoin closing price by date 2015-2020



Transform

- Aggregated the collective stock market prices to the average of all for each day
- This would line up with the simpler Bitcoin data set
- Renamed columns
- Merged the tables based on dates of the overlapping years (2015-2018)

In [1]:
```
import pandas as pd
```

In [2]:
```
stocks = pd.read_csv("data/all_stocks_5yr.csv")
```

In [3]:
```
stocks.head()
```

Out[3]:

	date	open	high	low	close	volume	Name
0	2013-02-08	15.07	15.12	14.63	14.75	8407500	AAL
1	2013-02-11	14.89	15.01	14.26	14.46	8882000	AAL
2	2013-02-12	14.45	14.51	14.10	14.27	8126000	AAL
3	2013-02-13	14.30	14.94	14.25	14.66		
4	2013-02-14	14.94	14.96	13.16	13.99		

In [4]:
```
grouped = stocks.groupby(by=['date']).mean()

grouped.head(20)
```

Out[4]:

	open	high	low	close	volume
date					
2013-02-08	58.068730	58.659286	57.802352	58.368654	4.643484e+06
2013-02-11	58.355556	58.686073	57.890072	58.272196	4.092442e+06
2013-02-12	58.321741	58.771200	57.954938	58.404071	5.122822e+06
2013-02-13	58.478904	58.899515	58.040928	58.485503	5.132086e+06
2013-02-14	58.289216	58.878329	57.871977	58.473144	5.48
2013-02-15	58.485982	58.932768	57.970143	58.437170	5.73
2013-02-19	58.475575	59.076565	58.088696	58.762733	5.27
2013-02-20	58.702446	59.020518	57.765783	57.915434	5.66
2013-02-21	57.765623	58.128038	57.007764	57.505777	6.04
2013-02-22	57.765128	58.342281	57.408570	58.088968	4.65
2013-02-25	58.332967	58.682924	56.980270	57.020570	5.76

In [9]:
```
close = grouped[['close', 'volume']]

close.head(20)
```

Out[9]:

	close	volume
date		
2013-02-08	58.368654	4.643484e+06
2013-02-11	58.272196	4.092442e+06
2013-02-12	58.404071	5.122822e+06
2013-02-13	58.485503	5.132086e+06
2013-02-14	58.473144	5.482383e+06
2013-02-15	58.437170	5.734999e+06
2013-02-19	58.762733	5.274365e+06
2013-02-20	57.915434	5.667893e+06
2013-02-21	57.505777	6.042141e+06

In [10]:
```
close.to_csv('data/cleaned_stocks.csv')
```

Load

- Created the table through SQL connection engine with SQL syntax
- Uploaded the merged table to postgres using sqlalchemy
- Having more structure data we used a more structured program

```
password = ""
```

```
connection_string = f"postgres:{password}@localhost:5432/Bitcoin-ETL"
engine = create_engine(f'postgresql://{connection_string}')
```

```
merged_table.to_sql(name='data', con=engine, if_exists='append', index=False)
```

The big idea…

To analyze that data based on date to summarize the correlation between the average stock price and Bitcoin price for each day

See if a trend could help predict the Bitcoin market and best buying and selling opportunities