Various Stock Market Analyses

Active Trading Strategy, Sector Portfolio Analysis, and Improved Stock Clustering

Dylan Alexander

Capstone Project for Springboard Intermediate Data Science in Python

Introduction

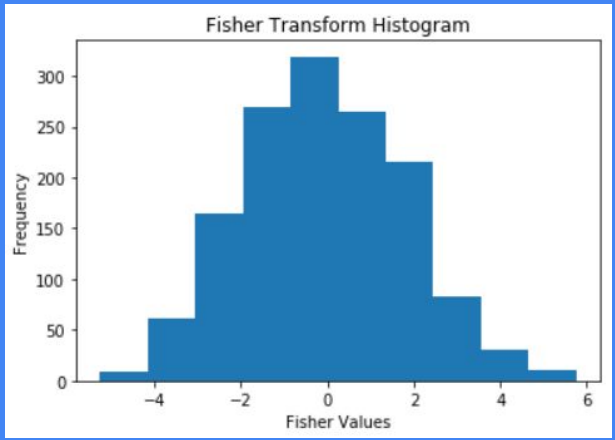
- Traditionally 11 sectors in S&P 500
- Wide range of strategies
- Data Science and algorithmic trading has taken the driver’s seat
- Covering 3 areas
  - Fisher Transform trading strategy for an active trader
  - Portfolio analysis for each sector
  - K-Means clustering to develop new clusters

Data Wrangling

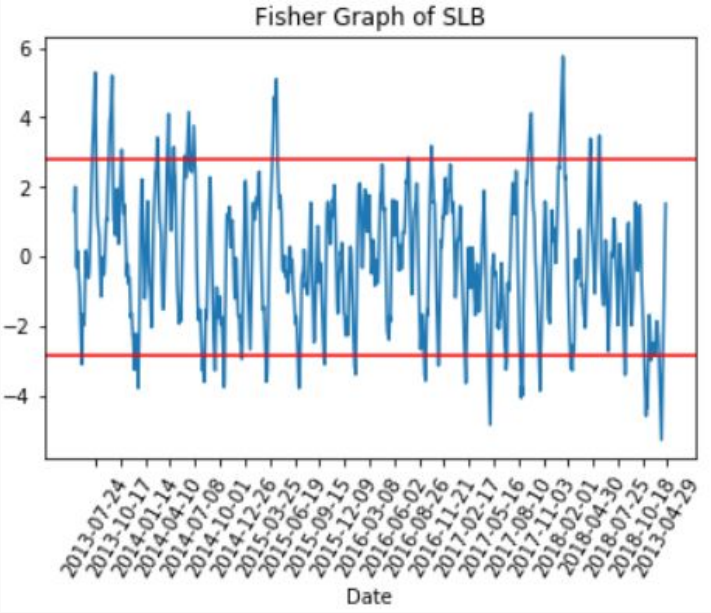
- Wrangled a total of 36 data sets
- One for prices and one for returns of 10 stocks for each sector
- From this, compiled prices and returns into two separate data sets containing all sectors
- Stocks selected based on highest volume on a random day
- Gathered data through the API on Tiingo.com
- Pivot(), concat(), pct_change(), and drop() functions

Fisher Transform Strategy

Created and introduced by John Ehlers in 2002, the Fisher Transform takes price as an input and "transforms" it into an approximately Gaussian distribution



- Schlumberger (SLB) stock price used in strategy test
- Strategy:
  - Buy if Fisher value < 1.5 standard deviations
  - Sell if Fisher value > 1.5 standard deviations
- 1.5 standard deviations = - 2.8146

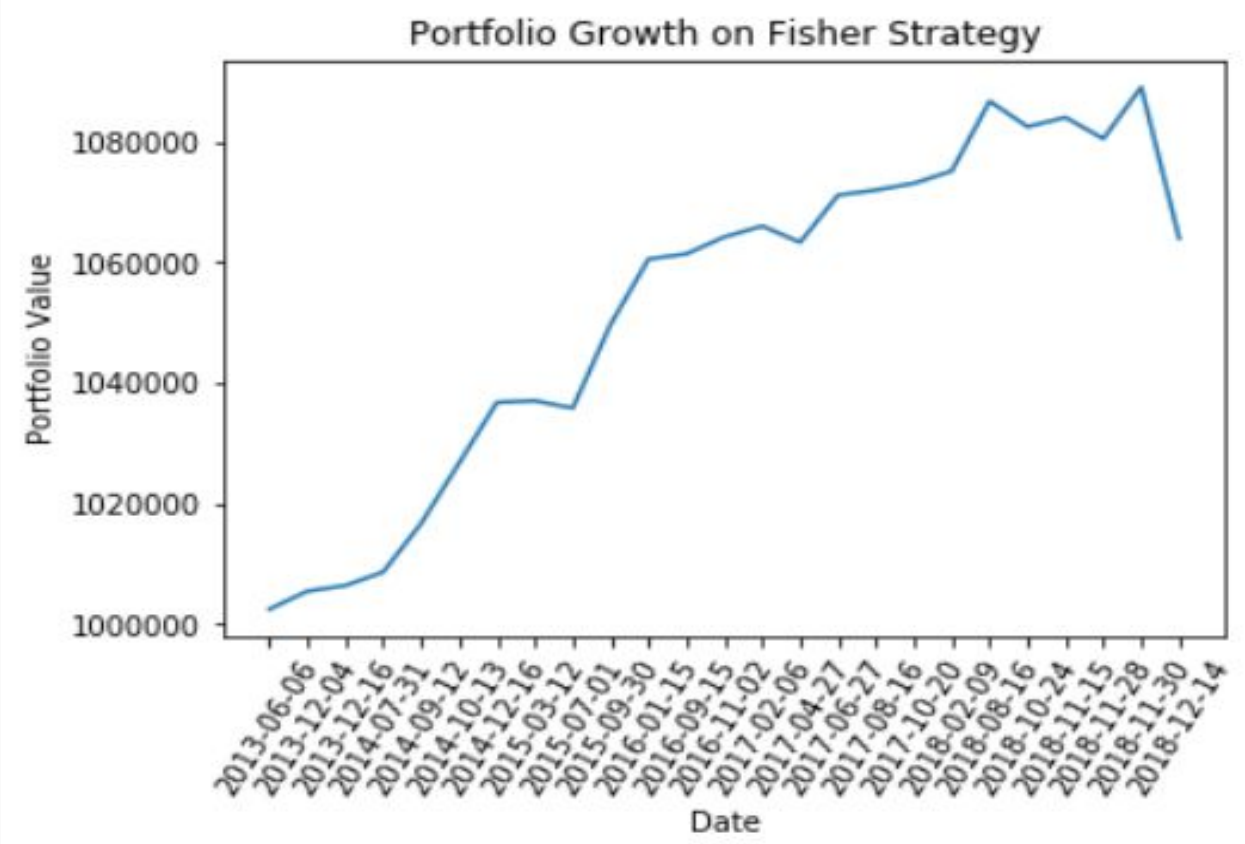


Fisher Transform Strategy

- DataFrame of outcome of each trade
- Mostly positive profit outcomes
- Cut off of data

	Start Port. Value	End Port. Value	End Date	Shares	Share Price	Trade Value	Profit per Share	Total Profit
2013-06-06	1.000000e+06	1.002528e+06	2013-06-07	4600.0	63.936309	294107.023182	0.549521	2527.795697
2013-12-04	1.002528e+06	1.005509e+06	2013-12-06	3900.0	76.697001	299118.304396	0.764510	2981.587132
2013-12-16	1.005509e+06	1.006435e+06	2013-12-20	3900.0	76.450952	298158.713135	0.237262	925.320144
2014-07-31	1.006435e+06	1.008687e+06	2014-08-11	3100.0	96.036831	297714.177547	0.726545	2252.289192
2014-09-12	1.008687e+06	1.016728e+06	2014-09-16	3300.0	90.900451	299971.488441	2.436580	8040.714912
2014-10-13	1.016728e+06	1.026730e+06	2014-10-17	3700.0	80.860673	299184.490797	2.703359	10002.428814
2014-12-16	1.026730e+06	1.036824e+06	2014-12-17	4200.0	71.646199	300914.036651	2.403395	10094.260613
2015-03-12	1.036824e+06	1.037089e+06	2015-03-19	4200.0	72.737616	305497.988241	0.062906	264.206316
2015-07-01	1.037089e+06	1.035860e+06	2015-07-10	4000.0	76.174972	304699.888107	-0.307230	-1228.920071
2015-09-30	1.035860e+06	1.049763e+06	2015-10-05	4900.0	62.725269	307353.818194	2.837507	13903.782989
2016-01-15	1.049763e+06	1.060588e+06	2016-01-22	5400.0	57.676290	311451.965342	2.004620	10824.945312
2016-09-15	1.060588e+06	1.061451e+06	2016-09-22	4400.0	71.711209	315529.318228	0.196085	862.775479
2016-11-02	1.061451e+06	1.064262e+06	2016-11-03	4300.0	72.607599	312212.674826	0.653618	2810.556486
2017-02-06	1.064262e+06	1.066072e+06	2017-02-10	4100.0	76.728914	314588.548885	0.441457	1809.972065
2017-04-27	1.066072e+06	1.063377e+06	2017-05-08	4600.0	68.834471	316638.567986	-0.585906	-2695.166284
2017-06-27	1.063377e+06	1.071192e+06	2017-07-03	5100.0	62.215074	317296.875098	1.532532	7815.911397
2017-08-16	1.071192e+06	1.072034e+06	2017-08-25	5200.0	60.644467	315351.226225	0.161820	841.464581
2017-10-20	1.072034e+06	1.073153e+06	2017-10-26	5300.0	60.585851	321105.010424	0.211067	1118.655618
2018-02-09	1.073153e+06	1.075197e+06	2018-02-20	5000.0	63.479536	317397.680400	0.408980	2044.899920
2018-08-16	1.075197e+06	1.086776e+06	2018-08-21	5200.0	61.294802	318732.969922	2.226583	11578.233985
2018-10-24	1.086776e+06	1.082546e+06	2018-11-07	6200.0	52.415984	324979.099955	-0.682268	-4230.061855
2018-11-15	1.082546e+06	1.084092e+06	2018-11-19	6800.0	47.541228	323280.351412	0.227423	1546.474227
2018-11-28	1.084092e+06	1.080562e+06	2018-11-29	7000.0	45.889942	321229.592111	-0.504285	-3529.995518
2018-11-30	1.080562e+06	1.089105e+06	2018-12-03	7200.0	44.594621	321081.272972	1.186553	8543.182429
2018-12-14	1.089105e+06	1.064039e+06	2018-12-31	8300.0	39.100000	324530.000000	-3.020000	-25066.000000

+ 9%


Portfolio Growth on Fisher Strategy
Portfolio Value
1080000
1060000
1040000
1020000
1000000
2013-06-06
2013-12-04
2013-12-16
2014-07-31
2014-09-12
2014-10-13
2014-12-16
2015-03-12
2015-07-01
2015-09-30
2016-01-15
2016-09-15
2016-11-02
2017-02-06
2017-04-27
2017-06-27
2017-08-16
2017-10-20
2018-02-09
2018-08-16
2018-10-24
2018-11-15
2018-11-28
2018-11-30
2018-12-14
Date

Communications Sector Portfolio Analysis

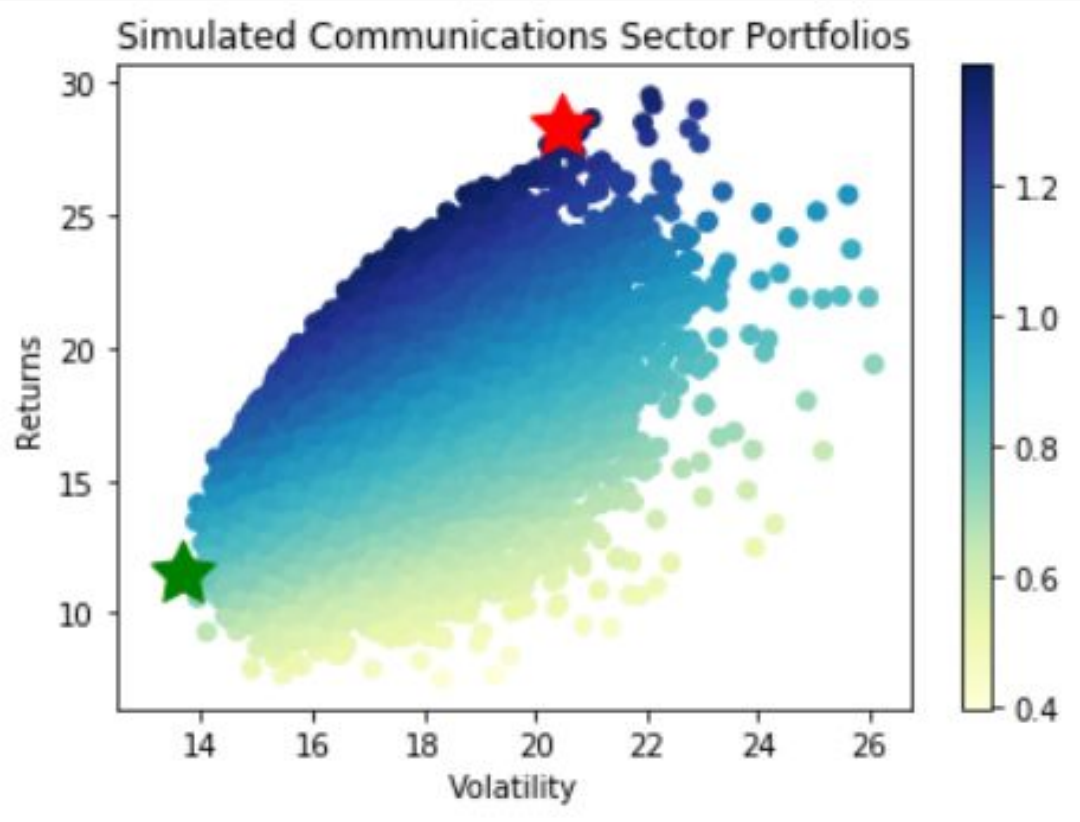
- How should an investor allocate their funds in a long term portfolio?
- Risk averse investor
  - Minimum volatility -> standard deviation
- Savvy Investor
  - Max Sharpe Ratio -> greatest return while carefully factoring in risk (volatility)

Sharpe Ratio(x) = (Rx - Rf) / StdDev (x)

Where:

x is the investment
Rx is the average rate of return of x
Rf is the best available rate of return of a risk-free security (calculated as zero for simplicity of demonstration)
StdDev(x) is the standard deviation of Rx

- Simulate 25,000 randomly weighted portfolios
- Red star = Max Sharpe Ratio
- Green star = Minimum volatility portfolio



Portfolio Allocations

Minimum Volatility Portfolio:

Portfolio Characteristics	% Allocation
ret	11.157127
stdev	13.590495
sharpe	0.820951
ATVI	3.290000
CMCSA	8.880000
CTL	2.440000
DIS	18.200000
FB	8.740000
FOXA	3.700000
NFLX	0.260000
T	24.600000
TWTR	1.880000
VZ	27.800000

Max Sharpe Ratio Portfolio:

Portfolio Characteristics	% Allocation
ret	28.361015
stdev	20.227419
sharpe	1.402107
ATVI	16.800000
CMCSA	9.930000
CTL	0.980000
DIS	11.900000
FB	20.200000
FOXA	7.430000
NFLX	24.000000
T	0.050000
TWTR	1.490000
VZ	6.940000

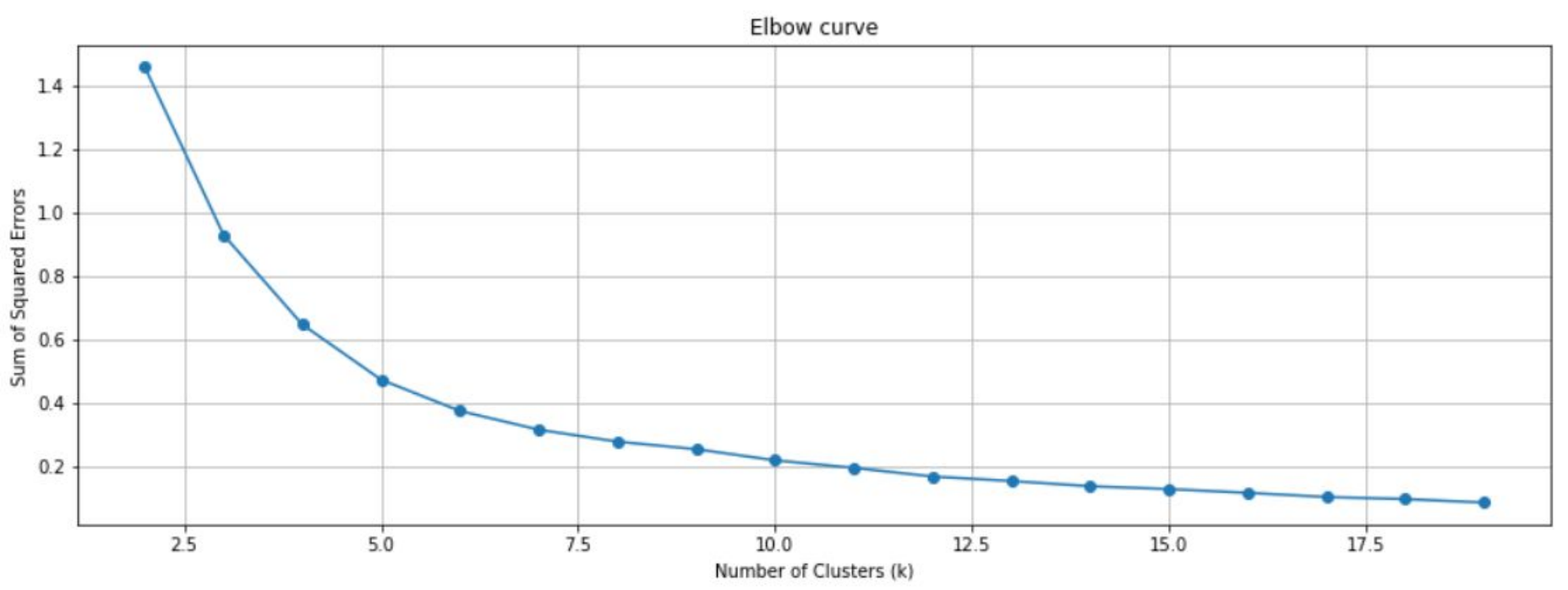
Stocks K-Means Clustering

The K-Means clustering algorithm is an unsupervised learning algorithm that groups a “cluster” of data points together based off similarities in their features. In this case, the features of interest are stock price returns and stock price volatility.

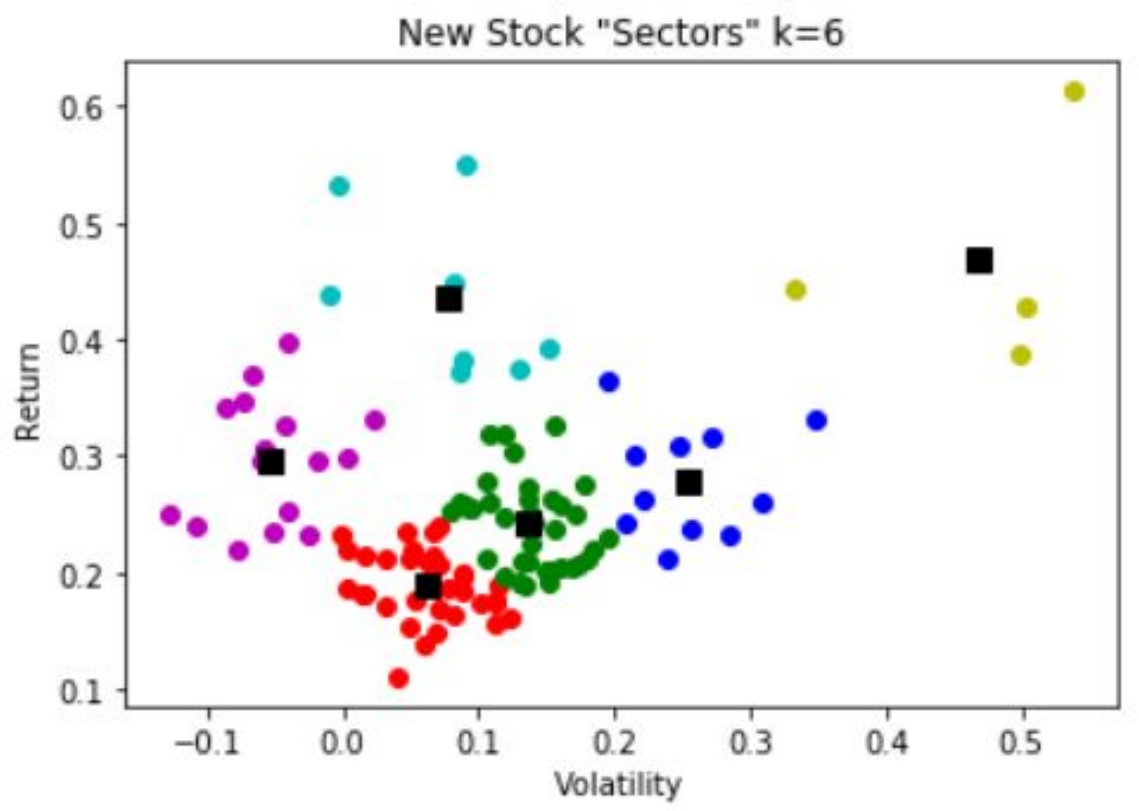
- Traditional stock sectors are divided based on similarities companies share in their resources, products, services and target customers
- Microsoft vs. Apple
- Clustering stock prices based on returns and volatility makes more sense when analyzing for investing and trading
- # of clusters = *k*

Elbow Method

- *k* is determined by where the rate of decrease of the Sum of Squared Errors slows down significantly
- *k = 6*



New “Sectors” (clusters)



Cluster	Ticker
0	11
1	4
2	33
3	38
4	16
5	8

Further Studies

- Fisher Transform Strategy
  - Test on other stocks
- Sector Portfolio Analysis
  - All of S&P 500
- Stocks K-Means Clustering
  - Test different allocations