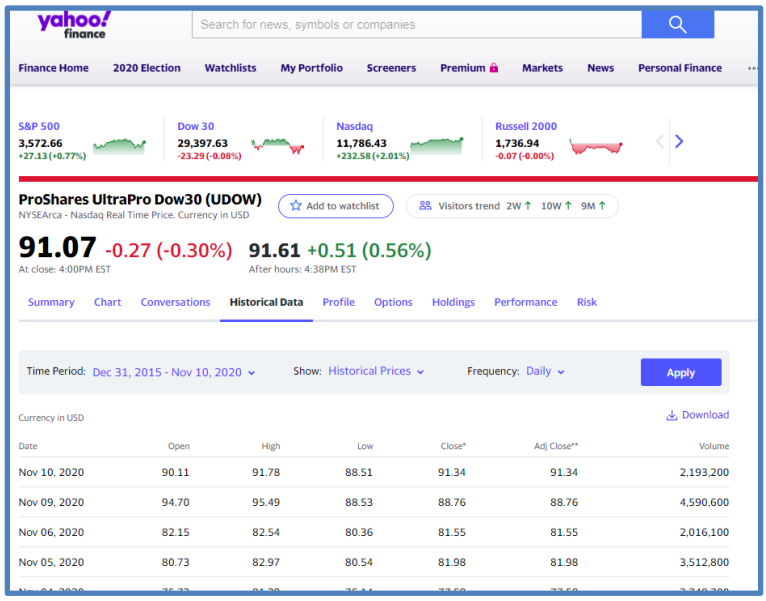
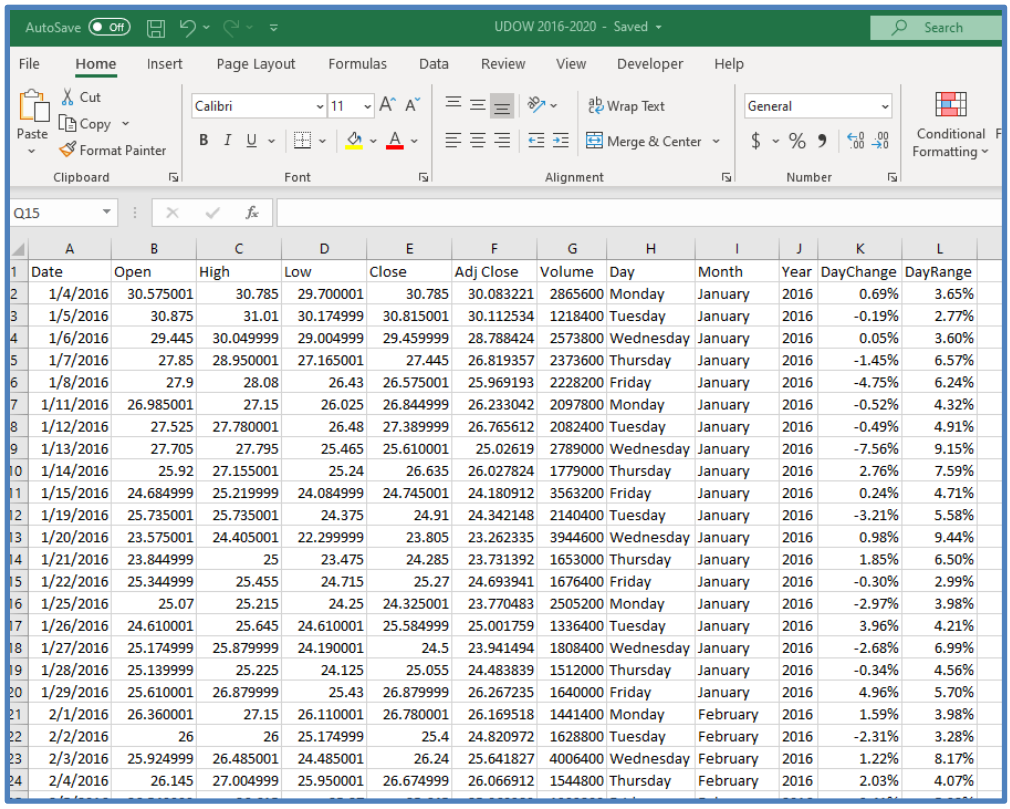
Dataset – Stock Price Data

I downloaded the daily stock price data from Yahoo Finance. I used the full years 2015 to 2019 and the year-to-date data for 2020. This data has columns for date, pricing data (open, high, low, close, and adjusted close), and volume of shares traded.

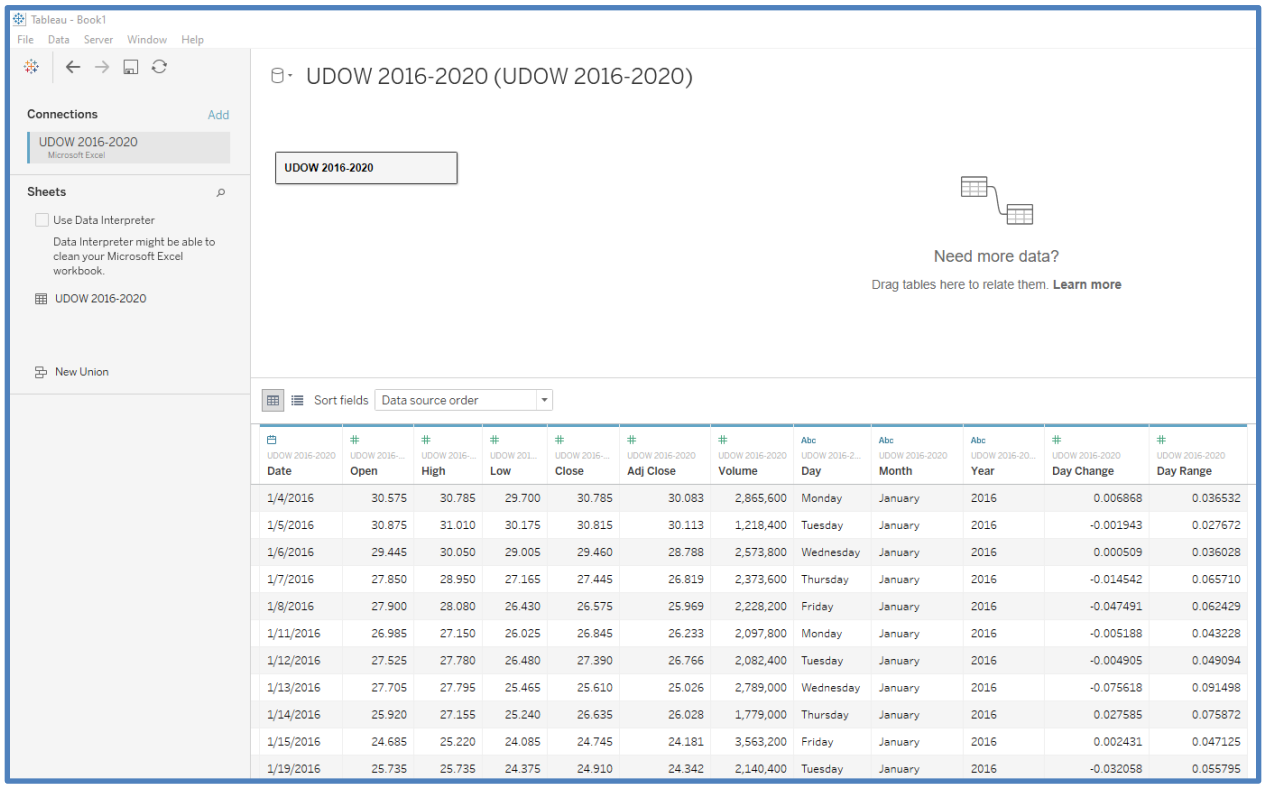


I downloaded the data into excel and added a few columns to make analysis easier. These columns included percent price change for each day and calendar data for each line including the day of the week, month, and year. These are columns H through L in the Excel worksheet below.



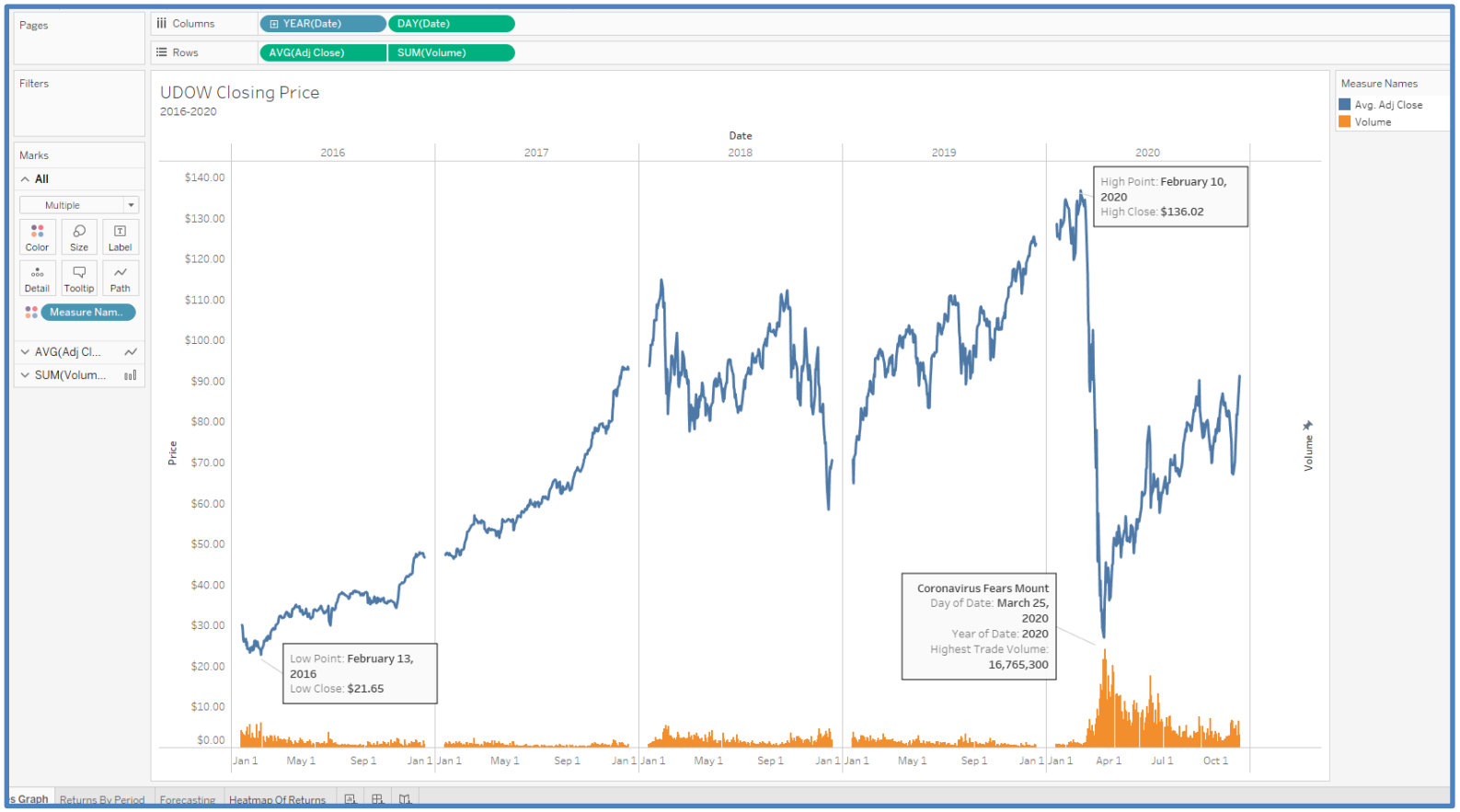
	Date	Open	High	Low	Close	Adj Close	Volume	Day	Month	Year	DayChange	DayRange
2	1/4/2016	30.575001	30.785	29.700001	30.785	30.083221	2865600	Monday	January	2016	0.69%	3.65%
3	1/5/2016	30.875	31.01	30.174999	30.815001	30.112534	1218400	Tuesday	January	2016	-0.19%	2.77%
4	1/6/2016	29.445	30.049999	29.004999	29.459999	28.788424	2573800	Wednesday	January	2016	0.05%	3.60%
5	1/7/2016	27.85	28.950001	27.165001	27.445	26.819357	2373600	Thursday	January	2016	-1.45%	6.57%
6	1/8/2016	27.9	28.08	26.43	26.575001	25.969193	2228200	Friday	January	2016	-4.75%	6.24%
7	1/11/2016	26.985001	27.15	26.025	26.844999	26.233042	2097800	Monday	January	2016	-0.52%	4.32%
8	1/12/2016	27.525	27.780001	26.48	27.389999	26.765612	2082400	Tuesday	January	2016	-0.49%	4.91%
9	1/13/2016	27.705	27.795	25.465	25.610001	25.02619	2789000	Wednesday	January	2016	-7.56%	9.15%
10	1/14/2016	25.92	27.155001	25.24	26.635	26.027824	1779000	Thursday	January	2016	2.76%	7.59%
11	1/15/2016	24.684999	25.219999	24.084999	24.745001	24.180912	3563200	Friday	January	2016	0.24%	4.71%
12	1/19/2016	25.735001	25.735001	24.375	24.91	24.342148	2140400	Tuesday	January	2016	-3.21%	5.58%
13	1/20/2016	23.575001	24.405001	22.299999	23.805	23.262335	3944600	Wednesday	January	2016	0.98%	9.44%
14	1/21/2016	23.844999	25	23.475	24.285	23.731392	1653000	Thursday	January	2016	1.85%	6.50%
15	1/22/2016	25.344999	25.455	24.715	25.27	24.693941	1676400	Friday	January	2016	-0.30%	2.99%
16	1/25/2016	25.07	25.215	24.25	24.325001	23.770483	2505200	Monday	January	2016	-2.97%	3.98%
17	1/26/2016	24.610001	25.645	24.610001	25.584999	25.001759	1336400	Tuesday	January	2016	3.96%	4.21%
18	1/27/2016	25.174999	25.879999	24.190001	24.5	23.941494	1808400	Wednesday	January	2016	-2.68%	6.99%
19	1/28/2016	25.139999	25.225	24.125	25.055	24.483839	1512000	Thursday	January	2016	-0.34%	4.56%
20	1/29/2016	25.610001	26.879999	25.43	26.879999	26.267235	1640000	Friday	January	2016	4.96%	5.70%
21	2/1/2016	26.360001	27.15	26.110001	26.780001	26.169518	1441400	Monday	February	2016	1.59%	3.98%
22	2/2/2016	26	26	25.174999	25.4	24.820972	1628800	Tuesday	February	2016	-2.31%	3.28%
23	2/3/2016	25.924999	26.485001	24.485001	26.24	25.641827	4006400	Wednesday	February	2016	1.22%	8.17%
24	2/4/2016	26.145	27.004999	25.950001	26.674999	26.066912	1544800	Thursday	February	2016	2.03%	4.07%

Next, I imported the data into Tableau to start the data visualization.



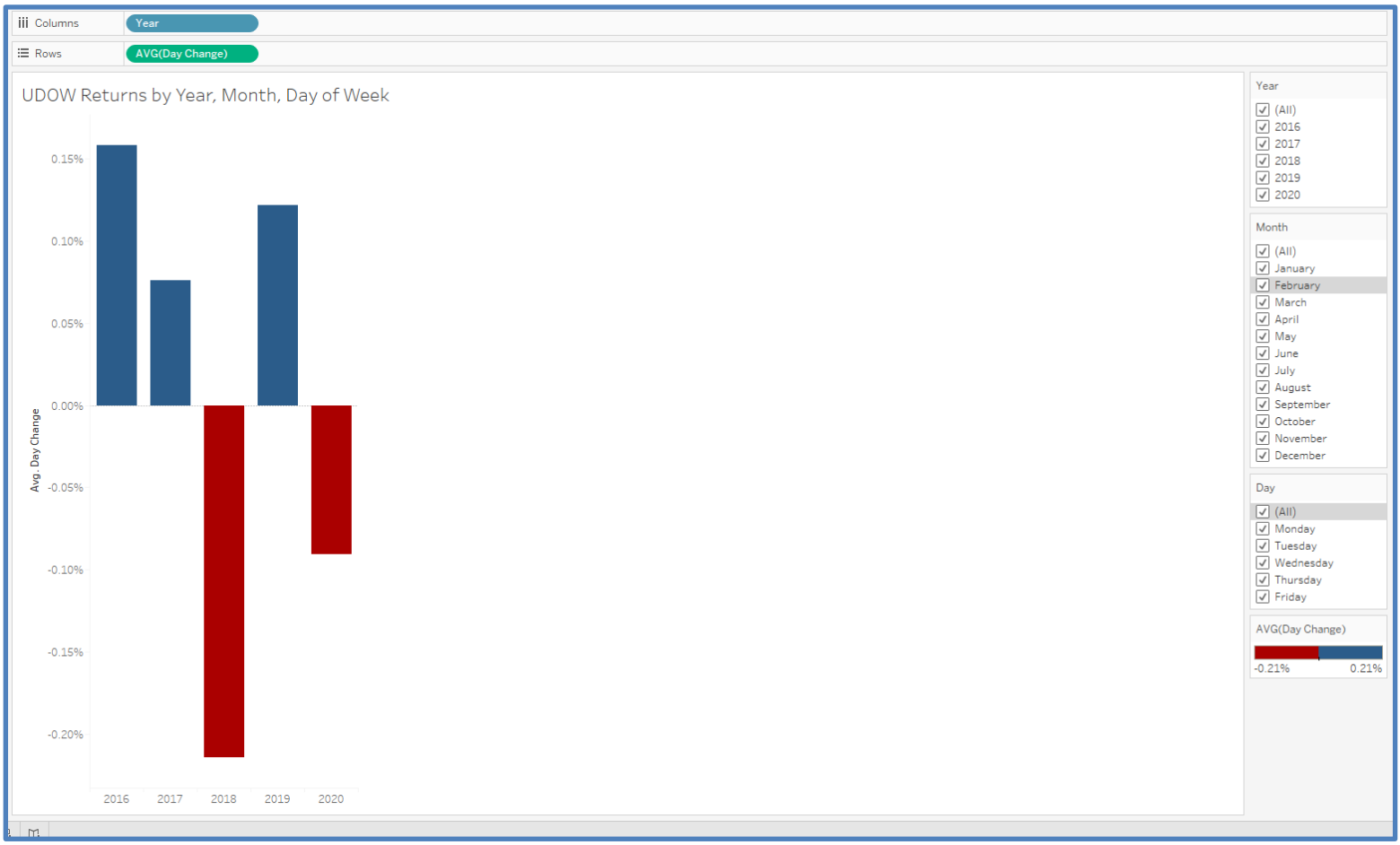
Date	Open	High	Low	Close	Adj Close	Volume	Day	Month	Year	Day Change	Day Range
1/4/2016	30.575	30.785	29.700	30.785	30.083	2,865,600	Monday	January	2016	0.006868	0.036532
1/5/2016	30.875	31.010	30.175	30.815	30.113	1,218,400	Tuesday	January	2016	-0.001943	0.027672
1/6/2016	29.445	30.050	29.005	29.460	28.788	2,573,800	Wednesday	January	2016	0.000509	0.036028
1/7/2016	27.850	28.950	27.165	27.445	26.819	2,373,600	Thursday	January	2016	-0.014542	0.065710
1/8/2016	27.900	28.080	26.430	26.575	25.969	2,228,200	Friday	January	2016	-0.047491	0.062429
1/11/2016	26.985	27.150	26.025	26.845	26.233	2,097,800	Monday	January	2016	-0.005188	0.043228
1/12/2016	27.525	27.780	26.480	27.390	26.766	2,082,400	Tuesday	January	2016	-0.004905	0.049094
1/13/2016	27.705	27.795	25.465	25.610	25.026	2,789,000	Wednesday	January	2016	-0.075618	0.091498
1/14/2016	25.920	27.155	25.240	26.635	26.028	1,779,000	Thursday	January	2016	0.027585	0.075872
1/15/2016	24.685	25.220	24.085	24.745	24.181	3,563,200	Friday	January	2016	0.002431	0.047125
1/19/2016	25.735	25.735	24.375	24.910	24.342	2,140,400	Tuesday	January	2016	-0.032058	0.055795

The first graph I created was a simple line plot of the 'adj close' price data. The price line is in blue. I also added volume data into this graph, on the bottom in orange. Seeing these two classes of information in conjunction is helpful to understand the story better. For example, there is a steep drop in price in March 2020. This corresponds with the highest volume day in these periods on March 25. If it were not obvious in price, the high volume would tell me that something very significant happened that day. Of course, that is when concerns about Covid-19 were mounting and the uncertainty caused the market to crash. As the graph shows, since then there have been several rallies and high-volume periods. It has been a volatile stetch, to say the least, but it is at least trending up. I annotated the Covid-19 caused dip, as well as the period high ($136.02 on February 10, 2020) and the period low ($21.65 on February 13, 2016). Like most stocks, UDOW had a period of gain from 2016 to 2017. 2018 was a tumultuous year followed by an exceptionally positive 2019.



Plotting the price line above was interesting, but I wanted to dive deeper into specific facets of the stock price performance. To do this, I created the following graph, which depicts average

daily returns by the year, month, and date, with filters for day of the week. As the above graph showed, the below confirms that 2016, 2017, and 2019 were positive (blue) years, while 2018 and 2020 were negative (red).



The filtering functionality of Tableau let me dive deeper. For example, Thursdays have had an average positive return in each year, with a 0.28% per day return in 2016 as the highest year.


UDOW Returns by Year, Month, Day of Week
Avg. Day Change
0.28%
0.26%
0.24%
0.22%
0.20%
0.18%
0.16%
0.14%
0.12%
0.10%
0.08%
0.06%
0.04%
0.02%
0.00%
2016
2017
2018
2019
2020
Year
(All)
2016
2017
2018
2019
2020
Month
(All)
January
February
March
April
May
June
July
August
September
October
November
December
Day
(All)
Monday
Tuesday
Wednesday
Thursday
Friday
AVG(Day Change)
-0.28%
0.28%

Also, November has been the only month with a positive average return each day of the month in every year. The returns were slightly less in later years, but the overall trend held true.

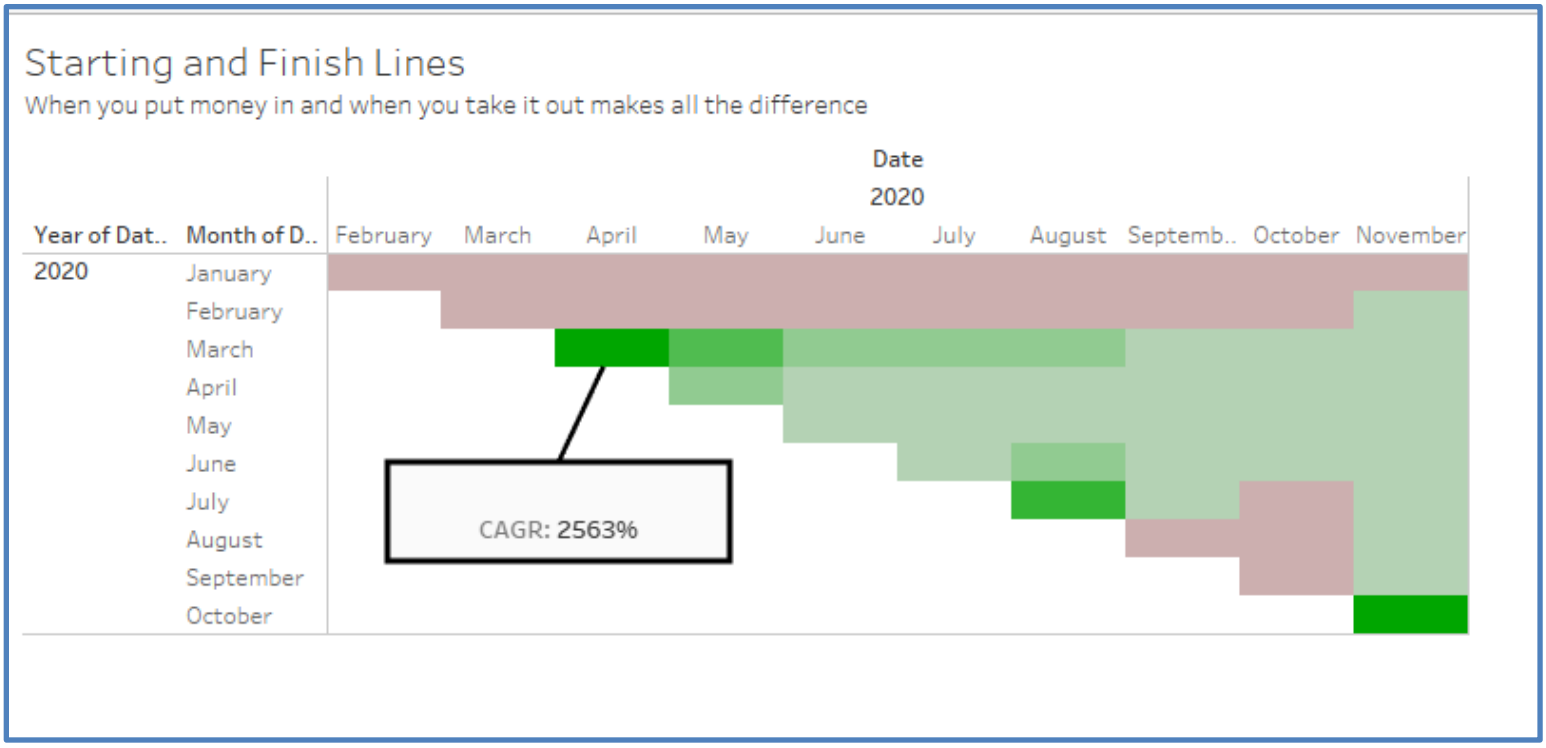
These are just two of the actionable insights that this analysis could provide. If I would have invested heavily in UDOW on Thursdays, and/or every day in November for the last five years, I would be a very, very rich man (or, at least richer than I am now). On the flip side, this analysis could point out months or days to not invest. Gaining these insights would not have been possible just by looking at the raw data.

As far as why these insights appear, it is unclear. Perhaps seasonality kicks in in November, boosting business for many of the companies in the Dow. Perhaps traders are extra happy the weekend is coming, so they are more willing to pay higher prices for UDOW on Thursdays, making average returns higher.

For the next steps, I focus on this idea of how *when* one invests can alter how well the investments do. In 2011, The New York Times published a heat map of returns for the S&P 500, which illustrated how return percentages depend both on when one invests money *and* when one takes out money ("In Investing, It's When You Start and When You Finish", 2011). With some help from Edupristine.com, I have recreated this heatmap and display it, and my steps below.

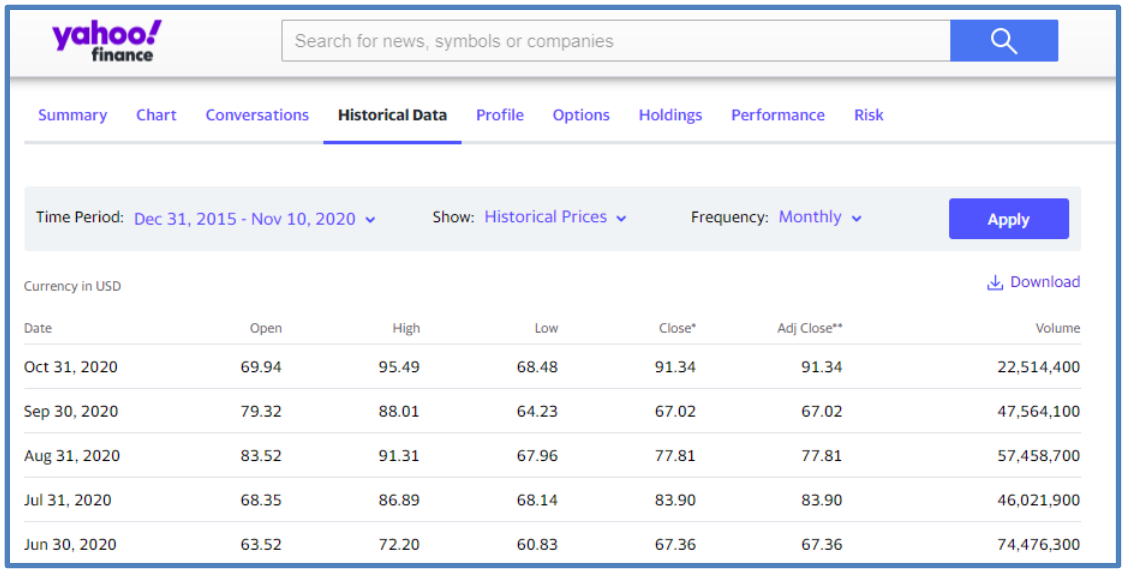
Directly below is the heatmap for 2020. For example, I annotated one square - the colored square corresponding to March on the vertical axis, and April on the horizontal axis is dark green. The annotation box says this month of April has a Compounded Annual Growth Rate (CAGR) of 2,563%. This means that if UDOW grew at the rate that it did in April for every

month for a year, the annual rate of return would be 2,563%, It is unlikely that the stock would grow 32% every month though, so looking to the right of the square provides more realistic numbers. The square for November of the same row has a CAGR of 224%, meaning if the stock grew for a year how it did monthly April through November 2020, then the annual return would be 224%. That rate of return is very good and shows that investing when the market is in a big upswing, can prove to be immensely successful.



But there is a reverse side to the above positive example too. Starting investing in UDOW in February and selling in November would give one and annual return of -28%. This shows that when you put money in is as important as when you take it out. Conventional finance wisdom says to leave money in the market for long stretches, which is still true because the long stretches allows the historically more frequent positive years to overcome the historically less frequent down years. But, if one gets the timing right, then their returns are significantly improved.

In the following figures, I show how I created the above heatmap. First, I needed to get the UDOW data in monthly format, not daily. For this I went to Yahoo finance again and imported it to another Tableau workbook after adding some columns of calculation in Excel.



yahoo! finance

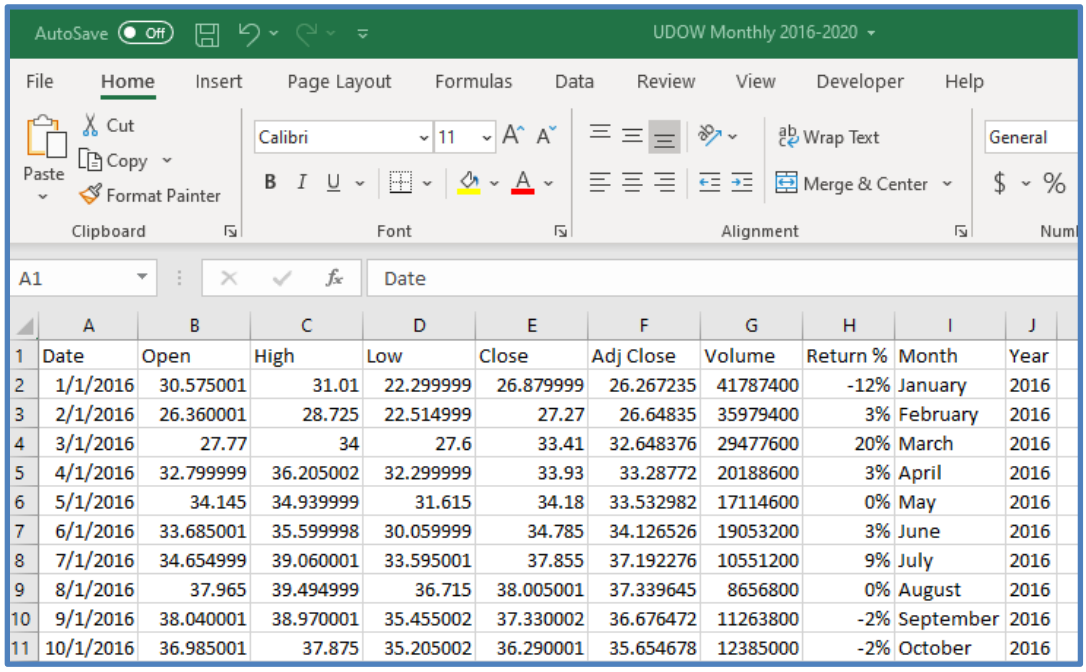
Search for news, symbols or companies

Summary | Chart | Conversations | **Historical Data** | Profile | Options | Holdings | Performance | Risk

Time Period: Dec 31, 2015 - Nov 10, 2020 Show: Historical Prices Frequency: Monthly Apply

Currency in USD Download

Date	Open	High	Low	Close*	Adj Close**	Volume
Oct 31, 2020	69.94	95.49	68.48	91.34	91.34	22,514,400
Sep 30, 2020	79.32	88.01	64.23	67.02	67.02	47,564,100
Aug 31, 2020	83.52	91.31	67.96	77.81	77.81	57,458,700
Jul 31, 2020	68.35	86.89	68.14	83.90	83.90	46,021,900
Jun 30, 2020	63.52	72.20	60.83	67.36	67.36	74,476,300



AutoSave Off — UDOW Monthly 2016-2020

File | Home | Insert | Page Layout | Formulas | Data | Review | View | Developer | Help

Cut Copy Format Painter Paste — Clipboard

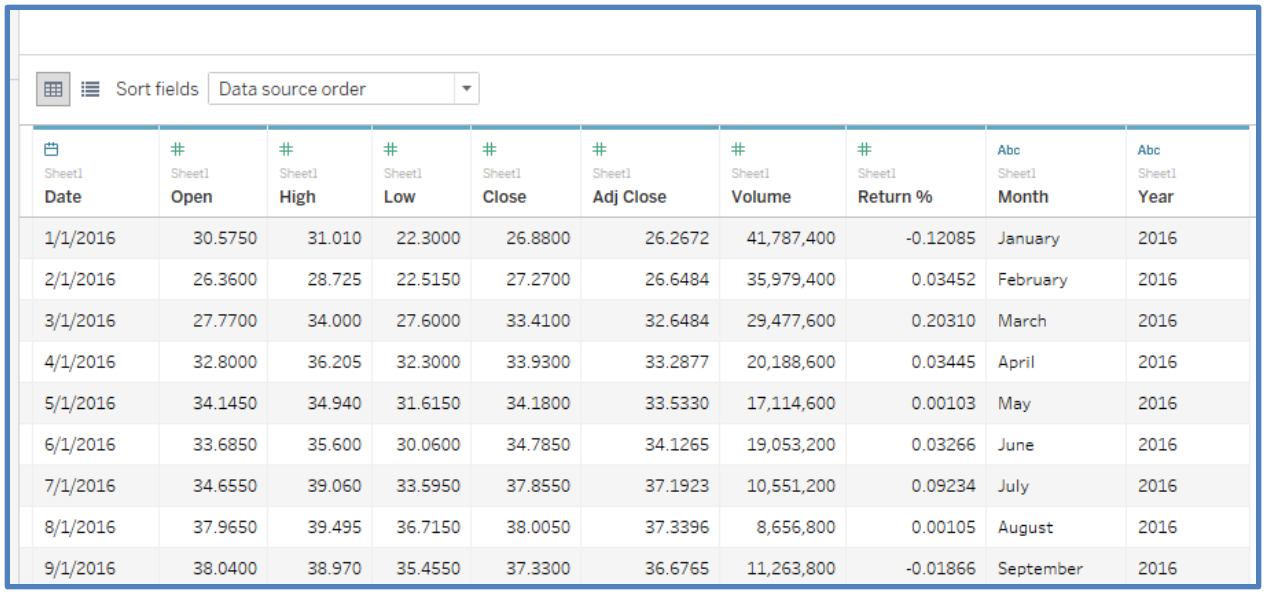
Calibri 11 — Font

Wrap Text Merge & Center — Alignment

General — Number

A1 Date

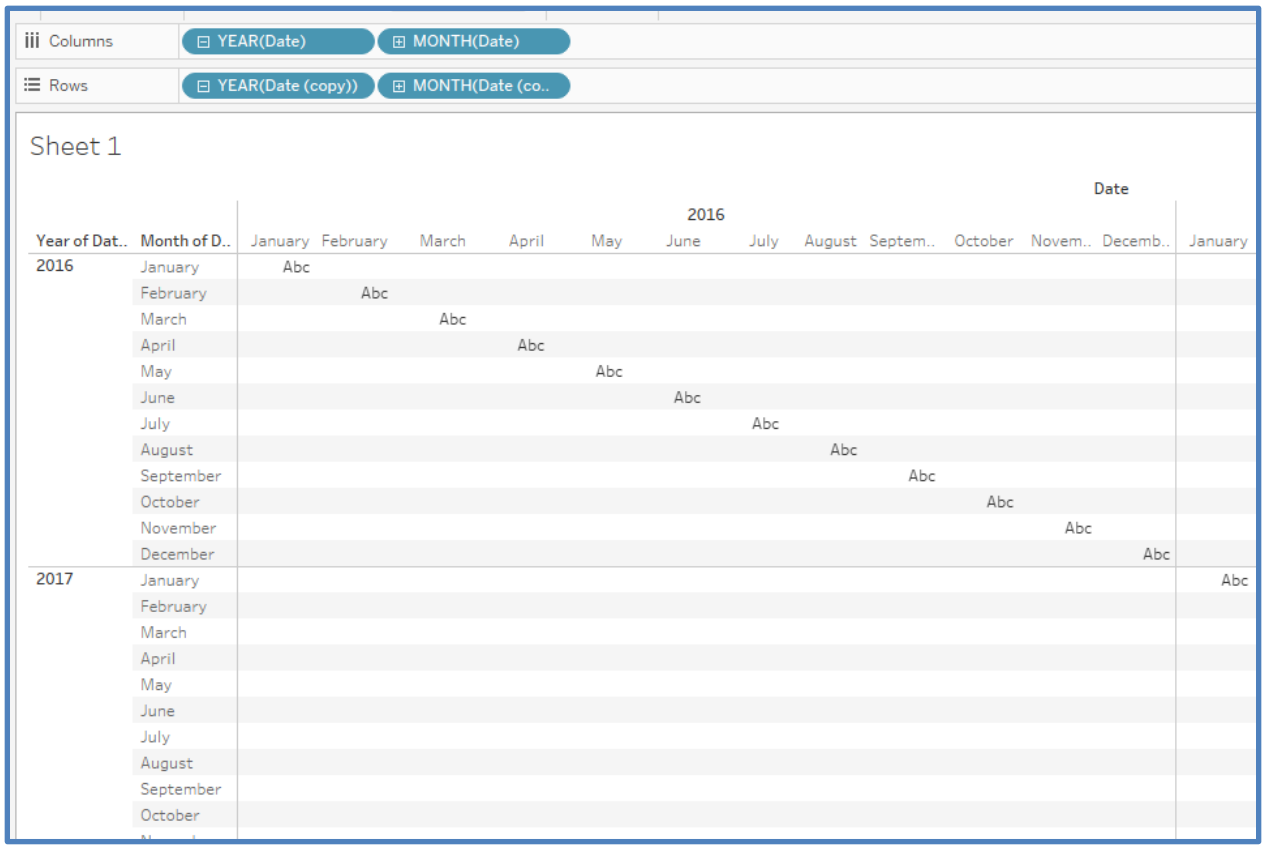
	A	B	C	D	E	F	G	H	I	J
1	Date	Open	High	Low	Close	Adj Close	Volume	Return %	Month	Year
2	1/1/2016	30.575001	31.01	22.299999	26.879999	26.267235	41787400	-12%	January	2016
3	2/1/2016	26.360001	28.725	22.514999	27.27	26.64835	35979400	3%	February	2016
4	3/1/2016	27.77	34	27.6	33.41	32.648376	29477600	20%	March	2016
5	4/1/2016	32.799999	36.205002	32.299999	33.93	33.28772	20188600	3%	April	2016
6	5/1/2016	34.145	34.939999	31.615	34.18	33.532982	17114600	0%	May	2016
7	6/1/2016	33.685001	35.599998	30.059999	34.785	34.126526	19053200	3%	June	2016
8	7/1/2016	34.654999	39.060001	33.595001	37.855	37.192276	10551200	9%	July	2016
9	8/1/2016	37.965	39.494999	36.715	38.005001	37.339645	8656800	0%	August	2016
10	9/1/2016	38.040001	38.970001	35.455002	37.330002	36.676472	11263800	-2%	September	2016
11	10/1/2016	36.985001	37.875	35.205002	36.290001	35.654678	12385000	-2%	October	2016



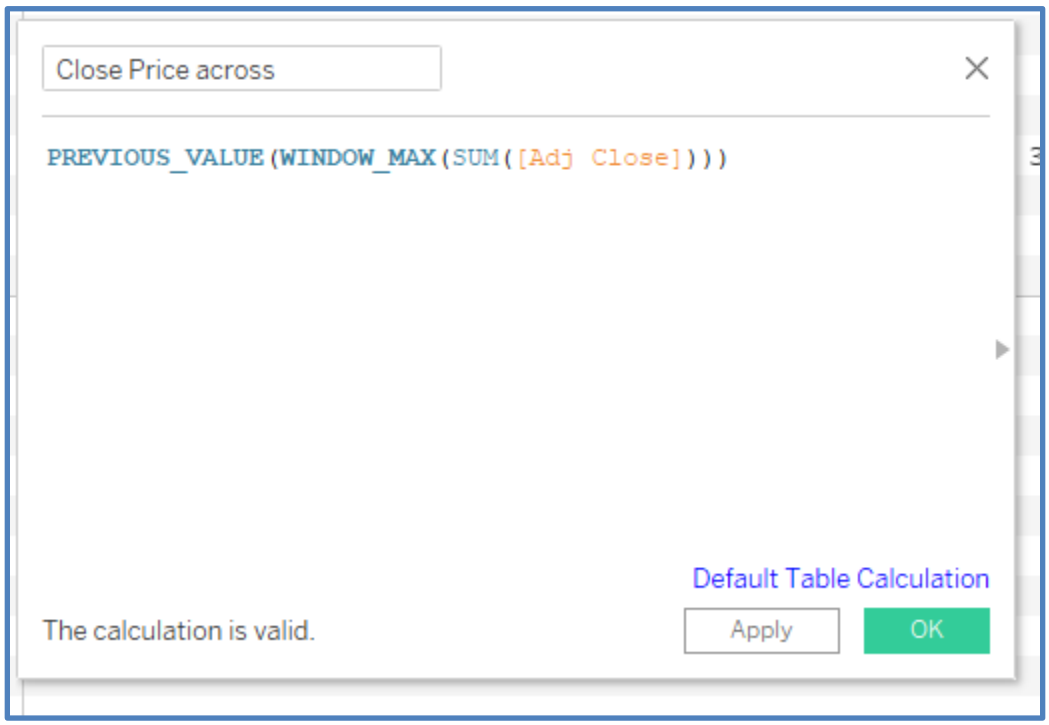
Sort fields: Data source order

Sheet1 Date	Sheet1 Open	Sheet1 High	Sheet1 Low	Sheet1 Close	Sheet1 Adj Close	Sheet1 Volume	Sheet1 Return %	Sheet1 Month	Sheet1 Year
1/1/2016	30.5750	31.010	22.3000	26.8800	26.2672	41,787,400	-0.12085	January	2016
2/1/2016	26.3600	28.725	22.5150	27.2700	26.6484	35,979,400	0.03452	February	2016
3/1/2016	27.7700	34.000	27.6000	33.4100	32.6484	29,477,600	0.20310	March	2016
4/1/2016	32.8000	36.205	32.3000	33.9300	33.2877	20,188,600	0.03445	April	2016
5/1/2016	34.1450	34.940	31.6150	34.1800	33.5330	17,114,600	0.00103	May	2016
6/1/2016	33.6850	35.600	30.0600	34.7850	34.1265	19,053,200	0.03266	June	2016
7/1/2016	34.6550	39.060	33.5950	37.8550	37.1923	10,551,200	0.09234	July	2016
8/1/2016	37.9650	39.495	36.7150	38.0050	37.3396	8,656,800	0.00105	August	2016
9/1/2016	38.0400	38.970	35.4550	37.3300	36.6765	11,263,800	-0.01866	September	2016

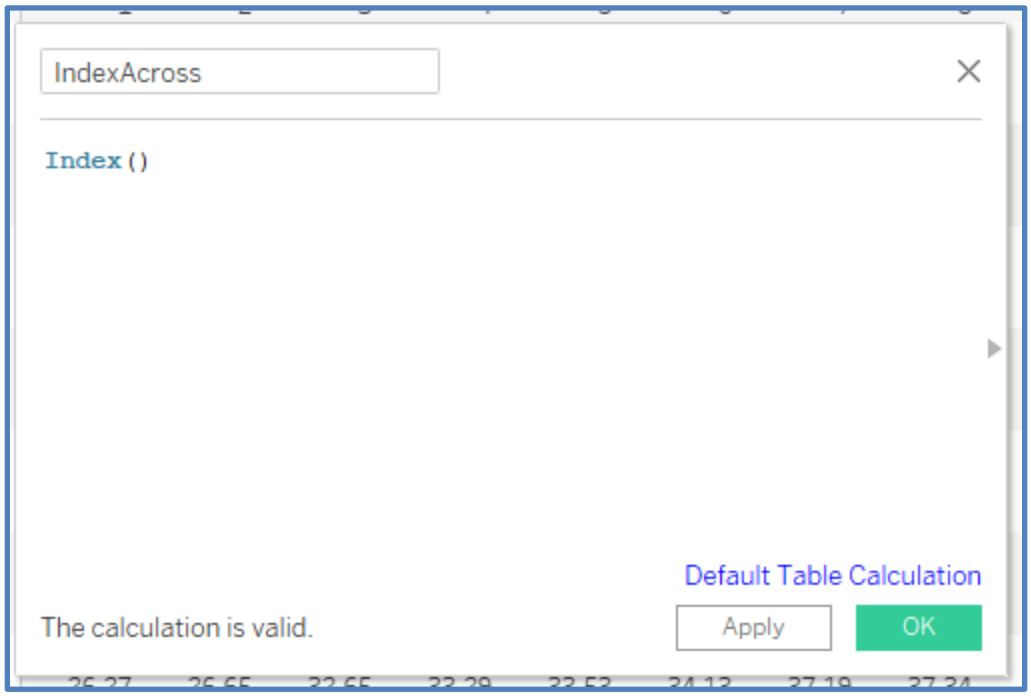
Then, in Tableau, I put the ‘adj close’ price data in the row and column, with price values in the Label mark.



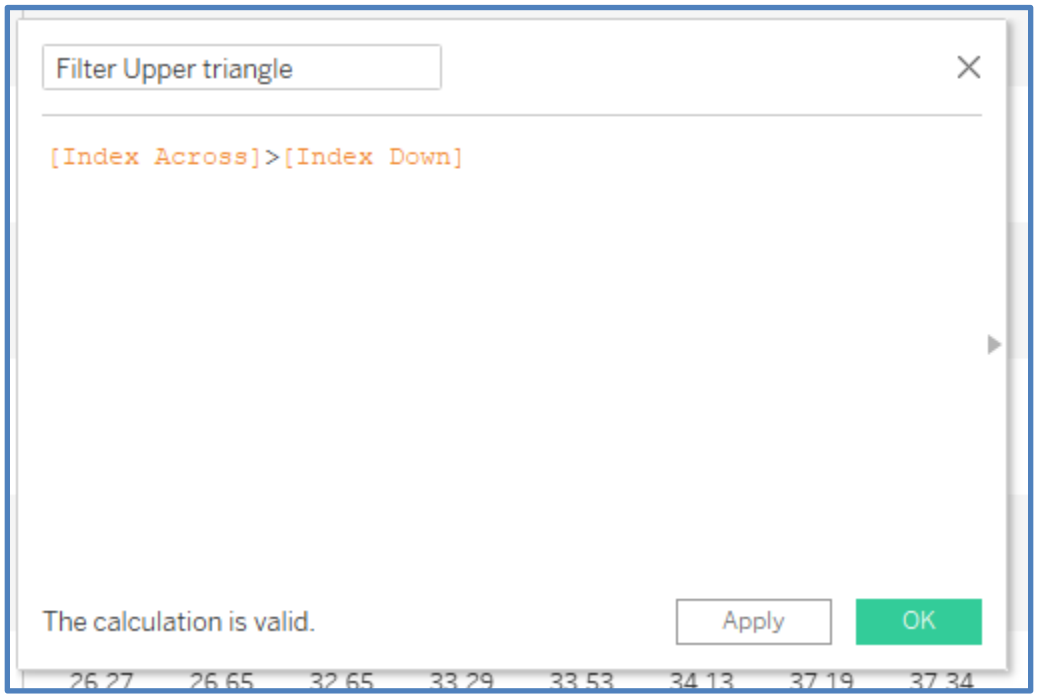
Then I created calculated fields to fill in the rows and columns of the table. Both calculations were the same as below, but one I set to calculated down the rows, and one across the rows.

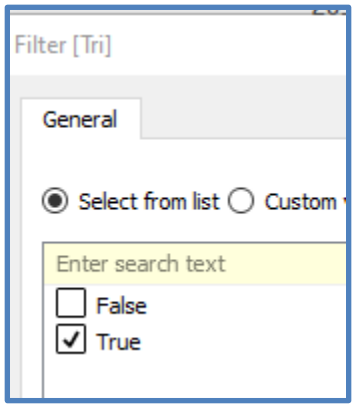


Then, I did the same thign for the Index. I need to keep track of the index later to eliminate the values I do not need from the chart. I created the below calculated field twice – once for down the table, and once for across. This function puts the row and column number into my table.



Then I used the index in the following calculated field that I applied as a filter. If the row number is greater than the column number, I want to display that cell. This step forms the right triangle shape of the table. I set the filter to show all "TRUE" values to finish this below.

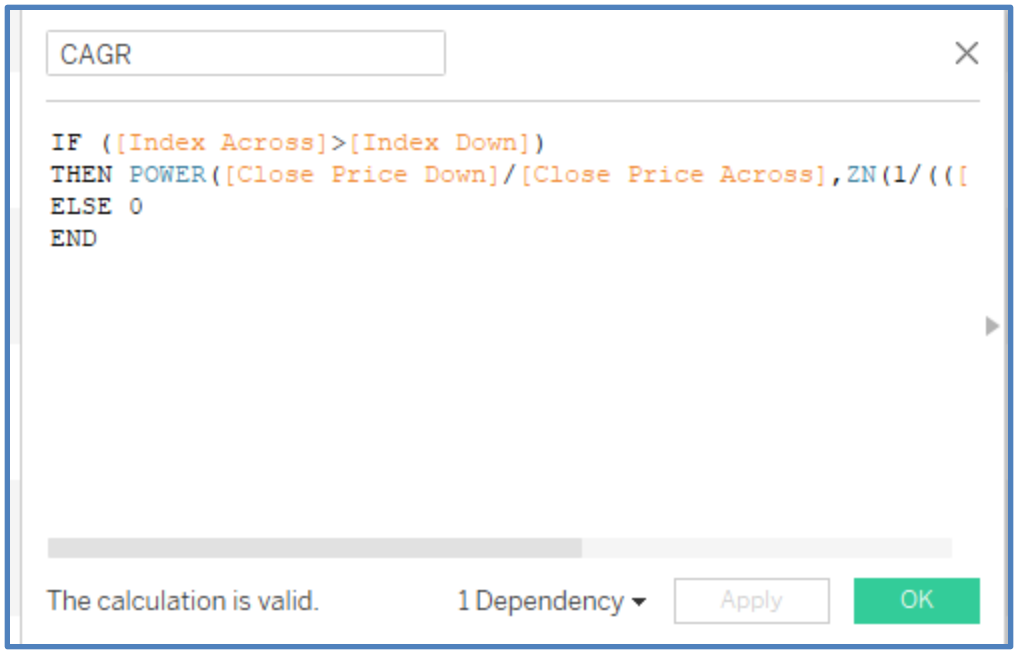




The result of indexing and the filter is the following right triangle shaped table.

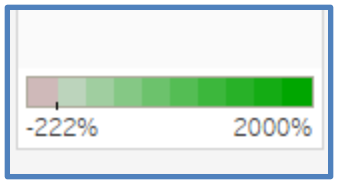
		2016							
Year of Dat..	Month of D..	February	March	April	May	June	July	August	Septer
2016	January	26.27 26.65 2 1	26.27 32.65 3 1	26.27 33.29 4 1	26.27 33.53 5 1	26.27 34.13 6 1	26.27 37.19 7 1	26.27 37.34 8 1	26. 36.
	February		26.65 32.65 3 2	26.65 33.29 4 2	26.65 33.53 5 2	26.65 34.13 6 2	26.65 37.19 7 2	26.65 37.34 8 2	26. 36.
	March			32.65 33.29 4 3	32.65 33.53 5 3	32.65 34.13 6 3	32.65 37.19 7 3	32.65 37.34 8 3	32. 36.
	April				33.29 33.53 5 4	33.29 34.13 6 4	33.29 37.19 7 4	33.29 37.34 8 4	33. 36.
	May					33.53 34.13 6 5	33.53 37.19 7 5	33.53 37.34 8 5	33. 36.
	June						34.13 37.19 7 6	34.13 37.34 8 6	34. 36.
	July							37.19 37.34 8 7	37. 36.
	August								37. 36.

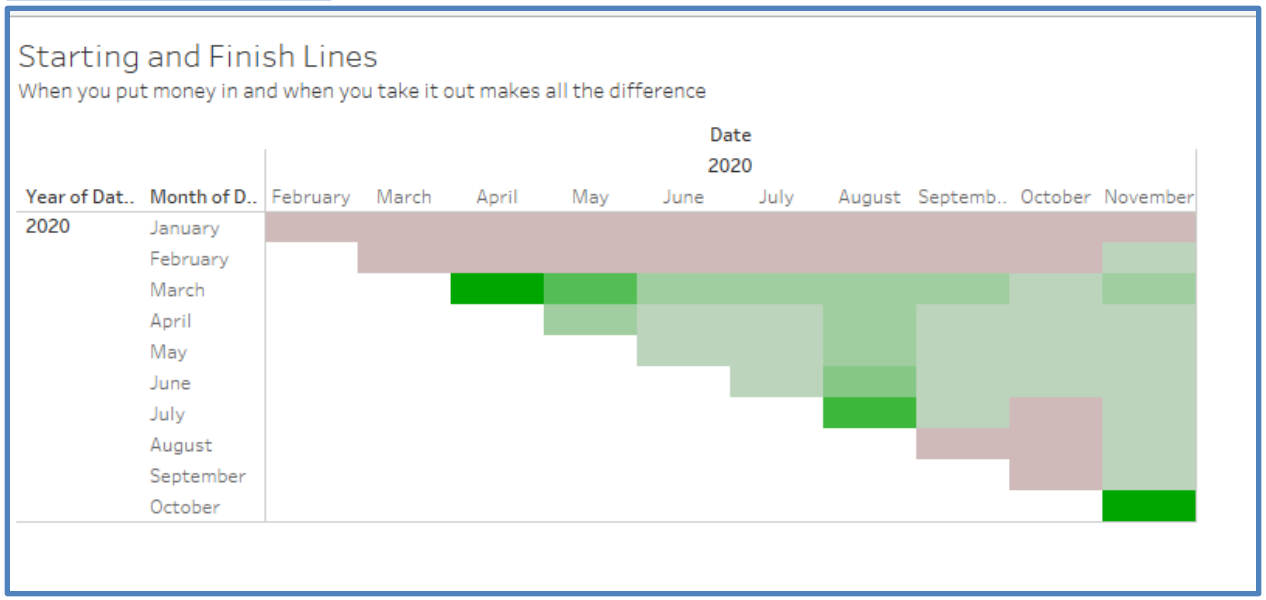
Next I create the calculated field that computes the CAGR for each month, and then I dsplay the CAGR value and remove the index values and the close prices from displaying.



Year of Dat..	Month of D..	February	March	April
2016	January	19%	269%	158%
	February		1044%	280%
	March			26%
	April			
	May			
	June			
	July			

Then I add the color formatting by adding the CAGR field to the color Mark box. The result is below.

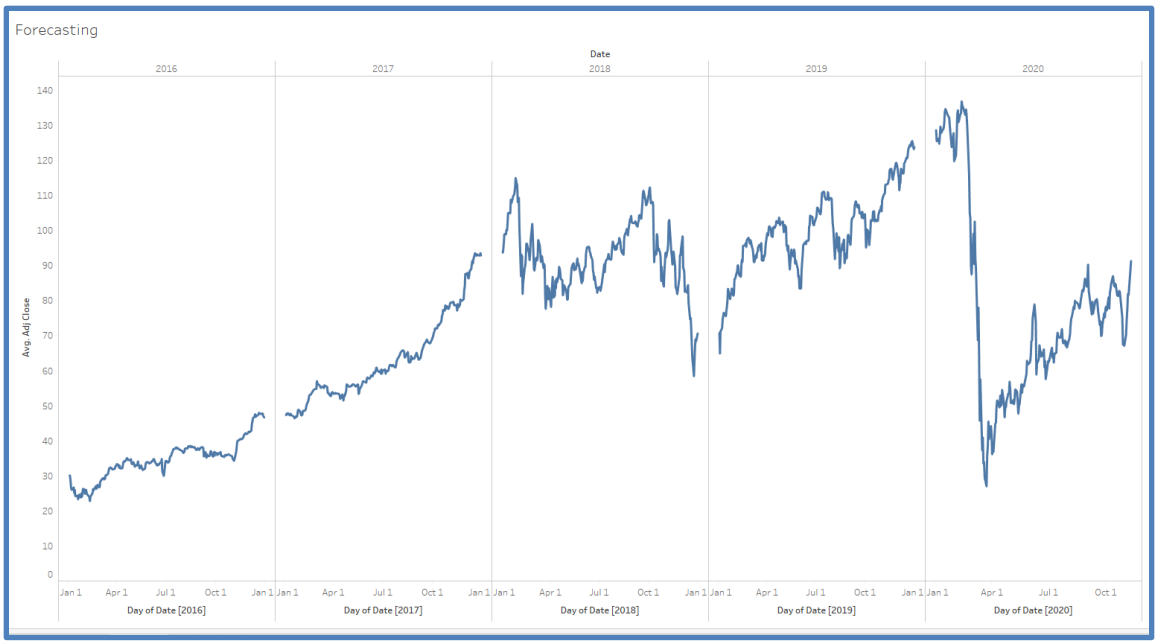




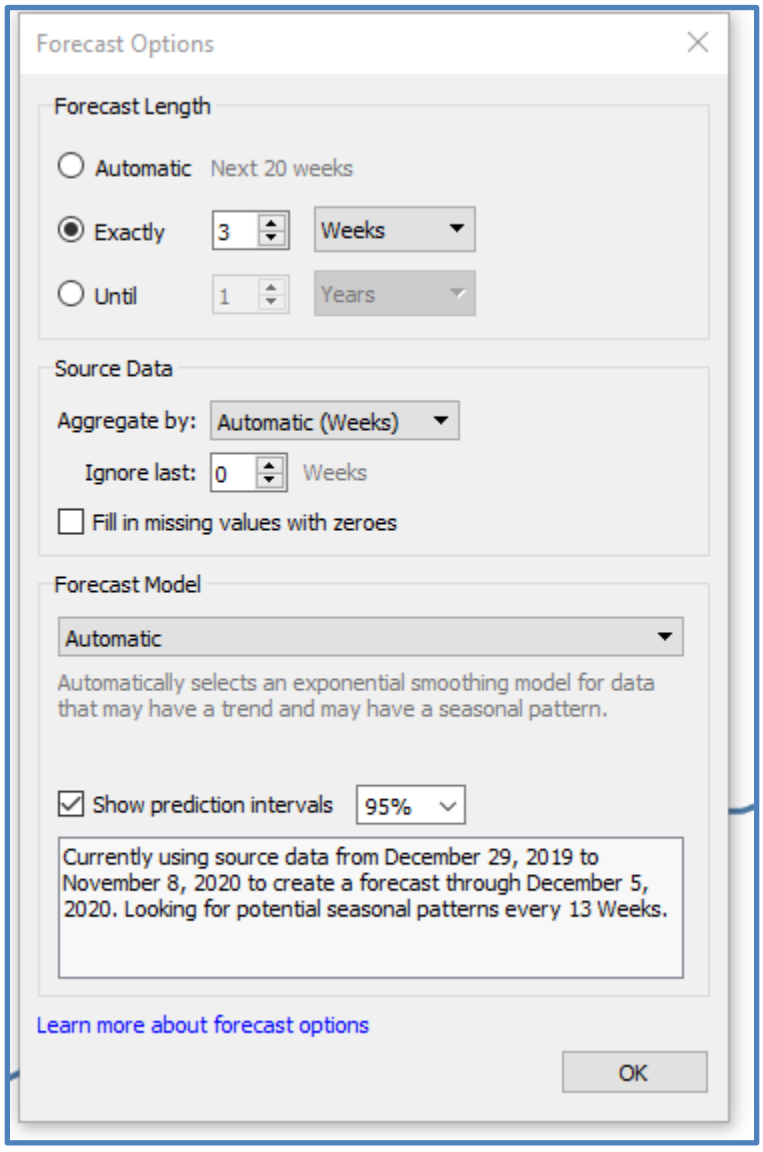
The process of learning how to create this chart was challenging but rewarding. It shows the immense importance of timing in investing in stocks. It reinforces the age-old adage of 'buy low, sell high'.

But how does one know when the price will be low enough to buy or high enough to sell? In the next steps, I utilize Tableau's forecasting feature to try and answer that question.

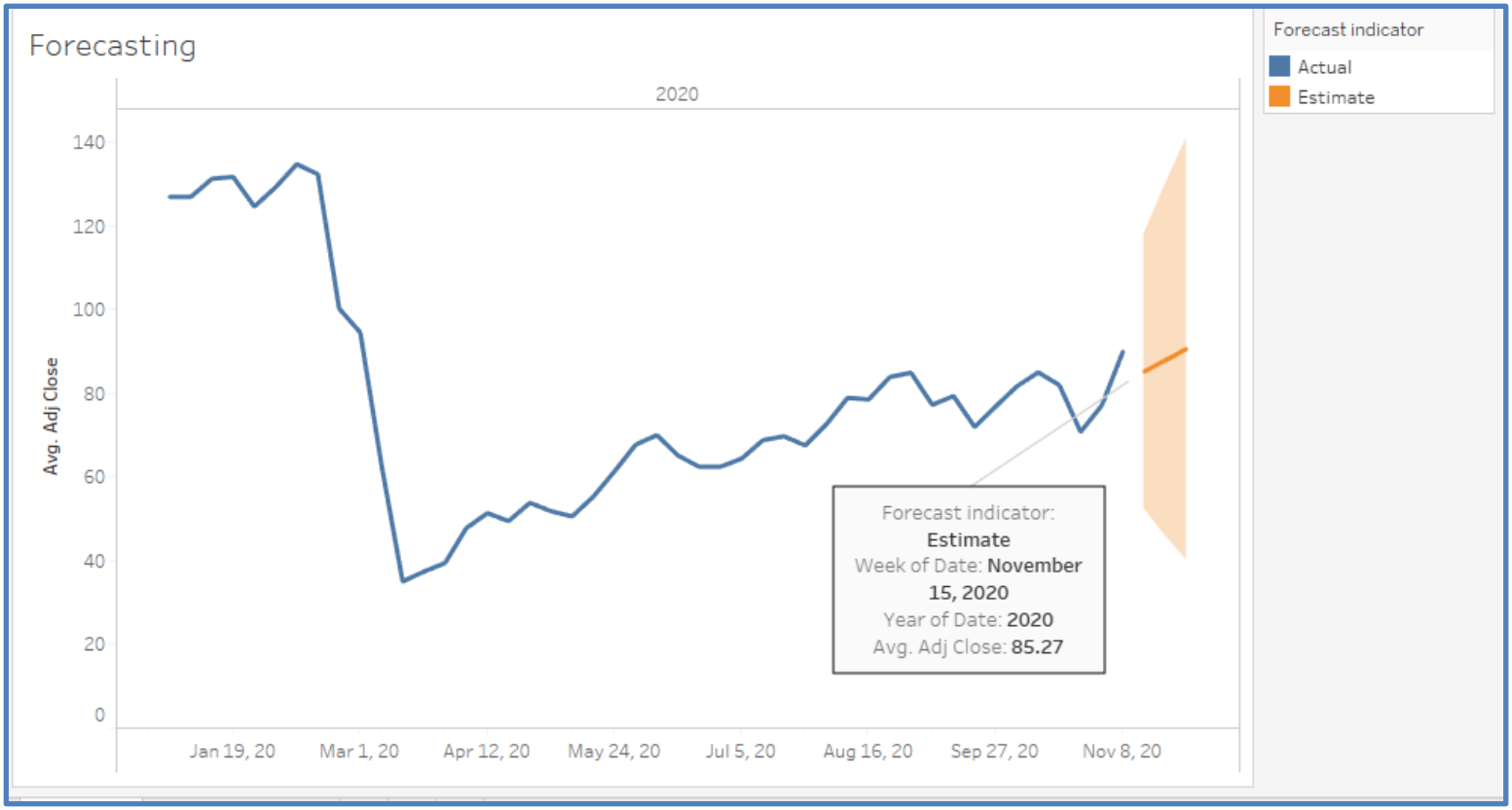
For this graph, I start with the simple price line graph from the first step of this lab. This is the daily stock price data for UDOW (The heatmap used monthly data).



Then I used the Analytics section and the Forecasting option to customize the forecasting settings. I wanted to forecast UDOW's price 3 weeks out using data up to the most recent week, and withing a range given a 95% confidence interval.



This produced the following forecast below. The stock price is currently around $90, but the forecast projects the stock to fall to $85.27 at the beginning of next week and then to rise for the next few weeks back up to $90.



Additionally, the forecasting function plots a large orange cloud representing the range of possible future values for UDOW. This is a wide range and does not inspire much confidence. It is basically saying the stock could go down by a lot, or it could go up by a lot. This may not be the best insight to make a decision on giving its wide range of possibilities. However, it is still an interesting functionality in Tableau, and I could see it being more applicable to a business in terms of projecting sales or expenses. One insight from this graph might be that forecasting stock price is too difficult and that there are always risks with investing. It is best to pay attention to many other factors in conjunction with computer forecasts.