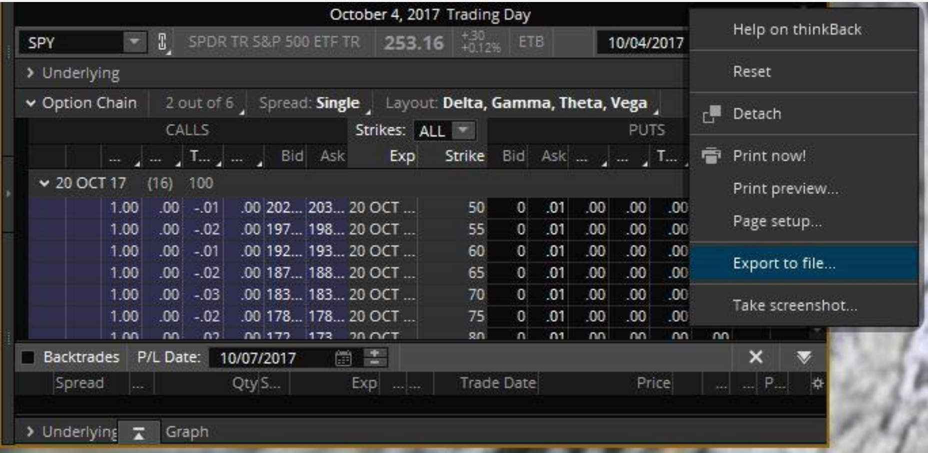
Data Acquisition



Option Chain | 2 out of 6 | Spread: Single | Layout: Delta, Gamma, Theta, Vega

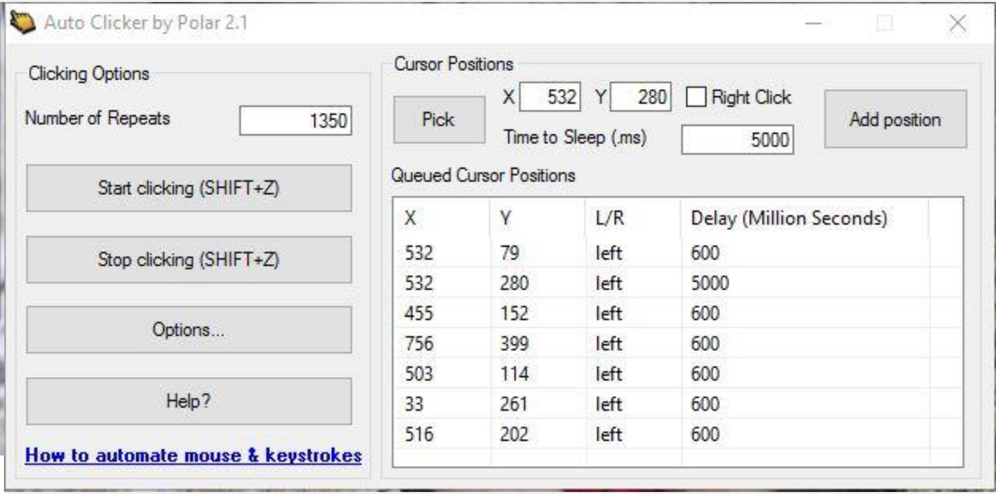
CALLS | Strikes: ALL | PUTS

20 OCT 17 (16) 100

...	...	T...	...	Bid	Ask	Exp	Strike	Bid	Ask	...	...	T...
1.00	.00	-.01	.00	202...	203...	20 OCT ...	50	0	.01	.00	.00	.00
1.00	.00	-.02	.00	197...	198...	20 OCT ...	55	0	.01	.00	.00	.00
1.00	.00	-.01	.00	192...	193...	20 OCT ...	60	0	.01	.00	.00	.00
1.00	.00	-.02	.00	187...	188...	20 OCT ...	65	0	.01	.00	.00	.00
1.00	.00	-.03	.00	183...	183...	20 OCT ...	70	0	.01	.00	.00	.00
1.00	.00	-.02	.00	178...	178...	20 OCT ...	75	0	.01	.00	.00	.00
1.00	.00	-.02	.00	172...	173...	20 OCT	80	0	.01	.00	.00	.00

□ Backtrades P/L Date: 10/07/2017

Spread	...	Qty	S...	Exp	...	...	Trade Date	Price	...	...	P...



Clicking Options

Number of Repeats: 1350

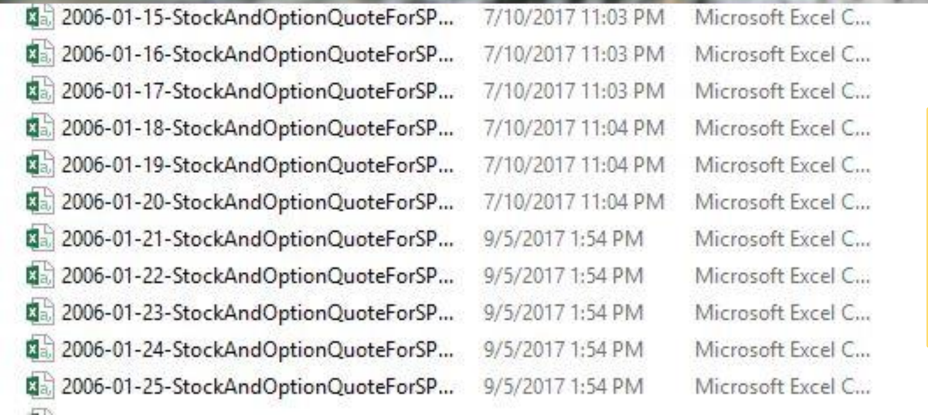
Cursor Positions

X: 532 Y: 280 □ Right Click

Time to Sleep (.ms): 5000

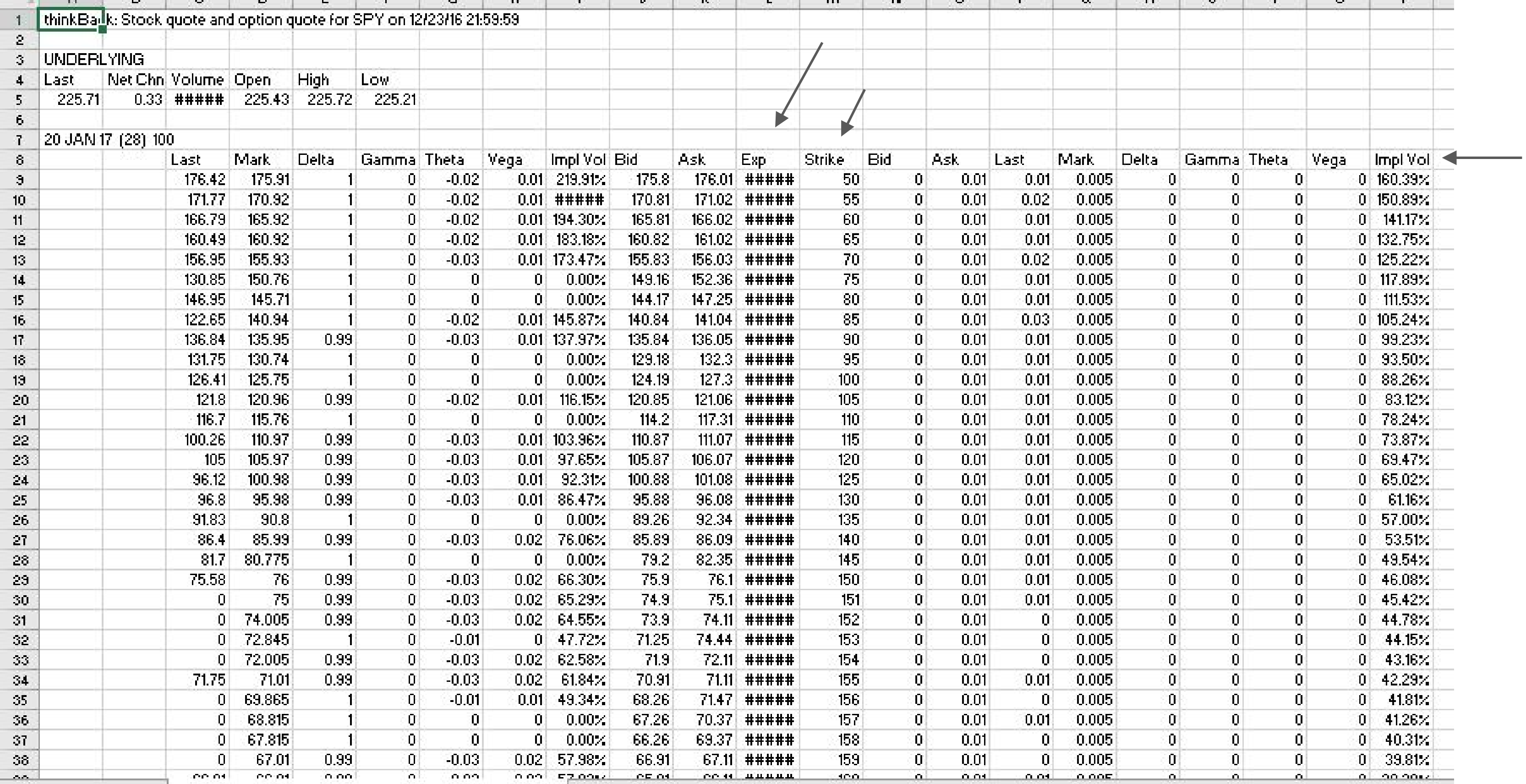
Queued Cursor Positions

X	Y	L/R	Delay (Million Seconds)
532	79	left	600
532	280	left	5000
455	152	left	600
756	399	left	600
503	114	left	600
33	261	left	600
516	202	left	600



2006-01-15-StockAndOptionQuoteForSP...	7/10/2017 11:03 PM	Microsoft Excel C...
2006-01-16-StockAndOptionQuoteForSP...	7/10/2017 11:03 PM	Microsoft Excel C...
2006-01-17-StockAndOptionQuoteForSP...	7/10/2017 11:03 PM	Microsoft Excel C...
2006-01-18-StockAndOptionQuoteForSP...	7/10/2017 11:04 PM	Microsoft Excel C...
2006-01-19-StockAndOptionQuoteForSP...	7/10/2017 11:04 PM	Microsoft Excel C...
2006-01-20-StockAndOptionQuoteForSP...	7/10/2017 11:04 PM	Microsoft Excel C...
2006-01-21-StockAndOptionQuoteForSP...	9/5/2017 1:54 PM	Microsoft Excel C...
2006-01-22-StockAndOptionQuoteForSP...	9/5/2017 1:54 PM	Microsoft Excel C...
2006-01-23-StockAndOptionQuoteForSP...	9/5/2017 1:54 PM	Microsoft Excel C...
2006-01-24-StockAndOptionQuoteForSP...	9/5/2017 1:54 PM	Microsoft Excel C...
2006-01-25-StockAndOptionQuoteForSP...	9/5/2017 1:54 PM	Microsoft Excel C...

The first step was to obtain data on historical option prices. This was done using a program called ThinkorSwim, and required the use of an autoclicker, as ThinkorSwim only allows the downloading of a single of data at a time. Data was downloaded for every day, including non-trading days (such as weekends and holidays).



thinkBack: Stock quote and option quote for SPY on 12/23/16 21:59:59

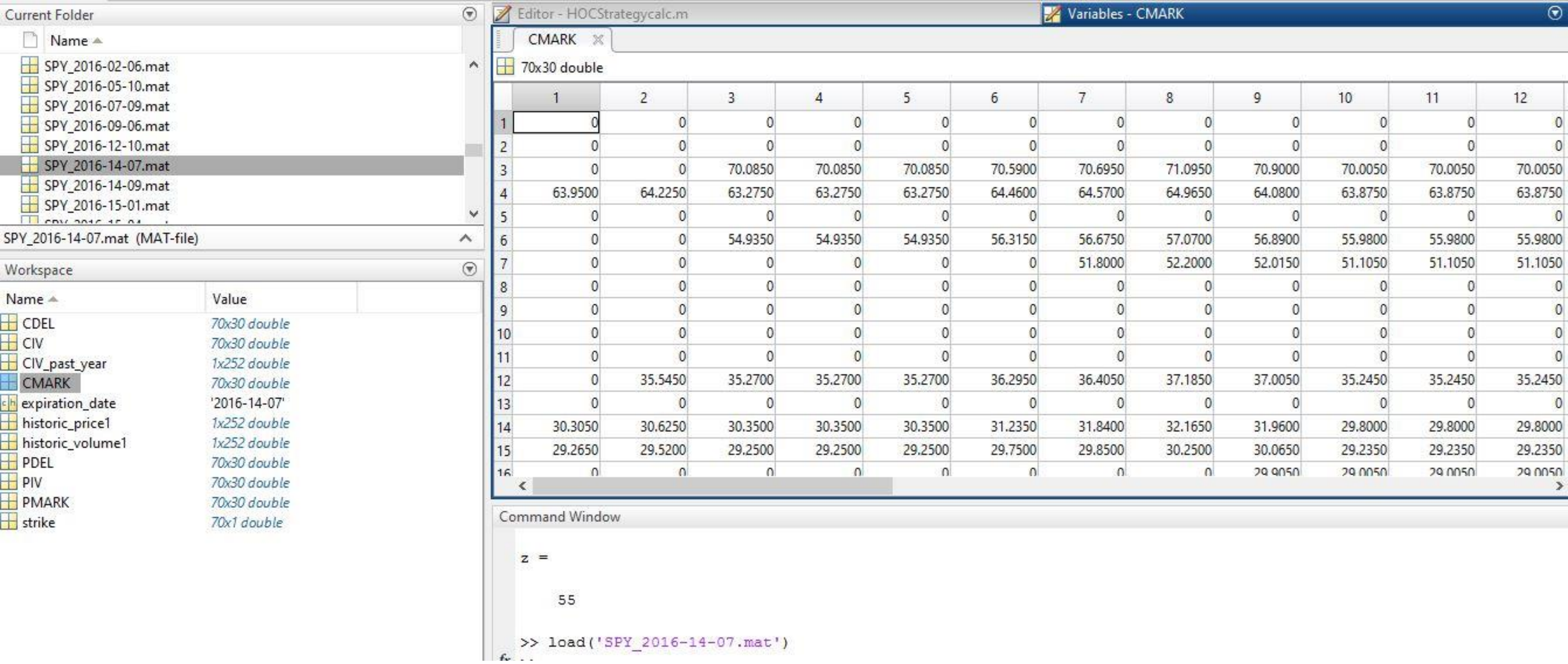
UNDERLYING

Last	Net Chn	Volume	Open	High	Low
225.71	0.33	#####	225.43	225.72	225.21

20 JAN 17 (28) 100

Last	Mark	Delta	Gamma	Theta	Vega	Impl Vol	Bid	Ask	Exp	Strike	Bid	Ask	Last	Mark	Delta	Gamma	Theta	Vega	Impl Vol
176.42	175.91	1	0	-0.02	0.01	219.91%	175.8	176.01	#####	50	0	0.01	0.01	0.005	0	0	0	0	160.39%
171.77	170.92	1	0	-0.02	0.01	#####	170.81	171.02	#####	55	0	0.01	0.02	0.005	0	0	0	0	150.89%
166.79	165.92	1	0	-0.02	0.01	194.30%	165.81	166.02	#####	60	0	0.01	0.01	0.005	0	0	0	0	141.17%
160.49	160.92	1	0	-0.02	0.01	183.18%	160.82	161.02	#####	65	0	0.01	0.01	0.005	0	0	0	0	132.75%
156.95	155.93	1	0	-0.03	0.01	173.47%	155.83	156.03	#####	70	0	0.01	0.02	0.005	0	0	0	0	125.22%
130.85	150.76	1	0	0	0	0.00%	149.16	152.36	#####	75	0	0.01	0.01	0.005	0	0	0	0	117.89%
146.95	145.71	1	0	0	0	0.00%	144.17	147.25	#####	80	0	0.01	0.01	0.005	0	0	0	0	111.53%
122.65	140.94	1	0	-0.02	0.01	145.87%	140.84	141.04	#####	85	0	0.01	0.03	0.005	0	0	0	0	105.24%
136.84	135.95	0.99	0	-0.03	0.01	137.97%	135.84	136.05	#####	90	0	0.01	0.01	0.005	0	0	0	0	99.23%
131.75	130.74	1	0	0	0	0.00%	129.18	132.3	#####	95	0	0.01	0.01	0.005	0	0	0	0	93.50%
126.41	125.75	1	0	0	0	0.00%	124.19	127.3	#####	100	0	0.01	0.01	0.005	0	0	0	0	88.26%
121.8	120.96	0.99	0	-0.02	0.01	116.15%	120.85	121.06	#####	105	0	0.01	0.01	0.005	0	0	0	0	83.12%
116.7	115.76	1	0	0	0	0.00%	114.2	117.31	#####	110	0	0.01	0.01	0.005	0	0	0	0	78.24%
100.26	110.97	0.99	0	-0.03	0.01	103.96%	110.87	111.07	#####	115	0	0.01	0.01	0.005	0	0	0	0	73.87%
105	105.97	0.99	0	-0.03	0.01	97.65%	105.87	106.07	#####	120	0	0.01	0.01	0.005	0	0	0	0	69.47%
96.12	100.98	0.99	0	-0.03	0.01	92.31%	100.88	101.08	#####	125	0	0.01	0.01	0.005	0	0	0	0	65.02%
96.8	95.98	0.99	0	-0.03	0.01	86.47%	95.88	96.08	#####	130	0	0.01	0.01	0.005	0	0	0	0	61.16%
91.83	90.8	1	0	0	0	0.00%	89.26	92.34	#####	135	0	0.01	0.01	0.005	0	0	0	0	57.00%
86.4	85.99	0.99	0	-0.03	0.02	76.06%	85.89	86.09	#####	140	0	0.01	0.01	0.005	0	0	0	0	53.51%
81.7	80.775	1	0	0	0	0.00%	79.2	82.35	#####	145	0	0.01	0.01	0.005	0	0	0	0	49.54%
75.58	76	0.99	0	-0.03	0.02	66.30%	75.9	76.1	#####	150	0	0.01	0.01	0.005	0	0	0	0	46.08%
0	75	0.99	0	-0.03	0.02	65.29%	74.9	75.1	#####	151	0	0.01	0.01	0.005	0	0	0	0	45.42%
0	74.005	0.99	0	-0.03	0.02	64.55%	73.9	74.11	#####	152	0	0.01	0	0.005	0	0	0	0	44.78%
0	72.845	1	0	-0.01	0	47.72%	71.25	74.44	#####	153	0	0.01	0	0.005	0	0	0	0	44.15%
0	72.005	0.99	0	-0.03	0.02	62.58%	71.9	72.11	#####	154	0	0.01	0	0.005	0	0	0	0	43.16%
71.75	71.01	0.99	0	-0.03	0.02	61.84%	70.91	71.11	#####	155	0	0.01	0.01	0.005	0	0	0	0	42.29%
0	69.865	1	0	-0.01	0.01	49.34%	68.26	71.47	#####	156	0	0.01	0	0.005	0	0	0	0	41.81%
0	68.815	1	0	0	0	0.00%	67.26	70.37	#####	157	0	0.01	0.01	0.005	0	0	0	0	41.26%
0	67.815	1	0	0	0	0.00%	66.26	69.37	#####	158	0	0.01	0	0.005	0	0	0	0	40.31%
0	67.01	0.99	0	-0.03	0.02	57.98%	66.91	67.11	#####	159	0	0.01	0	0.005	0	0	0	0	39.81%

End of day option chain data was contained in a .csv file with this structure. Each data point has a corresponding expiration date, strike price, and category. Data was assigned to a .mat file labeled by stock symbol and expiration date, to a matrix corresponding to category, to the row in that matrix corresponding to the data point's strike price, and the column in that file corresponding to the date that data point was recorded. Which expiration dates were included in a csv would change depending on the date, and more strikes would be added to a given expiration date as the date of the csv approached the expiration date.



Name
SPY_2016-02-06.mat
SPY_2016-05-10.mat
SPY_2016-07-09.mat
SPY_2016-09-06.mat
SPY_2016-12-10.mat
SPY_2016-14-07.mat
SPY_2016-14-09.mat
SPY_2016-15-01.mat

SPY_2016-14-07.mat (MAT-file)

Workspace

Name	Value
CDEL	70x30 double
CIV	70x30 double
CIV_past_year	1x252 double
CMARK	70x30 double
expiration_date	'2016-14-07'
historic_price1	1x252 double
historic_volume1	1x252 double
PDEL	70x30 double
PIV	70x30 double
PMARK	70x30 double
strike	70x1 double

	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	70.0850	70.0850	70.0850	70.5900	70.6950	71.0950	70.9000	70.0050	70.0050	70.0050
4	63.9500	64.2250	63.2750	63.2750	63.2750	64.4600	64.5700	64.9650	64.0800	63.8750	63.8750	63.8750
5	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	54.9350	54.9350	54.9350	56.3150	56.6750	57.0700	56.8900	55.9800	55.9800	55.9800
7	0	0	0	0	0	0	51.8000	52.2000	52.0150	51.1050	51.1050	51.1050
8	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0
12	0	35.5450	35.2700	35.2700	35.2700	36.2950	36.4050	37.1850	37.0050	35.2450	35.2450	35.2450
13	0	0	0	0	0	0	0	0	0	0	0	0
14	30.3050	30.6250	30.3500	30.3500	30.3500	31.2350	31.8400	32.1650	31.9600	29.8000	29.8000	29.8000
15	29.2650	29.5200	29.2500	29.2500	29.2500	29.7500	29.8500	30.2500	30.0650	29.2350	29.2350	29.2350
16	0	0	0	0	0	0	0	0	29.9050	29.0050	29.0050	29.0050

```
z =

    55

>> load('SPY_2016-14-07.mat')
```

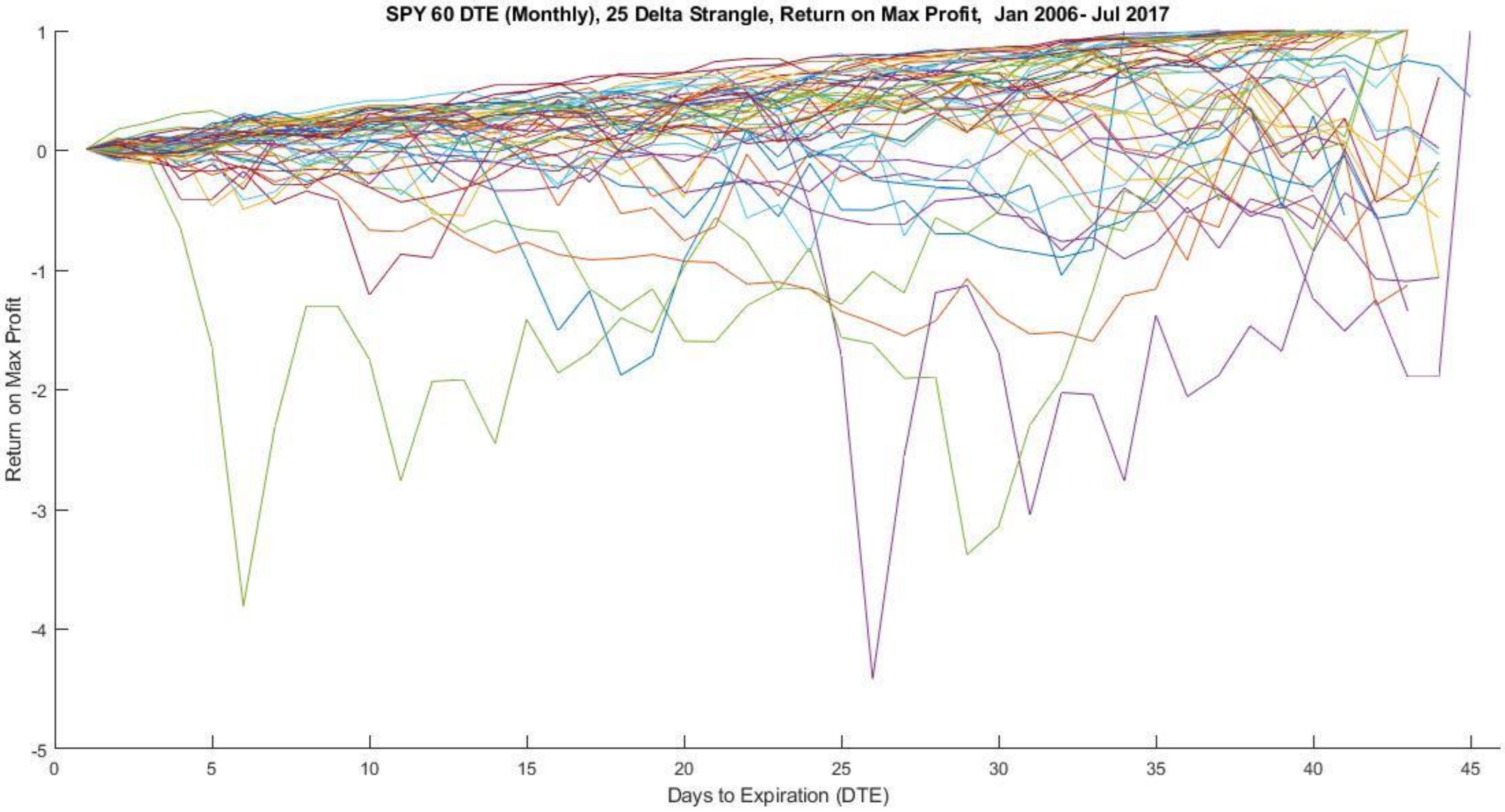
In the current folder section, .mat files labeled with the underlying stock symbol and expiration date are listed. The contents of the SPY 7/14/16 expiration data is displayed in the workspace. The call market prices, labeled CMARK in the workspace, is displayed. The zeros in the matrix correspond to strike prices that had not yet been available in the market. The strike prices are stored in the workspace variable labeled strike. Also included are delta and implied volatility values for both calls and puts, as well as the previous year of closing stock prices and stock volume for the underlying stock, as measured from the first day of data in the file.

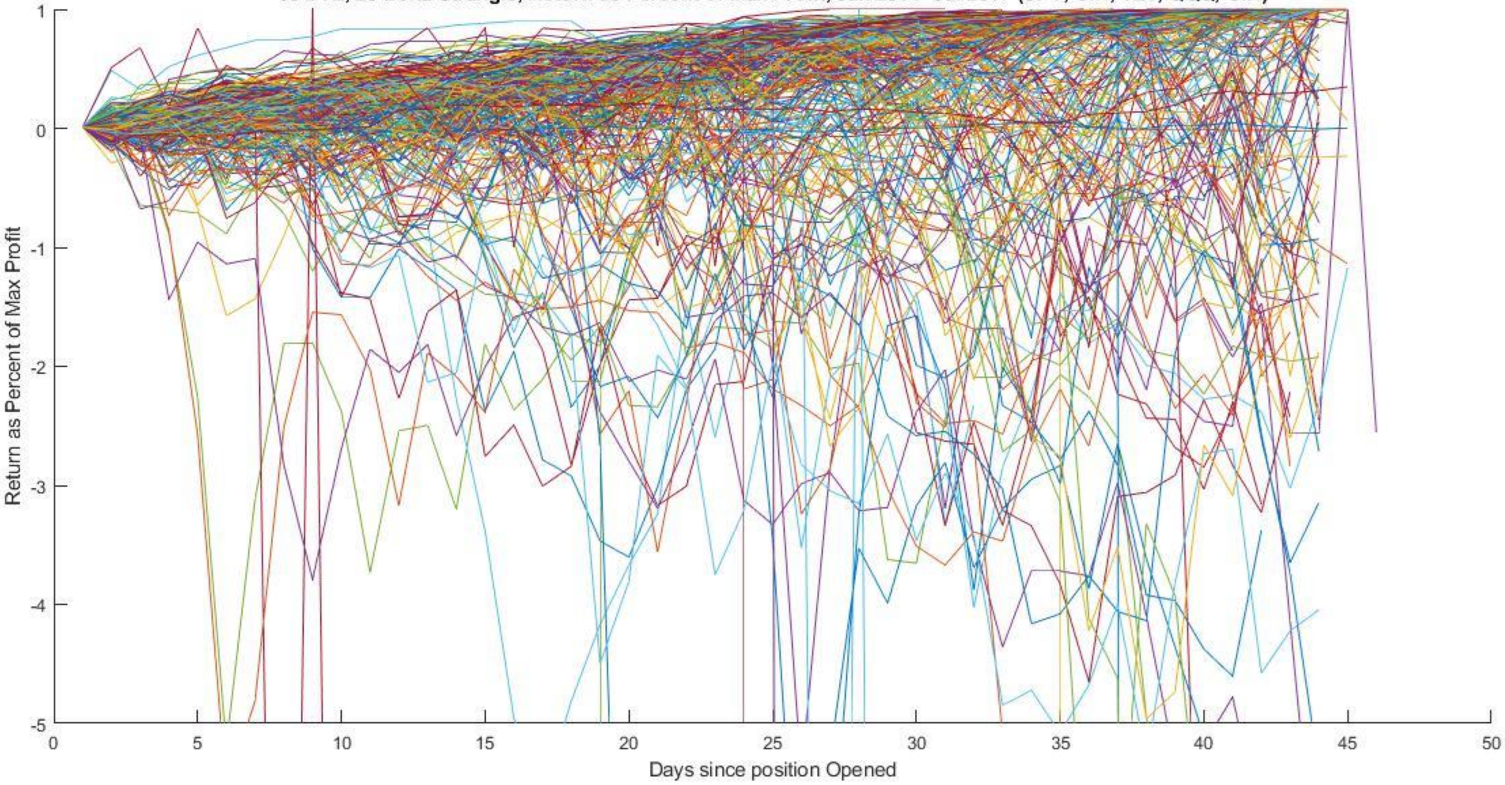
Discussion

Assigning data to a .mat file was achieved by finding all unique expiration dates in a given .csv, then either creating the .mat file if it didn't already exist, or loading the file if it did exist. While this is computationally ineffecent (organizing 1 year of data requires importing data from 365 .csv, and for each .csv file requires loading, saving and closing between 1 and 3 .mat files), it is the only way I could figure out how to do this.

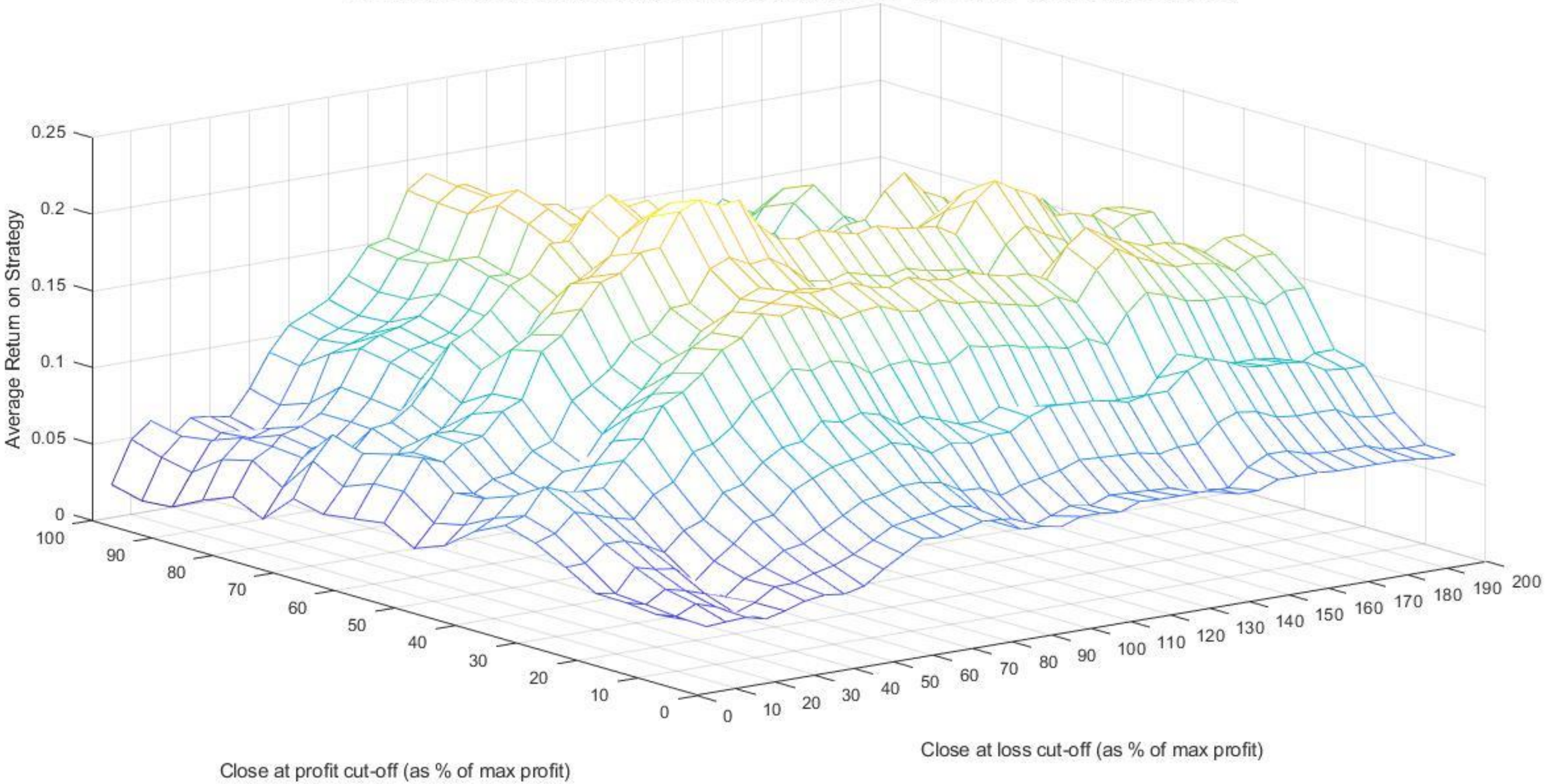
The real difficulty was figuring out how to organize the strikes. Essentially, each .csv could have different strike prices for a given expiration date, so the row indices in one .csv file would not necessarily correspond to the row indices in the next. In order to keep row indices organized, I would create a matrix with a preset range of possible strikes, assign new data to the rows in that matrix, and then once all the data was organized, I would remove rows with all zeros. While this may not be the most efficient way to accomplish this, it was a simple way to keep track of the row indices across .csv files. I was also able to apply this system in my real time option data downloader, which required the strike prices for hundreds of different underlying stocks, all of which would have different ranges.

I also had to figure out how to remove non-trading days from the data, since there wasn't any realistic way to program the auto clicker to skip non-trading days. This was accomplished by checking if the data for one day was identical to the data for the previous day, which would almost certainly not occur between two trading days.

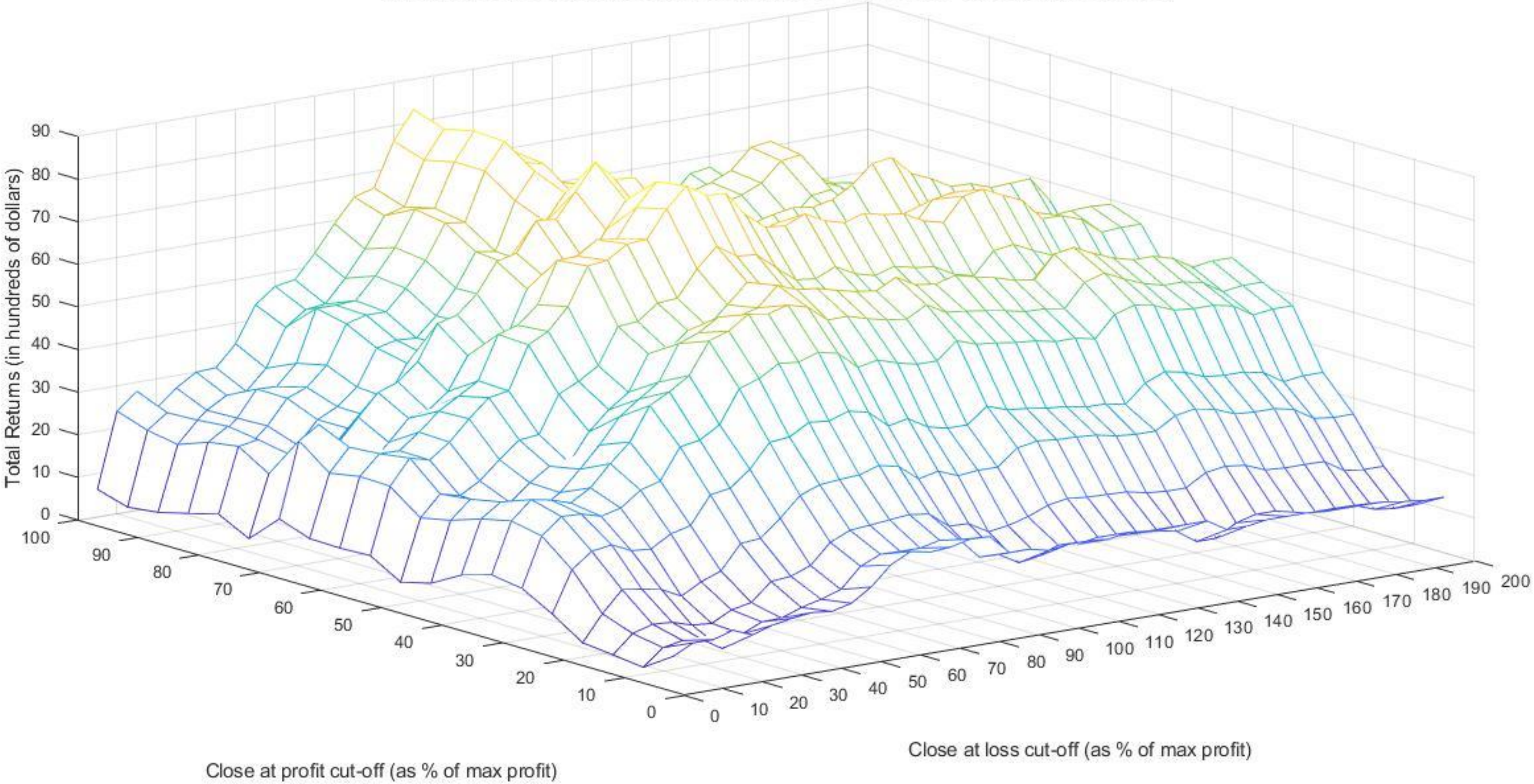

SPY 60 DTE (Monthly), 25 Delta Strangle, Return on Max Profit, Jan 2006- Jul 2017
Return on Max Profit
1
0
-1
-2
-3
-4
-5
0
5
10
15
20
25
30
35
40
45
Days to Expiration (DTE)



60 DTE, 25 Delta Strangle, Return as Percent of Max Profit, Jan 2011- Jul 2017 (SPY, OIH, TLT, QQQ, OIH)



60 DTE, 25 Delta Strangle, Average Percent Return, Jan 2011- Jul 2017 (SPY, OIH, TLT, QQQ, OIH)



60 DTE, 25 Delta Strangle, Total Return, Jan 2011- Jul 2017 (SPY, OIH, TLT, QQQ, OIH)