Assignment 1

Question 1:

Part 1 - generate_yearly_stats.py computes the mean and standard deviation for the sets R, R^- and R^+ of daily returns for my stock (COST) for each day of the week.

Part 2 - *note - all values for means and standard deviations are displayed as percentages (multiplied by 100) and rounded to three digits of accuracy. In the generate_yearly_stats.py file, both adjusted and unadjusted values are available.

2016								
	$\mu(R)$	$\sigma(R)$	$\lvert R^- \rvert$	$\mu(R^-)$	$\sigma(R^-)$	$\lvert R^+ \rvert$	$\mu(R^+)$	$\sigma(R^+)$
Monday	0.227	0.983	17	-0.740	0.928	28	0.822	1.032
Tuesday	0.046	0.792	22	-0.689	0.838	30	0.585	0.757
Wednesday	0.058	1.003	26	-0.719	0.926	26	0.834	1.073
Thursday	-0.170	1.444	32	-0.936	1.305	19	1.120	1.652
Friday	-0.081	1.138	27	-0.872	1.151	24	0.808	1.124

2017								
	$\mu(R)$	$\sigma(R)$	$\lvert R^- \rvert$	$\mu(R^-)$	$\sigma(R^-)$	$\lvert R^+ \rvert$	$\mu(R^+)$	$\sigma(R^+)$
Monday	0.030	1.049	21	-0.797	1.003	25	0.725	1.087
Tuesday	0.216	0.679	17	-0.529	0.671	34	0.589	0.683
Wednesday	0.263	0.808	22	-0.439	0.518	30	0.778	0.967
Thursday	0.142	1.352	24	-0.801	1.290	27	0.980	1.404
Friday	-0.222	1.687	25	-1.207	2.182	25	0.754	1.022

2018								
	$\mu(R)$	$\sigma(R)$	$\lvert R^- \rvert$	$\mu(R^-)$	$\sigma(R^-)$	$\lvert R^+ \rvert$	$\mu(R^+)$	$\sigma(R^+)$
Monday	-0.02	1.389	23	-1.061	1.471	25	0.938	1.31
Tuesday	0.046	1.148	26	-0.769	1.209	25	0.893	1.08
Wednesday	0.108	1.220	25	-0.775	0.995	25	0.991	1.411
Thursday	0.260	1.278	19	-1.013	1.232	32	1.016	1.304
Friday	-0.146	1.841	16	-2.036	3.006	35	0.717	0.899

2019								
	$\mu(R)$	$\sigma(R)$	$\lvert R^- \rvert$	$\mu(R^-)$	$\sigma(R^-)$	$\lvert R^+ \rvert$	$\mu(R^+)$	$\sigma(R^+)$
Monday	0.04	0.845	19	-0.777	1.028	29	0.575	0.7
Tuesday	0.234	1.05	21	-0.647	0.754	31	0.83	1.211
Wednesday	0.131	0.974	18	-0.864	1.208	33	0.673	0.818
Thursday	0.225	0.937	18	-0.822	0.978	32	0.815	0.913
Friday	0.136	1.188	27	-0.661	0.829	24	1.033	1.491

2020								
	$\mu(R)$	$\sigma(R)$	$\lvert R^- \rvert$	$\mu(R^-)$	$\sigma(R^-)$	$\lvert R^+ \rvert$	$\mu(R^+)$	$\sigma(R^+)$
Monday	0.241	2.363	22	-1.368	2.119	26	1.604	2.552
Tuesday	0.452	1.862	19	-1.167	1.482	33	1.384	2.049
Wednesday	0.259	1.347	20	-0.827	1.168	32	0.938	1.448
Thursday	-0.100	1.722	26	-1.316	1.876	25	1.164	1.546
Friday	-0.237	1.683	28	-1.087	1.529	21	0.897	1.869

Part 3 - For COST, there are more days with non-negative returns than there are days with negative returns.

Part 4 - To answer this question I created the file avg_up_and_down.py, which iterated through the csv of all days and sorted them into two lists- ups and downs. I then took the average of each of the lists to determine the average “up” day and “down” day, and found that COST loses more on a down day than it gains on an up day. There were 696 up days with an average return of 0.889%, and 560 down days with an average return of -0.911%.

Part 5 - The results are not the same across days of the week. On average, much more days are “up” days, accounting for 18 of the 25 averages with “down” days accounting for 7 of the 25. The average “up” weekday is roughly 0.173%, whereas the average “down” weekday is roughly -0.122%.

Question 2:

Part 1 - There is a visible pattern that the first half of the week tends to be positive, whereas the second half of the week tends to be negative.

Part 2 - Across the five averages of years, 4 of 5 Mondays, 5 of 5 Tuesdays, and 5 of 5 Wednesdays were positive and 2 of 5 Thursdays and 4 of 5 Fridays were negative.

Part 3 - Best and worst days to invest for each year:

2016 -	Best day: Monday	Worst day: Thursday
2017 -	Best day: Wednesday	Worst day: Friday
2018 -	Best day: Thursday	Worst day: Friday
2019 -	Best day: Tuesday	Worst day: Monday
2020 -	Best day: Tuesday	Worst day: Friday

Part 4 - The best/worst days to invest steadily change throughout the week but show a trend favoring the middle of the week. Best days seem to favor days that have weekdays before and after them, whereas the worst days tend to be either the first or last day of the week.

Question 3:

generate_all_stats.py computes the mean and standard deviation for the sets R, R^- and R^+ of daily returns for my stock (COST) and S&P-500 for each day of the week from 2016 to 2020.

*note - all values for means and standard deviations are displayed as percentages and rounded to three digits of accuracy. In the generate_all_stats.py file, both rounded and unrounded values are available.

Aggregate Table for 2016-2020								
	$\mu(R)$	$\sigma(R)$	$\lvert R^- \rvert$	$\mu(R^-)$	$\sigma(R^-)$	$\lvert R^+ \rvert$	$\mu(R^+)$	$\sigma(R^+)$
COST								
Monday	0.103	1.445	102	-0.967	1.419	133	0.925	1.47
Tuesday	0.199	1.193	105	-0.761	1.054	153	0.858	1.279
Wednesday	0.164	1.09	111	-0.719	0.982	146	0.836	1.165
Thursday	0.071	1.383	119	-0.987	1.406	135	1.003	1.362
Friday	-0.109	1.539	123	-1.094	1.771	129	0.829	1.287
SPY								
Monday	0.051	1.391	99	-0.861	1.712	135	0.721	1.111
Tuesday	0.121	1.146	112	-0.671	0.957	144	0.739	1.282
Wednesday	0.091	1.112	106	-0.737	1.249	151	0.672	1.004
Thursday	-0.002	1.168	117	-0.714	1.434	137	0.605	0.878
Friday	0.056	1.111	111	-0.726	1.112	141	0.671	1.114

Part 1 - For COST, Tuesday is the best day of the week and Friday is the worst day of the week. For SPY, Tuesday is the best day of the week and Thursday is the worst day of the week.

Part 2 - Surprisingly Tuesday is the best day of the week for both stocks, and they share a similar trend of the later half of the week being the worst time to invest.

Question 4:

Part 1 - Following the oracle's advice, my initial $100 with COST would become $45856.69 on the last trading day of 2020.

Part 2 - Following the oracle's advice, my initial $100 with SPY would become $11985.08 on the last trading day of 2020.

Question 5:

Part 1 - With a "buy-and-hold" strategy, my initial $100 with COST would become $263.92 on the last trading day of 2020.

Part 2 - With a "buy-and-hold" strategy, my initial $100 with SPY would become $203.81 on the last trading day of 2020.

Question 6:

Part 1 - oracle_revenge.py computes the results from the 3 scenarios of missing the 10 best days, realizing the 10 worst days, and swapping the 5 best days for the 5 worst days.

	COST	SPY
Oracle followed and correct results	$45856.69	$11985.08
a) Wrong results for 10 best days	$25057.72	$6849.92
b) Wrong results for 10 worst days	$24912.36	$6333.21
c) Wrong results for best 5 days and 5 worst days	$22361.15	$5648.89

Part 2 - For both COST and SPY, you would gain more by missing the best days as opposed to missing the worst days.

Part 3 - The results in part c show a reduction of roughly 50% compared to the results obtained in question 4.