# Executive Summary – Problem Set 1

Ethan Yixiao Wang

September 17, 2024

This project is a data analysis job based on portfolio market return, using Python. Pandas was used to analyze csv files of data and create dataframes.

The first task I accomplished is finding descriptive statistics for each size decile portfolio, in addition to the overall NYSE/Amex/Nasdaq value weighted(**VW**) and equal weighted(**EW)** monthly market return. The statistics I looked for were average, standard deviation, 25$^{th}$ 50$^{th}$ and 75$^{th}$ percentiles, as well as skewness, kurtosis, and normality in terms of skewness and kurtosis. These stats were first generated for the entire time frame of 1926/1–2023/12 in terms of monthly return.

| | mean | std | 25% | 50% | 75% | Skew | Kurt | Normal Skew | Normal Kurt |
|---|---|---|---|---|---|---|---|---|---|
| **VW** | 0.009298 | 0.05335 | -0.01834 | 0.012866 | 0.039376 | 0.127009 | 4.548521 | Normal skew | Abnormal kurt |
| **EW** | 0.011997 | 0.071624 | -0.020608 | 0.014076 | 0.04338 | 1.434828 | 11.801825 | Abnormal skew | Abnormal kurt |
| **DEC1** | 0.016649 | 0.103314 | -0.030125 | 0.008706 | 0.053341 | 2.911448 | 21.062482 | Abnormal skew | Abnormal kurt |
| **DEC2** | 0.012743 | 0.088045 | -0.02643 | 0.011005 | 0.046156 | 1.989091 | 15.525518 | Abnormal skew | Abnormal kurt |
| **DEC3** | 0.011314 | 0.079191 | -0.023149 | 0.010618 | 0.045624 | 1.898908 | 18.031784 | Abnormal skew | Abnormal kurt |
| **DEC4** | 0.010845 | 0.07449 | -0.024898 | 0.012692 | 0.045756 | 1.098939 | 8.921065 | Abnormal skew | Abnormal kurt |
| **DEC5** | 0.010974 | 0.07178 | -0.023185 | 0.012445 | 0.045607 | 0.860957 | 7.949995 | Abnormal skew | Abnormal kurt |
| **DEC6** | 0.01077 | 0.069802 | -0.022918 | 0.013955 | 0.04611 | 0.609754 | 5.823831 | Abnormal skew | Abnormal kurt |
| **DEC7** | 0.010783 | 0.066837 | -0.022955 | 0.014121 | 0.046035 | 0.420876 | 4.774982 | Normal skew | Abnormal kurt |
| **DEC8** | 0.009976 | 0.063043 | -0.022406 | 0.013584 | 0.044344 | 0.263687 | 4.828465 | Normal skew | Abnormal kurt |
| **DEC9** | 0.010367 | 0.059719 | -0.019397 | 0.014031 | 0.043307 | 0.180413 | 4.451826 | Normal skew | Abnormal kurt |
| **DEC10** | 0.008971 | 0.051136 | -0.017289 | 0.01246 | 0.038567 | 0.067216 | 3.981659 | Normal skew | Abnormal kurt |

**For normality, I used the rule of thumb which states that to be normal, skewness is in the range -0.5 to 0.5, and excess kurtosis is in the range -1 to 1.*

I then found data based on specific time frames. First I analyzed returns just for the month of January, and then I excluded January and analyzed February to December.

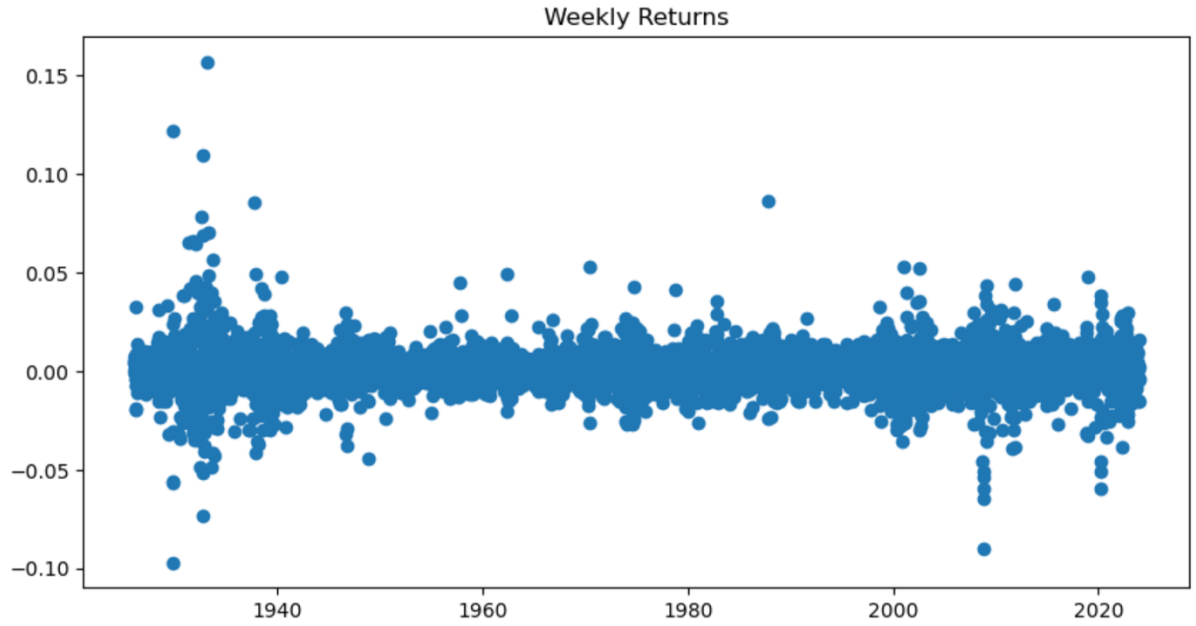
For data only in January:

| | mean | std | 25% | 50% | 75% | Skew | Kurt | Normal Skew | Normal Kurt |
|---|---|---|---|---|---|---|---|---|---|
| **VW** | 0.01492 | 0.047612 | -0.025711 | 0.01285 | 0.05091 | 0.255106 | -3.123942 | Normal skew | Abnormal kurt |
| **EW** | 0.052451 | 0.071533 | 0.005472 | 0.040898 | 0.089098 | 1.016121 | -0.92835 | Abnormal skew | Normal kurt |
| **DEC1** | 0.125989 | 0.127435 | 0.044327 | 0.087863 | 0.165502 | 1.521566 | 0.025105 | Abnormal skew | Normal kurt |
| **DEC2** | 0.084039 | 0.093661 | 0.022691 | 0.06337 | 0.11766 | 1.215687 | -0.714155 | Abnormal skew | Normal kurt |
| **DEC3** | 0.067666 | 0.083271 | 0.015437 | 0.05798 | 0.100829 | 1.375841 | 0.6315 | Abnormal skew | Normal kurt |
| **DEC4** | 0.055669 | 0.078866 | 0.006822 | 0.042505 | 0.091936 | 1.270813 | 0.135967 | Abnormal skew | Normal kurt |
| **DEC5** | 0.045818 | 0.070833 | -0.000207 | 0.040411 | 0.081865 | 1.002427 | -0.830596 | Abnormal skew | Normal kurt |
| **DEC6** | 0.040063 | 0.071105 | -0.007537 | 0.038468 | 0.085763 | 0.908915 | -0.929956 | Abnormal skew | Normal kurt |
| **DEC7** | 0.031943 | 0.066402 | -0.010131 | 0.022298 | 0.075532 | 0.966912 | -0.705494 | Abnormal skew | Normal kurt |
| **DEC8** | 0.024882 | 0.059835 | -0.01862 | 0.017879 | 0.066524 | 0.621835 | -1.802855 | Abnormal skew | Abnormal kurt |
| **DEC9** | 0.021778 | 0.054944 | -0.017102 | 0.021746 | 0.056763 | 0.619291 | -1.774536 | Abnormal skew | Abnormal kurt |
| **DEC10** | 0.010472 | 0.045895 | -0.024173 | 0.009303 | 0.044261 | 0.16941 | -3.310037 | Normal skew | Abnormal kurt |

For data only from February to December:

| | mean | std | 25% | 50% | 75% | Skew | Kurt | Normal Skew | Normal Kurt |
|---|---|---|---|---|---|---|---|---|---|
| **VW** | 0.008787 | 0.053832 | -0.018265 | 0.012866 | 0.038389 | 0.125614 | 4.944683 | Normal skew | Abnormal kurt |
| **EW** | 0.008319 | 0.070523 | -0.02322 | 0.012011 | 0.041247 | 1.526294 | 14.10127 | Abnormal skew | Abnormal kurt |
| **DEC1** | 0.006709 | 0.094846 | -0.035166 | 0.005262 | 0.042154 | 3.392964 | 32.278809 | Abnormal skew | Abnormal kurt |
| **DEC2** | 0.006262 | 0.084632 | -0.030315 | 0.008259 | 0.040646 | 2.194895 | 20.46289 | Abnormal skew | Abnormal kurt |
| **DEC3** | 0.006191 | 0.076825 | -0.026572 | 0.00811 | 0.040604 | 2.045993 | 22.51967 | Abnormal skew | Abnormal kurt |
| **DEC4** | 0.00677 | 0.07276 | -0.027245 | 0.010122 | 0.041953 | 1.094916 | 10.777925 | Abnormal skew | Abnormal kurt |
| **DEC5** | 0.007806 | 0.071055 | -0.02522 | 0.011303 | 0.042317 | 0.87213 | 9.251445 | Abnormal skew | Abnormal kurt |
| **DEC6** | 0.008107 | 0.069102 | -0.023873 | 0.011576 | 0.043448 | 0.585667 | 6.77013 | Abnormal skew | Abnormal kurt |
| **DEC7** | 0.008859 | 0.066574 | -0.023555 | 0.013289 | 0.043853 | 0.37825 | 5.368892 | Normal skew | Abnormal kurt |
| **DEC8** | 0.008621 | 0.063179 | -0.022513 | 0.013063 | 0.042477 | 0.245177 | 5.370639 | Normal skew | Abnormal kurt |
| **DEC9** | 0.009329 | 0.06005 | -0.019456 | 0.013583 | 0.042807 | 0.159927 | 4.861168 | Normal skew | Abnormal kurt |
| **DEC10** | 0.008834 | 0.051603 | -0.016629 | 0.01305 | 0.038078 | 0.062552 | 4.356082 | Normal skew | Abnormal kurt |

My second task was to get a weekly return series using data from the combined NYSE /AMEX /Nasdaq over the period 1926/1/2–2023/12/29. The weekly return was calculated through Wednesday closing prices, which was substituted with Thursday's, then Tuesday's, then a two week return if data is missing.
I created a weekly returns scatter plot to better visualize the data I generated from the code via matplotlib.




*Weekly returns 1926/1/2–2023/12/29.*

Finally, descriptive statistics including skewness, kurtosis, and normality were generated using the same methods and assumptions as previously stated.

| | mean | std | 25% | 50% | 75% | Skew | Kurt | Normal Skew | Normal Kurt |
|---|---|---|---|---|---|---|---|---|---|
| **VW** | 0.001003 | 0.011019 | -0.003569 | 0.001185 | 0.005496 | 0.970892 | 17.799758 | Abnormal skew | Abnormal kurt |
| **EW** | 0.001401 | 0.011195 | -0.002723 | 0.001652 | 0.005489 | 2.076421 | 41.856248 | Abnormal skew | Abnormal kurt |