

# Big Data in Finance: Part IV

## Homework 6: CRSP and Compustat

March 28, 2025

**Instructor:** Kirsten Burr, kburr26@gsb.columbia.edu

**TAs:** Dian Jiao (dj2526@columbia.edu), Varuun Deshpande (vdeshpande25@gsb.columbia.edu)

**Due:** April 3, 2025, by 11:59 pm. Late submission will not be accepted.

**Goal:** Explore the information CRSP and Compustat

**Delivery:** Please submit the homework using the template python file (*submission\_crspcompustat.py*) provided on canvas, along with the output graphs from below, to Gradescope.

Tips for submission: You are required to fill in code wherever you see a “TODO” in the python script. Look at the “main()” in the script to understand the flow, the sequence should match the homework questions below. Please do not change any of the names (of the files, any of the functions, output files and graphs), as well as the order in which values are returned from each function. Feel free to email the TAs if you have any questions or difficulties with this.

### Description

Despite significant market declines in the final weeks of 2024, US markets ended 2024 just shy of record best performances. The S&P 500-stock index gained more than 23% in 2024 after rising 24% in 2023, logging the highest consecutive back-to-back annual gain in nearly 30 years (CNN.com). The blue-chip Dow Jones Industrial Average gained 13% for the year, and the tech-heavy Nasdaq Composite gained almost 30% percent. Let’s look at what happened with stock returns using the data downloaded from CRSP and Compustat.

### Part I: CRSP stock returns

1. What was the cumulative average returns (market-cap weighted and equal-weighted) of CRSP stocks in 2024? (i.e. how much would an investor have by the end of 2024 had she invested \$1 in CRSP stocks at the beginning of 2024?)

2. Plot the market-cap weighted and equal-weighted cumulative average returns of CRSP stocks from the beginning of the year.
3. How much would the same investor have had she entered the market at the end of June, during which stock volatility was relatively higher?
4. What were the 10 top performing stocks in 2024?

Note: You should only consider ordinary common stocks (i.e., shrcd equal to 10 or 11) that are listed in one of the three main stock exchanges (i.e., exchcd equal to 1 (NYSE), 2 (NYSE MKT) or 3 (NASDAQ)). Remember to use the prior month's market-cap when market-weighting returns.

## Part II: Tech stock returns

1. How much would an investor have by the end of 2024 had she invested \$1 in a TECH portfolio (market-cap weighted and equal-weighted) at the beginning of 2024?
2. Plot the market-cap weighted and equal-weighted cumulative average returns of the Tech portfolio from the beginning of the year.
3. How much would the same investor have had she entered the market at the end of June?
4. What are the 10 top performing Tech stocks in the year of 2024?

Note: The closest definition of Tech firms according to Fama and French is Business Equipment industry, which includes Computers, Software, and Electronic Equipment. These are firms with SIC codes in the following ranges:

3570–3579

3660–3692

3694–3699

3810–3829

7370–7379

## Part III: High profitability stock returns

1. Given the fundamental data available to investors as of December 31, 2023, what are the 10 most profitable companies in 2023? Use operating profitability normalized by book-equity (OPBE) from the lecture notes as a measure of profitability.
2. How much an investor would have by the end of 2024 had she invested instead in a portfolio (market-cap weighted and equal-weighted) of these 10 most profitable firms' stocks at the beginning of 2024?
3. Compare your results to those from Part II. Which portfolio performed better? Why do you think this is the case?

4. Plot the market-cap weighted and equal-weighted cumulative average returns from the beginning of the year.
5. How much would the same investor have had she invested in these same 10 firms at the end of June?

Note: Publicly-listed companies must report their financials within 91 days after the end of the fiscal period. Most companies report on the deadline. Consider this when merging financial results to stock performance.