

Ali Zamani

1. Scraping SEC for inside trading information (Must do).

<https://github.com/AliZamani1995/scraping-SEC>

2. Data engineering and analyzing (Nice to have)

- a. The transaction with transactionPricePerShare=0 is removed.
- b. The average of transactions in each day is calculated.

AAPL		
	transactionShares	transactionPricePerShare(\$)
count	463	463
mean	3.569712e+04	212.180192
std	1.407571e+05	146.145895
min	1	7.475000
25%	2.882417e+03	117.635000
50%	9.461000e+03	153.630000
75%	2.500000e+04	241.273750
max	2.653560e+06	700.095000

AMZN		
	transactionShares	transactionPricePerShare(\$)
count	216	216
mean	4056.157254	1751.102345
std	9804.777329	1073.875966
min	2.222222	94.000000
25%	177.125000	940.162556
50%	401.678322	1690.274550

75%	2218.731343	2842.502095
max	61667.500000	3717.290000

NVDA		
	transactionShares	transactionPricePerShare(\$)
count	402	402
mean	4.607081e+04	105.627642
std	9.051567e+04	152.424762
min	1.500000e+01	3.155000
25%	3.087188e+03	13.842500
50%	1.511662e+04	18.600000
75%	6.350700e+04	166.063000
max	1.200000e+06	774.240000

TSLA		
	transactionShares	transactionPricePerShare(\$)
count	40	40
mean	1.692128e+05	661.409590
std	2.691568e+05	437.393091
min	3.684211e+02	17.000000
25%	2.455725e+04	165.960838
50%	5.869752e+04	888.549100
75%	9.678722e+04	1017.774400
max	1.068333e+06	1146.218600

The blue line is from <https://www.sec.gov> website.

The red line is from the Yahoo finance website.

Why are they different? [Cite: ChatGPT]

There can be a few reasons why the stock price listed on the SEC's website may be different from the real-time stock price. Here are a few possibilities:

- a. Time delay: The SEC's website may have a time delay in updating the stock prices. The stock prices listed on the website may be delayed by several minutes or even hours. Real-time stock prices, on the other hand, are updated constantly throughout the trading day.
- b. Different data sources: The SEC's website may use a different data source than the real-time stock price providers. This can lead to differences in the stock prices listed on the website.
- c. Different trading platforms: The stock prices listed on the SEC's website may be based on trading on a specific exchange, whereas real-time stock prices may be based on multiple exchanges. The differences in trading platforms can lead to differences in stock prices.
- d. Bid-ask spread: The difference between the highest bid price and the lowest ask price for a stock is known as the bid-ask spread. The stock prices listed on the SEC's website may be based on the bid or ask price, whereas real-time stock prices are based on the last traded price, which may be somewhere in between the bid-ask spread.

It's important to note that the stock prices listed on the SEC's website are generally accurate and reliable, but they may not always match real-time stock prices due to the reasons mentioned above.



