

MEDIAN OF PEER MULTIPLES

The Nasdaq data, and what the peers say a company is worth

Kerry Back

Thu Aug 27, 2026

Today

- 1 The data
- 2 Defining peers
- 3 Multiples
- 4 Exercise

THE DATA

The Sharadar Core US Equities Bundle

What is in it

Every NYSE, NASDAQ and NYSEMKT common stock, 2000 to today.

Where it comes from

Licensed by Rice Business from Nasdaq Data Link.

What it is

Nine tables online at motherduck.com

data.nasdaq.com/databases/SFA — the Sharadar Core U.S. Equity Bundle

Ask AI to do it. It will read `~/workspace/.claude/skills/finance-data/SKILL.md` to see how to do it.

The SKILL.md explains where the database lives, which table answers which question, and the traps in each one.

Ask for US stocks and it uses the bundle. Ask for something the bundle does not carry — short interest, bond yields, ... — and it routes to a free public source instead.

How You Reach It

What Is in There

TABLE	CONTENTS
sep	Daily prices — 33.4M rows, 2000 to today
daily	Daily market cap and valuation ratios
sf1	10-K and 10-Q data, 112 fields, plus precomputed ratios
sf2	Insider trades, from Form 4
tickers	Identity, classification, exchange, listing dates
actions	Mergers, splits, dividends, ticker changes
events	8-K event codes

The Table That Describes the Tables

indicators is the data dictionary: one row for every column in the bundle, with a definition.

It also documents some Sharadar tables that are not in the MotherDuck database — 13F institutional holdings, ETF prices, S&P 500 index membership.

	ANNUAL	QUARTERLY	TRAILING FOUR QUARTERS
As most recently restated	MRY	MRQ	MRT
As originally reported	ARY	ARQ	ART

Today we use **MRY** — we are describing companies as they stand. But in every **MR** row **datekey** is the period-end date copied over, so **MR** does not know when the numbers became public. Backtests use **ARY**, keyed on **datekey**, which is the real SEC filing date.

sf1 Has Dimensions

Which Stocks Are You Even Looking At?

FILTER	COMPANIES LEFT
Every ticker in <code>tickers</code>	15,607
Still listed — <code>isdelisted = 'N'</code>	4,355
Drop shell companies — 315 SPACs with no operations	4,039

Two filters, because the bundle has already done the rest. It holds only NYSE, NASDAQ and NYSEMKT common stock, and share classes are already collapsed — GOOG is not a row, it sits in GOOGL's `relatedtickers`. Against raw Sharadar you would have to strip ADRs, Canadian listings, preferred and secondary classes yourself.

closeadj and Periodic Returns

The skill tells the AI how to compute periodic returns using the `closeadj` variable, which adjusts for dividends as well as splits (like Yahoo Finance adjusted closing price).

So, we can say, for example, “compute monthly returns for all stocks from date1 to date2.”

PEER GROUPS

Six Ways to Say “Industry”

COLUMN	GROUPS	SCHEME
sector	11	Morningstar-style sectors
industry	151	Morningstar-style industries
sicsector	10	The SEC’s SIC divisions
sicindustry	437	SIC industry titles
siccode	437	The four-digit code itself
famaindustry	48	Fama and French’s 49 industries
All six sit in tickers.		

How Large are Peer Groups

GROUPING	GROUPS	MEDIAN GROUP	MEDIAN CELL	CELLS UNDER 10
sector	11	229	51	2%
famaindustry	49	50	10	50%
industry	143	16	4	84%
sicindustry	379	4	2	93%

The 4,039 live, non-shell companies, crossed with market-cap quintiles built from `daily.marketcap`. Note that nineteen companies have a blank `famaindustry` and will form a peer group of their own if not dropped.

FOUR MULTIPLES

What You Multiply by What

MULTIPLE	COLUMN IN <code>daily</code>	COMPANY'S OWN FIGURE IN <code>sf1</code>	GIVES YOU
Price / earnings	<code>pe</code>	<code>netinccmnusd</code>	Equity value
Price / book	<code>pb</code>	<code>equityusd</code>	Equity value
Price / sales	<code>ps</code>	<code>revenueusd</code>	Equity value
EV / EBITDA	<code>evebitda</code>	<code>ebitdausd</code>	Enterprise value

Value It Four Ways

- 1 Fix the peer set — same industry, same size bucket, same date — and count it
- 2 Take the median of each of the four multiples across those peers
- 3 Multiply each median by the company's own denominator
- 4 For EV/EBITDA, cross back from enterprise value to equity

Four implied equity values for one company. Compare each with the market cap it actually trades at.

What a P/E of -12 Means

Nothing. It means the company lost money.

A negative P/E is not a cheap stock. It is a stock with no E. Left in the sample it drags the median in the wrong direction and makes the cheapest-looking companies the ones losing the most.

EV/EBITDA fails the same way on negative EBITDA. P/B is close to meaningless where the assets walk out of the building at six o'clock — and it means something else again for a bank, where the balance sheet *is* the business. P/S always exists, which is not the same as always meaning something.

Filter out the negative values. We won't try to value them or to use them for valuing others.

Why the Median

THE MEAN

One company with a P/E of 900 — a real company that barely broke even — and the average for the industry is meaningless.

THE MEDIAN

Unmoved by the tail. With ratios, where the numerator is bounded below and the denominator can approach zero, this is not a refinement. It is the difference between a number and noise.

Report the count alongside every median. A median of four is a different object from a median of four hundred, and the table does not tell you which you have.

Deviations

Deviation = $\log(\text{market cap} / \text{implied value})$.

Positive means the market pays more for this company than its peers say it is worth.

Negative means the company is cheap compared to peers.

ORGANIZING YOUR WORK

CLAUDE.md

What is always true in this folder: what the project is, where things live, the conventions, and what never to do.

It is read at the start of every session in that folder, before you type anything.

Three people, three containers, one repository. `CLAUDE.md` is how a decision one of you makes on Thursday is still in force when somebody else opens the folder on Sunday. Everything you type into a prompt window instead dies with that conversation.

WHEN YOUR GROUP SETTLES SOMETHING:

Add to `CLAUDE.md`: peers are `famaindustry` crossed with market cap quintiles; minimum group size is eight.

You never open the file yourself. You say what you decided and the AI writes it down.

Tell Claude to Edit It

Keep a Record of Scripts and Outputs

PUT THIS IN `CLAUDE.md`

After running any script, append a row to `RUNS.md`: who ran it, the date and time, the script, and where the output went.

Keep the log in its own file. `CLAUDE.md` is read at the top of every session, so a hundred rows of history there is a hundred rows you would pay for on every prompt.

Use AI to Pull, Commit, and Push

Let GLM run Git for you. It will write better commit messages and resolve merge conflicts.

Data Files

- Keep your data files in your workspace so you don't waste time getting the data each time
- Parquet is the fastest file type for reading and writing and also the smallest
- Tell GLM to gitignore them. This means they won't be pushed to GitHub.
- Pushing large files to GitHub is slow, and GitHub rejects files > 100MB.

EXERCISE

Today's Work

- 1 Build the peer table — universe, industry grouping, size buckets, and the count in every cell
- 2 Take the four medians in each cell, and value every company four ways
- 3 Which median multiple matches the market the best?
- 4 Why do medians deviate from the market — is it asset turnover, margins, leverage, investment rate, ... — using the most recent 10K (sf1 with dimension = MRY)?

Step 1 is the one that takes the longest and the one everything else rests on. Do not let it be the step you rush.

The peer definition

Which grouping, which size buckets, what minimum count — and the cell counts that justify it.

Coverage

How many of the companies got a value under each of the four multiples, and which multiple lost the most.

What the deviations track

Profitability, growth, leverage — or nothing.

The disagreement case

The company where the four methods were furthest apart, and your diagnosis.

Five minutes, from the repository. Three groups drawn at random — so all of you are ready.

What to Present Tuesday