Smarter Stock Selections

Data Science for Investing

DISCLAIMER

My background is

Mathematics

Executive Pay

Data Science

Quantum Computing

ALL of my investment knowledge comes from books, lectures, conversations, self-study, reading financial statements, learned wisdom, painful experience, and lots of thinking

Tips for Investing

1) Invest In Yourself

   a) Learn new skills, increase your earnings potential

   b) Unlock new opportunities, protect yourself from competition

   c) Good health and exercise today (free) is worth (???) tomorrow

2) Take the 401(k) company match, always

   a) Consider maxing out your 401(k) / Roth contributions

   b) Same for ESPP (if offered)

3) Itemize your weekly, monthly, daily spending

   a) You spend far more on eating out / drinks / other nonessentials than you realize

4) Automate your savings (before you have a chance to miss it)

5) If you have extra cash, invest in a business!

Investing in a Business

1. Start your own business

100% ownership, 100% of the risk

Requires **time**, dedication, commitment

You work for yourself (and the IRS)

2. Buy ownership in a private business

Can be expensive ($10K, $100K, $1M)

You may be expected to **contribute**

Illiquid, hard to extract cash on demand

3. Buy ownership in a public business (shares of public stock)

Highly liquid ownership, very large market of buyers and sellers

You will NOT control the business, simply enjoy the upside (or downside)

Requires **very little time** or **lots of time**. MUCH CONFLICTING INFORMATION

Common Stock Valuation Methods

Recommended	
Discounted Cash Flows (DCF)	**Ratios and Multiples**
Theoretically sound, actual parameters fuzzy	P/E Ratio, price-to-sales, price-to-book
Predicts the **today** value of **future** *free cash*	Good for benchmarking, not quick decisions
Problem of "tuning to the answer"	Feels like signal, but sometimes...

| Not Recommended	
Industry/Technology Forecasts	**Pure Gut Feelings, Intuitions**
"Autonomous driving will be valuable"	*"I like the product, company must be good!"*
Of course, but HOW valuable? $10B $100B? $1T?	A kick in the Pavlov bells
Except Ford and Tesla, *every* American automaker has gone bankrupt	Not trustworthy
	• ...but sometimes?

Buy and Hold

Buy low ~~/ sell high~~ **AND HOLD**

Buying and selling investments
requires you to be right **twice**:

1) **Buy** at the right time
2) **Sell** at the right time

Why not buy a great business at a
good price and hold while it grows?
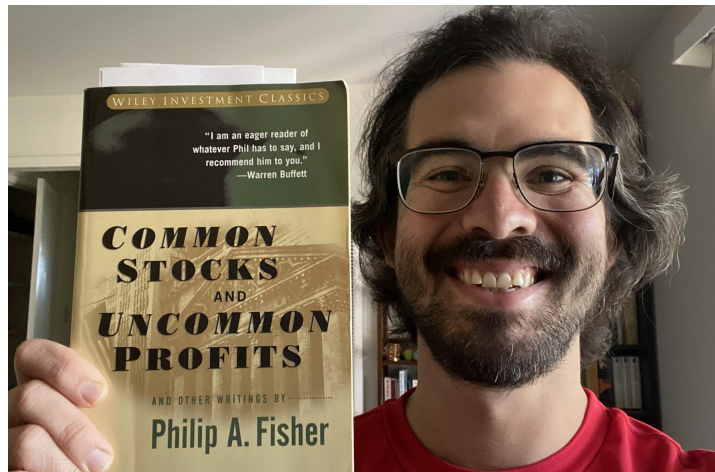
"Even in those earlier times [1920s, 1930s], **finding the really outstanding companies and staying with them** through all the fluctuations of a gyrating market **proved far more profitable** to more people than did the more colorful practice of trying to buy them cheap and sell them dear."

-Phil Fisher, *Common Stocks and Uncommon Profits*

Modeling the Market

Building Financial Data Models, EDA

Initial list: 3,600+ publicly-traded companies (stocks)

Quarterly (TTM) data from stockrow.com, based on ticker:

- Income Statement
- Balance Sheet
- Cash Flow Statement

Python won't create dynamic variables from a list

e.g. aapl_df, msft_df, intc_df

Instead used dictionaries for all dataframes, and meta-dictionaries, corresponding with each ticker

Data was relatively clean and uniform

Not every company reports every metric

For modeling purposes, extracted the most common (and useful) financials before combining data

Final list: 2,889 companies / 9-year average / 15 features / 104K rows

Classifying Industry "Sector"

There are 11 GICS sectors; do company financials align with these sectors?

Are sectors predictable?

Random Forest Model

With 50% train / test split, the model accurately predicted a majority of industry sectors

Sector Prediction	RF Train Score	RF Test Score
Tickers included	100%	96.9%
Without tickers	98.9%	91.2%

KMeans Model

Set to N = 11 clusters, these do not seem to naturally correspond with industry sectors, including 3 clusters with 1 (one) company each

Cluster label		Actual Sector Count	
5 ---- 72,547		Financials ------------------	19,571
1 ---- 11,545		Industrials ------------------	16,352
2 ---- 8,064		Information Technology --	15,257
7 ---- 5,743		Health Care ----------------	14,775
4 ---- 4,052		Consumer Discretionary -	13,026
8 ---- 1,587		Materials --------------------	4,959
0 ---- 608		Real Estate ----------------	4,724
6 ---- 160		Energy ----------------------	4,671
10 ---- 39		Communication ----------	4,382
9 ---- 39		Consumer Staples ------	4,196
3 ---- 39		Utilities --------------------	2,510

Visualizing KMeans Clusters

Clustering companies based on financials
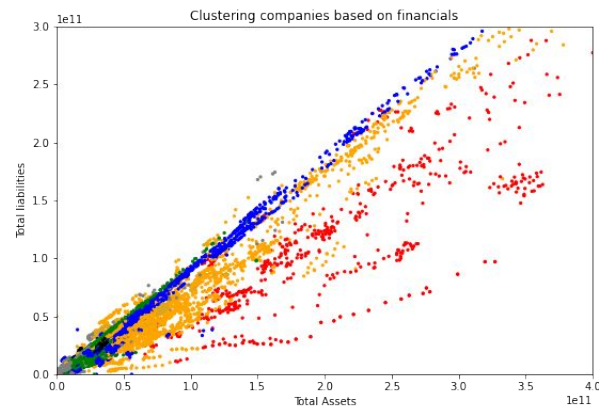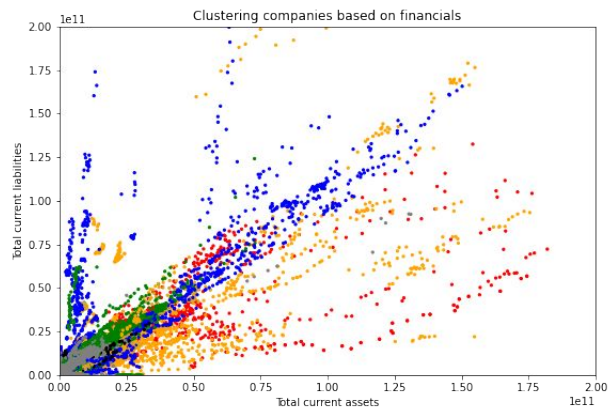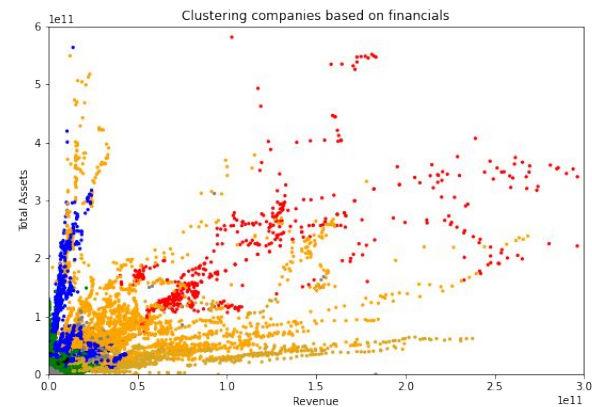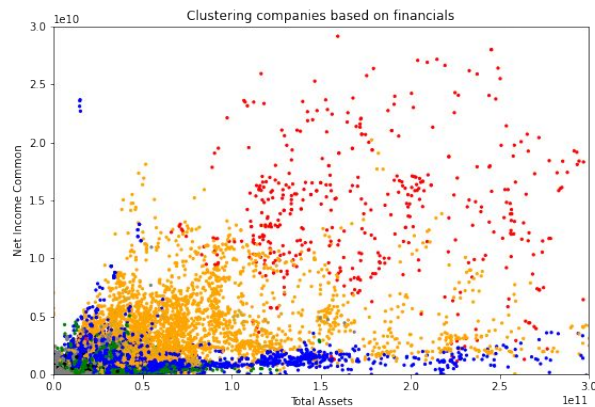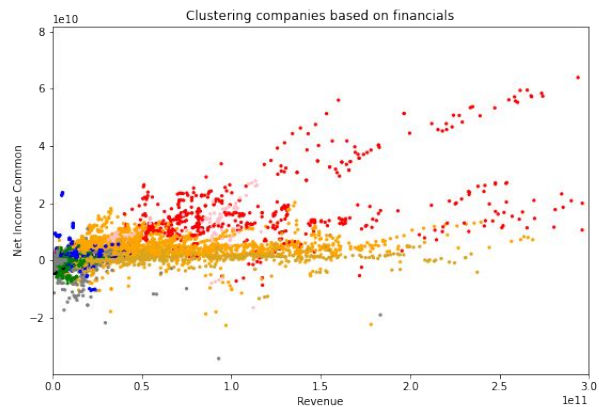
Valuing a Business

Discounted Cash Flows

The idea of *intrinsic value*:

The **present** value of the **future** stream of cash that an asset will give off between now and the endtimes

If a business generates $100M of "free" cash this year, **how much will it generate next year**, the year after, and so on?
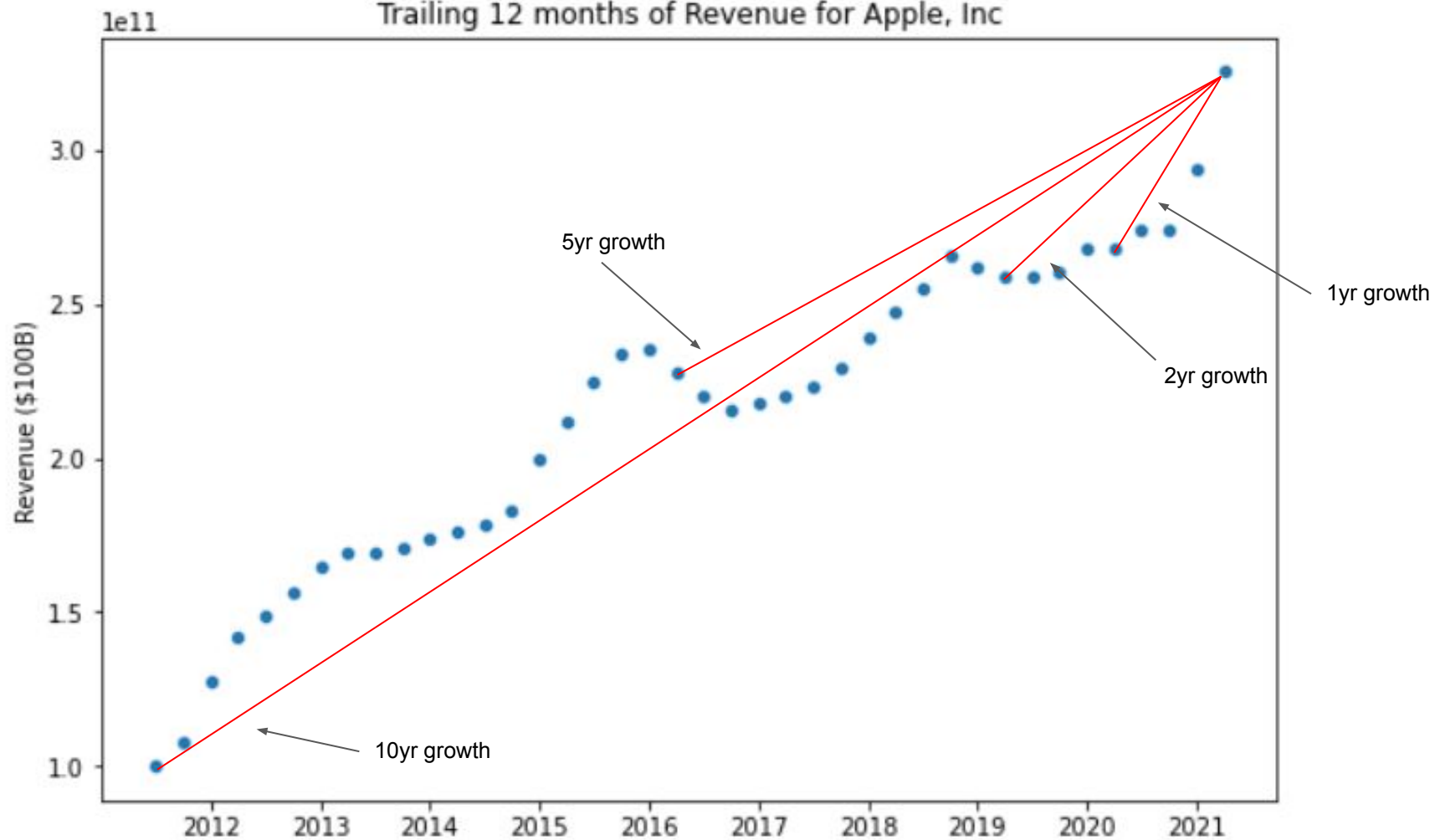
Requires a *growth* estimate. Revenue growth? Earnings growth? Dividends? Assets? Equity? ~~Stock price?~~
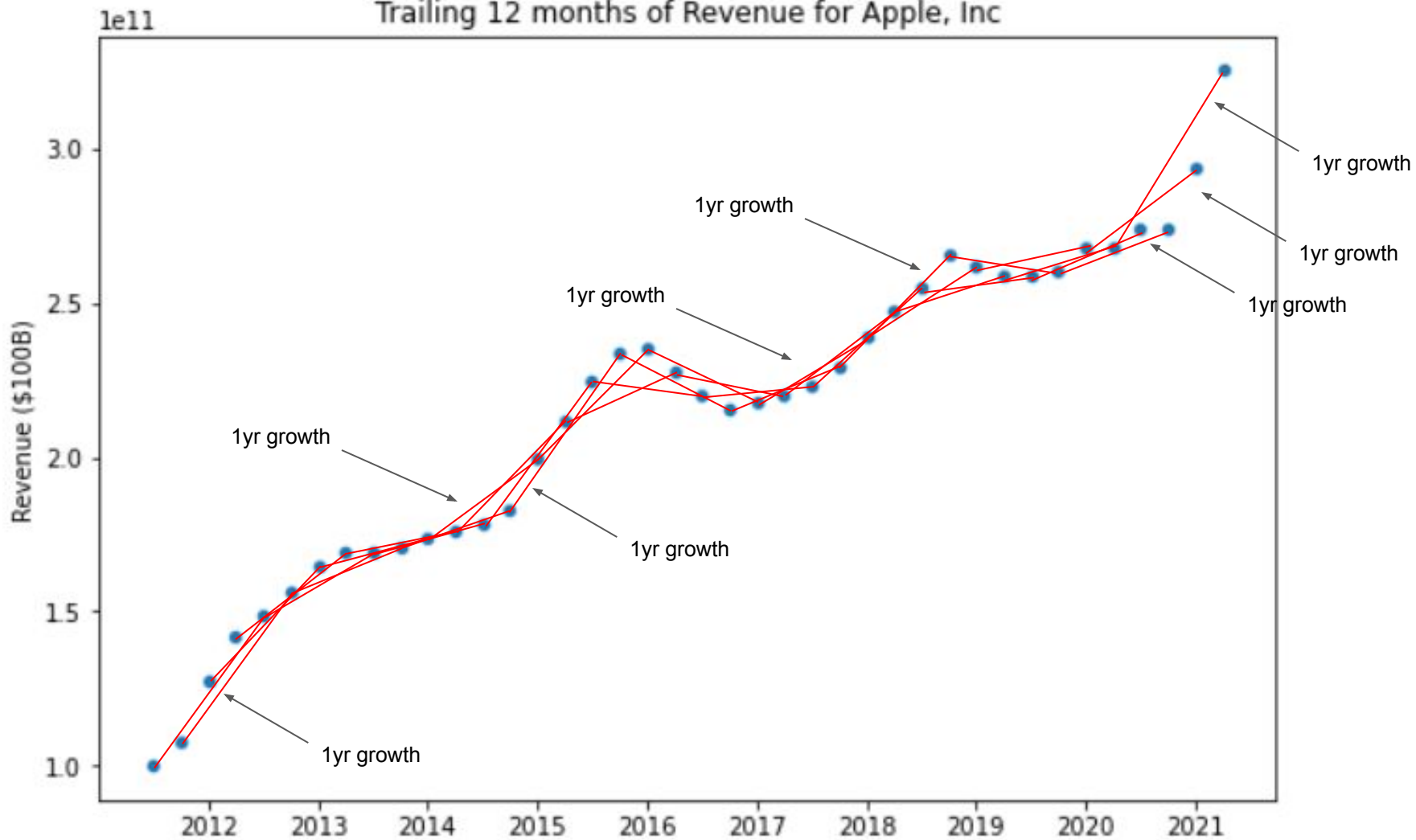
Year 1	$10,000
Year 2	$14,142
Year 3	$20,000
Year 4	$28,284
Year 5	$40,000
Year 6	$56,569
Year 7	$80,000
Year 8	$113,137
Year 9	$160,000
Year 10	$226,274
Year 11	$320,000
Total	**$1,068,406**

Example of 41% YOY growth

Trailing 12 months of Revenue for Apple, Inc

Trailing 12 months of Revenue for Apple, Inc

To the notebook!

Future / Ongoing Work

- Fix DCF/FCF models
  - Error handling for negative values and other issues
  - Possibly incorporate WACC as discount rate
  - Possibly incorporate beta (stock volatility)
  - Auto-update treasury rates
- Compare DCF projections to actual stock prices
  - I don't expect this to ever match; efficient markets are only roughly right
- Pull data directly from SEC.gov
  - Unlocks a much longer potential history of data
  - Currently not a simple or trivial task
- Reach out to stockrow.com for collaboration opportunities

Thank you!

To General Assembly

To Chuck and Varun and Grant

To my DSI cohort classmates

To everyone who contributed and shared thoughts

For more please visit:

https://github.com/sfreagin/SEC_testing

https://www.stockrow.com

[This page left intentionally blank]