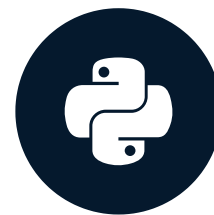


Cash flow statement

ANALYZING FINANCIAL STATEMENTS IN PYTHON



Rohan Chatterjee

Risk Modeler

Show me the money!

Accrual method of accounting

- Transaction recorded in income statement when it takes place, not when cash is exchanged

Cash flow statement

- Tells us business' cash and cash equivalents in a (financial) period
- Records transactions when cash is exchanged
- Entries in cash flow statement are linked to entries of the income statement and balance sheet
- However, it is still a distinct financial statement

Types of cash flow statement

Two ways of preparing a cash flow statement

- **Indirect method:** Based on the accrual method of accounting - accountant starts with net income (or loss) and adjusts up or down depending on transactions done in cash
- **Direct method:** Recording cash transactions when they take place

Since most businesses follow accrual-based accounting, we will focus on the **indirect** method.

Structure of cash flow statement

The cash flow statement is usually broken down into three parts:

- Cash from **operating** activities
- Cash from **investing** activities
- Cash from **financing** activities

Cash from operating activities

- Focuses on cash flows in the core business
- **Accounts receivable**
 - Customers to pay at a later date
 - Increase in accounts receivable is a cash outflow because more customers bought products on credit
- **Accounts payable**
 - Increase in accounts payable is a cash inflow

Cash flow statement of ABC	
In Thousands of US Dollars	
Cash flow from operations	
Net Income	2000000
Increase in accounts recievable	(15000)
Increase in accounts payable	1000
Net cash from operating activities	1986000

Cash flow from investing activities

- Focuses on cash generated (or lost) from investing activities of the company
- Investments that pay cash will be a cash inflow
- Purchase of property, plant, and equipment (with cash) is a cash outflow

Cash flow statement of ABC	
In Thousands of US Dollars	
Cash flow from investing	
Cash generated from investments	200
Purchase of Property, Plant and Equipment	(300)
Net cash from investing activities	(100)

Cash flow from financing activities

- Focuses on cash generated (or lost) from financing activities of the company
- Dividends paid in cash to the investors of the company is a cash outflow

Cash flow statement of ABC	
In Thousands of US Dollars	
Cash flow from financing	
Dividends paid	(150)
Net cash from investing activities	(150)

Cash flow statement all together

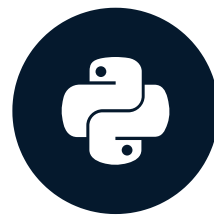
Cash flow statement of ABC In Thousands of US Dollars	
Cash flow from operations	
Net Income	2000000
Increase in accounts recievable	(15000)
Increase in accounts payable	1000
Net cash from operating activities	1986000
Cash flow from investing	
Cash generated from investments	200
Purchase of Property, Plant and Equipment	(300)
Net cash from investing activities	(100)
Cash flow from financing	
Dividends paid	(150)
Net cash from investing activities	(150)
<u>Net cash flow</u>	1985750

Let's practice!

ANALYZING FINANCIAL STATEMENTS IN PYTHON

Financial ratios from the cash flow statement

ANALYZING FINANCIAL STATEMENTS IN PYTHON



Rohan Chatterjee
Risk Modeler

Reading in JSON data

- Data from the wild does not always come in spreadsheets
- Sometimes it comes in the `JSON` ("JavaScript Object Notation") format
- Companies can share their financial statement information in `JSON`
- We can read `JSON` files into Python using `pandas`

```
cash_flow = pd.read_json("cash_flow_statement.json")
print(cash_flow.head())
```

	Capital Expenditures	Change In Cash	Change To Account Receivables
0	-10495000.0	24311000.0	245000.0
1	-7309000.0	-10435000.0	6917000.0
2	-11085000.0	-3860000.0	-10125000.0
3	-10708000.0	-10952000.0	-1823000.0
4	-13925000.0	-590000.0	-2812000.0

Cash flow to net income ratio

- Proportion of cash flow from operating activities to net income
- Operating activities are core activities of the business
- High ratio implies business generates sizable proportion of cash from operating activities

Formula:

$$\frac{\text{Cash flow from operating activities}}{\text{Net income}}$$

Operating cash flow ratio

- Proportion of cash flow from operating activities to current liabilities
- Measure of how many times company can pay off short-term obligations from cash generated from core business
- Ratio of more than one implies that a business generates more than enough cash to meet its short-term obligations

Formula:

$$\frac{\text{Cash flow from operating activities}}{\text{Current liabilities}}$$

Imputing missing values

- Data in "the wild" often has missing values
- Data from the numerator of a ratio might be available, but its denominator might be missing, or vice-versa
- Solution: impute missing data with data from other companies

Imputing missing values

- In the DataFrame named `dataset` shown, some entries of "Total Current Liabilities" are missing, indicated by `NaN`

Year	company	Total Current Liabilities	comp_type
2019	AAPL	NaN	tech
2020	AAPL	NaN	tech
2021	AAPL	1.255e+11	tech
2022	AAPL	1.540e+11	tech
2019	MSFT	6.942e+10	tech
2020	MSFT	NaN	tech
2021	MSFT	8.866e+10	tech

- Missing current liabilities for a company can be imputed using non-missing values for that company

Imputing missing values

- We fill in missing values with the average of non-missing values of the companies:

```
imputation = dataset.groupby("company")["Total Current Liabilities"].transform("mean")

dataset["Imputed Total Current Liabilities"] = dataset["Total Current Liabilities"].fillna(imputation)
```

- After imputing, `dataset` looks like:

Year	company	Total Current Liabilities	comp_type	Imputed Total Current Liabilities
2019	AAPL	NaN	tech	1.397e+11
2020	AAPL	NaN	tech	1.397e+11
2021	AAPL	1.255e+11	tech	1.255e+11
2022	AAPL	1.540e+11	tech	1.540e+11
2019	MSFT	6.942e+10	tech	6.942e+10

- Take percentiles to be conservative

Imputing missing values with percentiles

- Imputing a missing value with its 70th percentile worse non-missing value will give a more conservative imputation for it.
- Computing ratios with the more conservative imputation might be more prudent if the ratio is to be used in decision-making.
- Imputing missing values using 70th percentile, grouped over `company` :

```
imputation = dataset.groupby("company")["Total Current Liabilities"]\
    .transform(lambda x: np.nanquantile(x, 0.7))

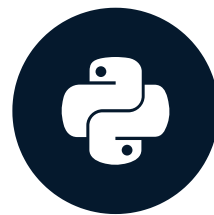
dataset["Imputed Total Current Liabilities"] = dataset["Total Current Liabilities"]\
    .fillna(imputation)
```

Let's practice!

ANALYZING FINANCIAL STATEMENTS IN PYTHON

Visualizing financial ratios for comparison

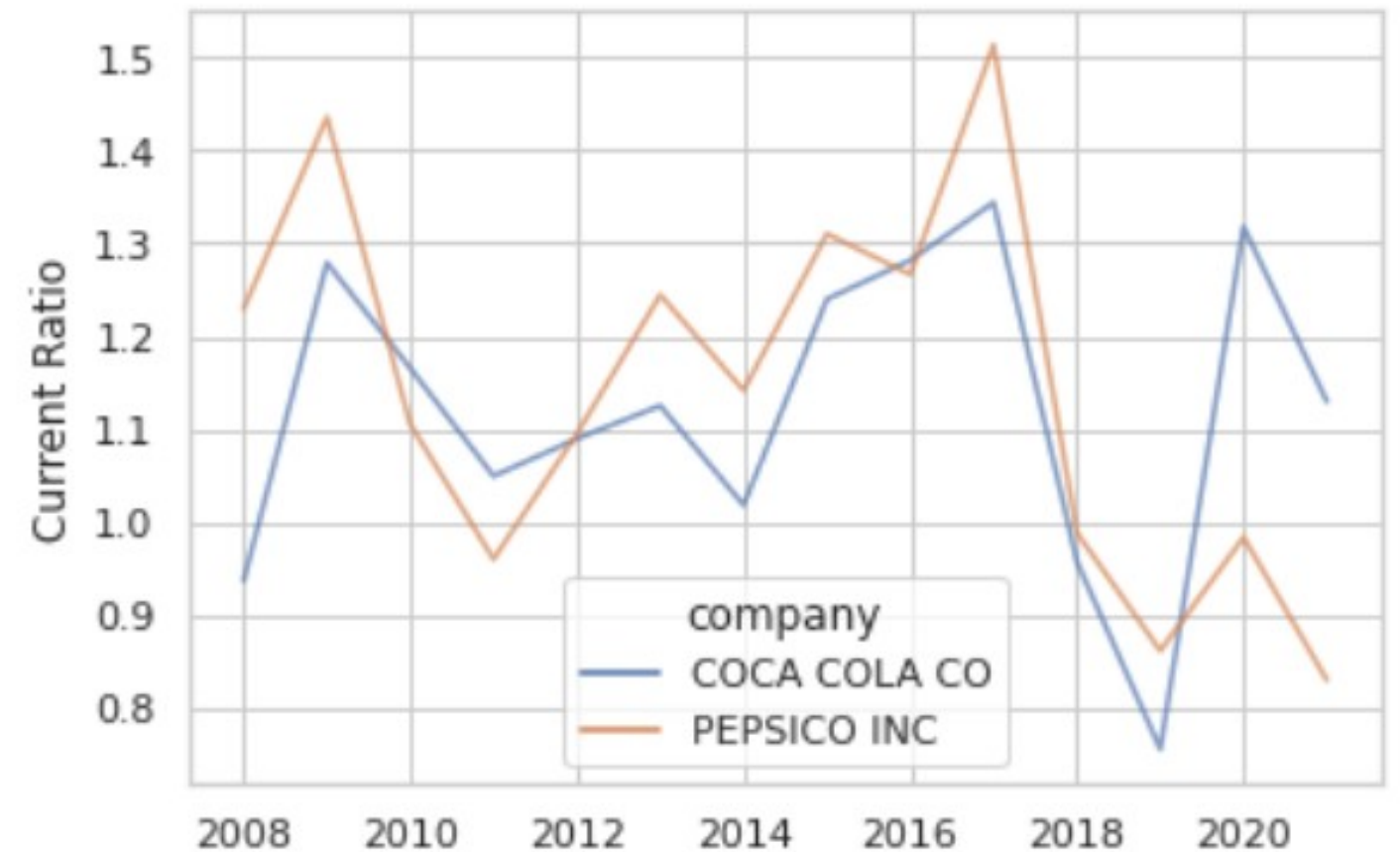
ANALYZING FINANCIAL STATEMENTS IN PYTHON



Rohan Chatterjee
Risk Modeler

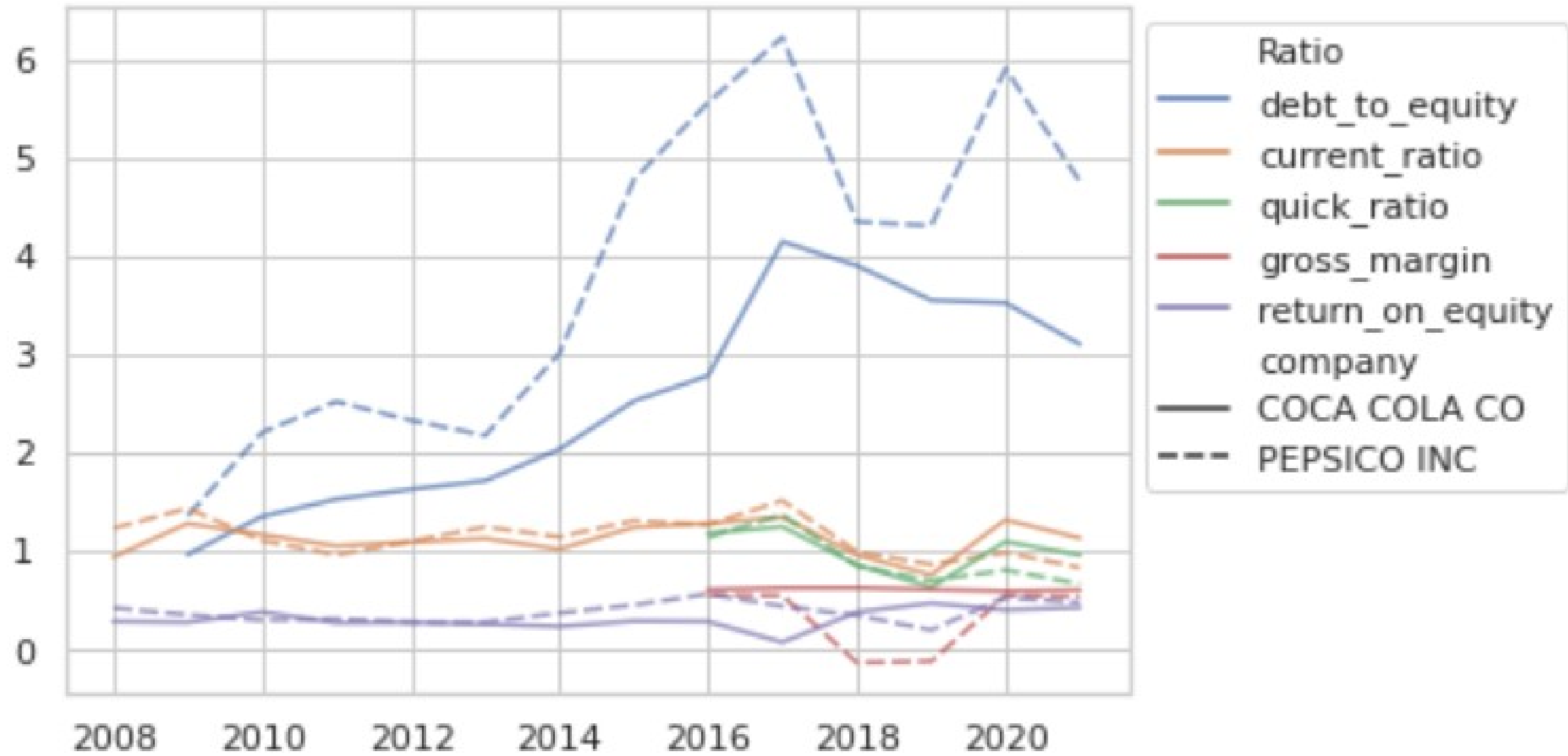
Plotting the ratios of two companies

```
sns.set_style("whitegrid")
ax = sns.lineplot(data=current_ratio,
                  x="Year", y='current_ratio',
                  hue="company", alpha=0.7)
plt.xlabel("Year")
plt.ylabel("Current Ratio")
```



Overcrowding when using sns.lineplot()

This graph has too much information in too little space, making it overcrowded



Introducing sns.relplot()

- First step is to melt the DataFrame to get longitudinal data.
- The "unmelted" DataFrame:

```
plot_df.head()
```

	company	Year	debt_to_equity	equity_multiplier	current_ratio	quick_ratio	gross_margin	return_on_equity
0	COCA COLA CO	2016.0	2.784147	3.784147	1.281848	1.181027	0.606693	0.283020
1	COCA COLA CO	2017.0	4.148547	5.148547	1.343863	1.246231	0.623356	0.073102
2	COCA COLA CO	2018.0	3.900536	4.900536	0.957973	0.857288	0.624569	0.378894
3	COCA COLA CO	2019.0	3.550919	4.550919	0.756720	0.631446	0.607712	0.469944
4	COCA COLA CO	2020.0	3.523343	4.523343	1.317718	1.094035	0.593112	0.401420

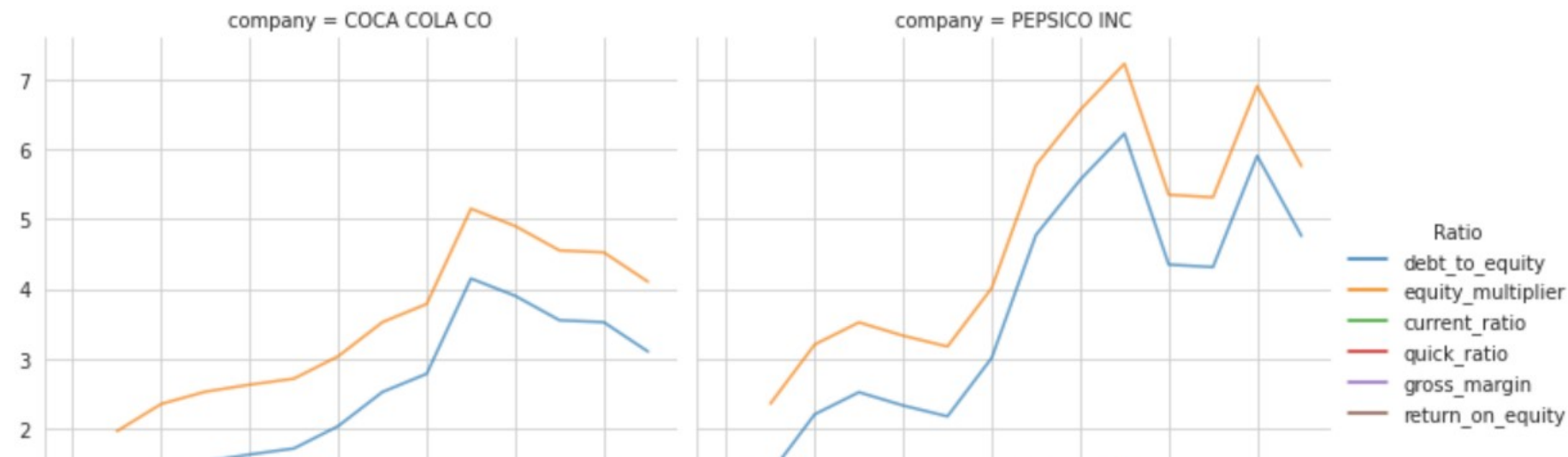
Now, melt the DataFrame

```
plot_df_melt = plot_df.melt(id_vars = ['company', 'Year'], var_name = "Ratio")
```

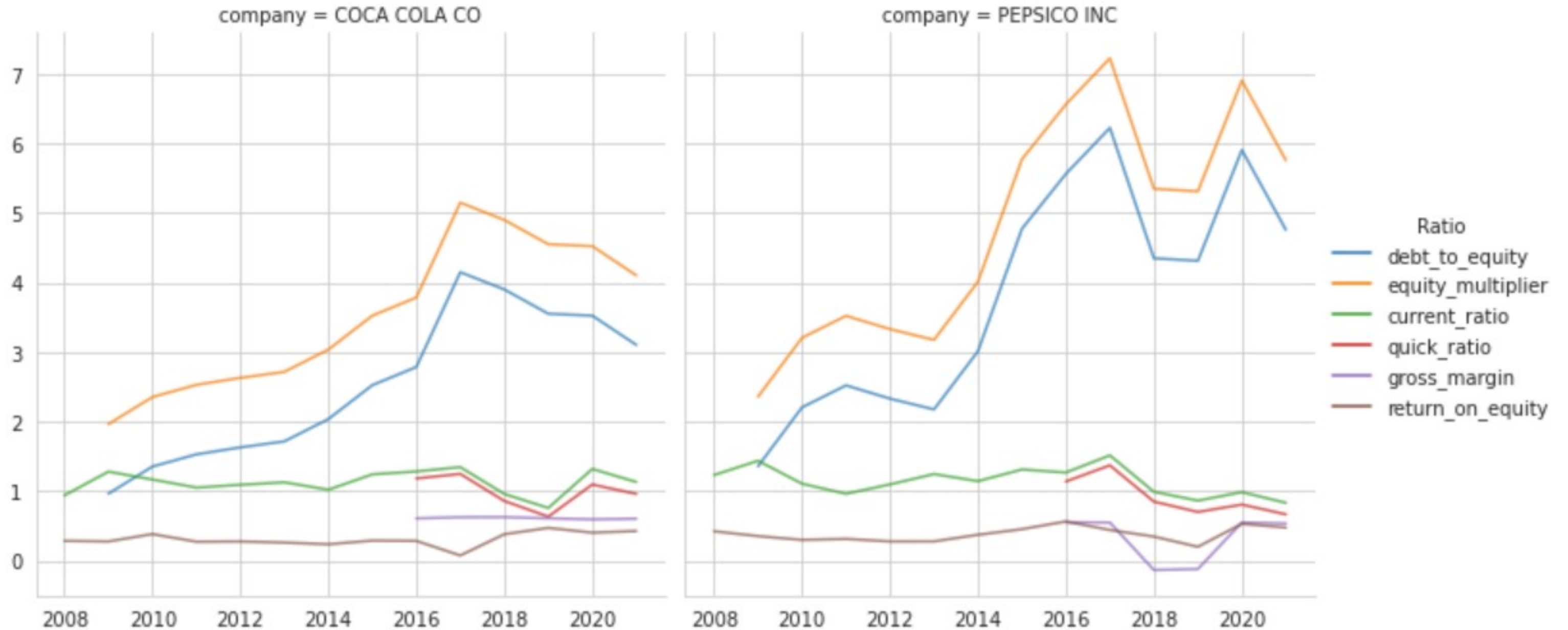
	company	Year	Ratio	value
0	COCA COLA CO	2009.0	debt_to_equity	0.962619
1	COCA COLA CO	2010.0	debt_to_equity	1.352063
2	COCA COLA CO	2011.0	debt_to_equity	1.528023
3	COCA COLA CO	2012.0	debt_to_equity	1.628057
4	COCA COLA CO	2013.0	debt_to_equity	1.714708
...	...	...	...	...
127	PEPSICO INC	2017.0	return_on_equity	0.439746
128	PEPSICO INC	2018.0	return_on_equity	0.344813
129	PEPSICO INC	2019.0	return_on_equity	0.197863

Plot ratios using sns.relplot()

```
ax = sns.relplot(data=df_melt, x="Year",  
                y="value", hue="Ratio",  
                alpha=0.7,  
                col='company',  
                kind='line')  
  
ax.set_ylabels('')  
ax.set_xlabels('')
```



A closer look at the plot



Let's practice!

ANALYZING FINANCIAL STATEMENTS IN PYTHON