

BA870 (SPRING 2022) – ASSIGNMENT #1

Instructions (3 pages):

Part (1): Create a Colab Notebook that uses the following URL:

<https://www.sec.gov/Archives/edgar/data/320193/000032019321000105/R2.htm>

to extract generic income statement data that can be used to create a complete 4-column Pandas Dataframe that contains the Income Statement data for Apple Inc (AAPL) for the years 2021, 2020 and 2019. An example of the original formatted HTML file for Apple is:

CONSOLIDATED STATEMENTS OF OPERATIONS - USD (\$) shares in Thousands, \$ in Millions	12 Months Ended		
	Sep. 25, 2021	Sep. 26, 2020	Sep. 28, 2019
Net sales	\$ 365,817	\$ 274,515	\$ 260,174
Cost of sales	212,981	169,559	161,782
Gross margin	152,836	104,956	98,392
Operating expenses:			
Research and development	21,914	18,752	16,217
Selling, general and administrative	21,973	19,916	18,245
Total operating expenses	43,887	38,668	34,462
Operating income	108,949	66,288	63,930
Other income/(expense), net	258	803	1,807
Income before provision for income taxes	109,207	67,091	65,737
Provision for income taxes	14,527	9,680	10,481
Net income	\$ 94,680	\$ 57,411	\$ 55,256
Earnings per share:			
Basic (in dollars per share)	\$ 5.67	\$ 3.31	\$ 2.99
Diluted (in dollars per share)	\$ 5.61	\$ 3.28	\$ 2.97
Shares used in computing earnings per share:			
Basic (in shares)	16,701,272	17,352,119	18,471,336
Diluted (in shares)	16,864,919	17,528,214	18,595,651
Products			
Net sales	\$ 297,392	\$ 220,747	\$ 213,883
Cost of sales	192,266	151,286	144,996
Services			
Net sales	68,425	53,768	46,291
Cost of sales	\$ 20,715	\$ 18,273	\$ 16,786

Your Colab Notebook should parse the HTML data to create and print a complete Pandas dataframe that has the following format on the next page:

Example dataframe output for Apple Income Statement:

	Income Statement Item for Apple	Sept 25 2021	Sept 26 2020	Sept 28 2019
0	Net sales	365817.00	274515.00	260174.00
1	Cost of sales	212981.00	169559.00	161782.00
2	Gross margin	152836.00	104956.00	98392.00
3	Operating expenses:	NaN	NaN	NaN
4	Research and development	21914.00	18752.00	16217.00
5	Selling, general and administrative	21973.00	19916.00	18245.00
6	Total operating expenses	43887.00	38668.00	34462.00
7	Operating income	108949.00	66288.00	63930.00
8	Other income/(expense), net	258.00	803.00	1807.00
9	Income before provision for income taxes	109207.00	67091.00	65737.00
10	Provision for income taxes	14527.00	9680.00	10481.00
11	Net income	94680.00	57411.00	55256.00
12	Earnings per share:	NaN	NaN	NaN
13	Basic (in dollars per share)	5.67	3.31	2.99
14	Diluted (in dollars per share)	5.61	3.28	2.97
15	Shares used in computing earnings per share:	NaN	NaN	NaN
16	Basic (in shares)	16701272.00	17352119.00	18471336.00
17	Diluted (in shares)	16864919.00	17528214.00	18595651.00
18	Products	NaN	NaN	NaN
19	Net sales	297392.00	220747.00	213883.00
20	Cost of sales	192266.00	151286.00	144996.00
21	Services	NaN	NaN	NaN
22	Net sales	68425.00	53768.00	46291.00

Part (2):

Your Colab parsing algorithm should be then be applied to the Income Statement for Microsoft (MSFT):

<https://www.sec.gov/Archives/edgar/data/789019/000156459021039151/R2.htm>

and your notebook should create and print a similar complete Pandas dataframe for MSFT for the years 2021, 2020 and 2019.

Notes:

- Use the example Colab notebook “*Assignment_1_Template.ipynb*” that is posted in the Lecture #2 folder on QustromTools to help you get started.
- You can use whatever libraries and code you wish to solve to assignment

DELIVERABLE (PDF file):

- Make sure you put your full name in a text cell at the top of your Notebook.
- Make sure your Notebook contains text and comments that fully describe how you solved Assignment #1.
- Create a PDF file of your executed Colab Notebook that contains the (i) **executed Python code**, (ii) **the output and results of your executed code**, and (iii) **all documentation in the text cells** that describe and document your solution in your Notebook.
 - One way to create a PDF file from Colab is to “Print” the Notebook to a PDF file.
 - Please make sure that the PDF file does not missing any items from your Notebook or “truncates” any output.
- Upload this PDF file under the Assignments tab on QuestromTools for “Assignment #1”
- Assignment #1 is due by 11:00pm on Tuesday, March 29, 2022.