Assignment 3: Data Cleaning and Analysis with Pandas

The objective of this assignment is to use Pandas to analyze a sample of data of corporate financial metrics and management language in earnings calls. The sample includes 100 companies randomly drawn from the S&P 500 index and their earnings calls in the fiscal year 2015 Q1. The finance dataset includes common corporate financial metrics, and the text dataset includes variables that measure management language characteristics. Please find more information in the data dictionary below. You will be asked to clean, merge, and analyze the finance and text datasets.

Data Dictionary

finance_data.csv:

- ticker: a company's trading symbol
- year: fiscal year
- quarter: fiscal quarter
- sic_two_digit: the company's industry code
- sales, revenue, net_income, total_assets: the company's common financial metrics

text_data.csv

- ticker: a company's trading symbol
- year: fiscal year
- quarter: fiscal quarter
- total_word_count: the number of words spoken by an internal speaker in the earnings call
- neg: negative sentiment in the speech by an internal speaker in the earnings call
- pos: positive sentiment in the speech by an internal speaker in the earnings call
- compound: overall sentiment in the speech by an internal speaker in the earnings call

Please follow the following steps to set up this assignment.

1. Download the files named finance_data.txt and text_data.csv from Canvas and put it in your Google Drive/Colab Notebooks folder
2. Click on Assignment3: Pandas starting Colab link
3. Make a copy of the starting notebook and connect it to your Google Drive
4. Run the first code cell to ensure you have connected to your Google Drive and are able to read the two datasets
5. Follow the instructions in the notebook and write code to answer the ten questions
6. Please note that the problems in this assignment are sequential. There is no need to load finance_data or text_data from the CSV files for each question.
7. For submission: **please download your notebook and submit the file to Canvas Assignment 3: Submission folder. Please also share your notebook with view access and paste the link as a comment when you submit.** Please also include a statement for the use / non-use of generative AI tools.