Overview

In class, we discussed Wall Street Journal articles based on academic research (see the references below) that used data analytics to detect when a company is potentially inflating its earnings per share number. The analytics work by highlighting when a company has fewer than a reasonable number of the integer 4 in the spot after the decimal point in diluted EPS reported in cents (for example 30.4 cents, 125.4 cents, 8.4 cents, etc.). On average, the integer 4, and all other integers, should appear directly after the decimal point of Diluted EPS represented in cents approximately 10% of the time, since each of the 10 digits has an equal likely chance of being listed. Thus, finding that a company routinely lists the digit 4 less frequently than 10% is a reason to suspect the company is manipulating its earnings. Why is this evidence of misreporting? Because the company will round its EPS to the nearest cent on the financial statements and 4 will round down while a 5 will round up. Thus, the company might be reporting a 5, 6, 7, 8, or 9 in that spot so that their EPS will be rounded up and higher than it should be. For example, true/unadjusted EPS of 100.4 cents will be rounded to $1.00 on the financial statements while true/unadjusted EPS of 100.5 cents will be rounded to $1.01. While that extra one cent may not seem like much, remember that EPS is earnings per one share of stock outstanding, which might be very, very large. For example, the average amount of stock outstanding in our data is over 165 million shares. Thus, $0.01 of earnings for 165 million shares is $1,650,000 of earnings.

We explored this data in class, but let’s use the same data to gain some additional insight. Specifically, suppose you are the researcher who first reports the phenomenon of too few 4's to the SEC (U.S. Securities and Exchange Commission), and you want to use your analytics skills to really help them understand this phenomenon. Follow the steps below to help make your point.

Instructions

Using R/RStudio or Alteryx (but not both), follow the instructions below.

* **1a**. Using the dataset we cleaned up and used in class (called "...EPS rounding\_after class" and provided for you as well as the data description sheet, "...Data Description Sheet\_EPS rounding\_after class.xlsx"), graph a bar chart that shows separate plots for the frequency of the EPS digit after the decimal point (called "digit\_diluted" in the dataset) based upon the quarters of the year (1-4). Thus, you will have four separate plots or one plot with 4 separate panels or facets.  
  ***Hint*:** We did this in class but had a separate plot for each year instead of for each quarter. For R, use `facet\_wrap()` and for Alteryx use "Batch".
  + If you're using **R**: use **[assign\_data\_EPS rounding\_after class.rds](https://canvas.illinois.edu/courses/21357/files/5243936?wrap=1" \o "assign_data_EPS rounding_after class.rds" \t "_blank)**[Download assign\_data\_EPS rounding\_after class.rds](https://canvas.illinois.edu/courses/21357/files/5243936/download?download_frd=1)file
  + If you're using **Alteryx**: use **[assign\_data\_EPS rounding\_after class.csv](https://canvas.illinois.edu/courses/21357/files/5243934?wrap=1" \o "assign_data_EPS rounding_after class.csv" \t "_blank)**[**Download assign\_data\_EPS rounding\_after class.csv**](https://canvas.illinois.edu/courses/21357/files/5243934/download?download_frd=1)file
* **1b**. Which quarter or quarters have the smallest difference in the frequency of integer 4 relative to integer 5? You don't need to calculate this difference, just look at the four plots and make your best judgment.
* **2a**. Next, create and examine these same sets of plots for just 1998. Thus, you will have four sets of graphs for that year.
* **2b**. Next, create and examine this same set of plots for just 2019. Thus, you will have four sets of graphs for that year.
* **2c**. What differences do you notice between 1998 and 2019 in the frequency of integer 4 and the frequency of integer 4 relative to integer 5?
* **3**. Next, use your analytics skills to find and report the worst 10 offenders of EPS rounding to the SEC.
  + Specifically, create a table of the 10 worst offenders following the criteria below.
  + You will need to create a table that has one observation per company. Thus, you will need to aggregate each company's results down to one row using a summarize function.
  + Additionally, you will only include a company in the table if the aggregated results from that company meet the following criteria:
    - The company has more than 56 observations (quarters/rows) in the original, "EPS rounding\_after class," dataset.
    - Less than 1.18% of all of the company’s observations have a 4 in the digit after the decimal point of EPS (e.g., in the `**EPS\_diluted\_cents`** column).
    - More than 11% of all of the company’s observations have a 5 in the digit after the decimal point of EPS (e.g., in the `**EPS\_diluted\_cents`** column).
    - ***Hint***: Create columns in your table for each of the three bullet points above and then use a filter function. One way to do this is to create a table with one observation for each company and the following columns: the total times the company had a 4 in the digit after the decimal point  of EPS (e.g., in  the `**EPS\_diluted\_cents`** column), the total times the company had a 5 in the digit after the decimal point  of EPS (e.g., in  the `**EPS\_diluted\_cents`** column), the total observations (quarters/rows) in the original dataset, the percentage of observations with 4 in the first digit after the decimal, and the percentage of observations with 5 in the first digit after the decimal.
* **3a**. Submit this table. At the minimum, the table (or a screenshot of the table) must contain the ticker (tic), company name (conm), percentages of observations with 4 in the first digit after the decimal, and percentages of observations with 5 in the first digit after the decimal.
* **3b**. What observations can you make about these companies?
* **3c**. Are the three companies charged by the SEC (FULT, HCSG, and TILE) in the following article in your table? A simple "yes" or "no" is fine. You do not have to answer this question. We will not grade your 3c.

      Maurer, M. (2021). "[Sec digs deeper into companies' EPS manipulation. (Links to an external site.)](https://www.wsj.com/articles/sec-digs-deeper-into-companies-eps-manipulation-11633870803)" *The Wall Street Journal.  
      Isn’t that really cool?! (You don’t have to answer that.)*

Data Files

* If you are using R, use **[assign\_data\_EPS rounding\_after class.rds](https://canvas.illinois.edu/courses/21357/files/5243936?wrap=1" \o "assign_data_EPS rounding_after class.rds" \t "_blank)**[**Download assign\_data\_EPS rounding\_after class.rds**](https://canvas.illinois.edu/courses/21357/files/5243936/download?download_frd=1)
* If you are using Alterxy, use **[assign\_data\_EPS rounding\_after class.csv](https://canvas.illinois.edu/courses/21357/files/5243934?wrap=1" \o "assign_data_EPS rounding_after class.csv" \t "_blank)**[**Download assign\_data\_EPS rounding\_after class.csv**](https://canvas.illinois.edu/courses/21357/files/5243934/download?download_frd=1)

Data descriptions can be downloaded: **[assign\_Data Description Sheet\_EPS rounding\_after class.xlsx](https://canvas.illinois.edu/courses/21357/files/5243931?wrap=1" \o "assign_Data Description Sheet_EPS rounding_after class.xlsx" \t "_blank)**[**Download assign\_Data Description Sheet\_EPS rounding\_after class.xlsx**](https://canvas.illinois.edu/courses/21357/files/5243931/download?download_frd=1)

Deliverables

If you're using **R**:

* You are submitting two files - the R notebook (**.Rmd**) and a knitted html file (**.html**).
* Include your response to the questions inside your R notebook.
  + Use a clear section heading in your R notebook to respond to each question asked in the instructions.
* Save your R notebook as **netid\_hw2.Rmd** (e.g., **johndoe1\_hw2.Rmd**)**.**
* Knit your R notebook into **netid\_hw2.html**(e.g., **johndoe1\_hw2.html**).
* Submit both the **netid\_hw2.Rmd** and **netid\_hw2.html** files on Canvas.

If you're using **Alteryx**:

* Save your workflow as **netid\_hw2.yxmd** (e.g., **janedoe1\_hw2.yxmd**).
* Include your response to the questions inside a separate .docx or .pdf file (e.g., **janedoe1\_hw2.docx** or **janedoe1\_hw2.pdf**).
  + Use a clear section heading in your report to respond to each question asked in the instructions.
* Submit both the **netid\_hw2.yxmd** and **netid\_hw2.docx** (or netid\_hw2.pdf) files on Canvas.

References

* Malenko, N., Grundfest, J., and Shen, Y. (forthcoming). [Quadrophobia: Strategic rounding of EPS data (Links to an external site.)](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1474668" \t "_blank). *Journal of Financial and Quantitative Analysis, forthcoming*
* Maurer, M. (2021). "[Sec digs deeper into companies' EPS manipulation. (Links to an external site.)](https://www.wsj.com/articles/sec-digs-deeper-into-companies-eps-manipulation-11633870803)" *The Wall Street Journal.*
* Michaels, D. (2018). "[SEC probes whether companies rounded up earnings per share. (Links to an external site.)](https://www.wsj.com/articles/sec-probes-whether-companies-rounded-up-earnings-1529699702)" *The Wall Street Journal*.
* Sebastian, D. (2020). "[SEC settles with Interface, Fulton Financial for violations related to EPS reporting. (Links to an external site.)](https://www.wsj.com/articles/sec-settles-with-interface-fulton-financial-for-violations-related-to-eps-reporting-11601314460)" *The Wall Street Journal*.