COVID-19 Twitter Reactions

Case Study Rubric

DS 4002 – Spring 2024 - Lily England
Due: May 02

Submission format:

- **Document containing all deliverables assigned below**
- **Google Colab .ipynb file**

Why am I doing this? In times of crisis, understanding how people feel, and what topics trigger the strongest reactions, is essential for crafting messages that inform, support, and unite. In this case study, you will use real-world social media data to uncover how sentiment shifted around different COVID-19 topics during the early shutdowns. You'll strengthen your ability to clean text data, filter based on real-world issues, create meaningful visualizations, and communicate your findings to a non-technical public health communications team.

What am I going to do?

You will work with a dataset of tweets from March–April 2020 that have already been labeled with:

- A sentiment category (Positive, Negative, or Neutral)
- A VADER sentiment score (numerical value)

You will:

- Select 2–3 keywords or phrases related to the COVID-19 shutdown (e.g., "Amazon," "lockdown," "masks," "restaurants," "Trump")
- Filter the tweets based on those keywords
- For each keyword:
    - Calculate the count of positive, negative, and neutral tweets
    - Calculate the average VADER sentiment score
- Create clear, simple visualizations to show how sentiment differed across topics
- Summarize your findings in a short, non-technical report aimed at public health communicators

Final Deliverables include:

- Summary report (1-2 pages) clearly describing your process, findings, and insights
- 3 visualizations
- Analysis code, properly organized and commented (Google Colab file)

Tips for success:

- Focus on clear, simple storytelling- don't overcomplicate it.
- Keep your visualizations clean and easy to understand
- Briefly acknowledge any biases

- Ask yourself, "*Would someone without a data background understand my conclusions?"*

How will I know I have Succeeded? You will meet expectations on Case Study rubric

Formatting	• 1-2 page PDF report containing:
	  o Name, Case Study, Date
	  o Title
	  o Summary
	  o Discussion Paragraph
	  o 3 visualizations
	• Code file (.ipynb)

Report	• Goal: This document should summarize the process you took and highlight conclusions
	• The document will contain:
	  o An explanation of your keywords and filtering process
	  o A summary of sentiment patterns across topics
	  o Clear conclusions aimed at a non-technical public health audience
	• 3 visualizations
	  o Bar Plot → Sentiment counts grouped by keyword
	  o Box Plot → Compare VADER score distribution per keyword
	  o Pie Chart → Sentiment breakdown for one keyword
	• Two page maximum
	• PDF format

Code	• Goal: Code should be easy to follow and easy to reproduce
	• Google Colab file (.ipynb)
	• Code will:
	  o Read in provided .csv
	  o Filter tweets by keyword
	  o Calculate counts and averages
	  o Generate at least three clear plots
	• Clearly organized with helpful comments

Acknowledgements: Special thanks to Jess Taggart from UVA CTE for coaching on making this rubric. This structure is pulled from Streifer & Palmer (2020).