Semester Project Research Essay

Mia Eberle | Jacob Fenn | Zane Lemaster

WRIT 20833

Shark Tank Around the World

Our project focuses on understanding how viewers talk about the sharks on Shark Tank and how that perception has changed over time and across different versions of the show. Shark Tank has been around for many years now, and because of that, the sharks have become major public figures whose personalities and decisions get talked about constantly. Online culture has also changed a lot during this time, becoming more opinionated and more global. These changes made us wonder whether people's reactions to the show have shifted in the same way. We originally expected the comments to be mostly negative, because online spaces often feel harsh or critical, but our earlier deliverables in this class showed us that assumptions about online negativity can be totally wrong. That motivated us to explore the topic on a much larger scale by studying thousands of comments across different countries. Our goal was to see what people actually say about the sharks and whether that changes over time, instead of assuming we already knew the answer.

The main research question guiding our project was how the perception of the sharks on Shark Tank has changed over time, and how the general perception of the entire show has held up and shifted with ever-changing cultures. This question matters because Shark Tank isn't just entertainment, but rather it influences how people think about entrepreneurship and business values, and the sharks themselves have become symbols of those ideas. Understanding how viewers respond to them helps us understand how public attitudes toward business, success, and even risk-taking change in different cultural moments.

To answer our question, we built a dataset of over 10,000 YouTube comments taken from 33 episodes across several international versions of the show, including Shark Tank US, Shark Tank Global, Shark Tank India, Shark Tank Australia, and Dragons' Den (United Kingdom). We scraped the comments of each one of these episodes, hoping to find a correlation between names, topics, and various patterns. We made sure to keep our focus strictly on the text, and kept an open mind for anything we might find upon data analysis. It was important to us that we handled the data ethically, since real people wrote these comments. We made sure not to misrepresent anyone's words and this is where we were able to utilize close reading to double-check when the computer might have misunderstood sarcasm, jokes, or tone.

Once our dataset was cleaned, we analyzed it using Python. We used pandas for organizing the comments, VADER for sentiment analysis, and Gensim's LDA for topic modeling. We also used term frequency analysis to see which words appeared the most often, especially names of sharks and recurring vocabulary related to business or pitching. Term frequency helped us notice right away which people or ideas dominated the conversation. Sentiment analysis helped us measure whether comments were mostly positive, negative, or neutral. This allowed us to track patterns in how viewers emotionally responded to each investor. Topic modeling was meant to reveal larger themes people were discussing, although it ended up being tricky because YouTube comments are usually short. Still, it gave us a general sense of what topics tended to form in different versions of the show.

While we were working with these tools, we also realized how much small decisions on our end shaped the output. For example, choosing to include or exclude certain words from our analysis could completely change what the algorithm focused on. Even deciding how many topics to generate in the LDA model changed the story the data seemed to tell. This made us more aware that this type of data analysis isn't just a neutral, automatic process. There's a lot of human choice built into it, and those choices influence what patterns appear and which ones stay hidden. Technology and coding resources are fantastic, and they all typically do their job very well, but the human mind behind it is just as important. Understanding this helped us read the results more carefully instead of assuming the computer was always right without running it by a pair of human eyes.

Putting all these methods together helped us see different layers of meaning in the comments. The computational tools gave us patterns we couldn't have found by hand, while close reading helped us make sure the data wasn't misleading. For example, a comment saying "Oh great job, Kevin, totally fair…" could be sarcastic, and a computer might label it as positive. Close reading would help to avoid that. This combination of methods made our analysis much stronger, because no single tool could tell the whole story on its own.

When we looked at the sentiment analysis results, we were surprised. Instead of seeing mostly negative comments like we originally expected, we found that over half of the comments were neutral and about a third were positive. Only a small percentage were negative. This pattern was consistent across different countries, which showed us that online audiences around the world respond to these shows in surprisingly similar ways. We expected bigger differences between countries, but the sentiment distributions were almost identical. This challenged our assumptions about how harsh people online really are and made us realize that viewers tend to be more supportive or neutral than critical.

Term frequency revealed another major insight. Across all shows, shark names came up constantly. In fact, people talked about the sharks even more than the entrepreneurs or products being pitched. This suggests that the investors are the true center of the show in the eyes of the audience. Even when a pitch was interesting, many commenters focused more on a shark's reaction, advice, or personality. This held true in positive comments, neutral ones, and especially negative ones. When viewers did criticize something, it was almost always a shark's decision or attitude rather than the entrepreneur or product. That showed us that negativity, when it appears, is usually directed at the investors because viewers feel strongly about the choices they make.

Topic modeling gave us a smaller amount of insight compared to the other methods, mainly because the comments were so short that the model couldn't form strong topics. However, it did still reveal that many comments clustered around investor behavior, specific shark personalities, and reactions to decisions made during the pitches. Even though the model didn't create perfect topics, it still supported the idea that the sharks dominate the discussion.

Overall, these results taught us that the online conversation around Shark Tank is a lot more positive and shark-focused than we expected. We thought people would mostly criticize the entrepreneurs or the products, but instead the investors took center stage in almost every category of analysis. We were also surprised that the tone of comments didn't shift dramatically across countries, even though cultural norms and business values differ. The clear language barrier in certain countries did skew our data, but overall the findings remained similar. This made us think more about how global media shapes conversations and creates similar audience behavior across cultures.

Reflecting on the process, this project showed us why computational tools and humanities skills are both necessary for understanding cultural data. The computational tools let us process thousands of comments and find patterns that our eyes alone would never catch. But the humanities side gave us the ability to understand tone, meaning, and nuance that algorithms simply can't understand. Together, they created a fuller picture of how people talk about Shark Tank and what those comments actually mean. It also made us more aware of the role of algorithmic classification. The tools we used, like VADER and LDA, categorized the comments in ways that shaped our understanding of the results. For example, if VADER labeled a comment as neutral, we tended to treat it that way even though a human reader might see subtle positivity or frustration. This raised an important question about what gets lost when people's complex expressions are reduced to a few categories. AI can't understand humor, emotional nuance, or cultural context the way humans can, so our interpretation still mattered a lot.

There were definitely limitations to our project. Our dataset definitely was not perfectly balanced across seasons or countries (because these factors varied), and sentiment tools struggled with sarcasm. Topic modeling also didn't work as well as we hoped because short comments don't give the model much to analyze. If we had more time, we would want to gather a larger and more evenly distributed dataset, possibly from one season at a time to make time-based comparisons more accurate. We would also want to use more advanced language tools that understand sarcasm better, and maybe compare YouTube comments to platforms like Reddit or TikTok to see if audiences behave differently there.

Even with these limitations, we feel confident in the big conclusions we found. People around the world talk about Shark Tank in surprisingly similar ways, the overall sentiment is more neutral and positive than expected, and the sharks are consistently the main focus of conversation. These findings show that while online culture can seem absolutely crazy, chaotic, and rude, people's reactions to popular shows like Shark Tank tend to follow clear patterns that become visible when key computational methods are utilized and studied.