2023 Game Day Analytics Challenge

Charchit Dahal, Sampoornaa Ghosh, Samruddhi Supekar, Jordan Contreras

Grad Team 6 - Stats BOMB

THE UNIVERSITY OF UTAH

1. Introduction

The Super Bowl is widely recognized as one of the world's most viewed live sports events. With millions of viewers tuning in from across the globe, it presents an ideal platform for brands to showcase their brands and services. This white paper aims to share insights about Super Bowl LVII advertisements as part of the Game Day Analytics competition at the University of Utah. Our study tries to understand the performance of advertisements, brand representation, and social media engagements, among other valuable metrics. The findings presented here will help marketers, advertisers, and educators get insight into the advertisement landscape in this year's Super Bowl.

2. Methods

The analytics competition organizers provided us with the Twitter dataset, which comprised a dataset of 2 million tweets taken during the Super Bowl. The data was divided into digestible bits using a python program, and it was ready to be uploaded to Domo [1], a data management and analytics platform. An extract, transform, and load (ETL) procedure was created in Domo to aggregate and clean the data, preparing it for analysis.

Domo's robust analytics tools were used to create time series plots, word clouds, and geolocation maps, among others, to examine the data. These cards were used to uncover trends and insights into the data, including engagement levels, popular brands, and celebrity influence. Lastly, the

findings were displayed in various visually appealing and educational visualization tools, which offered insights into the Super Bowl LVII advertising scene.

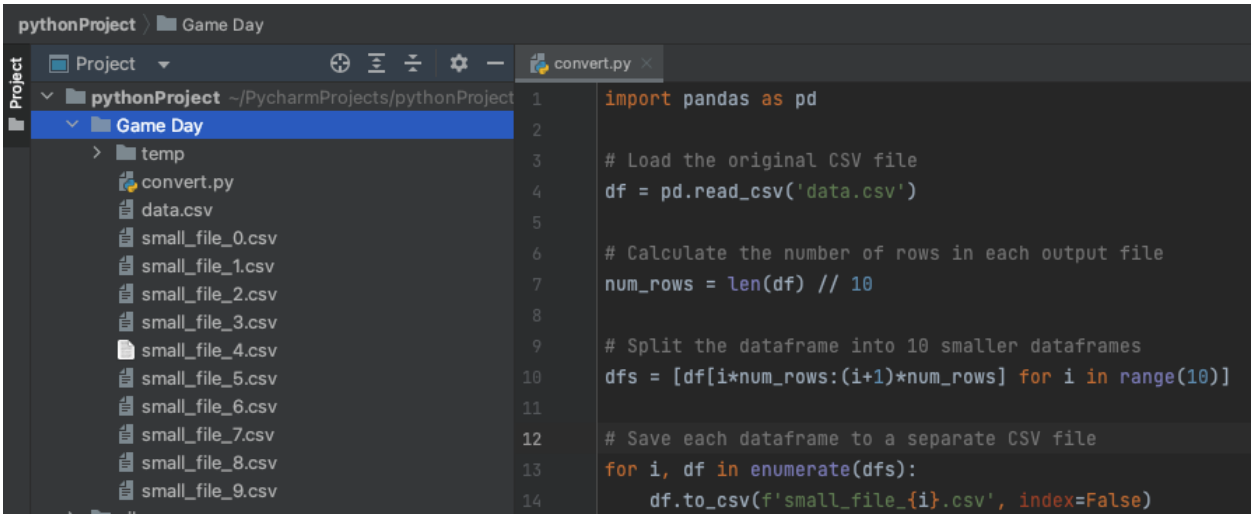

Fig 1. Example of Python code used to import the Twitter data set using the pandas package. Code will be provided in the References section.

In keeping with the spirit of transparency and open science, we are happy to offer all our analysis techniques in this white paper, including the Python code, ETL pipelines, and visualization available to the public. We hope that making these materials available to other professionals and future competition participants will promote collaboration and produce brand-new insight into the future Super Bowl advertising landscape. Access to these resources is available via our GitHub project, listed at the bottom of this paper [2].

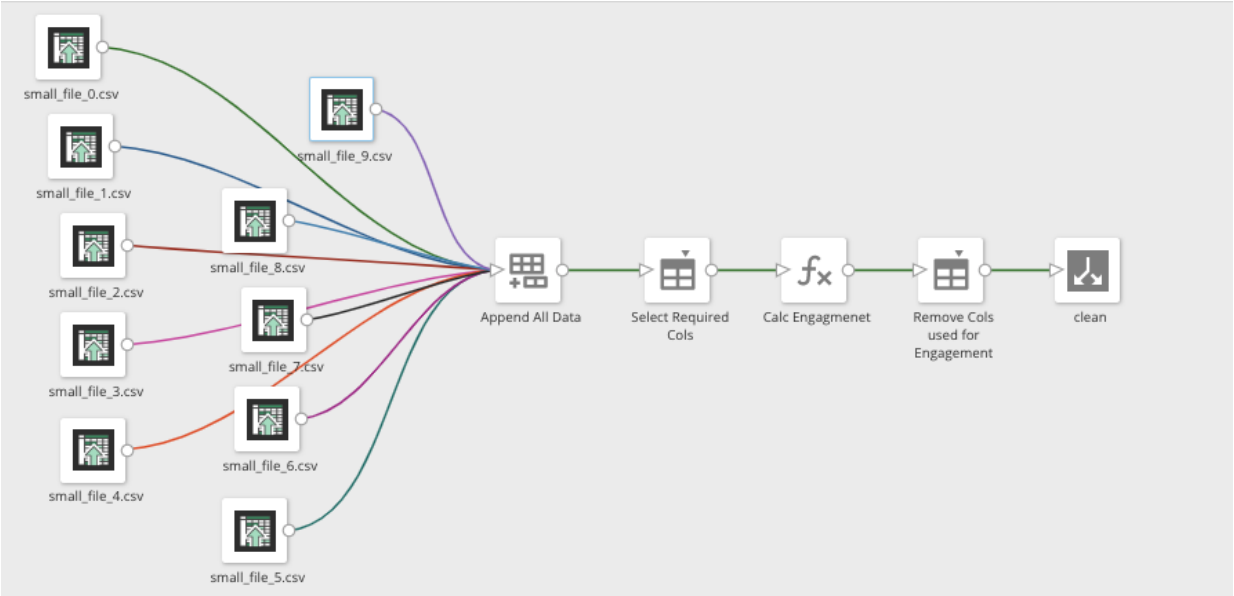

Fig 2. Schematic showing the workflow of the data reduction and cleaning process.

3. Results and Metrics

The Results and Metrics section contains a full explanation of the data analysis, key metrics, and conclusions drawn from the data. Location, Time Series, Engagement, Celebrity, and Sensitivity comprise these sections. We have used visual aids such as tables and charts to provide a comprehensive summary of the key findings. The key findings of our investigation are explained below.

3.1 Location

Our study found that the highest levels of interaction and viewership of the Super Bowl came from the states of Texas, California, and New York City. This indicates that these regions had a greater interest in the game and related content, possibly due to population size, cultural influences, or regional football team affiliations.

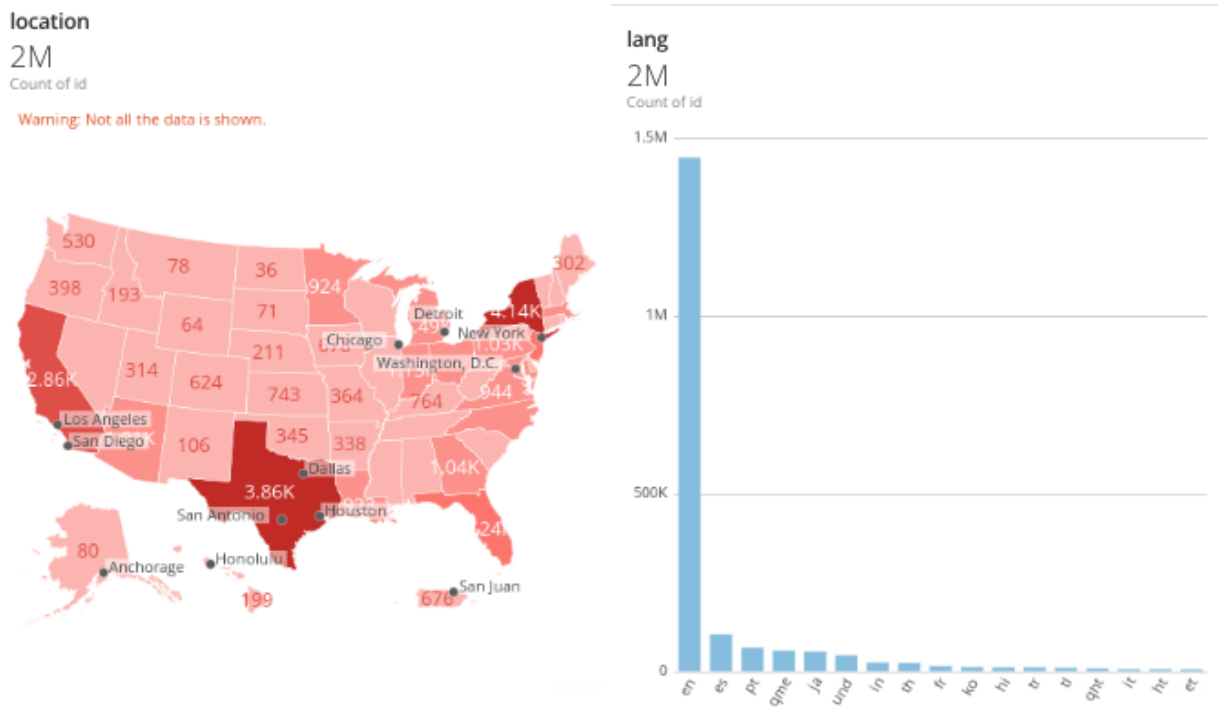

Fig 3. (Left) Density plot of Twitter engagement in the United States by state. (Right) Bar graph of languages used to engage in Twitter.

Furthermore, we discovered that most interactions and viewership were in the English language, with Spanish and Portuguese being the next languages of choice for those interacting with the content. This suggests that a significant portion of the audience (> 10%) prefers to consume and engage with content in languages other than English. Catering to these language preferences could be a valuable strategy for future Super Bowl marketing and advertising efforts.

3.2 Time Series

The time series research in Figure 4 shows that the volume of tweets grew during and after halftime, with the third quarter showing the highest activity–this is presented by the peak in Twitter engagement at around 3:00AM-UTC. The analysis found that the Tweets volume related to the Super Bowl increased after halftime, with the third quarter showing the highest activity and the first quarter coming in second. This information could be used by future advertisers or even the event organizer to price the ad slots differently.

The increase in tweet volume after halftime is likely because of the Super Bowl halftime show, a highly anticipated event often featuring famous musicians and entertainers. This can generate much buzz and excitement on social media, increasing tweet volume. We theorize that the higher activity during the third quarter could be due to this critical period in the game, where the outcome is often still uncertain, and fans are likely to be more engaged on social media. Thus, in order to reach a wider audience companies are encouraged to advertise during this time. The second-highest activity was during the first quarter which may correspond to the excitement and anticipation leading up to the game. The second quarter, similar to the first, also scores high on engagement and might be a more cost-accessible option for slightly less viewer response. Ads during the fourth quarter of the game show significantly less engagement, which may correspond to people paying closer attention to the game itself or being desensitized from an advertisement at this stage.

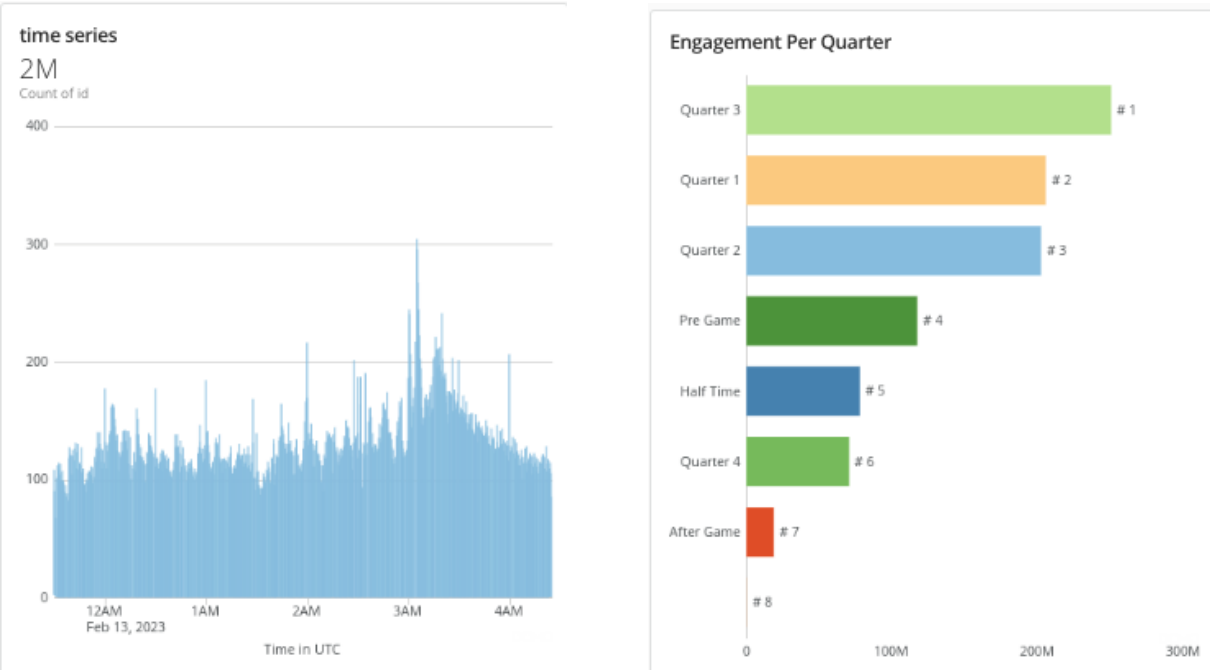

Fig 4. (Left) Time histogram of interactions in Twitter in Coordinated Universal Time (UTC) with a bin width selected to showcase the spikes in Twitter activity. (Right) Bar chart showing the number of interactions during relevant times before, during, and after the Super Bowl.

3.3 Engagement

Based on the number of ads and dollars spent, we discovered that some brands had higher levels of engagement. The above analysis found that certain brands had higher levels of engagement despite having similar ad placement and spending equal amounts of money on their ads. Several factors could explain why some brands had higher engagement levels than others. One possibility is that the ads were more impactful and creative and had a celebrity ambassador. Brands that created advertisements that stood out from the rest and resonated with viewers are more likely to generate higher levels of engagement.

The higher engagement levels resulted from effective pre- and post-game marketing efforts. Brands such as *Limit Break* and *Jeep - Electric Boogie* generated buzz and anticipation leading up to the Super Bowl. They were followed up with engaging post-game content and have been more successful in generating engagement and keeping their brand top-of-mind with consumers than other advertisements. *Disney's 100 Years Special Look* came a close third using the same tactics as above. Overall, the analysis of ad engagement levels suggests that a combination of factors, including ad creativity and marketing efforts, all played a role in determining the success of Super Bowl advertising campaigns.

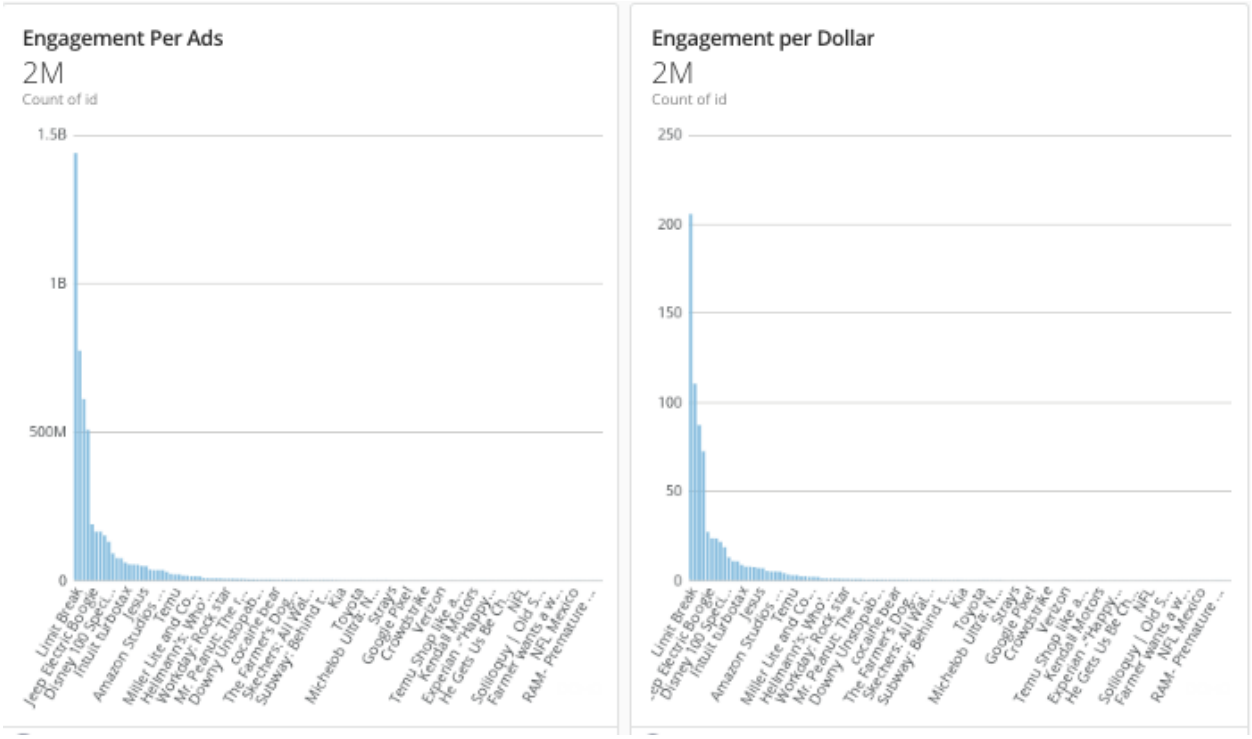

Fig 5. (Left) Bar graph showing the number of ad-related Twitter engagements. (Right) Bar graph showing the cost per ad.

3.4 Celebrity Impact

According to our findings, commercials featuring celebrities significantly increase audience reach and overall engagement. Celebrities are often well-known and well-liked by the audience, making viewers more likely to pay attention to an ad featuring them. The viewers may feel a personal connection or sense of aspiration toward the celebrity, which can create an emotional attachment to the ad and the advertised product.

Celebrity endorsements can also provide a sense of authority or expertise to the advertised product or service, increasing consumer trust and interest. Another critical factor is the emotional connection that viewers may feel toward a celebrity. This connection can make viewers more likely to engage with the advertised product or service and may lead to a long-term attachment to the brand. By featuring celebrities in their ads, brands have tapped into these emotional connections and created a more meaningful and memorable experience for viewers.

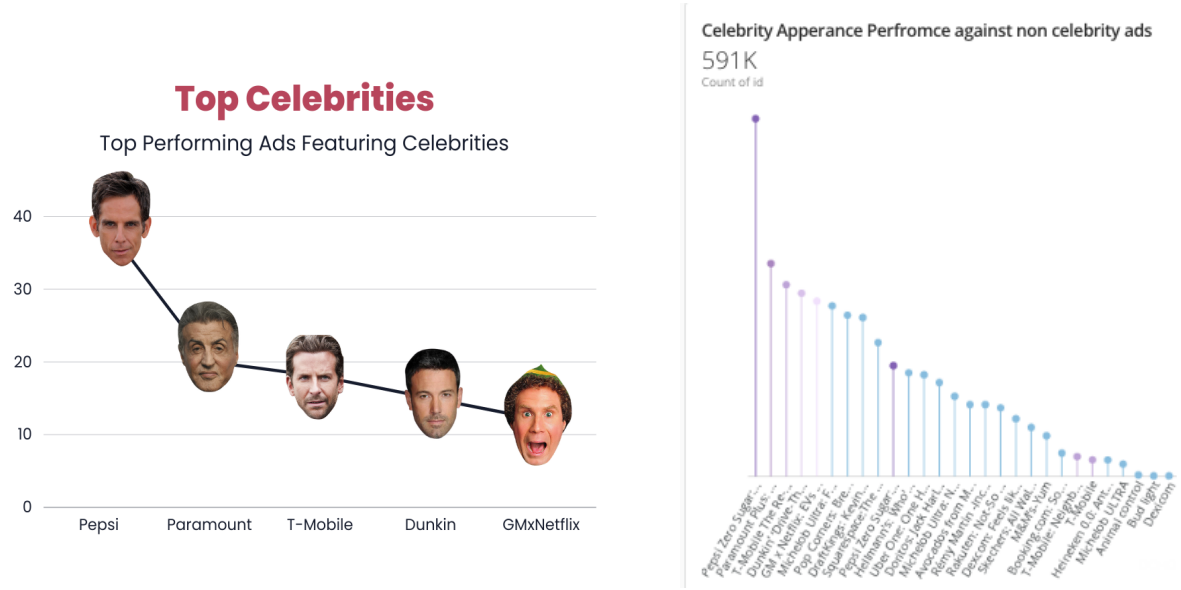

Fig 6. (Left) Percentage of celebrity-related Twitter engagement for the top 5 celebrity ads. (Right) Twitter engagement of all ads. The purple color corresponds to those ads that featured a celebrity.

Additionally, using known celebrities in ads generated buzz and excitement leading up to the Super Bowl. Fans of the superstar are more likely to tune in to see their favorite stars in action, which drives higher viewership and engagement. Younger celebrities also tend to have a higher chance of attracting an audience than older ones. *Pepsi Zero Sugar* had two advertisements with Ben Stiller and Steve Martin, respectively. Only the ad with Ben Stiller gathered engagement, becoming the top-performing celebrity advertisement. The analysis suggests that using celebrities in advertising during the Super Bowl is a highly effective strategy for increasing audience reach and engagement. By carefully selecting well-liked and relevant celebrities for the target audience, brands have created a memorable and impactful advertising campaign that resonates with viewers long after the game.

3.5 Keywords

We examined the most popular phrases and hashtags on social media during the event to future advertisers. Our keyword analysis revealed that the themes of winning and taste were responsible for the majority of tweets of industry ads during the Super Bowl LVII. The accompanying bubble graphs offer further information on these keywords. Similarly, the below table shows niche keywords or misspelled words that do not perform well.

Keyword	**Count**
giannisantetokounmpo	1
youcantdothat	1
slipins	1

In advertising, the primary objective is to generate interest in a brand. Consumers often use social media platforms to express their views and share their experiences with a brand, which can significantly impact brand perception. A consumer's level of engagement with a brand is directly related to their likelihood of engaging with it on social media. Therefore, our research focused on the number of keywords attributed to each brand to evaluate the effectiveness of the Super Bowl LVII ads.

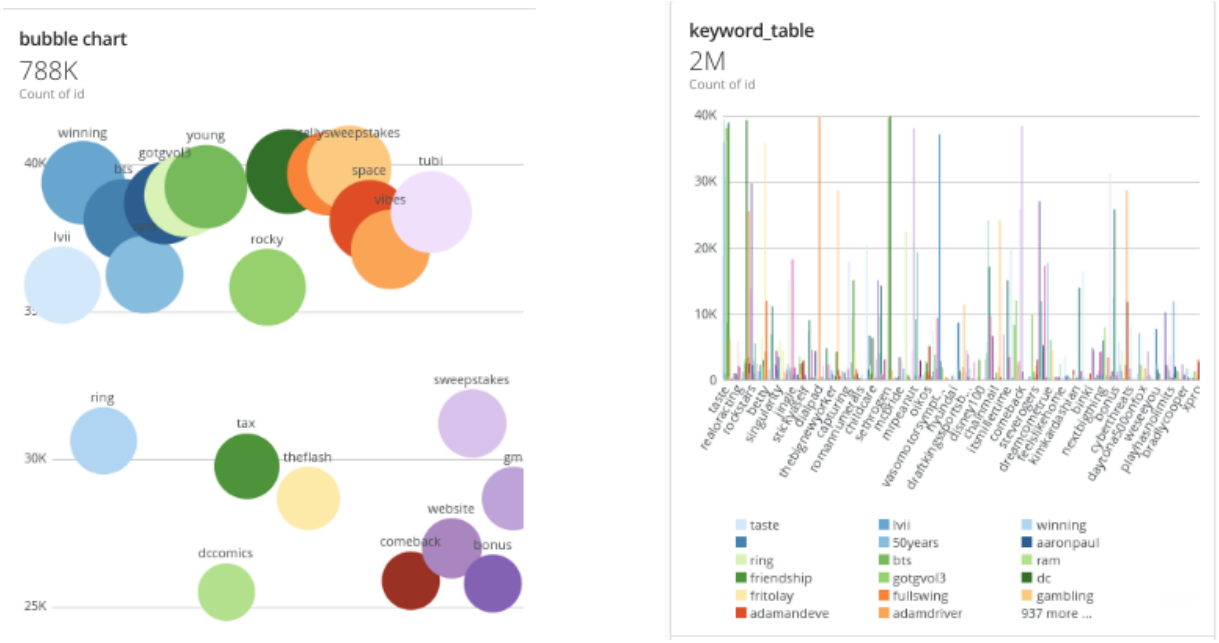

Fig. 7 (Left) Diagram showcasing the top 23 most popular keywords used on Twitter during the Super Bowl. The colors indicate different ads and the size as well as vertical location correspond to the number of times a particular word appeared. (Right) Bar graph corresponding to all the keywords in the Twitter data set.

3.6 Sensitivity

We also examined the potentially sensitive ad reaction from our dataset. With this information, future advertisers can anticipate potential unfavorable responses to their comments. Our sensitivity analysis of Super Bowl LVII advertisements allowed us to identify the most controversial ad based on audience reception.

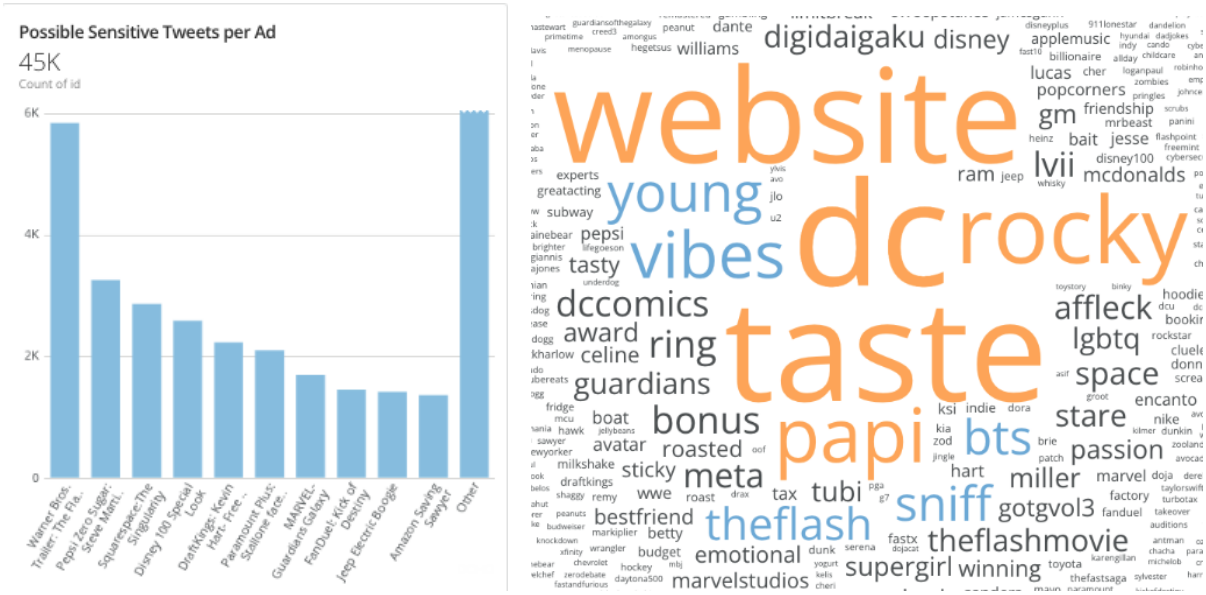

Fig. 8 (Left) Bar graph showing sensitivity responses in Twitter engagement deemed as unfavorable for the listed ads. (Right) Word cloud indicating the most used words in Tweeter data with unfavorable sensitivity responses.

We found that the Warner Brothers Flash commercial with Ezra Miller and the Pepsi Zero Sugar commercial featuring Steve Martin was the most divisive, generating significant negative feedback and criticism from viewers on social media platforms, respectively. These advertisements sparked much debate and were met with strong reactions from viewers. Negative perceptions from the public of these celebrities directly impacted the brand that they represented. These findings underscore the importance of responsible and thoughtful advertising in today's world, where social media can amplify adverse reactions and lead to significant brand damage.

We can identify the sensitive keywords to describe the advertisements by the above word cloud in Figure 8. Some of the common negative keywords can be seen together with famous brand names like DC or The Flash confirming that the viewers tend to associate bad celebrity images with the brands that they represent.

4. Conclusion

In conclusion, our examination of the Super Bowl LVII advertising environment yielded several significant findings regarding the potency and significance of the commercials that ran during the game. The important trends and patterns can influence future advertising tactics we found by evaluating data on geography, time series, interaction, celebrity presence, keywords, and sensitivity. The public study of our methods and tools can aid in a better understanding of the Super Bowl advertising environment and offer insightful advice for marketers and advertisers looking to develop more compelling ad campaigns. Overall, our study emphasizes the significance of data analytics in the advertising sector and the potential advantages of utilizing data-driven methodologies to maximize the effectiveness of advertising tactics.

5. References

[1] Analytics Tool - https://www.domo.com/
[2] GitHub Repo - https://github.com/charchitdahal/GameDay-Analytics-Challenge