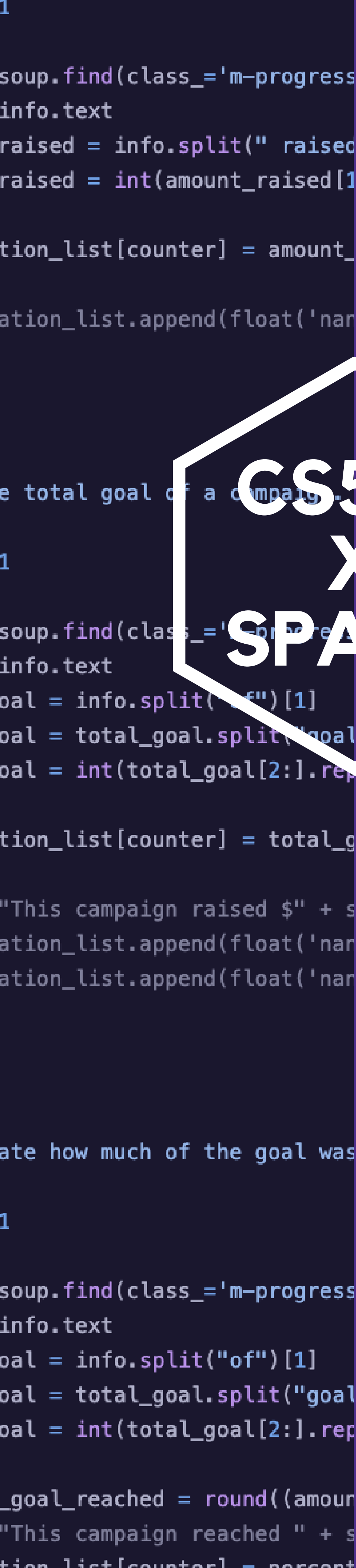
SPRING 2021



CS506 X SPARK

Deviance or Deservingness: Opioids, Morality, & Economic Precarity

PROPOSED BY

Jayden Font | Divya Gowravaram
Michalina Jadick | Nikita Jakkam | Della Lin

PROPOSED TO

Heather Mooney
Department of Sociology
Boston University

TABLE OF CONTENTS

INTRODUCTION & BACKGROUND

BACKGROUND

Sociology, broadly, looks at how macro-level systems, institutions, and ideologies shape life outcomes for individuals and groups. This project explores how the ongoing opioid crisis, which has killed over 760,000 people since 1999, impacts support people, care workers, and victims seeking aid on a public platform. To provide an idea of the scope of this issue, an estimated 10.3 million people aged 12 or older misused opioids in 2017, 9.9 million people misused prescription pain relievers, and 808,000 people used heroin. This project primarily addresses how individuals and groups frame their struggles in order to garner support by analyzing crowdfunded campaigns hosted on GoFundMe posted from 2010-2020. Campaigns are filtered using a variety of keywords related to drug use and overdose to explore how competing frames of drug use and addiction change over time. It is crucial to investigate how various factors associated with a campaign impact framing and campaign success in addition to exploring the different relational, moral, and affective appeals that are made to potential donors online.

This project has both policy implications and theoretical promise. Given the long-reaching effects of COVID-19 and the ongoing opioid crisis (which has been overshadowed by and accelerating since COVID-19 began), it will be important to understand how death, loss, and need are constructed by supporting people in times of (layered) crisis. This research represents a case to explore how morality and deservingness change over time and across populations.

GOAL & KEY QUESTIONS

The goal of our project is to analyze how drug use is framed. Specifically, how it is either a moral failing (deviance) or a medical condition in need of treatment (deservingness). Part of our goal is to also examine how the framing of drug use affects GoFundMe campaign success. Some of the key questions that we aimed to answer include:

1. What do we use to define a successful campaign and why?
2. What are some of the trends we see in campaign success over time?
3. How is the financial need for stigmatized conditions (from rehabilitation services to memorial/funeral costs) framed, and how does that vary across time and by population?

POTENTIAL INSIGHTS

With a preliminary understanding of the opioid crisis prior to analyzing our scraped data, we hypothesized that a threshold of 80% of a campaign's goal money raised may be counted as a successful campaign. The client also noted that the opioid crisis peaked in 2017, so we hypothesized that we'd see spikes in the number of campaigns and campaign activity in 2017 and 2020 (due to the coronavirus pandemic). In regards to financial needs for stigmatized conditions for the opioid crisis, we hypothesized that we'd see a higher campaign activity for campaigns launched in states with greater population density.

METHODOLOGY

DATA COLLECTION

In order to obtain the necessary data for our project, our team built a web scraper to pull campaign information from the GoFundMe website. Our scraper consists of three main Python files:

- ***gofundme_scrape.py:***
  Used to scrape campaigns associated with a key search term from the list of terms provided by our client. These URLs are added to *urls.csv*.
- ***campaign_info_bs4.py:***
  Used to scrape specific information for each campaign in *urls.csv* using both Selenium and BeautifulSoup
- ***data_formatting.py:***
  Used to organize our final list of campaigns into the format specified by our client (an Excel file with multiple sheets corresponding to campaigns associated with each keyword and a master Excel sheet with a list of all campaigns)

Our scraper utilizes the Selenium and BeautifulSoup Python packages to scrape a list of both dynamic and static content obtained from a campaign's HTML page. The information scraped is as follows (See Figure 0 below):

- campaign title
- campaign organizer
- beneficiary of the campaign
- date created
- location campaign is based in
- campaign description
- campaign tags
- number of followers, number of donors, number of shares for the campaign
- amount raised compared to the total goal (as a percentage)
- number of campaign updates
- number of campaign comments
- JSON information from the main JavaScript tag on each campaign page that includes: whether a beneficiary is a charity (T/F), whether campaign comments are enabled (T/F), whether campaign donations are enabled (T/F), whether a campaign is part of a team (T/F), whether a campaign is part of a business (T/F)

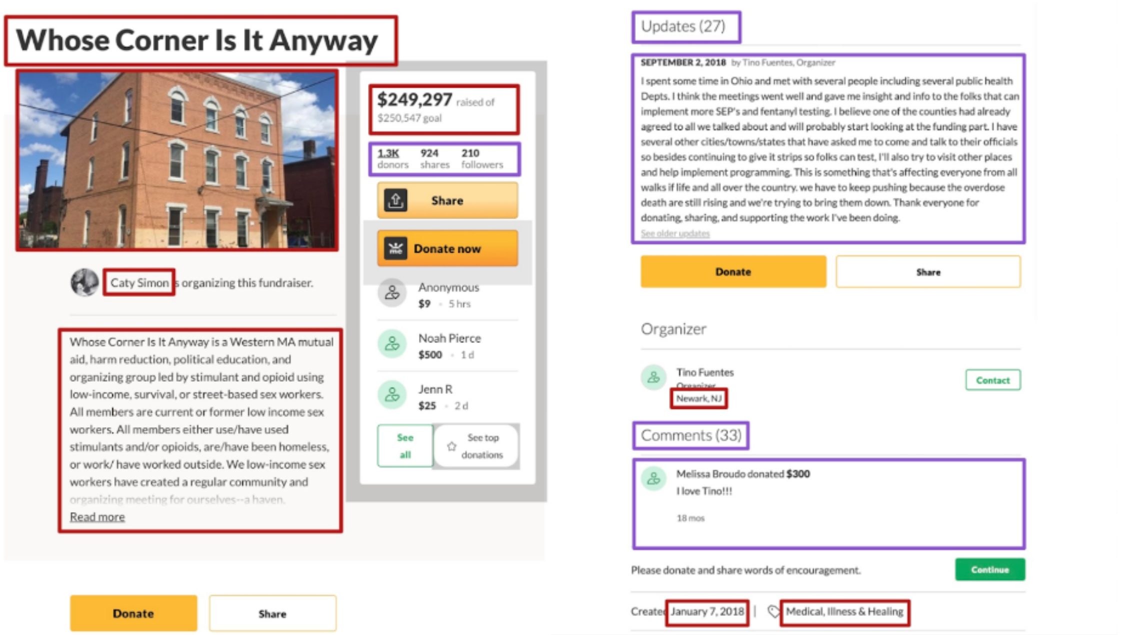


Figure 0 - Information in red scraped using BeautifulSoup & information in purple scraped using Selenium

Each keyword returned about 85 pages of search results, and each page displayed 12 campaigns. After removing campaigns based outside of the United States, our scraper outputted just over 8000 campaigns.

DATA PROCESSING

The data that was collected from GoFundMe required processing after each step of the web scraping. Because the initial URLs were collected by searching for keywords in GoFundMe, many of the campaigns that were found were duplicated under multiple keywords. Post-processing was implemented to ensure that each campaign we identified would only be scraped once and that all of the keywords that a campaign was found with were recorded for later use. All duplicate instances of URLs were collected in one row of the url.csv file with each keyword being imputed into a "keywords" column to keep track of all the applicable words (implemented in *./web_scraper/gofundme_scrape.py*).

In order to save time, the team divided up the work of scraping the campaigns. Once each team member completed their portion, the separate data files were merged into one full CSV file. After collecting the full set of data from GoFundMe for each of the campaigns, we needed to filter out some of the campaigns which were not going to be used for our analysis. Specifically, we wanted to remove international campaigns, rows that were completely empty (as they were the result of deleted campaigns that still had active links), and convert the campaign dates from Strings to Datetime objects for easier computational analysis using time as a metric. Once completed, we were left with a raw CSV file with all viable campaigns for analysis (implemented in *./web_scraper/merge_filter_csvs.py*).

Once the raw CSV file was ready, the last step of data processing was to isolate campaigns by keyword and reformat the raw data into an excel file based on the client's requests. Because we had already maintained a list of all relevant keywords for each campaign, we were able to regenerate the duplicated set of campaigns for each keyword. Once this was done, we could then group the campaigns by keyword and save each grouping as a separate Pandas DataFrame. An excel file was then generated with sub sheets that contained only the campaigns found for each keyword, as well as a sub sheet for all campaigns to make it easy to analyze either all of the data or only specific keywords (implemented in ./web_scraper/data_formatting.py)

FINDINGS & ANALYSIS

01

OVERARCHING QUESTION

How is drug use framed — a moral failing, or a medical condition in need of treatment? How does the framing affect success?

To analyze how drug use was framed and its effects on campaign success, we used a sentiment analysis model on the GoFundMe campaign descriptions. By examining the tone of campaign descriptions, we can understand how organizers present drug use cases, and we can also distinguish sentiments regarding keywords that were included in campaign descriptions and how those sentiments relate to campaign success. The model that we ended up using is VADER (Valence Aware Dictionary for Sentiment Reasoning).

VADER is a sentiment analysis model that produces a score pertaining to how neutral, positive, or negative text is and the “intensity” of that sentiment. VADER scores on a scale of -4 to 4, with -4 being text that is extremely negative and 4 being text that is extremely positive. The model is based on a dictionary that maps words to either positive, negative, or neutral sentiments and changes the intensity score based on the word’s relation to the text. VADER is also able to pick up on features of the text that are commonly found in social media, including Western emoticons and capitalization, and reflect that in its intensity scores. For this reason, we gravitated towards using this model as it is lightweight compared to other NLP techniques and we believed that GoFundMe descriptions would match the social media baseline of VADER as GoFundMe campaigns, like social media posts, are commonly shared amongst a user’s network.

By running sentiment analysis with VADER, we can get a sense of how a campaign is framed, positive or negative, and in turn, use that as new attributes to predict success using logistic regression or another model that better predicts based on our data.

Using the sentiment scores from VADER, we were able to cluster the data using a Gaussian Mixture Model to determine if there were any trends that clustering would pick up regarding campaign framing and success. We specifically sought to see if certain clusters contained more successful campaigns based on the total dollar amount raised and what percentage of the goal was reached, and whether or not certain keywords were more frequent in those clusters. When measuring success, the median for each clustering was used rather than the mean to avoid outlier effects. There were 8 clusters total, but two were excluded from this analysis since they only contained either one or two outlier campaigns. The results for the clusters were as follows:

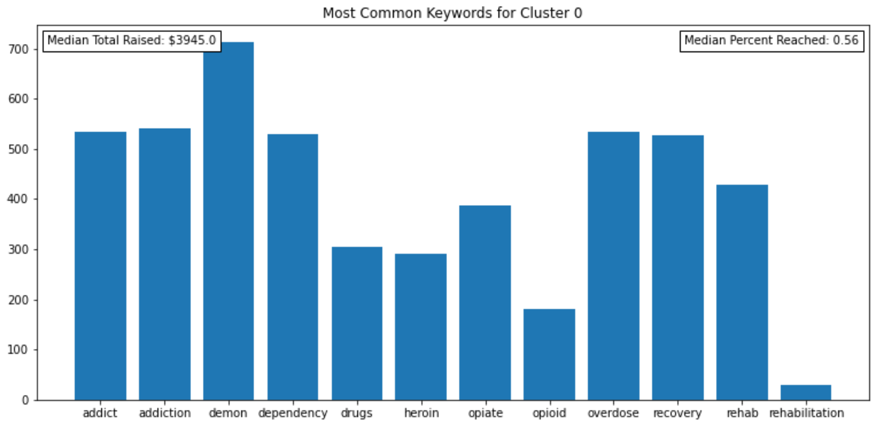


Figure 1.1 from "./analysis/Clustering.ipynb"

Cluster 0 had a median amount raised of $3945 and a relatively low median percentage reached of 0.56. Because both the total raised and the goal reached were small, it suggests that the campaigns in this cluster (which happened to be the largest cluster by far) were only moderately successful. The most common words for this cluster were “addict”, “demon”, “addiction”, and “dependency”. Relative to the other keywords, it is possible to state that these have a more negative sentiment.

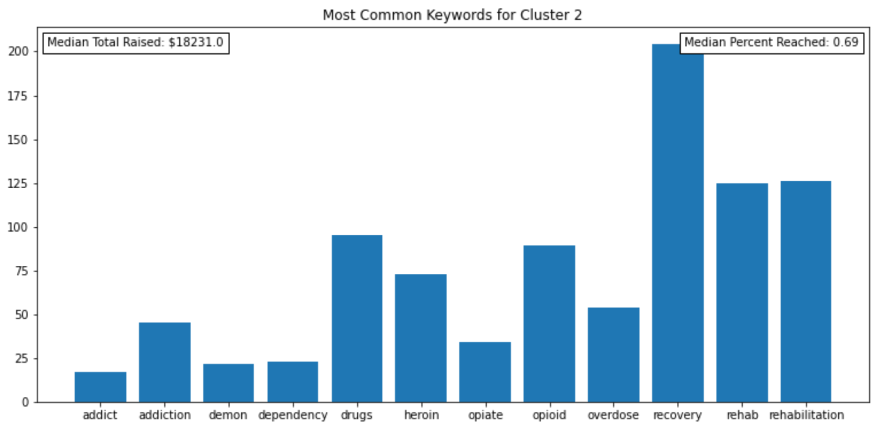


Figure 1.2 from "./analysis/Clustering.ipynb"

Cluster 2 had a large median total raised ($18231) and a relatively high median percentage reached (0.69). Because the total raised is high, the fact that the percent reached is high is indicative of these campaigns being very successful. The keywords most commonly used in this cluster were “recovery”, “rehabilitation”, “rehab”, and “drugs”. With the exception of “drugs”, these keywords can be seen as more positive than some of the others. “drugs” may be common because it is a descriptive word that likely showed up in successful campaigns because it provides context for the campaign’s purpose. Interestingly, the keywords that were most common in cluster 0 were some of the least common words in this cluster (“overdose” was also uncommon), implying that words that had a more positive connotation performed better while words with a worse connotation performed worse.

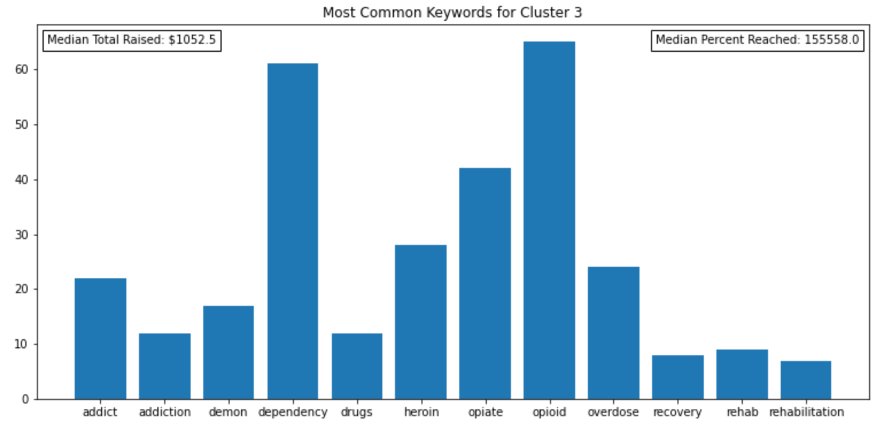


Figure 1.3 from "./analysis/Clustering.ipynb"

Cluster 3 had a small median total raised of $1052.50 and an unusually high median percentage reached of 155558. Upon further analysis, all of the campaigns in this cluster lacked a goal (either due to not being listed on GoFundMe or being lost when scraping), so it is most likely that these campaigns were grouped together due to the lack of some data skewing the percentage metric.

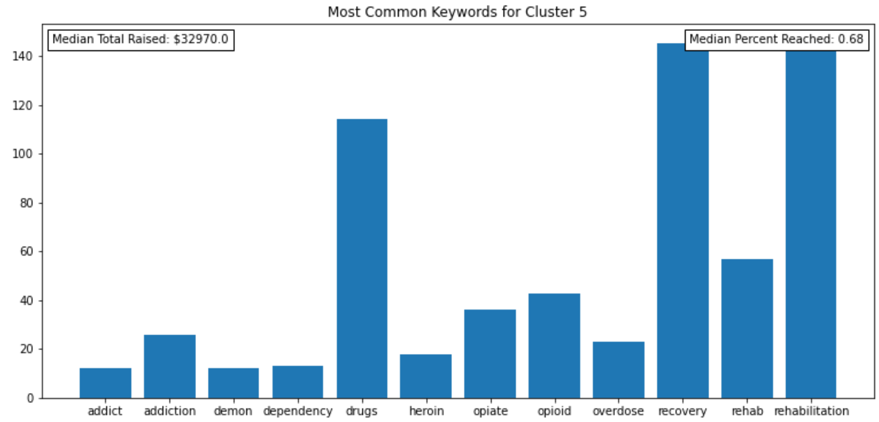


Figure 1.4 from "./analysis/Clustering.ipynb"

Cluster 5 was very similar to cluster 2. The median percent raised was 0.68 and the median total raised was $32970. Similar to cluster 2, the high amount of money raised and the high percentage of their goals reached suggest that these campaigns were also very successful (perhaps more so than cluster 2). The top keywords in this cluster were “rehabilitation”, “recovery”, “drugs”, and “rehab”. Like cluster 2, these keywords reflect words with positive sentiments and words that are descriptive (in the case of drugs). However, "drugs" could be considered a negative word depending on the context, so its inclusion as one of the most common words is interesting.

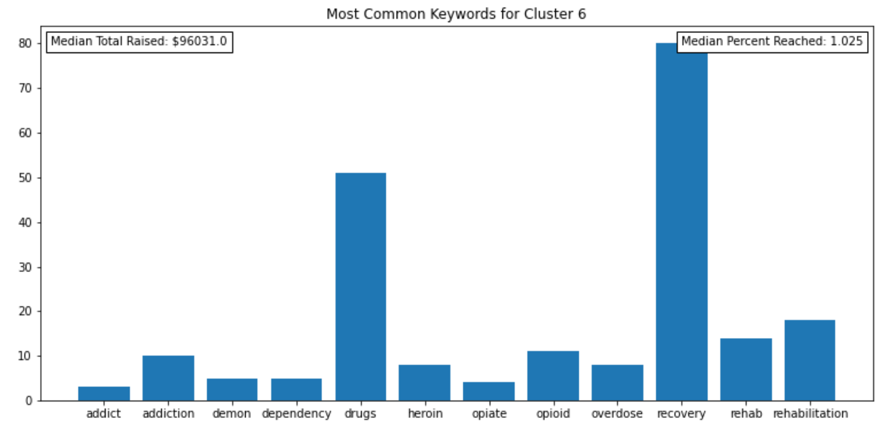


Figure 1.5 from "./analysis/Clustering.ipynb"

Cluster 6 is also similar to clusters 2 and 5, except both the median percentage of their goal reached (1.025) and their total amount raised ($96031) were much higher. It is possible that these were the most successful campaigns, although the percent raised value does suggest that not all of the campaigns had a goal listed. The top keywords are the same as for cluster 5, although interestingly “recovery” and “drugs” are drastically more common than the others.

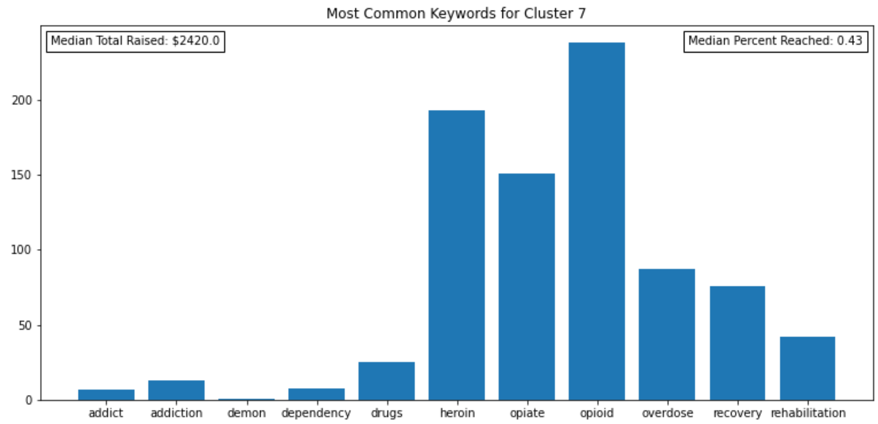


Figure 1.6 from "./analysis/Clustering.ipynb"

Cluster 7 was the worst-performing cluster, with a median percentage of 0.43 and a median total raised of just $2420. This suggests that the poor performance was due to campaigns not raising enough to meet their goals rather than missing data (which would reflect an unusual percentage). The most common words in this cluster were “opioid”, “opiate”, “heroin”, and “overdose”. These words can be seen as either having no strong connotation (i.e., more descriptive) or having a negative connotation depending on the context. It is possible that campaigns lacking strong connotative language and campaigns relying heavily on descriptive terms or negative terminology performed the worst.

To summarize the above results, campaigns that were more successful tended to include more positive-sounding words such as “rehabilitation” and “recovery”, while campaigns that used more negative terminology or more descriptive/dry language performed worse. This is in line with our plot showing the percent of the goal that was reached by keyword (see key question 3), which also demonstrated that campaigns with more positive keywords (in general) performed better. Combined, these results all suggest that framing drug addiction as a medical condition that someone heals from (“rehabilitation”, “recovery”, etc.) rather than a moral failure (“demon, “addict”, etc.) did improve campaign success. However, these relationships were not reflected 100% percent of the time, the relationships are not directly causational (GMM finds patterns but does not necessarily provide explanations for these results), and missing data may have influenced some results. More in-depth analysis may be needed to confirm these results. Additionally, the manner by which we identified certain keywords as “positive” or “negative” was a combination of sentiment analysis with VADER and human interpretation, meaning the conclusions that we made were partially subjective.

KEY QUESTION #1

02

What do we use to define a successful campaign, and why?

To assess the predictability of campaign success based on the set of attributes we have collected for our data, we built a logistic regression model using the quantitative attributes collected, focusing specifically on the year 2017, which saw a spike in campaigns and is known as a time where there was a peak in the opioid crisis. Labels representing campaign success were generated by defining a threshold in the percent reached of the campaign goal--anything above that threshold would be considered successful (encoded: 1) and anything below would be considered not successful (encoded: 0).

By building this model, we were able to tune the input definition of a ‘success threshold’ to more directly answer what an appropriate value would be and saw better model performance when using a threshold of 70% compared to the 80% we had previously theorized (specifically for the year 2017). The accuracy of the model is about 63%, which indicates that there is some predictive power behind the attributes we have extracted, which include scores that rank the descriptions and titles of campaigns as positive, neutral, or negative according to sentiment analysis performed (as described under our overarching project question), but there is room for improvement. Principal Component Analysis (PCA) was used to reduce the dimensions of the input attributes and address the issue of high correlation between different inputs, although this did not have a significant effect on the overall accuracy of the model (see Figure 2.1), neither did normalizing the features, and it also takes away our ability to measure the effects of individual attributes. We believed by generating more attributes that relate to the sentiment behind a campaign (for example, assigning weighted scores to each keyword used as another indication of how positive or negative a campaign is framed) we can further improve the model.

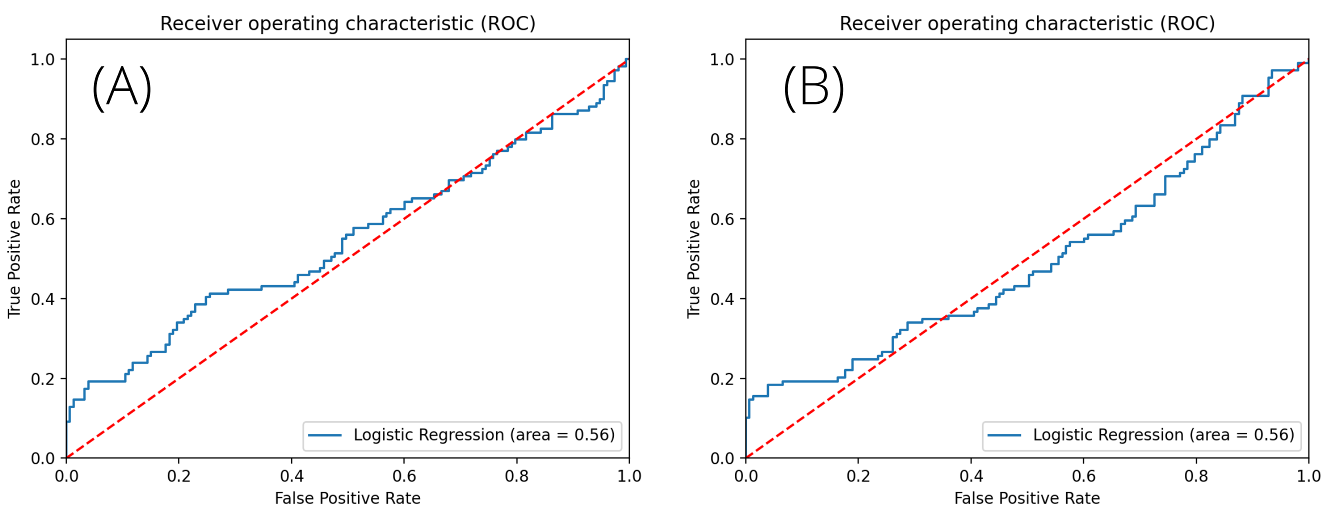


Figure 2.1 from "./analysis/logistic_regression.py". (A) Results from logistic regression model using compressed attributes, PCA resulted in 4 compressed attributes (B) Results of the logistic regression model using uncompressed attributes (no PCA).

Based on our discussions with the client, we decided to run with this idea of finding a way to assign some weighting to the keywords and reflect the connotations we believed revealed themselves in the clustering analysis, where more 'positive' key terms like 'recovery' and 'rehabilitation' were associated with greater campaign success. Then, we could compute a weighted sum of all the keywords associated with a particular campaign, an attribute that may be helpful to the model in its attempt to predict 'success'.

We figured if we could somehow rank the key terms and designate a score to each that represents the degree to which that term overtly references drug abuse, more broadly to rank them from positive connotations to negative connotations, this would capture the insight we were looking for.

To do this, we went through each cluster of campaigns (discussed in Section 1) and for each key term:

- multiplied the mean fraction of the total goal raised in that cluster by the total number of times that key term appeared in the cluster (to account for how much that key term was associated with that level of success)
- averaged across all relevant clusters
- normalized those values

The final weight ranking we hoped would show that higher-ranked terms are more 'positive' compared to lower-ranked terms.

WORDS RANKED	Average	Normalized
Recovery	88.51	1
Rehabilitation	55.6825	0.59974395
Heroin	55.45	0.59690915
Drugs	51.52375	0.54903754
Opioid	51.06625	0.54345938
Rehab	34.84	0.34561748
Opiate	29.2425	0.27736881
Overdose	24.6275	0.22109948
Addiction	16.1425	0.11764437
Dependency	8.31875	0.02225169
Demon	7.22375	0.00890068
Addict	6.49375	0.0056

For the most part, the final ranking seems to match our impression of the positivity versus negativity of the words, with 'recovery' ranked highest and terms like 'demon' 'addict' etc. ranked lowest.

An extra function was added to *./web_scraper/merge_filter_csvs.py* to append a column for the weighted sum over all of the keywords associated with a campaign score. After re-running the logistic regression and generating some summary stats, we found that although the overall accuracy of the model was not substantially improved by the addition of the weighted sum, this individual attribute contributed most significantly to the prediction power of the model, indicated by a significant p-value <0.05

To compare the results from 2017 to more recent years like 2020 and 2021, we trained the same model on the subset of data for each of these years, as well. Using the optimized 'success threshold' for 2017, we first noticed that the accuracy was much lower than before. It turned out that increasing the 'success threshold' from 0.7 to 0.8 for 2020 improved the performance of the model to the same level as that seen in 2017 (with an accuracy of approximately 63%). When the same was done for 2021, the parameter for 'success threshold' had increased even more to 0.9 in order to get comparable results.

Although we could get comparable accuracies around 60-63% by adjusting the input parameter of 'success threshold' between the different years, which revealed interesting patterns over time, we believe there is still room for improvement in the performance of the model overall.

Nevertheless, these results gave us an indication that as time has gone on, the bar for what defines success has risen. From here, we looked more deeply into what trends we could observe over time in order to explain why the model behaves this way when trained in more recent years.

03

KEY QUESTION #2

What are some of the trends we see in campaign success over time?

Some of the trends we uncovered were in relation to the keywords included in campaigns. After getting all of our data, we looked for trends describing the popularity of keywords over the years. This finding would also connect to our analysis on campaign success relating to each keyword and campaign success over time. If certain keywords were more used during certain years, then we can more likely predict/verify campaign success during those years based on earlier models. Because the number of campaigns differed each year, we normalized the number of campaigns in relation to each keyword by year. To do so, we divided the number of campaigns associated with a keyword by the total number of campaigns that year. Based on the results, we found that harsher keywords associated with drug use (for example, ‘overdose’, ‘addict’, ‘demon’, etc.) were described more in campaigns in the early and mid-2010s. In contrast, words associated with drug use recovery (for example, ‘rehab’ and ‘recovery’) were much more popular in the late 2010s and early 2020s. There was also a peak in 2017 when it comes to the keyword opioid, which follows our client’s prediction since that year was the height of the opioid epidemic.

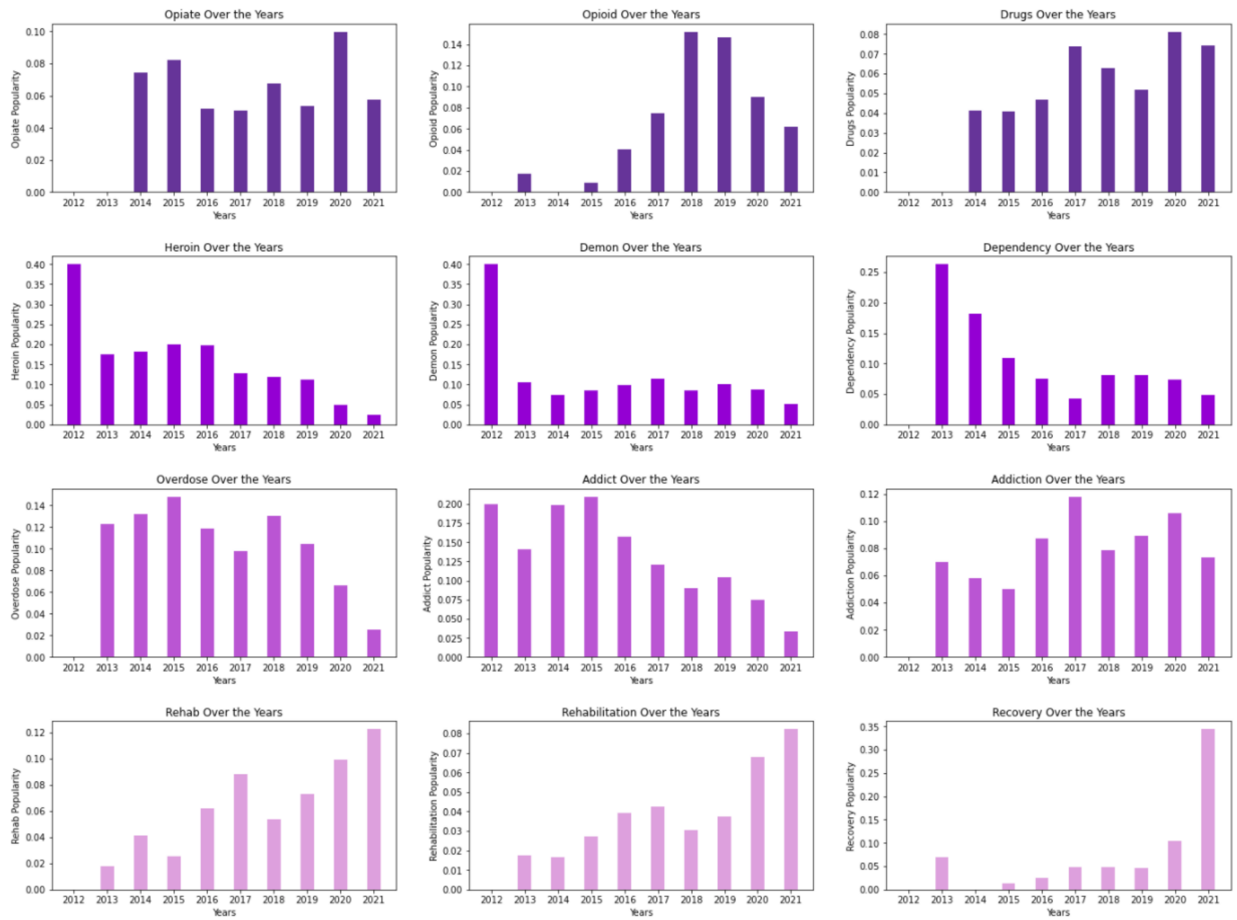


Figure 3.1 from "./analysis/figures.ipynb"

Additionally, we looked at the average amount of money raised each year. This would also help us get a better understanding of how strongly correlated keywords and campaign success were. We used the median method of calculating the average to avoid severely skewed results from outliers.

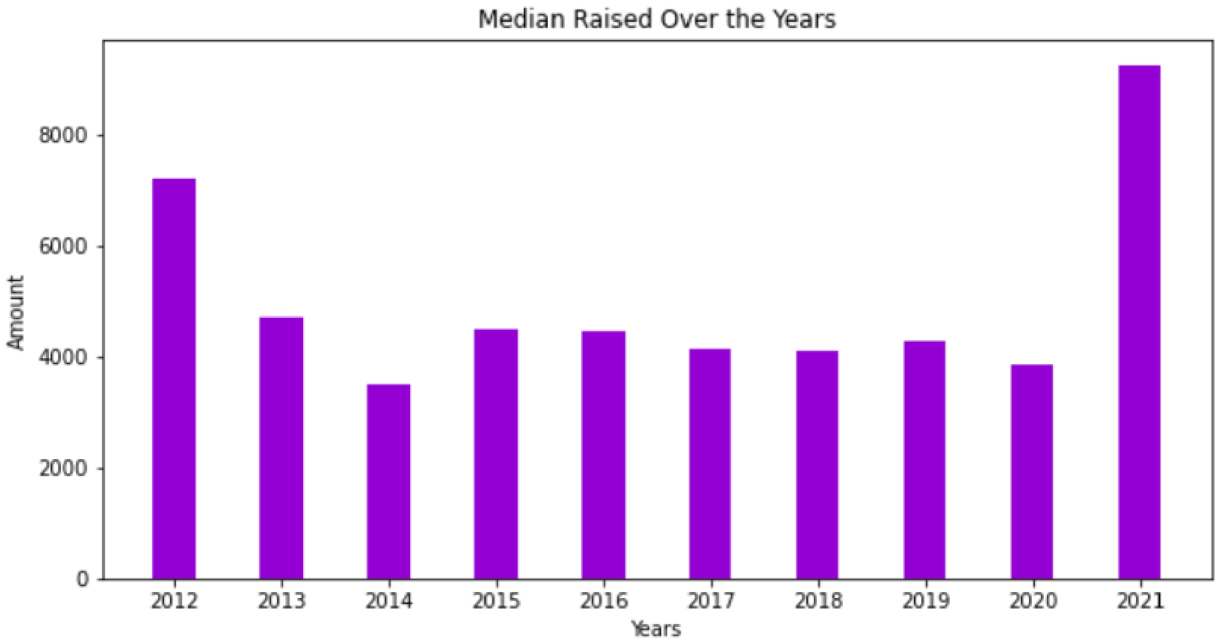


Figure 3.2 from "./analysis/figures.ipynb"

These results correlate to trends we found in other analyses. The peak in 2021 corresponds to the high number of campaigns related to the keywords “recovery,” “rehab,” and “rehabilitation,” which were keywords found in clusters with high percentages of success rates.

04

KEY QUESTION #3

How is the financial need for stigmatized conditions framed? How does it vary across the population?

Based on figure 4.1, we see a higher concentration of campaigns in more populous states (CA, NY, FL, TX). Variation in numbers can be explained by the population density of urban areas in said states. This may indicate how the financial need for opioid treatment, whether classified as deviant or deserving, is deemed as more necessary in more urban areas. This idea is further reinforced in the Map of All Campaigns figure, as the campaign density is heavier in the coastal US. Also, in figure 4.2, we see the total number of campaigns per year spike in 2017 and 2020. The client mentioned that the US opioid crisis peaked in 2017, but the global spike in 2020 may be attributed to the coronavirus pandemic and its effect on opioid users.

Number of Campaigns by State

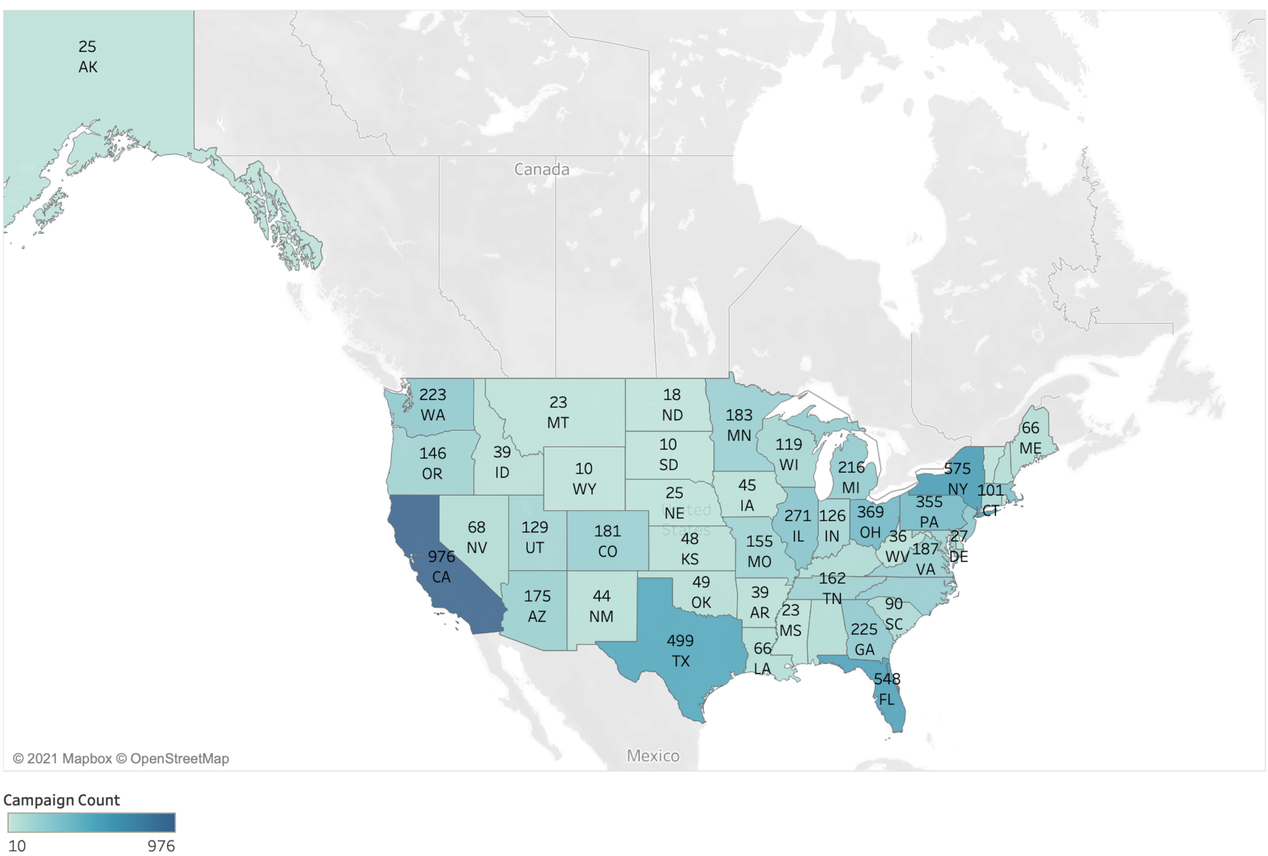


Figure 4.1 from Tableau

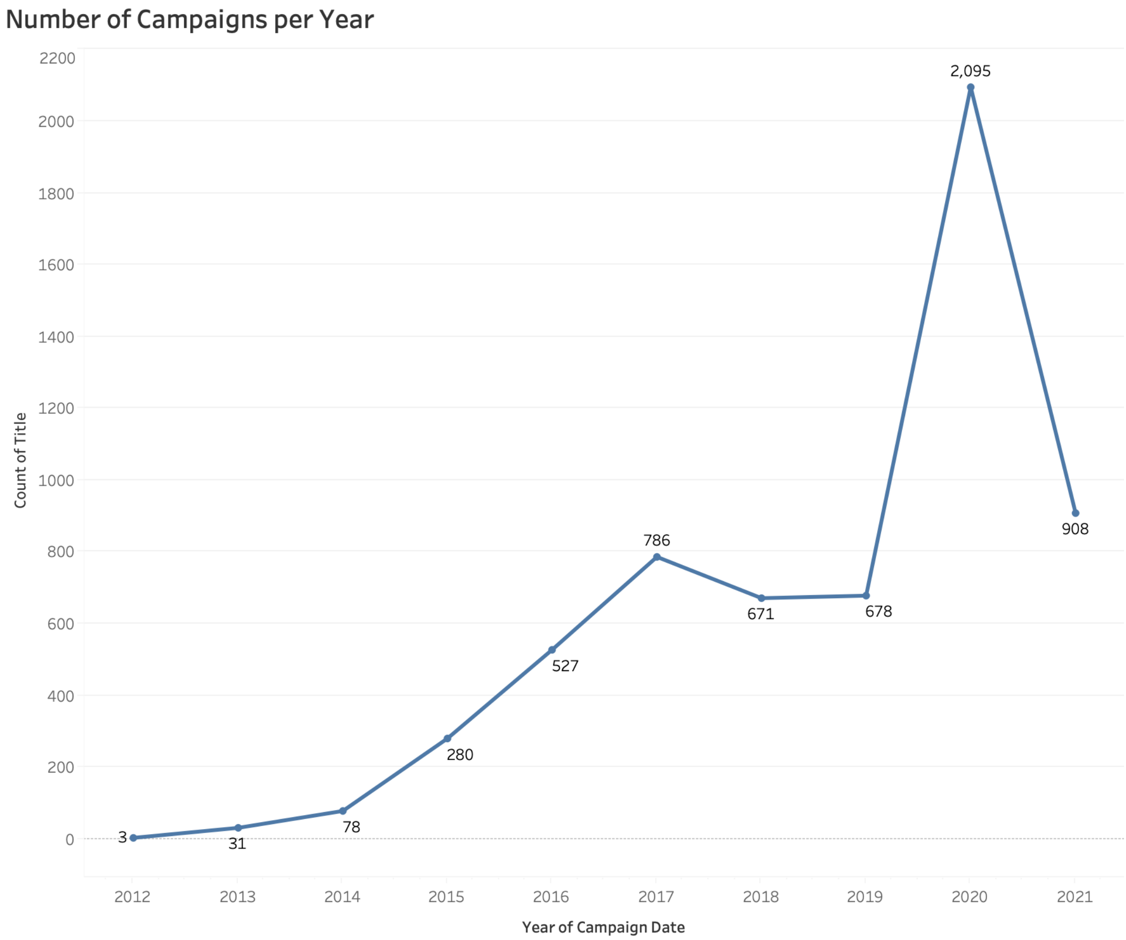


Figure 4.2 from Tableau

Based on figure 4.3, campaigns that pertain to more positive keywords (like rehabilitation and recovery) or refer to opioids in general terms (like drugs) are more successful at raising funds when compared to campaigns that pertain to more negative keywords (like demon and overdose) or refer to opioid addiction explicitly. The keyword “rehabilitation” is particularly striking as it has the least amount of campaigns associated with it but has roughly the same amount of campaign success (0.15% of the money raised) as the keyword “addiction” (0.16% of the money raised) despite the fact that addiction had the highest amount of campaigns associate with it.

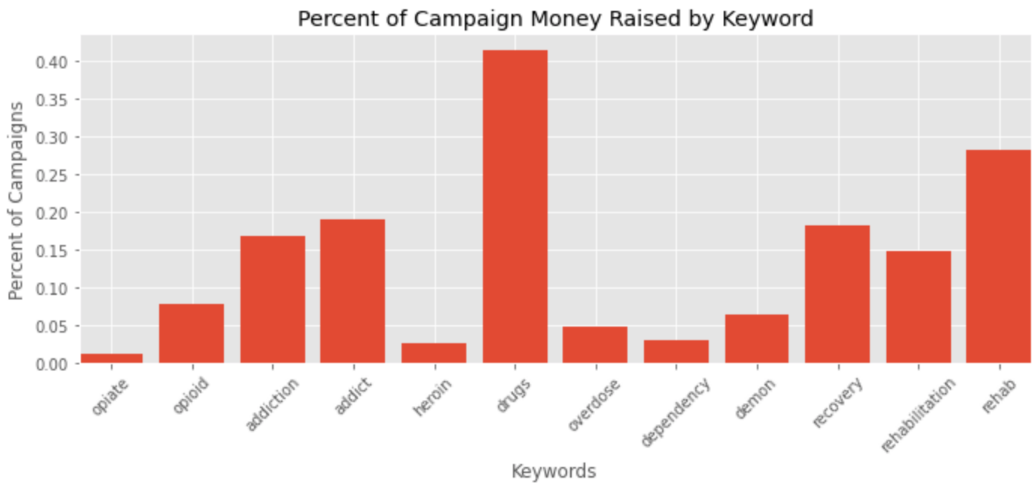


Figure 4.3 from "./analysis/percent_campaign_money_by_keyword.ipynb"

One of the ways we tried to figure out which words were more “positive” or “negative” was to use an NLP model (VADER) for sentiment analysis where keywords with descriptions with higher scores had more positive sentiments than others. Keywords like “rehabilitation” and “rehab” had more positive sentiment scores than words like “opioid” and “opiate”, but it is important to consider that the number of campaigns associated with a keyword could affect these sentiment analysis scores.

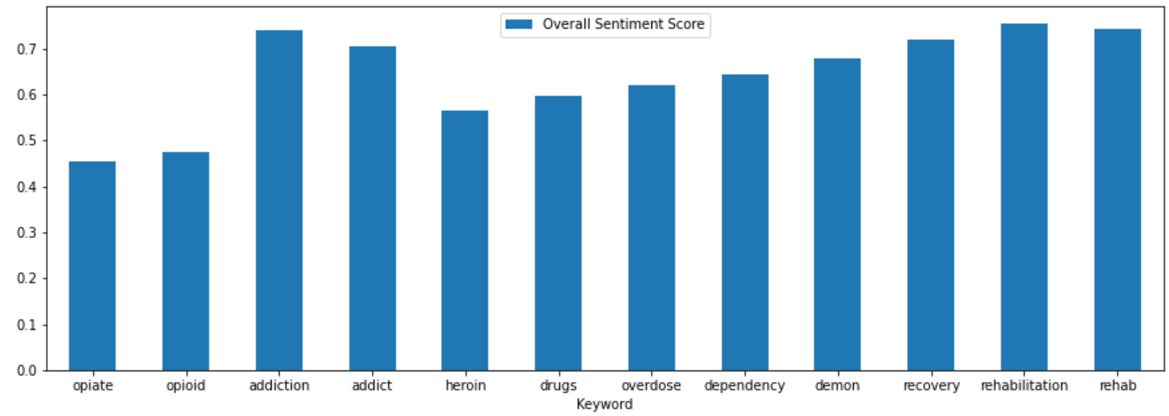


Figure 4.4 from "./analysis/processing_sentiment_locale_features.ipynb"

Based on figure 4.5, we can see that there is a spike in campaign success in rural areas. This may be due to the fact that rural areas tend to have a closer and smaller community. Therefore, people are more likely to know each other, which would prompt them to donate, on average, a larger amount of money. Rural areas may also have smaller goals than campaigns from metropolitan areas, so goals are more likely to be reached.

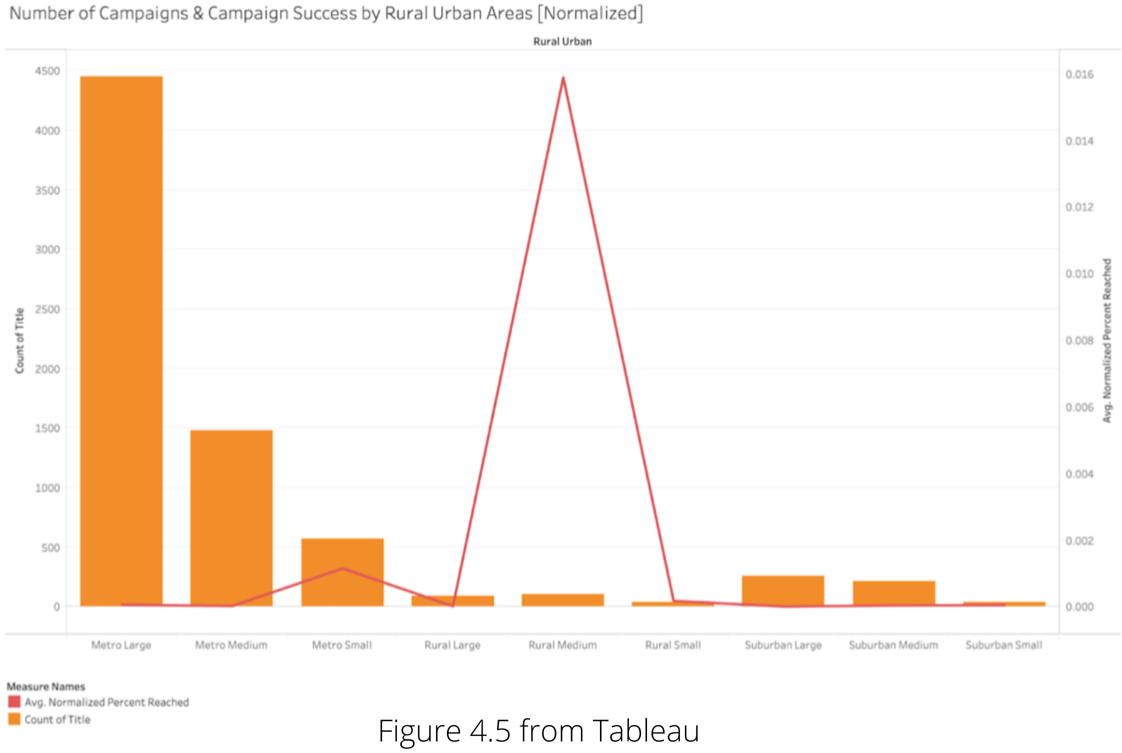


Figure 4.5 from Tableau

PROJECT EVALUATIONS

DATA LIMITATIONS

One of our main challenges with the data was dealing with errors that resulted from running our scraper on GoFundMe's website. Since we retrieved GoFundMe campaigns from the website by searching for campaigns based on keywords, some of the campaigns that were further in the search results had been closed and the URLs for these webpages were no longer functional and hence contained no data. We also encountered different types of connection errors when running our scraper that resulted in missing information for a few of our campaigns. Another limitation we encountered was that there was no way for us to select campaigns based on year so search results for keywords were skewed towards campaigns in earlier years. We were also provided a larger list of keywords from our client, but for the scope of this project, we scraped campaigns associated with keywords that our client had indicated to us were of high priority. Lastly, due to the nature of the keywords in our list and the fact that campaigns were obtained from GoFundMe search results, there is a high likelihood that campaigns that were not associated with opioid addiction were scraped as part of our data collection and analysis. Our client is aware of this and will take this into consideration when conducting her own content analysis and selective discourse of the campaigns.

PROJECT COMPLEXITY

Unlike many of the other projects that were available, we were starting out with no data and therefore had to implement this project nearly from scratch. Given the challenges associated with web scraping, we relied heavily on combining different tools and techniques to gather all of the data that we needed as quickly and comprehensively as possible. The process of formatting and filtering the scraped data also required careful consideration of how the data was going to be used for analysis. Once we began the analysis of the data, we needed to develop our own ideas for how to approach the data and determine what tools were viable to answer the questions we were focused on. We took many avenues to answer these questions: implementing sentiment analysis to quantify the underlying meaning inside the text of the campaigns, clustering to understand the trends between campaign success and language, logistic regression to understand if the data was able to be used to predict success and to determine which features were the most impactful, temporal and geographic analysis to understand trends in keyword popularity, campaign frequency, and success.

ACKNOWLEDGEMENTS

Special thank you to Heather Mooney, Jane Feng, Professor Lance Galletti, Greta Bruce & the team at BU Spark!

SPARK!

BOSTON UNIVERSITY