

| |
|--|
| Weekly Project SCRUM Report |
| Date: 4/1/2021 |
| Project Name: Deviance or Deservingness? Opioids, Morality, and Economic Precarity |
| Team Members: Jayden Font, Della Lin, Nikita Jakkam, Michalina Jadick, Divya Gowravaram |
| <p>Tasks Completed This Week:</p> <ul style="list-style-type: none"> - Added code to scraper to get campaign photo URLs - Moved code to get # of campaign updates and # of campaign comments into the main scraper file - Met with Dharmesh to ask about preventing connection errors / scrape detection while scraping - Adding anti-scraping measures (randomized sleep, user-agent for requests) to scraper to prevent the 403 connection errors we have been getting when trying to scrape - Solidifying plans for data analysis - Split up the scraper to run among the 5 of us overnight to spread out processing <ul style="list-style-type: none"> - Some of us had errors which means it's taking an extra day to fully scrape everything (breaking the list down into smaller batches), but have enough to move forward with filtering and identifying preliminary trends/correlations |
| <p>Plans for Next Week:</p> <ul style="list-style-type: none"> - We want to start creating data visualizations such as: <ul style="list-style-type: none"> - Plotting the number of campaigns by year to visualize trends - Plotting the number of campaigns by keywords to see which keywords have more campaigns - Plot the average donation goal reached for each keyword - Generating a pairplot with Seaborn to check for correlations in data - After collecting all data, we will need to drop all international campaigns and blank campaigns - Sort URL dates for data analysis on trends - Meeting again with Dharmesh to fix connection issues when scraping |
| <p>Obstacles and Questions:</p> <ul style="list-style-type: none"> - Need to talk to Heather about our idea ranking and classifying the key terms to see if we can identify trends in more overt vs. covert references to drug abuse. Would it be useful to try to come up with a ranking and try to ID trends from there? Is that too subjective? |
| <p>Met with the client recently? When is the next meeting with the client?:</p> <ul style="list-style-type: none"> - Meeting for March 31st was replaced by email updates - Meetings are biweekly so next meeting with client is April 7th |

URLs to be scraped:

Mica: 0 - 2000

Della: 2000 - 4000

Jayden: 4000 - 6000

Nikita: 6000 - 8000

Divya: 8000 +