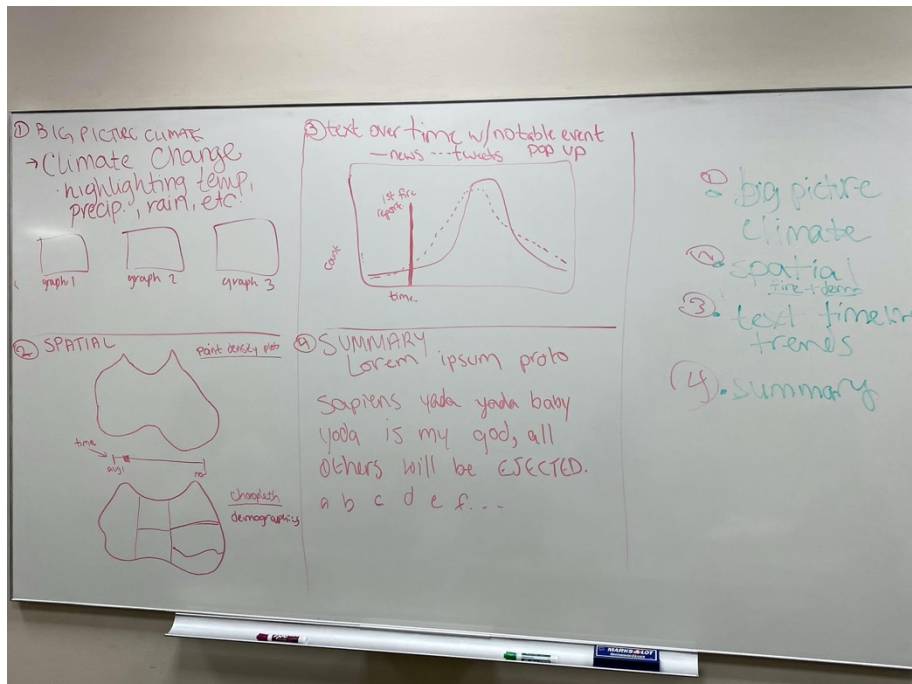


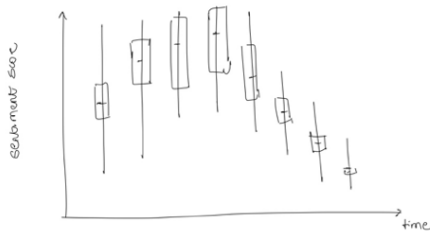
Australia Fires Project Book

Group K

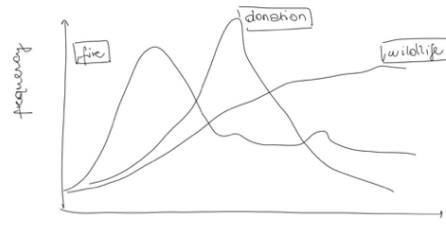
Initial Overview:



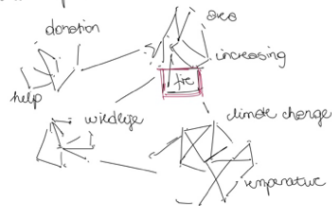
1) how the sentiments of texts have changed over time



2) Topical analysis: top 10 most frequently used words



3) Network analysis
What topics are mentioned? to use any specific points?



How our plans changed after the initial proposal:

Changes to Overall Presentation Layout

- At the start, we envisioned telling the story via a single scrolling page, moving from weather to fire to text data. However, after creating all of our plots, we realized that the original form would have been quite extensive and, frankly, too long to expect a reader to push through. Instead, we changed the overall layout of the Shiny app to have an initial landing page which introduces the topic/story and gives a quick guide to the forms of data/analysis available in the app. We then used click-through links and a set of tab buttons to link to separate pages for each of the three forms of data and analysis. We then chose to elaborate and tell the story from the three different perspectives via each of the pages. As a result, the overall structure allowed for a much more concise presentation while also allowing for sufficient elaboration on each of the plots to improve interpretability.

Changes to weather analysis

- We really wanted to show how climate change played a role in the wildfires but quickly realized it would be challenging to show different weather metrics on the same graph. In order to achieve this goal we used a weather index for each metric with the year 2010 as the baseline. This allowed us to look at the metrics together and compared how they had changed. We hadn't originally planned to include a data table but after realizing the complexity of each weather metric it seemed like a good way to show each of the metrics in a more granular way. Lastly, we wanted to show how weather changed over time but decided that it would be hard for the reader to see how different this year was from last. Therefore we created a variable that looked at the percent change of each weather metric in 2019/2020 vs 2018/2019 season to identify which metrics might be contributing to the climate change the most.

Changes to Spatial Visualizations

- For the spatial visualizations, several changes occurred to the designs as the project developed. First, instead of hex bins, we chose to use a combination of choropleth maps and clustering techniques to offer two separate visualizations for understanding the location of fire occurrences. We also had the chance to work with Australian census shape files and made use of the various levels of granularity offered to offer different levels of visualization. The fire data was also limited to the month of January, so we chose not to include a time subsetting option as it would have added very little value for the excess needed to incorporate the feature. We did, however, stick with the choropleth map for the population data and again made use of the census shape files.

Changes to NLP section

- Compared to the original plans, we had to change the NLP and visualization part of the project as one of our group members has dropped out. We were originally going to look at a news dataset and twitter but the news dataset was challenging to download in a manageable way. For the final project we kept the Twitter dataset.

- Compared to the initial plans, two out of the three planned graphs are available on our website and the third was attempted but as the tweets did not show significant topical changes over time, that graph has been eliminated. The sentiment boxplot and the network analysis has been slightly modified, but the results came out as expected. Additionally, the sentiment of the tweets was further analyzed by creating a graph that detects sentiments beyond positive and negative and the results of that graph provided a linkage between temperature changes and the fire events as surprise as a sentiment was the least significant. Furthermore, the most used positive and negative words were plotted to understand the factors that drives each side.