COMPSCI 590V - Project Proposal

Authors: *Nikita Bhalla, Apurva Bhandari, Deepanshu Pariyani, Maggie Zhang*

ABSTRACT

Wildfires are a natural part of many ecosystems, providing essential cleansing of the forest floors from the debris. However, because of the rising temperature of the earth, reducing rainfall, high winds and many other manmade and natural reasons, Australian fire seasons lately have been unprecedented with fire seasons being longer, severe and more frequent. The goal of our project is to design and implement visualization systems for analysing the spread and impact of the bushfires that occurred across the Australian region from late 2019 to early 2020. Along with demonstrating the presence and spread of fire across different geographical areas, we are also going to show its impact on the humans, wildlife and the overall climate of the area in order to emphasize the magnitude of the devastating impact this event has created.

We believe the findings and insights from our work could be helpful for the local communities, environmentalists and government agencies in answering a multitude of questions. Our work in particular will help gain insights into why this fire season is so much worse than the previous ones based on climate related variables like temperature and rainfall. Also, our visualizations can be helpful in coping with the future fire seasons and prevent them from becoming disastrous. We also aim to help concerned users and any other third parties in providing an assessment of the overall damage caused by the fire.

INTRODUCTION

1.1 DATA AND PROBLEM SELECTION

Problem Selection:

Bush fire season is nothing new to Australians. They've shaped the nature of the continent over millions of years. Bushfires have killed approximately 800 people in Australia since 1851— almost 5 people per year. This year alone killed at least 34 people, an estimated one billion animals and burnt an estimated 46 million acres. Research has shown that human-caused climate change created high-risk conditions which made the widespread burning at least 30 percent more likely than in a world without global warming.

While it is clear that 2020 is the worst fire season that Australia has ever seen, how bad is it compared to previous years? Has the world seen a bushfire like this before? What made this summer so calamitous? How are these fires distributed over time? Should containing the fires be a top priority now? How do these fires impact the local as well as the global environment?

These are a few intriguing questions that make this problem worth exploring. Our visualizations also serve as a reminder that we, humans, may have been one of the causes behind this disastrous natural event.

Data Selection:

The brightness data at different locations and at different times is a publicly-sourced data provided by NASA**[1]**. It is near real-time data collected from the satellites using two instruments, MODIS (M6) and VIIRS(V1). NASA's Terra and Aqua satellites produce snapshots of events happening on ground. The satellites detect fire hotspots by measuring the infrared radiation emitted by the blazes. The available data is for the time period - Sept 1, 2019 to Jan 31, 2020. The dataset has 2 files sourced from MODIS C6 and VIIRS 375m instruments aboard the satellites with attributes ranging between 14-15 in number providing information about the intensity of fire at a particular location and contain 148k and 106k records in total respectively.

The credibility and the wide number of features available made us choose this dataset. Also, the fact that the data is near real-time will enable us to do a real-time analysis of the fire.

Apart from this we have used an open source data available on github,TidyTuesday**[2]**, which contains climate related dataset like temperature and rainfall sourced from Australian Bureau of Meteorology (BoM). The data ranges over the time period - 1858 to Jan 2020 letting us analyse trends over a long period of time. Additionally, the fact that it contains data till 2020 makes it the perfect dataset for our trend evaluation.

The attributes available in both of these datasets will help us in answering questions pertinent to the disaster management team, the local community, and the environmentalists.

Acknowledgement: *We acknowledge the use of data and imagery from LANCE FIRMS operated by NASA's Earth Science Data and Information System (ESDIS) with funding provided by NASA Headquarters.*

1.2 USERS AND TASKS IDENTIFICATION

We identified three main users of our visualizations which include but are not limited to:

1. Local Residents
2. Environmentalist Specialists
3. Disaster Management Organizations/Government Agencies

a) Local Residents:

Major Concerns of the local residents may include:

- Understanding the rate of human casualties due to the fires to determine if continued residency is still safe
- Examining the periphery of fire-affected regions as well as the rate and direction of the fires' expansion to assess if an evacuation is needed and when
- Evaluating the air quality and temperature fluctuations to prepare any necessary supplies or equipment should they remain at home

b) Environmentalists:

Major Concerns of Environmentalists may include:

- Analyzing the number and intensity of fires and their effect on long term climate change to slow down or prevent future natural disasters
- Comparing the trends of fire damage and spread of the Australian fires with other geographical locations, such as the Amazon rainforest fires, to assess the success of fire containment and suggest improvement in the government's containment plan where needed
- Examining the effect of the fires on wildlife and forest/land cover of the region over the years.
- Understanding if the human-caused climate change created a high-risk environment increasing the likelihood of these fires.

c) Disaster Management Organizations/ Government Agencies:

Major Concerns of Disaster Management Organizations and Government Agencies may include:

- Assessing the distribution of fires across areas in order to evaluate and prepare evacuation strategies
- Studying the status of the Australian fires over the years to determine if it is a serious growing trend that needs to be prioritized over other government goals
- Understanding the rate of human casualties due to the fires to determine if continued residency for the local residents is still safe

- Evaluating Air Quality index to help make long-term plans to fix it

In Section 1.3 we break down these high-level goals into more specific tasks.

1.3 DATA AND TASK ABSTRACTION

Data Abstraction

MODIS/VIIRS DATA: This publicly-sourced data is provided by NASA collected from satellites using two instruments, MODIS(M6) and VIIRS(1). For our visualizations we used data for both Australia and Brazil.

Label	Description	Range	Cardinality	Type
Latitude	Location coordinate (Latitude) of the fire. The point is centered at a radius of 1km from the fire location.	[-42.8,-9.39]	N/A	Quantitative
Longitude	Location coordinate (Longitude) of the fire. The point is centered at a radius of 1km from the fire location.	[113,153]	N/A	Quantitative
Brightness -	This indicates the channel 21/22 brightness of the fire at a particular location measured in Kelvin.	[300,507]	N/A	Quantitative
Scan	The scan value depicts the spatial-resolution in the east-west direction of the satellite's scan.	[1,4.8]	N/A	Quantitative
Track	The track value depicts the spatial-resolution in the north-south direction of the satellite's scan.	[1,2]	N/A	Quantitative
Acq_date	Acq_date is the date of MODIS acquisition	09/01/2019 - 01/31/2020	NA	Quantitative
Acq_time	Time of acquisition/overpass of the satellite in UTC.	0000 - 2400	NA	Cyclic
Satellite	This data attribute tells which satellite is the MODIS data sourced from.	[Aqua, Terra]	2	Categorical
Instrument	This data attribute tells which instrument is the data collected from,VIIRS or MODIS	[VIIR, MODIS]	2	Categorical
Confidence	This value helps understand the quality of individual hotspot/fire pixels, based upon some intermediate quantities used in the detection process..	[0 - 100]	N/A	Quantitative
Version	Identifies the collection and source of data processing. Sources are either Near Real-Time (NRT) or Standard Processing (collection only).	[6.3, 1.0, 6.0NRT, 1.0NRT]	4	Categorical
bright_t31	Brightness temperature is a measure of photons at a particular wavelength received by the spacecraft presented in units of temperature.	[266 - 400.10]	N/A	Quantitative (Sequential)

frp	Fire Radiative Power of the pixel measured in megawatts (MW)	[0 - 7,400]	N/A	Quantitative (Sequential)
Daynight	Daytime fire or Nighttime fire	[D, N] D = Day, N = Night	2	Categorical

Temperature & Rainfall: This climate data was gathered from the Australian Bureau of Meteorology (BoM) where a number of weather stations were chosen, based on their proximity to major Australian cities such as Sydney, Perth, Brisbane, Canberra, and Adelaide.

Rainfall data(179k rows × 11 columns):

Label	Description	Range/Values	Cardinality	Type
Station_code	Station Code	There are 7 unique station code: 9151, 23011, 40383, 40913, 66062, 70351, 86232	7	Categorical
City_name	City name	Rainfall data was sourced from - Subiaco, Sydney, Melbourne, Brisbane, Canberra, Adelaide	6	Categorical
Year	Year	1858 - 2020	N/A	Sequential
Month	Month	1 to 12	12	Cyclic
Day	Day	1 to 31	31	Cyclic
Rainfall	Rainfall in millimeters	Min: 0.0, Max: 327.6, Mean: 2.398, Median: 0.0	N/A	Diverging
Period	How many days was it collected across	Min: 1, Max: 23	15	Ordinal
Quality	Certified quality or not	'Y', 'N'	2	Categorical
Lat	Latitude	-31.96, -34.92, -27.51, -27.48, -33.86, -35.31, -37.83	7	Categorical
Long	Longitude	115.79, 138.6 , 153.05, 153.04, 151.21, 149.2 , 144.98	7	Categorical
Station_name	Station name	There are 7 different stations in the data. Eg: "Melbourne Botanical Gardens", "Brisbane" etc.	7	Categorical

Temperature Data(528278 rows × 5 columns):

Label	Description	Range/Values	Cardinality	Type
City_name	City name	N/A	7	Categorical
Date	Date	1910-01-01 to 2019-05-31	N/A	Sequential
Temperature	Temperature in Celsius	-11.5 to 48.3	N/A	Diverging
Temp_type	Type can be either “min” or “max”. For each day, min and max temperatures are specified.	“min” & “max”	2	Categorical
Site_name	Actual site/weather station name. For example “Melbourne (Olympic Park)” where the city is Melbourne.	N/A	7	Categorical

Along with the aforementioned data, we plan to utilize the following data to supplement the key data:

Air Quality Index (AQI): Sourced from the World Air Quality Index (WAQI) project team, this data records the levels of particulate matter/pollution, O_3, NO_2, SO_2, and CO in the air over a period of time.

Human Bushfire Casualties: This data reports the number of human deaths due to bushfires at a given Financial Year and the percentage of deaths that year contributed to the total bushfire deaths between 2010 and 2020.

Task Abstraction

1. Local Residents:

Q1. Human casualties: “*I want to see the rate of human casualties and understand if it is a serious growing trend”*
- Derive trends
- Locate attributes
- Compare values
- Identify correlations

Q2. Fire expansion rates: *“I want to know the rate of the fire and direction expansion in order to know when to call an evacuation.”*
- Discover features
- Browse dependencies
- Summarize dependencies

Q3. Air Quality Index (AQI): *“I want to see the trends of the AQI to predict and prepare if the fires continue to get worse”*
- Present trends
- Locate attribute
- Identify value

Q4. Change in temperature: *“I want to see the trends in temperature fluxuations”*
- Present features
- Browse outliers
- Identify extremes

2. Environmentalists:

Q1. Long term Climate Change: *“I want to see if the intensity and number of the fires have had a significant effect on the long term climate changes”*
- Derive trends

- Locate attributes
- Identify correlations

Q2. Deforestation/Wildlife: "*Do the fires have a significant impact on the wildlife population and the vegetation cover?"*
- Discover Correlation
- Locate attributes
- Compare values

Q3. Fire trends in Australia vs. Brazil (Amazon Rainforest): *"I want to see the fire trends of other geographical locations to understand if other locations have successfully minimized damage/wildfires or are they in the same boat as we are?"*
- Derive trends
- Locate attribute
- Compare values

3. Disaster Management Organizations/ Government Agencies:

Q1. Fire density/distribution: "*I want to know which areas have the highest distribution of fires in order to understand the critical areas for evacuation"*
- Present features/shape
- Explore outliers
- Identify distribution

Q2. Australian Fires over the Years: "*I want to observe the status of the fires and if it has become worse over time"*
- Discover trends
- Locate attribute
- Compare value

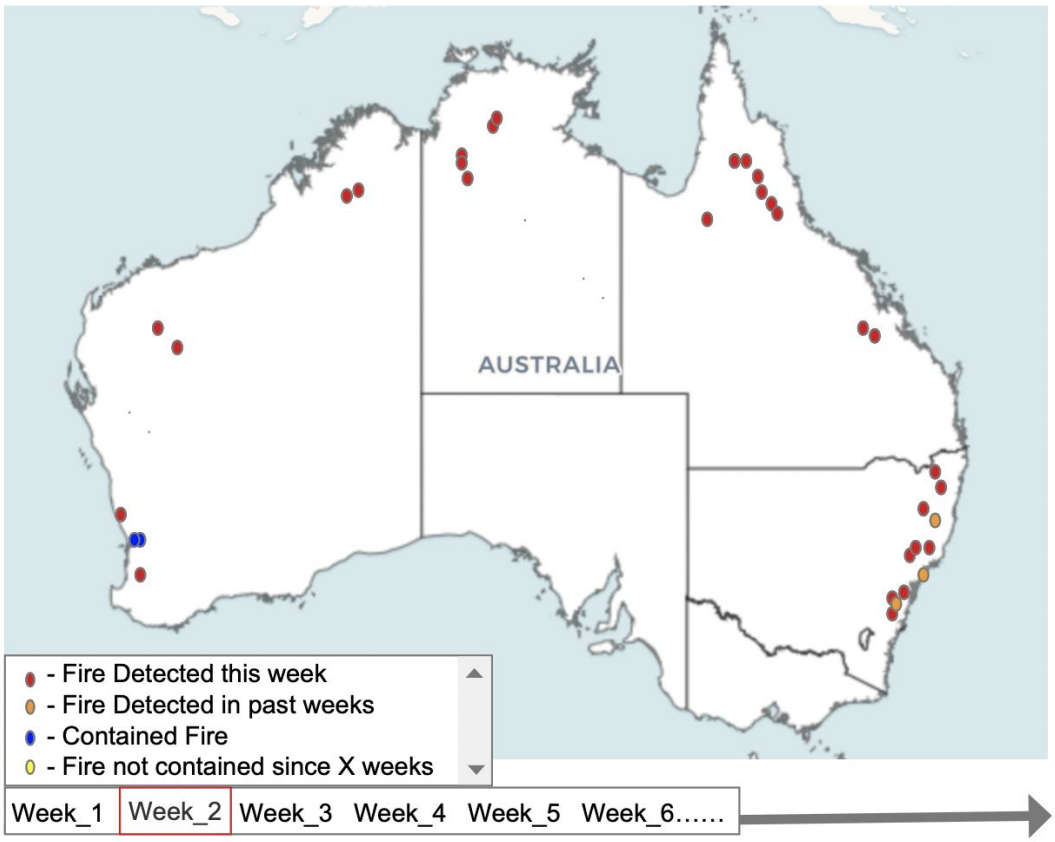
Q3. Human casualties: "I want to see the relationship between the casualty rates caused by the fires and the number/intensity of fires over the years"
- Discover trends
- Locate attribute
- Compare value

Q4. Air Quality: *"I want to observe the air quality index over the years to prepare long-term plans to fix it"*
- Discover trends
- Locate attribute
- Compare value

1.4 DESIGN

To achieve the goal of answering the above mentioned questions and user specific tasks, we propose the following three visualizations along with the design choices and reasoning as to why they are the best choice for this particular situation:

1. **Heat Point Map** - The main questions that we are trying to address via this visualization are
   - Expansion of Forest Fire and the direction of expansion in Australian Region
   - Understanding or assessing the distribution of fire across different Australian Regions in order to evaluate and prepare evacuation strategies
   - How much of the detected forest fires have been contained



Description: An Australian map as the base layer in which the red dots point out the fire detections and are being plotted for every week. If an old detected fire has been contained, the color changes to blue and if it crosses x weeks of time, it becomes yellow. And there is a transition from red to yellow. As and when the fire becomes old and not contained, it keeps becoming more orange and ultimately becomes yellow if not controlled for more than x months. Choosing this geographical map and points as the marks and different colors for the points gives us an indication of how many forest fires are actually being detected and how many are being contained and where exactly some plan of action is required. As these points keep on increasing, these cumulative points look more like an area, giving a good idea of the expansion of fire, direction of the expansion and the areas affected.

Marks - Points and Areas; Channels - Colour, Position

Interactivity: Observations can be made over weeks as the user can choose a week to visualize all the fires detected till that week.

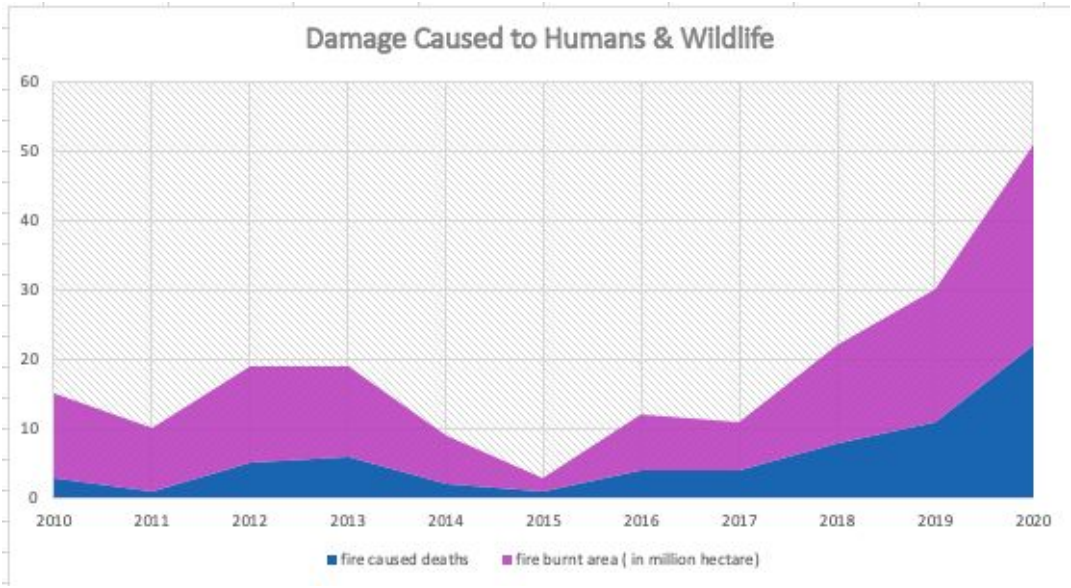
Comparison:

- Heat Point map offers higher levels of recognition compared to heat map in terms of accurate readings
- It is better than a point map, as the circles can overlap each other which is more layered.

2. **Area Chart :**

The main questions that we are trying to address via this visualization are

- The concerns of the local community as well as environmentalists interested in knowing about the extent of damage caused to human and animal lives because of the fire.
- Damage to infrastructure and forest/land cover over the years.



Plot title: Fire caused human deaths and burnt area for past 10 years

Description: The visualization will depict two area charts in a figure showing the yearly trends of human and animal deaths caused due to bushfires. For showing comparison of population i.e, humans v/s animals, an area chart is a good choice, since the filled area can indicate a countable amount instead of just rate of change. Also, since we are comparing among just 2 variables, the design choice is justified, otherwise this kind of visualization wouldn't have been suitable if we had many variables to plot.

Marks - Lines and Areas; Channels - Dimensions, Colour

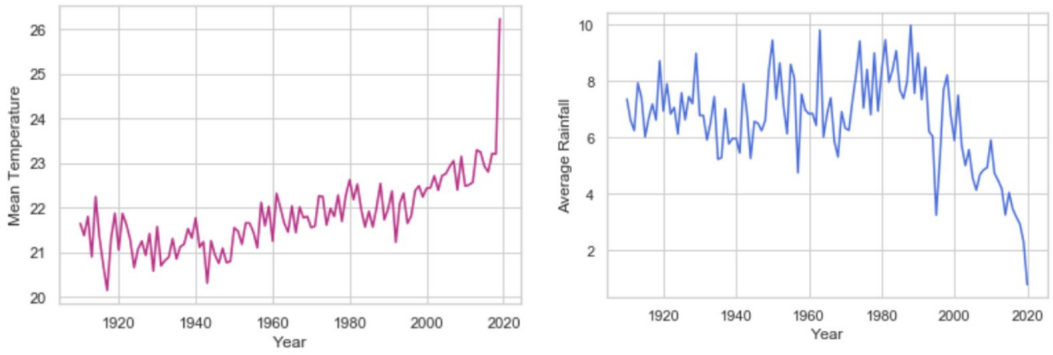
Interactivity:

- Highlighting of the selected region on hovering over the above graph.
- it will show the y-axis values for both the variables making it easy to make year on year comparisons.

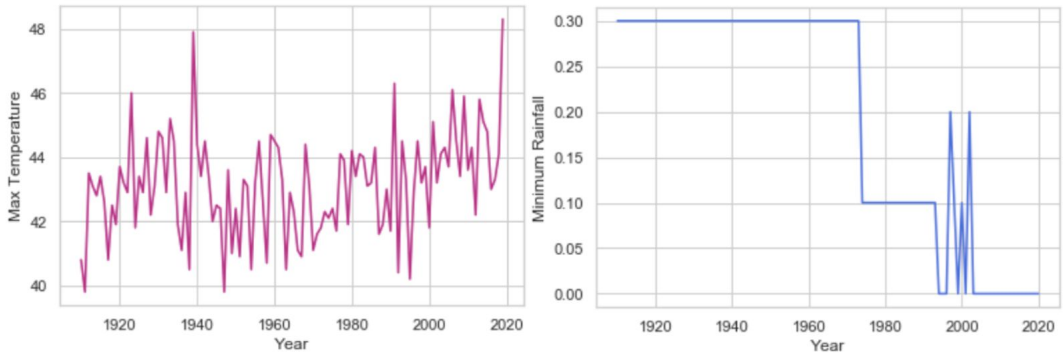
Comparison:

- Since, we are showing YoY trend comparison here, hence it makes sense to use an area chart that effectively depicts time series variation. Other kinds of visualizations like scatter plot/pie chart etc would not have highlighted the comparison better.

3. **Line Chart:** - The main questions that we are trying to address via this visualization are
   - Trends in Air quality Index
   - Changes in Rainfall and Temperature over the years
   - Understanding the forest fires with respect to global climate change



Plot title: Mean Temperature & Mean Rainfall over a particular year for 100 years.



Plot title: Max Temperature over a particular year for 100 years

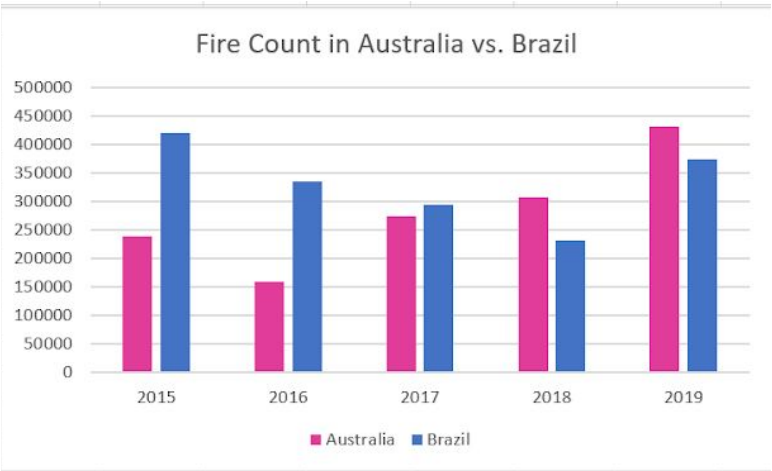
Description:
The visualization will depict line charts to show trends of climate related variables over the years. We show rainfall and temperature side-by-side to facilitate comparison of trends between the two.
Marks - Lines; Channels - Colour
Interactivity*:* When we hover over the two charts, it shows the value that corresponds to the y-axis value. That makes it easy to compare across the two charts.
Comparison:

- Area chart: We need to just show trends over years, hence, showing area would be an unnecessary addition.
- Pie chart: A line chart better facilitates comparison over different years, as compared to a pie chart.
- Bar Chart: There are a lot of data points on the x axis. Too many bars can seem like a clutter. A line does the perfect job of capturing the trend.

4. **Bar Chart**

The main questions that we are trying to address via this visualization are:

- Comparison of number of fires between Australia and Brazil over the years.
- Visualize the increase/reduction in number of fires to pinpoint extremes.



Plot title: Annual count of Fire in Australia vs. Brazil for past 5 years

Description: The visualization will present the fire detection counts of both Australia and Brazil side-by-side so that comparing the number of fires for each country respectively can be easily visualized.
Marks- line; Channels - color, height, x position
Interactivity: Upon viewing the visualization, the bars rise in chronological order by country, allowing the user to see the changes year by year

Comparison:

- Line chart: We would use a bar chart over a line chart to show not only the trends, but also emphasize the magnitude of the number of fires, marked by the height of the bars
- Pie chart: We want to analyse the trend over a linear progression of time, meaning that marking time is important as well.

REFERENCES

1. https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms
2. https://github.com/rfordatascience/tidytuesday/blob/master/data/2020/2020-01-07/readme.md
3. https://www.bbc.com/news/world-australia-50951043
4. https://www.nytimes.com/interactive/2020/01/02/climate/australia-fires-map.html
5. https://firms.modaps.eosdis.nasa.gov/active_fire/#firms-txt
6. https://www.preventionweb.net/news/view/70445/
7. https://aqicn.org/data-platform/register/
8. https://cse442-17s.github.io/Citiviz/final/#ninthPage
9. https://en.wikipedia.org/wiki/2019%E2%80%9320_Australian_bushfire_season
10. https://github.com/rfordatascience/tidytuesday/blob/master/data/2020/2020-01-07/readme.md