

**Parikh Malav Arunkumar**

**30708974**

**Lab 04 Monday 8 AM**

**Fahimeh Sadat Saleh and Pratik Bhumkar**

**Introduction**

Wildfires have been a dominant force of causing destructions across the world since long and wildfires in United States of America (USA) are a common phenomenon. Though wildfires are a natural disaster, there are multiple reasons due to which wildfires occur and global warming is one of them. Global warming has been one big reason due to which wildfires are occurring frequently as it has resulted in extreme weather conditions which causes wildfires. A statistic about wildfire suggest that, California is one of the states which has the highest number of wildfires across USA and in recent years it has increase five folds and the intensity of these fires has been severe. Wildfires cause damage to vegetation as it starts in an area of combustible vegetation, but vegetation is not the only thing which is affected due to it. It causes damage on multiple fronts and thus, this report is aimed to draw some comparisons and show trends of wildfires and its impacts and its relation to various other things.

**The visualization depicts:**

The number of wildfires over the years in all the states of the United States. And how it is related to temperatures. Depiction of temperatures being proportional to number of wildfires. Subsequently, how the rise in wildfire numbers results in higher release of CO2 and which causes respiratory diseases in those states more and lead to death. Moreover, societal impacts like business closures, GDP and over all population are shown with the number of wildfires to see whether there is a direct relation between them.

The motivation behind exploring the wildfires is to see how wildfires have a domino effect and the need to act against it. The motivation is to see how in the recent years, the phenomenon of wildfires has impacted on multiple fronts. The intended audience for the visualization are us: The People. The visualizations tries to depict how everything is related to each other and how we as humans need to come together and act against the changing climates and global warming to reduce life threatening events like wildfires which cause damages far and wide.

**Design**

The five-design sheet methodology is used to depict the various visualizations alternatives that have been brainstormed before realizing the final design of the visualization. As the name suggests five sheets have been prepared to go through the various design and presentations of the data to reach to a design conclusion. The first sheet is the brainstorming sheet where multiple ideas whatever comes to mind related to the visualization has been depicted. Based on those ideas filtering is done to remove the duplicate ideas and keep the ideas which are best suited to visualize and implement. Categorize by ordering the ideas and then organizing the ideas by combine and refine. (View the five-design sheet in the Appendix section)

The sheets that follow the brainstorming sheet namely: Sheet 2, 3 and 4 are the one which depict the ideas by visualizing all the questions that one wants to answer in 3 different ways, each sheet representing one way of answering the question. Based on these three sheets, the final design sheet is realized.

The **sheet 2**’s (initial design 1) first representation is a bar plot which can be changed based on different years just to describe the number of wildfires over years in every state of USA. This design is followed by a scatter plot depicting the relation between temperatures and number of wildfires answering whether they are directly proportional to each other. This is followed by a grouped bar chart showing CO2 emissions levels of different states alongside the number of deaths caused by respiratory diseases. The fourth visualization answers the business impacts due to wildfires using a two-way mirrored bar plot depicting the number of wildfires on one side and the number of business closures on the other. The last and the final visualization shows the relation of wildfires, GDP and population through a multi-line chart where each line is of a different color. Tooltips have been added to show the information while hovering over the graphs.

**Pros:** Easy to implement, simple and clean designs.

**Cons:** Repetitive graphs, look very much easy.

The **sheet 3**’s (initial design 2) first representation is a multi-line chart showing the number of wildfires over the years in various states which can be selected by dropdown. The second graph depicts the radar chart of temperature and wildfires overlapping each other which answers the relation between temperature and wildfires. The CO2 emission levels are visualized using a circular packing graph where the size of each circle is relative to the emission levels. The number of deaths are represented by a grouped bar chart of wildfires and deaths by respiratory disease of every state. And the last visualization of a line chart is used to display the relation of GDP and population with wildfires.

**Pros:** Radar chart is the highlight of this idea, circular packing too is innovative and attractive

**Cons:** Repetitive line charts.

The **sheet 4**’s (initial design 3) first representation is a choropleth map indicating the number of wildfires of different state by circle markers and each circle having radius based on the number of wildfires. Another choropleth map is used to answer the CO2 emissions and death data. The business closures and wildfire relation is depicted using a circular dual bar plot.

**Pros:** New idea of the circular bar plot. Choropleth map looks good.

**Cons:** Choropleth map cannot be depicted due to lack of geo-json data.

The final sheet i.e. **the realization sheet** uses the bar plot of first sheet depicting the wildfire numbers over the years for all the states which change with the change in year. The second relation of wildfire and temperature is shown using the radar chart from sheet 3. This is followed up by the circular packing chart of the CO2 emissions from the sheet 3 again. Alongside that a bar plot representing the number of deaths against wildfires is used. The business impacts question is answered by the radial dual bar chart of sheet 4 which depict the number of business closures for each state against corresponding number of wildfires. The last and the final visualization depicts the line graph of GDP, population and wildfires answering the impact of wildfires on both of them over the years.

**Pros:** Varied type of graphs are used without any repetition, very much informative, supports the whole narrative with interactive features like tooltips and click events, filtering based on dropdowns and legends is used to make it interactive.

**Implementation**

The represented visualization is made using d3 library of java-script alongside html for content and CSS for styling. D3 is used to provide visually informative and attractive presentations of data. The story can be gone through by the provided nav bar which shows different angles of the story and subsequently, the impacts a wildfire can cause. Multiple datasets are used to complete this visualization. The data sets are a combination of multiple parameters which have been wrangled using R and MS Excel. Each dataset comprises of data of various different years and each having different notations and ways of representing data. Combination of D3’s different versions have been used to bring out the real visualization. Nesting of data for representations using d3.nest() and load.csv() to load data, color palettes of various categories, functions to bring out interaction are used. Multiple tooltips and interactions are added to go through the visualization easily and informatively. Each page can be navigated through by clicking on the navbar options and each page represents a different angle connected to each other so, it is better to go page by page. A conclusion page is added to sum up the whole story represented in numbers and bars and lines and charts in the previous pages to justify the domino effect of wildfires.

**User Guide**

The visualization can be viewed by opening any of the html files but, it is advised to start with “WilfireInNumbers.html” as it is the introductory page and there after each visualization can be visiting by navigating using the navbar options. Another recommendation would be to go through it page by page as the order is depicted to see the story behind.

Each page has a guide section at the bottom which represents how the visualization can be interacted with making it easy for everybody to go through it. A good way to start is by hovering over the graphs to see the resulting interaction.

**Conclusion**

The visualization performed on the dataset clearly depicts domino effect of global warming on multiple fronts. The findings justify that there is a direct relation between temperature and wilfires. As the temperatures rise, the number of wildfires too rise due to the dried up land and burnt vegetation. Further, the findings took on from the previous findings to justify that wildfires release large amounts of CO2 into the atmosphere which gives rise to more and more respiratory diseases and subsequently resulting in more deaths due to the same. The trends showed that regions having higher number of wildfires faced more deaths due to respiratory diseases which clearly justify our question of a direct relation between the two. Subsequent visualizations depict the business impact and societal impact of the wildfires. The findings tried to see a pattern where wildfires caused more business closures but a clear trend could not be identified between the two as the reason of business closures will not only closures will not only depend on the wildfires but due to more than one clear factor and also our findings are from the years after the great recession of 2008 which turned the economies upside down which do not do justice as keeping wildfires as the sole reason for the impacts.

**Reflection**

The learnings from this project suggest that every data has some story behind it and upon visualizaing the data some clear patterns can be found out which help understand the story telling behind the data clearly. The project gave a better understanding of how closely the destruction of our nature by us humans affects none other than ourselves bringing unavoidable circumstances for the mankind as whole. It is utterly important to preserve our nature and rightly justify that protect nature and nature will protect you. Alongside that using languages like javascript and the use od D3 to represent the visualizations makes the process a learning curve in the world of visualization.

The certain limitations of theses visualizations are due to lack of all datasets giving data about the same year range which makes our findings year range small and hence cannot clearly identify and justify the trends like in population and gdp which more data availability could have provided.

**Data Sources**

1. **Population dataset** contains information about population by years and zip code

<https://data.world/lukewhyte/us-population-by-zip-code-2010-2016>

1. **Zip code dataset** contains data about every zip code. Its location, state and every related data. <https://simplemaps.com/data/us-zips>
2. **Number of Wildfires** contains month, year and state wise data of number of wildfires and acres burned [https://www.ncdc.noaa.gov/societal-impacts/wildfires/month/12?params[]=acres&params[]=fires](https://www.ncdc.noaa.gov/societal-impacts/wildfires/month/12?params%5b%5d=acres&params%5b%5d=fires)
3. **Wildfires by state** contains data about number of wildfires and acres burned in every state every year <https://www.iii.org/table-archive/23284>
4. **GDP** contains data about GDP of US, GDP per capita, inflation and unemployment rate

<https://www.imf.org/external/pubs/ft/weo/2019/02/weodata/weorept.aspx?sy=1980&ey=2019&scsm=1&ssd=1&sort=country&ds=.&br=1&pr1.x=21&pr1.y=11&c=111&s=NGDP_RPCH%2CNGDP%2CNGDPPC%2CPCPIPCH%2CLUR%2CBCA_NGDPD&grp=0&a=#download>

1. **CO2 emissions by state** contains year and state wise data of CO2 emissions

<https://www.eia.gov/environment/emissions/state/>

1. **Respiratory Disease** contains data about the number of deaths caused by respiratory diseases across states in different years

<https://www.cdc.gov/nchs/pressroom/sosmap/lung_disease_mortality/lung_disease.htm>

1. **Temperature by State and Year** contains data about state wise and year wise temperatures

<https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data>

1. **Business Data (Table Name: State)** contains data about every state wise and year wise new business establishments, closures, job creations, job closings and all related data

<https://www.census.gov/programs-surveys/bds/data/data-tables/legacy-establishment-characteristics-tables-1977-2014.html>

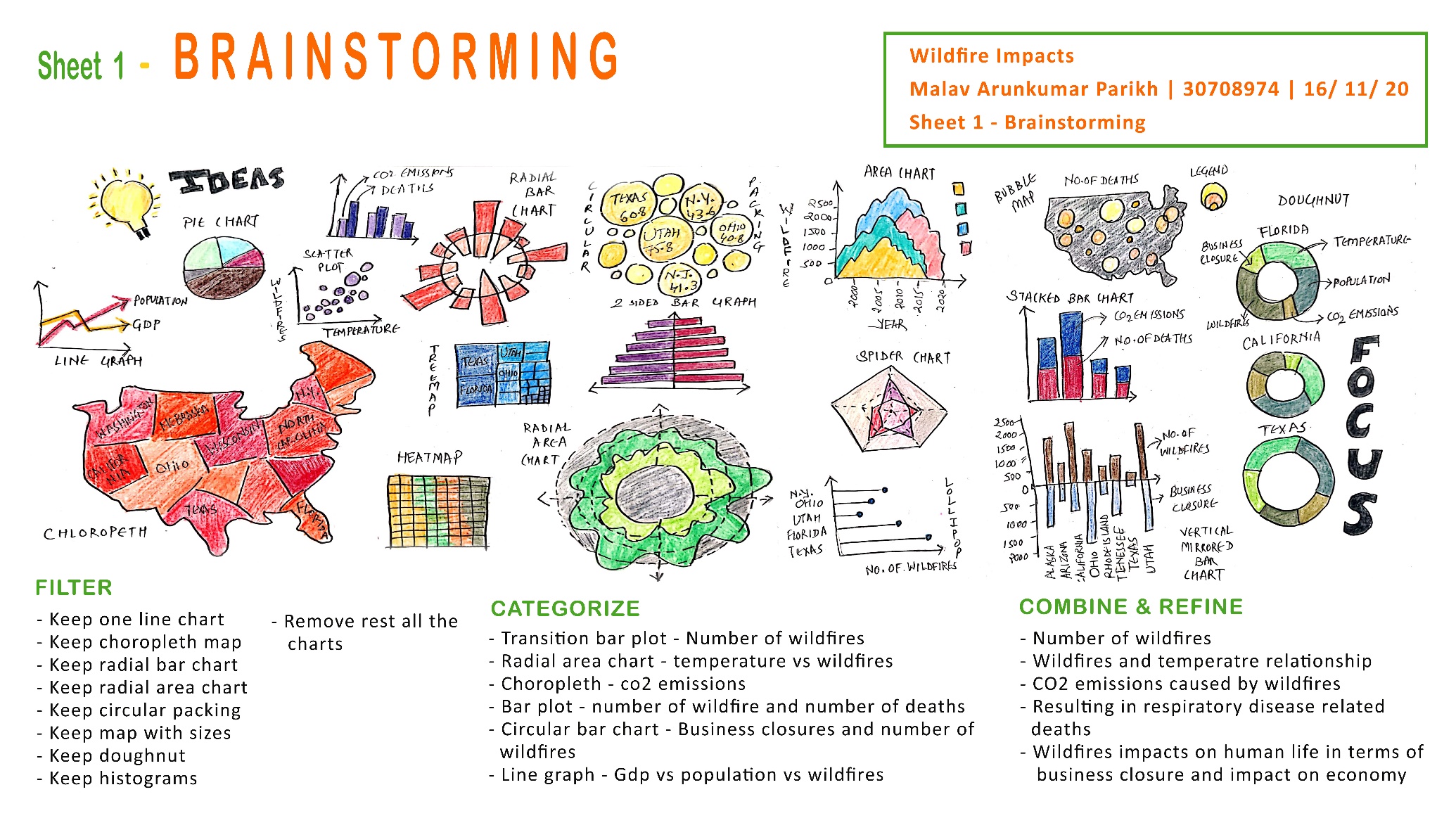
**Bibliography**

[1] D3: Data Driven Documents. “Home.” Why GitHub? <https://github.com/d3/d3/wiki/> (Accessed Nov. 16, 2020).

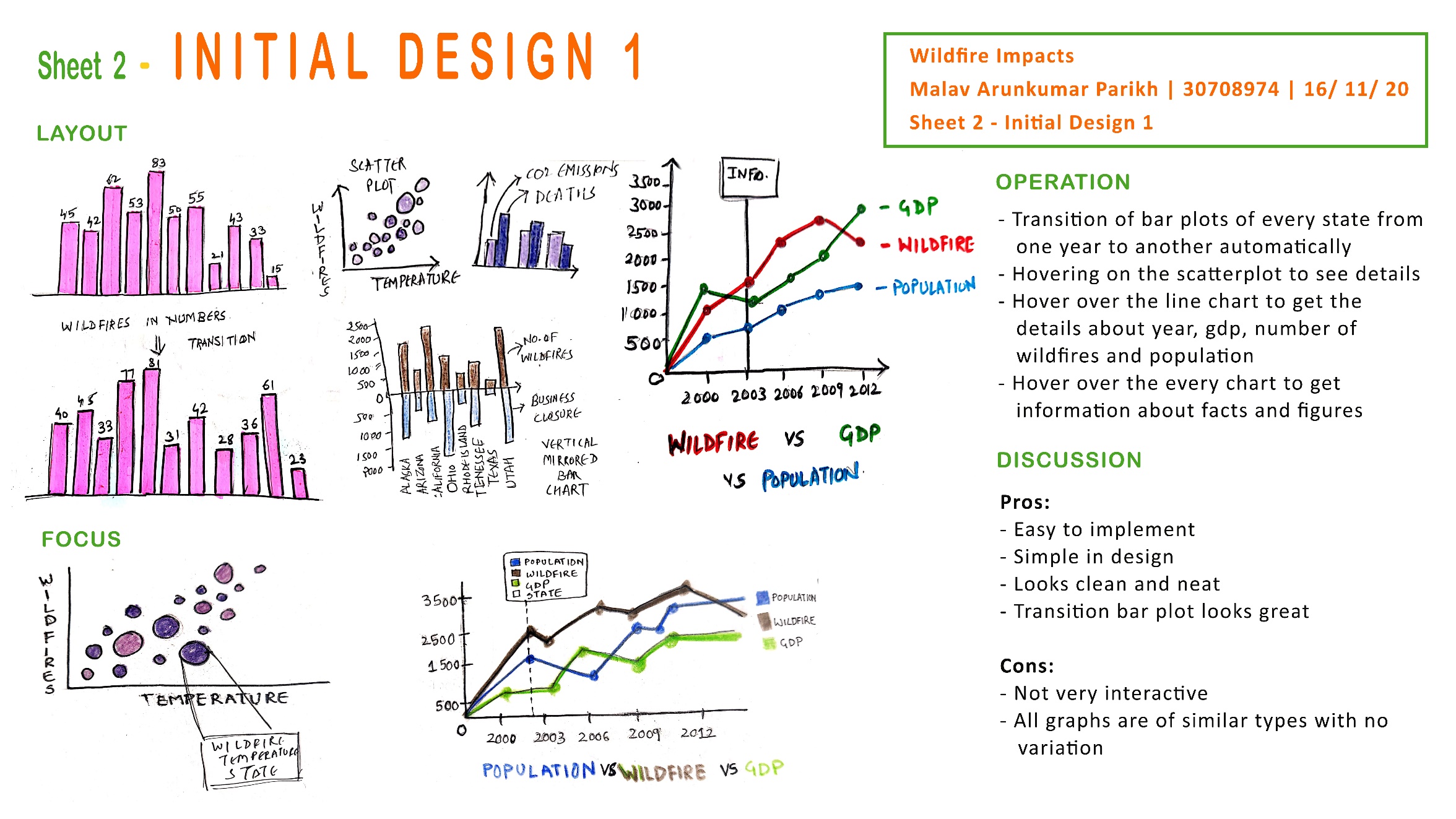
[2] DATA.GOV. “The home of U.S. Government’s open data.” DATA.GOV. <https://www.data.gov/> (Accessed Sep. 18, 2020).

**Appendix**

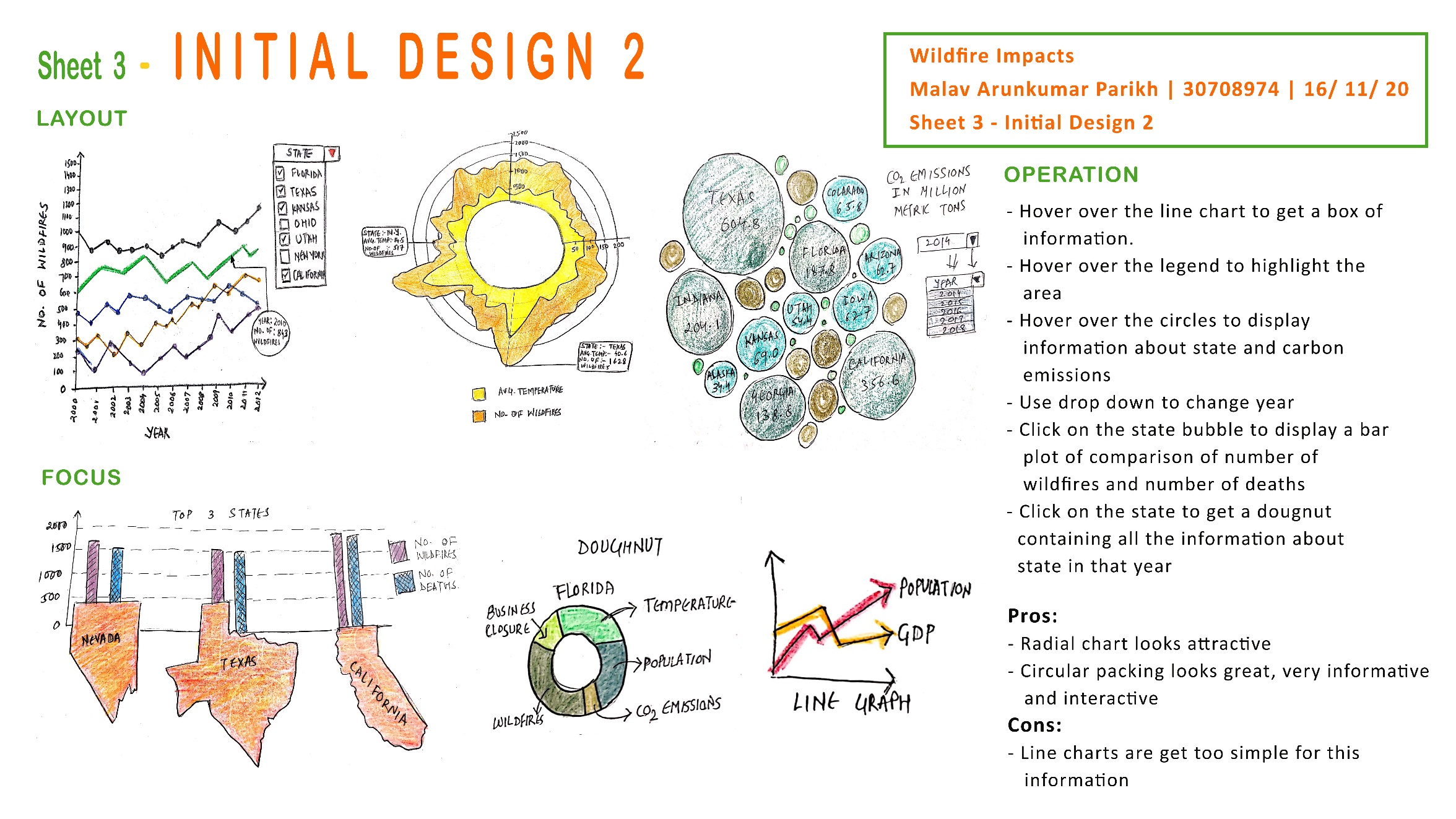
**Brainstorming**



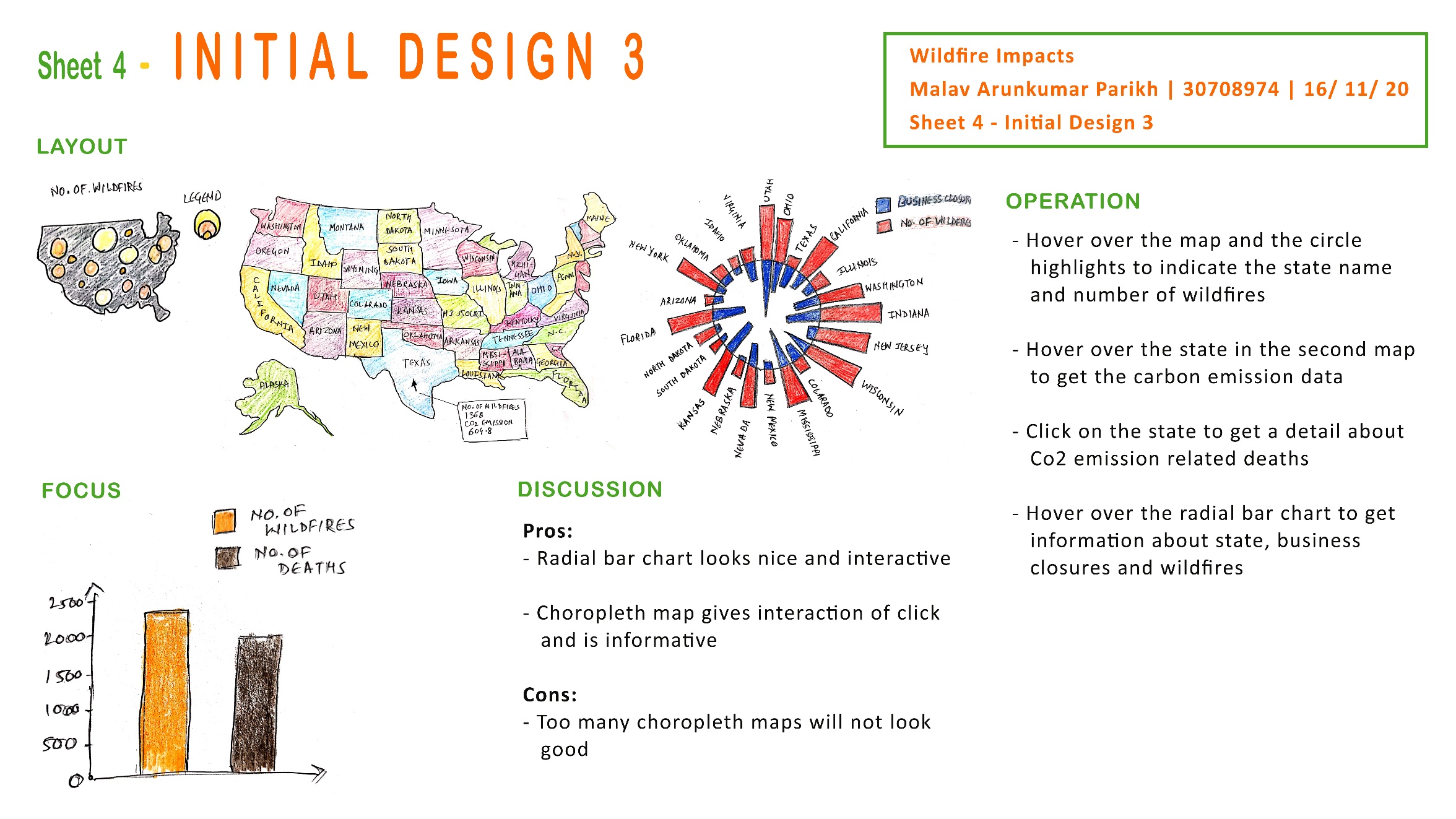
**Initial design 1**



**Initial design 2**



**Initial design 3**



**Realization Design**

