CS5764 Design Sprint Process Book

Leader: Asmita Hanchate (asmi0604, asmi0604@vt.edu)
Group Member 1: Hemayet Ahmed Chowdhury(hemayetahmedc,hemayetahmedc@vt.edu)
Group Member 2: Darshan Vekaria (dvekaria,dvekaria@vt.edu)
Group Member 3: Amith Reddy Nalla Venkat (amithreddy,amithreddy@vt.edu)

Team Agreement

- How will you communicate? [technology, rules, …]'
  -Official communication regarding updates and design decisions will be communicated via WhatsApp and meetings will be conducted on the Zoom platform.
- How and when will you meet? [regular meeting time]
  -Meetings will be held every Thursday and discussions will be done post lectures on Wednesday. Meetings can also be held as and when required before the submissions for completion of pending work.
- How will you collaborate on implementation? [GitHub, process, …]
  -We will collaborate through emails or if required, create a private repository on GitHub with all the team members as collaborators. Code can be shared through any medium as long as it is being smoothly integrated into one final project.
- How will you deal with non-performing members?
  -Non-performing members will be sent 2 warnings and will be given an appropriate amount of time to get done with their tasks. Even after this, if the team member does not perform, this matter will then be discussed with the TA or the professor.

Signatures: Asmita Hanchate
Date: 03/18/2022
Signatures: Hemayet Ahmed Chowdhury
Date: 03/18/2022
Signatures: Darshan Vekaria
Date: 03/18/2022
Signatures: Amith Reddy Nalla Venkat
Date: 03/18/2022

Dataset chosen - US Accidents (2016 -2021)

Link to the dataset -

https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents?select=US_Accidents_June20.csv&ref=hackernoon.com

Map

1. Who is your audience? Come up with **at least three** options and pick one target audience.
   Our audiences are -
   - US Traffic Control and maintenance
   - National/Regional/Local news outlets
   - General Public, especially drivers.

2. Describe your target audience in more detail. What do they know? What are their interests? What visualization literacy do they have? At what level of detail will you present information to them?

   -We believe that the US Trafic Control and Maintenance would benefit from our visualizations to recognize a pattern in accidents and take appropriate measures to curb them. Their interests would be to maintain the traffic even possibly in extreme conditions to avoid loss of life.

   -News outlets at all levels will benefit from these visualizations as they will be able to bring awareness in terms of road safety on a much broader scale.

   -Our main audience is the general public, especially drivers because it will help them make out a trend to take precautionary measures. E.g., if there is a trend that “x” state has cold winters and an increased number of accidents during the season, drivers from that region might be able to take better measures before heading out.

3. What questions about your data will be interesting for your audience? Come up with a list of interesting questions that your audience may have about your data. The more, the better, but your team should come up with **at least ten questions**.

1. Are there more accidents in warmer or colder areas?
2. Which 5 states have the highest number of accidents?
3. Does New York show up in the data? If yes, why is the count lower if this is the most populated city.
4. Among the top 100 cities in a number of accidents, which states have the highest number of these cities?
5. What time of the day are accidents most frequent?
6. Which days of the week have the most accidents?

7. Which months have the most accidents?
8. What is the trend of accidents year over year? Is it increasing or decreasing?
9. When are accidents per unit of traffic the highest?
10. What could be the reason for the highest accident-prone cities/areas having these many accidents?

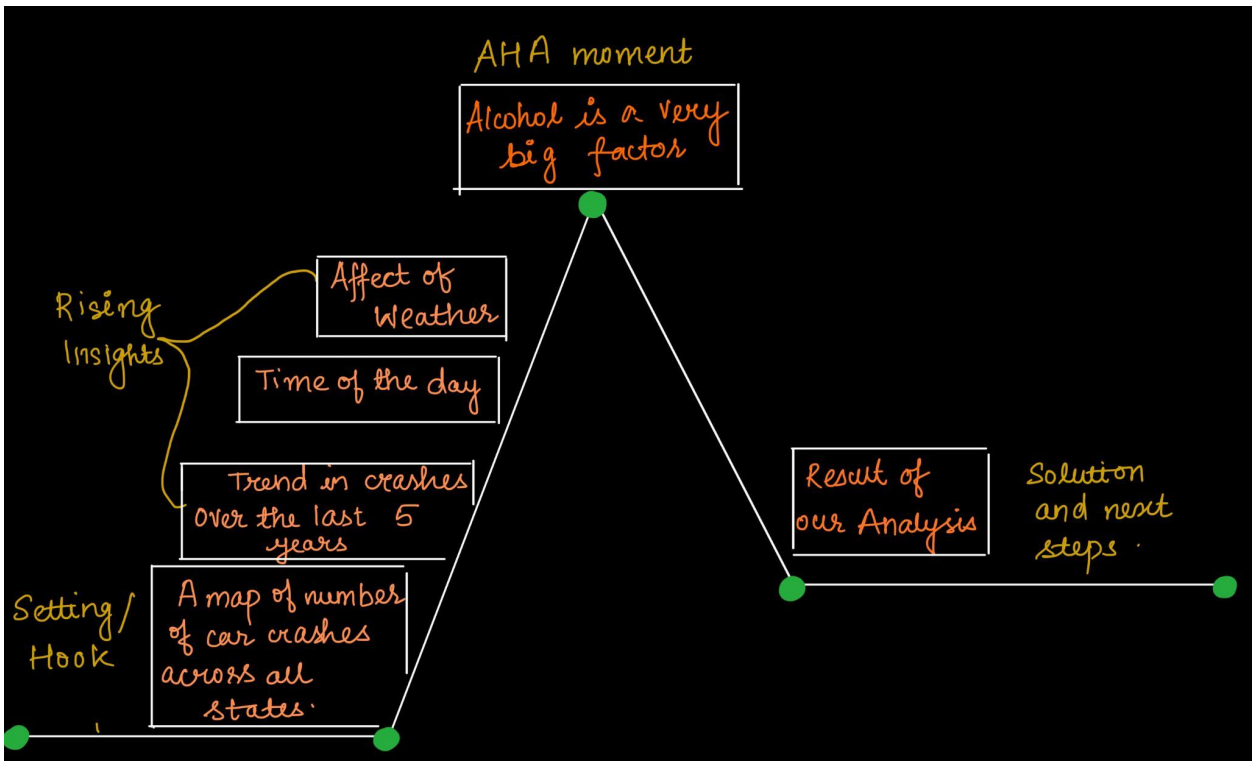
4. What data do you have? Give a brief description of each attribute and its data type (categorical, ordinal, or quantitative) in your process book. It's OK if you are not sure about the data type for some attributes - you can simply describe them (e.g., geographic location).

```
 #   Column                Non-Null Count    Dtype

---  ------                --------------    -----

 0   ID                    1516064 non-null  object Ordinal

 1   Severity              1516064 non-null  int64  Quantitative

 2   Start_Time            1516064 non-null  datetime64[ns]

 3   End_Time              1516064 non-null  object Quantitative

 4   Start_Lat             1516064 non-null  float64 Ordinal

 5   Start_Lng             1516064 non-null  float64 Ordinal

 6   End_Lat               1516064 non-null  float64 Ordinal

 7   End_Lng               1516064 non-null  float64 Ordinal

 8   Distance(mi)          1516064 non-null  float64 Quantitative

 9   Description           1516064 non-null  object None

 10  Number                469969 non-null   float64  Quantitative

 11  Street                1516064 non-null  object  None

 12  Side                  1516064 non-null  object  None

 13  City                  1515981 non-null  object  Categorical

 14  County                1516064 non-null  object  Categorical
```

15	State	1516064	non-null	object	Categorical
16	Zipcode	1515129	non-null	object	Categorical
17	Country	1516064	non-null	object	Categorical
18	Timezone	1513762	non-null	object	Categorical
19	Airport_Code	1511816	non-null	object	Categorical
20	Weather_Timestamp	1485800	non-null	object	Categorical
21	Temperature(F)	1473031	non-null	float64	Quantitative
22	Wind_Chill(F)	1066748	non-null	float64	Quantitative
23	Humidity(%)	1470555	non-null	float64	Quantitative
24	Pressure(in)	1479790	non-null	float64	Quantitative
25	Visibility(mi)	1471853	non-null	float64	Quantitative
26	Wind_Direction	1474206	non-null	object	Categorical
27	Wind_Speed(mph)	1387202	non-null	float64	Quantitative
28	Precipitation(in)	1005515	non-null	float64	Quantitative
29	Weather_Condition	1472057	non-null	object	Categorical
30	Amenity	1516064	non-null	bool	Categorical
31	Bump	1516064	non-null	bool	Categorical
32	Crossing	1516064	non-null	bool	Categorical
33	Give_Way	1516064	non-null	bool	Categorical
34	Junction	1516064	non-null	bool	Categorical
35	No_Exit	1516064	non-null	bool	Categorical
36	Railway	1516064	non-null	bool	Categorical
37	Roundabout	1516064	non-null	bool	Categorical

```
38  Station                1516064 non-null  bool    Categorical
39  Stop                   1516064 non-null  bool    Categorical
40  Traffic_Calming        1516064 non-null  bool    Categorical
41  Traffic_Signal         1516064 non-null  bool    Categorical
42  Turning_Loop           1516064 non-null  bool    Categorical
43  Sunrise_Sunset         1515981 non-null  object  None
44  Civil_Twilight         1515981 non-null  object  None
45  Nautical_Twilight      1515981 non-null  object  None
46  Astronomical_Twilight  1515981 non-null  object  None
```

Storyboard



Explanation : The primary goal of our project is to analyze the factors that cause the most accidents across the states in the USA to raise awareness amongst the general public, especially drivers, of any common patterns present in our data.

The Hook: Our first visualization is a heat map of the number of accidents across all 52 states in the USA. This is to grab the attention of the audience about the severity of the accidents.

Rising Insights: In the rising insights section, we investigate the factors that affect car crashes the most.

- Time of the day: During which period of the day do accidents happen the most.
- Effects of weather: Does weather affect the frequency of car crashes
- Crashes over gender and age: How do gender and age of the victim affect the crashes

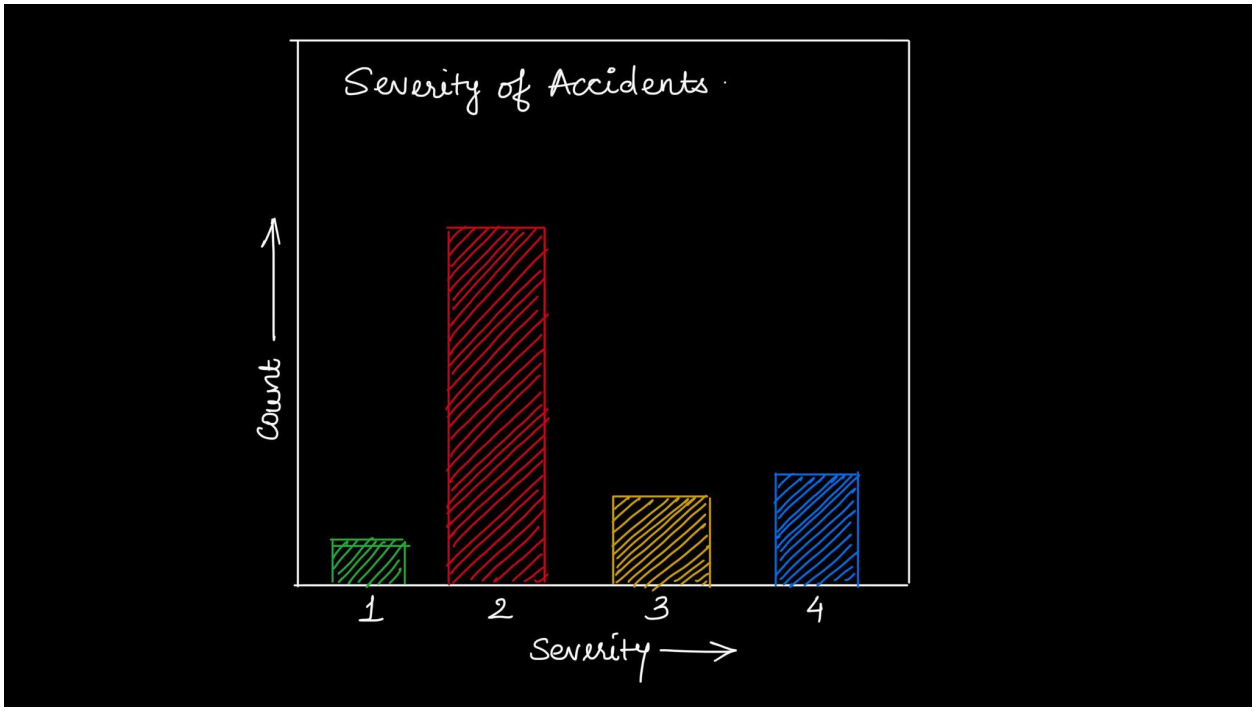
AHA! Moment (Main Message): Our data will show that perhaps the strongest indicator of accidents is the use of alcohol. This is surprising because about 1/3rd of the accidents in the USA takes place while driving under influence, indicating how we should be taking drinking and driving a lot more seriously than we are doing now.

Solution and Next Steps: A summary of all the factors that influence car crashes and perhaps the necessary steps we can take to raise further awareness.

Sketch

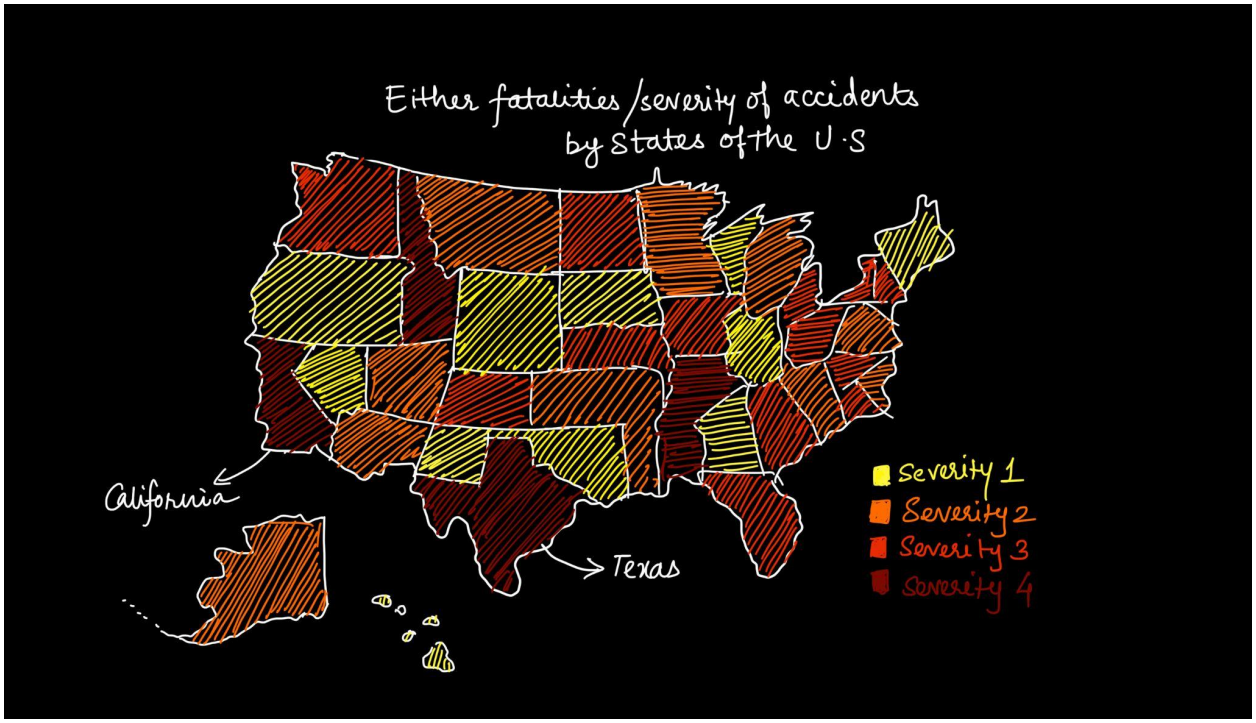
Asmita Hanchate

1.



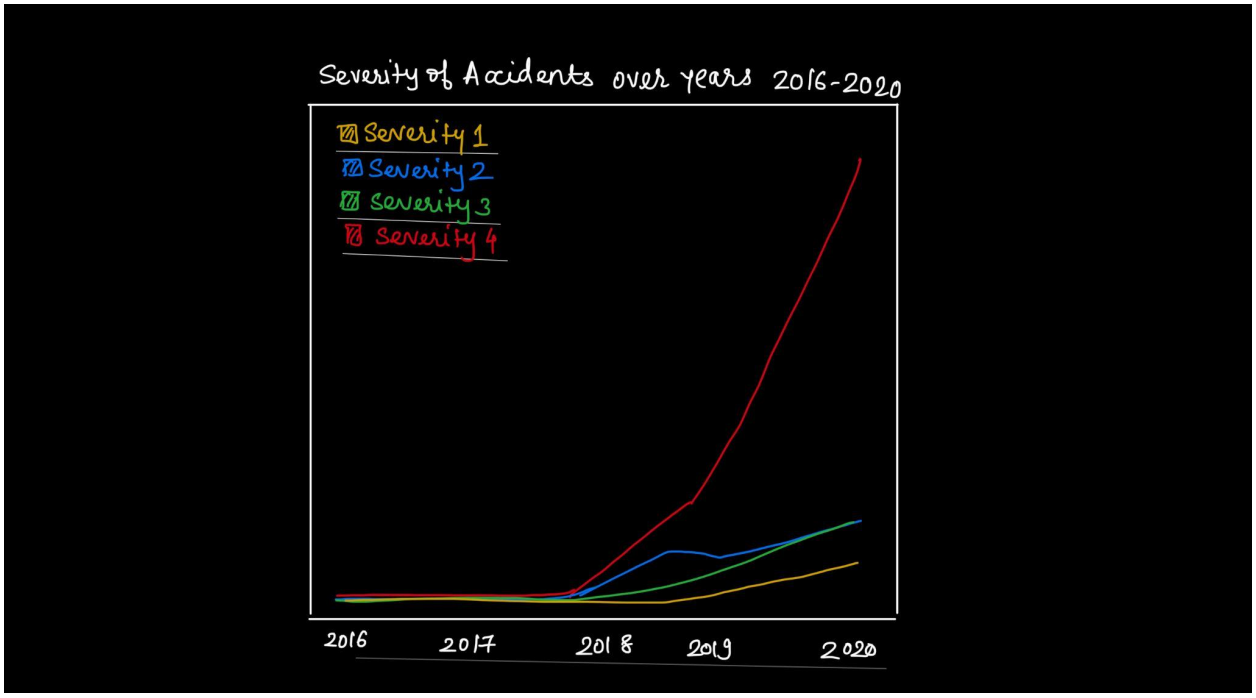
Message - Number of Accidents categorized by severity. [MID -1]

2.



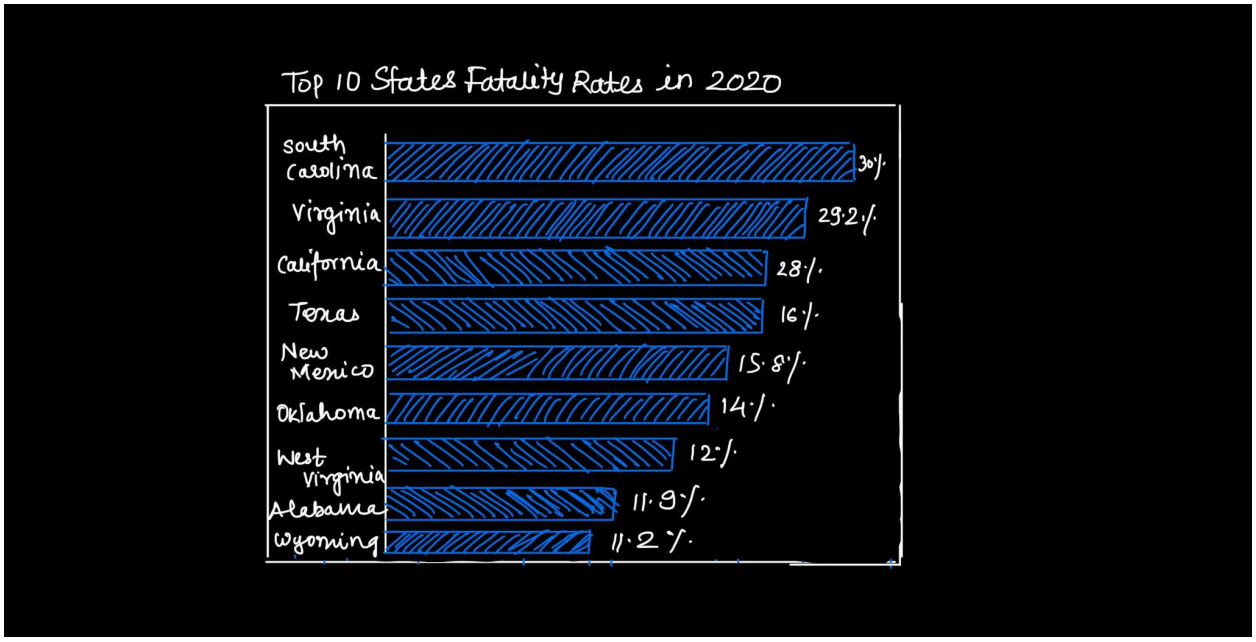
Message - Fatalities or Number of Accidents by severity categorized by states. [MID -2]

3.



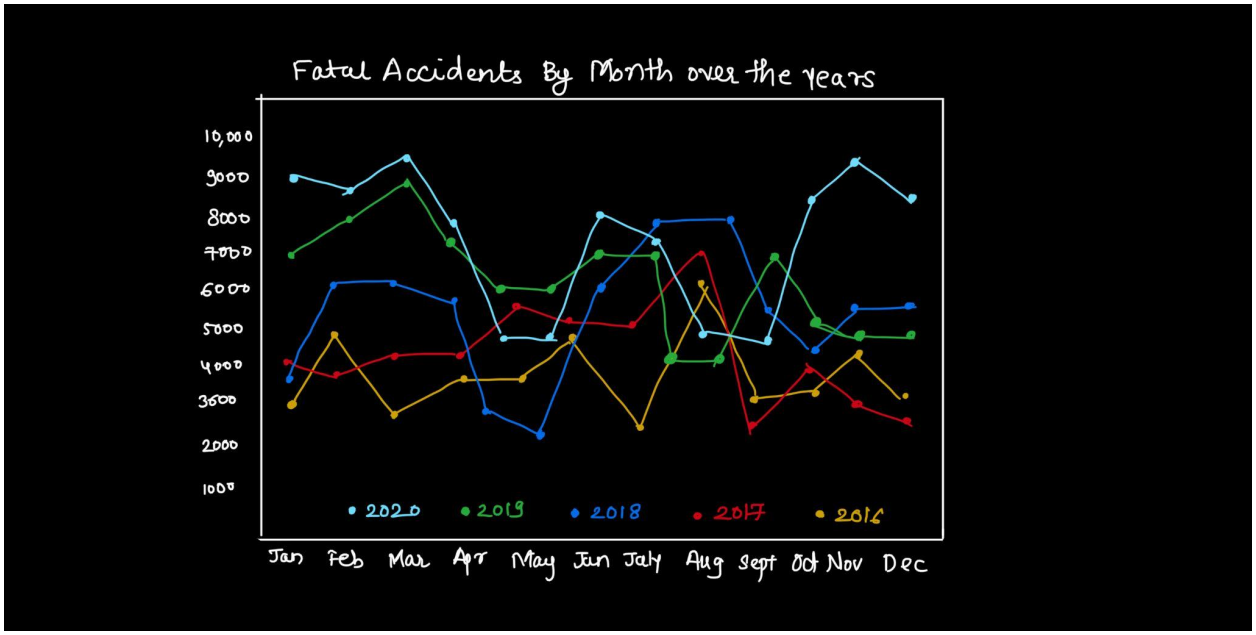
Message - Severity of accidents is categorized by years from 2016-to 2020. [MID -3]

4.



Message - Top 10 states categorized by a count of fatalities. [MID -4]

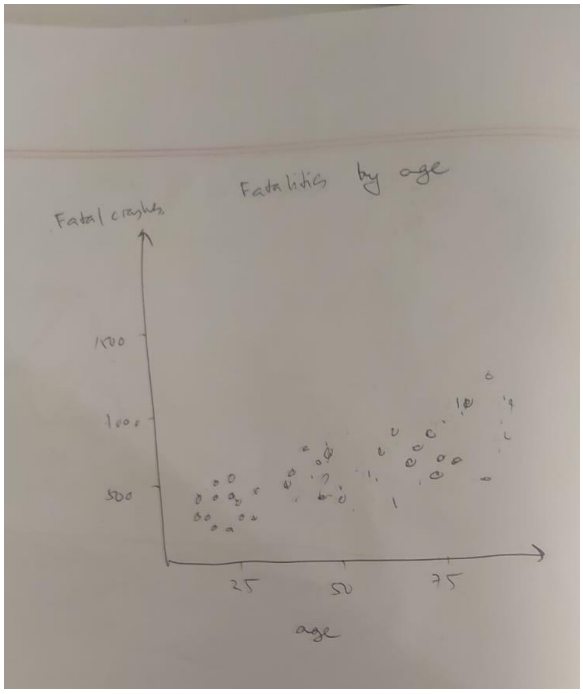
5.



Message - Count of accidents categorized by months.[MID -5]

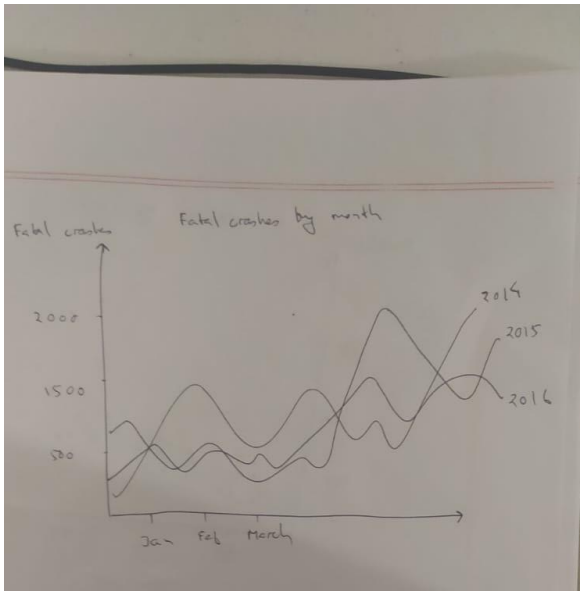
Hemayet Ahmed Chowdhury

6.



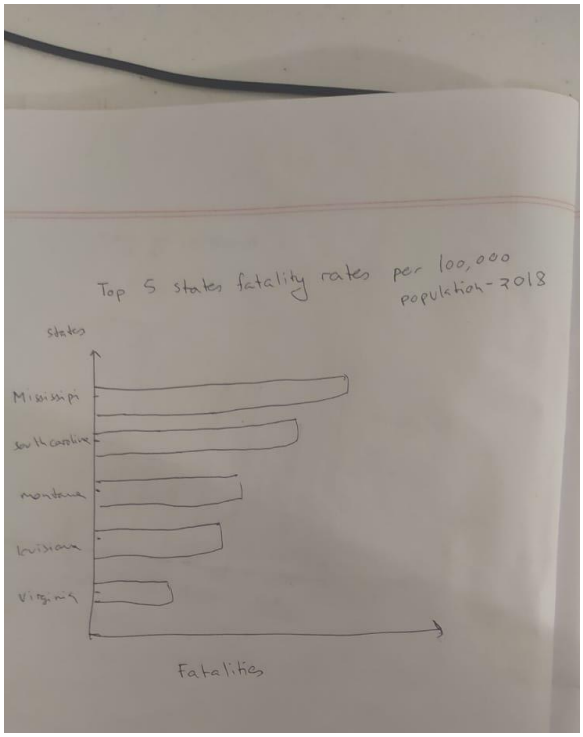
Message: Number of fatalities by age of the victims. [MID -6]

7.



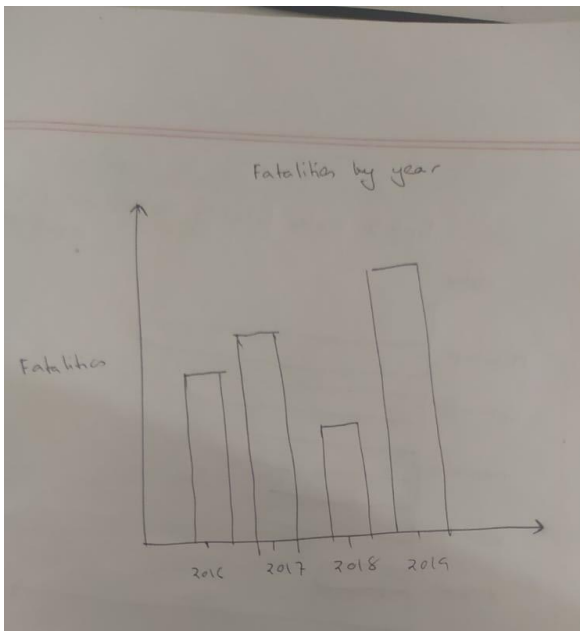
Number of fatalities by month over the course of 3 years. [MID - 5]

8.



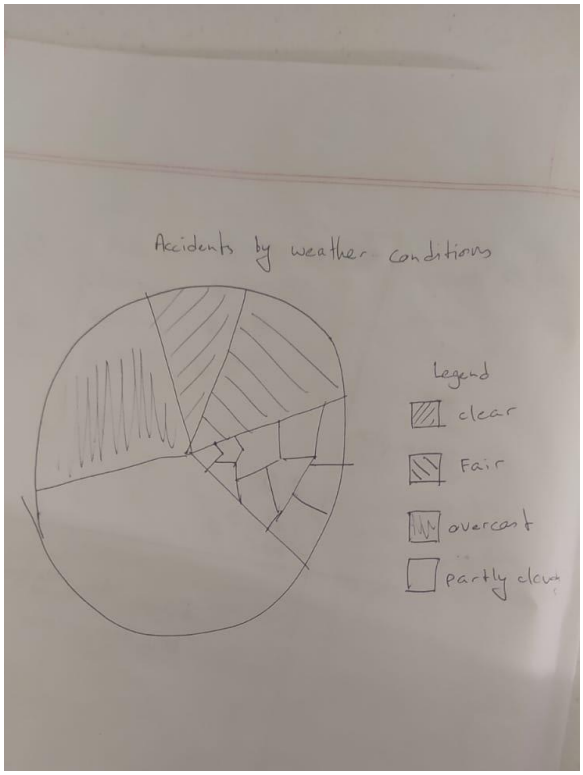
Top 5 states with the most fatalities. [MID -4]

9.



Total Fatalities by year. [MID -3]

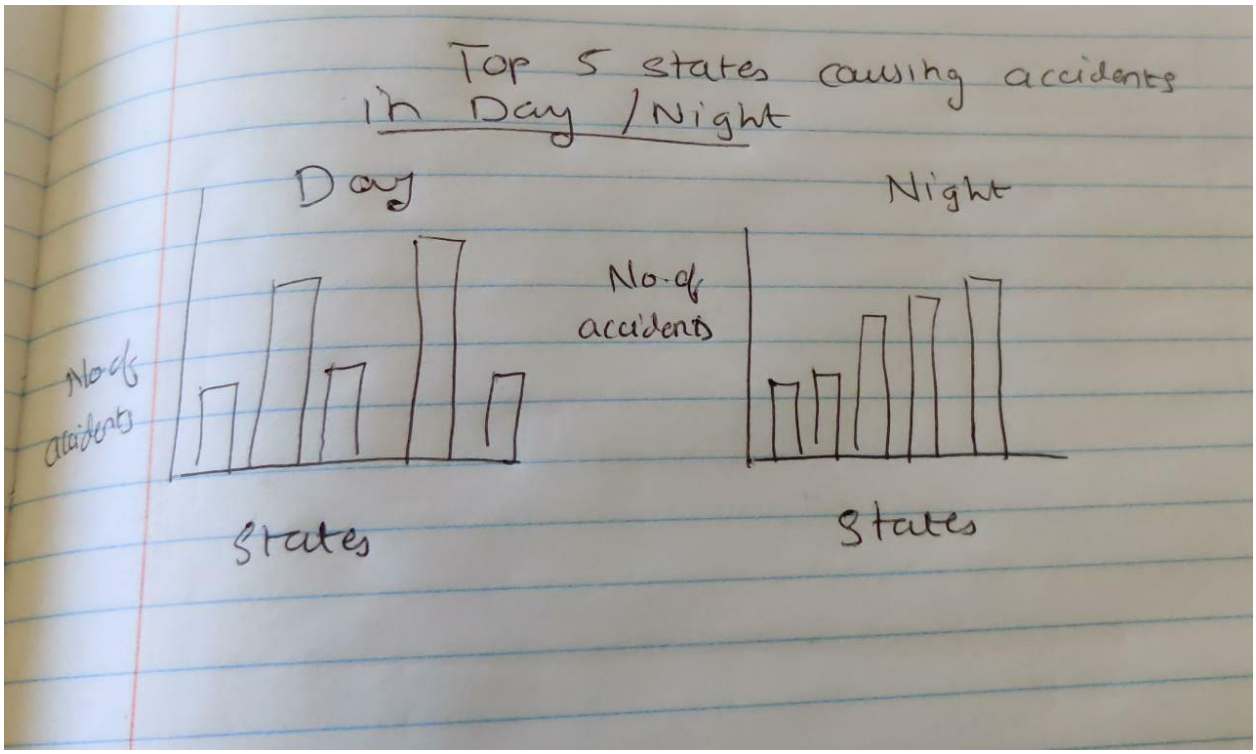
10.



The number of accidents in specific weather conditions. [MID -7]

Darshan Vekaria

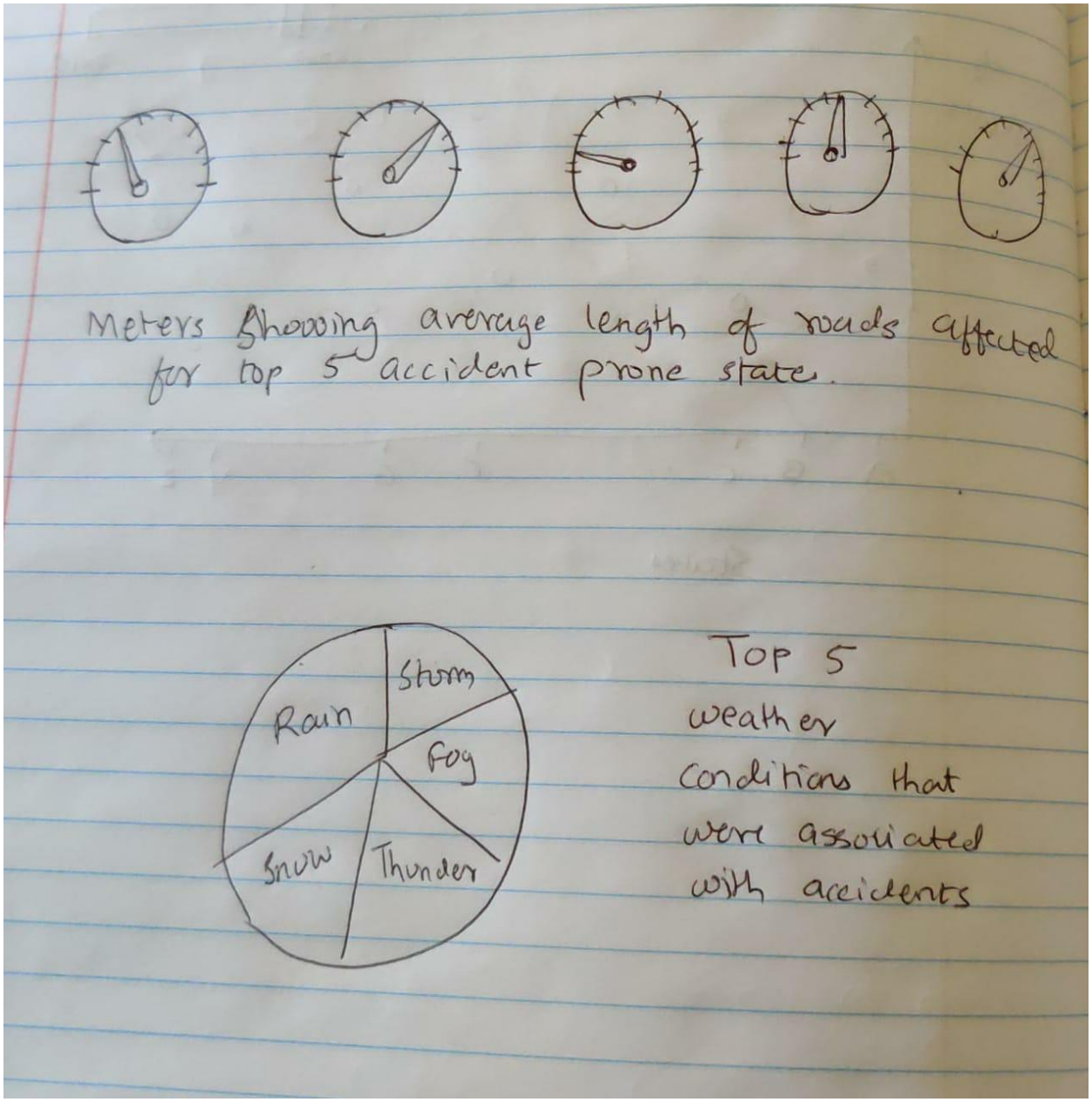
11. Day/Night distribution of accidents for the states with the most number of total accidents. [MID -8]



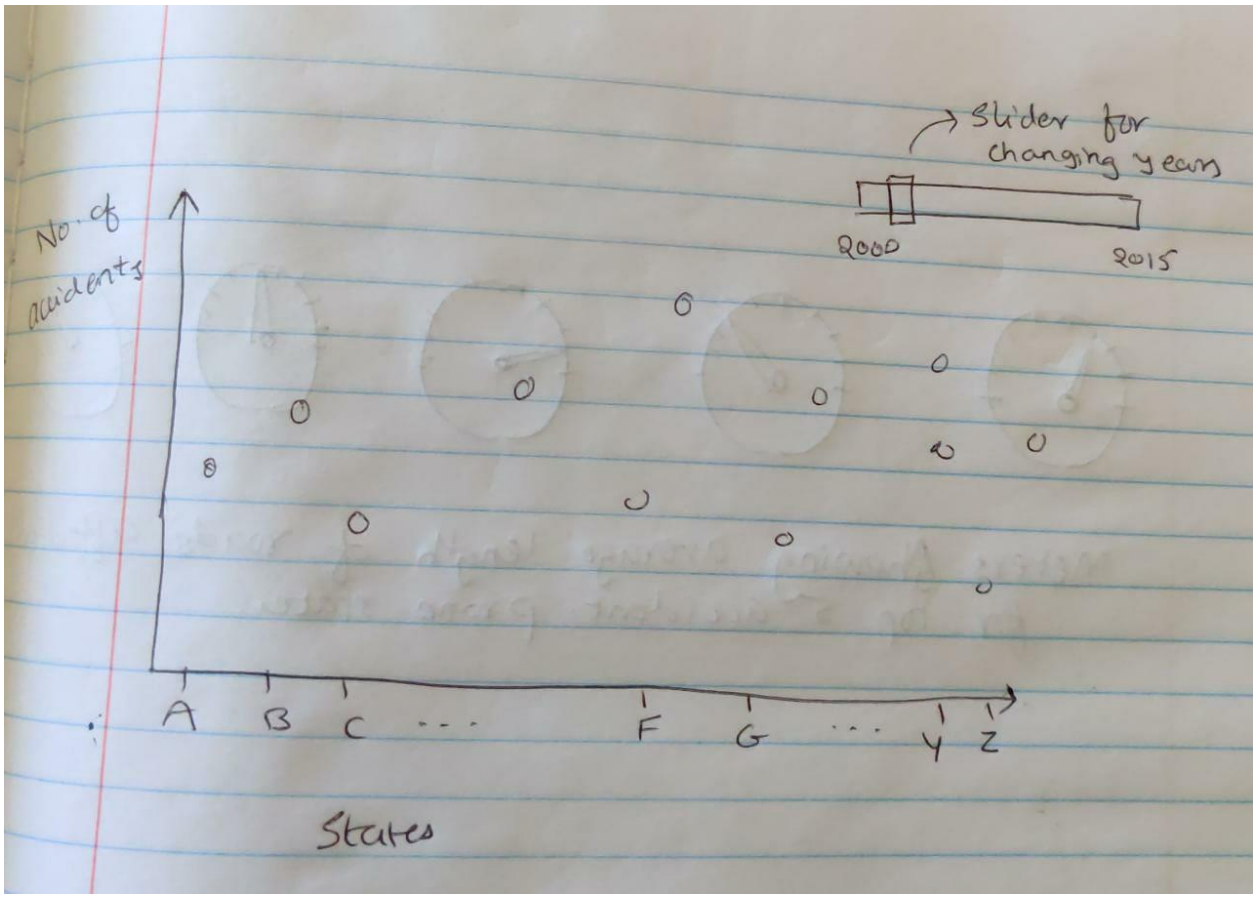
12 and 13

12: Average length of road affected for top 5 accident-prone states. [MID -9]

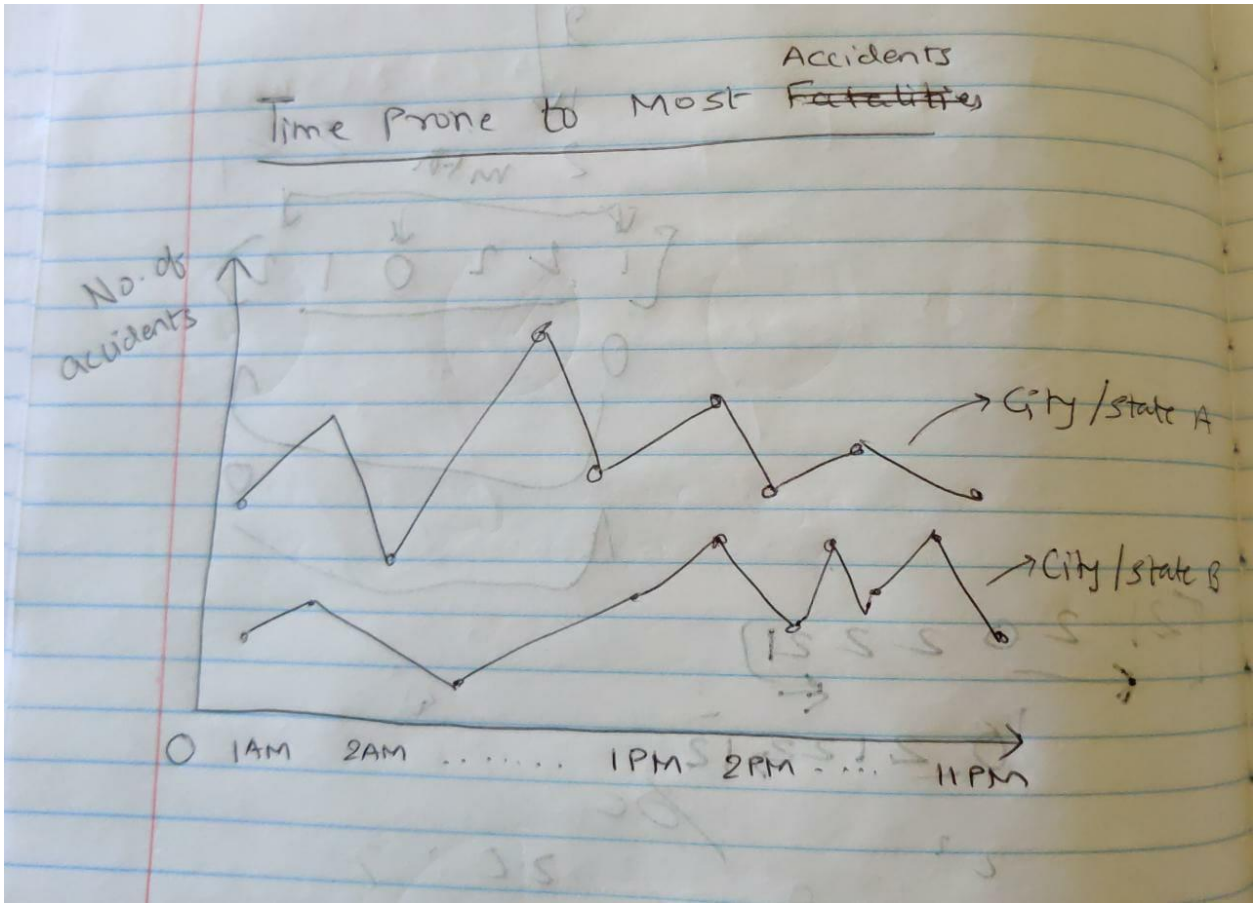
13: Distribution of weather conditions contributing to most accidents. [MID - 7]



14. Change in accidents over the years for various states. [MID - 2, MID - 3]

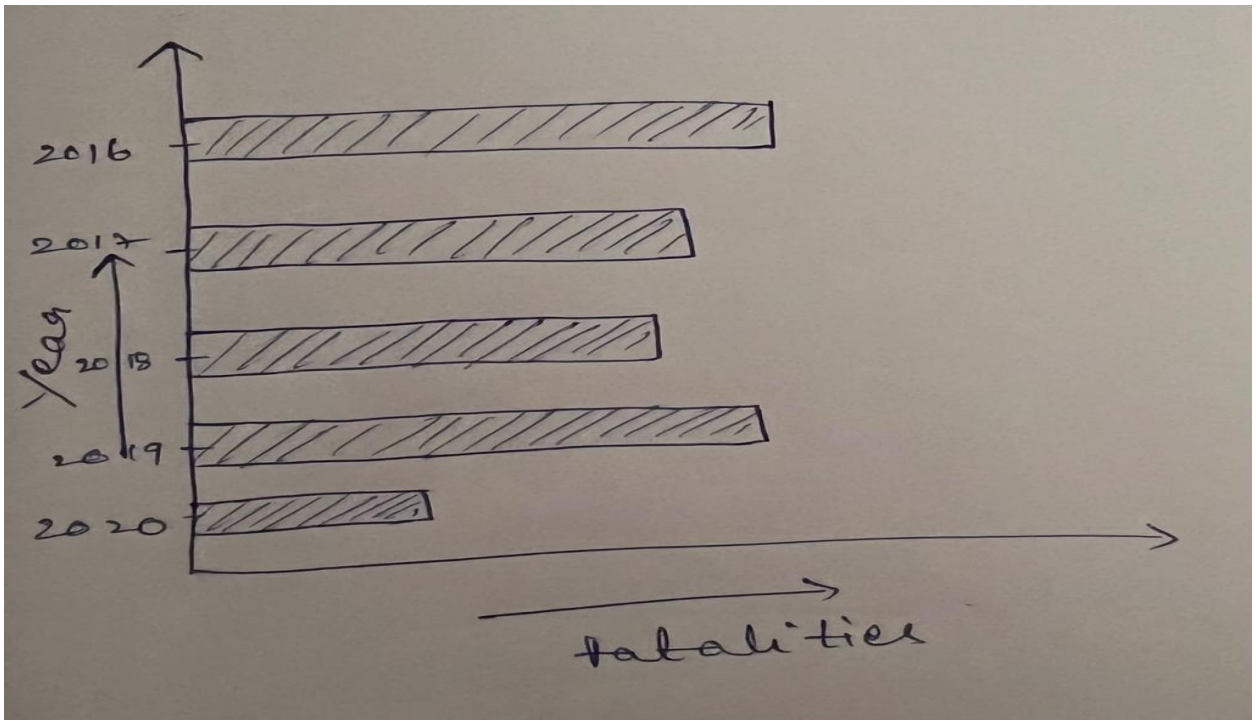


15. Distribution of accidents according to the hour of the day. [MID -5]

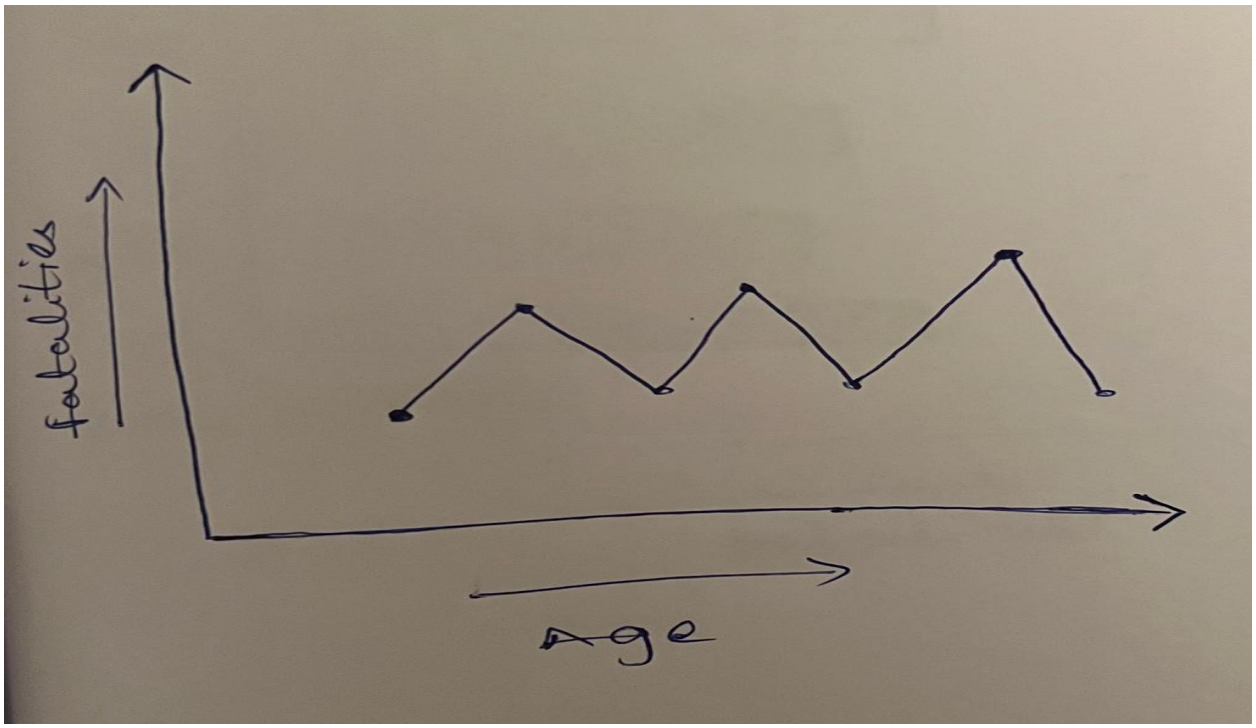


Amith Reddy Nalla Venkat

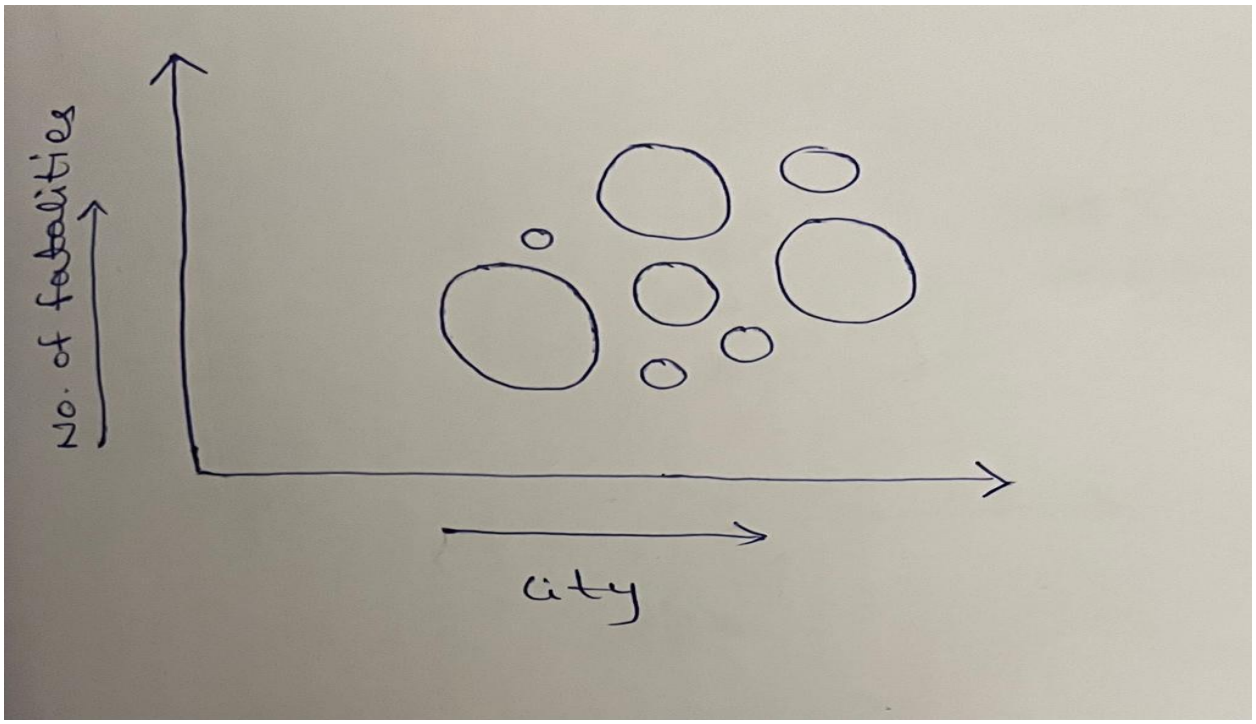
16. Increase or decrease in fatalities over the years [MID - 3]



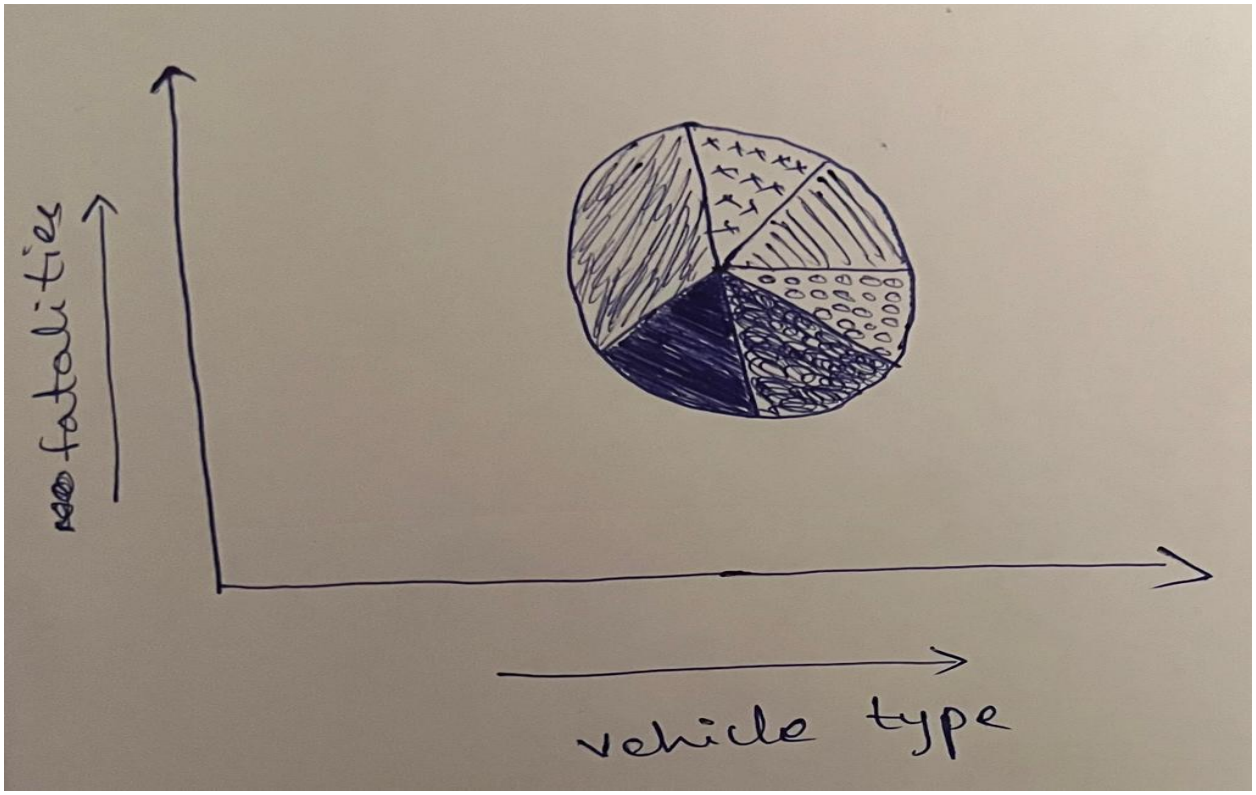
17. The number of fatalities by age [MID - 6]



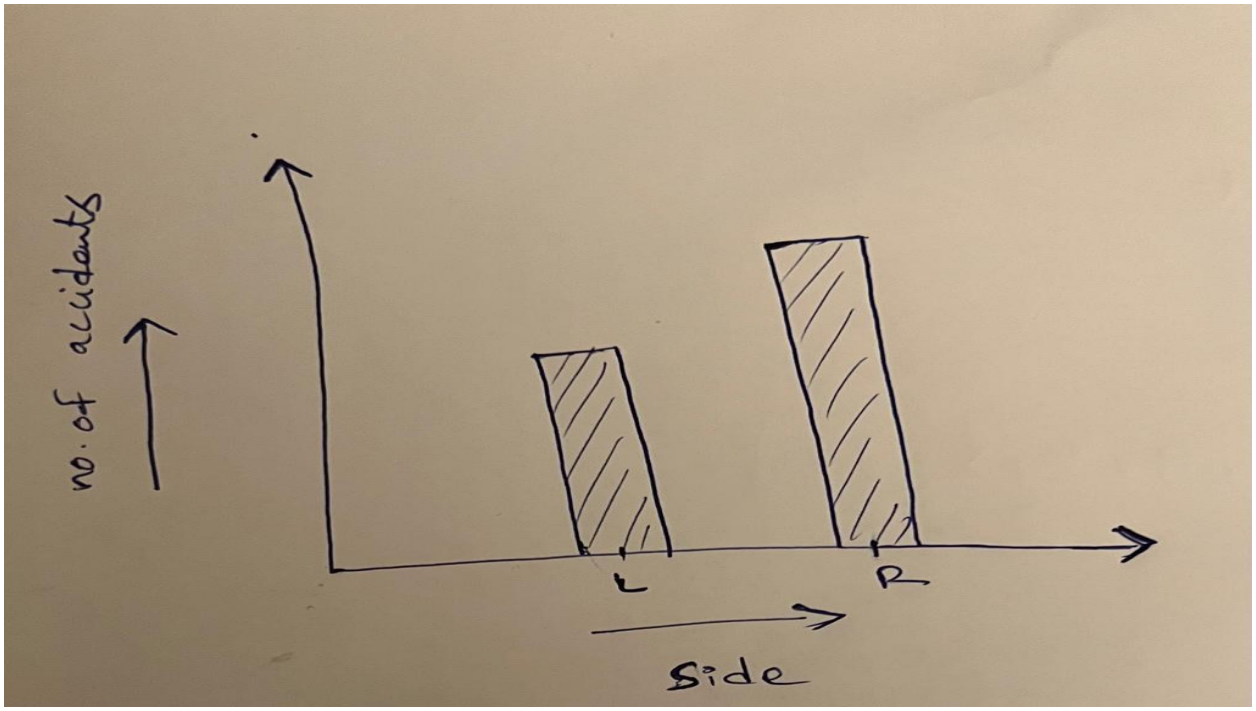
18. The number of fatalities by city [MID - 10]



19. Pie chart visualizing which vehicle type has more fatalities [MID - 11]



20. Bar chart showing which side of the road has more accidents involved [MID - 12]



Decide

Sketch ID	Message ID	Author
1	MID 1	AH
2, 14	MID 2	AH, DV
3, 9, 14,16	MID 3	AH, HC, DV,AR
4, 8	MID 4	AH, HC
5, 7, 15	MID 5	AH, HC, DV
6,17	MID 6	HC,AR
10, 13	MID 7	HC, DV
11	MID 8	DV
12	MID 9	DV
18	MID 10	AR
19	MID 11	AR
20	MID 12	AR

Top Picks :

We picked sketches 2,4,5,9,10 after a round of voting. These sketches are the most relevant to creating a coherent storyline from the data showing how different factors affect car crashes across the United States. In-depth details of the reasoning behind each sketch are added in the explanation section of our storyboard arc.

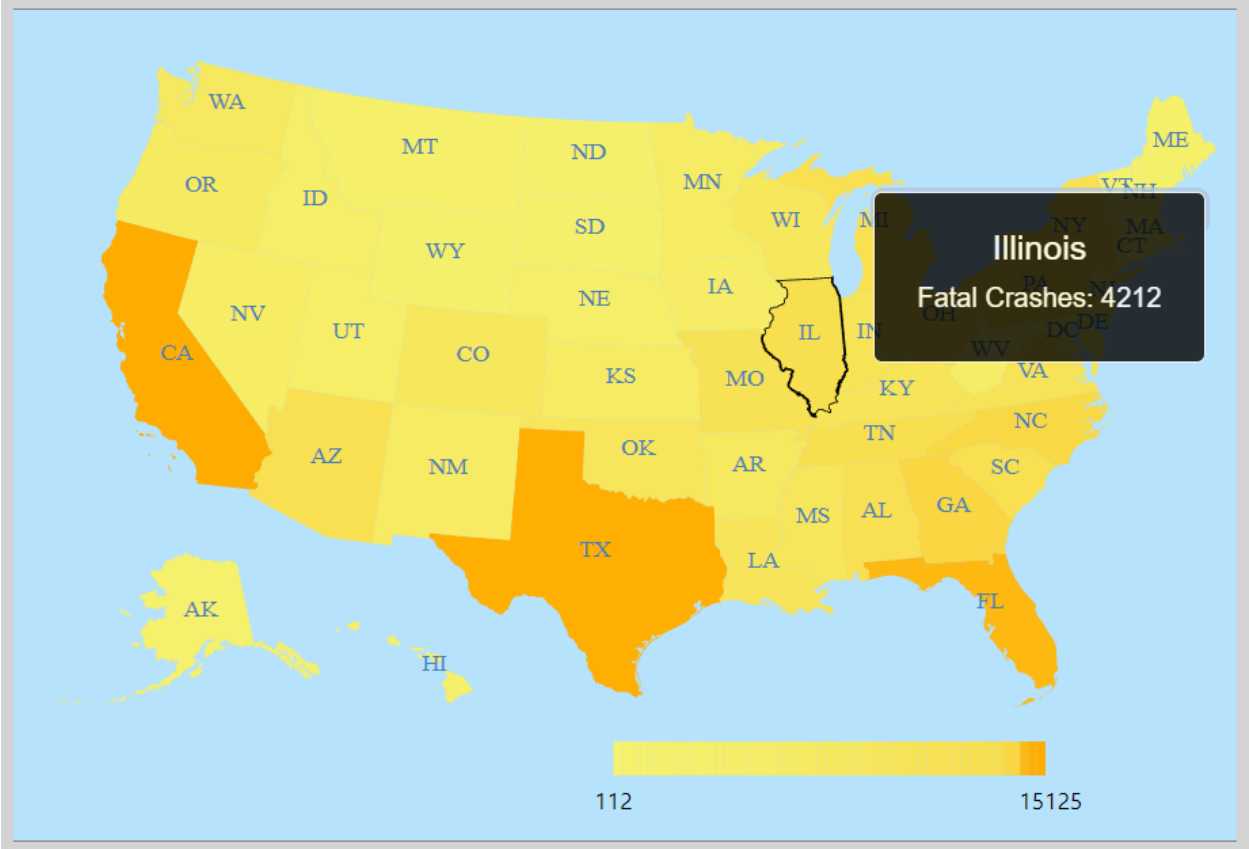
Prototype No. 1

Link to clean datasets

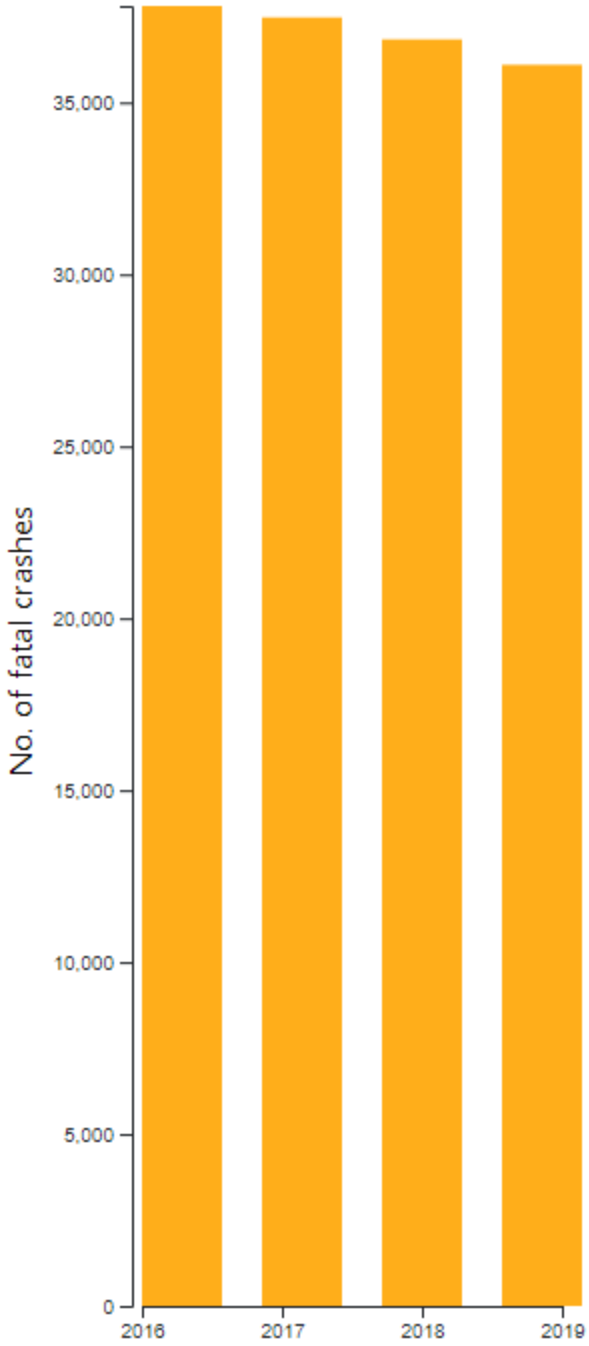
Fatal car crashes from 2016 to 2019 for all states of the U.S
fatal_year

Nature of U.S Accidents over the years
US_Accidents

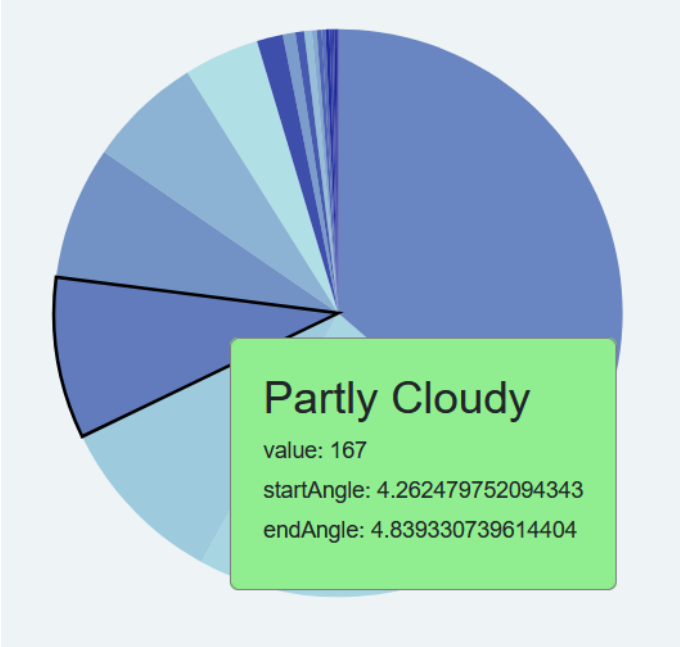
D3 Visualizations



- This will be our hook to allow the audience to get a good understanding of the severity of fatal crashes across all states in the US from 2016-2019.
- This visualization is interactive and will affect other complimentary bar charts to the side to display information about specific states onClick.
- Tooltip is implemented with opacity and other styling elements.
- Title and labels will be added with more decoration and prettifying.
- Brushing for the years 2016-2019 will be added

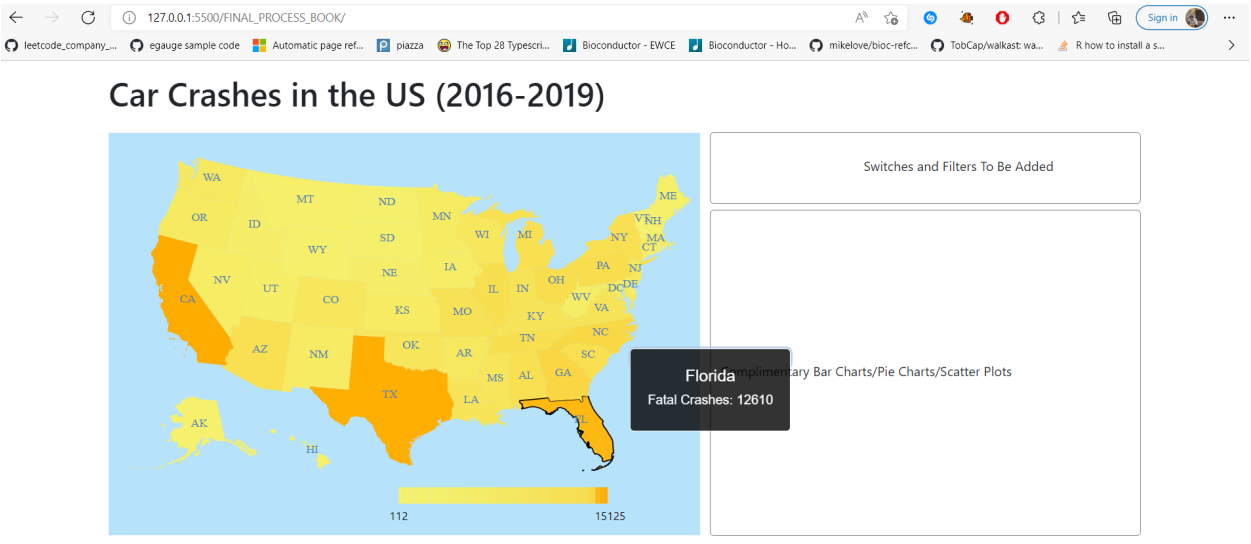


- This will be one of the rising insights of our story. It shows the number of fatal crashes over all the states in the U.S categorized by year.
- This visualization will be interactive and will change once the user clicks on a specific state. The data from 2016-to 2019 will be displayed for the selected state.
- A tooltip will be added along with the title and other information for decoration.



- This visualization will represent the number of accidents categorized by the weather conditions of the day.
- This will dynamically change as we select different states.
- Title, labels, tooltip, and other information will be added.

Rough Webpage design



- Current Web Page setup.

- Bootstrap styling is set up to be used.
- Each box will act as a dynamic component, moving in and out of the screen.
- Carousels will be used for boxes when necessary.
- Left and Right directional buttons will be added on both sides to facilitate the story points.

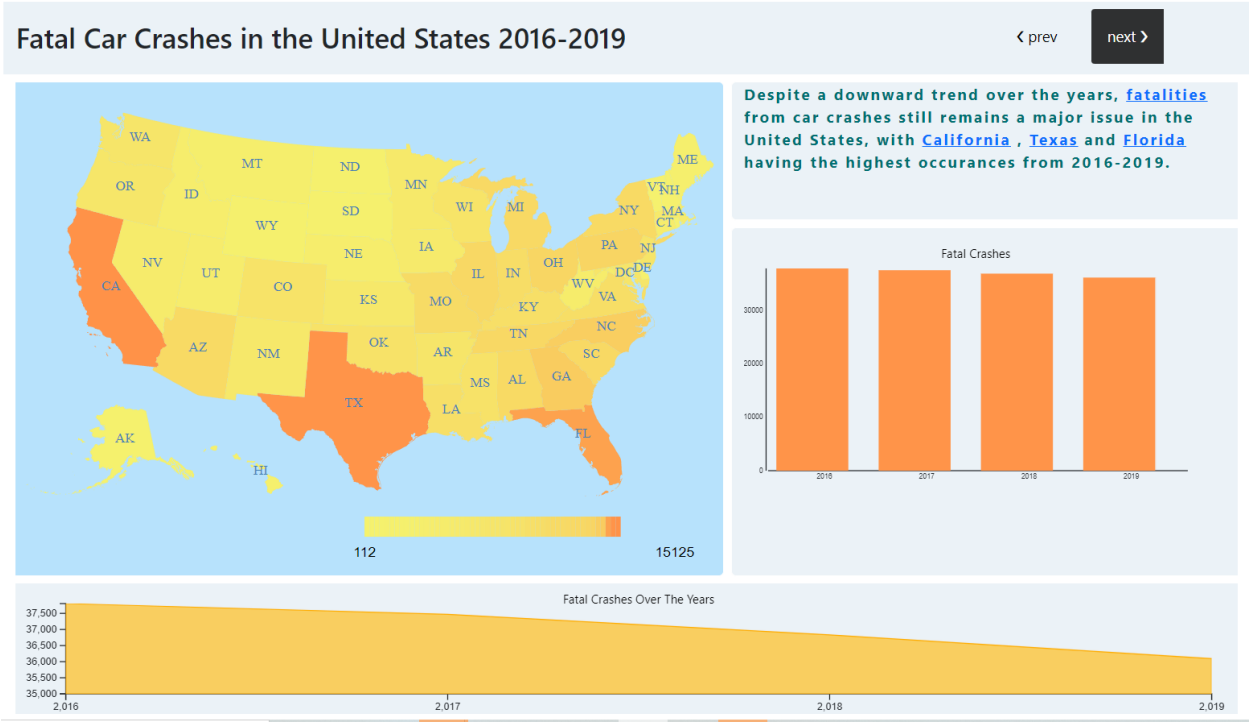
Prototype No. 2

Name of the students that worked on the submission:

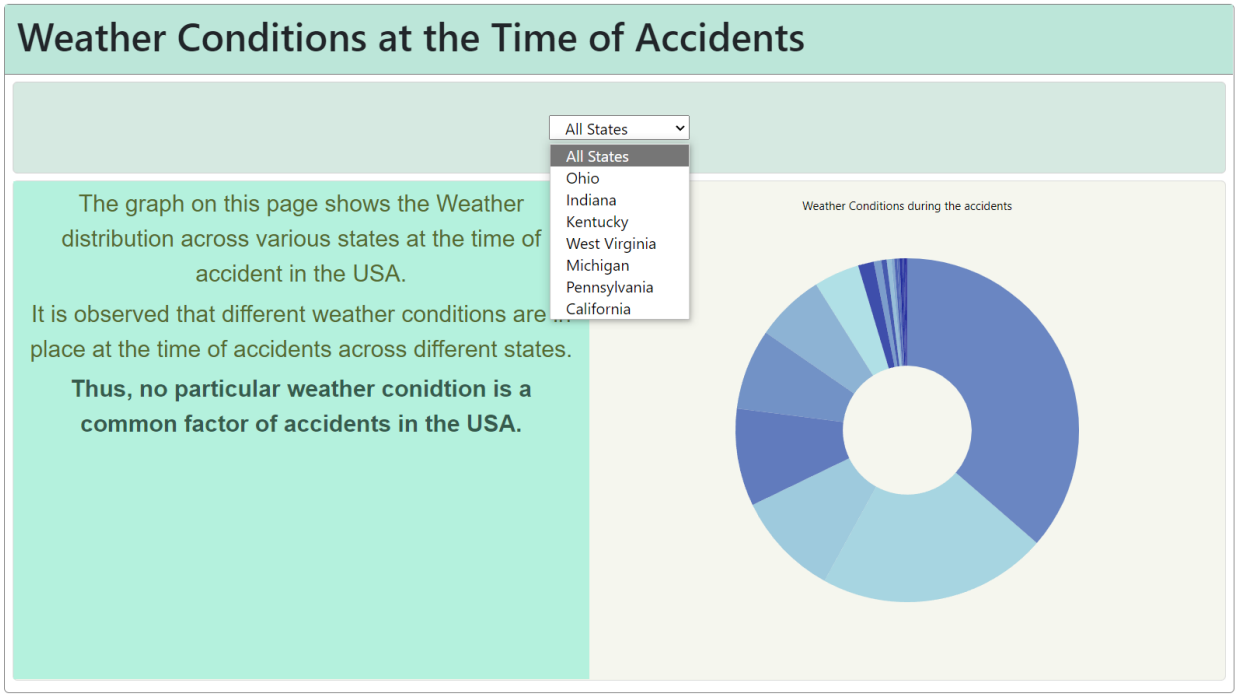
Hemayet Ahmed Chowdhury,
Darshan Vekaria,
Amith Reddy Nalla Venkat,
Asmita Hanchate

Screenshots to be added

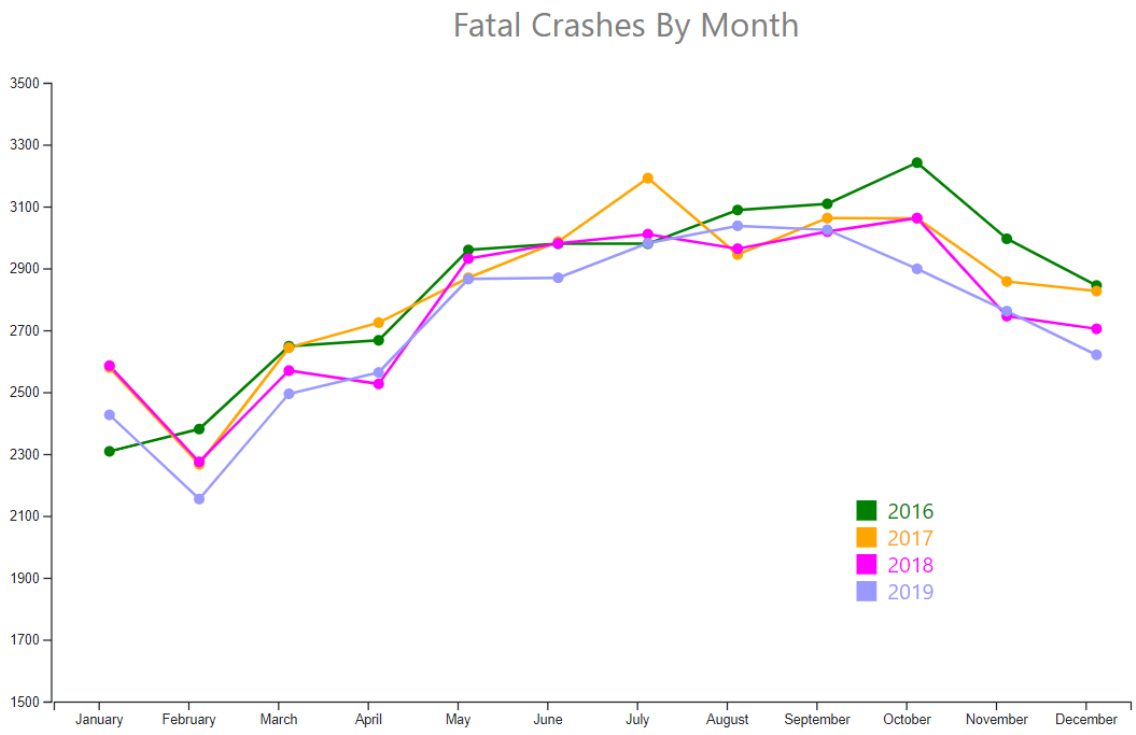
Slide 1 - Complete Visualization



Slide 2 - Complete Visualization (just the data for individual states to be added.)



Slide 3 (partial implementation. Base for the innovative chart)



Structure of the Web page-

1. Start by opening **index_main.html.** This is our main web page framework with slide mechanisms added. This also contains our visualization 1 and overall architecture.
2. **index_weather.html :** Visualization 2. Second slide of our data story. To be integrated with the main framework.
3. **Index_by_month.html** : Visualization 3. Third slide of our data story. To be integrated with the main framework.

All data files are stored in the data folder, with some dummy testing data. JS files are still under development.

Index_Main file describes a Map based visualization of the crashes that took place from 2016 to 2019**.** This visualization has a USA heat map wherein the intensity of color represents the number of crashes that took place in the respective states. The visualization has an area chart as well. The user can brush over this chart to get an updated view of the map. After brushing, the map will represent the crashes over the time period of brushing. The user can hover over the map to get a tooltip that describes the details of the hovered state. The slide also has a bar chart containing fatal crashes over the years.

Index_Weather file describes the weather condition involved at the time of the crash. A donut chart is used to display the same. A state-based dropdown selector is also used which lets the user explore state-wise weather distribution at the time of accidents. It can be deduced that different weather conditions are in place at the time of accidents across different states. Thus, it is found that no particular weather condition is a common factor in accidents in the USA.

Index_By_Month represents a line chart showcasing the statistics of crashes for different states over a span of time. Even this chart will be enabled by a state filter through which users can select to see the trend for different states individually.

(Note: This chart, moving forward, will serve as our novel visualization. We intend to create a scatter plot eventually wherein each bubble will represent the state and its size will be the number of crashes. There will be a slider to change the year. Upon moving this slider, the position and size of the bubble will change to mark the evolution of the number of crashes in that state over the years.)

Prototype No. 3

Innovative Visualization 1

Description:
Our innovative visualization is a **Dynamic Scatterplot with a Timeline Slider and Autoplay** which shows the **number of car crashes** that take place on **every hour of the day,** categorized by **months of the years 2016-2019.**

Purpose: This visualization gives the user an understanding of which hours of the day cause the most accidents. For further details, the user can also use the slider to go through months to understand which part of the day causes more accidents in different months.

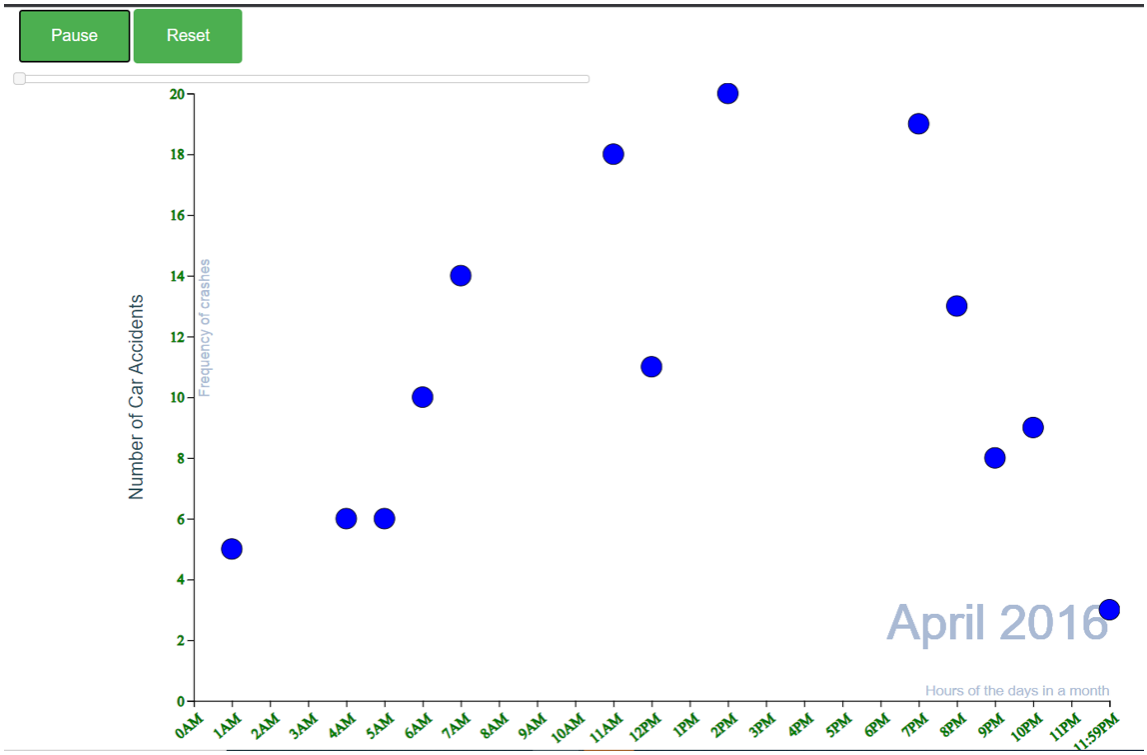
How it fits in the story: This slide is to give the user an understanding and raise awareness of how the time of the day contributes to accidents and which time to be aware of the most.

Details: The **slider** marked inside the **blue box** allows the user to scroll through the different months and see how the rate of accidents was for every hour of the day. The **play** button makes this visualization automatically transition from month to month over the years. The **play** button converts to a **pause** button after play is clicked which allows users to pause the video-like visualization at any frame and **analyze** the frame better. A tooltip is also attached for the user to understand the details of the data better. The **red box** shows the month the user is currently in. The **scatter dots** transition from frame to frame upon the play button click, to create a timeline video-like visualization so the user understands how the data changes shape over time.

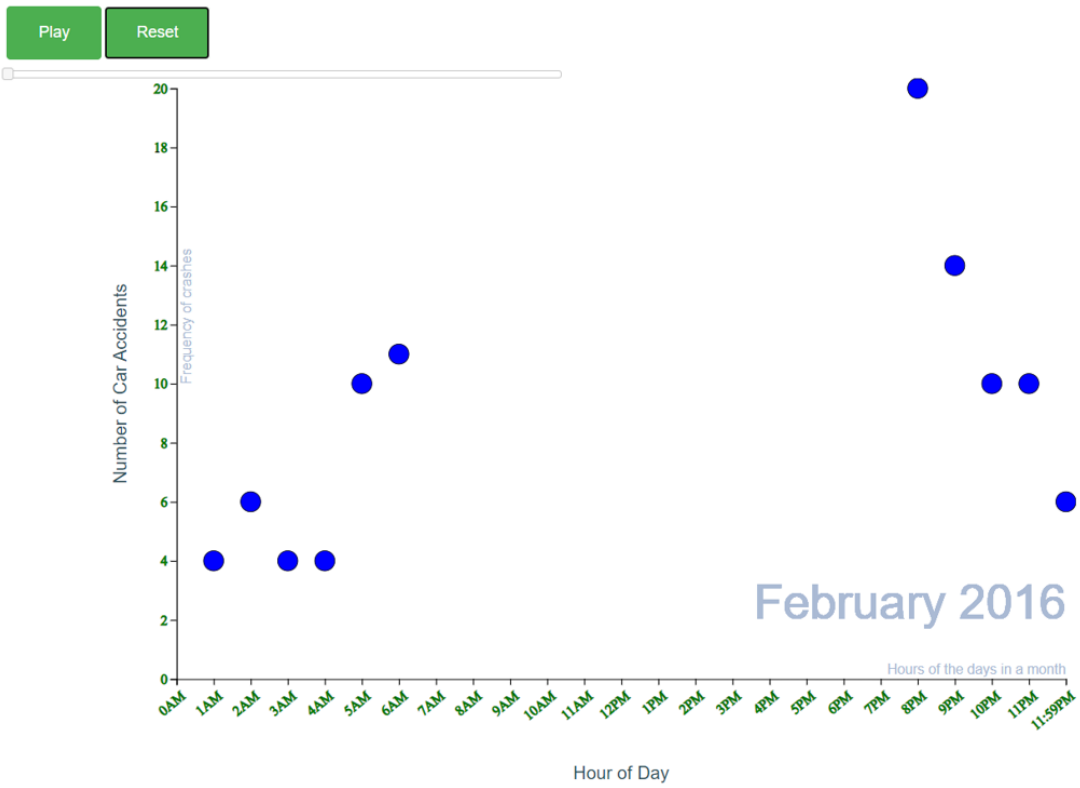
Here is the video link showcasing the dynamic visualization: https://drive.google.com/file/d/1DMqZRJcBSpF8phPuYN-5AYKwNDYQPi_a/view?usp=sharing

Innovative Visualization Frame 1: (interactions implemented, the full dataset will be attached)

Innovative Visualization Frame 2 :

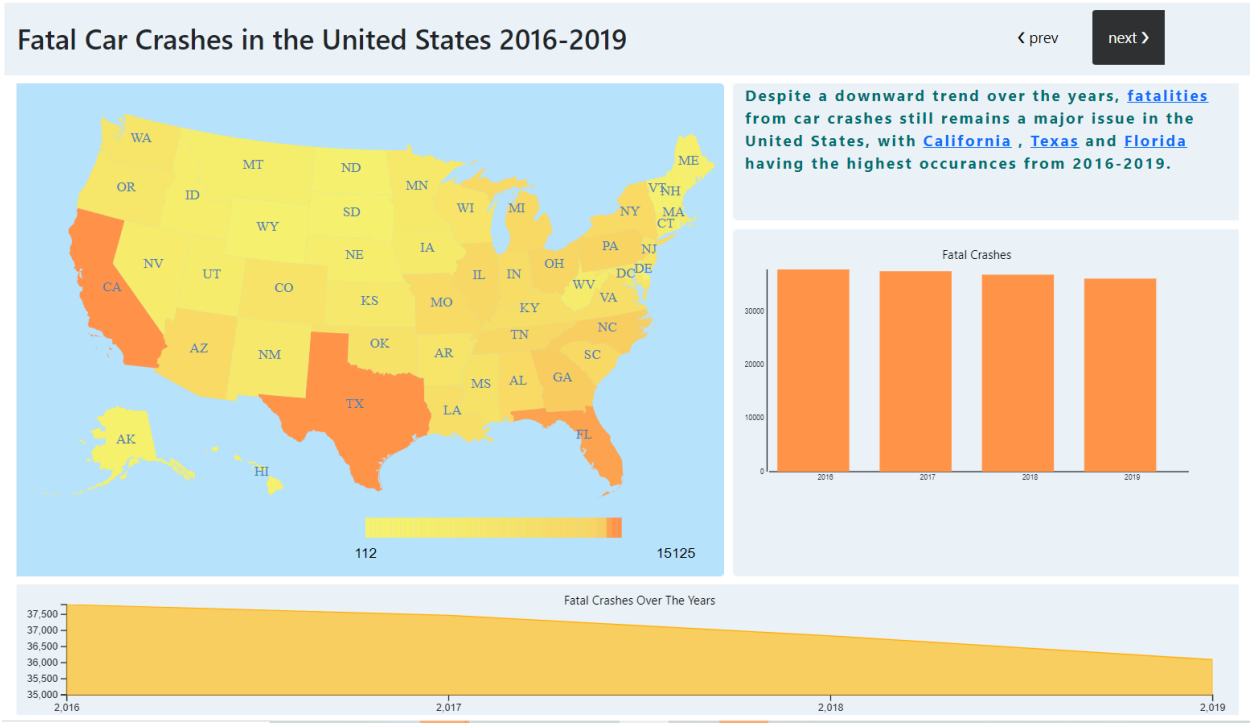


Innovative Visualization Frame 3 :



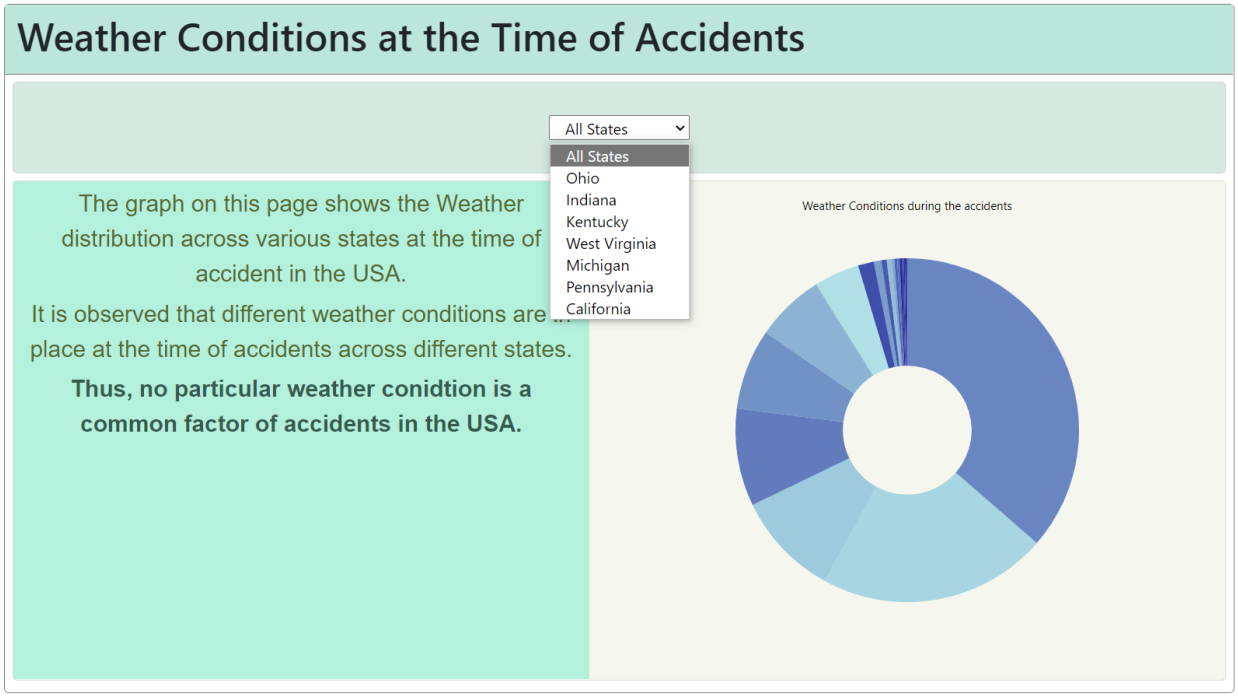
Other Visualizations

Slide 1 - Fully Implemented (Interactions fully implemented)



Description: This is a Map based visualization of the crashes that took place from 2016 to 2019**.** This visualization has a USA heat map wherein the intensity of color represents the number of crashes that took place in the respective states. The visualization has an area chart as well. The user can brush over this chart to get an updated view of the map. After brushing, the map will represent the crashes over the time period of brushing. The user can hover over the map to get a tooltip that describes the details of the hovered state. The slide also has a bar chart containing fatal crashes over the years.

Slide 2 - Partially implemented (larger dataset will be used for all states)



Description : This visualization describes the weather condition involved at the time of the crash. A donut chart is used to display the same. A state-based dropdown selector is also used which lets the user explore state-wise weather distribution at the time of accidents. It can be deduced that different weather conditions are in place at the time of accidents across different states. Using the tooltip, observations can be drawn about which weather affects car crashes the most.

Slide 3 - Fully implemented (with state-wise filtering)

Description : This visualization represents a line chart showcasing the statistics of crashes month-wise for different states over a span of time (2016-2019) .The chart is enabled by a state filter through which users can select to see the trend for different states individually over the above span of time.

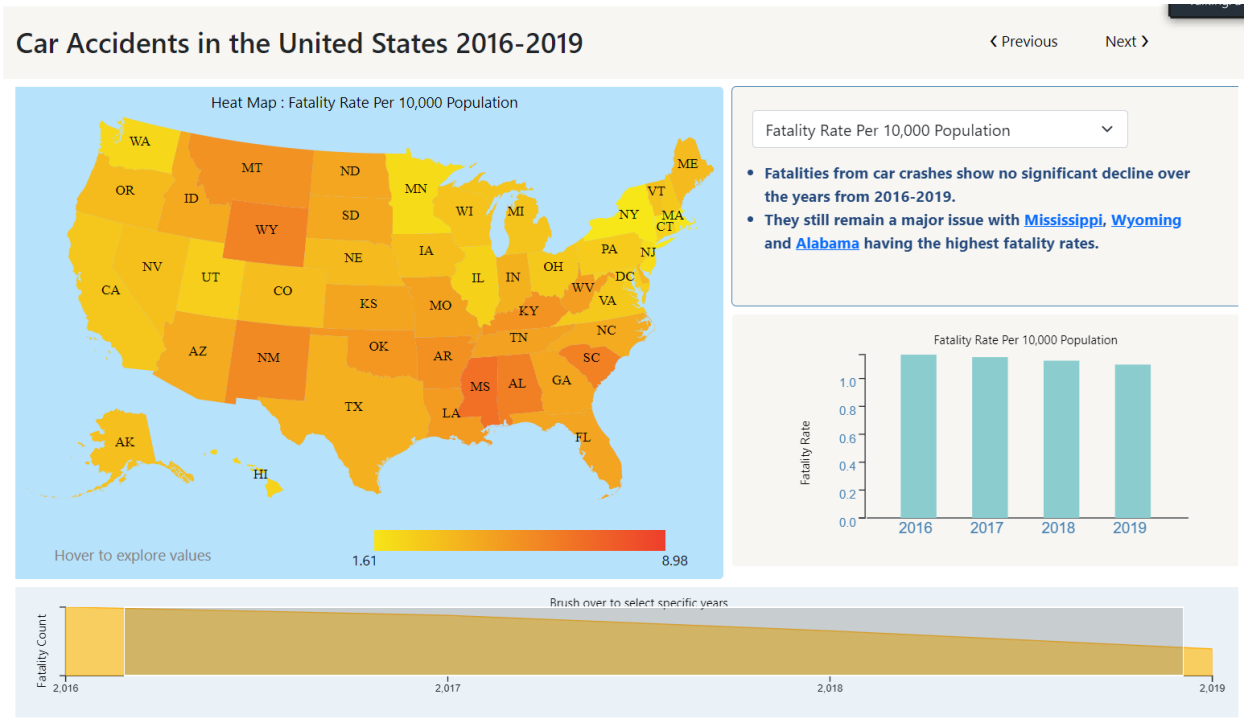
Prototype No. 4

Data and Usage:

Copy all the content given in the drive link: https://drive.google.com/drive/folders/1a45P3k7G2DhfuEifEgEAhk7DbT4jHuq4?usp=sharing into your data folder of the JS project. There are no subdirectories, all individual files should be directly in the data folder.

(As instructed, we had to keep the data separate because of its huge size. The project has been working well on the machines of all team members. Please let us know in case you run into any issues.)

Visualization 1



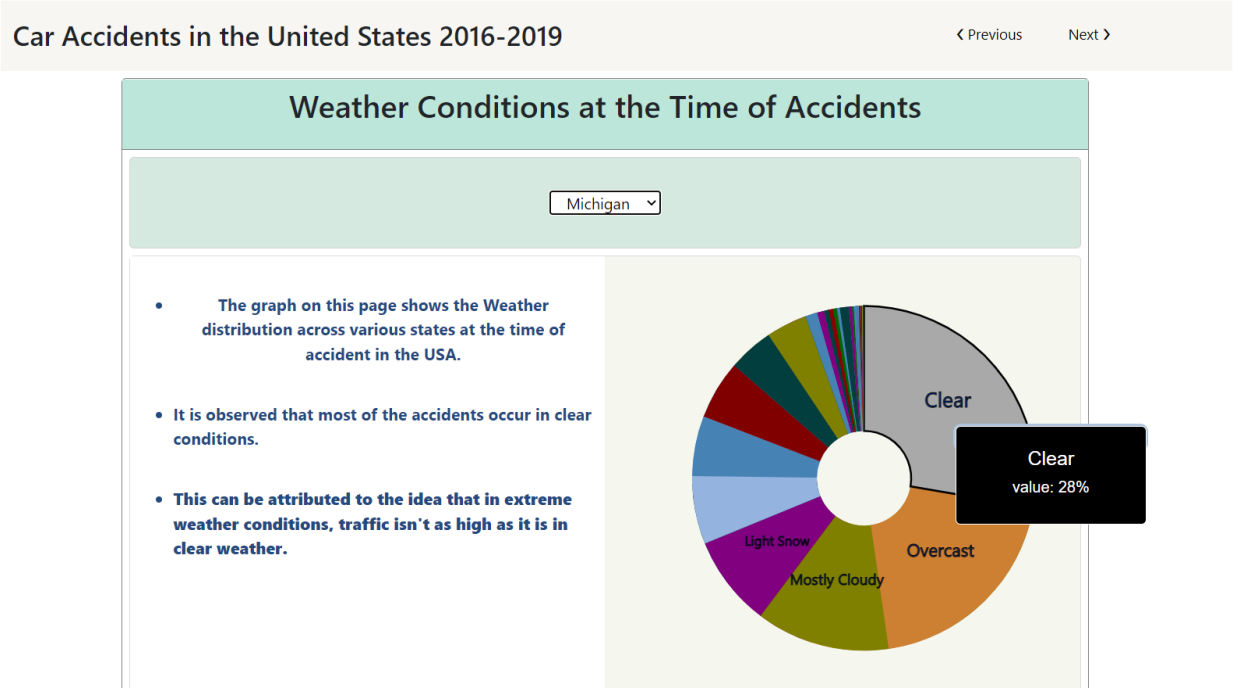
Description: Our first visualization is a **choropleth map** representing the number **of crashes** categorized by **states** across a span of **time(2016-2019)**. Below is a timeline from 2016-2019 where the user can **select a time frame** to analyze the number of crashes across all states for that time period. To the right, is a **bar chart** displaying the **trend** of crashes from 2016 to 2019.

Purpose: The purpose of this visualization is to give the user an **overview** of fatal crashes over a time period **(2016-2019)**. It is sort of an entry point for the user to get a headstart on analyzing different factors influencing the crash.

How it fits in the story:
It is the Hook of our story which gives the user an overview of the entire story we are trying to portray.

Details:
The **map dynamically changes** on selection of a specific **time period** in the timeline provided below it. We also have a filter where the user can select if they want to see relative values i.e., number of crashes by population of that state or absolute values of number of crashes that have occurred in that state.

Visualization 2



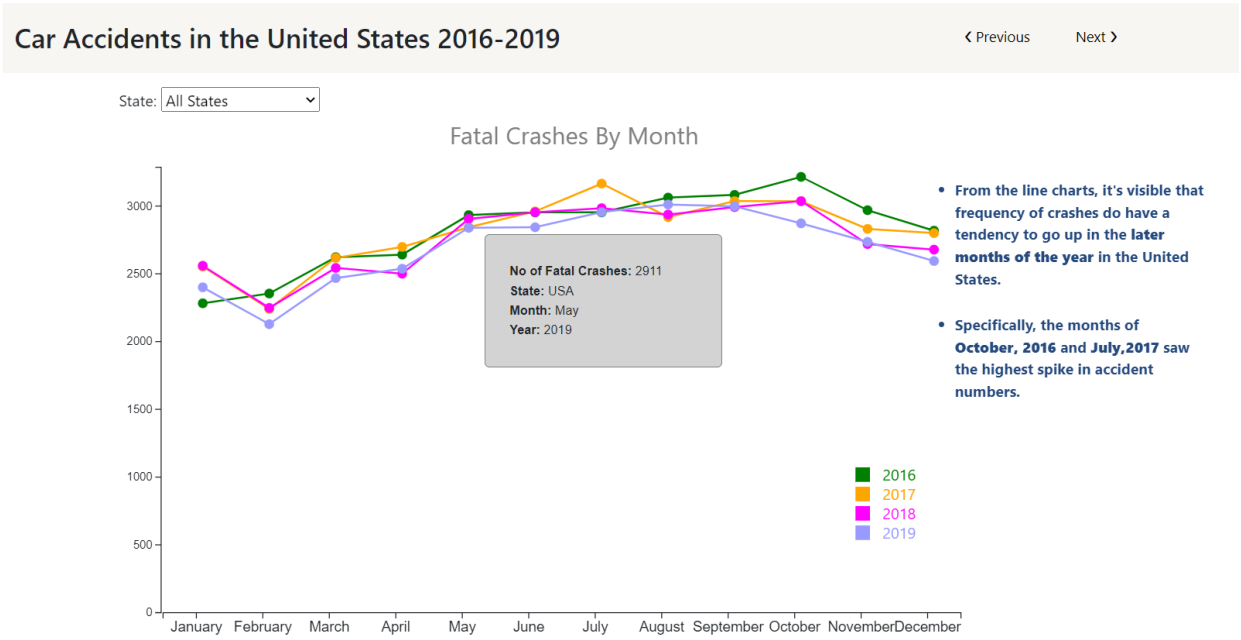
Description: Our second visualization is a donut chart which represents the weather condition at the time of the crash categorized by the states.

Purpose: The purpose of this visualization is to make the user help draw a conclusion if the weather condition played an important role in causing the crash.

How it fits in the story: The state-wise filter gives the user a chance to explore weather being a factor in influencing the car crash. The user can explore and draw their own conclusion on which factors they find important and be aware of what weather they should drive carefully in.

Details:
The visualization has a **state-wise filter** where the user can select which state they want to see the data for. The donut chart gives a representation of the number of crashes categorized by the weather condition at the time of the crash. A tooltip has been added which shows the **percentage of crashes** for that specific weather condition.

Visualization 3



Description:
Our third visualization is a line chart which shows the number of fatal crashes by month of the years 2016-2019 for all the states.

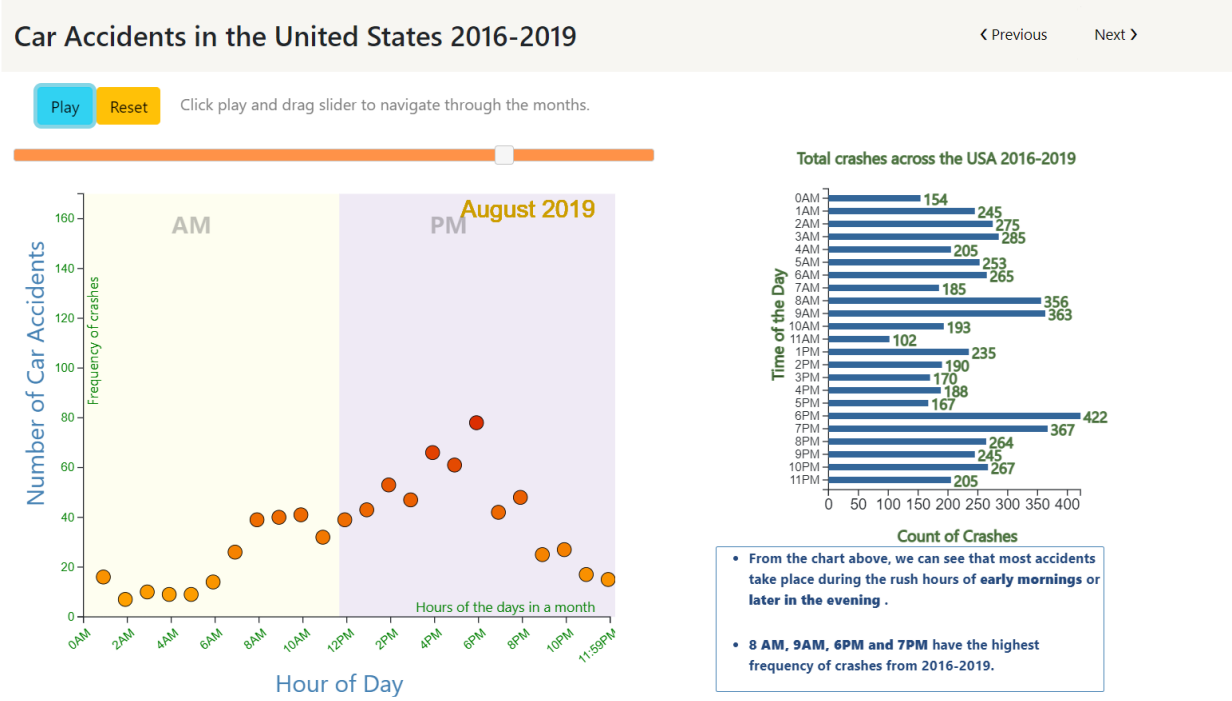
Purpose:
It gives us an understanding of which month of the year 2016-2019 of a state has more crashes involved and what is the trend of the crashes over the months. It also visualizes which year has more number of fatal crashes involved for a specific month of a state.

How it fits in the story:
This visualization is to give users a clear understanding of which month contributes to the most number of fatal crashes so that they can focus on a specific month to minimize the crashes.

Details:
We have multiple line graphs where each **line graph** represents **fatal crashes** for each of the years **2016-2019**. A legend has been added to the visualization to get an understanding of which line is for a specific year (2016-2019). A **filter** is added so that the user can **select the**

state and the data will be filtered based on state and will be visualized accordingly.The axes are created dynamically to get the visualization better when a particular state is selected by the user. A **tooltip** is also added that includes number of fatal crashes, month,state and year so that when the user moves the cursor over the graph the user will be able to get to know the number of crashes of the state and also helps better in understanding the data.

→ Visualization 4 (Innovative Visualization 1)



Description:
Our first innovative visualization is a **Dynamic Scatterplot with a Timeline Slider and Autoplay** which shows the **number of car crashes** that take place on **every hour of the day,** categorized by **months of the years 2016-2019.**

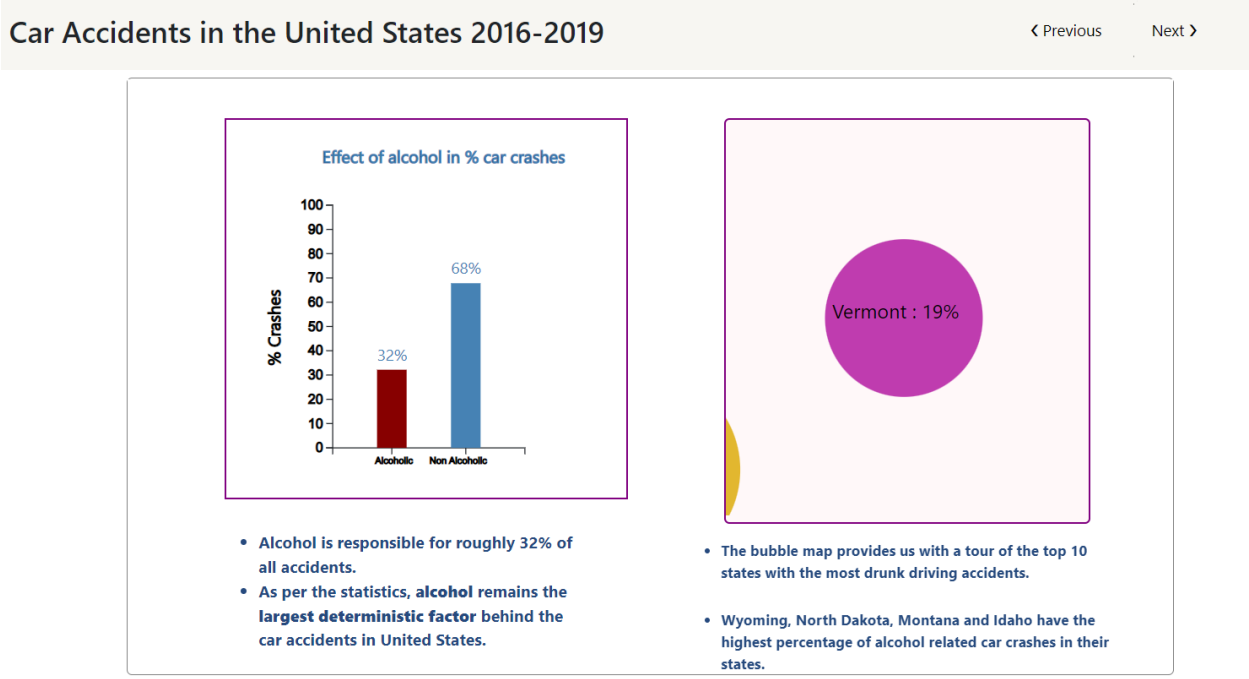
Purpose: This visualization gives the user an understanding of which hours of the day cause the most accidents. For further details, the user can also use the slider to go through months to understand which part of the day causes more accidents in different months.

How it fits in the story: This slide is to give the user an understanding and raise awareness of how the time of the day contributes to accidents and which time to be aware of the most.

Details: The **slider** marked inside the **blue box** allows the user to scroll through the different months and see how the rate of accidents was for every hour of the day. The **play** button makes this visualization automatically transition from month to month over the years. The **play** button converts to a **pause** button after play is clicked which allows users to pause the video-like visualization at any frame and **analyze** the frame better. A tooltip is also attached for the user to understand the details of the data better. The **red box** shows the month the user is currently in.

The **scatter dots** transition from frame to frame upon the play button click, to create a timeline video-like visualization so the user understands how the data changes shape over time.

→ Visualization 5 (Innovative Visualization 2)



Description:
Our second innovative visualization is a **Bubble chart with Auto Zooming** which shows the **number of crashes under the influence of alcohol**. There are 10 bubbles, each representing a **state** containing the **percentage of crashes** that are **alcohol-related** categorized by **months of the year (2016 - 2019).**

Purpose: This visualization portrays the grave consequences of driving under the influence of alcohol. It showcases how alcohol is a huge factor contributing to the majority of car crashes.

How it fits in the story: This visualization is our AHA moment to bring the user closer to a conclusion. The story analyzes a lot of factors that may potentially have been causing fatal crashes and no other factor gives a conclusive answer to the question as much as alcohol does. It outweighs other factors by comparing alcohol-related crashes with the total number of crashes.

Details: This visualization shows a zooming enabled bubble chart that zooms over the top 10 states having the highest percentage of crashes due to alcohol across the US. Supplementing it

is a bar chart that depicts the impact of alcohol-driven crashes, which clearly shows, that alcohol causes around 1/3rd of accidents in the US.

Test

Tester Name: Ashish Bhat, Nidhi Priya, and Surendrabikram Thapa
Tester Email: ashishbhat@vt.edu, nidhipriya@vt.edu, surendrabikram@vt.edu

Answers/Suggestions from both the testers have been summarized below.
Suggestions from the TA (Leonardo Pavanatto Soares) have also been considered and incorporated in our data story.

General Observations from the think-aloud study:

- All visualizations are interactive and very engaging.
- The data presented is appropriate and easy for the tester to grasp and understand.
- The story has been presented clearly and the main message we want to convey is clear.
- The flow of the story was intuitive and all visualizations were appropriate for the kind of data being presented.

What does the tester like about your data story?

- All factors that could have potentially led to the fatal crash were covered in the data story.
- It was very clear and intuitive to follow through with and led to the main message we wanted to convey.
- The state-wise filter enabled the tester/user to explore everything in detail.

What improvements does the tester point out?

Following are the improvements that the testers have suggested -

- The Introduction slide for the story needs to be more engaging and catchy.
- In the first visualization, the font for the map title should be enlarged for readability.
- Instructions for brushing should be mentioned clearly.
- The states highlighted in blue and underlined should be mentioned in another color to avoid the confusion of them being a hyperlink.
- The carousel should not cycle back to the main page to avoid confusion on where, to begin with.
- Years on the X-axis in the first visualization should be mentioned without the commas.

Was the intended key message clear to the tester? Why or why not?
The intended message was clear to the tester. Both, the key findings and the main message have been clearly presented in the last slides of our data story to aid the users. The visualizations and the ease of flow of the data story also aid the users in understanding the key message clearly.

Did the tester get your next steps or call to action? Why or why not?
The testers were able to grasp the next steps/ call to action as the carousel has been provided for the user to go to every step of the story. The only instances where the testers were not clear on interacting with the visualizations were-

- The instructions for brushing in the first visualization were not clear.
- The carousel circled back to the first page for the story and the tester could not identify it.

Final Project - Wrap Up

We have created several visualizations to convey our story on Car accidents in the USA. We have ensured the use of smooth transitions for all interactive visualizations.

We have created **two innovative visualizations** in order to **effectively convey** our story and **engage** the audience.

Github Deployment :

The fully functional website with complete data is deployed here :
Link : https://asmi06.github.io/

Data and Usage:

For running the code submitted on canvas locally, copy all the content given in the drive link: https://drive.google.com/drive/folders/1W2rNVs1Jr5EkA6pHc1_FyBBiqUBvdRGJ?usp=sharing

into your data folder of the project. There are no subdirectories, all individual files should be directly in the data folder.

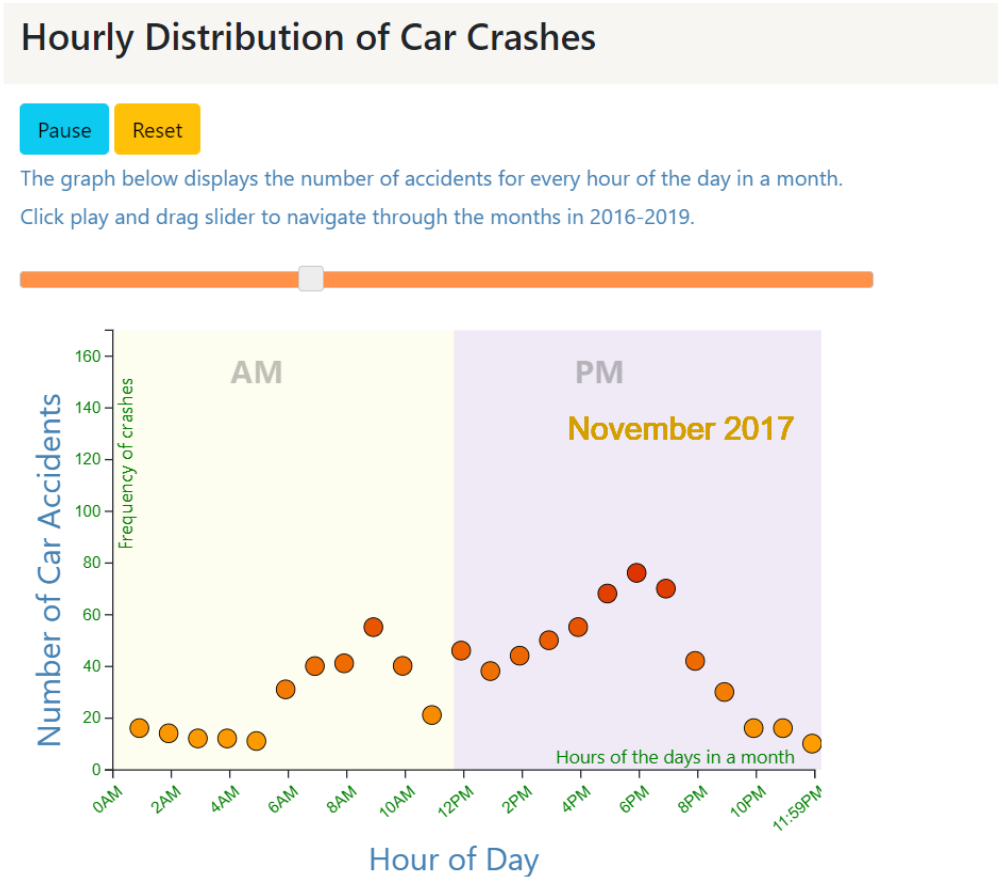
(As instructed, we had to keep the data separate because of its huge size. The project has been working well on the machines of all team members. Please let us know in case you run into any issues.)

Github repository (code) : https://github.com/Asmi06/Asmi06.github.io

Video Link :

https://drive.google.com/file/d/1xAdDbm39bvDMuZJTluOsuyInhXNZ_t8N/view?usp=sharing

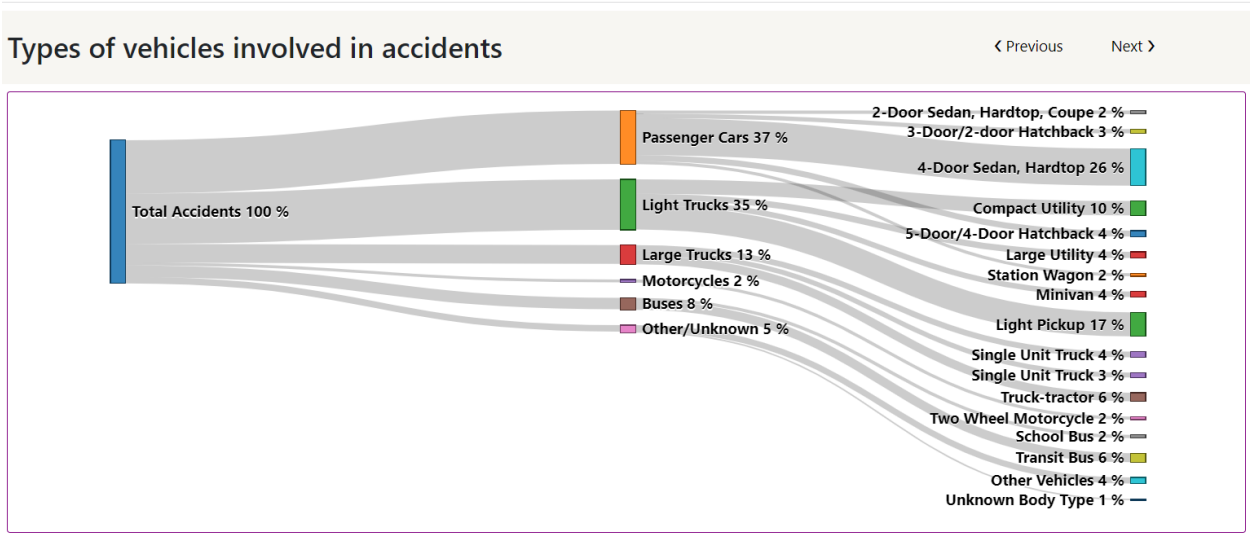
Innovative Visualization 1:



The purpose of this **interactive dynamic scatter plot visualization** is to help the viewer check which hours of the day are frequently facing crashes for calendar months across the years 2016 - 2019.

A viewer can use the Play/Pause button to see the dynamic scatter plot showcasing the number of accidents at each hour of the day. The user can drag the slider back and forth if they wish to see the statistics for a particular calendar month. The smooth transition of the scatter plot lets the user observe a changing pattern in the number of accidents for each month. Reset Button enables the user to start again from the year 2016.

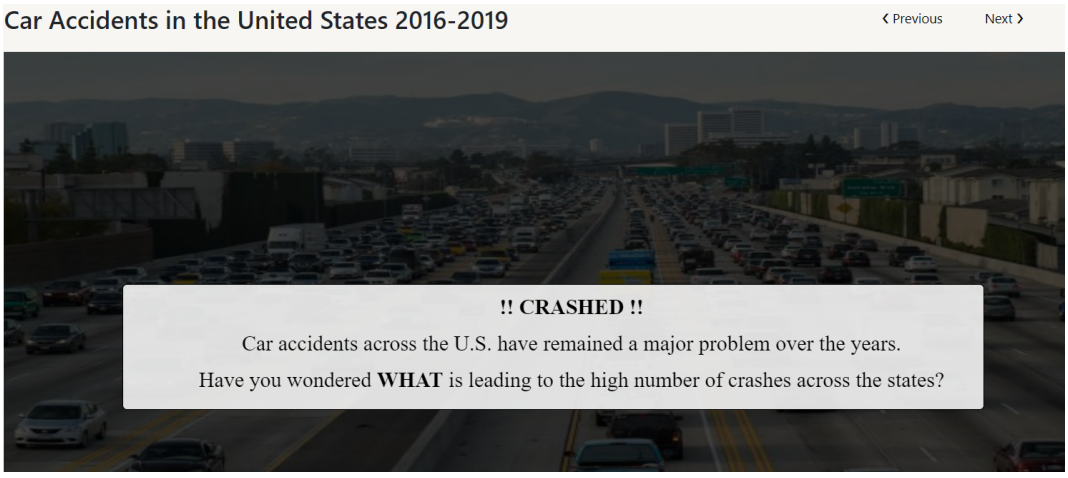
Innovative Visualization 2:



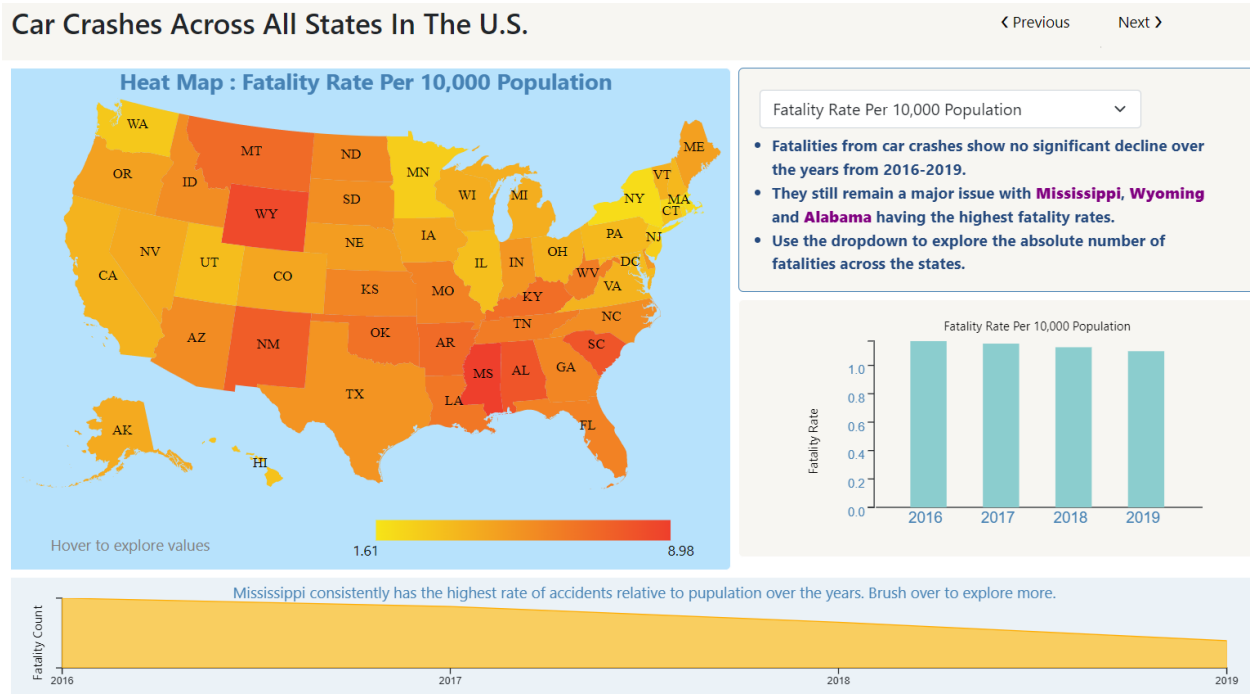
- The sankey graph explains the percentages of categories and sub categories of vehicle types that contributed to the car crashes from 2016-2019.
- It is found Passenger cars contribute the most to the total number of accidents followed by light trucks.
- Upon **deep dive** into subcategory, it is seen that **4-Door Sedan (26%)** and **Light Pickup trucks (17%)** contribute significantly to the crashes.

This is a **Sankey Chart** which lets the audience see the distribution of types of vehicles involved in car crashes across the USA. Each represents the type of vehicle followed by the percentage count of its involvement in accidents. Each type of vehicle is further divided into subcategories to enable a deep dive of the scenario to users. The nodes are connected by pipelines to enhance cognition and readability.

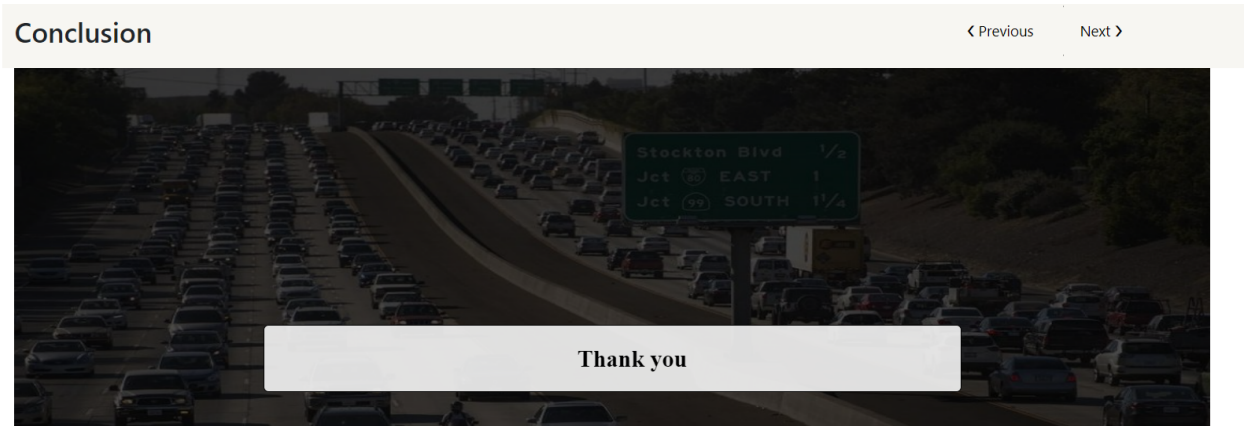
Improvements Done after Testing Phase:



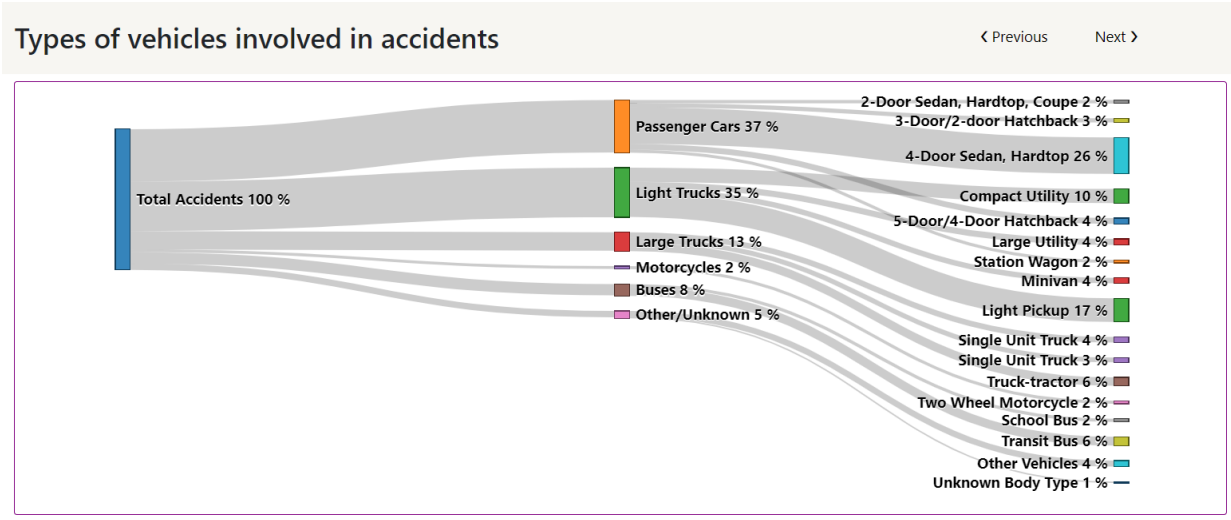
Made introduction page more engaging



Updated fonts and color for better readability and added clear instructions for brushing



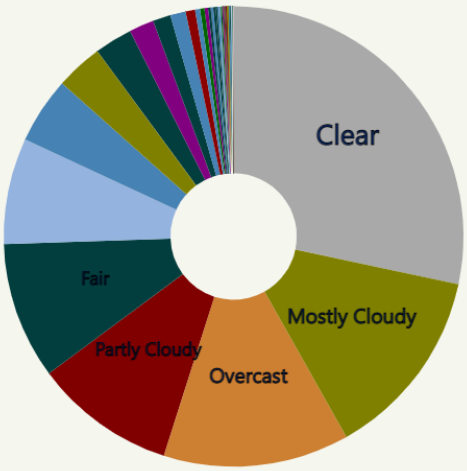
Added an End Page to Conclude the story and stopped its lopping back to first page.



Removed the Bubble Zoom chart (submitted in previous submission) and added this sankey chart for effective storytelling and use engagement.

All States

- **The graph on this page shows the Weather distribution across various states at the time of accident in the USA.**
- **It is observed that most of the accidents occur in clear conditions.**
- **This can be attributed to the idea that clear weather is more common throughout the year in general and traffic is also heavier on clear days.**



Added Conclusion of Clear Weather Conditions being dominant for most cases