INVESTIGATION OF AVIATION ACCIDENTS

(WWW.AVIATIONCV.COM, 2016)

12-JUN-2019

Data Visualization Project

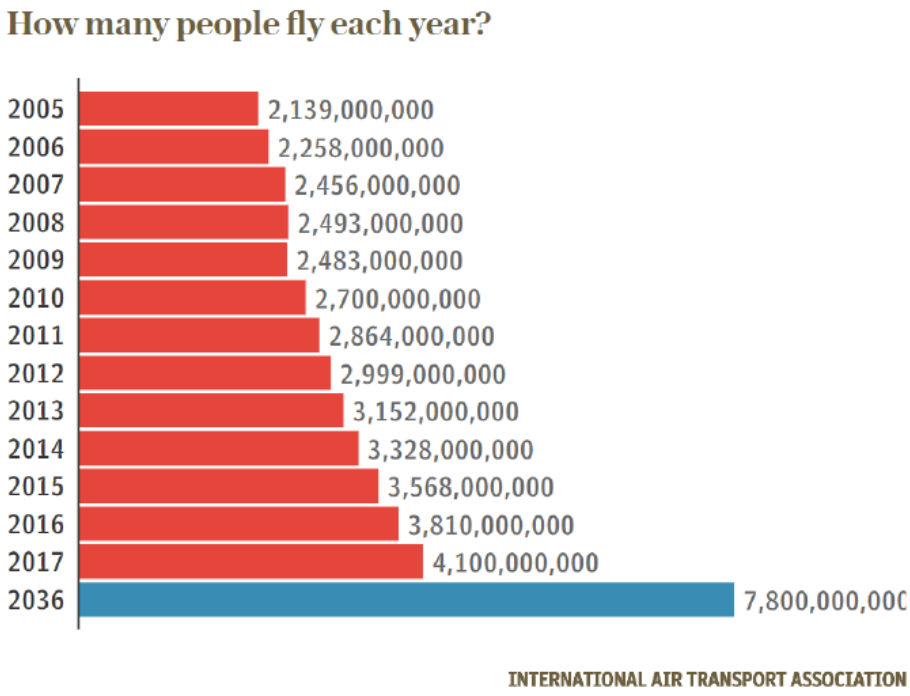
Name: Arunava Munshi	Organization: Monash University

TABLE OF CONTENTS

INTRODUCTION

Motivation

Aviation industry in today's world gainfully enjoys its commercial success with the ever increasing number of customer choosing to fly. There are myriads of reasons for such customer behavior such as increased income level of passengers, hassle-free & shorter journey, international relations, advancements in aviation technology etc. Commercial flights flew nearly 4 billion passengers in the year 2017 itself that is more than twice the statistics 12 years ago (Telegraph, 2017). According to International Air transport association, almost 8 billion people are going to fly in 2036. So, considering this huge boom in this industry, air safety has been gradually becoming an immensely important matter for the aviation authority. Our project thus, tries to visualize different factors affecting aviation accidents from the available NTSB data on Kaggle and wishes to provide some comprehensive grounds to be consider by aviation industry.



(www.iata.org, 2019)

Project Description

The project aims to do an interactive visualization on NTSB Data for Aviation Accidents and tell a story on the history of aviation accidents happened so far and different factors majorly affecting these accidents. From the exploration project previously we could certainly dig out these factors, which we shall show within this project through interactive visualization. Please note that, here we assumed, only the fatal accidents are the areas of concerns for the aviation authority and our project will show three of such most important factors leading to these accidents. These are as follows –

1. Aircraft makers and models as a factor for accidents
2. Location as a factor for accidents
3. Different primary and secondary reasons as factors for accidents

Dataset Used

We had already done with the data wrangling part before starting this project. Because data wrangling was the part of our previous exploration project. So after doing the wrangling on the main data, we could come up with the dataset of our need for this visualization project.

File Names	Brief Description	Size
Aircraft_Makers_No_Of_Accidents.csv	Contains year-wise accident information for different aircraft maker companies and their respective models involved into accidents.	23204 rows x 5 columns
Location_Wise_Accidents.csv	Contains year-wise locations details such as country, state, city, zip, geocode etc. information for different accidents.	21717 rows x 10 columns
Reasons_For_Accidents.csv	Contains year-wise primary, secondary reasons for different air crashes.	31613 rows x 9 columns

DESIGN

Description of the visualization design process

Below is the detailed explanation of how I came up with the best visualization designs to depict the different aspects related to aviation accidents. I have used the five design sheets methodology to come up with the final design through a series of alternative designs. From the huge NTSB data, I could filter out 6 factors that are our main areas of consideration to investigate the trend of aviation accidents from 1948 to 2017. I will now discuss about their contribution to the accidents in various extents. The factors are:

1. Aircraft makers
2. Aircraft models
3. Location of accidents in states
4. Location of accidents in cities
5. Primary reasons behind the accidents
6. Secondary reasons behind the accidents

Sheet 1

IDEAS:

I have initially drafted mini design ideas in Sheet 1 considering all the factors that play a role in the aviation accidents. I have designed a bar chart which depicts the number of accidents caused by the various aircraft makers. For representing which aircraft models have seen the highest and lowest number of accidents, a line graph is drawn. Each line chart is prepared for an aircraft maker where the x-axis shows the aircraft models for that particular maker and the y-axis shows the total number of accidents caused by those models.

The accident prone zones are represented as dots in the map, which shows the cities of USA. I have thought of showing the total number of accidents that have taken place in each state of USA through a bar graph.

There are just 5 primary reasons which are depicted through a bar chart against the number of accidents that was caused due to each reason in the y-axis. Each primary reason has been categorized into several secondary reasons, which has also been shown in a line chart. The x-axis labels the secondary reasons and the y-axis labels the number of accidents caused due to each secondary reason.

FILTER:

The filtering criteria among the huge NTSB data remains those 3 factors already mentioned above.

CATEGORIZE:

I have categorized 2 relevant factors in each of the layouts for a meaningful pictorial representation. I have depicted aircraft maker and its respective models together which will give an understanding to the viewer which aircraft makers or models have the highest or lowest accident counts.

Similarly, I have categorized the states and their respective cities of USA into a separate layout because both will give an idea about the most and the least accident prone areas.

Also, representing the primary and the secondary reasons separately doesn't make sense because both are inter-related and showing both simultaneously in another layout would help us in a better understanding of the main reasons that contribute to the accidents and should be taken care of later in the future.

COMBINE AND REFINE:

Thus we combined and refined 2 factors together in 3 different layouts for better visualization, which will be discussed in Sheets 2, 3 and 4.

VISUALIZATION IDEA:

Thus our visualization idea narrows down to the following in design sheet 1:

1. Visualize aircraft makers and their respective models along with the number of accidents into one group
2. Visualize the states and cities of USA along with the number of accidents into one group
3. Visualize primary and secondary reasons for the accidents along with the number of accidents into one group

Sheet 2

LAYOUT 1:

PANEL 1:

As discussed in sheet 1 that the aircraft makers and their respective models will be combined together, so I have decided to go with the stacked bar graph visualization where each bar will represent the aircraft maker and the divisions of each bar will represent its models against the number of accidents in y-axis that has been caused by those models. Here, I have introduced an input value as 'Years' so that the user/ viewer has the flexibility of viewing the rate of accidents for one or more number of years. The input value is designed as checkbox, each representing years from 1948 to 2017. According to the years selected, the stacked bar graph would change depending on whether the aircraft makers/models were present during that time period.

PANEL 2:

The map is refined to present visualization with dotted points representing location of accidents derived from geocodes. The input is same as that used in Panel 1. On selecting the specific years, only those locations will be shown as dots on the map where accidents have occurred during those selected years.

PANEL 3:

For primary and secondary reasons, I have thought of the stacked bar chart where each bar will represent the primary reason and the divisions of each bar will represent its secondary reasons against the number of accidents in y-axis that has been caused due to the secondary reasons. The input is the years in checkbox. The number of accidents will vary based on the time period chosen.

OPERATIONS:

One can select years (one or more) from the checkboxes and the corresponding graphs or maps will appear on the screen. One can also select and deselect the years from the checkboxes and the data for those years will appear accordingly. For any changes done in the checkboxes, the visualizations change dynamically.

FOCUS:

Stacked bar chart can plot multiple data on one bar, allowing multiple items to be plotted efficiently. Map was also an area of focus because it can provide an idea on the location, be it state or city.

DISCUSSION:

All the designs were rejected for the following reasons:

1. It does not provide a higher level of interactivity.
2. Plotting multiple values or huge amount of data on one bar makes the model look clumsy and less readability among viewers.
3. Plotting geocodes on map make the points indistinguishable because the points are too close to each other.
4. So, now I have further refined all the designs with an aim for better visualization, as discussed in Sheet 3 below.

Sheet 3

LAYOUT 2:

PANEL 1:

Considering the visualization for aircraft makers as a bar chart, as derived from design sheet 1, the second level of visualization for aircraft models have been planned to be a facet grid bar chart. So, for each maker, all the available models will be displayed at once on the screen, showing the number of accidents that has occurred during the range of year selected. The input value for the year range is designed as dropdown, each value representing a range of years from 1979 to 1987, 1988 to 1997 and so on till 2017. According to the years selected, the both the bar graphs would change depending on whether the aircraft makers/models were present during that time period.

PANEL 2:

The map is refined to represent a city wise dotted visualization as the location of accidents in USA. The input is same as that used in Panel 1. On selecting the specific dots, a bar chart with zip codes in the x-axis and number of accidents in the y-axis will be displayed for the selected years.

PANEL 3:

For primary reasons, I have thought of a simple bar chart where each bar will represent the primary reason. On selecting each bar, a line graph will be displayed which will give us a visualization for its respective secondary reasons against the number of accidents in y-axis. The input is the years in a dropdown list, as described above. The number of accidents will vary based on the time period chosen.

OPERATIONS:

One can select only one year range. For each panel, the second level of visualization is generated dynamically from the previous level. In Panel 1, the second level of bar chart is a facet grid bar chart that gets generated upon clicking the bars in the first level. Similar functionality is implemented for the other panels as well.

FOCUS:

Facet grid bar chart is interesting as it can convey huge information for multiple data at once. Interactive visualization within charts is very powerful and efficient.

DISCUSSION:

Some of the designs were rejected due to the following reasons:

1. We cannot select a wider range of years as input.
2. Facet grid bar chart for representing the aircraft models proved to be even worse for interactive visualization because the screen looks clumsy with too many data and the scales cannot be read properly as they are not clearly visible.
3. Plotting only the cities on map still looks cloddish.
4. Panel 3 does not visualize the year wise trend at all.

So, now I have further refined all the designs with an aim for better visualization, as discussed in Sheet 4 below.

Sheet 4

LAYOUT 3:

PANEL 1:

Continuing with the visualization of aircraft makers in a simple bar chart, I have changed the representation of their respective models will from the faced grid bar graph to bubble chart where each bubble denotes the aircraft models. The same is being shown in the legend as well. Here, the visualization has been against the years in x-axis and the number of accidents in y-axis that has been caused by those models. Here, I have introduced a slider input for selecting the range of years so that the user/ viewer has the flexibility of viewing the rate of accidents for a shorter or longer duration of time period. According to the years selected, the bubbles change their position along with the number of accidents.

PANEL 2:

The choropleth map is chosen to be the best form of visualization to represent the location of accidents state-wise in US. The input is same as that used in Panel 1. On selecting the specific range of years, only those locations (states) will be shown as dots on the map where accidents have occurred during those selected years. On clicking each state, their corresponding cities where accidents have occurred, will be displayed as a bar chart in the second level.

PANEL 3:

For primary reasons, I have finalized the radar chart for its visualization, depicting the percentage of accidents that have occurred due to each reason. On selecting a primary reason, an area graph will be displayed which will represent its respective secondary reasons in the legend against the number of accidents in y-axis and the years in x-axis. The input design for the years in slider input. The number of accidents will vary based on the time period chosen.

OPERATIONS:

One can select year range from slider input. In Panel 1, I have kept the visualization of aircraft maker as a bar graph. On clicking any of the maker, a second level of representation of aircraft models as a bubble chart used with legends (showing the model names) appears on the screen. In Panel 2, I have used choropleth map to represent the states (accident zones) in USA. On clicking any of the state, a bar chart appears showing the number of accidents in the cities for that particular state. In Panel 3, radar chart seems to be the most appropriate visualization for depicting primary reasons and their respective percentages on the number of accidents. Clicking on each of the primary reason would give us an area chart showing its respective secondary reasons accountable for the number of accidents during the selected time period.

FOCUS:

Bubble plot, radar chart and area map are efficient ways of plotting data. Also, choropleth map can assume the states of US as the accident regions.

DISCUSSION:

Some of the designs were rejected due to the following reasons:

1. Bubble chart looks uncoordinated as there are many models for one maker.
2. Only radar plot in Panel 3 provides the percentage of primary reasons that contribute to the accidents, though it does not show year wise trends.

3. Now, some of the designs that have been prepared are rejected and some are finalized for the final visualization, as discussed in Sheet 5 below.

Sheet 5

FINAL LAYOUT:

A Shiny Dashboard is prepared to display the final visualization designs for all the 3 panels in 3 tab items, namely 'Accidents on Aircraft Makers', 'Accident Locations' and 'Reasons for the Accidents'. The input for all the 3 panels is the range of years from 1948 from 2017 on Slider Range Shiny Widget. Let's discuss the final layout designs below:

PANEL 1:

The first panel or the tab item is 'Accidents on Aircraft Makers'. I have introduced a slider input for selecting the range of years from 1948 to 2017. I have come up with the visualization design of aircraft makers in a simple bar chart and the representation of their respective models through a heatmap. The bar chart shows the number of accidents in y-axis and the aircraft makers in x-axis with aircraft maker names as legend. On clicking on any of the bars, the heatmap is generated which has all the years (chosen in the slider range) on x-axis and the number of accidents as the y-axis label. The legend shows the number of accidents within a specified range with color for the respective models. According to the years and the aircraft maker selected, the heatmap changes accordingly and dynamically.

PANEL 2:

The second panel or the tab item is 'Accident Locations'. The best form of visualization to represent the location of accidents statewise in US is the choropleth map. The input is same as that used in Panel 1. On selecting the specific range of years, only those locations (states) will be shown as boxes on the map where accidents have occurred during those selected years. The colour of the states vary according to the proportion of the accidents (represented in the legend) that had occurred in those locations. On clicking each state, their corresponding cities (on x-axis) where accidents have occurred and the number of accidents (on y-axis) in those cities, will be displayed as a colored bar chart. The name of the cities separated with different colors appear on the legend of the bar chart.

PANEL 3:

The third panel or the tab item is 'Reasons for the Accidents'. For primary reasons, there have been 2 visualization designs. One is the radar chart which depicts the percentage of accidents that have occurred due to each reason. This chart tries to combine the proportion of accidents for different reason into one chart. The other is the line chart which shows the number of accidents that have occurred during the years selected. The lines are the primary reasons which are depicted in various colors in the legend. On selecting a primary reason from the line chart, an area graph will be displayed which will represent its respective secondary reasons in the legend against the number of accidents in y-axis and the years in x-axis. The input design for the years in slider input. The number of accidents will vary based on the time period chosen. The area graph has colored legend for different reasons.

DETAIL:

The platform used to build the shiny app for demonstrating the visualization designs is R, version 3.5.1. The libraries used are shiny, shinydashboard and plotly. The number of dashboard panels prepared is 3. In all the panels, a slider range has been designed for selecting the range of years but the other plots are different in each panel. In the first panel, a bar chart and a heat map has been plotted. In the second panel, a choropleth map and a bar chat has been plotted. In the third panel, there are 3 plots – line chart, radar chart and area graph.

OPERATIONS:

1. In panel 1, for the default range of years from 1948 to 2017, a bar chart is displayed representing the number of accidents caused for all the aircraft makers. An user can change the range of the years by sliding the bar from both the ends and the bar chart will change accordingly. After that, upon clicking on each bar, a heat map is generated representing the number of accidents that has been caused for the aircraft models of the chosen aircraft maker over the period of time.
2. In panel 2, a choropleth map of USA is generated with the colour density proportion to the number of accidents that occurred in the states of US. Upon clicking on the states, citywise accident details are displayed through a bar chart dynamically.
3. In panel 3, a line chart showing the yearwise trend of accidents due to primary reasons and a radar chart visualizing the proportion of the primary reasons accountable for the aviation accidents are generated for a default range of year as soon as the third tab 'Reasons for the Accidents' is chosen. Upon clicking any of the primary reasons from the line chart, an area chart is displayed showing the number of accidents yearwise for the respective secondary reasons.

ZOOM/FOCUS:

1. In panel 1, the focus is on the heat map which gets generated from the bar chart.
2. In panel 2, the proportion of the accidents in choropleth map according to the colour density is focused.
3. In panel 3, the area chart is focused to display information on secondary reasons, the number of accidents and the chosen years.

IMPLEMENTATION

Description of the Tool

Tool Name	Programming Language	Version
RStudio	R	R version 3.5.1
Libraries: 1. Shiny: Does the structure for the interactive visualization. 2. Shinydashboard: A special feature in Shiny for creating dashboards. 3. Plotly: An very powerful library for interactive visualization		

Implementation Reasons for Technology

We chose R as our programming language because of the following reasons –

1. It is open source, so easily available.
2. R can do reproducible research and it is very in R to handle errors.
3. R provides an extremely easy way to do data wrangling, so any reformatting of data could be easily done in R.
4. Through R, we can do super advanced visualization very easily. Because of its easy implementation, R is my choice for visualization.

5. R can easily extend itself to specific needs and into specific platform.

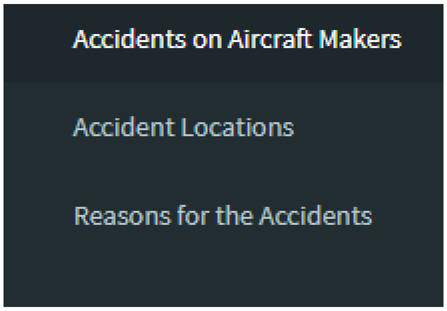
We used 'shiny' package because we can do interactive visualization through it, our primary aim for this project. 'shinydashboard' is a special library in R to create dashboards, which are very useful in creating lucrative interactive visualization application. Lastly, we used 'plotly' because through plotly we can do better interactive visualizations within charts. For example, if we need to create another detailed level chart from a relatively simpler one by adding some 'mouse hover' or 'mouse click' events, we can do it through plotly very efficiently.

Implementation Reasons for Visualizations

The reasons for and the way of implementation of different selected have been discussed in the design phase in detail. The pictorial illustration is given into the user guide.

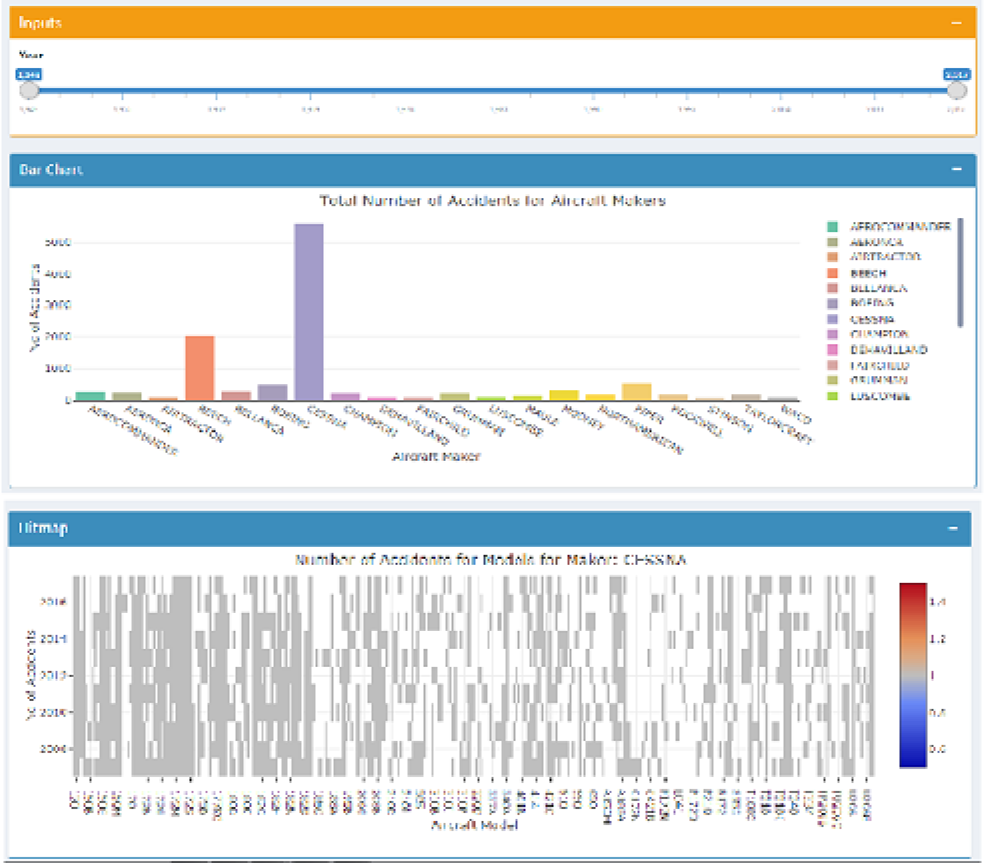
USER GUIDE

1. At first, the shiny application needs to be downloaded from the Moodle, unzipped saved somewhere in the desktop computer.
2. Within the unzipped folder we will have 3 aforesaid input files and one shiny application. We should keep them within the same folder.
3. The shiny app should be opened from the same unzipped folder and opened in RStudio. The working directory in RStudio should also be set to the same unzipped folder location.
4. The shiny app is now ready to run and it should be run from RStudio by clicking green 'Run App' button in RStudio.
5. Once the app starts running, it will open a shiny dashboard panel with three options like below. The 'Accidents on Aircraft Makers' option guides us to the panel for visualizing the accident data linked to different aircraft makers and their respective models. The 'Accident Locations' option guides us to the panel for visualizing the accident data linked to different locations of accident. 'Reasons for the Accidents' option guides us to the panel for visualizing the accident data linked to different reasons for accidents.

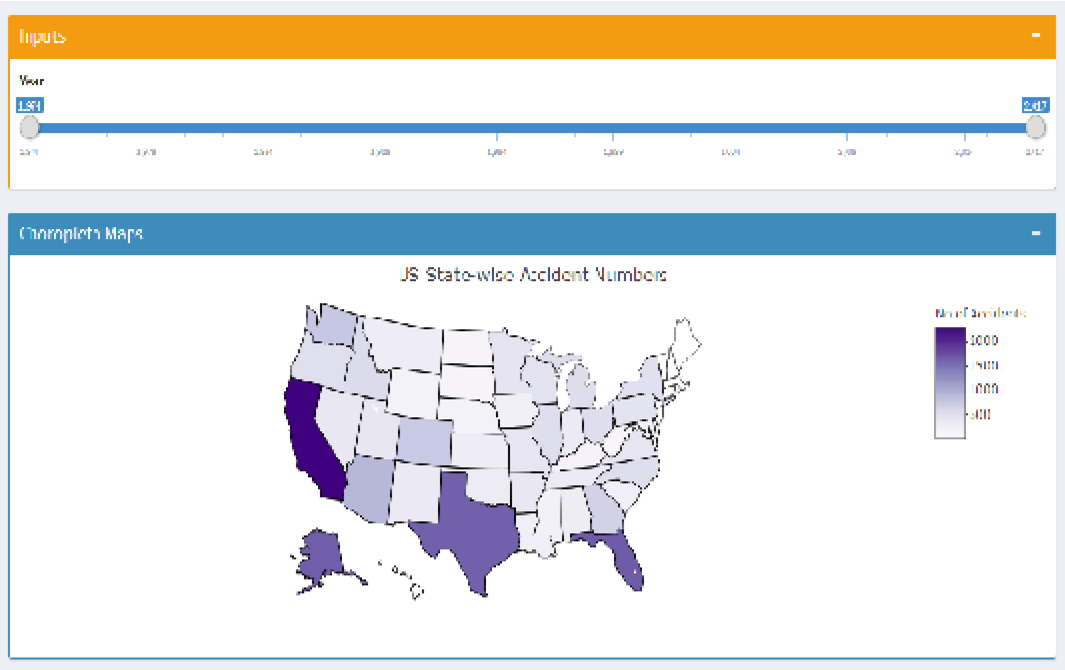


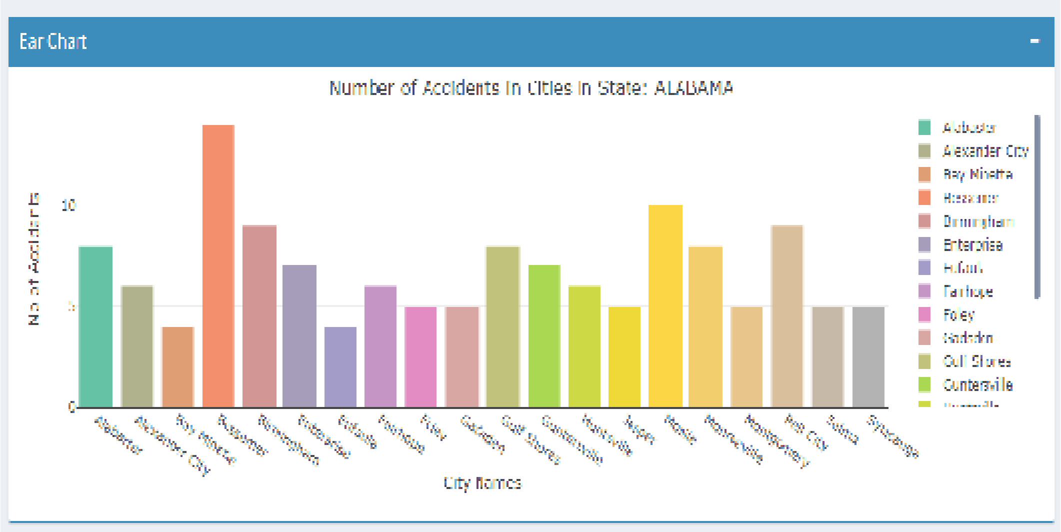
6. If we click on to 'Accidents on Aircraft Makers', we go to the below tab where slider input for a year range and a bar graph appears, showing top 20 aircraft makers involved in accidents against the number of accidents for the aircrafts from these makers. We can select the preferable year range from this slider input and accordingly the bar graph updates itself. From the legend of the bar graph we can select or deselect any maker. If we hover on each bar, the details are shown for that particular bar. If we now click on a bar for a particular maker, a heatmap appears below showing the year-wise accident trends of all models for that maker with that specific year range. For our convenience, we only show the trend for the last 10 years for the specified year range. Hovering on

the heatmap shows the number of accidents occurred for that particular model for the given year on Y axis.

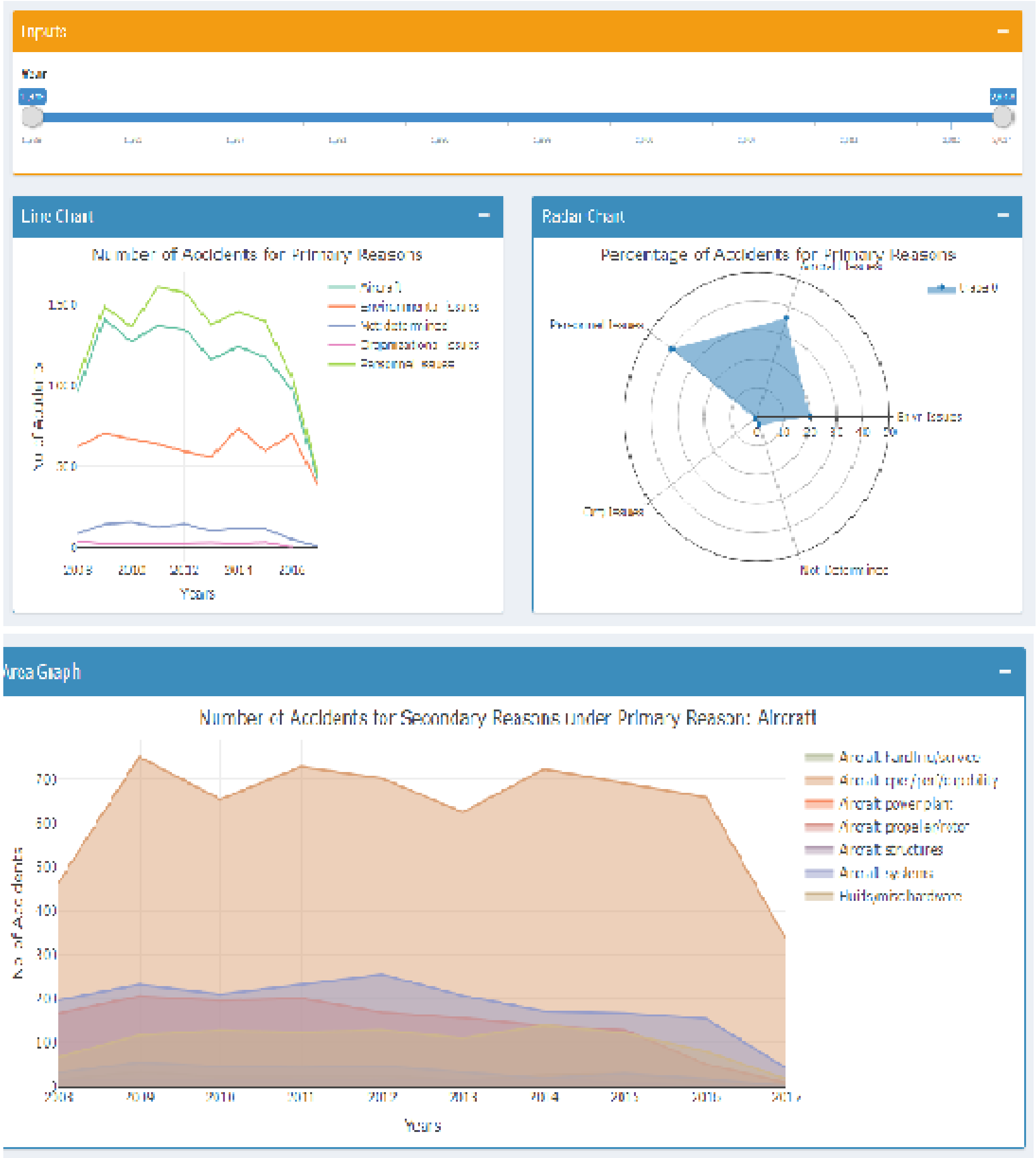


7. If we click on to ‘Accident Locations’, we move to the below tab where slider input for a year range and a cloropleth map appears, showing different US states (As the location data only for US is available here) as accident prone zones with the respective number of accidents occurred during that time period. We can select the preferable year range from this slider input and accordingly the map updates itself. If we hover mouse pointer on each state, respective accident details appears. The color density in line with the legend on the right hand side shows how worse the state is as an accident prone zone. If we click each state, a bar graph appears below showing the cities with their respective number of accident within the given time period. Hovering our mouse on each bar of the bar graph shows the details of the accidents. We can also select or deselect the cities from the legend of the bar graph.





8. If we click on to ‘Reasons for the Accidents’, we move to the below tab where slider input (which we can adjust as we want) for a year range and two charts appear. The first one is a line chart showing the year-wise trends in the number of accidents for five primary reasons and second one is the proportion of the accidents for the same five reasons within a radar chart. From the line graph, we can select or deselect our preferred reasons from the legend. We can select the preferable year range from this slider input and accordingly these two charts update themselves. If we hover mouse pointer on the lines of the line graph, we get the respective details of the accidents. If we click in each line, an area map appears below showing the year-wise accident details of the secondary reasons under those primary reasons. If we hover mouse pointer on the area graph, respective accident details comes up. We can also choose our preferred secondary reason from the legend.



CONCLUSION

The achievement of this project is illustrated from the above said story that gives us a tour with the entire visualization. From the first panel, we understand how and to what extent different models of several aircraft makers are involved into accidents over the years. From this, we can generate insight on which model series from which aircraft maker are more involved into accidents and accordingly the authority can make decision on whether to impose a ban or more routine inspections on these models. Accordingly, from the second panel we saw different accident prone ones of U.S.A in terms of states and cities. By checking the trends in accidents in different city and states, the aviation authority can commence inspections in those regions to get even more nuanced insights of the reasons why those locations are notorious for air crashes, accordingly the flights can be diverted over the relatively safer ones (depending on the number of accidents again) to avoid any future possibilities. From the third panel, we know what are reason mostly associated with the aviation accidents and by looking into those reasons, the authority can do something to prevent the reasons to occur, allowing a safer journey of the passengers.

From this project we learned several aspects of visualization. Firstly, we learned interactive visualizations using R shiny package using plotly. We now know how to generate plots interactively using different mouse events (hover, click). We came to know about five design sheet methodology in this project to design visualizations in a systematic way. This technique helps a developer to be on the right track, lets the designer compare different designs for the same visualizations to choose the best one. However, I think there are certain areas that could have been better achieved with this visualization project. We could have tried to find the interactions among the main three factors for our visualization here and could come with interactive charts showing the inter-relations among them. For example, we could have tried looking how different aircraft models are associated with accident reasons or how different reasons of air accidents change with the location. But overall, this project will help an aviation expert to chalk out the future strategies in the aviation industry.

BIBLIOGRAPHY

Clark, C. (2017). *NTSB Aviation Accidents Through 2017*. Retrieved 2019, from www.kaggle.com: https://www.kaggle.com/cwclark/ntsb-aviation-accidents-through-2017

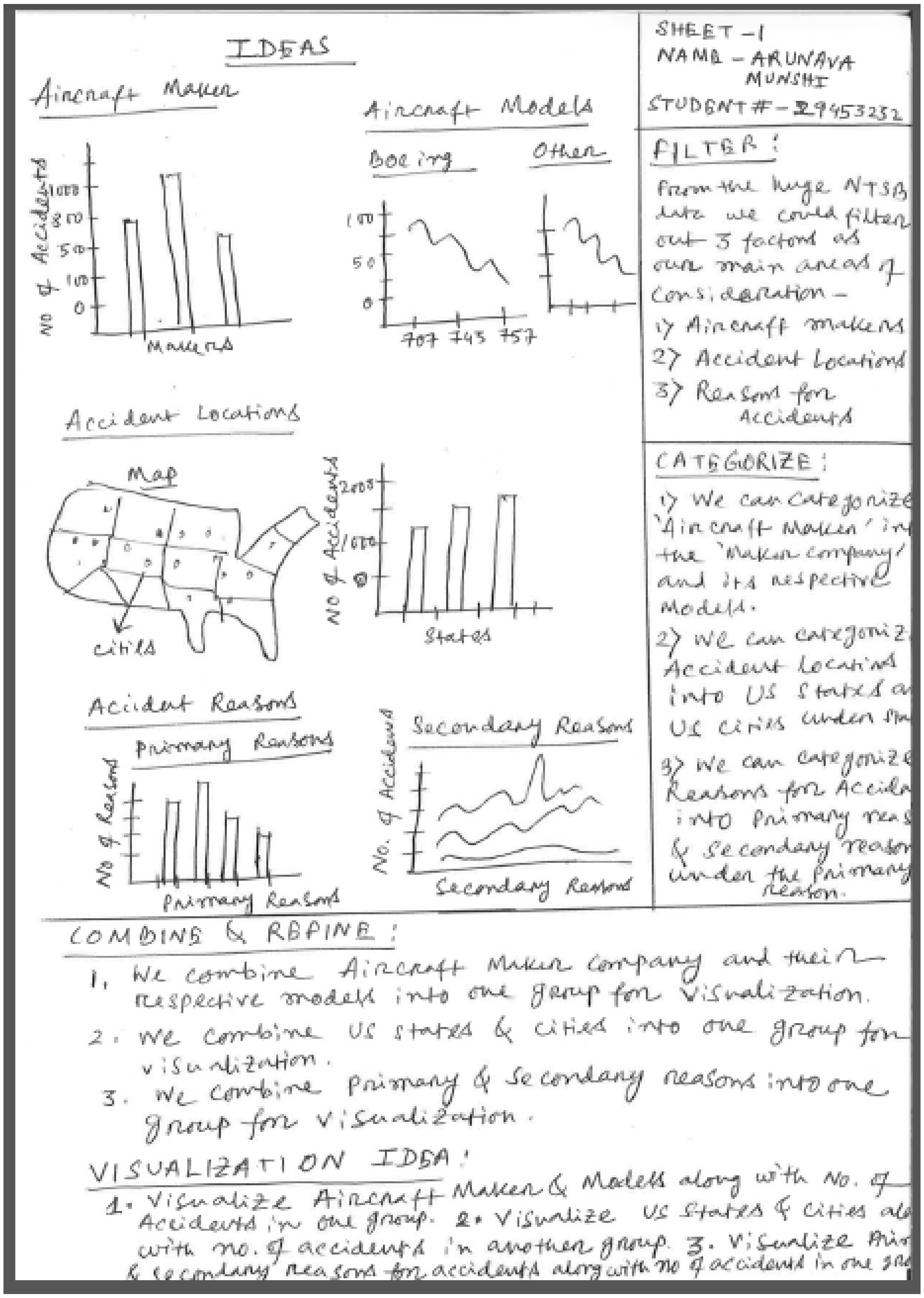
Telegraph. (2017). *How many planes are there in the world right now?* Retrieved 2019, from Telegraph.co.uk: https://www.telegraph.co.uk/travel/travel-truths/how-many-planes-are-there-in-the-world/

www.aviationcv.com. (2016, Jan 25). *Deadliest Accident in Aviation History*. Retrieved Apr 28, 2019, from www.aviationcv.com: https://www.aviationcv.com/aviation-blog/2016/deadliest-accident-in-aviation-history

www.iata.org. (2019). *International Air Transport Asociation*. Retrieved 2019, from www.iata.org.

APPENDIX

The five design sheets are attached below –



SHEET – 1
NAME – ARUNAVA MUNSHI
STUDENT # – 29453232

FILTER:

From the huge NTSB data we could filter out 3 factors as our main areas of consideration –

1) Aircraft makers
2) Accident Locations
3) Reasons for Accidents

CATEGORIZE:

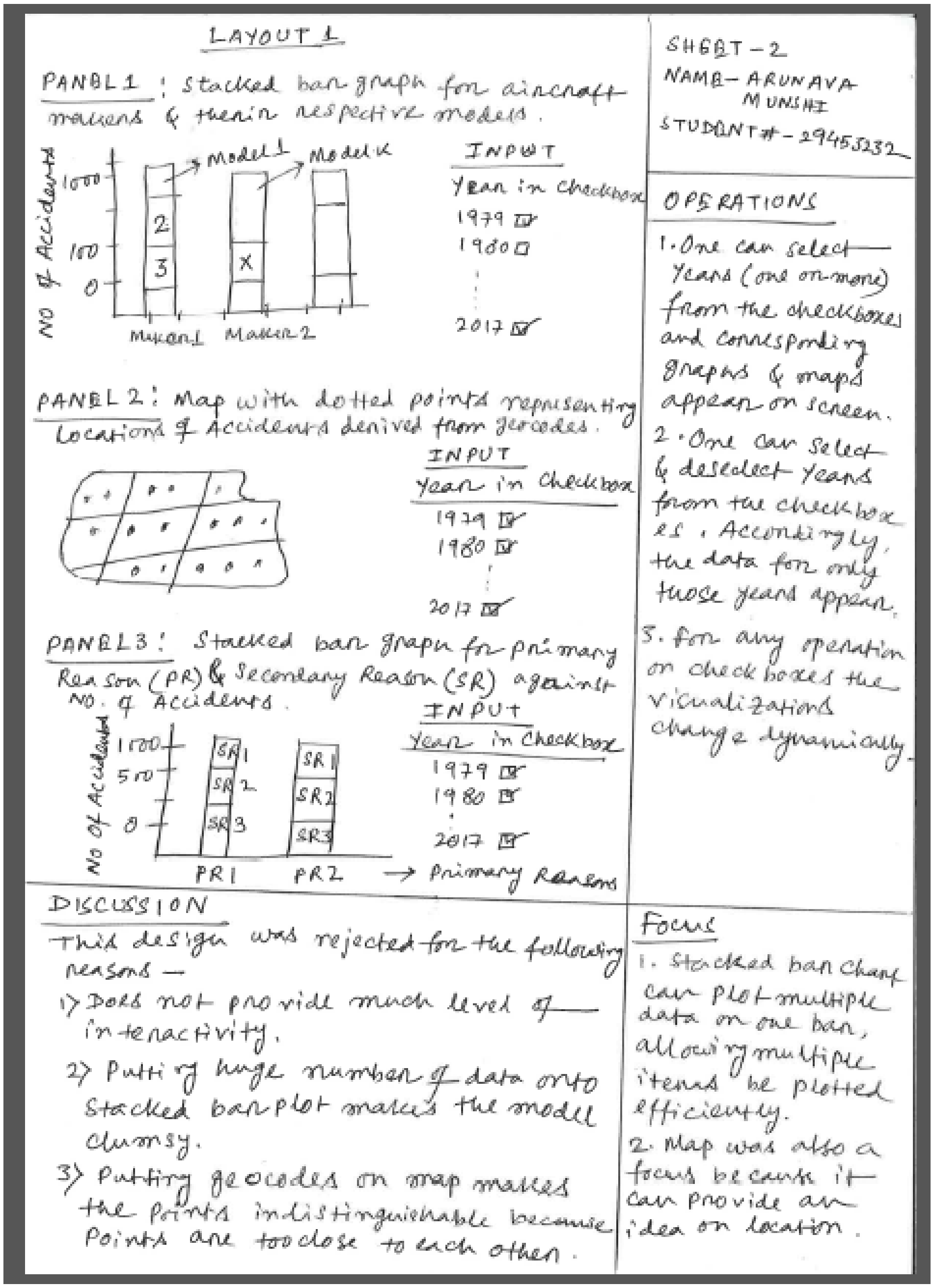
1) We can categorize 'Aircraft Maker' in the 'Maker company' and its respective Models.
2) We can categorize Accident Locations into US States a US cities under the
3) We can categorize Reasons for Accide into Primary reas & secondary reason under the Primary reason.

COMBINE & REFINE:

1. We combine Aircraft Maker company and their respective models into one group for visualization.
2. We combine US states & cities into one group for visualization.
3. We combine Primary & Secondary reasons into one group for visualization.

VISUALIZATION IDEA:

1. Visualize Aircraft Maker & Models along with No. of Accidents in one group. 2. Visualize US States & cities al with no. of accidents in another group. 3. Visualize Prir & secondary reasons for accidents along with no of accidents in one gr



LAYOUT 1

SHEET – 2
NAME – ARUNAVA MUNSHI
STUDENT# – 29453232

PANEL 1 : Stacked bar graph for aircraft makers & their respective models.

No of Accidents
1000
100
0
2
3
X
Model 1
Model K
Maker 1
Maker 2

INPUT

Year in Checkbox
1979 ☑
1980 □
⋮
2017 ☑

PANEL 2 : Map with dotted points representing locations of Accidents derived from geocodes.

INPUT

Year in Checkbox
1979 ☑
1980 ☑
⋮
2017 ☑

PANEL 3 : Stacked bar graph for primary Reason (PR) & Secondary Reason (SR) against No. of Accidents.

INPUT

Year in Checkbox
1979 ☑
1980 ☑
⋮
2017 ☑

OPERATIONS

1. One can select Years (one or more) from the checkboxes and corresponding graphs & maps appear on screen.
2. One can select & deselect Years from the checkboxes. Accordingly, the data for only those years appear.
3. For any operation on check boxes the visualizations change dynamically.

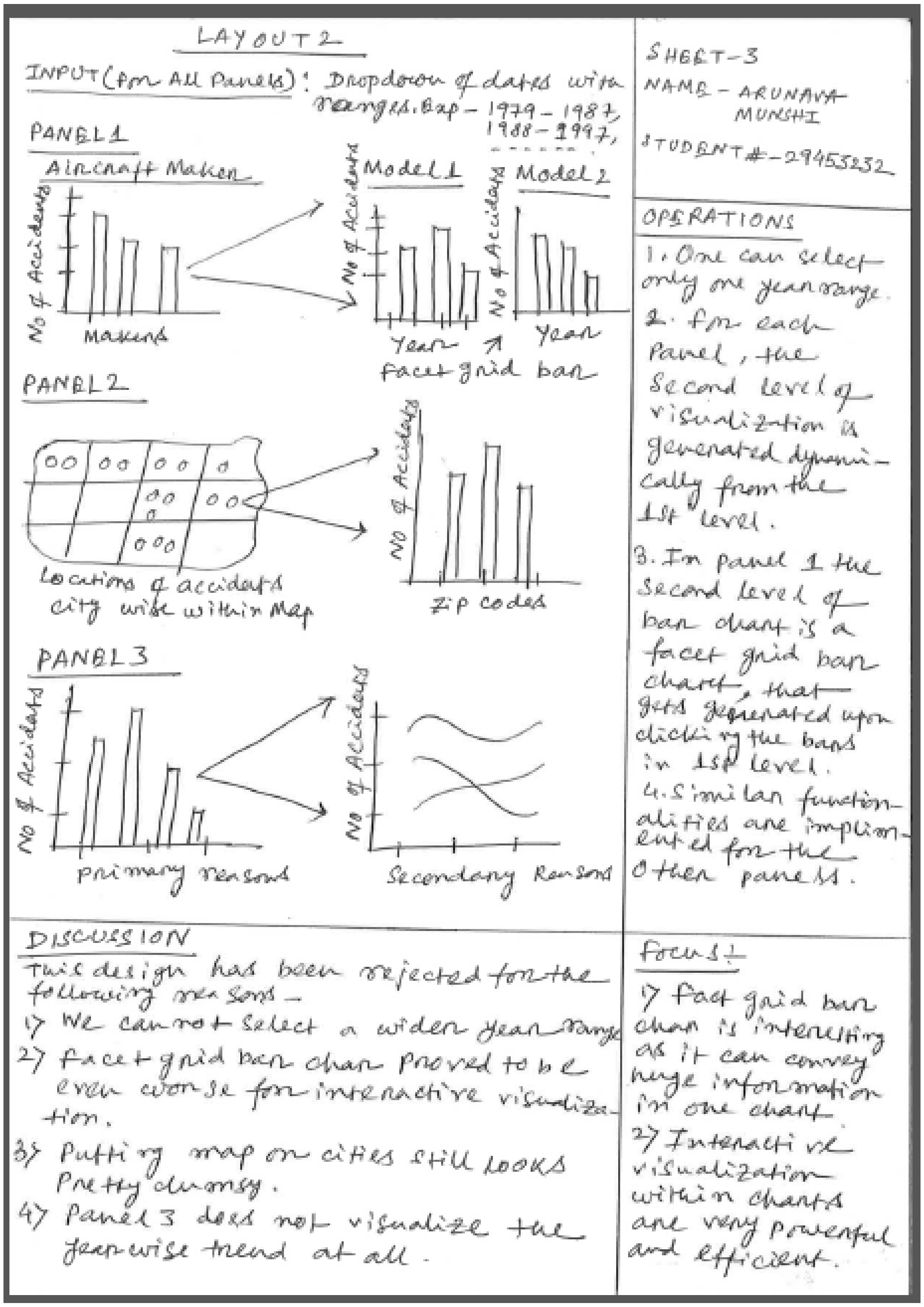
DISCUSSION

This design was rejected for the following reasons –

1) Does not provide much level of interactivity.
2) Putting huge number of data onto stacked bar plot makes the model clumsy.
3) Putting geocodes on map makes the points indistinguishable because points are too close to each other.

Focus

1. Stacked bar chart can plot multiple data on one bar, allowing multiple items be plotted efficiently.
2. Map was also a focus because it can provide an idea on location.



LAYOUT 2

INPUT (for All Panels): Dropdown of dates with ranges. Exp - 1979 - 1987, 1988 - 1997, ------

SHEET - 3
NAME - ARUNAVA MUNSHI
STUDENT # - 29453232

OPERATIONS

1. One can select only one year range.
2. for each Panel, the Second level of visualization is generated dynamically from the 1st level.
3. In panel 1 the Second level of bar chart is a facet grid bar chart, that gets generated upon clicking the bars in 1st level.
4. Similar functionalities are implemented for the other panels.

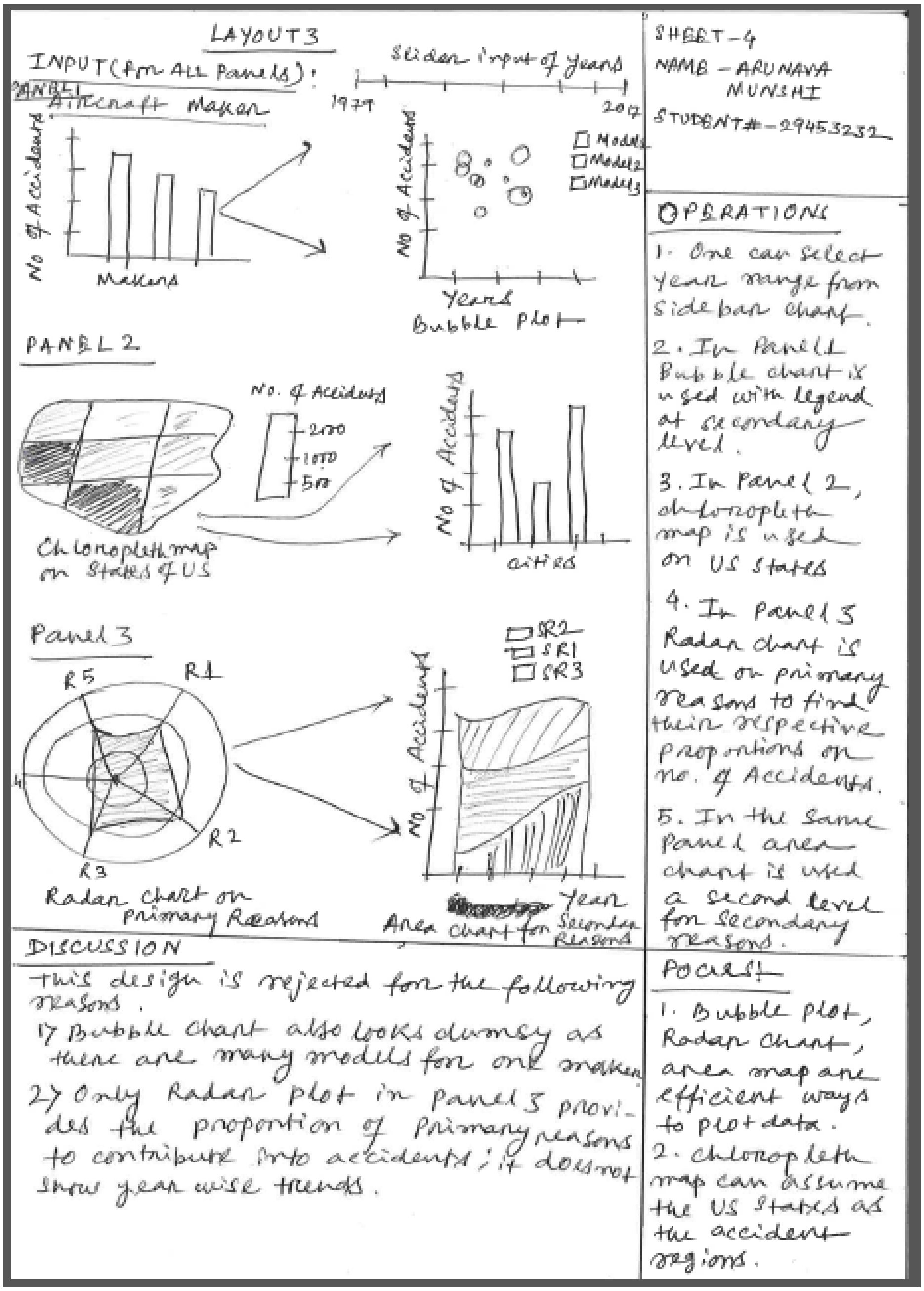
DISCUSSION

This design has been rejected for the following reasons -

1) We can not select a wider year range
2) facet grid bar chart proved to be even worse for interactive visualization.
3) Putting map on cities still looks pretty clumsy.
4) Panel 3 does not visualize the yearwise trend at all.

Focus :-

1) Facet grid bar chart is interesting as it can convey huge information in one chart
2) Interactive visualization within charts are very powerful and efficient.



SHEET-4

NAME - ARUNAVA MUNSHI

STUDENT#-29453232

OPERATIONS

1. One can select year range from sidebar chart.

2. In Panel1 Bubble chart is used with legend at secondary level.

3. In Panel 2, chloropleth map is used on US states

4. In Panel 3 Radar chart is used on primary reasons to find their respective proportions on no. of Accidents.

5. In the same Panel area chart is used a second level for secondary reasons.

FOCUS!

1. Bubble plot, Radar chart, area map are efficient ways to plot data.

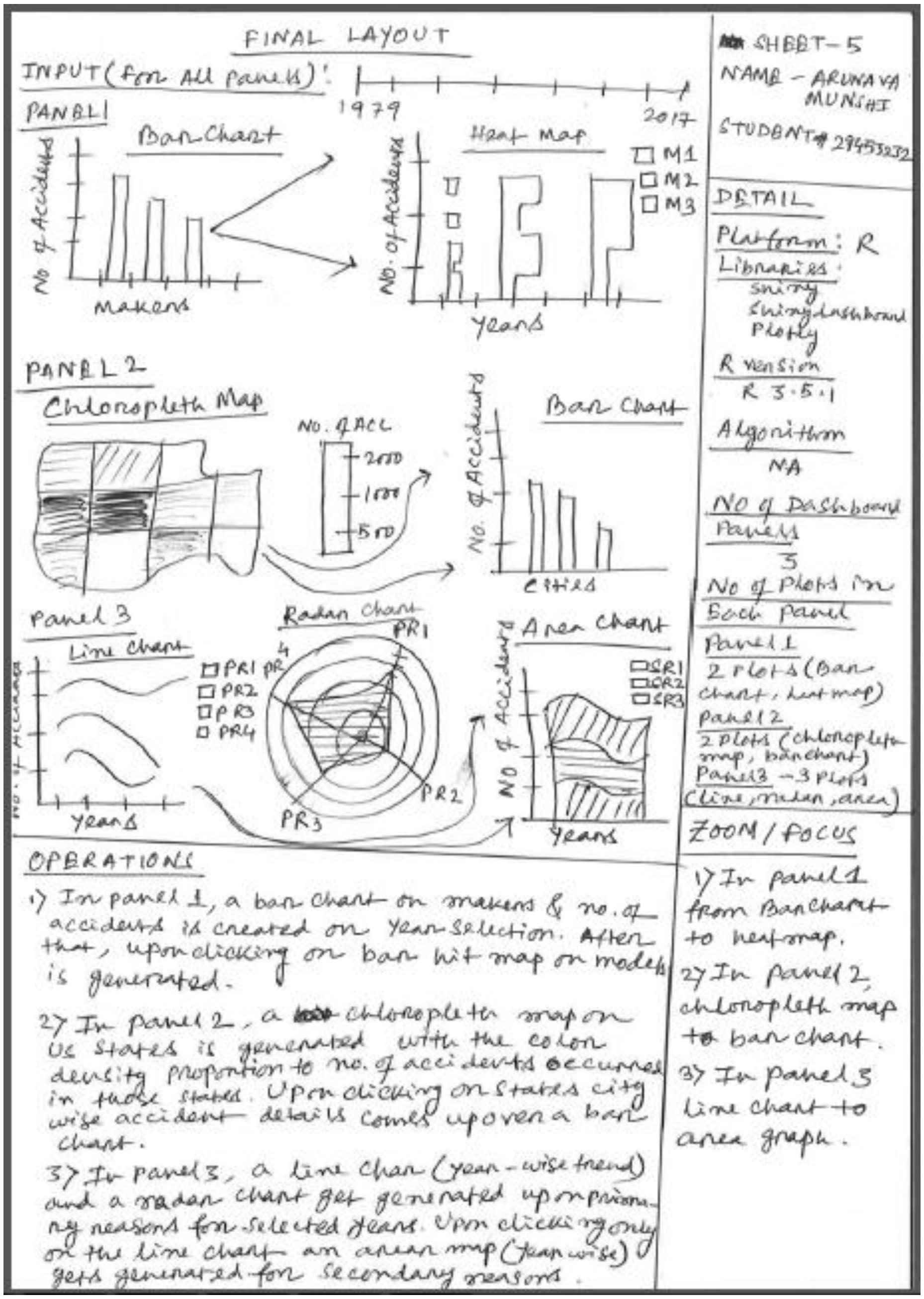
2. chloropleth map can assume the US States as the accident regions.

DISCUSSION

this design is rejected for the following reasons.

1) Bubble chart also looks clumsy as there are many models for one maker

2) Only Radar plot in panel 3 provides the proportion of Primary reasons to contribute into accidents; it does not show year wise trends.



FINAL LAYOUT

INPUT (for All panels):

1979
2017
PANEL1
Bar Chart
No of Accidents
Makers
Heat Map
No. of Accidents
Years
□ M1
□ M2
□ M3
PANEL 2
Chloropleth Map
NO. of ACC
2000
1000
500
Bar Chart
No. of Accidents
Cities
Panel 3
Line Chart
NO. of Accidents
Years
□ PR1
□ PR2
□ PR3
□ PR4
Radar Chart
PR1
PR4
PR2
PR3
Area Chart
No of Accidents
□ SR1
□ SR2
□ SR3
Years

SHEET-5

NAME - ARUNAVA MUNSHI

STUDENT# 29453232

DETAIL

Platform: R

Libraries:
shiny
shinydashboard
Plotly

R version
R 3.5.1

Algorithm
NA

No of Dashboard Panels
3

No of Plots in Each panel

Panel 1
2 plots (Bar chart, heatmap)

Panel 2
2 plots (chloropleth map, barchart)

Panel 3 - 3 plots (line, radar, area)

ZOOM / FOCUS

1) In panel 1 from Barchart to heatmap.

2) In panel 2, chloropleth map to bar chart.

3) In panel 3 line chart to area graph.

OPERATIONS

1) In panel 1, a bar chart on makers & no. of accidents is created on Year-Selection. After that, upon clicking on bar hit map on models is generated.

2) In panel 2, a chloropleth map on US States is generated with the color density proportion to no. of accidents occurred in those states. Upon clicking on states city wise accident details comes up over a bar chart.

3) In panel 3, a line char (year-wise trend) and a radar chart get generated upon priomary reasons for selected years. Upon clicking only on the line chart an arear map (year wise) gets generated for secondary reasons.