Dasom Eom &
Fitsum Seyoum
Dr. Alex Endert
CS 4460
November 30, 2018
Project 5

# Description of Project 5

For project 5, we chose the dataset of Aircraft incidents. The data in the Aircraft Incidents file contains major and minor incidents of Aircraft from all over the world. The period of the incidents is from 1995 to 2016. Here are some of the analytic tasks which our visualization supports:

1) Overview first, zoom and then details on demand: At first glance, a user of our visualization can see the overview distribution of the incidents in the map of the world. This can give the user the spread and depth of the incidents in different countries. The representation allows the user to zoom in on items of interest into a specific incident. The user can then hover over the dots to read the detail information of the prefered incident.

2) Retrieve values: Our visualization allows the user to answer questions like what is the number of Aircraft incidents in a given country? How many incidents does a given aerospace company had during the period?

3) Filter Extremum: Which year had the highest number of incidents during the given period? Which airline has the highest number of incidents from 1995 to 2016? Here we were not able to calculate the percentage of the incidents for each aerospace company, to compare one another, because we were not able to find the total number of aircraft on the sky during the incident time for each aerospace company. As a result, we only compare the total number of incidents from the given CSV file.

4) Characterize incident distribution over time: Again our visualization allows the user to see the incident distributions from 1995 to 2016. Using the bar chart which we used to display the incidents distribution, the user can learn how the incidents fluctuate over time. Besides, we can see that the highest number of incidents were in 2010 and the lower number of incidents were in 2016.

During the design process we wanted to challenge the visualization with geojson and wanted to see what would happen with this dataset. So firstly, we visualized the dataset on the world map using Topojson. Unfortunately, some of the tuples had null values for coordinates of longitude and latitude, and those data could not be displayed on the map. Through the visualization on the map, we were able to see that many points were drawn on the United States (Figure 1). We made it possible for the user to zoom in, out, or drag the map to see it better

(Figure 2). Also, when the user hovers over the data of interest, the color of the datum is changed to red, a tooltip pops up and shows detailed information about the event (Figure 3).

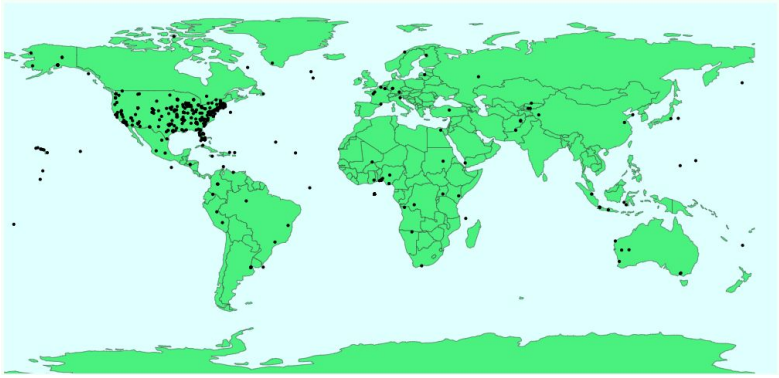

*Figure 1. The visualization of Aircraft Incidents on a map using Topojson*

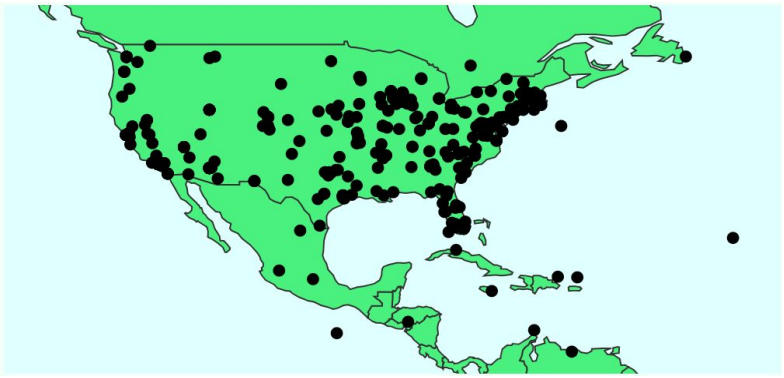

*Figure 2. The magnified scene of the world map visualization focused on United States*


*Figure 3. Hovering over a circle marks it red and shows the detailed information*

Information visualization through a world map was effective in helping users to grasp the data with spatial sensibility, but as mentioned earlier, in a given dataset with lot of data not appearing on the map, we felt the need to visualize them in a graph with quantitative measures. The dataset had attribute of "country" and "state", but a lot of countries do not have value for state attribute, so we visualized the number and percentages of aircraft incidents by country on a bar chart (Figure 4). Since we had so many countries to represent on the chart and we wanted to present all of them on a screen with the world map, we drew vertical bar chart with each bar

having a short width length. Also, our bar chart summarizes each data set in a visual form. Surprisingly, the bar chart showed that most plane accidents were concentrated in the US.

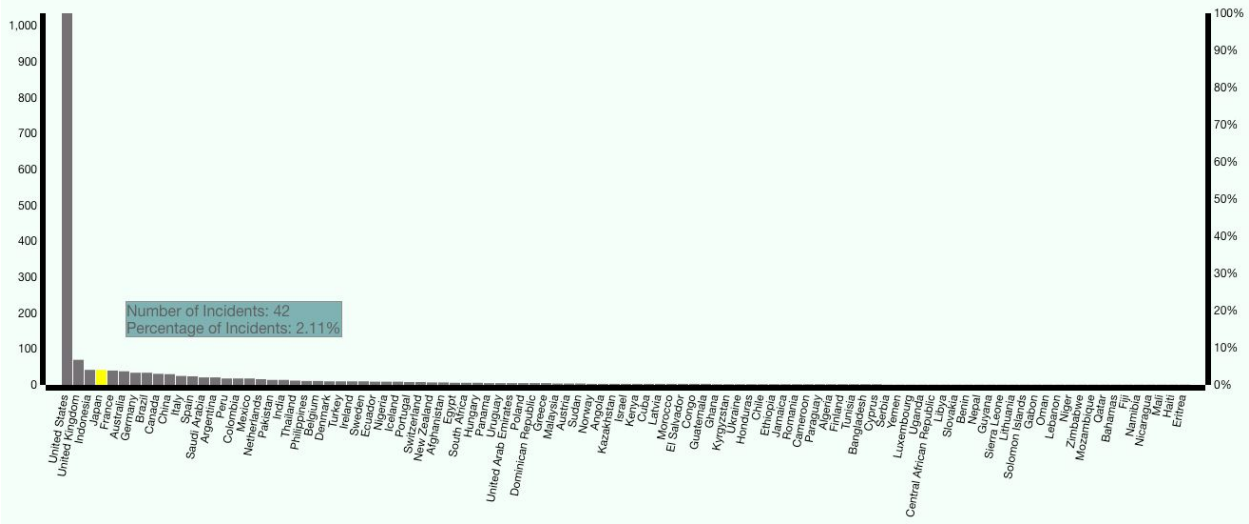


*Figure 4. The Number(Percentage) of Aircraft Incidents by Country Bar chart*

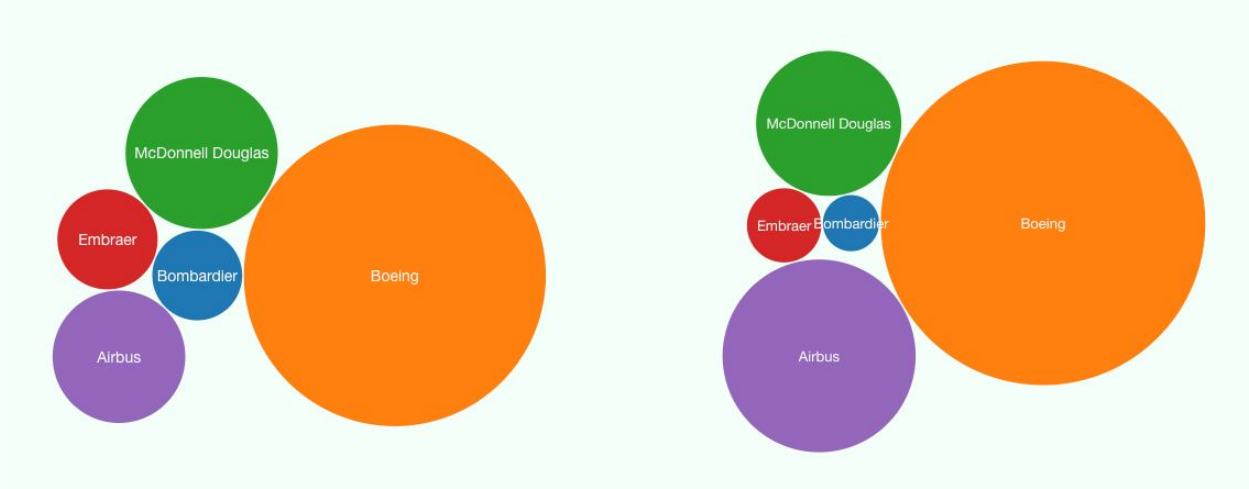


*Figure 5. the number of Incidents by Make VS the number of Fatal Injuries by Make*

The next thing we wanted to plot was to see if there is any correlation between the make of the aircraft and the number of incidents or the fatality. There were only five different aircraft makes in the given dataset, and we wanted to plot it in bubble chart so that users can see the general trends between the incident and aircraft make (Figure 5) next to each other. This bubble chart does not take into account how many planes from each make are in use on the market, and how many of them were involved with accidents. Thus, we should not make an assumption that if the bubble is big then the company makes poor aircrafts.

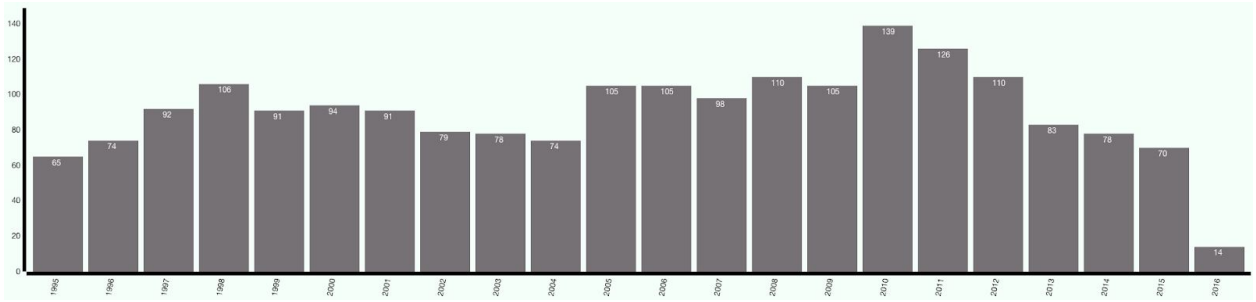


*Figure 6. The bar chart of the number of Incidents by year in time history*

Lastly, we wanted to see if the development of flight technology had improved the flight skills so that there were fewer airplane accidents recently, or if there were more airplanes crashes as traveling by plane were becoming more common as time went on. Even though we originally expected to see gradual ascending trend or descending trend over time on the bar chart, but the bar chart showed us that there was no one-sided trend. The largest number of plane accidents occurred in 2010, and after that, the number of accidents decreased. Especially, the number of incidents greatly reduced from 2012 to 2013, there might be some reasons on this sudden change. Since the dataset had only had data from 1995 to mid-2016, so the 2016 bar graph was not exactly visible. Also, we marked the number of incidents in each year on top of each bar.