CS 4460 - P5
My Ngoc Nguyen

Nhu Nguyen

Aircraft Dataset

For this assignment our team decided to use the aircraft dataset. The aircraft dataset provided many different attributes for the list of aircraft incidents, but our focus were on finding the causation and correlation between location, board of flight and airplane planes and the severity of the incidents. We wanted to see what kind of impact each attribute had on the likelihood of the incident.

We used a map, a pie chart, a tree map to analyze the impact of each attribute on the incidents. To analyze the location, we used the map to plot against the longitude and latitude of where the planes crashed. This helped us to quickly recognize the clusters where aircraft incidents are most likely to happen. This graph also helps us to identify if geographical features are one of the leading causes of aviation accidents. We are curious to see the risks involve in being flown over one specific region over another. Even though we think aircraft incidents are caused by human error, we don't want to disregard the locality factor. The dots represent each incident are color-coded to show the severity of each incident, we picked red for the color of fatal accident because this is a bold color and it grabs attention quickly. Red is an intense, bold color and it is also often associated with danger. We were able to utilize the principles for design that we learned in lecture. Choosing red makes "target detection" easier for pre-attentive processing, as we know that viewers are more likely to be interested in the fatal plane incidents. We also included a description of each color, so viewer can quickly determine the severity of the incident at each location. We used the "hue and shape" principle so viewer can quickly distinguish the severity of each incidents. We chose to only show the incidents that took place in the US, because based on our research, US is the country with highest number of fatal civil airliner accidents, so it is more effectively to show the US map instead of the world map. Since we can plot the longitude, latitude easily and still keep the details at a good scale. Viewer can also click on each dot to see more information about the accident such as location, airport name, date and airplane model. By represented the incidents by dots, we used Gestalt principles, specifically similarity and connected because viewer can quickly notice the cluster of accidents and also the severity of the accidents that happened in those clusters, which is one of the main task taxonomies. This can help viewers to answer two analytical questions: where the aircraft incidents most likely to take place and what the severity at those places. One interesting observation is that all fatal accidents happen in the North East and some part of the South, and no fatal accident has occurred in the West. But since we are not certain that the dataset has the record of all aircraft incidents ever happened and there needs to be an in-depth analysis on the geographical factor, the observation is not a strong claim to link the causation to aircraft incidents.

To get a more detailed breakdown of the total number of fatal, serious and uninjured passengers of each state and each aircraft make and model, we chose to use a pie chart to easily show the percentage, amount of the number of passengers involved in the accidents. We used a dropdown menu, so viewer can select the state that they are interested in. The total number of fatal, serious

and uninjured passengers are filtered by states. The second pie chart shows in break-down percentage filtered by the aircraft model of the selected state. This can help viewer to quickly identify the aircraft model. Since we want to keep the locality is the focus, the aircraft make is filtered by state and not just by their own. Because it is unlikely that a specific aircraft model is the cause of the accidents as each was designed and tested by professional standard and a safety and engine check are also required before take-off. If we showed the aircraft incidents by aircraft make and model without taking in consideration of the locality, it could be misleading if one aircraft model has a higher percentage than others and cause viewers to make a quick assumption about its causation. It could be that the model is the most popular and the most used. This visualization helps viewers to achieve tasks such as retrieve value, filter, find extremum and characterize distribution.

We used a treemap to visualization aircraft incidents by aircraft damage, board phase of flight, make and model. We were curious to see how much of an impact the board phase of flight in terms of correlation to the accidents. The phase that requires most of the pilot's attention are takeoff and landing, during the flight it is now common that aircraft are flown by autopilot. So even if the dataset doesn't include pilot errors to be one of the most common causations, our research has shown that more than half of aircraft incidents happened are linked to pilot error. So our team used the board phase of flight as a tool to help us answers the analytical questions about pilot errors.

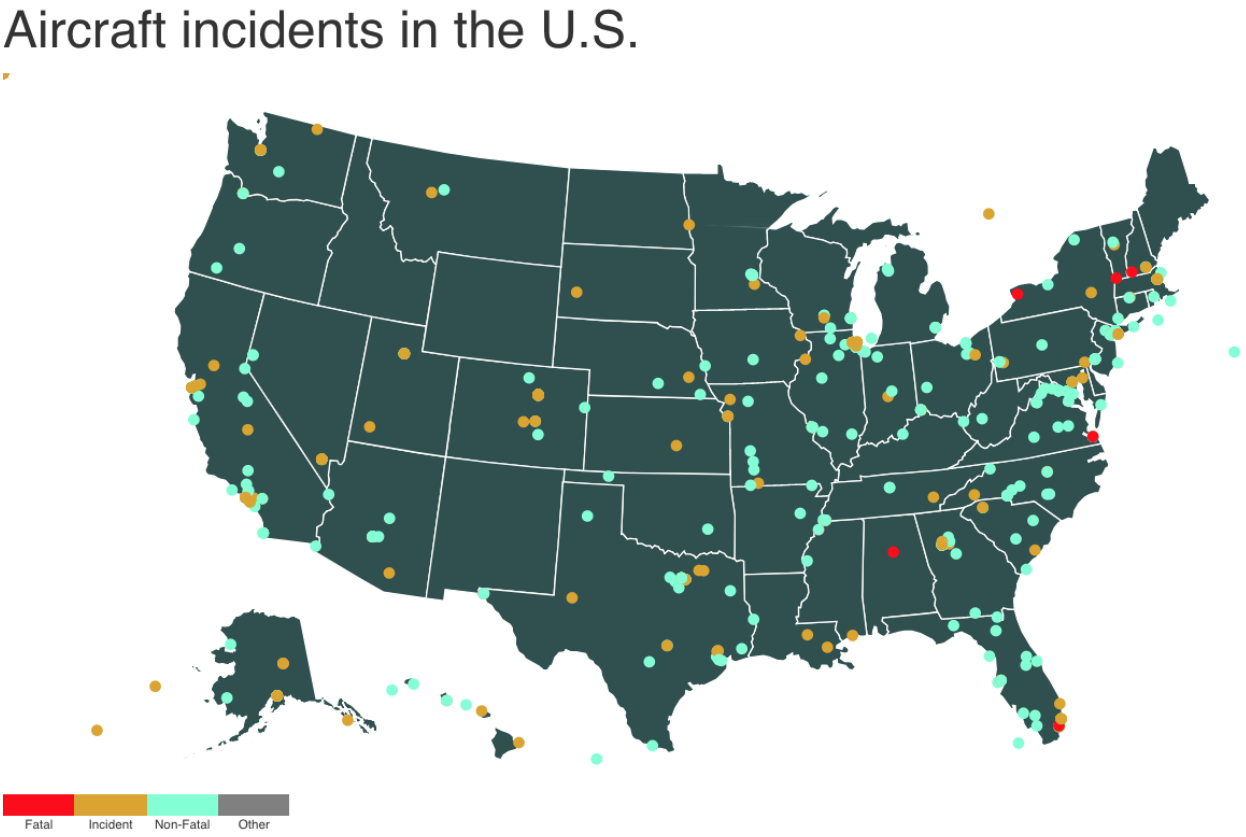

Figure 1: Aircraft Incidents plotted by longitude, latitude.

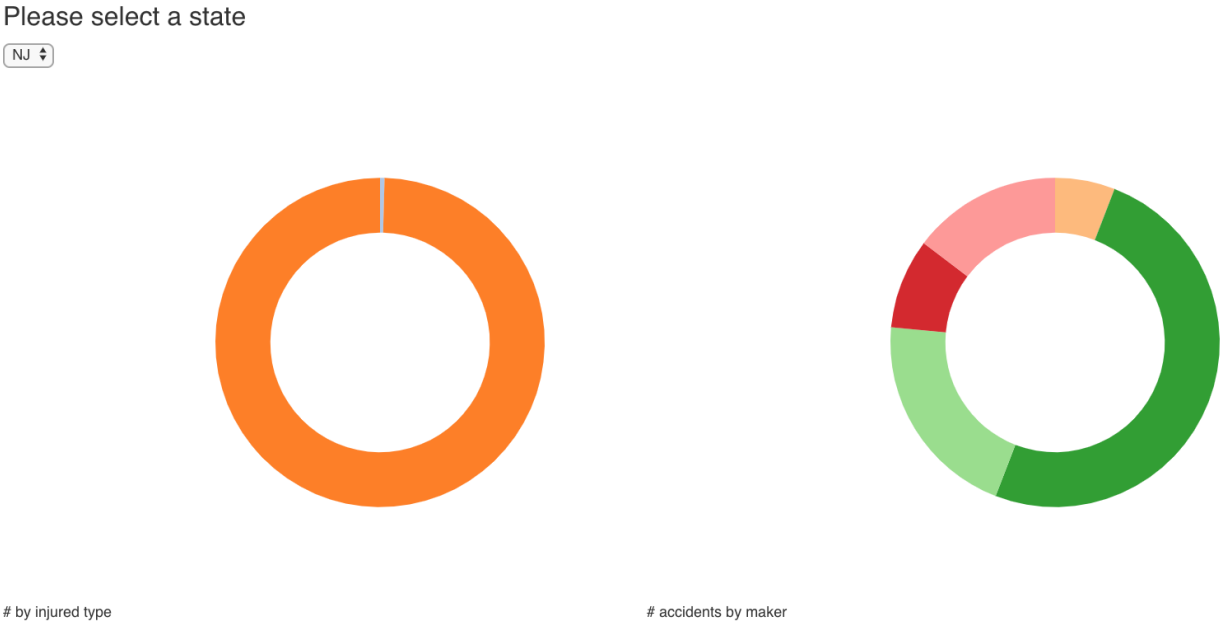

Figure 2: Pie charts to represent total fatal, serious and uninjured passengers based on the state and the aircraft make and model.

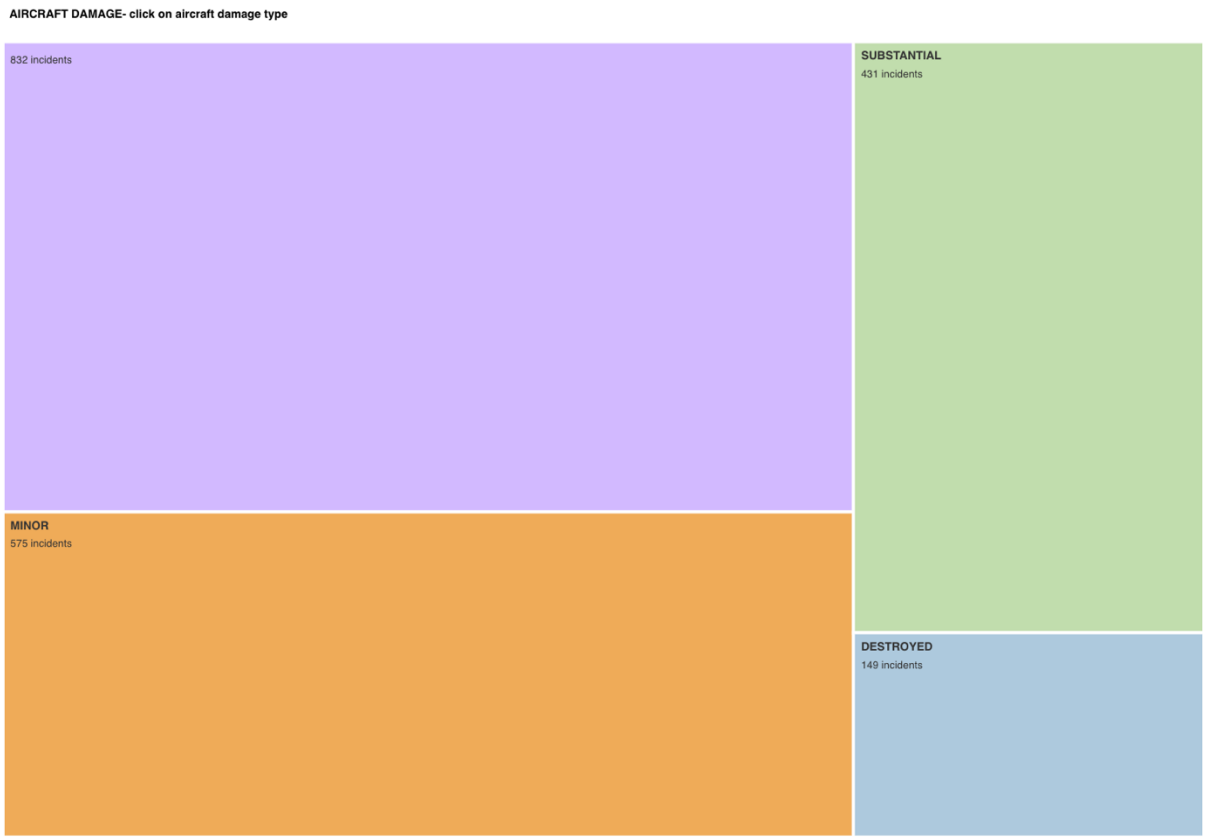

Figure 3: Tree map visualizes the number of incidents based on aircraft damage, board phase of flight, make, model and the frequencies of the numbers of passengers involved.