MAYDAY: A Visual Exploration of the FAA Accidents/Incidents Dataset

Ken Chan, Diviteja Guntamadugu, and Efrayim Zitron

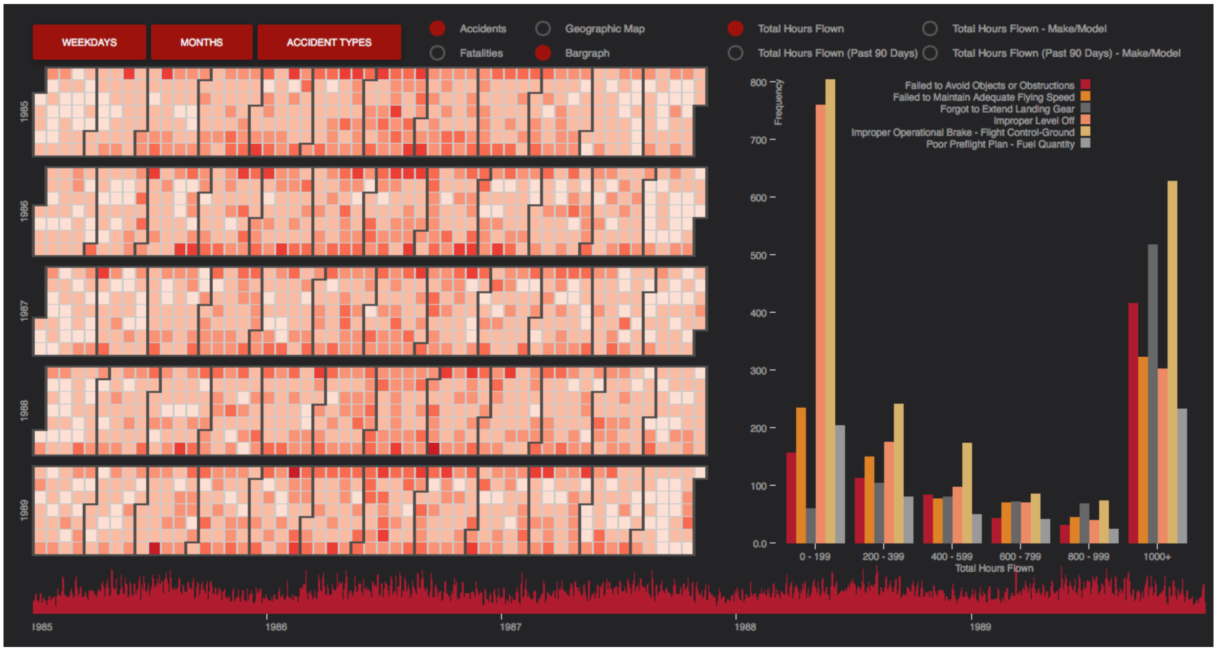

WEEKDAYS
MONTHS
ACCIDENT TYPES
● Accidents
○ Fatalities
○ Geographic Map
● Bargraph
● Total Hours Flown
○ Total Hours Flown (Past 90 Days)
○ Total Hours Flown - Make/Model
○ Total Hours Flown (Past 90 Days) - Make/Model
1985
1986
1987
1988
1989
Frequency
800
700
600
500
400
300
200
100
0.0
Failed to Avoid Objects or Obstructions
Failed to Maintain Adequate Flying Speed
Forgot to Extend Landing Gear
Improper Level Off
Improper Operational Brake - Flight Control-Ground
Poor Preflight Plan - Fuel Quantity
0 - 199
200 - 399
400 - 599
600 - 799
800 - 999
1000+
Total Hours Flown
1985
1986
1987
1988
1989

Fig. 1. MAYDAY: Years 1985-1989. The calendar heatmap on the left depicts number of accidents per day in the 5-year period. The grouped bar chart on the right displays the amount of occurrences a particular type of accident has with respect to pilot experience

Abstract—The FAA does not dedicate much resources on analyzing accident data involving small, non-commercial aircraft. Potentially valuable findings are hidden beneath large quantities of accident data, with no viable way to parse and extrapolate said data. We present *MAYDAY*, a visual tool aimed to alleviate this difficulty by culling the comparatively irrelevant data and displaying the relevant information in an easily accessible visual format. We focus on presenting an overall calendar heatmap displaying frequency of accidents in 5-year ranges and several more selective side components to depict relationships within the data. Thus, temporal trends and previously uninvestigated correlational relationships will become far more readily observable to domain users and researchers. In this paper we address our approach to determining which attributes are most relevant to domain researches, as well as our decisions of the final visual interface.

Index Terms—FAA Accident Database, Part-91, Accident frequency, visual exploration

1 Introduction

The FAA collects a wealth of information pertaining to all aircraft accidents that occur within the continental United States. Unfortunately, when it comes to accidents pertaining to small, non-commercial aircraft (Part 91 Operator Regulations), the FAA does not allocate the resources to thoroughly analyze the data to determine useful metrics and statistics. Analyzing this data to find meaningful correlations, especially with respect to accidents regarding private aircraft, is a fairly daunting task. However, if this data can be presented and visualized in an effective manner, very useful information can potentially be gleaned from it. In particular, determining whether there are observable seasonal or temporal trends, or whether statistical correlations between the pilot's experience, accident type, weather conditions, and geographic location of the accident can be very useful for various entities within the private aviation industry. Most important, it can be helpful for determining what safety protocols and policies need to be enacted to further mitigate the probability of an accident occurring. Our main interface with the domain is with the AOPA (Aircraft Owners and Pilots Association), and they have recently responded back with valuable input into what specific correlational and causal relationships they would like us to focus on. As such, we have selected a subset of the fairly large dataset which we believe can lead to very illuminating observations about the statistical distributions of the accidents. We believe this information could be very helpful to the members of this domain. Under the suggestion of our adviser, Jay Koven, an experienced pilot himself, we have decided to limit the data to Part 91 Operator regulations, which target small, non-commercial aircraft.

- *Ken Chan is with NYU Polytechnic School of Engineering. E-mail: ken.chan@nyu.edu.*
- *Diviteja Guntamadugu is with NYU Polytechnic School of Engineering. E-mail: diviteja@nyu.edu.*
- *Efrayim Zitron is with NYU Polytechnic School of Engineering. E-mail: efrayim.zitron@nyu.edu.*

2 Related Works

Tyler Fox, Mary Ann Howell, Michael Senatore, and Saji Varghese, who were students at Drexel University, explored [1] the FAA Aviation Accident Database and created several visualizations to search for trends and find correlations/relationships within these records. Such visualizations include word clouds, word trees, sparklines, and pie charts to serve as substantial evidence supporting their findings and claims. For instance, they aggregated the words found in the actual narratives of each accident to construct a word cloud that indicates the words frequently found and utilized. Another example involves a word tree visualization that illustrates hierarchically the most common words found adjacently to comprehensively give insight into the leading causes of these accidents. Based on this visualization and the magnitude of the size of the words, they determined that pilot errors were commonly caused by a failure to maintain control of the aircraft because of insufficient airspeed and not receiving proper clearance from airport personnel during standard landing procedures. Additionally, they found aviation accidents are primarily caused by general equipment problems involving the engine.

Though accidents can be caused by various factors, the scope of their work is limited to variables involving the build of the aircraft, visibility conditions and the text associated with describing the primary cause. Our tool is geared to identify temporal trends, and relationships between pilot experience, accident type, weather conditions, and geographic location. Additionally, the series of visualizations created by this group is not very interactive and dynamic, but rather, highly static. They built multiple tools, one for each use case, using Tableau, Many Eyes, Word It Out, Google Public Data Explorer, and Gephi. We will develop a single interactive visualization tool to explore all the factors in detail instead of having several different static visualizations.

Another paper [3], by Zohreh Naziri, George Donohue, and Lance Sherry , analyzes the patterns between aircraft incidents and accidents pertaining to commercial flights (part-121). Incidents, or events that result in serious bodily injury or aircraft damage, occur far more frequently than accidents happen. Merely studying accident data, which is a comparatively small sample size, does not allow for effective pattern recognition. By determining the causes of incidents, and associating them with accidents, the team attempted to accurately determine the most relevant factors that contribute to an event.

Our work differs from theirs in several respects. First, their research focused on commercial aircraft and we focused on events concerning non-commercial aircraft. Second, our tool focuses on different things, such as identifying temporal trends, pilot experience, and weather conditions. More importantly, our primary intention is to create an exploratory visualization tool, and although Naziri, Donohue, and Sherry incorporated some charts and graphs in their results to summarize their findings, creating a visualization tool was not their goal.

A team at Berkeley [2] developed a data-driven tool that displays aircraft accidents ranging from 1919 to 2014 (they utilized the Aviation Safety Network database to obtain the dataset for this tool). The purpose of this tool was not to analyze the type or cause of an accident, or what factors may have contributed to it, but merely to visualize the number of accidents that occurred over a specified timeline, along with the phase of the flight in which the accident occurred.

3 Data Analysis and Abstraction

Our dataset comes from the FAAs Accident/Incident Event database. The FAAs Aviation Accident/Incident Database is a complex relational database that stores in-depth details about each aviation accident that has occurred from 1975 to the present consisting of approximately 200,000 events. Each event consists of nearly 200 highly categorical attributes describing various aspects of the accident. Due to the incredibly large number of attributes, we had to be very selective in what attributes were selected for presentation. The final tool incorporated attributes pertaining to pilot experience, date of the accident, number of fatalities, weather conditions, type of accident, and geographic location. We believe, along with our adviser and domain user, that these attributes bear more relevance than most of the others. There are a number of other attributes which we decided had relevance, but were only able to include them to the extent that the user can see a list of these attributes when selecting a specific event(s). However, we were unable, either due to time or space constraints, to include them in the final iteration of the tool, as actual graphical representations of the data.

Table 1. Attribute Details

Name	Type	Range	Size	Description
Date of Accident	Quantitative	1-1-1975 to 12-31-2014	—	Date accident occurred
Fatality Count	Quantitative	˜0 to 50	—	Frequency of fatalities on a given day
Accident Count	Quantitative	˜0 to 60	—	Frequency of accidents on a given day
Accident Type	Categorical	—	68 types	Type of accident
Weather Conditions	Categorical	—	63 types	Weather conditions during the accident
Accident Location	Categorical	—	50 types	State in which accident occurred
Hours Flown	Quantitative	˜0 to 25000	—	Number of total hours flown
Hours Flown Make/Model	Quantitative	˜0 to 10000	—	Number of total hours flown with aircraft of same make and model
Hours Flown 90 days	Quantitative	˜0 to 1000	—	Number of total hours flown in past 90 days
Hours Flown	Quantitative	˜0 to 500	—	Number of total hours flown with same make/model in past 90 days
Primary Causes	Categorical	—	116 types	Primary cause of accident
Aircraft Make	Categorical	—	300 types	Manufacturer name
Aircraft Model	Categorical	—	1375 types	Aircraft model number
Airframe Hours	Quantitative	˜0 to 10000	—	Number of hours airframe experienced
Age	Quantitative	˜0 to 75	—	Age of pilot
Pilot Certification	Categorical	—	116 types	Pilot's flying certification level
Visibility	Categorical	—	13 types	Visibility level at time of accident
Light Conditions	Categorical	—	6 types	Light conditions at time of accident

Here, we briefly describe the main attributes which were used in the visualization tool. There are others which appear in the tool, but have a more secondary role, are thus just listed in the table. The quantitative attributes that are used in the final tool consists of the date of the accident, the number of accidents, the number of fatalities, and attributes related to pilot experience. The number of accidents is not an attribute provide by the FAA, but is one we created by aggregating the data by date. Similarly, to determine the total number of fatalities on a given day, the data was also aggregated by date, and the sum of the fatality frequencies from this subset of data for a given day was calculated to yield the total frequency for that particular day. Pilot experience is determined by the number of hours has flown, and is split into four categories. The first is the number of total hours the pilot has flown; the second is the total number of hours flown using the particular make and model of aircraft that was used during the accident; the third is number of hours flown within the 90 days prior to the accident; and the fourth is number of hours flown within the past 90 days using the particular make and model of aircraft used in the accident.

Attributes pertaining to weather condition, geographic location, and accident type are highly categorical. There are approximately twenty different weather conditions, ranging from light snow to hurricane. The geographic location used for the visualization was the state in which the accident occurred. Ideally, we would have wanted to use the actual coordinates of the accident, but many accident did not have that information provided. The final, and possibly most important attribute, is the type of accident, of which there over 60. This attribute is very important because we believe that seeing accident type graphed vs pilot experience can prove to be useful to the domain.

Table 1 has a list of the final attributes, that were used in the final iteration of the visualization. Most of them are not esoteric and are understood at least on a superficial level by a layman. Most of the attributes are categorical or quantitative in nature. We determined what data to cull by analyzing the dataset and remove attributes that we either believed to be relatively inconsequential or redundant. This data includes information such as the more technical attributes of the aircraft/engines (such as type of landing gear or type of engine), and the non-primary contributing factors of the incident. To specifically list all the culled data would be too cumbersome, as more than 150 attributes were removed.

4 Task Analysis and Questions

Guidance for our project stemmed from two sources: Our adviser, Jay Koven, and our contact at the AOPA (Aircraft Owners and Pilots Association), David Kenny. There was much overlap in what was wished to be visualized. There was much interest in identifying temporal or seasonal trends, potential relationships between accident type, pilot experience, and weather conditions. Or tool would be developed in an attempt to identify and present such trends and patterns. As such, our task questions were developed as follows:

- Are there any seasonal or long-term trends with regard to accident frequency?
    - If there are any, do the trends hold when accident type, pilot experience, and weather conditions are taken into account?
- Is there a significant correlation between the total number of flying hours a pilot has logged and the type (as well as quantity) of accidents over a period of time?
    - If so, does that same statistical correlation apply when looking at the number of hours flown in the 90 days prior to the accident?
- Is there a significant correlation between the total number of flying hours a pilot has logged with the specific make and model of aircraft used in the accident and the type (as well as quantity) of accidents over a period of time?
    - If so, does that same statistical correlation apply when looking at the number of hours flown in the 90 days prior to the accident?
- What effect do various weather conditions have on the frequency of events/accidents?
    - Is there a geographic relationship between accidents and types of weather conditions? Is there a prevalence of a specific weather condition contributing to accidents?

5 Visualization and Interaction Design

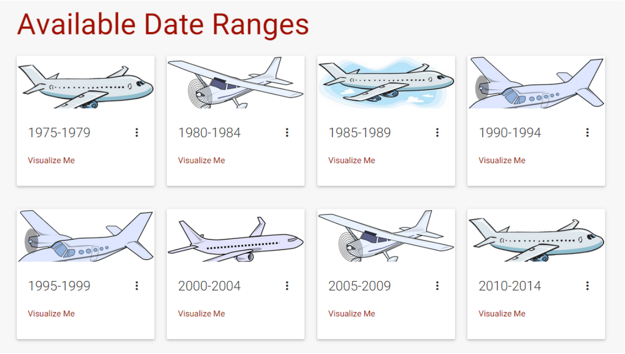


Fig. 2. MAYDAY landing page

The landing page (Fig 2) is the initial portal to the visualization tool. The FAA incident/accident dataset is divided into multiple .csv files based on a given time period (every five years since 1975). This enables the user to limit the scope of the temporal analysis to a certain time period. For instance, the user might be interested in visualizing a recent accident dataset as opposed to an older accident dataset. Additionally, the user might have a strong background in history, and potentially correlate historical events to certain patterns in accident and fatality frequencies. Dissecting the dataset using this particular method helps to reduce the amount of data loaded within the browser, which expedites the rendering of the individual visualization panels within the dashboard.

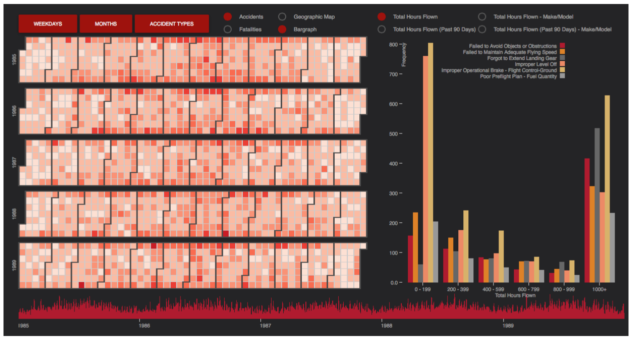

WEEKDAYS
MONTHS
ACCIDENT TYPES
● Accidents
○ Fatalities
○ Geographic Map
● Bargraph
● Total Hours Flown
○ Total Hours Flown (Past 90 Days)
○ Total Hours Flown - Make/Model
○ Total Hours Flown (Past 90 Days) - Make/Model
1985
1986
1987
1988
1989
Frequency
800
700
600
500
400
300
200
100
0.0
Failed to Avoid Objects or Obstructions
Failed to Maintain Adequate Flying Speed
Forgot to Extend Landing Gear
Improper Level Off
Improper Operational Brake - Flight Control-Ground
Poor Preflight Plan - Fuel Quantity
0 - 199
200 - 399
400 - 599
600 - 799
800 - 999
1000+
Total Hours Flown
1985
1986
1987
1988
1989

Fig. 3. MAYDAY overview.

The visualization (Fig 3) itself is broken up into several components. The first is the calendar heatmap, which is intended to give a temporal view of accident/fatality frequency. The second is a grouped bar chart which displays the number of accidents with respect to accident type and pilot experience. The third is a topoJSON map which color-codes state by accident frequency, and shows a breakdown of weather conditions for each state. There is a barchart on the bottom which depicts accident values for the 5-year period while acting as a timeline scale as well. Whatever time frame is selected on the slider will be reflected in the grouped bar chart, treemap, and topoJSON map as well.

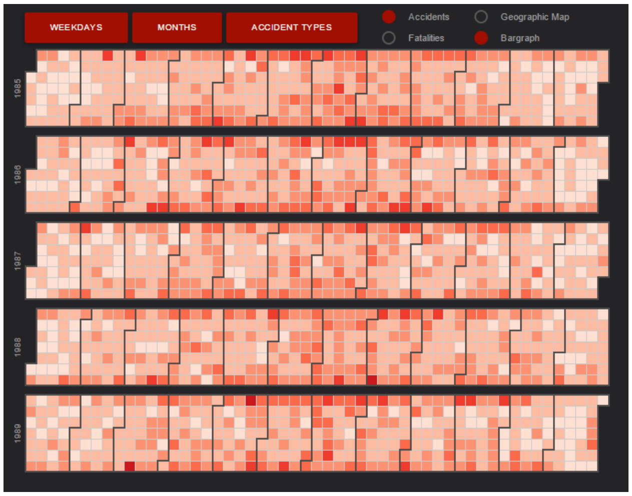


Fig. 4. Calendar heatmap. The higher the saturation, the more accidents occurred on that day.

A calendar heatmap (Fig 4) is generated to visualize seasonal and long-term trends. Accident data within a given year is divided into years, as represented by each labeled strip of twelve polygon shapes with a black background, which individually represent the months within the year. Furthermore, the boxes within these polygons represent a particular day. The saturation spectrum of red encodes the number of accidents or fatalities, as chosen by the user via radio buttons, experienced for each day with higher levels of saturation indicating instances of massive casualties and lower levels of saturation indicating less catastrophic instances. This calendar heatmap can be utilized as a general overview of the dataset to guide the user towards the multiple panels on the right-hand side to determine any trends caused by pilot-based, and condition-based variables. The user can hover over any of the small, colored boxes to see a small text box with granular details such as exact fatality and accident total counts for its particular day of a month and year combination, but clicking on the box alters the visualizations within the right-hand side panels to restrict the data displayed to a given day and modifies the slider at the bottom of the dashboard, which is driven by a brush control, to reflect this selection.

An alternative approach involves rendering a series of five bar graphs instead of five rectangular strips for the calendar heatmap to visualize the accident and fatality frequencies over the five year period. However, despite this method having several advantages, particularly familiarity and easier consumption of the data, it ultimately proved to be ineffective because of the minuscule widths of the individual bars (365 bars per bar graph), which was found to be inappropriate for granular interactions such as hovering for determining precise counts and clicking for reducing the data loaded into the other visualization panels to a subset for the particular day selected.

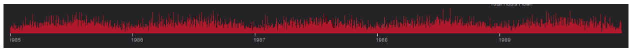


Fig. 5. Timeline Scale: Depicts bar-chart alternative to heatmap and serves as a time selector

A timeline scale (Fig 5) at the bottom of the dashboard allows the user to restrict the data displayed within the right-hand side visualizations to those within that particular time period. Tick marks located beneath the bars on the horizontal x-axis mark the beginning and ending of each year. Furthermore, the time period selected will be reflected upon the calendar heatmap by highlighting the colored boxes representing days that fall within this time period and fading out the remaining boxes. Additionally, the implementation of the timeline scale incorporates the bar-graph alternative to provide a comprehensive overview of the frequency trends over the five year period, which enables the user to recognize noticeable spikes in the frequency or time periods with constant low frequency levels.

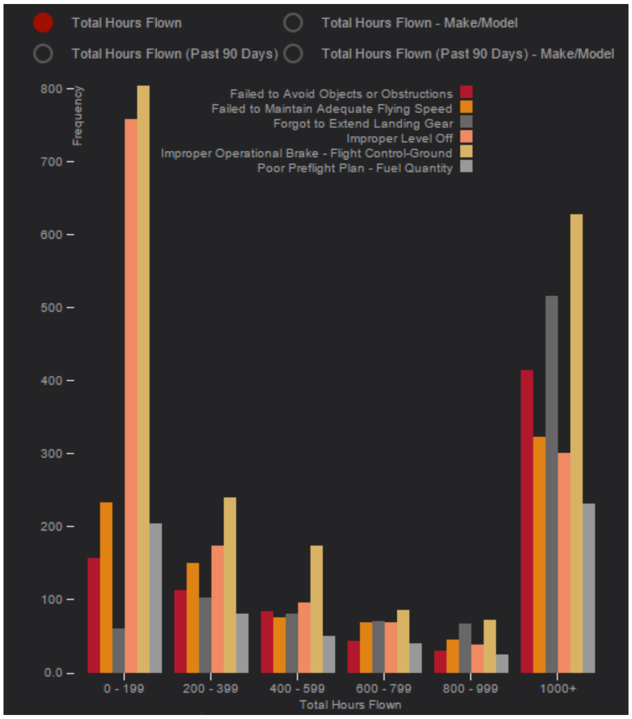
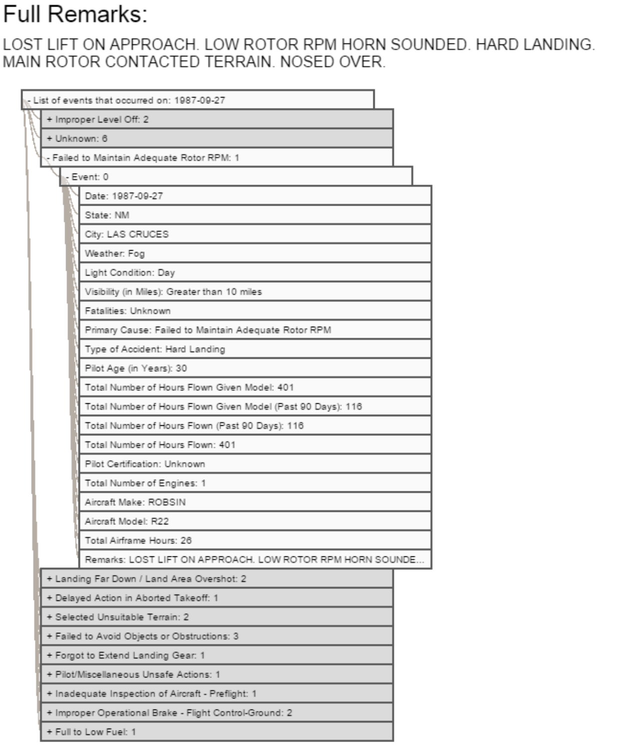


Fig. 6. Grouped Barchart: Depicts accident frequency with respect to accident type and pilot experience

The grouped bar chart (Fig 6) displays the amount of occurrences a particular type of accident has with respect to the number of hours flown by the pilot, whether or not dependent on the aircraft's make and model (this restricts the pilot to only the most recent aircraft operated if the pilot has operated multiple vehicles since the last time the hours were logged). The data was aggregated into six particular bins, each encompassing hourly values that fall within a certain interval: 0 to 199 hours, 200 to 399 hours, 400 to 599 hours, 600 to 799 hours, 800 to 999 hours and 1000+ hours. The hourly variables are total hours flown, total hours flown given a make/model combination, total hours flown in the past ninety days and the total hours flown in the past ninety days given a make/model combination, and the x-axis of the grouped bar chart can represent any one of these variables at any given moment via radio buttons. A maximum of six different accident types can be selected from the corresponding dropdown menu to be visualized in the grouped bar chart. This limit on the number of accident types is enforced due to the limited real estate available on the screen. Some use cases include a user wishing to determine which accident types were most prevalent in one bin as opposed to another and examining similar accident types (improper usages, inadequate resources, etc.) for comparing and contrasting their frequencies within the individual bins. Additionally, if the user decides to compare the frequencies of certain accident types across different total hours flown variables, then the filtered options are carried over to this newly generated bar chart with the newly adjusted x-axis. Limiting the scope of the bar chart to a particular time period via the bottom timeline slider is reflected correctly, and carried over to all the other total hours flown options.

Initially, a scatterplot matrix was utilized to portray this data, but this idea was eventually discarded since numerous tests showed the scatterplots to be dense with many points with no recognizable clusters and unreadable for the dashboard's purposes.



Full Remarks:

LOST LIFT ON APPROACH. LOW ROTOR RPM HORN SOUNDED. HARD LANDING. MAIN ROTOR CONTACTED TERRAIN. NOSED OVER.

- List of events that occured on: 1987-09-27
+ Improper Level Off: 2
+ Unknown: 6
- Failed to Maintain Adequate Rotor RPM: 1
- Event: 0
Date: 1987-09-27
State: NM
City: LAS CRUCES
Weather: Fog
Light Condition: Day
Visibility (in Miles): Greater than 10 miles
Fatalities: Unknown
Primary Cause: Failed to Maintain Adequate Rotor RPM
Type of Accident: Hard Landing
Pilot Age (in Years): 30
Total Number of Hours Flown Given Model: 401
Total Number of Hours Flown Given Model (Past 90 Days): 116
Total Number of Hours Flown (Past 90 Days): 116
Total Number of Hours Flown: 401
Pilot Certification: Unknown
Total Number of Engines: 1
Aircraft Make: ROBSIN
Aircraft Model: R22
Total Airframe Hours: 26
Remarks: LOST LIFT ON APPROACH. LOW ROTOR RPM HORN SOUNDE...
+ Landing Far Down / Land Area Overshot: 2
+ Delayed Action in Aborted Takeoff: 1
+ Selected Unsuitable Terrain: 2
+ Failed to Avoid Objects or Obstructions: 3
+ Forgot to Extend Landing Gear: 1
+ Pilot/Miscellaneous Unsafe Actions: 1
+ Inadequate Inspection of Aircraft - Preflight: 1
+ Improper Operational Brake - Flight Control-Ground: 2
+ Full to Low Fuel: 1

Fig. 7. Treemap which shows some additional information for accidents that occurred on day(s) selected.

The treemap visualization (Fig 7) categorizes a subset of the data, which is based on the time period selected, ordered by an accident's primary cause. This component is contained within the side-panel, and users can view a more detailed list of factors of a specific event. The primary purpose of this panel was to provide the numerous details about any accident (date, visibility conditions, remarks, etc.) in an organized manner. If the slider is adjusted to a specific range, then the side panel displays all events that occurred during that time-span, ordered by accident type. This is not essential for our task abstractions, but we believe that it can be useful to the domain users since if a particular event is selected by the user, then the user can gain more in-depth insight into the details of the accident. Additionally, the user can move the cursor over the remarks to view the full description.

The final component is the topoJSON map (Fig 8), which is a map of the United States where all the states have been outlined. The stauration spectrum of red encodes the number of accidents or fatalities that occurred within a specific state. Higher levels of saturation indicate that the state has a very high accident/fatality count.

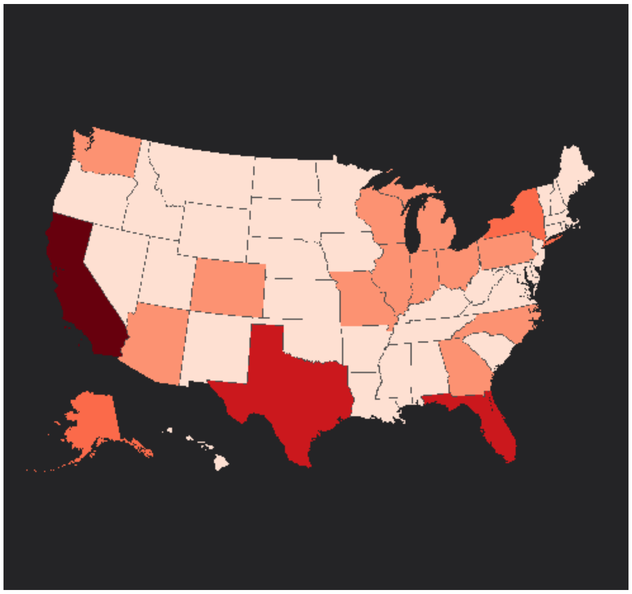


Fig. 8. TopoJSON Map: Shows a topoJSON map of the United States, where higher color values correlate to states with a high accident count

When the user clicks on a state, the topoJSON map zooms into the state, and a bar graph of the six most recurrent weather conditions for accidents within this state is generated (Fig 9). The bars are ordered with the tallest one, which corresponds to the most frequent weather condition, being positioned to the left, and the smallest one, which corresponds to the sixth most frequent weather condition, being positioned to the right.

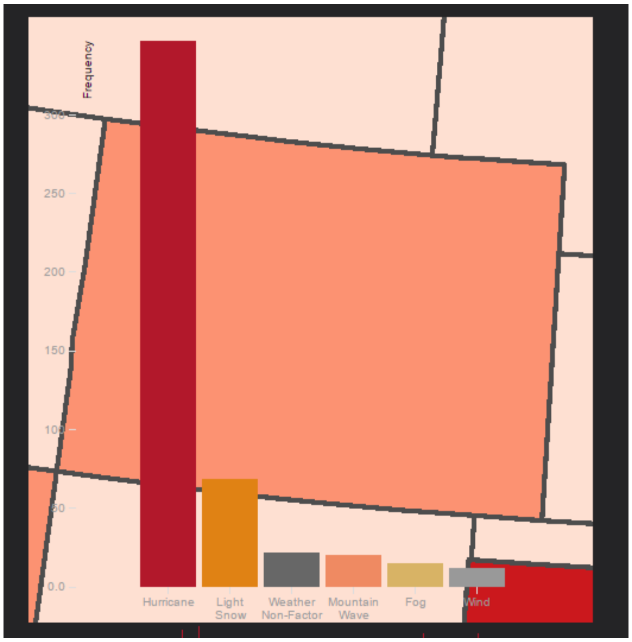


Fig. 9. TopoJSON Map Zoom: For a selected state, presents a bar chart which depicts accident count with respect to various (top 6) weather conditions

6 Findings and Insights

Visual exploration played a significant role in the discovery of several temporal patterns in regards to the relationship between the frequency of accident types and total hours flown given certain circumstances. Specifically, the visualization provides supporting evidence to reinforce some preconceived ideas relating to the seasonal trends, geographical location and factors as a contributor to the occurrence of accidents, etc.

- *Finding 1: Seasonal/Temporal Trends:* After designing the calendar heatmap, it was immediately noticeable that there was a tendency for accidents to occur during the summer months. Additionally, weekends generally featured more accidents than weekdays. See Fig 10. This was expected, as more people fly during vacation and warmer weather but nonetheless interesting to see it visualized so glaringly.

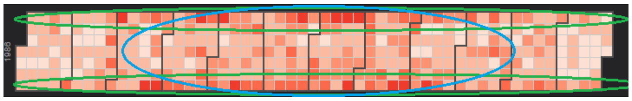


Fig. 10. Higher accident count during summer months and weekends

- *Finding 2: Overall Experience Trend:* What became very apparent upon making the barchart was the general trend of decreasing accidents with respect to increased experience (Fig 6). Although it appears that there is a spike in the 1000+ hours bin, that is only because that bin encompasses pilots with up to 25000 hours of experience. Overall, the more experience a pilot has lends to lower accident frequency.

- *Finding 3: Geographic Accidents:* Highly populated states with warmer weather (California, Texas, Florida, etc.) were found to have a larger number of accidents compared to other states (Fig 9). In practically all states, the most prevalent poor-weather condition for accidents has been hurricane weather. This seems to indicate that pilots are less concerned about the risks of flying in hurricane weather even though it is evidently quite dangerous.

- *Finding 4: Total Hours Flown - Make/Model:* The bin for 0 to 199 total hours flown given a make/model combination was found to have a disproportionately larger frequency of accidents than the other bins for primarily any accident type grouping. This is an expected outcome, as we expect pilots with less experience with a particular make and model to be more prone to errors and mishaps. What is more interesting is that some types of accidents occur with a higher frequency with pilots with greater general experience, but less experience flying a specific model. than pilots with less general experience. This may imply that experienced pilots may be more prone to encounter a specific event when flying an aircraft that they lack extensive experience with, than pilots who have overall less experience overall. As an example, see the 'Forgot to extend landing gear' accident type in the 1985-1989 dataset (grey bar, observable in Fig 11 and Fig 12).

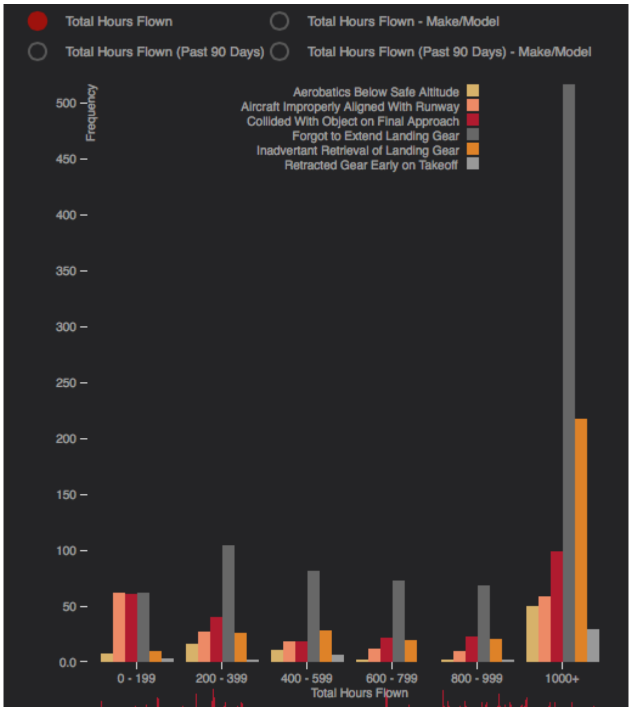


Fig. 11. 1985-1989 dataset; Total hours flown; Forgot to extend landing gear

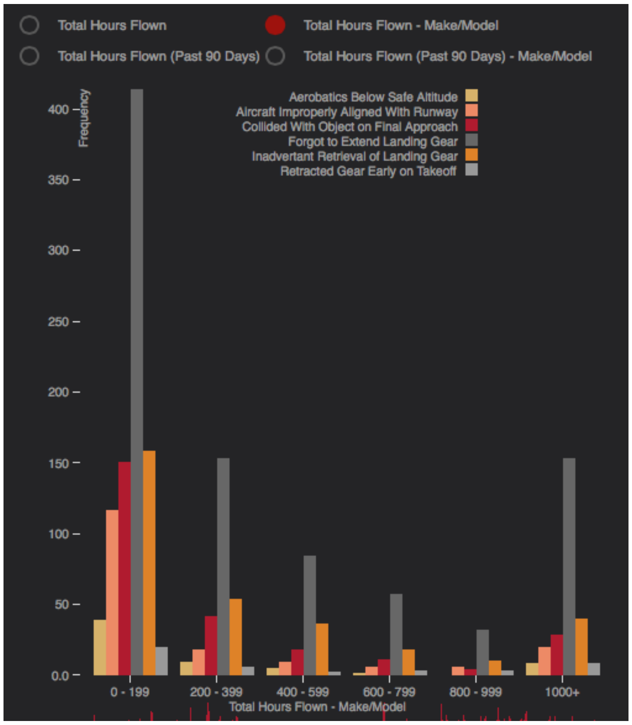


Fig. 12. 1985-1989 dataset; Total hours flown - Make/Model; Forgot to extend landing gear

- *Finding 5: Frequency of Reported Accident Types Over Time):* The frequency of the accident types reported has decreased over time. Particularly, the 1980 to 1984 dataset featured the most accident types reported, as shown by the initial view of the grouped bar chart with the six most reported accident types. Also, no accident within the 2010 to 2014 dataset had a reported accident type.
- *Finding 6: Accident Types Involving Improper Operation):* The frequency of the accident types reported as ”Improper...” was found to be substantially larger than others reported as ”Poor Preflight Plannng”, ”Inadequate...”, etc. This implies human error during the in-flight phase caused a majority of the accidents.

7 Limitations and Future Work

The primary limitation for our work was the sheer quantity of factors that the FAA records for each individual accident. Ideally, we would be able to include as many of the relevant factors directly into the visualization tool, but due to to space and time constraints, we were unable to do so. Specifically, visualizing attributes such as airframe hours, make/model of plane, light conditions, visibility, etc., would be very useful in analyzing aircraft accidents. Our rough mockups included visualization components that would graphically display these data.

Additionally, we would focus on expanding our existing components to enhance analytical potential. Specifically, we would somehow leverage our existing real estate to allow for bins in our grouped bar chart. We believe having the bins quantized with smaller values would be helpful, especially when viewing pilot with minimal experience. Seeing smaller bins for pilots with 0-200 hours of flying may allow for different patterns to emerge.

8 Conclusion and Lessons Learned

One of the first, and most important lessons we learned, was the sheer difficulty in attempting to visualize and analyze such a large dataset. We spent innumerable hours juts trying to decide which attributes would be more valuable than others, and would not heavily skew the results. Additionally, due to the highly categorical nature of most of the attributes, we were forced to find clever ways to visualize the data. This resulted in many discarded ideas, which we determined would not adequately serve our, or the domain's, interests. We considered several potential designs that we though would be 'cool' and highly sophisticated. In the end, we discovered that sometimes a simpler view can yield the most promising results.

As consumers, we do not pay attention to the details of various visualizations that are presented to us, and we take for granted all the minute details that the content creator has to consider prior to releasing his tool. Details like component position, color selection, and data encoding need to have significant forethought in their design. Through trial and error, we hope we managed to select the ideal components and encoding to successfully visualize the data.

Acknowledgments

The authors wish to thank the AOPA, Jay Koven and Professor Enrico Bertini for their guidance and insights in understanding the domain and developing an effective, intuitive visualization tool for addressing the necessary task abstractions.

Links

- GitHub Repository:
  http://www.github.com/NYU-CS6313-Projects/FAA-Visualization
- MAYDAY Website:
  http://maydayviz.azurewebsites.net/
- Video:
  https://vimeo.com/128098177

References

[1] T. Fox, M. A. Howella, M. Senatore, and S. Varghese. Visualizing the faa aviation accident database.

[2] H. Jannah, S. Agrawal, and S. Siddiqui. Talespin: Visualizing airplane accidents over time.

[3] Z. Naziri, G. Donohue, and L. Sherry. Analyzing relationships between aircraft accidents and incidents.