Visualizing the FAA Aviation Accident Database

Tyler Fox, Mary Ann Howell, Michael Senatore, and Saji Varghese
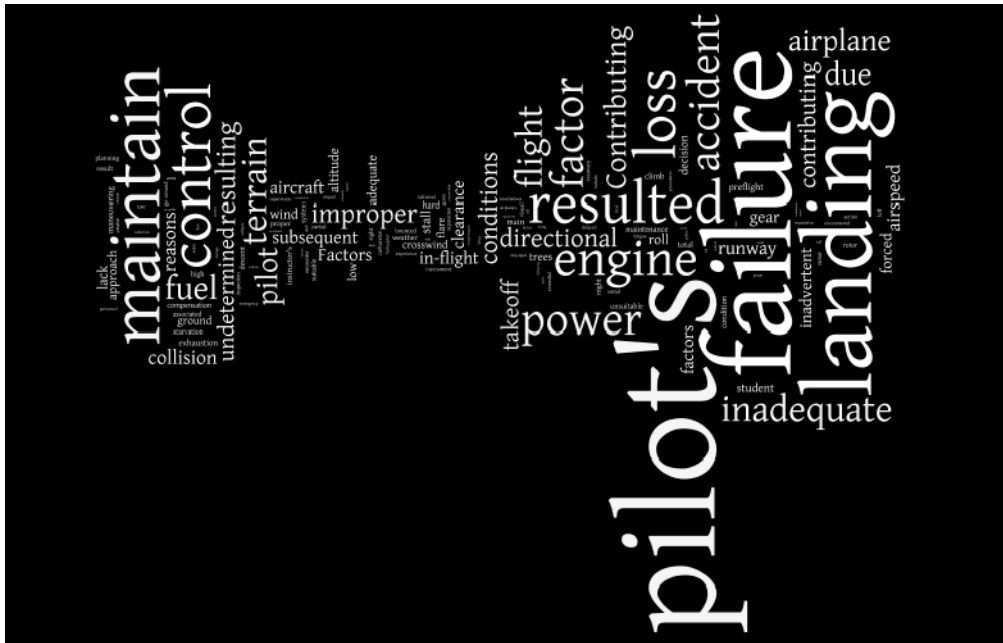
Figure 1:Word Cloud from Combined Narratives (generated in Many Eyes)

Abstract—The Federal Aviation Administration (FAA) Aviation Accident/Incident Database is a complex relational database that stores in-depth details about each aviation accident that has occurred from 1962 to the present. This paper demonstrates how information visualization techniques can be used to search for trends, and find correlations and relationships in the accident records. Visualizations are created in Many eyes, Word It Out, Tableau, Google Public Data Explorer, and Gephi, and demonstrate how the data related to aviation accidents that occurred inside the 50 United states from 1990 to 2010 can be explored.

Index Terms—FAA, information visualization, aviation safety, accident investigation

◆

1 INTRODUCTION

The rather chilling word cloud visualization above depicts what is often the underlying cause of aviation accidents: pilot error. (Commercial flights, slightly over 50%, general aviation, year-by-year in the 70 percentile range [1], [2].) But what are the events and factors underlying these and other aviation accidents?

The FAA Accident Database records hundreds of details surrounding more than 68,000 accidents. This database is open and accessible to the public, ripe for information mining to answer questions, tell stories, and search for insights that could lead to improvements in aviation safety.

The FAA Accident database is an excellent candidate for exploration using Information Visualization techniques. Chaomei Chen, in his book: *Information Visualization: Beyond the Horizon*, defines Information Visualization as "computer generated interactive graphical representations of information, and also "the process of producing information visualization representations" [3].

1.1 Issues

Accidents are rarely the result of one factor, they are almost always caused by an unfortunate chain of events [4]. Analyzing these multiple chains of events to isolate common causes is a challenge. But the chain of events can be searched for re-occurring themes, and the possibility of the elimination of one unfortunate event could help break potential future chains, and keep a deadly sequence from forming again.

Another challenge is the sheer massiveness and the complexity of the database itself. There are twenty tables with multiple columns, many of which have many-to-many relationships with records from other tables. Column headings are abstract and often use codes whose meanings are listed in other tables. Searching the database with simple queries can be difficult, especially if you are not sure what your are looking for.

1.2 Our Approach

The FAA accident database is a complex relational database that stores hundreds, sometimes thousands of details about each aviation accident. Designing an information visualization tool to explore the entire dataset at once is a daunting task. This paper demonstrates how subsets of related information can be pulled out and examined for possible trends and correlations.

Information visualization can help investigators see at a glance what other influences may be common that increase the likelihood of accidents, or affect pilots ability to react effectively. For instance, are there trends relating to environmental effects such as visibility, light conditions, or wind velocity? What is the size of plane that is most commonly involved in accidents? Are accidents more common on particular days of the week? What airports seem to be the deadliest?

1.2.1 The FAA Aviation Accident/Incident Database

The FAA accident database contains records from 1962 to the current date [5]. The database includes accidents and some incidents (close calls) that involve all types of aircraft and that incurred within the United States, its boundaries, or its territories. The database contains two key tables (aircraft and events) which between them, have over 150 different fields of information. 18 additional tables support these two key tables. Currently, residing in this structure are over 68,000 event records for separate accidents [6].

In order to work with a manageably-sized dataset, our project generally uses a subset of the FAA Aviation database that includes fatal accidents that occurred within the 50 US states, from 1990 through 2010. These accidents include all types of aircraft flying under all types of flight-rule categories.

2 INFORMATION VISUALIZATION EXAMPLES

The visualizations that follow were created in Many eyes, Tableau, Google Public Data Explorer, and Gegraph, and demonstrate how the data related to aviation accidents from 1990 to 2010 can be explored. These tools are all freely available on the Internet.

2.1 Exploring Connections Between Aircraft Manufacturers and Injury Accidents

Traveling by air is the fastest means to travel, but like all travel options, it has risks. Accidents can occur due to bad weather, equipment malfunctions, human error, fault of the pilot, fault of the crew, or a combination of events. Tableau and Word It Out were used to look for trends and relationships between aircraft manufacturers and any accident that involved injuries.

2.1.1 Tools Used

Two tools were used to investigate connections between injury accidents and aircraft manufacturers:
- Tableau Public
  "... is a free service that lets anyone publish interactive data to the web. Once on the web, anyone can interact with the data, download it, or create their own visualizations of it" [7].
- Word It Out
  Is a beta applet that allows users to visualize text according to the importance of the words used. The more important the word, the larger font, allowing users to summarize large documents at a glance [8].

2.1.2 Finding Relationships with Tableau Public

Tableau Public includes a wide range of visualization techniques, such as bar graphs, line charts, pie charts, and geographic charts with longitudes and latitudes. Tableau Software provides many options, allowing users to match visualization displays with their data type. Ron Powell of the Business Intelligence Network stated that "Tableau's visual analysis provides the ability to address unanticipated questions without waiting weeks by performing on-the-fly calculations, comparative visual analysis, and ad hoc question and answer sessions" [9].

Tableau supports Ben Shneiderman's "Information Visualization Seeking Mantra", which states that users should be able to overview the data, zoom in on particular data sets, choose which data to visualize, and filter out unwarranted information. Continuing with the Mantra, Tableau also allows users to seek relationships between datasets and allows multiple options for extraction [10]. Although Tableau has many benefits, it can be a complex software tool, especially for the less experienced user. Supporting documentation is very limited.

Here, Tableau Public is used to capture which aircraft manufacturers and models were involved in the most injury accidents.
The top five manufacturers involved in aircraft injuries are:
- Piper: involved in 3,241 injuries.
- Beech: involved in 1,876 injuries,
- Boeing: involved in 986 injuries,
- Bell: involved in 482 injuries.

Cessna, Piper, Beech, Bell, and Boeing manufacture five of the most popular and widely-used aircraft. Most Cessna, Piper, and Beech models are small aircraft, consisting of four-six seats, a single engine, with a high-fixed wing. The Bell models are helicopters and the Boeing models are commercial jet liners. Figure 2 shows the five manufacturers and their models that were involved in the most injury accidents. Each slice represents the number of injuries per model, the bigger the slice, the higher the number of injuries.

Figure 2 shows that:
- PA-28 models were involved in the highest number of injury-causing accidents for Piper.
- 172 models were involved in the highest number of injury-causing accidents for Cessna.
- A36 models were involved in the highest number of injury-causing accidents for Beech.
- 206B models were involved in the highest number of injury-causing accidents for Bell.
- 747-131 models were involved in the highest number of injury-causing accidents for Boeing.
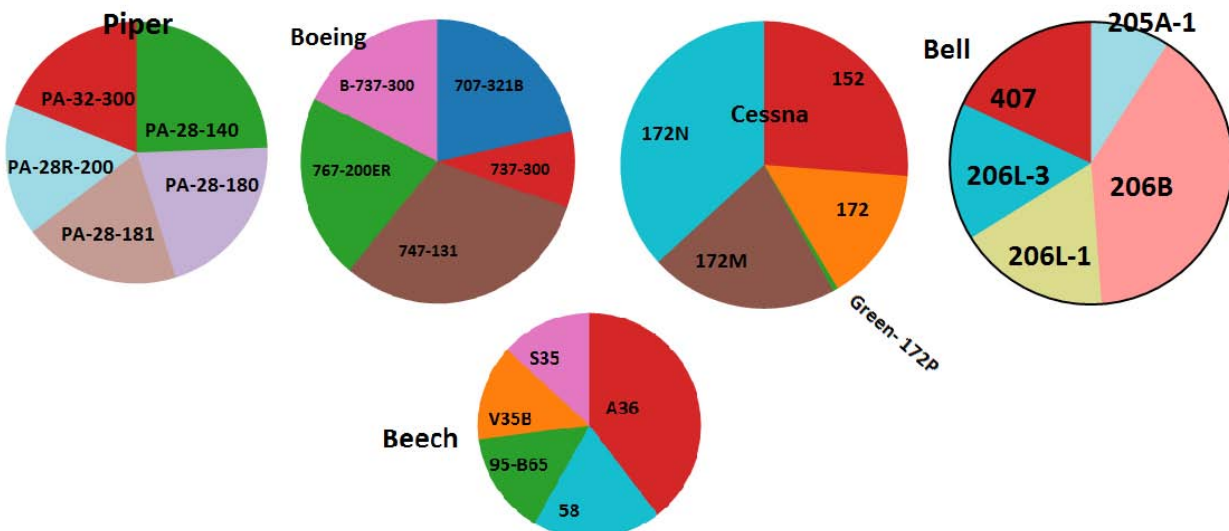

Figure 2: Manufacturer: Models Representing Highest Injuries

2.1.3 Checking the Findings Using Word It Out

Tag clouds, or word clouds, add an aesthetic piece to a web site, are simple to create, and are effective in relating their message. Word It Out generates a word cloud from inputted text. The word cloud may include hundreds of words, and the larger the word, the more frequently it was used.

Word It Out allows several customization options, such as changing colors and font size, to give the final graphic a more appealing and artistic look. These clouds provide an excellent way to sum up what's important in the text, and can help users narrow their searches.

We used word clouds to further examine the findings in Figure 2. Are the aircraft models displayed in the pie charts inherently dangerous machines, or are they just the most widely-used aircraft?

After aggregating the narratives of accident causes by manufacturer, and then feeding in the aggregated narratives to Word It Out, we found that the most used words for all models and manufacturers were: *Pilot, Maintain, Failure, Flight, Condition, Weather.* (See Figures A-1 through A-5 in the Appendix.) On closer look, the word clouds show that most accidents were caused by pilot error, air traffic control error, improper maintenance, and weather conditions, not malfunctions from aircraft design or manufacturing problems. Using multiple visualization tools can help investigators double check trends, patterns, and relationships between the data being investigated, and can assist investigators determine the key points of the causations. The clouds used in this section were generated from a subset of manufacturer-related data, See 2.3 "*Analyzing Narrative Text: An In-depth Look into Accident Factors*" in the Appendix for more information on of exploring the causes of accidents using word clouds.

2.2 Investigating Fatal Aviation Accidents Over a Range in Time

Since the aviation accident database includes dates and locations, total fatalities can be tracked over time, state by state. Animations are an efficient method of finding trends and patterns over a range of time, and can dramatically accentuate the fluctuations of data. Sparklines, tiny line graphs designed to be used inline with text, can also be used to quickly display variations over time.

2.2.1 Using Google Public Data Explorer to display time-variant animations

In 2010, Google released the beta version of Google Public Data Explorer, a free information visualization tool. Google allows registered users to upload datasets and publish them as information visualizations. These information visualizations can be kept private, shared among a designated group of users, or made public for anyone to access [11].

After importing a groomed subset of the aviation fatal accidents dataset, Google Public Data Explorer displays data points over time, with three views:

- Line graph
  Displays the number of fatalities on the Y axis and years on the X axis. States can be selected and fatality numbers compared against each other and against the US total for each year.
- Bar chart
  Displays the number of fatalities on the Y axis and a bar for each state on the X axis. States can be selected and fatality numbers compared against each other. When the blue arrow is clicked, an animation displays that shows how the bar sizes fluctuate over time. You can hover over a bar to get the state's name and its fatalities total.
- Map chart
  Displays an animated view of the change in total fatalities for each state over time, indicated by bubbles of varying sizes. You can hover over a bubble to get the state' name and its fatalities total.
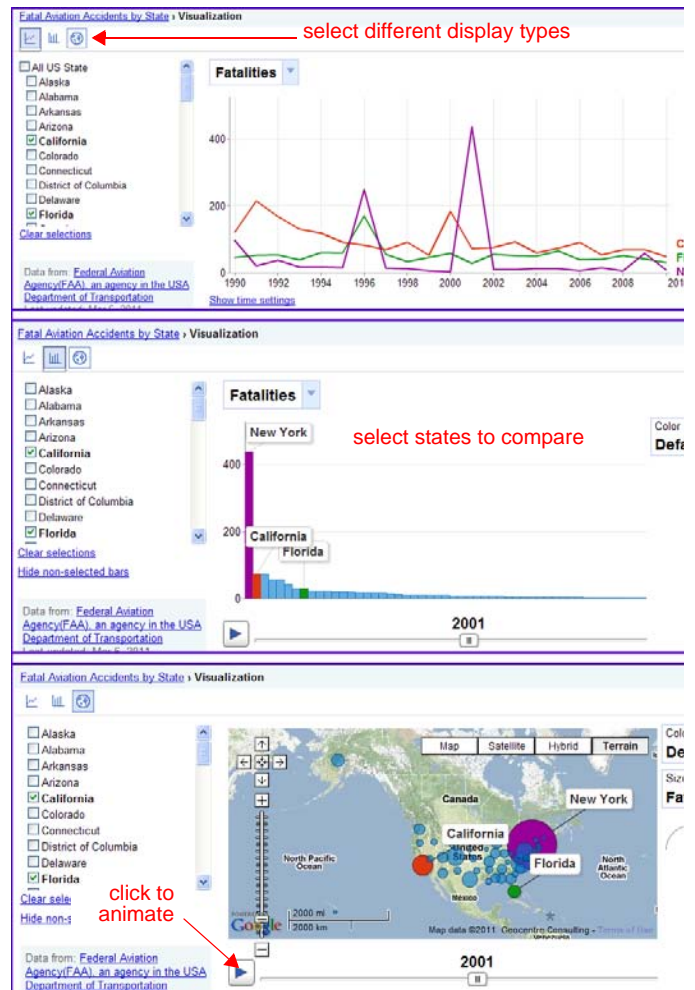

Figure 3: Three types of Google Public Data Explorer charts showing aviation fatalities per state from 1990 to 2010

The three screen shots above demonstrate how a dataset can be explored. The displays pictured here are frozen in time to 2001, comparing aviation fatalities for California, Florida, and New York. You can easily see the spike for New York state caused by the tragic results of the combined 9-11 and the American Airlines Flight 587 (Jamaica Bay) accidents.

This information visualization can be used to investigate many questions whose answers may help plan safety measures [12]. For instance, are accidents increasing or decreasing over the years for a particular state? Do other states fluctuate in a similar way? If so, might there be a connection? Questions raised from the big picture such as displayed in Figure 3 may lead to more detailed investigations. For instance, what type of aviation is responsible for the most accidents? Are there trends showing safety problems with uncertified aircraft?

In the United States, the FAA establishes federal aviation regulations (FARs) for each type of aviation, for example, Agricultural (often crop dusters), Air Carrier (commercial scheduled aircraft), and General Aviation (often smaller, private aircraft) [13]. The accident database includes the FARs type for each accident. Using Google Public Data Explorer, it can be see that General Aviation has the most fatalities through the years. Animations show comparisons between the fatality rates among types of aviation, and how they change over the years. (See Figure A-6: "*Fatal accidents by type of aviation (flight regulations)*" in the Appendix.)

General Aviation aircraft may be certified aircraft, built by manufacturers under tight controls, and certified by the FAA, or they may be uncertified, built from a kit. Google Public Data Explorer can be used to look for trends over the years in fatality rates of uncertified airplanes used in General Aviation, and compare the numbers against certified airplanes. The information visualization for this dataset shows that while the fatality rate for certified airplanes shows a steady downward trend, the fatality rate for uncertified airplanes has stayed fairly constant. Closer investigation may lead to improved safety measures for designing and building airplanes from kits. (See Figure A-7: "*Fatal accidents: certified aircraft and homebuilt aircraft*" in the Appendix.)

2.2.2   The Google Public Data Explorer dataset

Google created the Dataset Publishing Language (DSPL), the dataset format used for all Google Public Data Explorer information displays. DSPL is an XML file with metadata that includes information about the creator and the source of the data. The DSPL XML file contains tables that link the stand-alone CSV (Comma-Separated Value) files containing the data. The data needs to be organized and aggregated ahead of time. For instance, the fatality totals for each state needed to be calculated and listed in the CSV files, at this time, Google Public Data Explorer is not designed to take a CSV file with a list of accidents with state and time information, and then sum the total values [14].

2.2.3   Using Google Public Data Explorer

Google Public Data Explorer is challenging for the first time user. The application is still in Beta, and there are only basic instructions available. There is an FAQ page, a tutorial, and a brief Developer's Guide that focus on helping a user put together a DSPL dataset. But at this time, there is little conceptual information or strategy advice about how to organize and design an information display. The DSPL forum Google started in February 2011, helps fill this gap. The forum is manned by the project engineers and supplies excellent support, usually responding to questions and problems within hours [15].

The interface supports users' ability to overview the information set, zoom in, and filter information, which are the basic user tasks described by Shneiderman in *The Eyes Have It: A Task by Data Type Taxonomy for Information Visualization* [10], but the interaction is still a bit rough and it may take a few minutes to get your bearings when exploring more complex information display sets. Google has a lot of plans for this tool, and is constantly fixing bugs and adding features. At this time the tool is a little uneven, but looks very promising.

Notes:
- For more information about the dataset prepared for Google Public Data Explorer, see "*Google Public Data Explorer: Example datasheet set*" on page 7 of the Appendix.
- To explore this information visualization online, go to:
 http://www.google.com/publicdata/
explore?ds=afo75ph633085_&ctype=l&strail=false&nselm=h&met_y=fatalities&h
l=en&dl=en#ctype=l&strail=false&nselm=h&met_y=fatalities&scale_y=lin&ind_y=
false&rdim=state&idim=state:CA:FL:NY&hl=en&dl=en

2.2.4   Using Sparklines to display time variations

Sparklines are tiny graph lines used as embedded inline word graphics, designed to replace several sentences needed to describe changes over time. Edward Tufte proposed the design and use of these "datawords" and christened them "sparklines" [16].

Table 1:   Sparklines showing variations in fatalities per state from 1990 to 2010.

Sparklines	State
	Alaska high: 55 low: 7
	Alabama high: 31 low: 2
	Arkansas high: 22 low: 4
	Arizona high: 41 low: 22
	California high: 214 low: 48
	Colorado high: 56 low: 7
	Connecticut high: 12 low: 0

The variations in the table above show the fluctuations in the number of fatalities over the last twenty years for the first five states. The blue dot marks the highs, the red dot marks the lows, the green dot is the latest number.

While sparklines are efficient and concise, and can save time and space, users should be given an initial, one-time introduction into what they mean. And if a group of sparklines are shown together, they may be misleading to the uninitiated, because the scales (Y and X axis) may differ for each sparkline. The scale needs to be fitted to each dataset so that the fluctuations of highs and lows are discernible at a quick glance. For instance, in the table above, California has many more fatal accidents in any given year than Connecticut. California's sparkline uses a Y-axis scale of 0 to 250 fatalities, while Connecticut uses a scale of 0 to 25. It would not be accurate to try and compare the two sparklines against each other, except for rate of change.

The sparklines in the table above show that while each state has its ups and downs, California, Colorado, and Alaska have the most dramatic changes, which could indicate a need for more investigation. What are the incidents behind these peaks? What was their context? What measures could be taken to help prevent the context from forming again?

2.2.5   Creating Sparklines

These sparklines were generated with a free online utility created by Joe Gregorio. The generator is simple to use. You cut and paste a comma-separated list of values into a field, then select the height in pixels, the min and max limits, colors for the markers, and the spacing between data points. The sparkline displays immediately, and you can copy it for your use [17].

2.3 Analyzing Narrative Text: An In-depth Look into Accident Factors

FAA Accident/Incident database provides a narrative description of the accidents. These narratives provide key information regarding causes and factors leading up to the accident. For our research, keywords were extracted from these narratives and grouped into categories to gain a better understanding of the accident. Based on the narratives, events from the databases were grouped into the following categories: aircraft equipment failure, inadequate maintenance or installation of the aircraft, errors made by the pilot, and errors made by the ground-support crew. An additional category holds records of accidents where the cause was undetermined for various reasons including missing wreckage or lack of vital information.

Categorizing these events in the database provides a better method of filtering the events that are otherwise lost in the huge database. In addition to categorizing, illustrating this data into visualization provides further clarity into underlying reasons for fatal accidents. Information visualization can also avoid the common threat of information overload by focusing in on the important data.

Due to the emerging interest in information visualization, many tools are now available to decipher large data sets into smaller chunks for easier comprehension. IBM's Many Eyes provides a variety of tools including ones that can translate textual data into visual interpretations.

Word Cloud is a popular tool offered by Many Eyes to display word frequencies. The word size in the resulting display increases based on the frequency of the word count. Figure 1 shows a Word Cloud illustration of the accident narratives. Without having to read much further into the details of the database, it is obvious that pilot's failures play an important factor in aviation accidents.

Phrase Net [18] is another tool that shows how words are related to each other. Using pattern matching between the words, Phrase Net shows the prominent words in the text and their association with other words.

Based on the narratives on the event reports, human error caused by pilot is the number one cause for fatal accidents. Processing these narratives into Phrase Net provides a clearer picture of how certain words relate to others forming the underlying reason for the accident. Phrase Net also allows users to choose the word sequence filter that connects one word to another. See Figure A-9: "*Phrase Net Pilot Error*" in the Appendix. to see the different patterns of keywords derived from Phrase Net using the pilot-error narratives.

Many Eyes has another tool for visualizing text called Word Tree. As the name suggests, Word Tree [19] arranges the words in a tree-like order with prominent words shown in larger font size. Visualizing the accident data using Word Tree gives an instant insight into the major causes. It almost spells out the various causes in plain text. Based on the illustrations in Word Tree (see Figure 4), it is obvious that the major factor that contributes to pilot error is failure to maintain control of the aircraft due to lack of clearance or inadequate airspeed following an intentional maneuver or during landing procedures. This visualization gives clear indication of areas that need more training and better policies. Weather-related incidents also contribute to the pilot's failure to maintain control over the aircraft. However, most weather-related accidents are due to the pilot's poor judgment on intentionally flying into adverse weather or ignoring warnings. Other factors that cause pilot error include lack of experience, inadequate preflight planning and improper fuel management.

The next major cause of aviation accidents is failure of aircraft equipment including engine failure, loss of power, and age fatigue of equipment. Based on the narratives, the factors leading to equipment failure is not always determined due to severely damaged parts. Improper maintenance and incorrect installation of parts are also leading causes for equipment failure. A small percentage of aviation accidents are also attributed to support crew including ground crew and air traffic controller.

Many Eyes tools follow Ben Shneiderman's visualization mantra of overview, zoom and filter, and details on demand [10]. The basic overview provides the general display of information as a starting point. Clicking a keyword or choosing a connecting word provides the ability to filter out data and see details of the results.

Although categorizing and visualizing the common reasons for fatal injuries gives a generalized perspective, it fails to display the complete reality regarding aviation accidents. To gain a focused perspective, additional layers were added to the visualization. Based on the aircraft category and number of seats, further classifications were made to group incidents to small, medium, and large aircraft. Helicopters, immobility airplanes, gliders, and small aircraft with less than 15 seats were categorized as small size. Aircraft with seats ranging from 15 to 100 were classified as medium size. Aircraft with more than 100 seats were categorized as a large size. After adding this additional layer, the major reasons for accidents remain the same, but interestingly, the order of the reasons for accidents changed among the different-sized aircraft.

Based on our research, overall fatal accidents are higher in smaller aircraft while larger airplanes have fewer incidents. The major cause remains to be pilot error in smaller aircraft, while most of the accidents in larger airplanes were caused by equipment failures (see Figure A-8: "*Fatal accidents by number of seats*" in the Appendix). This visualization supports the idea that smaller aircraft are operated by amateurs and pilots with less experience who are more likely to have accidents. Pilots of larger commercial aircraft have more experience and have gone through vigorous training which helps them to maintain better control over their aircraft, reducing their chances for accidents. Also, large aircraft are equipped with automation that typically controls most of the navigation and operations of air flight.

Visualization helps uncover the underlying data in large datasets. By interacting with the visualization, we are also able to obtain just the necessary information from the vast collection of data. Using visualization in the FAA database can be valuable for investigators to pinpoint leading causes for accidents, for equipment manufacturers to conduct better quality control, and for the flight instructors to provide more training for the most vulnerable areas of flying.

Figure 4: Word Tree visualization of accident narratives

2.4 Examining Natural Conditions Surrounding Accidents

The natural conditions surrounding aviation accidents between 1990 and 2010 are worthwhile considerations when assessing the cause of these accidents. While there are many factors involved, this section addresses natural conditions and their possible roles. Three factors are evaluated in determining trends related to natural causes and accidents: light conditions, wind velocity, and visibility.

2.4.1 Tools

Gephi is an interactive visualization and exploration platform for all types of networks and complex systems [20]. This tool allows users to analyze data to discover patterns and facilitate reasoning. Dynamic and hierarchical graphs, were used to apply visualization techniques to the aviation accident dataset. Exploratory analysis with this tool allows connections to be made between natural conditions and fatal aviation accidents.

Many Eyes, an experimental data exploration tool developed by IBM, was also leveraged in order to observe the dataset with different visual approaches. With its simplistic interface and uncomplicated visualizations, it is much easier to work with compared to other tools [21].

2.4.2 Findings

Light conditions play a somewhat small role in the cause of aviation accidents. When examining the data visualization in Figure 5, it can be seen that over 77% of fatal accidents occur during daylight (DAYL, blue color) hours. The second highest occurrences of fatal accidents (13%) occur during near dark (NDRK, tan color). At first consideration, it seems surprising that most accidents happen in daylight, but since most flights occur during daylight hours, it follows that those numbers are much higher. It is safe to assume that there are heavier influences on accident rates than light conditions; the information visualization confirms that low-light conditions are not a primary cause for fatal aviation accidents.

Using Many Eyes, a bubble chart was created from the dataset to examine relationships between the levels of damage (destroyed, substantial, minor, etc.) and the light conditions. (See Figure A-12: "*Relationships between aircraft damage and light conditions*" in the Appendix.) In all scenarios, regardless of light conditions, the aircraft was usually destroyed. In fact, across the board, each lighting condition column seemed to be directly proportional to one another with damage levels decreasing from destroyed to substantial to none.
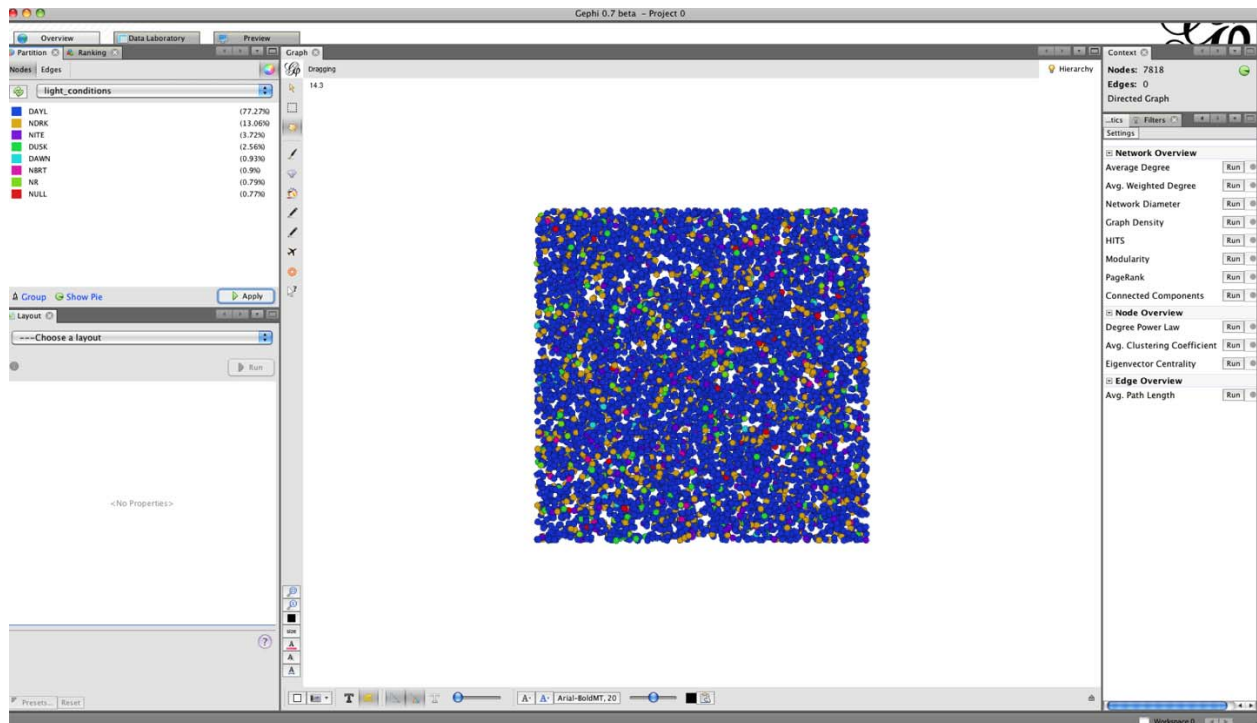

Figure 5: light conditions by percentage during fatal aviation accidents (1990-2011)

Visibility also plays a role in fatal aviation accidents, but again, the information visualization demonstrates that it is not a predominant one. Over 46% of fatal accidents occurred with 10 miles of visibility present, which is a more than reasonable amount of visibility. (See Figure A-10: "*Visibility (in statute miles) and fatal accidents*" in the Appendix.) Even fewer accidents occurred while there were less than 5 miles of visibility. In fact, only 2.29% of fatal accidents occurred with 1 mile of visibility existing. While visibility could play a contributing role to fatal aviation accidents, it is by no means a primary factor. Instrument-rated pilots (IFR) are certified to fly relying on the use of the instruments in the cockpit, and need little visibility. Pilots certified for flying under FAA visual flight rules (VFR) are not permitted to fly when the visibility is less than three miles.

Similarly, wind velocity at the time of accident seems to play a relatively small role in the overall scheme of fatal aviation accidents. In viewing the visualization of the data set, one can see that over 12%

of the accidents reported a present wind velocity of 0 knots, which is the highest percentage related to this factor. (see Figure A-11: "*Wind velocity (in knots) and fatal accidents*" in the Appendix.) This statistic demonstrates the relatively small impact that wind velocity has on accidents. In fact, as the accident percentages decrease, the wind velocity numbers increase, further supporting this point. Wind velocity would have to be extremely high to play a predominant role in fatal aircraft accidents.

2.4.3 Comments

The 2010 NALL Report (the Air Safety Institute's review of general aviation aircraft accidents occurring in 2009) supports these findings [22]. The report shows that more than 60% of fatal accidents in 2009 were pilot related. Regardless of the natural conditions present, the pilot seems to be the primary cause for accidents, although in some cases elements such as visibility, light conditions, and wind velocity were contributing factors. While some adverse factors might exist at

the time of the accident, in many cases, it was the pilots' improper responses that caused the accidents.

While Gephi was useful as an exploratory tool, I did find it to be cumbersome and unable to handle large datasets in a timely manner. I often could not determine if the visualization was being created, or if the application was hanging up. It was able to provide insights regarding the data in some cases, but I would have liked to build additional relationships using the tool. Many Eyes was very easy to use and allowed me to build some of these relationships that I was looking for. However, it also did not provide me with timely feedback while visualizations were being computed. Patience seemed to be the key when using these tools.

3 CONCLUSION

Information Visualization helps uncover the underlying data in large datasets. It can help investigators determine what influences are in common that increase the likelihood of accidents. Because there are many factors involved with aviation accidents, it is important to be able to visualize the data from different perspectives. By utilizing multiple visualization tools, analysts can double check trends, patterns, and relationships between the data being investigated, as well as assist investigators in determining the key points of the causations.

Our data shows that most accidents were caused by pilot error. While there may have been other contributing factors such as air traffic control error, improper maintenance, and natural conditions, ultimately the pilots' inability to adapt caused the fatal accidents.

Accidents have been a part of aviation history since the beginning. Advanced technology and better training has brought down the number of accidents and injuries over the past decade. However, aviation fatalities continues to be a major area that needs further studies. Using visualization, our research explored various factors related to these accidents. Visualization tools helps us to look at the data in a different perspective. Although some factors like weather are not completely avoidable, aircraft designers can use this information to build aircraft with resistant materials and build technology to better handle such situations. Additional training and stricter regulations can help tackle the human error factor of accidents. As amateur enthusiasts fill the airspace with custom built aircraft, it is important to study the cause and effects of past accidents to create a better future for aviation.

REFERENCES

[1] AOPA, Joseph T. Nall yearly reports, http://www.aopa.org/asf/publications/nall.html. 2011

[2] Air Crash and CVR Database, http://www.planecrashinfo.com/cause.htm. 2011

[3] C. Chaomei, *Information visualization: Beyond the horizon*. 2nd. Singapore: Springer, 2004.

[4] FAA, *Human Error Analysis of Commercial Aviation Accidents: Application of the Human Factors Analysis and Classification System*, http://www.hf.faa.gov/docs/508/docs/HFACS2001Jb.pdf. 2011

[5] FAA Aviation Accident Database and Synopses, http://www.ntsb.gov/ntsb/query.asp. 2011

[6] FAA Downloadable Accident Database, http://www.ntsb.gov/avdata/. 2011

[7] Tableau Public Software, *Tableau Public FAQ*, http://www.tableausoftware.com/public/faq. 2011

[8] WordItOut. *Word It Out*. http://worditout.com/. 2011.

[9] Powell, Ron, *Tableau Reveals the Benefits of Visual Analysis for Business Intelligence*, *PR Web* (2007), http://www.prweb.com/releases/Visual_Analysis/Business_Intelligence/prweb531906.htm. 2011

[10] Ben Shneiderman, The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. In *Proceedings of the IEEE Symposium on Visual Languages*, pages 336-343, Washington. IEEE Computer Society Press, 1996. http://citeseer.ist.psu.edu/409647.html

[11] Google public data explorer, *Data visualizations for a changing world*, http://www.google.com/publicdata/home. 2011

[12] Google public data explorer, *Fatal Aviation Accidents by State*, http://www.google.com/publicdata/explore?ds=afo75ph633085_&ctype=l&strail=false&nselm=h&met_y=fatalities&hl=en&dl=en#ctype=m&strail=false&nselm=s&met_s=fatalities&scale_s=lin&ind_s=false&ifdim=state&hl=en&dl=en. 2011

[13] wikipedia, *Federal Aviation Regulations*, http://en.wikipedia.org/wiki/Federal_Aviation_Regulations. 2011

[14] DSPL: Dataset Publishing Language (Labs), *DSPL FAQ*, http://code.google.com/apis/publicdata/faq.html. 2011

[15] Google Public Data Explorer, *DSPL Discussion Forum*, http://code.google.com/apis/publicdata/forum.html. 2011

[16] E. Tufte, *Beautiful Information*. Connecticut: Graphics Press LLC, 2006.

[17] J. Gregorio, *BitWorking*, http://sparklines.bitworking.info/. 2011

[18] Many Eyes, *Phrase Net Guide*, http://www-958.ibm.com/software/data/cognos/manyeyes/page/Phrase_Net.html. 2011

[19] Many Eyes, *Word Tree Guide*, http://www-958.ibm.com/software/data/cognos/manyeyes/page/Word_Tree.html. 2011

[20] Gephi, http://gephi.org. 2011

[21] Many Eyes, *Bubble Chart Guide*, http://www-958.ibm.com/software/data/cognos/manyeyes/page/Bubble_Chart.html. 2011

[22] AOPA, *2010 Joseph T. Nall Report, http://www.aopa.org/asf/publications/10nall.pdf*. 2011