AVIATION ACCIDENT ANALYSIS USING TABLEAU

NAME: Anusha Manjappa

CIN: 305851746

CIS5270 Project1:

Aviation Accident Analysis

(2012-2016)

Anusha Manjappa

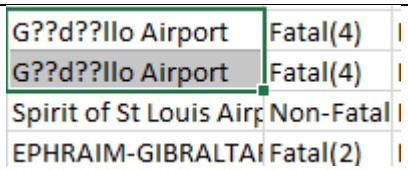
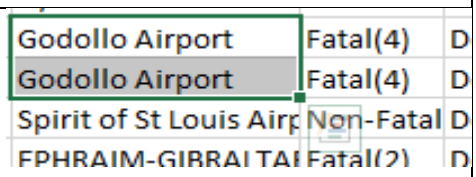
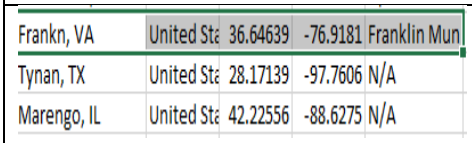
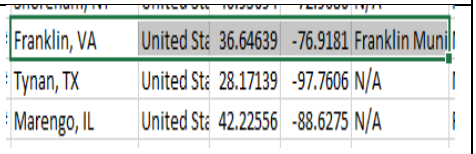
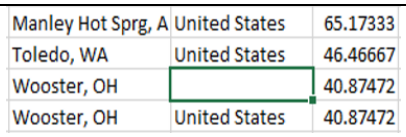
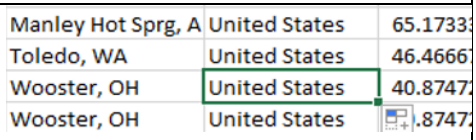
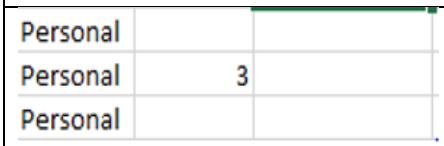
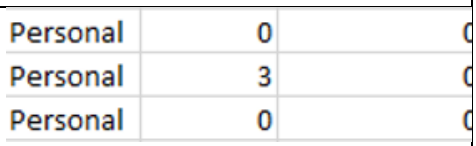
A) Dataset URL's:

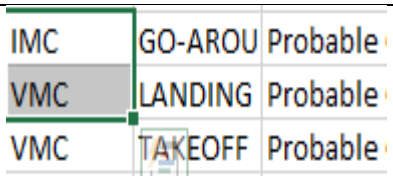
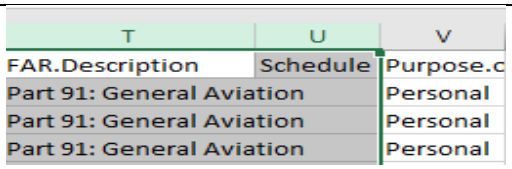
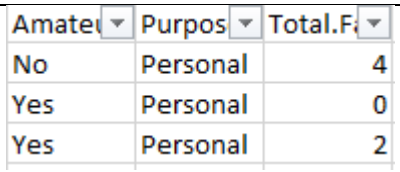
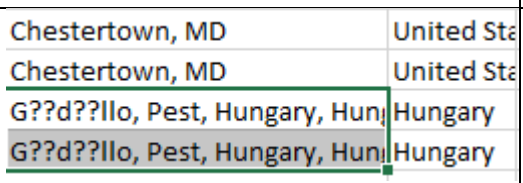
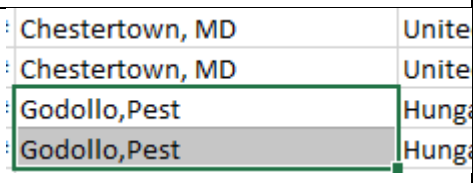
https://www.kaggle.com/khsamaha/aviation-accident-database-synopses

http://www.ntsb.gov/investigations/data

The dataset contains information from 1962 and later about civil aviation accidents and selected incidents within the United States, its territories and possessions, and in international waters. The data accumulated in the NTSB (National Transportation Safety Board) database about aviation accident provides a platform to investigate the various details of the accidents like injury severity, aircraft damage, total fatal injuries, total serious injuries, total minor injuries, the exact location n where it occurred, the weather conditions that prevailed then, etc. Due to this huge data volume and to concentrate more on the most recent incidents, I have selected a part of the dataset to analyze the incidents that have occurred from the year 2012-2016. On the whole I, have considered 24 columns and 6213 rows for my analysis for the years 2012-2016.

B) Data Cleaning:

Scope/Problem	Dirty Data	Cleaned Data/ Remarks
1. Misspellings	 The airport name contained junk values as shown.	 Corrected the error as the airport is named after the city Godollo.
	 The state field had the wrong spelling of “Franklin” in it.	 The name was corrected as shown based on the airport name in the same row.
2. Missing Values	 Missing Country name.	 Country name filled in based on the state field.
		

	In case of no injury of a particular type ,instead of putting count as 0 , it was a blank	Replaced the blank cells with the value " 0" , indicating no injuries of that type.
3. Cryptic values, Abbreviations		Contains Abbreviations for the "Weather conditions" field. UNK: Weather condition unknown VMC: Visual meteorological conditions. IMC: Instrument meteorological conditions.
4. Unwanted Columns	 The columns Far.Description and schedule were not required for the analysis.	 As shown, those columns were deleted.
5. Misfielded values		 Removed those unwanted values.

	State column contained Country names and junk values.	

C) Data Visualizations:

1) What is the count of total number of fatal, minor and serious injuries along with the total number of aviation accidents that have happened monthly in each year considered?

The count of total number of fatal, minor and serious injuries along with the total number of aviation accidents that have happened on a monthly basis in each year considered :-

Year of Event.Date	Month of Event.Date	Total.Fatal.Injuries	Total.Minor.Injuries	Total.Serious.Injuries	
2012	January	7	8	3	44
	February	18	13	14	62
	March	29	19	14	90
	April	16	23	18	93
	May	20	32	17	124
	June	30	33	17	129
	July	28	38	31	151
	August	31	28	21	143
	September	41	47	21	143
	October	33	17	16	91
	November	53	36	13	85
	December	35	22	12	79
	Total				1,234
2013	January	37	11	7	59
	February	27	17	15	70
	March	41	27	15	94
	April	33	37	20	88
	May	33	31	14	126
	June	57	55	50	158
	July	45	210	92	149
	August	51	43	29	159
	September	28	50	23	129
	October	61	22	34	114
	November	37	15	19	72
	December	35	31	9	94
	Total				1,282
2014	January	20	24	6	75
	February	32	8	18	98
	March	34	19	20	94
	April	36	40	24	108
	May	45	35	26	142
	June	63	49	21	168
	July	56	45	37	169
	August	75	71	36	157
	September	52	41	26	132
	October	42	31	26	121
	November	32	29	24	69
	December	50	13	11	63
	Total				1,366
2015	January	23	29	19	78
	February	17	15	30	62
	March	50	32	23	119
	April	44	28	11	96
	May	32	29	18	125
	June	49	34	30	144
	July	50	52	36	160
	August	63	57	39	160
	September	44	36	40	130
	October	63	13	30	115
	November	33	20	21	84
	December	37	20	18	72
	Total				1,345
2016	January	24	15	21	76
	February	49	22	11	88
	March	27	18	17	84
	April	34	33	23	105
	May	43	44	13	128
	June	35	45	21	156
	July	51	49	18	133
	August	52	48	26	141
	September	56	32	24	113
	October	29	19	13	91
	November	19	11	9	47
	December	24	5	5	29
	Total				1,171

[Tools used: Calculated fields, Dates, Totals]

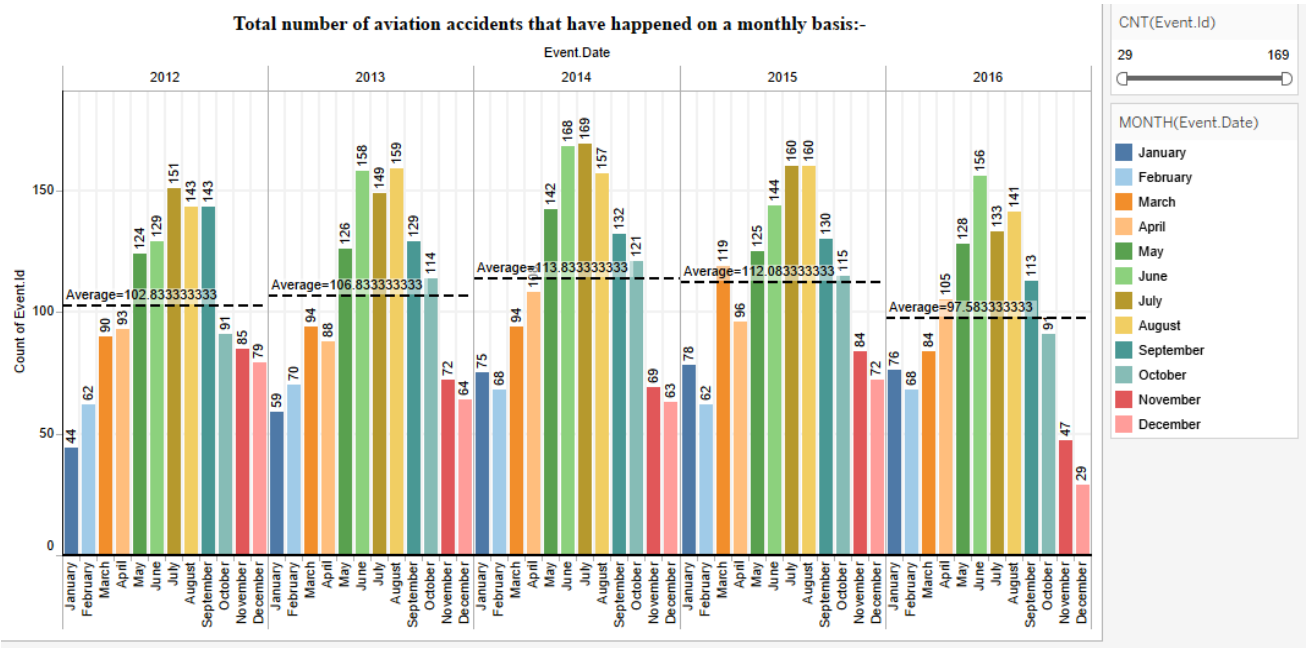
The above text tables depict the following measures for each month of the years from 2012-2016.

- Total Number of Fatal Injuries
- Total Number of Minor Injuries
- Total Number of Serious Injuries

This helps us know the severity of each incident that has occurred. It also consists of the count of the total number of accidents happened in each month of all the years from 2012-2016 and the total number of accidents per year. The total number of accidents per year has been calculated with the help of a calculated field, Number_Of_Events= COUNT ([Event.Id]). This returns the count

of event ids which is a unique attribute on a monthly basis and totals it to get the number of accidents per year.

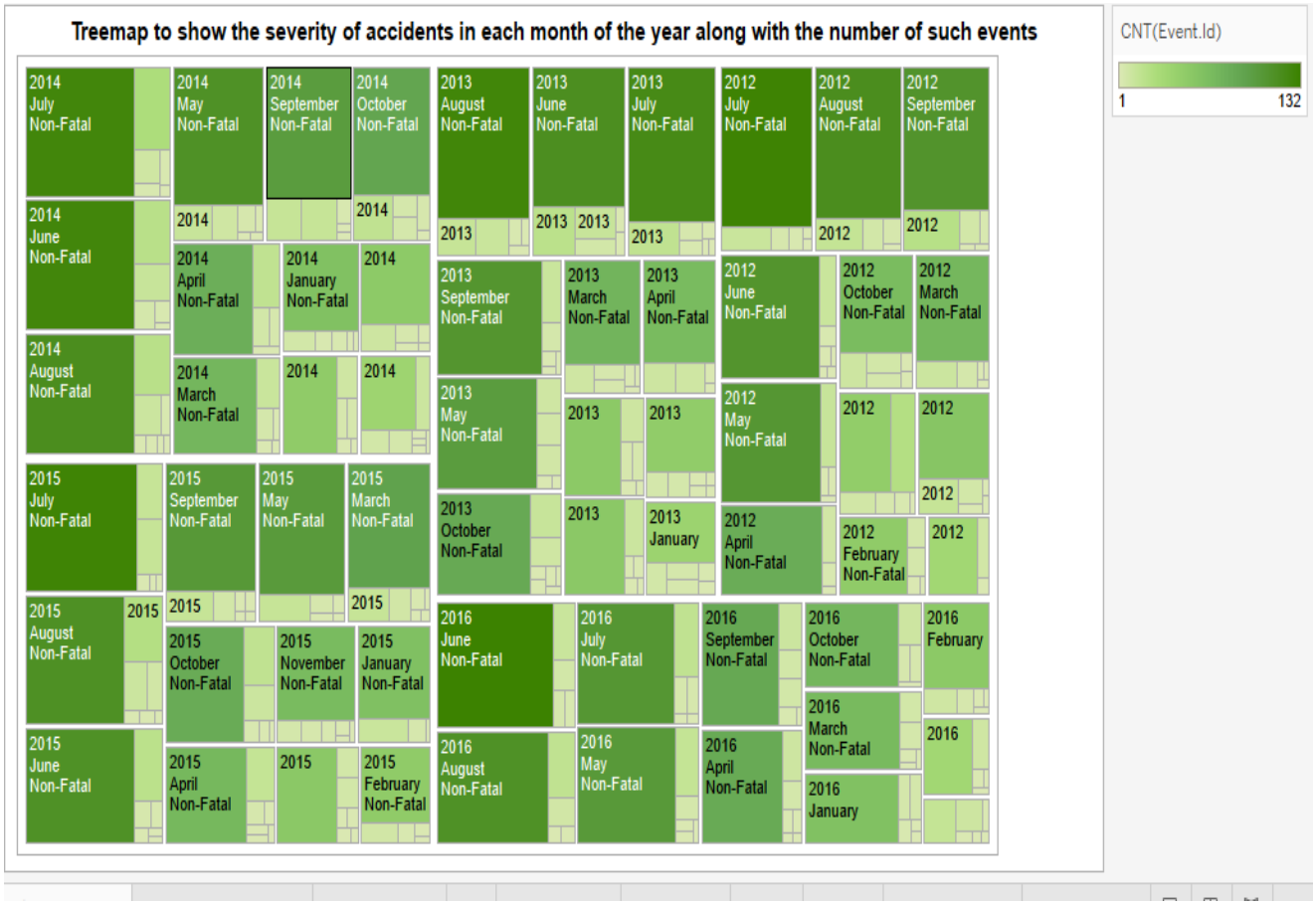
2) **Analysis of total number of aviation accidents that have happened on every month of the considered years and the average of the number of accidents per year.**



[Tools used: Calculated fields, Dates, Reference Line]

The above analysis using a side-by-side bar chart shows the count of the total number of aviation accidents that have happened per month of all the years considered with the help of a calculated field, Number_Of_Events= COUNT ([Event.Id]). The different months are all displayed in different colors as shown in the filter label. The average of the number of events per year has also been displayed using the reference line to have a rough idea about the number accidents happening. The average number of accidents was the highest in 2014 and lowest in 2016 which says that there is a considerable decrease in the number of accidents. It can also be seen that the lowest number of accidents was in Dec 2016 where there were just 29 accidents and the highest was in July 2014 where there were 169 incidents which were reported.

3) **Analysis of total number of aviation accidents that have happened on every month of the considered years and the average of the number of accidents per year.**

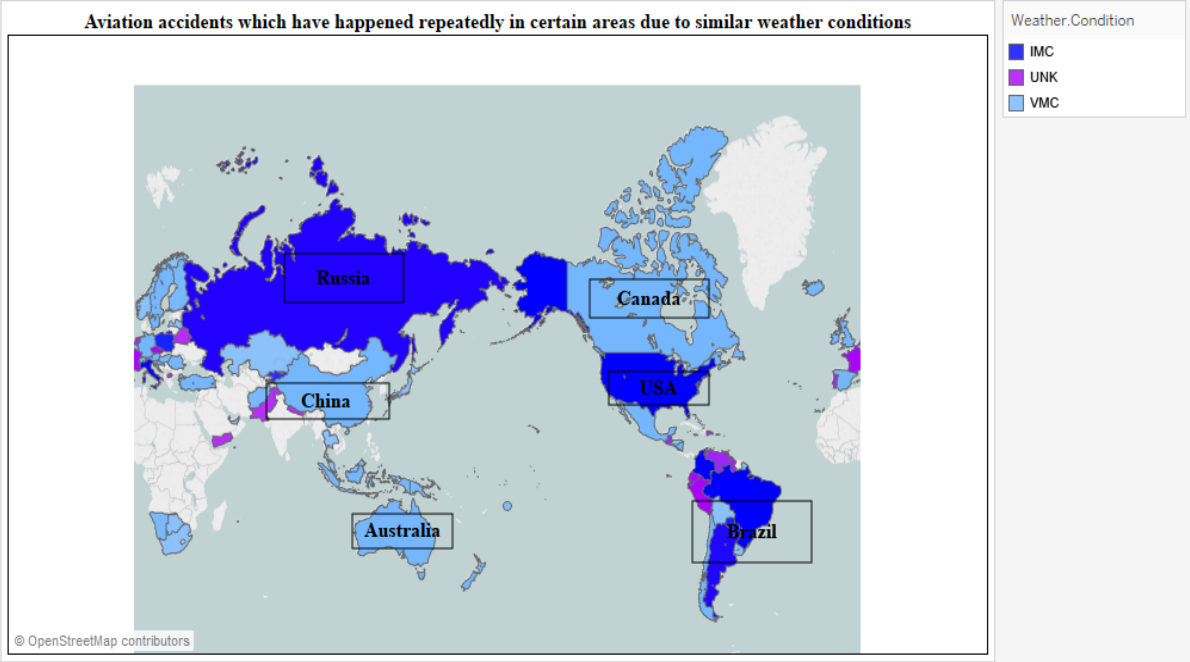


[Tools used: Calculated fields, Dates, Tree Maps]

I have demonstrated the severity of each incident in the form of a tree map here. The ones in dark color are the non-fatal accidents. The color goes lighter as the fatality increases. I have also demonstrated the number of such events that have occurred and that can be distinguished by the size of the blocks in the tree-map. More the number of incidents, bigger is the block of the tree-map. I have also labeled the respective years on them to make it more understandable. The number

of such events that have occurred is again calculated by a calculated field which counts the number of such events. This helps us analyze the severity of the incidents and to see if they have reduced over time.

4) Which are the aviation accidents which have happened repeatedly in certain areas due to similar weather conditions?



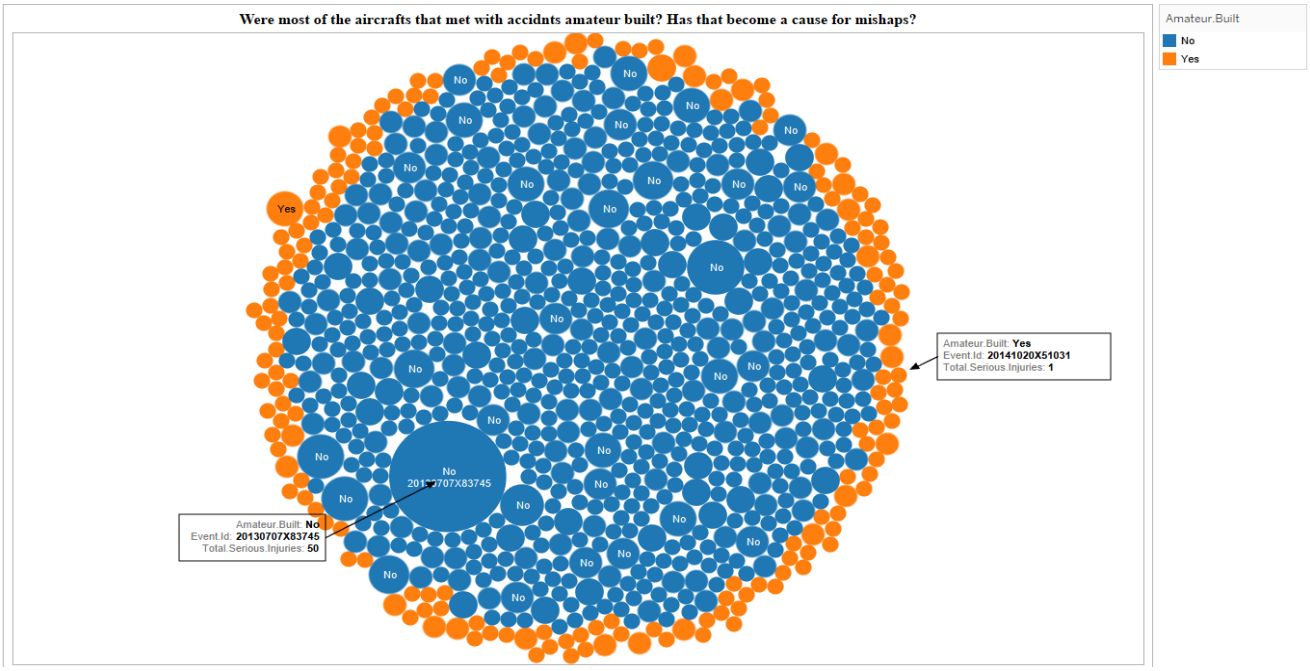
[Tools used: Geographic map]

This visualization illustrates the areas where accidents have happened due to similar weather conditions. As seen in the area in and around Russia, US, Brazil and Argentina the failure of the aircraft was mainly due to the weather condition of IMC, this describes the weather conditions that require pilots to fly primarily by reference to instruments, and therefore under Instrument Flight Rules (IFR), rather than by outside visual references under Visual Flight Rules (VFR). Therefore, pilots must be more careful in such situations as they do not have outside visual references. In areas like Canada, Australia and China accidents have occurred mostly in the weather condition

of VMC. It also describes the severity of accidents which is found to be fatal as pointed out in the diagram. The weather conditions are displayed in different colors for a better understanding.

5) **Were most of the aircrafts that met with accidents amateur built? Is it a huge number when compared to the aircrafts that were not amateur built? Has that become a cause for mishaps?**

a) Packed Bubble Chart

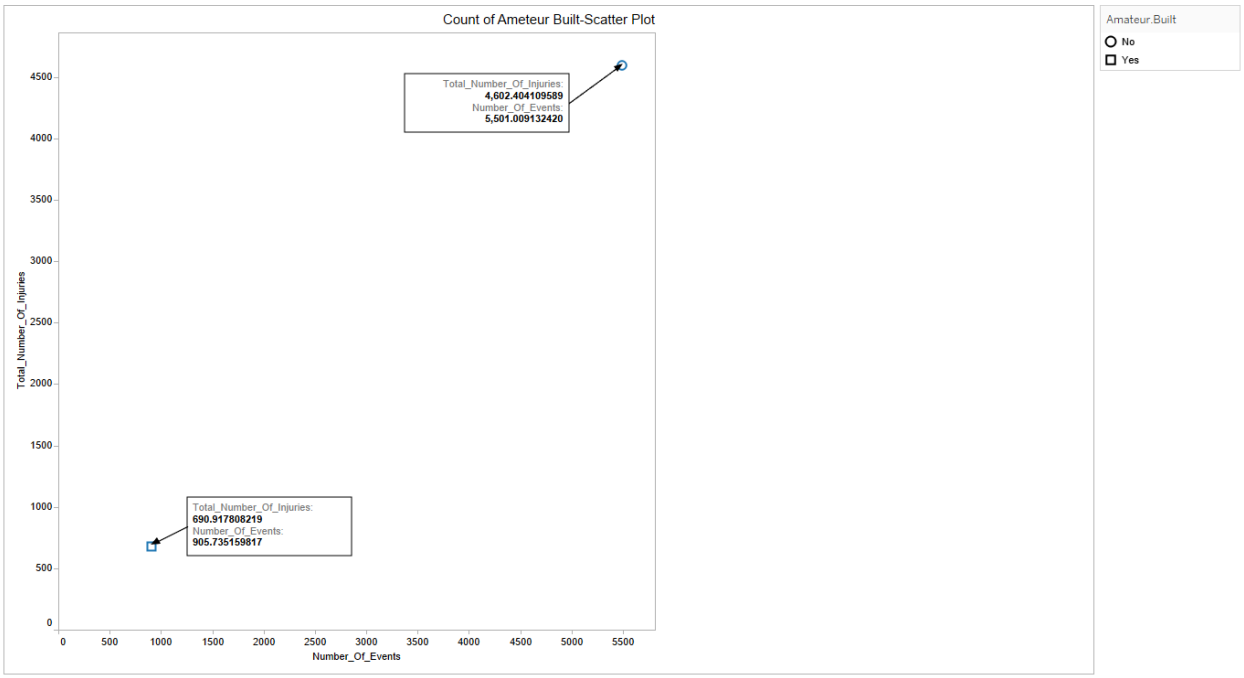


[Tools used: Packed Bubbles Chart]

Amateur-built aircraft can be built from scratch, based on original or established designs, or from a kit. If a kit is used, the builder must have completed most of the build for it to be considered an amateur-built aircraft. Experimental and amateur-built aircraft include any aircraft that is fabricated and assembled by anyone for their own education and/or recreation. During the analysis process, it was found that the number of amateur built aircrafts that failed were very less in number

as shown in the figure. The ones shown in orange are all failures of amateur built aircrafts and the ones in blue were not built by amateurs. Along with the separate colors used, I have also labeled the total number of serious injuries that occurred due to that mishap. The sizes of the bubbles vary per number of serious mishaps that have occurred. As seen, I have pin-pointed the smallest and biggest bubble which indicate 1 serious injury and 50 serious injuries respectively. Thus, it can be inferred that organizations are taking good protective measures about the amateur built aircrafts and that is not a reason for the accidents that have occurred.

b) Scatter Plot Chart

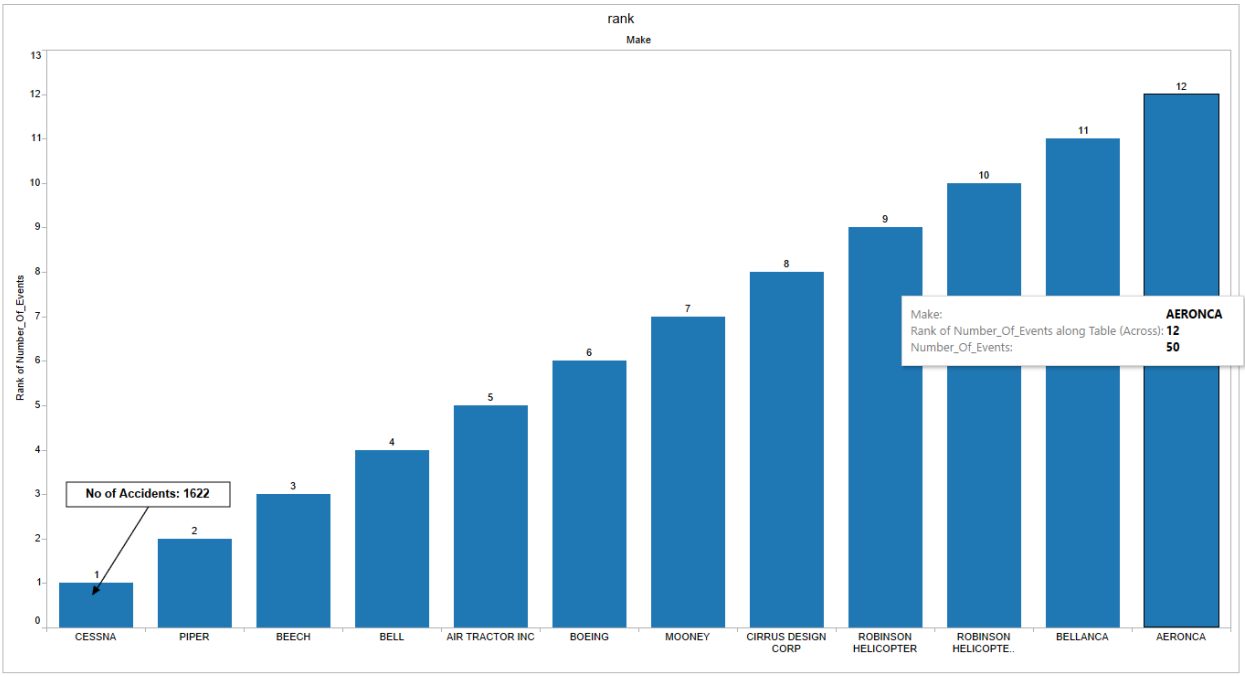


[Tools used: Scatter Plot, Calculated Fields]

In addition to the above bubble chart this above visualization uses two calculated fields, Number_Of_Events and Total_Number_Of_Injuries in a scatter plot to find out the exact count of accidents by amateur built aircrafts which is 905 and the total number of injuries which is 609. This is a very small amount in comparison to the accidents and injuries of the non-amateur built

aircrafts which is 5501 and 4602 respectively. Therefore, once again we can infer that it has not been a cause for the huge number of mishaps.

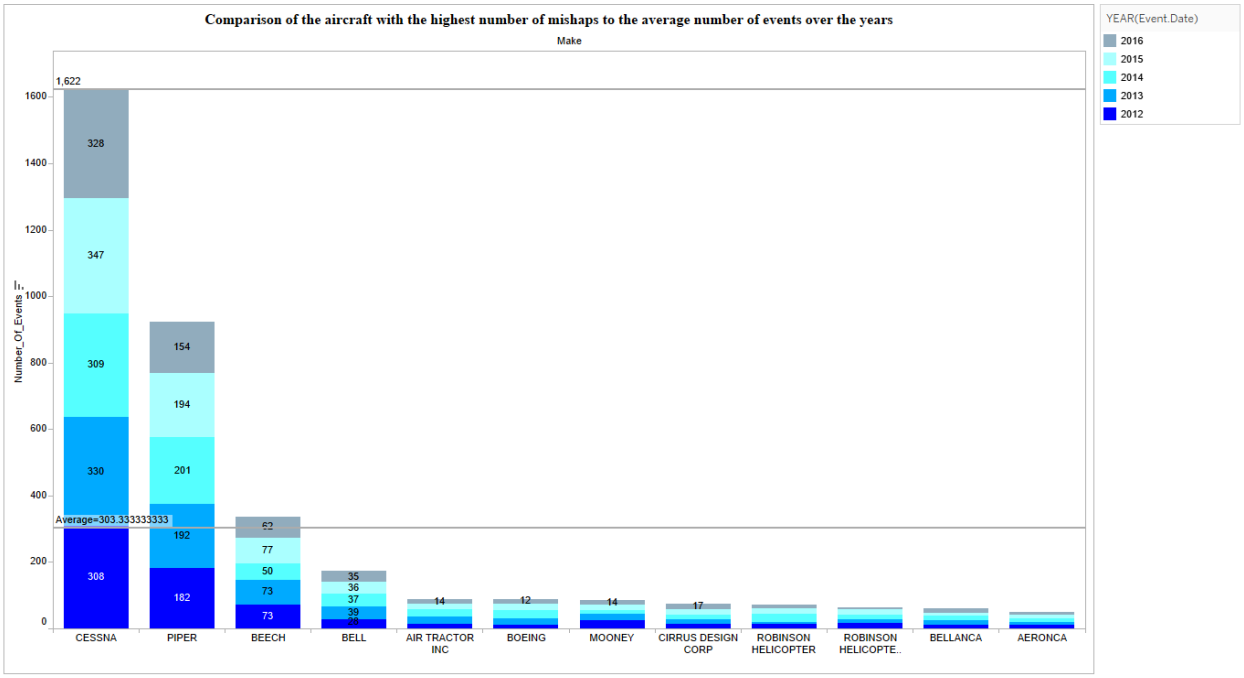
6) Which airline has experienced the most number of failures? Which are the top most airlines going through many failures?



[Tools used: Calculated Fields, Rank]

This Visualization shows the Top airline makes which have accounted for maximum number of airline disasters. I have set a filter to get the Airline makes who have accounted to about 50 and above accidents and ranked. The aircraft maker CESSNA has had a total of 1622 mishaps in the last 5 years itself and it has been highlighted in the above visualization. Next is the line is PIPER with a total number of 923 accidents. I have used a calculated field Number_Of_Events to calculated the number of accidents per aircraft make. This clearly explains the aircraft make to be chosen and not chosen in the future.

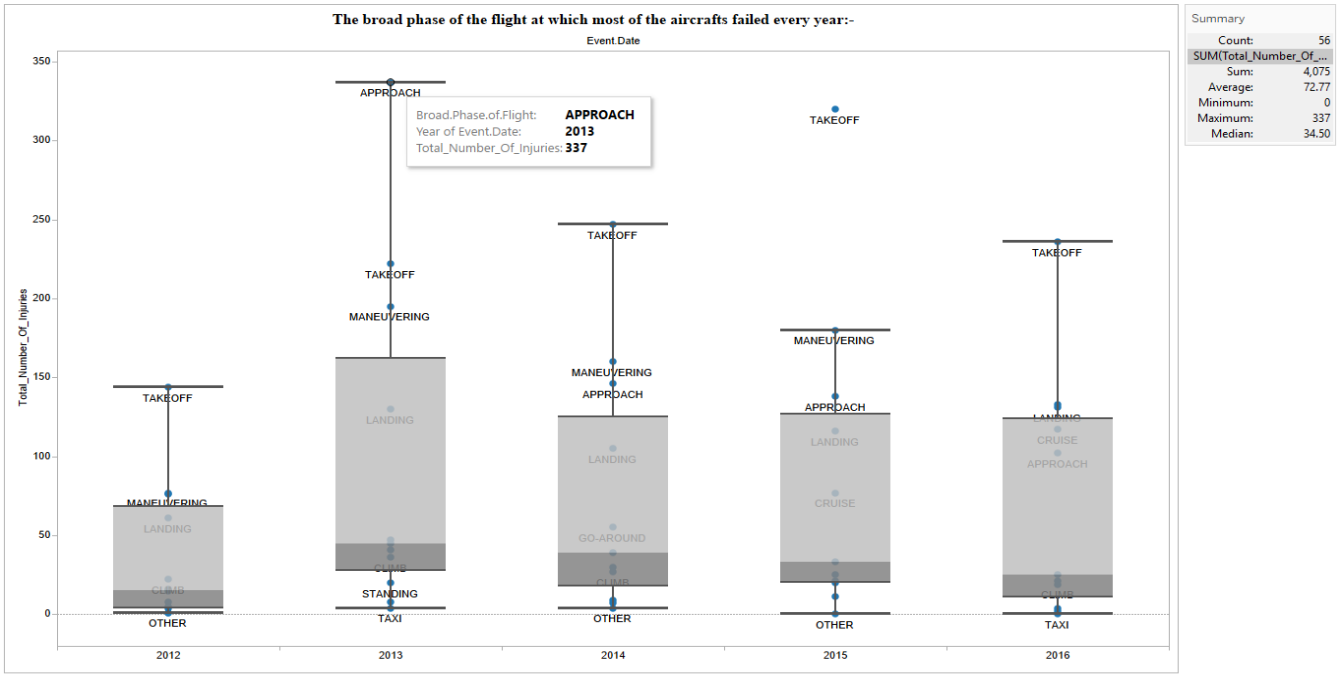
7) **Comparison of the aircrafts with the highest number of mishaps to the average number of events over the years**



[Tools used: Calculated Fields, Dates, Reference Line]

In this Visualization, in addition to the Top airline makes which have accounted for maximum number of airline disasters, I have shown the year wise split up of the number of disasters with different colors for better analysis. The make CESSNA has experienced the most number of disasters which is a total of 1622 as shown by the reference line. This is now split up as 328 disasters in 2016,347 in 2015, 309 in 2014, 330 in 2013 and 308 in 2012. Similar kind of break up has been shown for all other aircraft makes also. This has also been ordered in descending order for better understanding. A reference line indicates that the average number of mishaps for the last 5 years has been 303 and there are three aircraft makes who have experienced failures way beyond this value. Thus, measures can be taken to avoid these aircraft makers.

8) The broad phase of the flight at which most of the aircrafts failed every year.

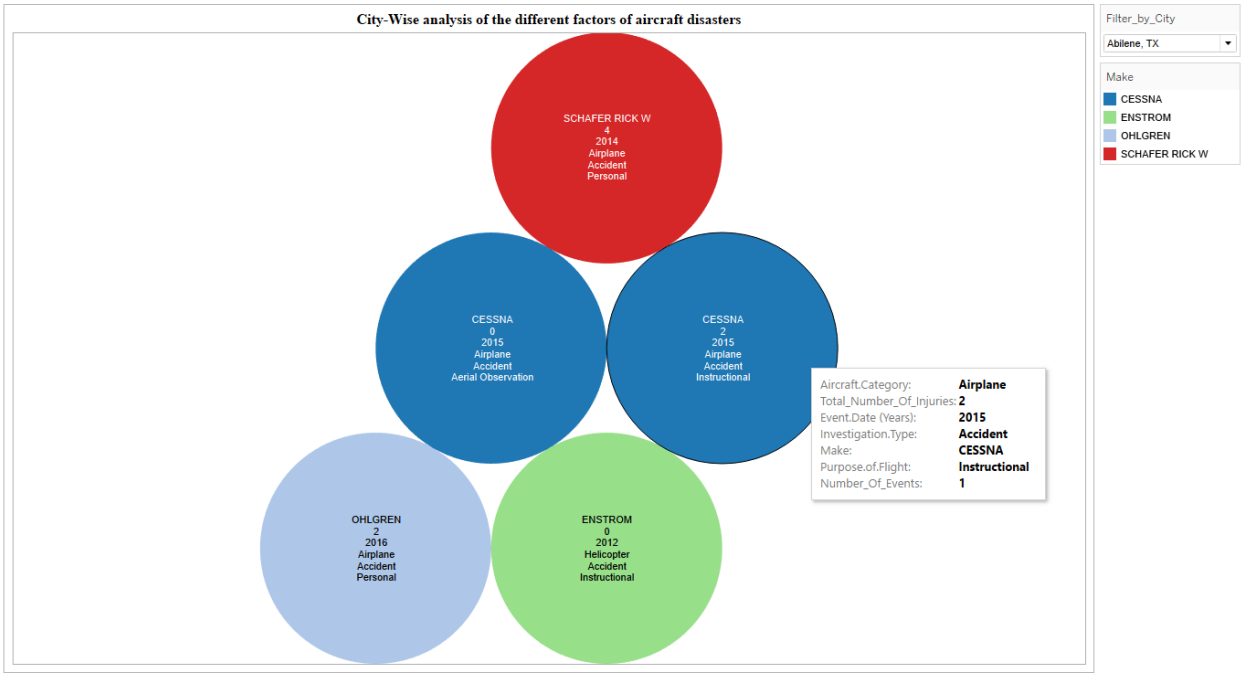


[Tools used: Calculated Fields, Dates, Box & Whisker plot]

Broad phase of flight refers to a period within a flight at which the accident happened. It helps them to investigate the cause of the accident and to see what should be done to avoid this in the future. As shown in the visualization, the highest number of injuries, i.e. 337 in the year 2013 have occurred at the "Approach" broad phase of the flight. The second highest in row with 320 injuries in the year 2015 stands the "Take off" broad phase of the flight. The total number of injuries has been got from a calculated field Total_Number_Of_Injuries= [Total.Fatal.Injuries]+ [Total.Minor.Injuries]+ [Total.Serious.Injuries]. The boxplot is a convenient way of graphically depicting groups of the total injuries through their quartiles. The lines extending vertically from the boxes (whiskers) indicates variability outside the upper and lower quartiles. The bottom and top of the box are the first and third quartiles, and the band inside the box is the second quartile

(the median). The entire summary of every plot is listed out on the right in the summary plot when the cursor is pointed to the respective observation.

9) City-Wise analysis of the different factors of aircraft disasters.

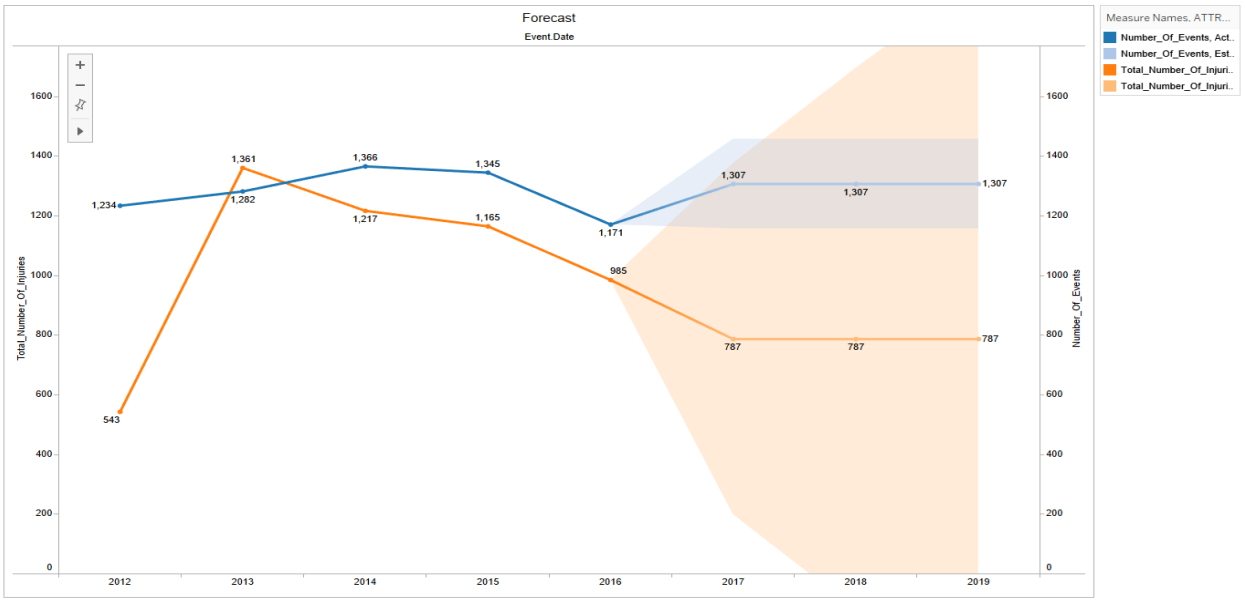


[Tools used: Calculated Fields, Parameters, Packed Bubbles chart]

This Visualization shows City wise details of all the airplane disasters like Aircraft Category, Total Number of Injuries, Year of the Event, Investigation Type, Make of flight, Purpose of flight and number of accidents. All these fields are again grouped per the make of aircrafts and put in different colors. Calculated fields are used to calculate the total number of injuries and total number of accidents. I have created parameters Filter_By_City and Filter_For_Selecting_City to select the state and city we want to analyze. For instance, in the above visualization when pointed at one of

those blue bubbles, it says that an airplane made by CESSNA met with 1 accident in the year 2015 with one injury and the purpose of that flight was instructional.

10) A forecast on the number of injuries and total number of accidents in the upcoming years.

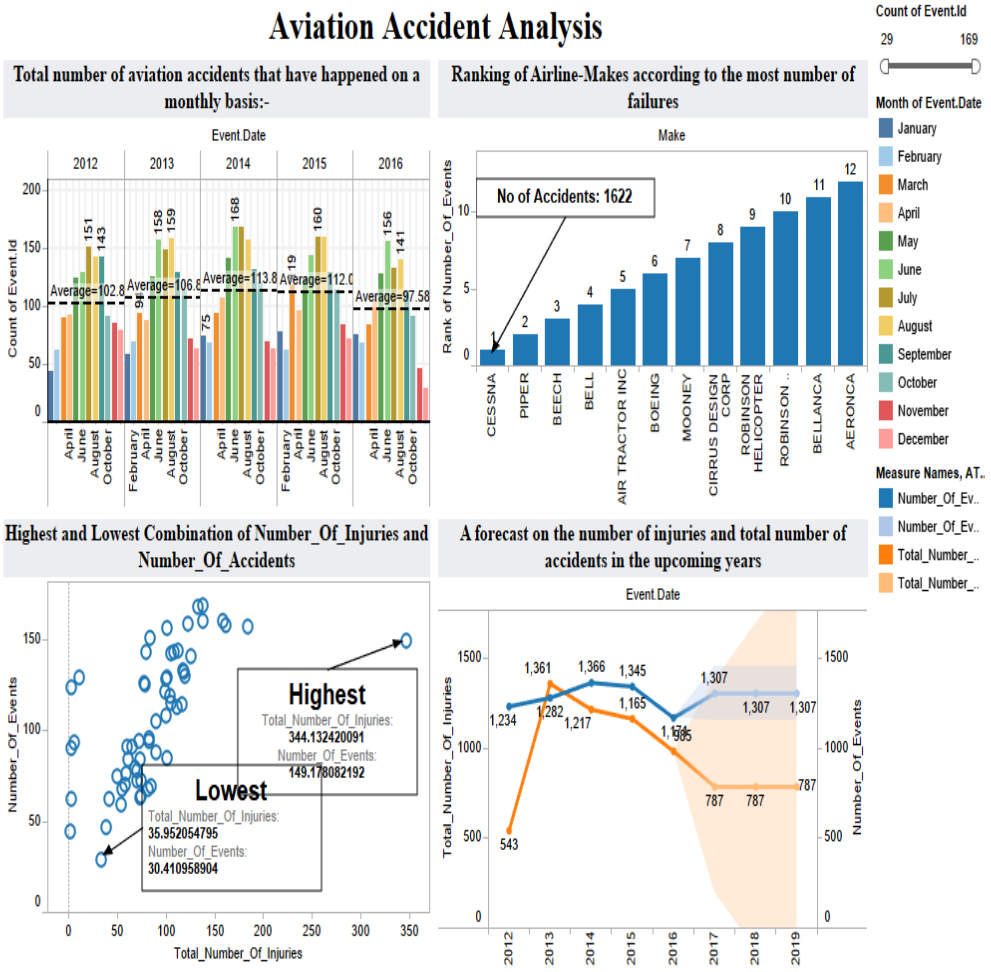


[Tools used: Calculated Fields, Dates, Forecast Trend Lines, Dual Axis Chart]

This Visualization shows the predicted values of the fields Total_Number_Of_Injuries and Total_Number_of_Accidents for the upcoming 3 years. The field Total_Number_Of_Injuries is a calculated field got by adding fields [Total.Fatal.Injuries], [Total.Minor.Injuries], [Total.Serious.Injuries]. The field Number_of_events which gives a count of all the accidents is also a calculated field got by the expression COUNT([Event.Id]). Total_Number_Of_Injuries field is on the left y-axis and is given in Orange. Number_of_events is on the Right y-axis and is given in Blue. The forecast lines show that the estimated number of flight accidents could be around 1307 overall which is slightly lesser than the number of accidents that have happened in the

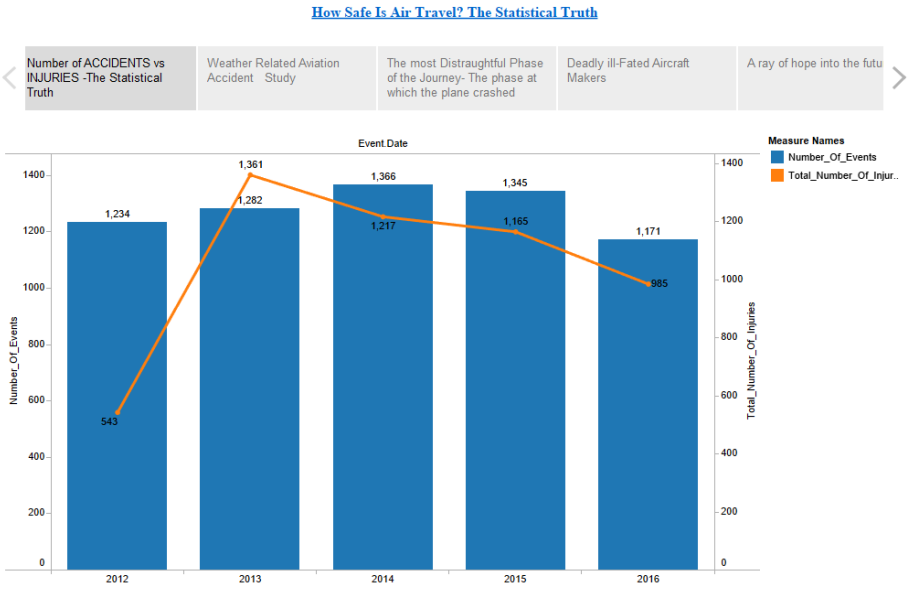
previous years. The number of injuries in the future years has been forecasted to be 787, which is reasonable when compared to the injuries happened in the past which was more than 1000 in most of the cases.

D) Dashboard:


Aviation Accident Analysis
Total number of aviation accidents that have happened on a monthly basis:-
Event.Date
2012
2013
2014
2015
2016
Count of Event.Id
Average=102.8
Average=106.6
Average=113.8
Average=112.0
Average=97.58
April
June
August
October
February
Ranking of Airline-Makes according to the most number of failures
Make
Rank of Number_Of_Events
No of Accidents: 1622
CESSNA
PIPER
BEECH
BELL
AIR TRACTOR INC
BOEING
MOONEY
CIRRUS DESIGN CORP
ROBINSON HELICOPTER
ROBINSON ..
BELLANCA
AERONCA
Count of Event.Id
29
169
Month of Event.Date
January
February
March
April
May
June
July
August
September
October
November
December
Measure Names, AT..
Number_Of_Ev..
Number_Of_Ev..
Total_Number_..
Total_Number_..
Highest and Lowest Combination of Number_Of_Injuries and Number_Of_Accidents
Number_Of_Events
Total_Number_Of_Injuries
Highest
Total_Number_Of_Injuries:
344.132420091
Number_Of_Events:
149.178082192
Lowest
Total_Number_Of_Injuries:
35.952054795
Number_Of_Events:
30.410958904
A forecast on the number of injuries and total number of accidents in the upcoming years
Event.Date
Total_Number_Of_Injuries
Number_Of_Events
2012
2013
2014
2015
2016
2017
2018
2019

E) Story telling:

Flights have been one of the most preferred ways of travel for long and short distances these days. When you can reach a place nearly 400miles away in about roughly just 1.5hours by flight while you would need about 12 hours by road, anyone would prefer the airways. It's easy to think of airplanes as a mode of transportation, but the aircraft itself is just one part of a wider air transportation system that includes airline operations, maintenance, crews, airports, security, weather services, and the air traffic control system. So, when we consider the safety of air travel, we must look separately at different countries and regions.



The number of fatalities attributed to air travel has been declining rapidly over the past years. Considering that the number of flights and passenger miles has increased nearly as quickly, the actual rate of fatalities continues to fall dramatically, by any measure. There are now approximately three times as many passenger miles flown as there were 30 years ago, but about one half the number of fatalities, on average [1].

According to the data from NTSB, there were 543 injuries from 1234 accidents in 2012 out of which 316 where minor injuries which shows that there were not many serious accidents this year although the count of accidents was high. 2013 produced 1361 injuries in 1282 accidents out of which there was a total of 485 fatal injuries,549 minor and 327 serious injuries. 2014 had 1217 injuries worldwide from 1366 accidents where there was a total of 537 fatal injuries,403 minor and 277 serious injuries. 2015 had 1165 injuries worldwide from 1345 accidents where there was a total of 537 fatal injuries,403 minor and 277 serious injuries and the most recent 2016 had 985 injuries from 1171 accidents. So yes, 2014 was a bad year in terms of total accidents and total number of fatalities too. The last most fatal airline accident in the United States was Asiana Airlines Flight 214 which crashed in San Francisco on July 6, 2013. It resulted in three fatalities out of 307 passengers and crew on board. Prior to that, the last accident was the February 12, 2009 crash of Colgan Air Flight 3407 near Buffalo, New York, which resulted in the deaths of all 50 people on board [2].

Experts follow and determine what went wrong and work out how to prevent these horrific tragedies from happening again. They examine the wrecks and official records, and hear from eyewitnesses, passengers and aviation experts as they reconstruct some of the most tragic disasters in aviation history. Air Crash Investigation looks at what went wrong and how future disasters can be averted [3]. One of the most important factor they consider is the weather condition that prevailed during the accident. In aviation, a visual meteorological condition (or VMC) is an aviation flight category in which visual flight rules (VFR) flight is permitted—that is, conditions in which pilots have sufficient visibility to fly the aircraft maintaining visual separation from terrain and other aircraft. Whereas IMC describes weather conditions that require pilots to fly primarily by reference to instruments, and therefore under Instrument Flight Rules (IFR), rather

than by outside visual references under Visual Flight Rules (VFR) [4]. In countries like Russia, US, Brazil and Argentina the failure of the aircraft was mainly due to the weather condition of IMC, this describes the weather conditions that require pilots to fly primarily by reference to instruments, rather than by outside visual references under Visual Flight Rules (VFR). Therefore, pilots must be more careful in such situations as they do not have outside visual references. In countries like Canada, Australia and China accidents have occurred mostly in the weather condition of VMC.

Another important factor considered for investigation is the broad-phase-of-flight. The entire journey can be put into six phases, namely Taxi, takeoff, climb, cruise, descent and final approach/landing. Almost half of all accidents occur during the final approach and landing stages. Fatal accidents are also likely to occur during the climbing stage. If the aircraft left the gate with undetected faults, these may become apparent during the climb, as the first stage taking place off the ground, and could prove dangerous. Most accidents and fatalities take place during the departure (take off or climb) and arrival (approach or landing) stages. During these phases aircraft are close to the ground and in a more vulnerable configuration than during other flight phases. Accordingly, in the past 5 years, most of the fatal flight injuries have happened at the “Approach” phase of flight. We have a total of 337 fatal injuries accounted at this phase.

One important measure taken to avoid aviation accidents is to avoid the aircraft makers having the most number of mishaps or to make them improve their standards. In the past five years, it was found that the aircraft maker CESSNA has seen 1622 accidents. PIPER stands second with 923 accidents. With all these analysis, we have finally predicted the trend of the number of accidents and the number of injuries in the future years to come. There is a considerable decrease in the number of injuries to 787 injuries per year but the number of accidents hasn’t seen much of

a dip in the count. With all these analysis, investigations and predictions hopefully better measures are taken to avoid airplane mishaps in the future. Flying is one of the safest forms of transport. But what happens when tragedy strikes? From human error and accidents to mechanical faults and design flaws, the success of aviation history is punctuated with disaster and catastrophe. It's rare, but it does happen.

References:

[1]. NTSB.” History of The National Transportation Safety Board” national transportation safety board, https://www.ntsb.gov/about/history/Pages/default.aspx

[2]. The Points Guy “How Safe Is Air Travel? The Statistical Truth”, https://thepointsguy.com/2015/02/how-safe-is-air-travel-the-statistical-truth/

[3]. National Geographic, “Air Crash Investigation”, http://www.natgeotv.com/uk/shows/natgeo/air-crash-investigation

[4]. Weather-Related Aviation Accident Study, http://www.asias.faa.gov/i/2003-2007weatherrelatedaviationaccidentstudy.pdf