# Cluster Analysis on Survival Rate on Air Disaster

27/04/2022

# I. Overview

Over 100 years, a large number of airplanes crashed due to different causes. As time and technology arise, people put more effort into more complex means of transportation. All sources of transportation include ground, water, and air for the most common uses for people to travel from place to place. One concerning question that appears to be most eye-catching is how safe is the transportation we choose? Undouptfully, transportation disasters are happening every second around the world, and the most fatal with the lowest survival rate post-disaster is traveling by aircraft. The safety of mankind while in the air requires high attention and precautions; especially for those large airplane companies and airports that have the responsibility of ensuring people's safety while in the air. Therefore, comparing the data from the twenty century to the twenty-first century, is there a rising number of airplane crashes due to more people traveling or decreasing number of airplane crashes due to technological improvement with more advanced aircraft manufacturers? These data and analysis on a variety of facts and factors that cause airplane crashes can be an asset to airports, aerospace companies, governments seeking to improve the passenger's air-safeness, own reputations, and decrease the number of deaths caused by air crushes to wider people.

# II. Research Question

What are the similarities and differences between airplane crashes with high survival rates and crashes with high fatalities? The segments will then be explored to discover the possible differences based on the properties of aircraft (Aircraft model, Capacity) or causes of crashes(Damage Type, incident category).

Statistical Hypothesis:

$$H0 : \kappa = 1 vs. H1 : \kappa = k$$

where $\kappa$ denotes the number of clusters present in the dataset.

H0: All factors of aircraft have same effect on survival rate.

H1: Different factors of aircraft have different effect on survival rate.

# III. Data Cleaning Summary

## A. Data Source:

We will use Aircraft Accidents, Failures & Hijacks Dataset from Kaggle that shows worldwide accidents and hijackings involving passenger airliners, corporate jets and military transport aircraft since 1919. Airliners are considered here aircraft that are capable of carrying at least 12 passengers. The original dataset is from the ASN Safety Database, updated daily, contains descriptions of over airliner, military transport category aircraft and corporate jet aircraft safety occurrences since 1919.

In the raw dataset, there are 23519 incidents recorded in the dataset and 23 variables including incident information, descriptions of aircraft, and information about onboard passengers. The detail of the variables is shown in References.

## B. Data Preparations:

In order to make the data more suitable for analysis, the following data cleaning procedures were made.

### 1. Missing Values:

By glancing the dataset, there are many missing values displaying as ‘?’, ‘-’ and ’ ’. After filtering these out, we replace these values with NA. According to Table 4: Missing Variables in reference, a few variables have more than 50% missing values: “Time”, “Aircraft_Engines”, “Ground_Casualties”, “Collision_Casualties”.

I only removed Aircraft_Engines and Time, because we thought the variables "Ground_Casualties" and "Collision_Casualties" record casualties causing by hitting the ground and collision. Although there are only 1% incidents causing casualties on ground and by collision, we cannot ignore these fatalities.

Moreover, since clustering algorithms use data on all variables, a missing observation on one variable will cause the entire row of data to be ignored for analysis. Therefore, it is important to impute missing values. We are going to make use of the mice package with the default method, predictive mean matching.

After imputing the missing values, there are still some character variables which are not able to impute, so we decided to omit these data for further analysis.

**2. Keeping useful variables:**

We deleted eleven variables that are not relevant to this research question. The rest of the variables (Aircraft_Damage_Type, Aircraft_Model, Aircraft_Operator, Incident_Category, etc.), help to analyze plan crashes, and compare segments of different crashes with different survival rates in order to generate any similar patterns.

**3. Changing Data Types:**

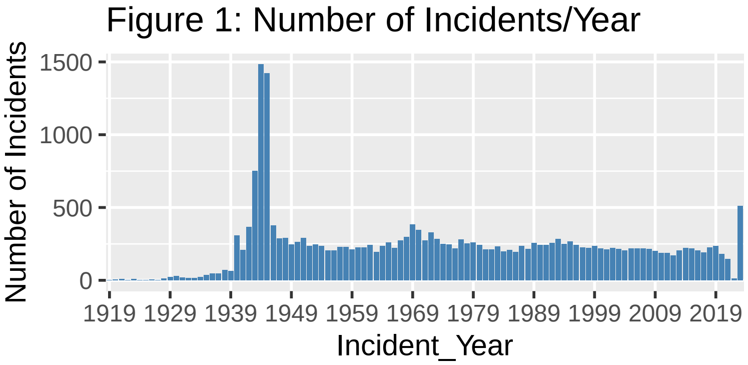
For variables Collision_Casualties, Ground_Casualties, Onboard_Total, these character data are actually describing number, so we extract the number out and change them to numeric data. For NA, we changed to 0. For incident_date, we changed character to date format.

**4. Adding more variables:**

Since our research question is to investigate the survival rate, we create survival rate by calculating the ratio of total survivors(total occupants - total onboard fatalities) and total occupants. In addition, we calculate the total fatalities which were resulted by hitting the ground or collision. However, these variables also contain missing value, we decided to impute them with mean.

**5. Remove outliers**

Figure 1: Number of Incidents/Year shows the number of incidents occurred per year from 1919 to 2022. As we can see from the figure, despite that there is a peak around Year 1943 to Year 1947, the number of incidents happened each year are similar. During the period of Year 1943- Year 1947, it is well know that this was the period of World War II, so that this time the number of aircraft crashes significantly differs from other years. Since outliers may cause variability or indicate experimental error, we decide to exclude the outliers for further analysis. In total, we remove 29 years of data, which are considered as outliers based on the boxplot.



## C. Variables:

After data cleaning process, the cleaned dataset consists of 11639 observations, and 19 variables, as shown in the table below.

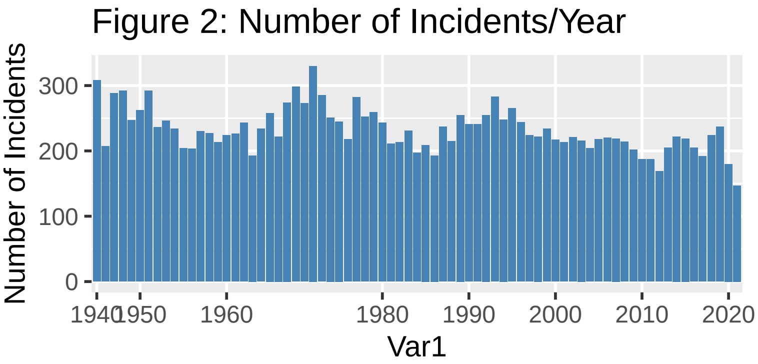
Table 1: Variables

| | | | | |
|---|---|---|---|---|
| Aircaft_Damage_Type | Aircaft_Model | Aircaft_Operator | Aircraft_Phase | Collision_Casualties |
| Departure_Airport | Destination_Airport | Fatalities | Fatalities_Onboard | Ground_Casualties |
| Incident_Category | Incident_Date | Incident_Location | Incident_Month | Incident_Year |
| survival_rate | Survivors_Onborad | Total_casualities | Total_Occupants | Aircaft_Damage_Type |

## D. Data Visualization:

The overall aircraft crash data plots are provided below.

Figure 2: Number of Incidents/Year shows the number of incidents occurred per year from 1940 to 2021. As we can see from the figure, after removing the outliers, the number of incidents happened each year are similar. However, with increasing transportation needs, the number of flights are increasing dramatically from 1940 to 2021, which indicates that ratio of incidents may drop dramatically over time.



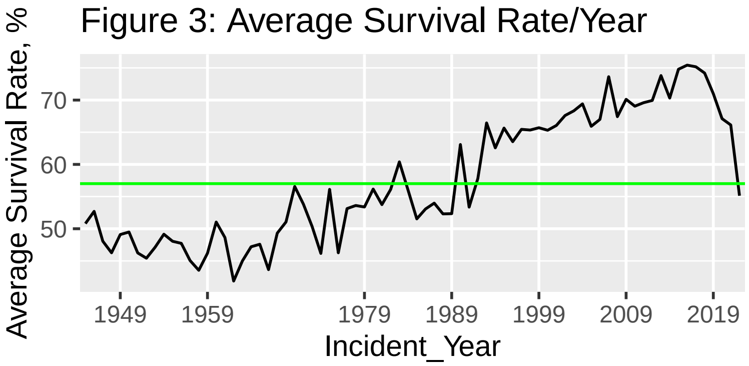
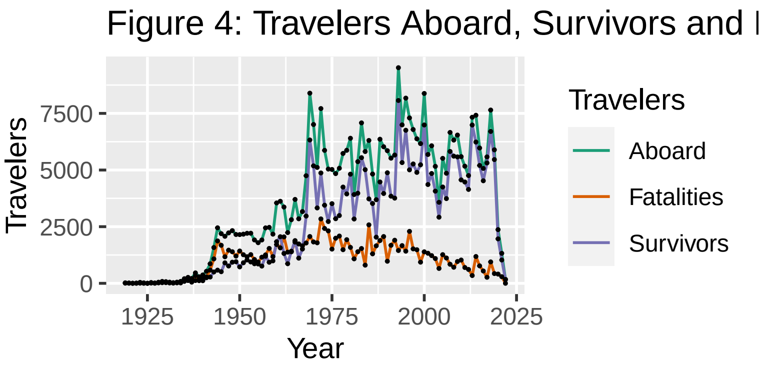


Figure 3: Average Survival Rate/Year shows the average survival rate for each year. The green line represents the mean of survival rate from 1919 to 2022, which is 57.01%. As we can observe from Figure 2, from 1919 to 1929, the trend appears to fluctuate widely, while for almost 40 years from 1929 to 1969, the average survival rates become flatter but all fall below the mean. Then, after 1969, the survival rate has continued to increase. However, there is a sudden drop in recent years, which worth more investigation.



As we can see from Figure 4: Travelers Aboard, Survivors and Fatalities Per Year, the total number of travelers aboard airplane crashes declined in the past two decades. In the last 50 years, fatalities also declined as would be expected with the decline of travelers involved in crashes. The gap in the difference between the number of fatalities and survivors has increased over the years. This indicates that the survival rate over time has became more and more higher.

# IV. Analysis

## Clustering

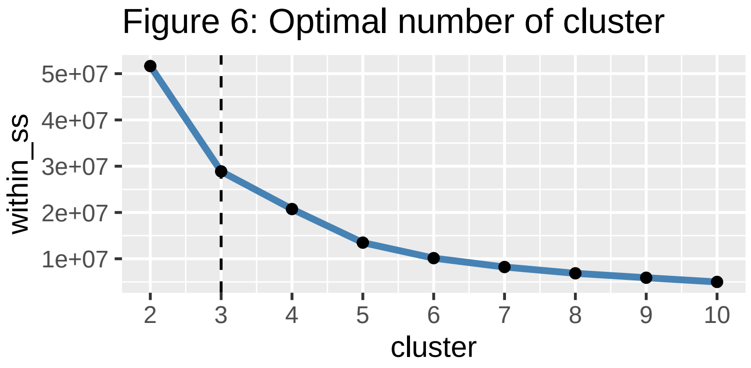
In order to solve our research question, first, we decided to cluster the data to segment crashes into low and high survival rates in order to discover the possible differences between survivable crashes and fatal crashes. The method we used is K-Means Clustering, which attempts to find groups that are most compact, in terms of the mean sum-of-squares deviation of each observation from the multivariate center (centroid) of its assigned group. This disorganized approach to clustering produces similar quality of clusters to hierarchical clustering but much faster.

Based on the rules of k-means, it doesn't allow the missing value, so we omit all the NAs from the dataset.

The k means clustering analysis technique applied to find similarities between airplane crashes based on: the number of people aboard, fatalities, survivors and survival rates. Twenty year intervals were applied to the clustering analysis model in the event technical advances that took place over the years may have contributed to increasing or decreasing crash events.

We tried four different K means cluster sizes of 2, 3, 4, and 5 were made along with a visualization comparison.

As shown in References:Figure 5, all four clusters provided distinctive groups with little overlap when depicted in a 2 dimensional space. To aid in determining the optimal number of clusters to us, the elbow method using the total within sum of squares plotted against the number of values of k was applied. We compare the total within sum of squares for a set of 10 cluster solutions. Ideal number of clusters is inferred from a sudden change in the line graph. As we can see from the Figure 6: Optimal number of cluster, three clusters will be used to segment the airplane crashes.



In order to examine the cluster result of airplane crash, we generated a list of mean values for travelers aboard, fatalities, survivors and survival rates. Moreover, to identify the capacity of aircraft, the minimum and maximum values for travelers aboard would also be included.

Table 2: Aboard, Fatalities, Survivors and Survival Rates

| Cluster | Plane_Crashes | Min_Aboard | Max_Aboard | Mean_Aboard | Mean_Fatalities | Mean_Survivors | Mean_SurvivalRate |
|---|---|---|---|---|---|---|---|
| 1 | 1084 | 91 | 524 | 165.80443 | 9.7961255 | 156.008303 | 95.834585 |
| 2 | 5230 | 0 | 113 | 19.40268 | 0.8279159 | 18.574761 | 95.880277 |
| 3 | 5325 | 1 | 217 | 19.14291 | 18.0507042 | 1.092207 | 5.650122 |

Inspecting the mean values by Aboard, Fatalities, Survivors and Survival Rates, the clusters can be defined by the number of travelers aboard and their survival rates.

- Cluster 1 - travelers ranged from 91 to 524 with an avg. of 165 aboard and had a high mean survival rate of 95.83%
- Cluster 2 - travelers ranged from 0 to 113 with an avg. of 19 aboard and had a high mean survival rate of 95.88%
- Cluster 3 - travelers ranged from 1 to 217 with an avg. of 19 aboard and had a high mean survival rate of 5.65%

Note: The average survival rate for all crashes in the data set was 57.01%.

These clusters can be referred to as follows:

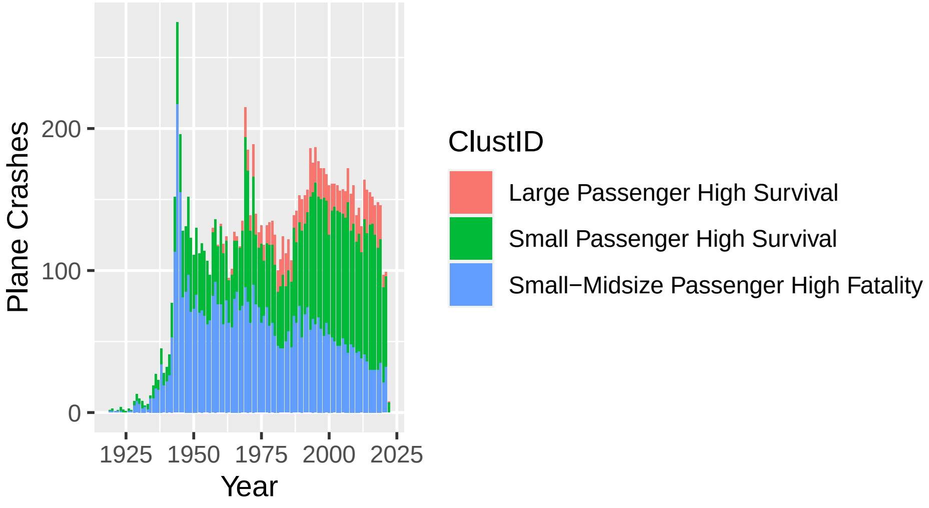
Large Passenger High Survival = Cluster 1

Small Passenger High Survival = Cluster 2

Small-Midsize Passenger High Fatality = Cluster 3

The next step is to evaluate survival rate for each cluster as per their cluster profile. Eventually, We would like to know if the survival rate is influenced by any significant factors.

Figure 7: Plane Crashes Per Year by Cluster



The chart clearly shows that Small Passenger Crashes involving planes with less than 113 occupants historically make up the bulk of plane crashes. There is a peak during Year 1940 to Year 1945. Moreover, Small-Midsize Passenger planes with less than 217 occupants are half less than the Small Passenger Plane. Large passenger plane crashes averaging more than 165 occupants begin to appear in the 1960's.

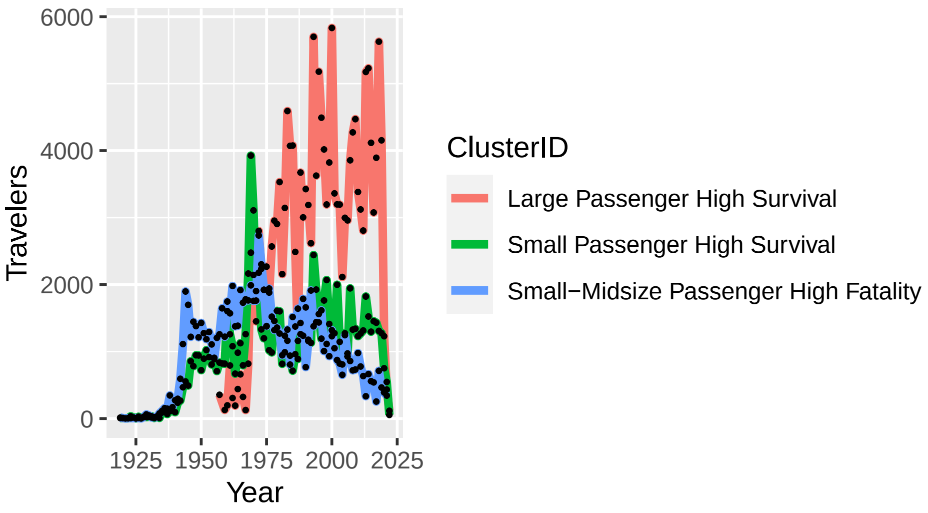


Figure 8: Travelers Aboard Airplane Crashes Per Year b

As the years progressed and large capacity planes increased, the annual number of travelers aboard large airplanes that crashed began to exceed the annual number of travelers aboard small to midsize and small planes that crashed. This is especially evident from the 1970's onward.

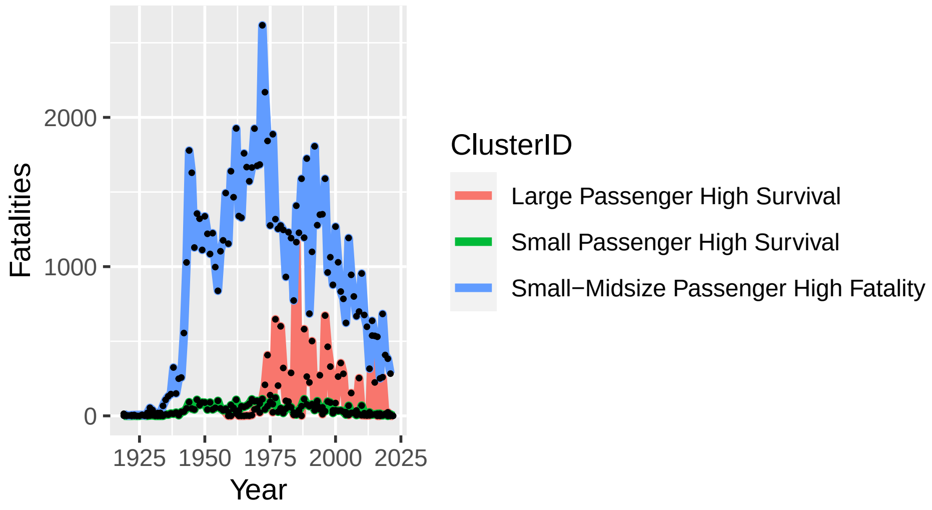


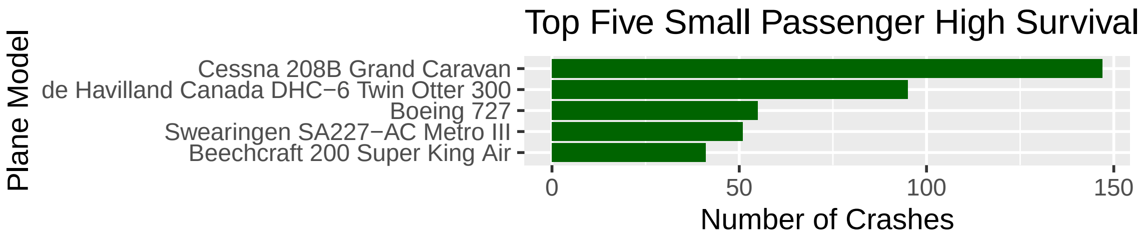
Figure 9: Airplane Crash Fatalities Per Year by Cluster

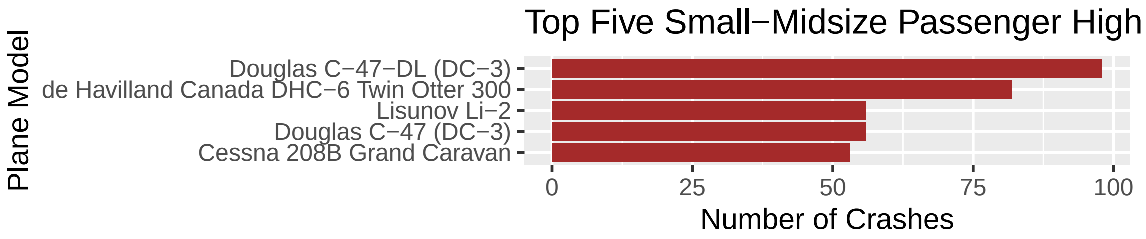
As expected small plane crashes with high survival rates ("Small Passenger High Survival") had the lowest annual fatalities.

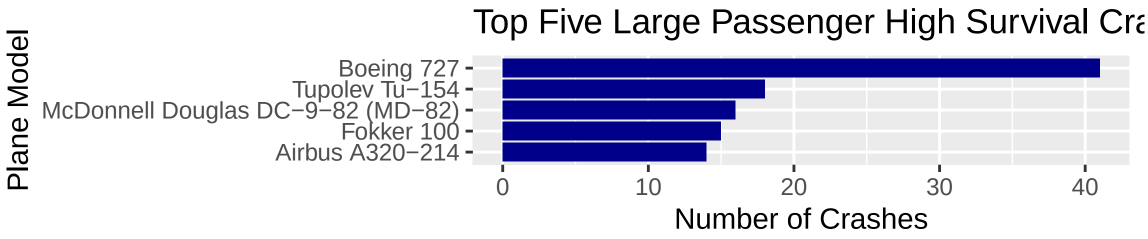
Also, as expected small-midsize plane crashes with high fatality rates ("Small-Midsize Passenger High Fatality") had the highest annual fatalities over years.

Starting in the 1960's we also see large plane crashes represented by the "Large Passenger High Survival" cluster (red line) begins to raise but still recognize as low fatality rate.

### Comparing Survival Rates on Three Size Passenger Planes







- American made Cessna 208B Grand Caravan and Canadian made de Havilland Canada DHC-6 Twin Otter 300 models appeared most frequently among the top 5 small passenger plane crashes with high survival rates.
- American made Douglas models and the Canadian made de Havilland Canada DHC-6 Twin Otter 300 models has most number of crashes among the top 5 small-midsize passenger plane crashes with high fatalities. After glancing top 10, Douglas models make up the most among the top 10 small-midsize passenger plane crashes with high fatalities.
- For large passenger airplane, the model most frequently involving in the air crashes, are Boeing 727.
- Based on the results, we could tell American made Douglas models may result in high fatality rate.

Next, we try to examine the result for Airline operator. As shown in References: Figure 10:

- Both USAF and USAAF Airline appeared most frequently among both the top 5 small passenger plane crashes with high survival rates and high fatality rates.
- Large passenger plane are mostly operated by American Airlines and Delta Airlines.
- Based on this result, we cannot determine whether airline operator will affect the survival rate or not.

Furthermore, let's examine the factor incident category. The resulted plot is shown in References Figure 11.

- Small passenger clusters and small-midsize passenger both have the highest accidents of a hull loss.
- Compared to large passenger clusters, the proportion of hijacking is significantly higher than in small and small-midsize aircraft. In addition, large passenger also has the highest proportion of repairable damage.
- In conclusion, for incident category, there is also not obvious difference between two groups.

Last but not least, the result of aircraft_damage_type is shown in References: Figure 12.

Table 3: Common Words and Network of Common Words

| word | count | word | count | item1 | item2 | n | item1 | item2 | n |
|---|---:|---|---:|---|---|---:|---|---|---:|
| united | 3651 | united | 3627 | america | united | 3154 | america | united | 3118 |
| america | 3155 | america | 3118 | united | america | 3154 | united | america | 3118 |
| russia | 763 | russia | 752 | kingdom | united | 463 | kingdom | united | 481 |
| canada | 486 | canada | 495 | united | kingdom | 463 | united | kingdom | 481 |
| kingdom | 463 | kingdom | 481 | ca | united | 334 | ca | united | 335 |
| brazil | 397 | brazil | 398 | ca | america | 334 | ca | america | 335 |
| colombia | 339 | base | 351 | united | ca | 334 | united | ca | 335 |
| ca | 335 | colombia | 344 | america | ca | 334 | america | ca | 335 |
| base | 331 | ca | 336 | fl | united | 289 | united | fl | 319 |
| afb | 309 | fl | 319 | fl | america | 289 | america | fl | 319 |

- By comparing the 3 clusters, all the cluster aircraft_damage_type “Damaged beyond repair” ranked the highest, except for the large passenger cluster.
- Small-midsize passenger has the highest destroyed damage type compared to the other two clusters.
- The large passenger cluster contains the highest substantial damage type.
- As a result, for aircraft damage type, damaged beyond repair makes up the most crashes, but there is not obvious pattern for differences between high survival rates and high fatality rate.

## Text Mining

Next, we decided to use text mining technologies to discover which country has the highest number of incidents. The variables we used to perform text mining are Departure_Airport and Destination_Airport. Based on the airport they departed or landed, we could discover the most common location that incidents happened.

Each review is tokenized into words and then pivoted to a tall format. dplyr functions are used to summarize, sort, and filter the top 10. Not surprisingly the list contains some prepositions and articles and words that don’t convey much meaning. Such words are called stopwords. In addition, “airport”,“international”, and “air” are used very often in Airport locations. So, we want to remove these words, which they are not too meaningful for most audiences.

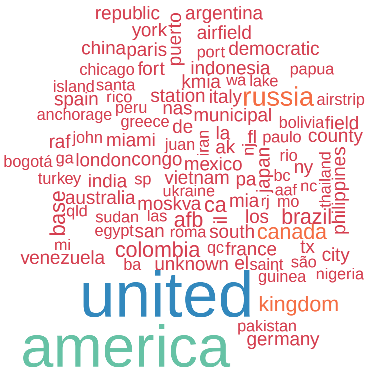
Table3: Common Words & Network of Common Words.

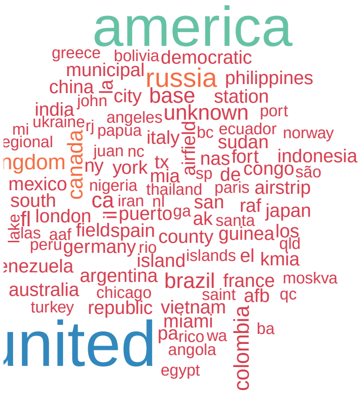
The left four columns are common words: the first two columns are from Departure_Airport, and the second two columns are words from Destination_Airport.

The rest columns are network of common words: the first three columns are from Departure_Airport, and the second three columns are words from Destination_Airport.

According to Table 3 as shown above and the word cloud graph as shown below, “united” and “america” are most common words in the dataset, then “russia”, “canada” and “kingdom are followed. After glancing the dataset, we found that United States of America and United Kingdom both begin with United, that is why United are so many records.

Figure: Word Cloud Plot. The left are words from Departure_Airport, and the right are words from Destination_Airport.





As a next step, we want to examine which words commonly occur together. We can use pairwise_count() from the widyr package to count how many times each pair of words occurs together in a Airport field.

According to Table 3 above and Figure 13: network plot below, "United" associated with "america" are most common, indicating that United States has highest number of incidents. The following is United Kingdom. Then, next few rows are "ca", "fl" and "ak", which are states of United States. The last two rows is "afb", which is Armed Forces Air Force, they mostly belong to Department of Military.

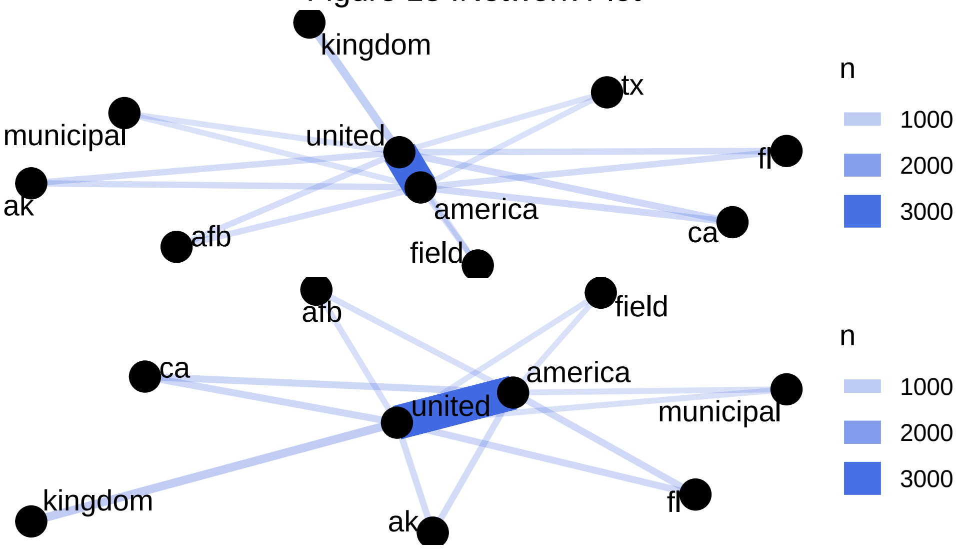


Figure 13 :Network Plot

Last but not least, we may want to the significance of each factors, so we try to fit the regression model. However, due to the fact that these factors has more than 1000 levels, they are not meaningful to fit the regression.

## V. Conclusion

Although the occurrence of air crashes contains many unknown factors, some of which are factors that cannot be predicted by human beings, we can still rely on scientific and technological means to help human beings reduce the occurrence of air disasters. The findings provided results to answer the question and discover reasonable patterns. In this instance, cluster analysis was used as a tool to further explore different segments under different factors that successfully identified patterns in three groups of survival and fatality rate and contributed insights to causes to similar yet different survival and fatality rate.

By identifying different clusters, we analyzed whether there are any significant factors that influence survival rate. We separated all plane crashes to 3 groups: Group 1 which travelers ranged from 1 to 200 has low mean survival rate of 4%, Group 2 ranged from 87 to 524 travelers with an average of 162 aboard had a low mean survival rate of 96%, Group 3 with travelers ranged form 0-98 with an average of 15 aboard had a mean survival rate of 93%. In addition, clearly small passenger crashes with 98 occupants are at most among all crashes, and large passenger plane crashes with more than 162 occupants appear after 1960.

Furthermore, for factor aircraft model, American made Douglas models appears most frequently in the top 10 small-midsize passenger plane crashes with high fatalities. For factor aircraft operator, USAF and USAAF are most frequently involved in the crashes for both high survival rate and high fatality rate, so it is hard to see the differences. For incident category, there is also not obvious difference between two groups. For aircraft damage type, damaged beyond repair makes up the most crashes, but there is not obvious pattern for differences.

Next, by identifying the country that has the most number of incidents, we utilized text mining technology, the result shows that United States has the highest number of aircraft crashed, and then United Kingdom.

Despite the clear pattern of the findings, there are some limitations that can be spotted while analyzing the result. For instance, before 1960 there were nearly no large passenger plane crashes, this could be due to the reason that before 1960, the aero technology was not mature enough to carry that large amount of passengers, therefore, most crashes would focus on small-sized passengers. Not only that, there were also military aircraft models included in the data, since military aircraft can only be occupied by one to two people, therefore, that could also be part of the reason why small-mid sized crashes accounted for more before 1960.

There are other considerations to keep in mind that when interpreting the result of cluster analysis, it does not help to predict or model based on current data; it helps to find a pattern from current data through visualization. The use of the findings could help to have a clearer view on how survival rates change and spread under different factors in order to help readers understand what can be reduced and what can be improved to eventually enhance the safety of passengers while in the area

# VI. Recommendation

Based on the result we analyzed, we can conclude that an airplane with a large capacity has a higher survival rate. On the other hand, an airplane with a lower capacity has a lower survival rate. Therefore, we recommend that passengers try to choose the largest aircraft possible when traveling. In spite of the result-based recommendation, logical reasoning should support our recommendation. Indeed, the row data not only reports commercial flights but also reports military, corporate jet, and training aircraft. Therefore, the recommendation is not necessarily true that larger aircraft are safer.

With the development of science and technology, the models of aircraft are also constantly updated. The latest aircraft have a faster flight speed and can reach longer flight distances than before, also accompanied by higher safety performance. As clarified by Julie O'Donnell, a spokesman for Boeing, where the 50s and 60s fatalities occurred roughly every 200,000 flights, today "fetal accidents [occur] less than once in every two million flights". National Transportation Safety Board also proves Julie O'Donnell's explanation, in fact, passengers on flights now have a 95% chance of surviving an accident now.

From our observation of the data and the above, we can see that the cause of the aircraft impact may not only be a problem with the aircraft itself but may be accompanied by many human factors. There were 130 hijackings between 1868 and 1972. "Some situation goes really bad, One of the most infamous pre-9/11 hijackings was TWA Flight 847, in which Hezbollah terrorists hijacked a plane bound from Athens to Rome, sought out passengers with "Jewish-sounding" names, and forced a two-week-long hostage situation at an airport in Beirut." But since 911, there have been only 50 hijackings in the last 20 more years, and none of them happened in the United States. That is because the Airline Company changed the security and airplane construction. It is another way to minimize the chance of air flight accidents by human factors.

# VII. References

Table 4: Missing Variables

| | x |
|---|---|
| Incident_Date | 0.0 |
| Aircaft_Model | 0.0 |
| Aircaft_Registration | 3.5 |
| Aircaft_Operator | 0.0 |
| Aircaft_Nature | 2.0 |
| Incident_Category | 0.0 |
| Incident_Cause.es. | 0.0 |
| Incident_Location | 2.0 |
| Aircaft_Damage_Type | 0.0 |
| Date | 0.0 |
| Time | 30.5 |
| Arit | 0.0 |
| Aircaft_Engines | 27.5 |
| Onboard_Crew | 0.0 |
| Onboard_Passengers | 0.0 |
| Onboard_Total | 8.5 |
| Fatalities | 0.0 |
| Aircaft_First_Flight | 11.5 |
| Aircraft_Phase | 0.0 |
| Departure_Airport | 19.5 |
| Destination_Airport | 19.5 |
| Ground_Casualties | 49.5 |
| Collision_Casualties | 49.5 |
| Total_Occupants | 14.5 |

```
## 'data.frame':    23519 obs. of  23 variables:
##  $ Incident_Date       : chr  "03-JAN-2022" "04-JAN-2022" "05-JAN-2022" "08-JAN-2022" ...
##  $ Aircaft_Model       : chr  "British Aerospace 4121 Jetstream 41" "British Aerospace 3101 Jetstrea
##  $ Aircaft_Registration: chr  "ZS-NRJ" "HR-AYY" "EP-CAP" "RA-64032" ...
##  $ Aircaft_Operator    : chr  "SA Airlink" "LANHSA" "Caspian Airlines" "Cainiao, opb Aviastar-TU" ..
##  $ Aircaft_Nature      : chr  "Domestic Non Scheduled Passenger" "Domestic Scheduled Passenger" "Dom
##  $ Incident_Category   : chr  "Accident | repairable-damage" "Accident | repairable-damage" "Acciden
##  $ Incident_Cause.es.  : chr  "Airplane - Engines, Airplane - Engines - Prop/turbine blade separatio
##  $ Incident_Location   : chr  "near Venetia Mine..." "Roatán-Juan ..." "Isfahan-Shah..." "Hangzhou-X
##  $ Aircaft_Damage_Type : chr  "Substantial" "Substantial" "Substantial" "Destroyed" ...
##  $ Date                : chr  "Monday 3 January 2022" "Tuesday 4 January 2022" "Wednesday 5 January 
##  $ Time                : chr  "08:10" "ca 12:00" "17:07" "04:40" ...
##  $ Arit                : chr  "03-JAN-2022" "04-JAN-2022" "05-JAN-2022" "08-JAN-2022" ...
##  $ Aircaft_Engines     : chr  "2 Garrett TPE331-14GR-805H" NA "2 CFMI CFM56-3C1" "2 Soloviev PS-90A"
##  $ Onboard_Crew        : chr  "Fatalities: 0 / Occupants: 3" "Fatalities: 0 / Occupants:" "Fatalitie
##  $ Onboard_Passengers  : chr  "Fatalities: 0 / Occupants: 4" "Fatalities: 0 / Occupants:" "Fatalitie
##  $ Fatalities          : int  0 0 0 0 0 0 0 0 0 2 ...
##  $ Aircaft_First_Flight: chr  "1995-05-19  (26 years 8 months)" "1985" "1992-09-18  (29 years 4 mont
##  $ Aircraft_Phase      : chr  "Landing (LDG)" "Landing (LDG)" "Landing (LDG)" "Standing (STD)" ...
##  $ Departure_Airport   : chr  "Johannesburg-O.R. Tambo International Airport (JNB/FAOR) , South Afri
##  $ Destination_Airport : chr  "Venetia Mine Airport (FAVM) , South Africa" "Roatán-Juan Manuel Gálve
##  $ Ground_Casualties   : chr  NA NA NA NA ...
##  $ Collision_Casualties: chr  NA NA NA NA ...
##  $ Total_Occupants     : num  7 19 116 8 0 NA 3 NA 12 2 ...
```

|| || || ||

### Figure 5

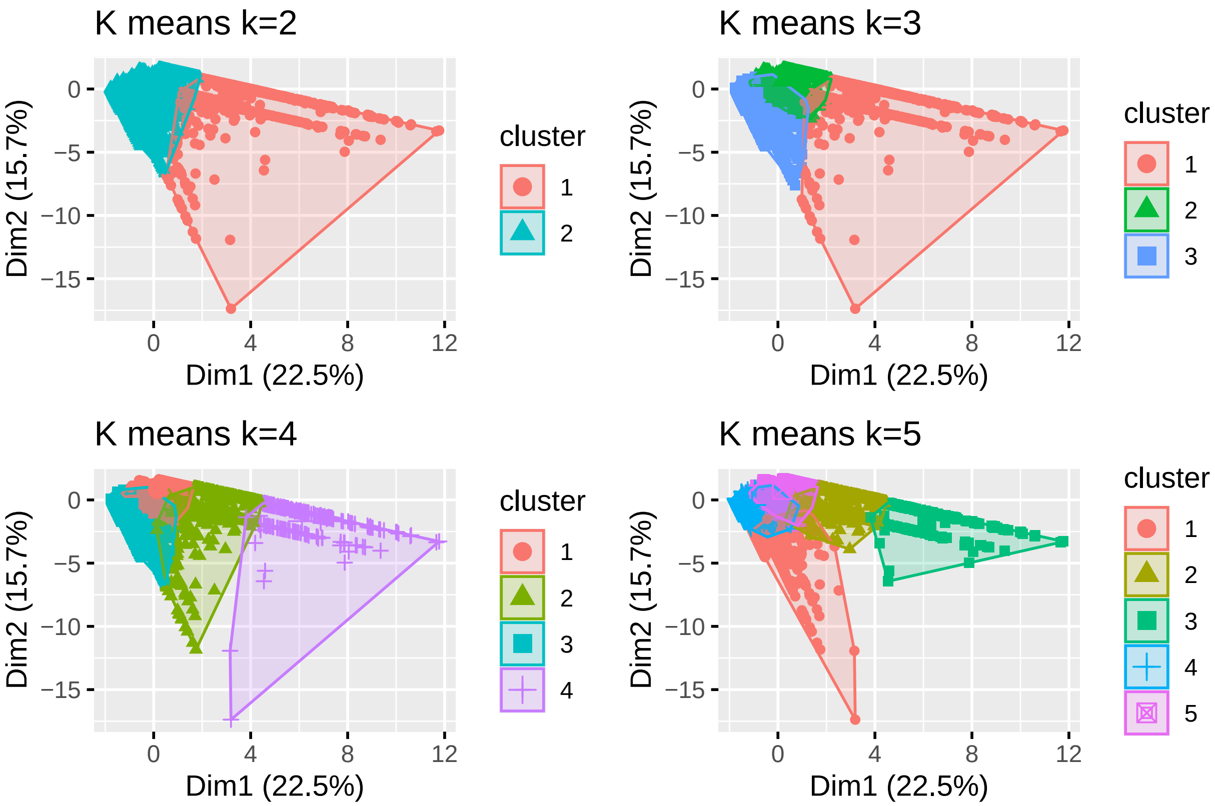
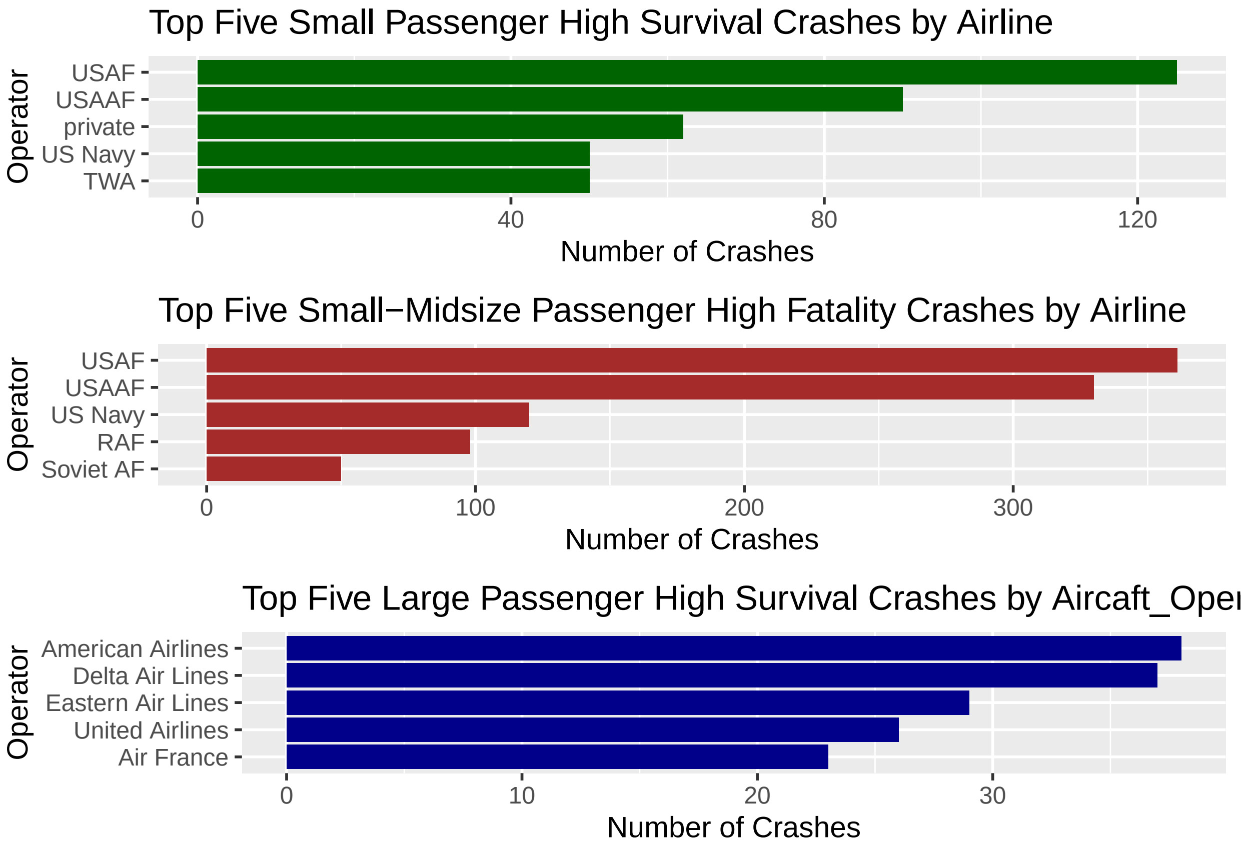


Figure 5: K means cluster sizes of 2, 3, 4, 5

**Figure 10**



Top Five Small Passenger High Survival Crashes by Airline

Top Five Small−Midsize Passenger High Fatality Crashes by Airline

Top Five Large Passenger High Survival Crashes by Aircaft_Oper

**Figure 11**

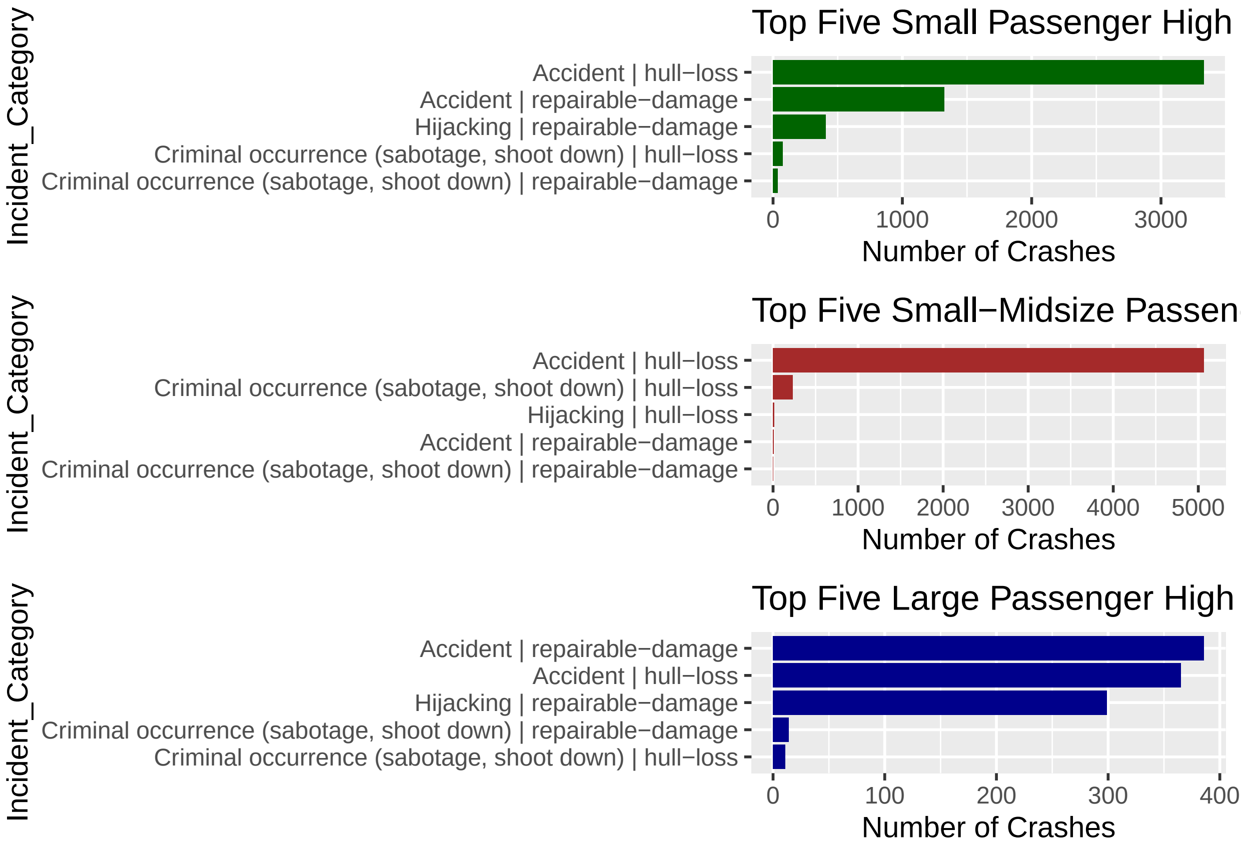
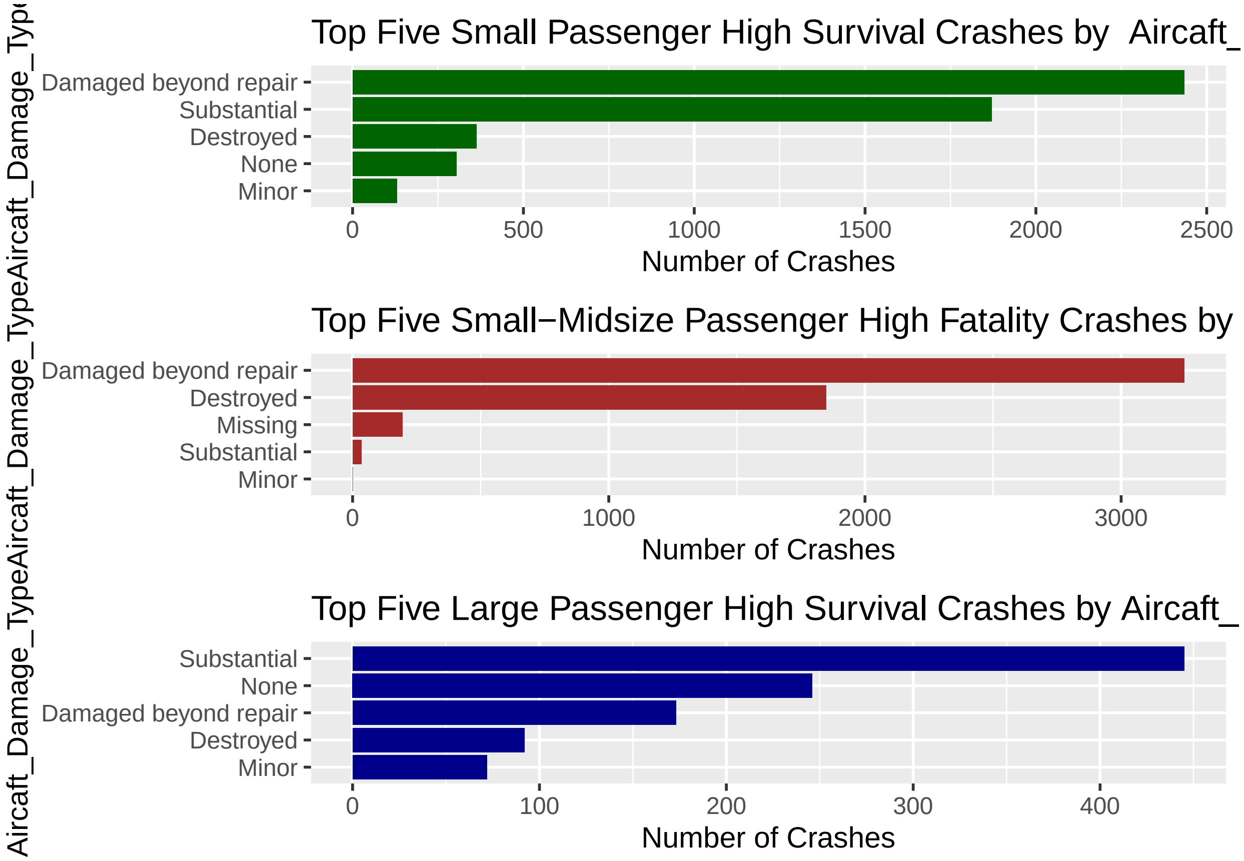

Top Five Small Passenger High
Incident_Category
Accident | hull−loss
Accident | repairable−damage
Hijacking | repairable−damage
Criminal occurrence (sabotage, shoot down) | hull−loss
Criminal occurrence (sabotage, shoot down) | repairable−damage
0
1000
2000
3000
Number of Crashes
Top Five Small−Midsize Passen
Incident_Category
Accident | hull−loss
Criminal occurrence (sabotage, shoot down) | hull−loss
Hijacking | hull−loss
Accident | repairable−damage
Criminal occurrence (sabotage, shoot down) | repairable−damage
0
1000
2000
3000
4000
5000
Number of Crashes
Top Five Large Passenger High
Incident_Category
Accident | repairable−damage
Accident | hull−loss
Hijacking | repairable−damage
Criminal occurrence (sabotage, shoot down) | repairable−damage
Criminal occurrence (sabotage, shoot down) | hull−loss
0
100
200
300
400
Number of Crashes

**Figure 12**



Top Five Small Passenger High Survival Crashes by  Aircaft_
Damaged beyond repair
Substantial
Destroyed
None
Minor
0
500
1000
1500
2000
2500
Number of Crashes
Top Five Small−Midsize Passenger High Fatality Crashes by
Damaged beyond repair
Destroyed
Missing
Substantial
Minor
0
1000
2000
3000
Number of Crashes
Top Five Large Passenger High Survival Crashes by Aircaft_
Substantial
None
Damaged beyond repair
Destroyed
Minor
0
100
200
300
400
Number of Crashes
Aircaft_Damage_TypeAircaft_Damage_TypeAircaft_Damage_Typ