

DJEBAR - FERRANDI - TISSOT

5ISS

2019 - 2020

Study of aircraft incidents/accidents in the US

What are the different parameters that cause accidents ?

Introduction

As we know, the airplane is the safest mode of transportation in the world, but it is not immune to malfunction. While in the event of a problem, a train can simply stop, for the plane a minor issue can quickly lead to disaster. We therefore first wanted to draw up a report on the evolution of the number of deaths by aircraft over the last few decades before trying to highlight the factors that play the most important role in the creation of incidents. For this we have chosen to use the NTSB Aviation Accident Dataset, which summarizes all aircraft incidents/accidents in the US since 1947. This dataset is constituted of 82635 records organized in 31 columns. There are informations related to the coordinates of the accident/incident, the number of fatal injuries, the aircraft manufacturer...

Data cleaning

Due to the difficulty to obtain reliable informations when an aircraft crashes, informations are missing in this dataset. This is why there are many "N/A" cells, but the large number of rows allows us to drop this value without consequences.

We also had to clean the aircraft manufacturer column because some names are written in uppercase while others are written in lowercase. And some company names are ended with "corp.", "corporation" or "company".

Dataset analysis

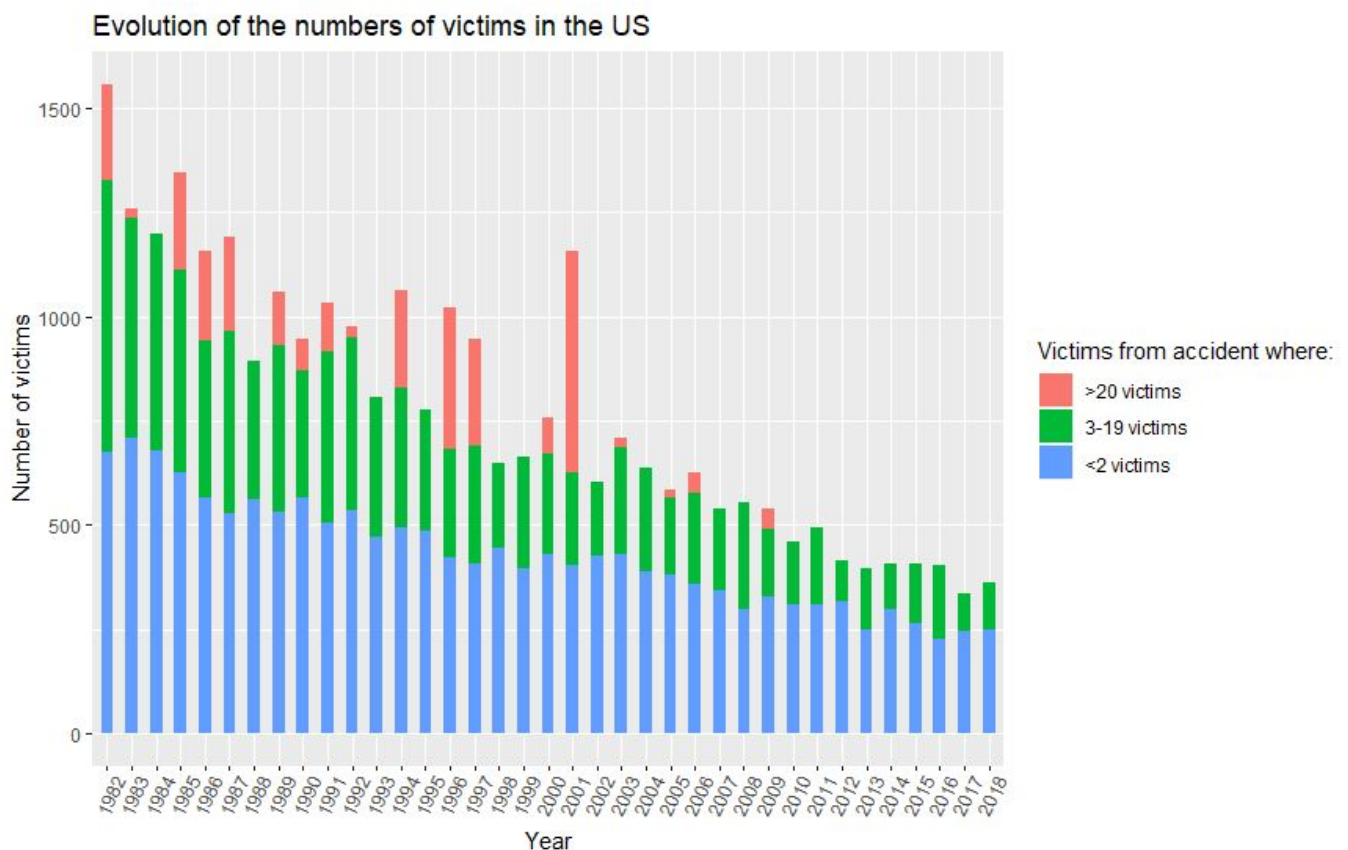
To try to answer our query, we have processed the data in order to make 5 graphs as meaningful as possible. First we make a report on the evolution of the number of deaths these last decades to know what we are talking about, then with the following graphs we try to highlight parameters which, we think, could have an incidence in the crashes of planes in the US such as weather conditions or location.

Chart 1: Evolution of the numbers of victims over the time in the US

The chart below depicts the evolution of the number of victims in the U.S. since 1982. We can see a more or less continuous downward trend. There was around 1500 victims in 1982 while there was less than 500 in 2018.

The colours represent the number of people who died in an accident involving a certain number of victims. In red are the victims from accident where more than 20 people died. This is likely crashes of commercial flights. We can see a general downward trend. We can notice a peak in 2001 which corresponds to the 9/11 attacks. After this date, the number of victims significantly decreases probably due to enforced security.

Since 1982, victims of commercial flights are only a small part of the total victims in the US. In fact, small and medium aircrafts crashes are nearly responsible of all victims.



Now that we have seen an overview of the evolution of the victims in the US, we can focus on the parameters of a crash.

Chart 2: Number of fatal accidents for different phases of flight

With this chart, we can see that the most accidental phases leading to a fatality are MANEUVERING, CRUISE, TAKE OFF and APPROACH. The aircraft damages represented using colours show us that the consequences of an accident are often a destruction of the plane. This is especially the case when the aircraft is in CRUISE phase.

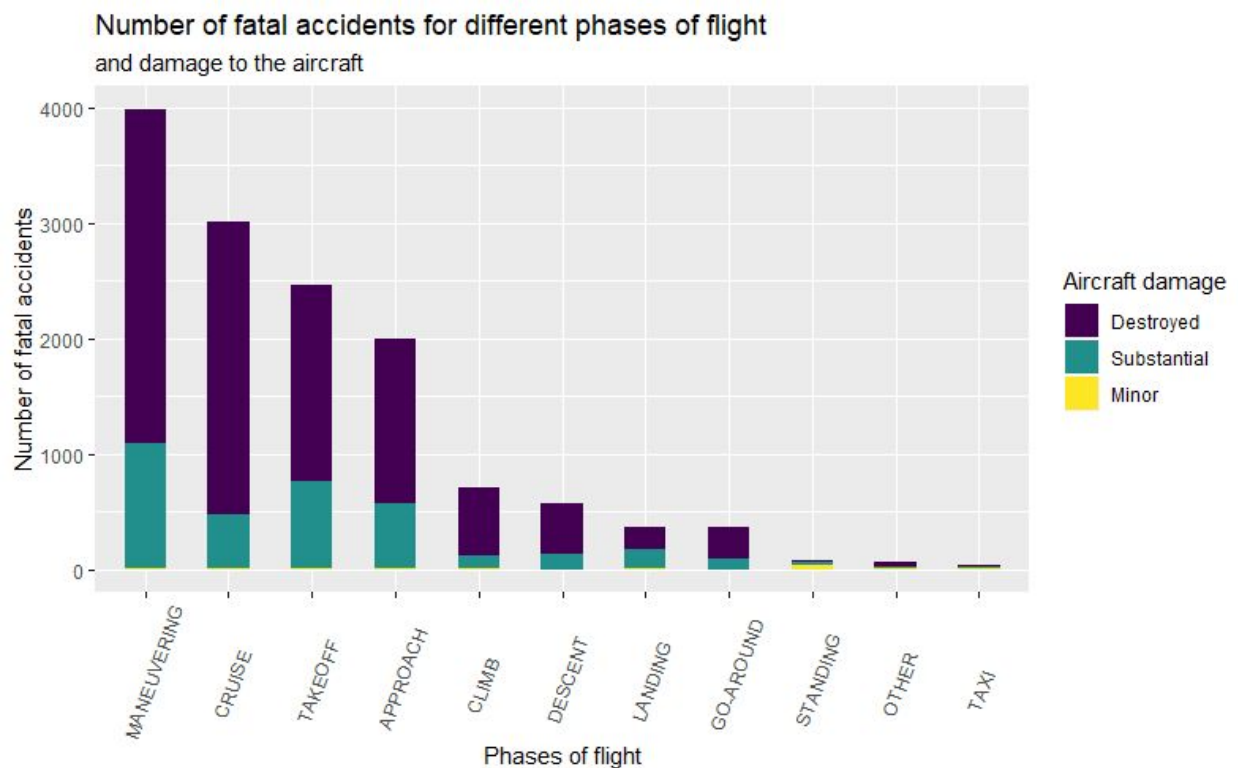


Chart 3: Total fatal injuries per month separated by weather conditions

We can also raise the question of the weather. Before explaining our analysis, it is necessary to understand the language used by airplane pilot. IMC means *Instrument Meteorological Conditions* in other world pilot has to fly only with plane instruments due to bad weather condition like cloud. VMC, *Visual Meteorological Condition*, means that pilots can fly with sufficient visibility. If we look at the chart, there are more fatal injuries in August and July than others due to a pick of flight during the holidays. In October, January or February the proportion of fatal flights with IMC is much higher than normal and accounts for almost half of the fatal flights with VMC. However, it is difficult to say whether the weather played a role in the crashes.

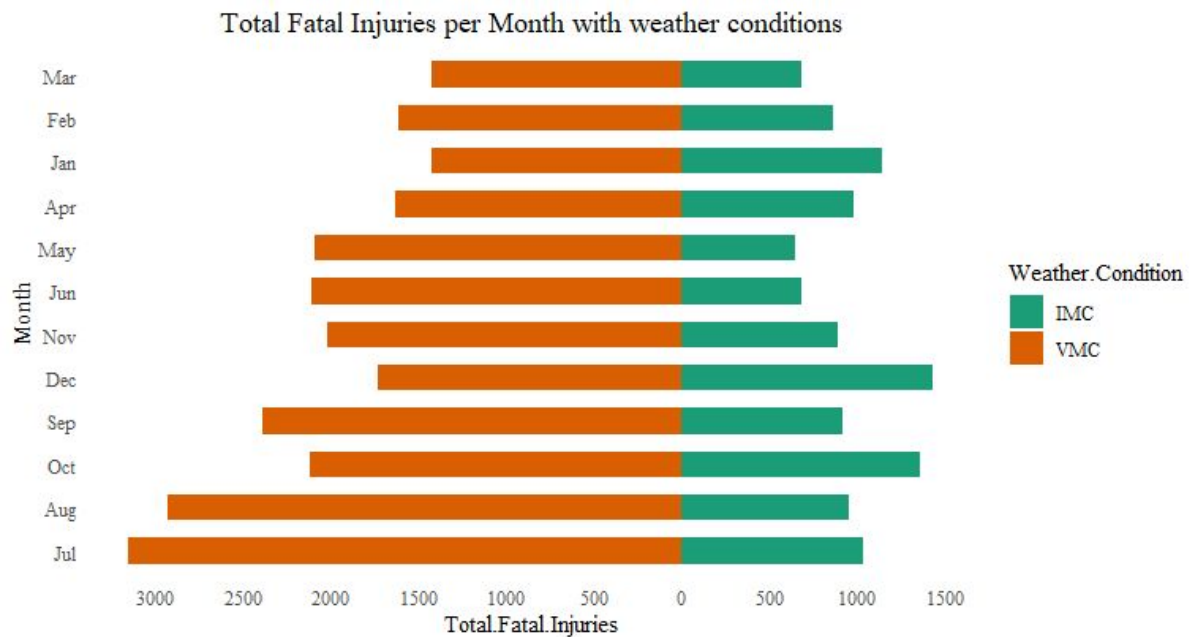


Chart 4: Death rate according to the aircraft damage

We can see with the violin chart below that the death rate is either near to 0 or near to 1. It means that when a crash occurs either all passengers die or all passengers survive. When the damages are minor or substantial, people are likely to survive. When the aircraft is destroyed, it is more distributed : some crashes are fatal while others not.

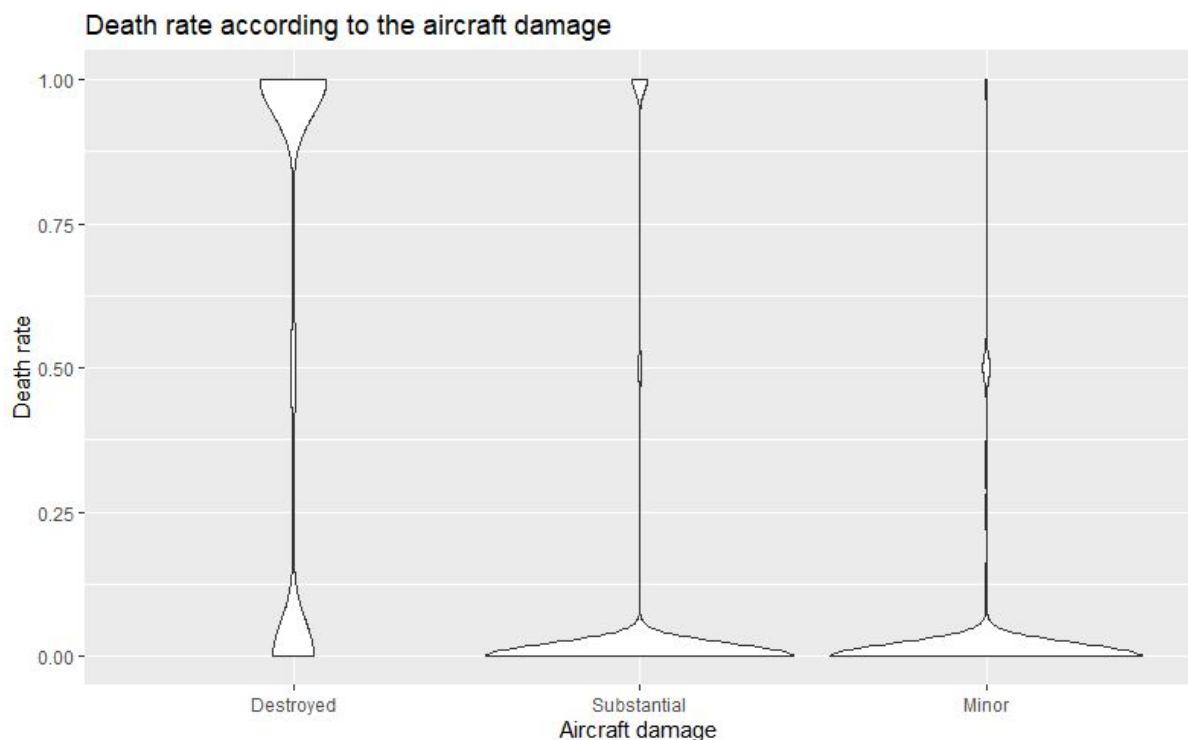
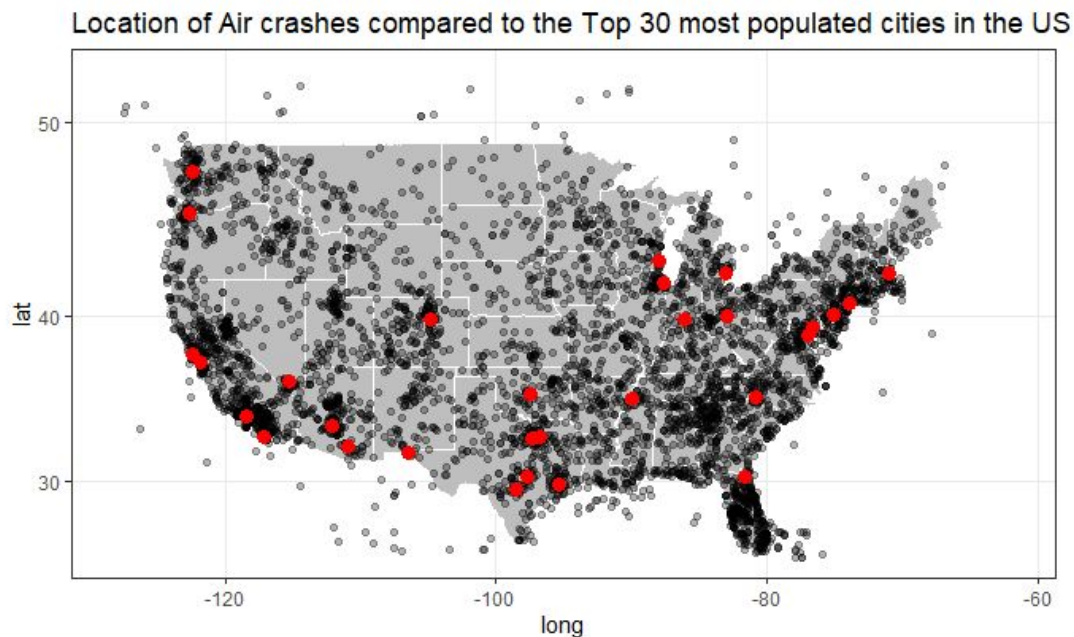


Chart 5: Location of Air Crashes in the US

As shown on the map, crashes generally take place near populated cities because there are more flights in the area and as we saw in chart 4 accidents often happen at TAKE OFF or APPROACH phases when the plane is close to an airport.



Conclusion

Even if the general number of victims of air crashes is decreasing with time in the US, the number of small planes crashes is holding steady. Death most often occurs during the phases of manoeuvring, take-off, approach and during which, in the event of an incident, the aircraft is often destroyed. Summer vacations are the deadliest in terms of victims and during the winter the number of deaths while flying with instruments increases. In case of minor or substantial damage the chances of survival are very high, whereas the opposite is true if the aircraft is destroyed. Finally, most aircraft crash in densely populated areas, which is consistent with the distribution of fatalities during flight phases.