The question I chose to try to answer is “has the number of plane crashes increased over the years?” It has appeared to me, over the past few years I have heard about plane crashes on a more regular basis. I decided to take a look to see if this is a real thought or if there is a different reason I may be hearing about them more often, ie. My age, social media, more media coverage, etc.

The original data set I chose had a lot of great information and had data going back to the early 1900s. However, after I started trying to analyze the data, I realized it wasn’t set up for great analyzation. There wasn’t a way to compare apples to apples. I later found another data set that allowed me to look at data that was more easily comparable. It broke down total flights by flight hours and gave information such as how many total aircrafts were active in a given year.

After looking at the data I found my original assumption that plane crashes (with fatalities) are more prevalent today than there were in previous years was not correct. By looking at the scatter plot ‘Fatalities by Year’, you can see a steady decrease since the late 1990s. I then looked at active aircraft in the late 90s to see if there was a decrease to correlate to the decrease in fatalities. There was not a decrease, but rather an increase. I also found the PMF of the Total Fatality Rate and compared “older” data to more recent data. I compared the PMF of the Total Fatality Rate from 1960-2005 and 2006-2017. The histogram of the PMF showed a lower fatality rate from the recent category than the older category.

I spent a lot of time working on an old data set trying to make it work, instead of looking for something more effective. I wish I would have changed data sets sooner, to have more time to focus more on the new data set. While the new data set had a lot better information that I was looking for, it would have been helpful to have a fuller data set. From 1960-1990 only one data point was provided for the decade. Whereas from 1995-2017 data was provided each year. This could have potentially skewed the results.