Project Task 1 Supporting Documentation

For this task, I wanted to analyze the metrics that describe airline safety. Based on the provided data, we observe sporadic total fatalities among airlines from 2000-2014, and there does not appear to be a dependency on an airline’s Available Seat km Per Week, which is somewhat indicative of how many passengers that an airline can service. There are even some airlines that have had no fatalities during the period.

I was able to find some data for U.S. Airline safety, and it appears that the safety has improved over the span of years since the 1980’s since the trend of the graphs for both number of fatalities and accidents have gone down. Just one note about the data I found for the year 2001 during the 9/11 attacks, 531 fatalities occurred based on other sources of information. However, I did not notice this count included in the data that I found. I am not sure if these numbers were omitted due to the nature of the incidents, but I did not find any documentation to explain the discrepancy.

It has been common to compare the safety of different modes of transportation against each other, so I have included metrics for automobiles in the U.S. These comparisons should be analyzed closely as they are not as straightforward. For example, there are more automobiles and drivers on the road than there are planes and pilots in the sky, and people usually travel more frequently in cars than in airplanes. Airplanes usually have a far bigger capacity than automobiles, so airplane crashes usually get more attention because more people are impacted. If we compare the raw numbers, we see that there are so much more vehicle fatalities in the U.S. compared to U.S. airline fatalities by a huge magnitude. I tried to indicate this in the graphs by putting the y-tick labels in the graphs in bold. We easily can observe the same for the graphs showing airline and automobile accidents that there are so much more automobile accidents than airline accidents per year.

For my graphs, I made sure to add titles and labels for the axis. I used colors to accommodate audience members that may be colorblind. When comparing airline and automobile data, I used consistent colors in the graphs so that the comparisons could be understood better. For the airlines data, I used horizontal bar graphs because of the number of airlines. I used line graphs for the rest of the graphs because they show continuous variables over a span of years.

Data Sources:

https://catalog.data.gov/dataset/accidents-and-accident-rates-by-ntsb-classification-1995-through-2014-for-u-s-air-carriers

https://catalog.data.gov/dataset/accidents-fatalities-and-rates-1995-through-2014-for-u-s-air-carriers-operating-under-14-c-2e5ec

[https://cdan.nhtsa.gov/tsftables/tsfar.htm#](https://cdan.nhtsa.gov/tsftables/tsfar.htm)

Research References:

Silver, Nate. "Should Travelers Avoid Flying Airlines That Have Had Crashes in the Past?" <https://fivethirtyeight.com/features/should-travelers-avoid-flying-airlines-that-have-had-crashes-in-the-past/>

Difference Engine: Up, up and away. Retrieved from https://www.economist.com/babbage/2013/01/07/difference-engine-up-up-and-away

Github Repository: https://github.com/cvibanez/DSC640\_Project