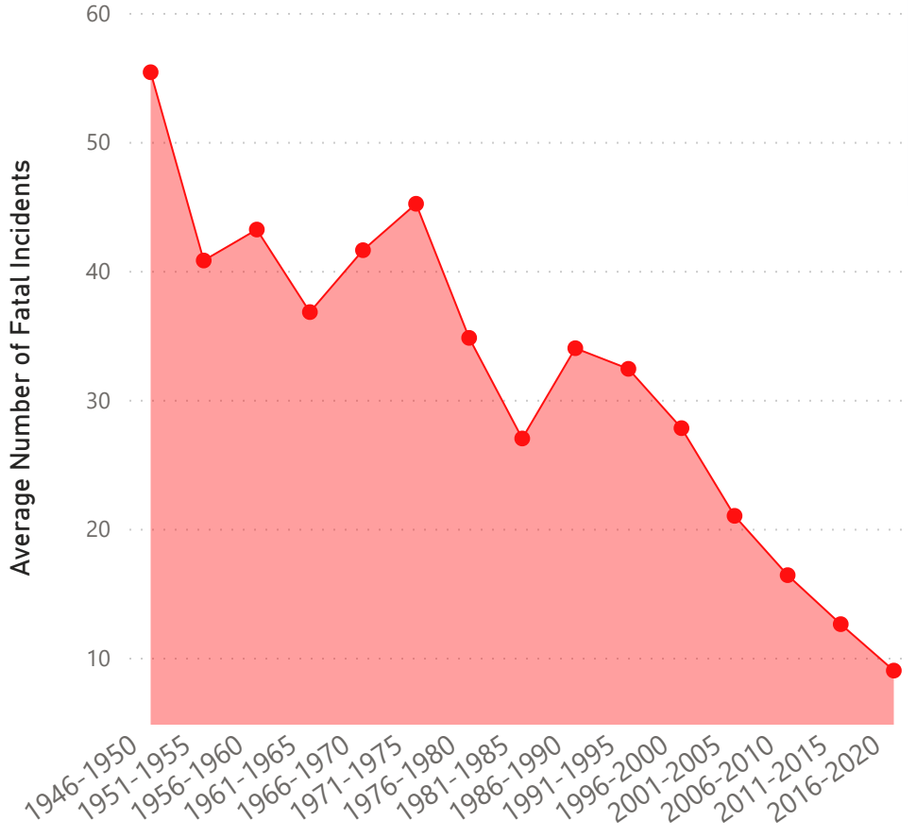
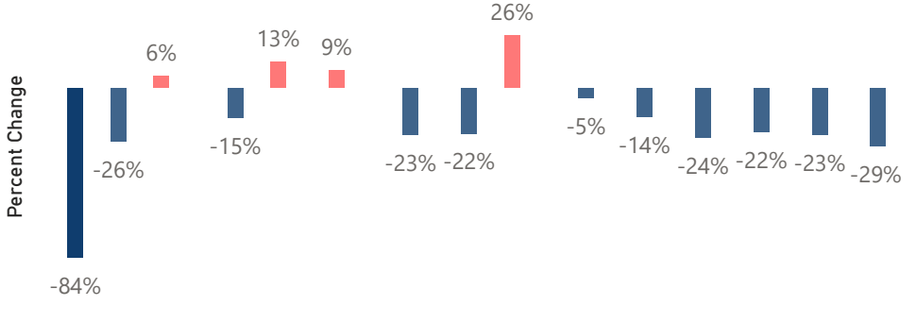
NewJet Air Safety Data Hub for Data Scientists:

Airlines: Average Number of Fatal Incidents in 5-year intervals





Average Number of Fatal Incidents 1995-2020

16.85

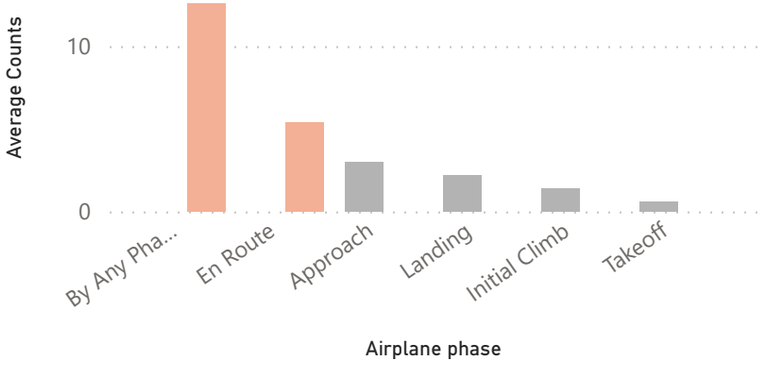
Available Seat Miles (Millions) 1995-2020

881381

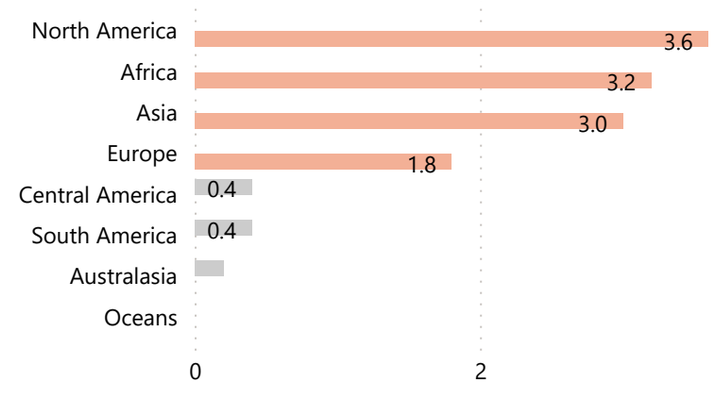
Fatal Incidents by Airline Company: 1985 to 2014

Airline Name	Sum of Incidents
Aeroflot	15
Delta / Northwest	14
United / Continental*	10
US Airways / America West*	9
American	8
China Airlines	7
Ethiopian Airlines	7
Air France	6
Garuda Indonesia	5
Korean Air	5
Pakistan International	5
Philippine Airlines	5
TAM	5
Thai Airways	5
Turkish Airlines	5
Egyptair	4
Avianca	3
Malaysia Airlines	3
Over 29 years	**159**

Average Counts of Fatal Incidents by Airplane Phase: 2017-2021



Average Number of Fatal Incidents by Continent: 2017-2021



Milestone 1

Kristie Kooken

DSC 640-T301

September 24, 2023

The dashboard that I created for the Data Science Team has seven visualizations in total. The purpose of my dashboard was to communicate high level information regarding flight safety by way of detailing incidents of fatalities for airlines. I wanted to be able to communicate the seriousness of the topic by using red to fill in either the color for the plot area, the column of fatality counts by airline or when the percent change goes in the non-improvement direction. Likewise, I wanted to convey positive or more neutral information using blue. This is shown for the text visualizations (a light blue to stand out more) and the blue shading for the percent change plot where an improvement is shown as well as the airline names (neutral). My visualizations do not have borders and often have a white background. For the bar charts, I emphasized the overall total (when present) as well as the higher counts for categories in a similar red/orange color while deemphasizing the lowest and zero counts to gray bars.

I wanted to have an overview of safety over time and used a line chart for this, however, I was interested in showing the percent change between the 5-year intervals as I thought it was interesting to see the continual drop in incidents during 1986 to 1990 interval and beyond. There were different historical events, such as hijackings in 1986, and after this peak, a steady, sustained decrease in incidents after 9/11 where cockpit rules changed to prevent hijackings as well as other safety measures were emphasized. Likewise, I included the list of airline names with counts of fatal incidents to allow the team to compare our company performance to other similar sized airlines as a high-level, first-pass look at individual company data. It was important to emphasize the many millions of miles airlines travel and how few safety events have taken place, on average, during the same timeframe. Also, I thought it was important to have a brief overview of where in world and where on the airline's journey fatal incidents occur. While there

is much more to analysis, I am hoping my high-level summary gets the team interested in next steps with this data.

Ethical choices I made with these data include using averages over many years of data – an average can smooth out differences between the years however, those differences could be very important such as new safety requirements that became the norm or improvements in technology that created a safer plane but not necessarily safer pilots (as an example). More work would have to be done in order to ensure averages were the best way to present the data. At times, when there is a lot of data, there needs to be a method for how to reduce it down in order to have a meaningful or understandable visualization and for these visualizations, this is what I used.