Flight Viz Project – readme

- Dataset
  Dataset used: Flight Delay dataset.
  US domestic flight in 2007. Source: Harvard dataverse (https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/HG7NV7)
- Summary of findings:
  - Short haul flights are more frequent than long haul flights.
  - Early arrivals are more common than early departures.
  - Carrier delays are the highest among delays and carrier cancellations are the most frequent among cancellations.
  - NAS delays are the main reason behind arrival delays for flights that departed on time.
  - Best on-time performance goes to Hawaiian Airlines and Aloha Airlines.
- Key insights of the presentation:
  The deck starts with airline associated delays for each of the 20 airlines in the dataset, then moves on to display a breakdown of flight cancellation categories. Moving on, we see a correlation matrix for all types of delays that clearly communicates relationships between different types of delays. Continuing with delays, next slide plots an x-y scatter of arrival delays against departure delays, then a zoom in on the dark spike noticed on the y-axis (arrival delay). To find out the main contributor for this spike, we plot a correlation heat map which shows that NAS delays are the main culprit here.
- References:
  - Understanding the Reporting of Causes of Flight Delays and Cancellations | Bureau of Transportation Statistics (bts.gov)
  - IATA delay codes - Wikipedia