STATS 3355.001 Team 1 Project Proposal
Ved Nigam, Hiba Khan, Monique Picott

Dataset Information

Our dataset, found on www.transtats.bts.gov, an official website run by the United States Department of Transportation, is a collection of data compiled by The Bureau of Transportation Statistics. As our primary focus, we have decided to observe the effects of COVID-19 on airline efficiency. The aviation data library contains 27 databases, and out of these, we have chosen to focus on the dataset Reporting Carrier On-Time Performance Data in the state of Texas during August. We will be utilizing data from the years 2019 and 2020 which each contain 123,148 and 78,056 records, respectively. Total records that will be considered will be 201,204 records. Each dataset will contain the same information detailed through 21 variables. This dataset encompasses statistics on major airlines, their scheduled flights, delays, diversions, etc. The link to the database we downloaded the dataset is:
https://www.transtats.bts.gov/Fields.asp?gnoyr_VQ=FGJ

Background and Motivation

The COVID-19 pandemic brought a sudden halt to the ability for people to travel freely. Many airlines took heavy hits and had to completely re-establish their protocols for flights to accommodate for COVID regulations. In 2020-present database records, the effects of the pandemic in airline efficiency show clearly. We chose this specific dataset info as our topic of interest because we are intrigued by how the pandemic affected flight corporations on a major scale, and how they were able to adapt and power through to avoid bankruptcy and return to pre-pandemic efficiency/revenue.

Multiple Potential Interconnected Questions

- Which airline was operating the most during pandemic years?
- Is there a particular city that had the most travel during the pandemic?
- What was the most common reason for flights being delayed during the pandemic timespan?
- Which particular airline experienced the most delays in this time period?
- Is there a particular day that was most favorable for travel during this time?
- Which month during the pandemic had the most flights?
- Which month during the pandemic had the least flights?
    - Note: Flights became very cheap at the beginning of the pandemic
- Have flight numbers returned to their pre-pandemic level of efficiency?
- What years can we classify as affected by the pandemic? (COVID-19 is still relevant/the pandemic had no official 'end', so when did the noticeable effect of the pandemic disappear?)
- Was there a specific airport where most diverted flights were sent during this time period?

NOTES:

DL - DELTA
AA - AMERICAN AIRLINES
UA - UNITED AIRLINES
WN - SOUTHWEST
NK - SPIRIT