Flight Delay Analysis

Omar Shatrat, Chloe Li, Siyan Li, Grace Xie, Yan Wang

Table of contents

Business Problem/Objective

Our group has been engaged by the U.S. Department of Transportation. They are trying to create policy to alleviate delays in air travel and want some insights they can use to identify the root of air travel inefficiency and delay.

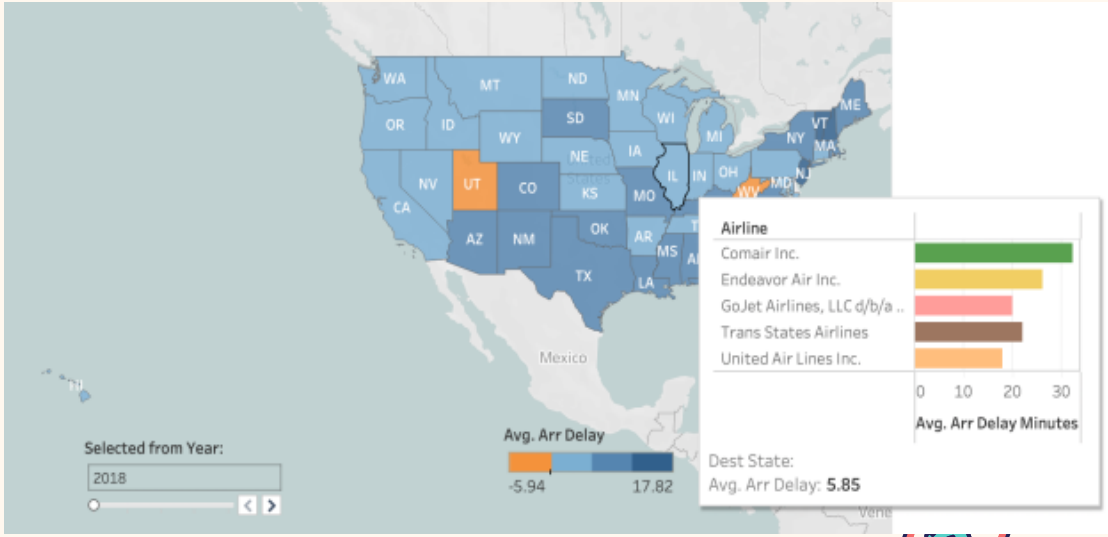
The Department would like to get a better understanding of how, where, and when delays occur so that they can tailor their policies. They would also like to understand how delay trends changed during the COVID pandemic.

Data Sources Considered, Evaluated, Selected

- Explanation: We wanted a dataset with temporal, geographic, and latent variables. These relationships between features and our target values are important because they help us answer the business question in a comprehensive way and from multiple perspectives.
- Source: https://www.kaggle.com/datasets/robikscube/flight-delay-dataset-20182022
- Overview: there are 61 columns in total. It spans years from 2018-2022 and each year has hundreds of thousands of rows of data.
- Because the data set is extremely large, we will randomly sample 20,000 records from each year in the dataset for our visualization.

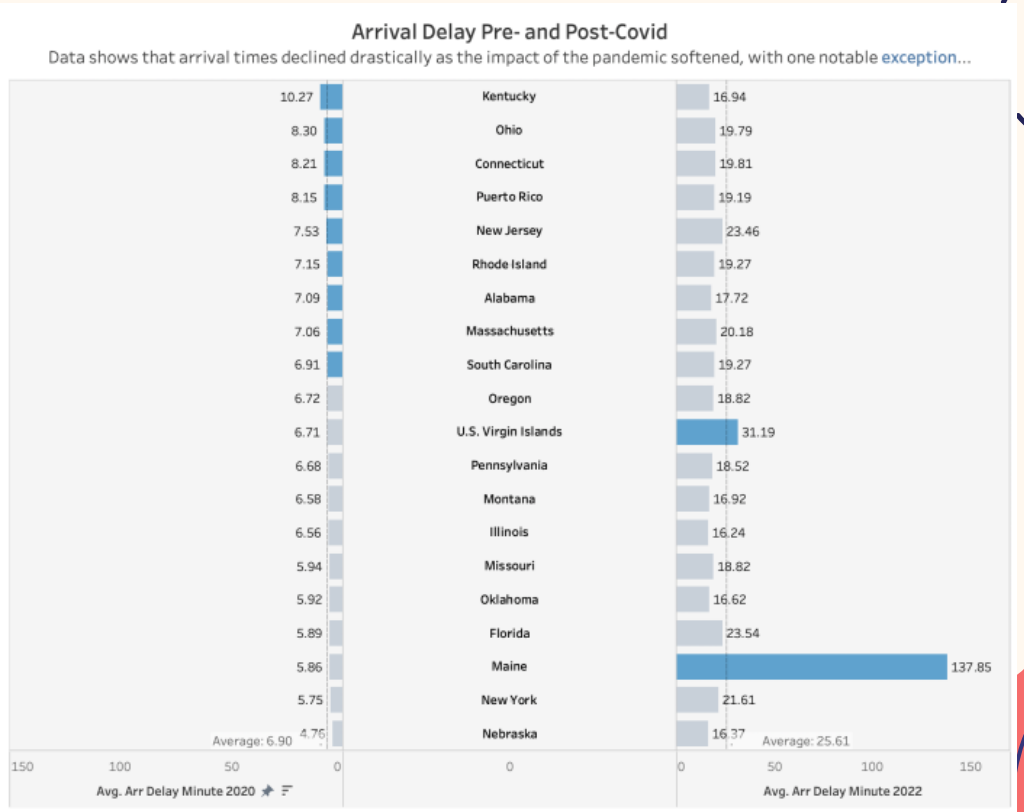
Interactive U.S. Airline Performance Dashboard

- Parameter 1: Adjust time range from 2018 to 2022
- Parameter 2: Select Top 5 Airlines by highest average delay time
- Calculations: True | False from the parameter representation
- As we hover on different states, horizontal bar charts will display the average delay time and delay minutes airline in those states
- Some popular travel states do not have longer delay times -Florida and California



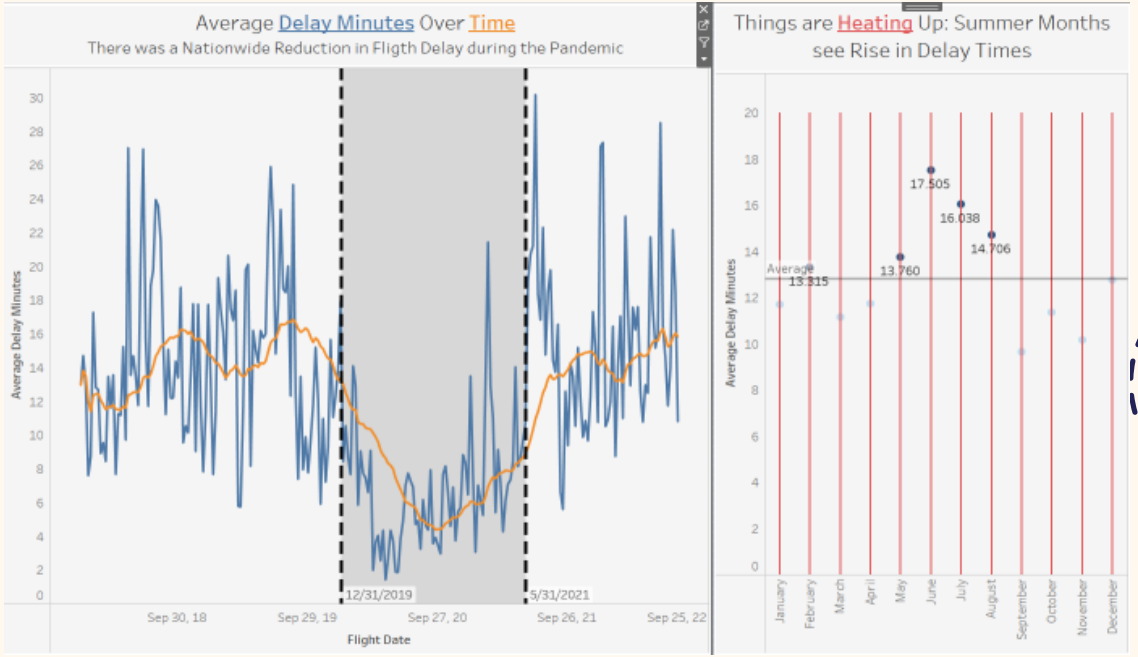
Arrival Delay Pre- and Post-Covid

- The chart shows the arrival delay minutes before and after Covid for the top 20 states with the most significant differences.
- Each bar is color-coded based on whether the state's average arrival delay is above or below the overall average.
- The average arrival delay pre-Covid was 6.9 minutes, while in the post-Covid period, it increased to 25.61 minutes.
- Republic Airlines at Portland International Jetport (PWM) recorded the most arrival delay minutes.



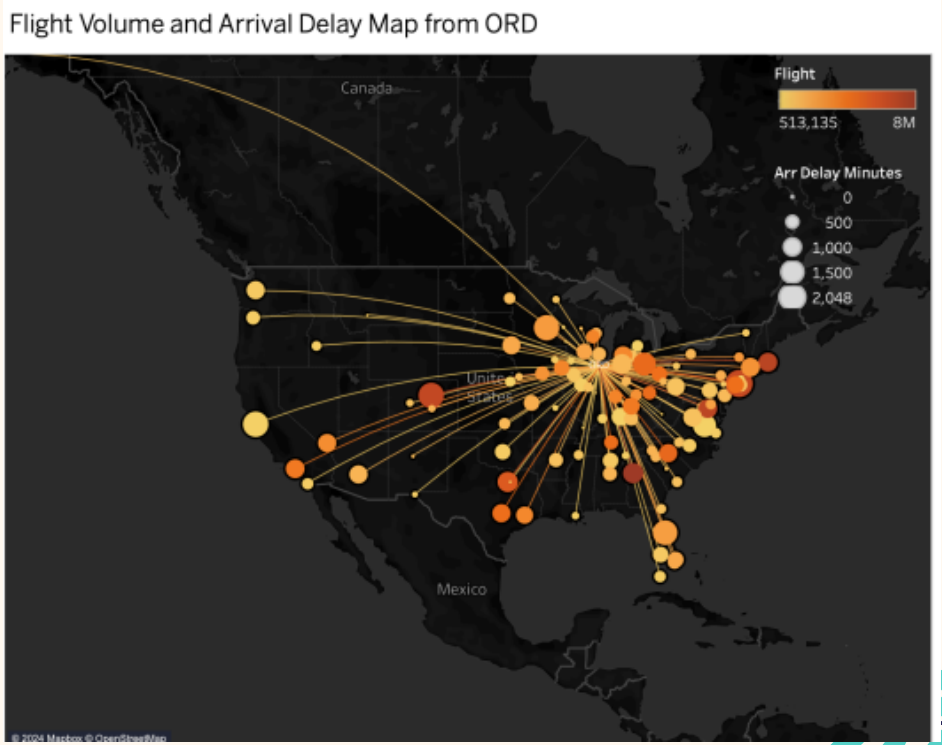
Arrival Delay Analysis: 2020 (COVID) vs. 2022 (Normal)

- The chart illustrates the temporal trends in delay minutes. We see during COVID that delay time fell, even in the summer when delay time are usually at their highest.
- The shaded gray region is considered to be the most intense COVID period.



Flight Path

- The map illustrates the flights originating from ORD to various destinations.
- The lines' color intensity on the map reflects the volume of flights, with darker shades representing heavier air traffic.
- Each destination's average delay is conveyed by the size of the dot on that location, where a larger dot equates to a longer delay.
- The map serves as a practical guide for passengers to anticipate expected delays at different airports, which can be essential for managing travel schedules and connections.



Sankey Chart

- The map creatively analyzes delay time broken up by airport and airline.
- The percentages are of total delay hours and are also broken up by airport and airline
- ORD experiences heavy delays on United and American flights

I created many parameters and calculated fields to make this chart, including a sigmoid function (to create curves in the bars), custom binning, and a window average of the sum of departure delay times

Next Steps/ Insights for DOT

- There is a Huge Republic Airlines Problem in Maine; Identify Source of Bottleneck
- See What Lesson Can be Learned from COVID; Do More Rigorous Investigation into Summer Discrepancies
- Chicago Struggles when Flying South and on United Flights, this is Most Efficient Place to Create Change

Thanks!

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**