STAT436 Homework 2

Shiny app: https://stevenren.shinyapps.io/homework_2/

Data: https://uwmadison.box.com/shared/static/7ux2kexdoetugbxd5xdp8642620l0snd.csv

Flight Delays and Cancellations

Introduction

The original data was found from Kaggle which contained a simple random sample from all 29,380,334 flights recorded between January 2019 to August 2023. As an out-of-state student who also enjoys traveling, I found this data to be useful in analyzing which days of the year are busy and result in delays and cancellations and helps with the travel planning process by identifying airlines with the highest and lowest delay or cancellation chances.

What are some interesting facts that you learned through the visualization. Provide at least one unexpected finding.

An interesting yet intuitive finding is that when observing the large airports in the US, we can easily notice the spikes in delays and cancellations around seasonal holidays or events. Christmas up to the time after the New Year is the busiest time of the year, which can cause many more delays or cancellations. Some more observed spikes appear to be seasonal, such as school starting and other major holidays. The one unexpected finding was the lack of flights in 2020 from March 17 to April 22 due to Covid-19 pandemic restrictions which I never knew about before.

How did you create the interface? Were there any data preparation steps? What guided the style customizations and interface layout that you used.

The data was prepared by first deleting columns that were not going to be used, then using mutate to make the date column formatted correctly with Shiny's functions. When performing calculations, I made sure to ignore any NA values. The shiny app utilizes fliudPage and sidearLayout for the overall format, titelPanel is used to display the title, sidebarPanel is used to position the various input options, finally, the mainPanel is used to display the 2 main plots.

My goal was for all possible inputs to be located on one area of the app, so the users do not have to be inputting all around the app. As a result, I chose to have it all within a side panel on the left side. Some features that I was unfortunately unable to implement currently include displaying additional information upon selecting or brushing columns of the plots. I am open to user feedback in adding more requested features.

What is the reactive graph structure of your application?

The reactive graph structure of my shiny application first consists of a filter for the selection of a date range, a selection filter to choose a year, selection filters for departure airport or city name, two checkboxes for displaying/hiding delay and cancellation information, and finally optional selection filters of arrival airport or city name. If a specific year is chosen, it will filter by

year only, if “No selection” is chosen, filter by date range only. Users are allowed to select multiple airports or cities in the selection filters for a broader search. A time series bar plot is created that corresponds to the inputs displaying the percentage of flights delayed or canceled over time. A text additionally outputs the information related to the most and least delayed airline, and the most and least cancelled airline.