Tableau Data Visualizations of Flight delays

MOHAMED HAYTHAM

DATA ANALYST

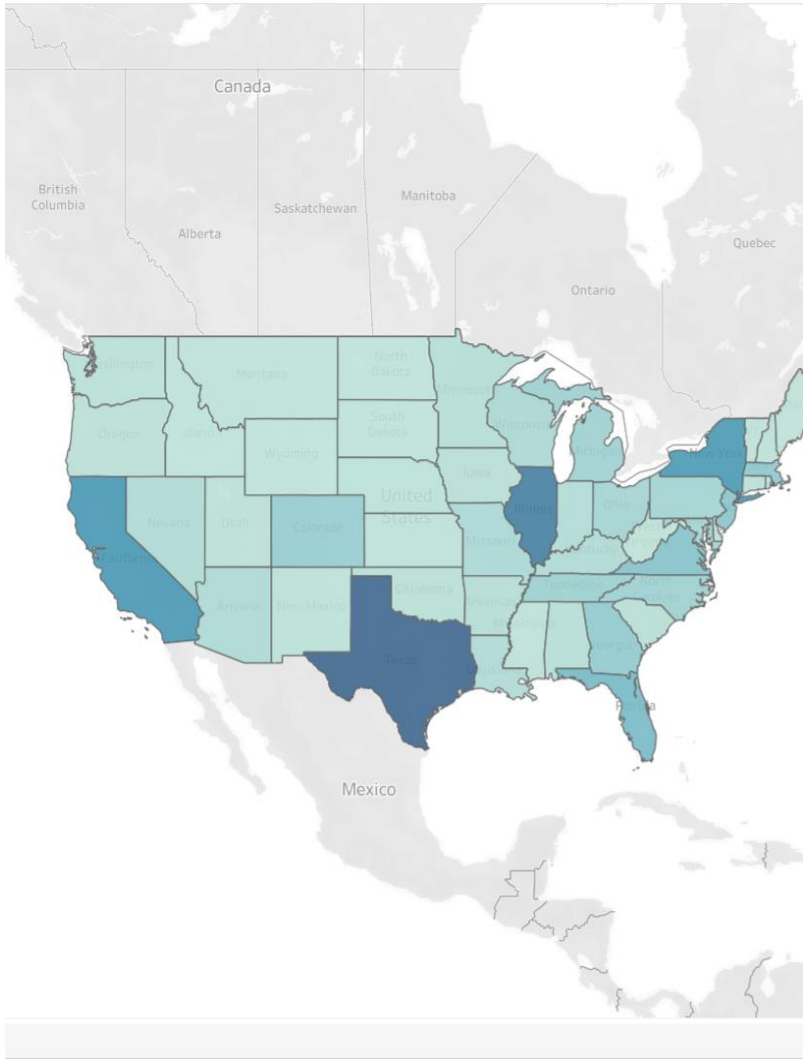
EDA and Data Cleaning

Before I went to tableau, I looked at the data trying to know more about it

Here's what I did:

- removed 3 unnamed columns that weren't in the Kaggle dataset
- renamed values in the 'CANCELLATION_REASON' column in the "flights" dataset from A,B,C to understandable reasons which can be more useful
- an HTML file (jupyter notebook) is provided for all mentioned above

cancellations

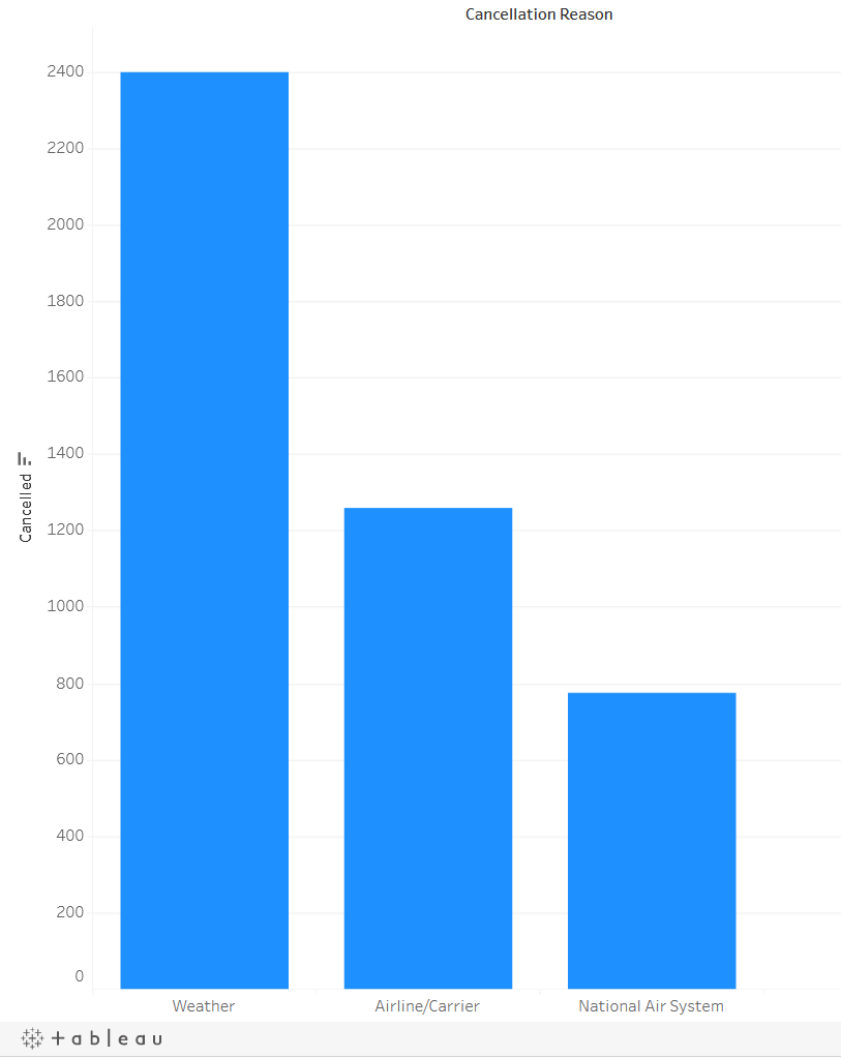


Sheet 1

- First, I looked at the relationship between states and cancellations, to find that Texas has the most flight cancellations while West Virginia has the least flight cancellations

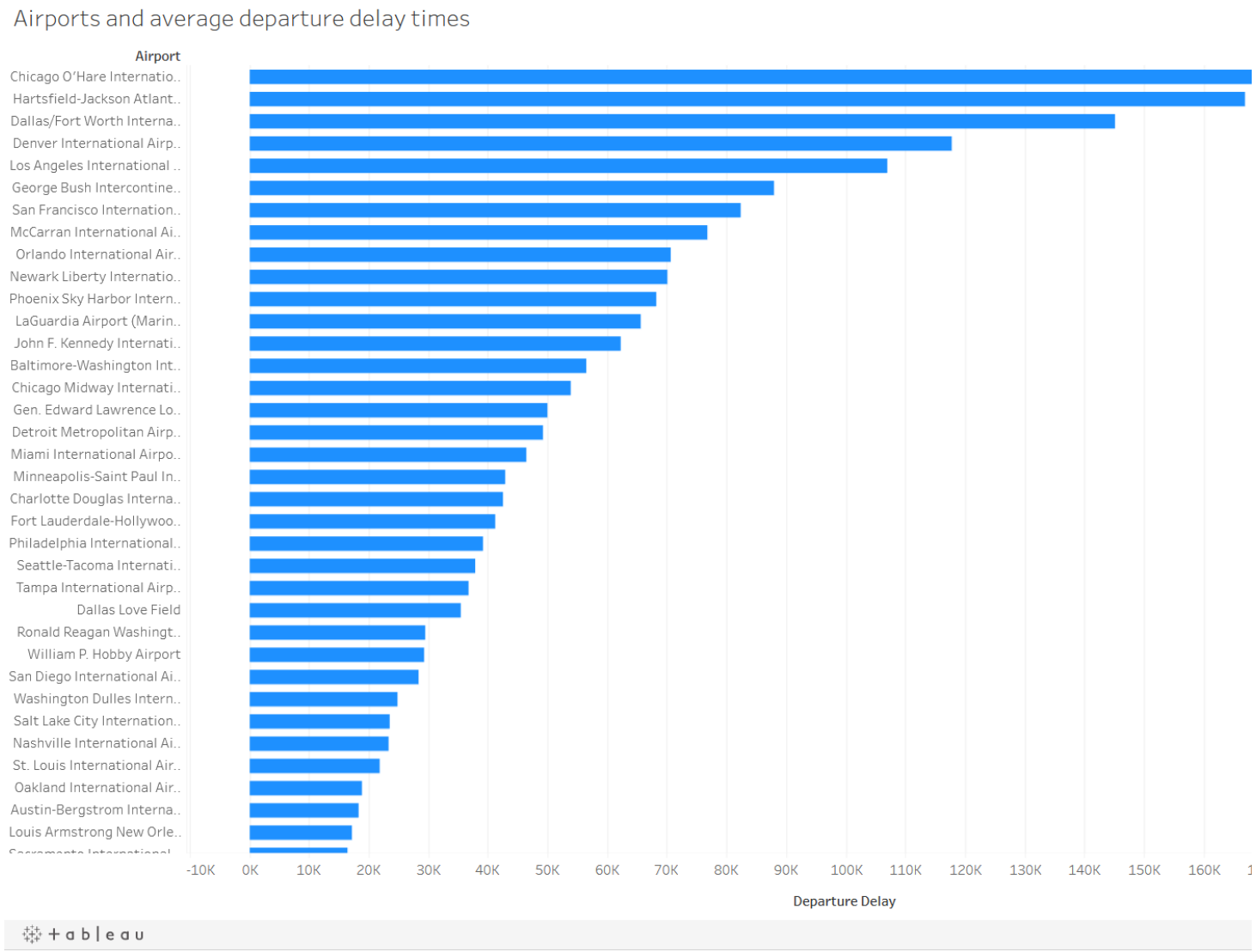
- Sheet 1 link

what causes cancellations?



Sheet 2

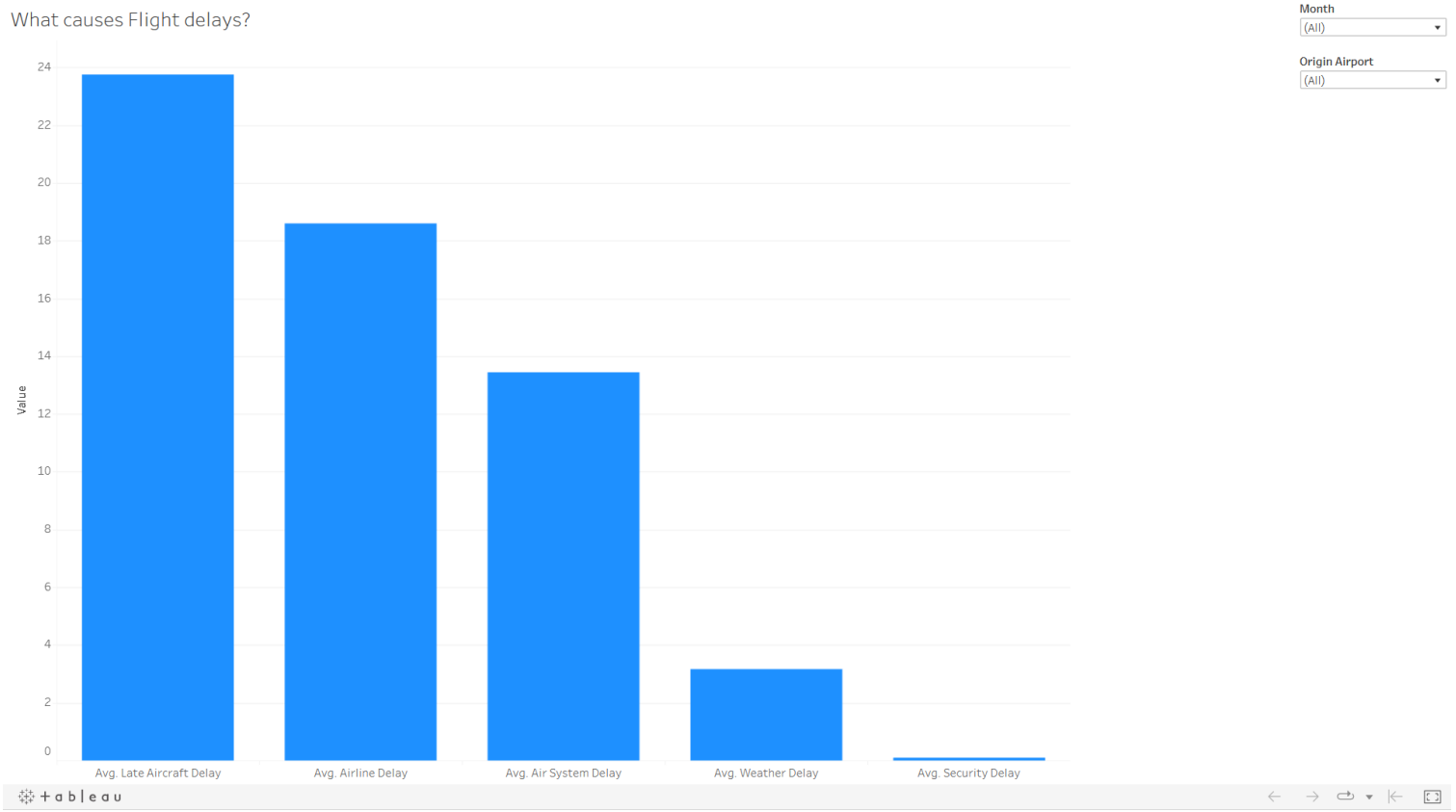
- Then, I got curious to know what caused these cancelations to find out that the weather was the biggest reason
- Sheet 2 link



Sheet 3

- After that, I looked at airports and average departure delay times and the worst airport in terms of departure delay was Chicago O'Hare International Airport

- Sheet 3 link



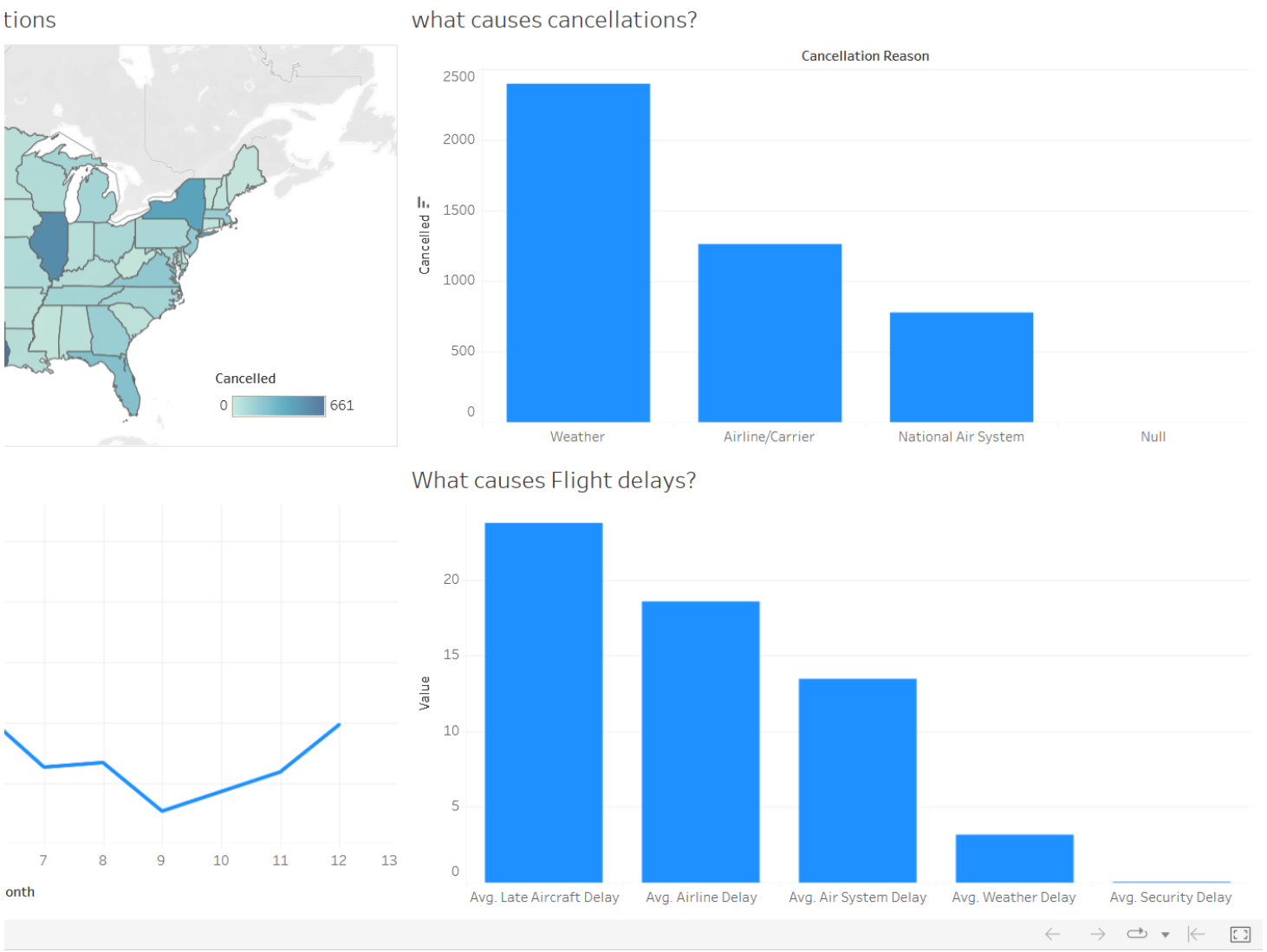
Sheet 4

- Then, I looked at the reasons of these delays and what causes delays, and I found out that what caused delay the most was late aircraft delay

- Sheet 4 link

Sheet 5

- Lastly, I wanted to know what time of year has most cancellations, and I found that at February flights get cancelled the most (1058 cancelled flights)
- Sheet 5 link



Dashboard

And I put everything together into a useful, clean dashboard

- dashboard link

Design

- In all of my visualizations I went with the dodger blue color as t’s a cool color, comfortable to look at and won't be a problem for colorblindness
- Didn’t use 3D plots as they can be misleading and not very helpful
- Used a color legend in map plot to determine max from min
- didn’t use multiple colors for each bar in bar plots
- overall tried my best to make the visualizations as clean, simple and informative as possible

Resources

N/A