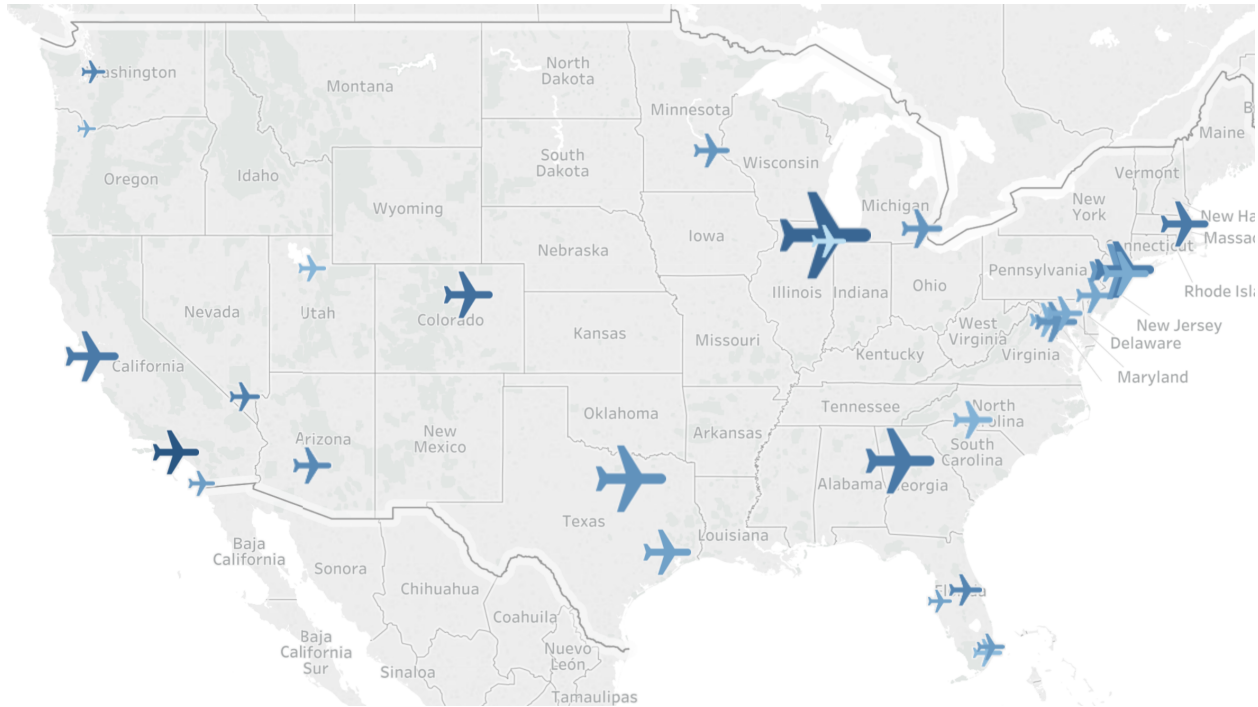# Performance Analysis of Airlines and Airports in USA

A Data Visualization Story created using Tableau

Dilip Rajkumar

May 09th, 2018

Udacity - Data Analyst NanoDegree

TABLEAU STORY LINKS

- First Draft - https://public.tableau.com/shared/GTGHQNQ99?:display_count=yes

- Final Version - https://public.tableau.com/shared/JX27N2GXP?:display_count=yes

SUMMARY

Air Travel in the USA has grown rapidly over the last few decades. Passengers experience delays caused by Airport Protocol Procedures, Airline Career or other factors. In this data visualization project we used Air Traffic Data for the last 10 years (from 2008 - 2018) from the US - Bureau of Transportation Statistics. The focus of this explanatory visualization is to investigate performance metrics such as Flight Arrivals , Cancellations and various delays by airport and airline carriers, so that the airline passenger can make insightful decisions for their next air trip. Some of the key findings are:

- ORD (Chicago), ATL (Atlanta) and DFW (Dallas) airports constantly top the list across most measures. For some measures like *Security Delay* and *National Aviation System Delay* we see LAX (Los Angeles) and SFO (San Francisco) airports respectively, appearing in the Top 3 list. These are also the most busiest airports in terms of passenger traffic.
- Looking at airline carriers, the general conclusion we can arrive at is that airline careers with high number of *Flights Arrived* (i.e a high volume of air routes) also seem to have the highest delays. Southwest Airlines Co., SkyWest Airlines Inc, American Airlines Inc. , United Airlines Inc., American Eagle Airlines Inc, US Airways are all airlines with both high Flight Arrivals and Arrival Delays.
- *Arrival Delay* was found to be the aggregate of all the 5 delays like *NAS Delay*, *Security Delay*, *Carrier Delay*, *Late AirCraft Delay* and *Weather Delay.* Amongst these 5 delays, *Late AirCraft Delay, Carrier Delay* and *NAS Delay* had the highest percentage of total contribution to Arrival Delay.

DESIGN

The entire visualization was created in a user driven approach with *dynamic axis selections*. This helps the user to investigate the dataset at a finer granularity, such as across different points in the timeline (months and years) or for different *measures*.

Whenever possible a color palette (Oranges and Blues) suitable for Color Blind people were used, but occasionally we also used non color blind colors for better contrast and visual aesthetics.

FEEDBACK

I showed the visualization to a colleague in the university to solicit feedback. The following were the points gathered from his feedback.

1. **Horizontal Bar Chart Vs. Radial Bar Chart**
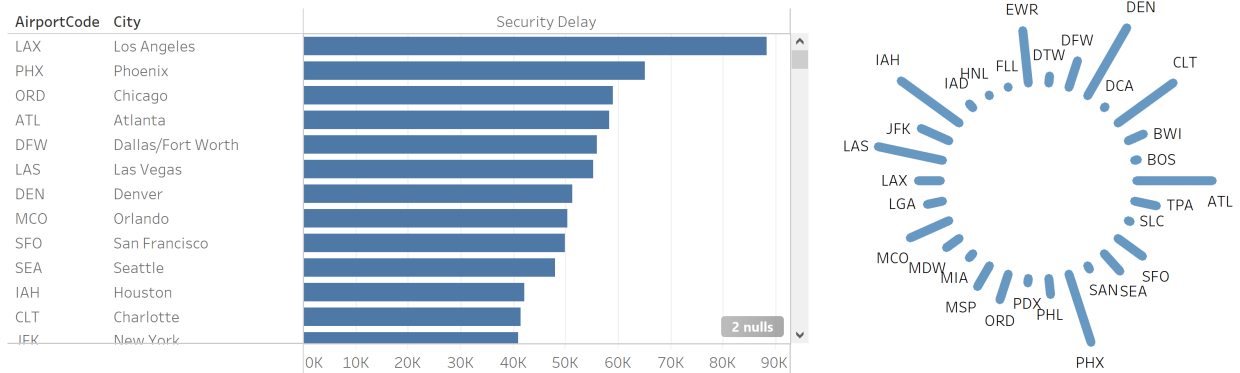
To showcase the performance metrics of various airports (categorical variables) I made both a Horizontal Bar chart and a Radial Bar chart. My friend suggested though the order of values and detail is very clear in the horizontal bar chart, the *radial bar chart was preferable* because:

- A Radial Bar Chart can can pack more data density for the same screen space. For eg, we see data for 30 airports compared to just 15 airports in the Horizontal Bar Chart (*left*).
- The general idea of value range and scale is still maintained in the Radial Bar Chart and the hover tool shows the exact value when the mouse pointer moves a radial bar.
- A Radial Bar Chart is not widely used and also has more visual appeal when animating time series data.

2. **Geo Map Plot vs. Bubble Plot**
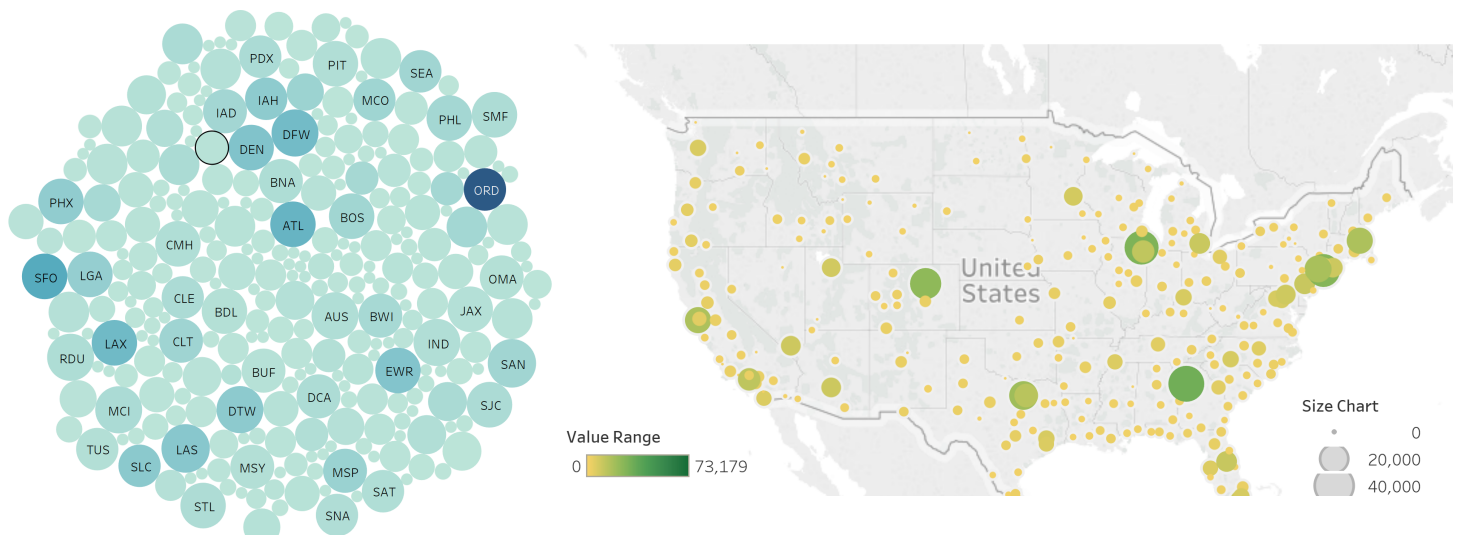
A Bubble plot with varying bubble sizes and colors was first considered to show the various performance metrics values of all the airports. However this did not convey the geographic location of the airports and consequently my friend suggested using a Geo Map plot to show the location of all the airports and over lay the measure values using circle makers.

3. **Device Resolution**

Originally i designed all the story dashboards for a HD resolution of 1920 x 1080 pixels, as my laptop had a 4K screen resolution. But when i asked my friend who to view my visualizations, they looked a little out of order and required scrolling. My friend's laptop was having an older laptop with a smaller screen size and a lower screen resolution (1366 x 768).

4. **Intro and Conclusion Slides**

An Introduction Story point giving a short summary introduction to the visualization story and a conclusion story point suggesting a *call of action* to the reader was suggested by my friend.

CHANGES AFTER FEEDBACK

- Removed Bubble Plot of Airports and replaced it with Map Plot.
- Removed Horizontal Bar Chart of Airports Performance Measures and used a Radial Bar Chart to show data for top 30 airports only as many smaller airports don't have any significant values for many measures. Radial Bar Chart was deemed more aesthetic.
- Changed the Device Setting and resolutions of all the Dashboards from Full HD to 1280 x 800, so that the visualizations are optimized for viewing across smaller screens with lower resolutions.
- Added Introduction and Conclusion story points to improve the flow of information and aesthetic appeal of the Tableau Story.

RESOURCES

- http://www.tableaulearners.com/radial-barchart-in-tableau/
- http://www.vizwiz.com/2012/07/tableau-tip-dynamic-axis-selections.html
- http://kb.tableau.com/articles/issue/dynamic-axis-titles
- https://public.tableau.com/en-us/s/blog/2017/02/mapping-real-time-flights-advanced