

## **Project 4: Build Data Dashboards – based on the data of On-time Performance of US Domestic Flights Operated By Large Air Carriers in 2015**

**By Phyllis Yu Yu Mo**

In this project, I made use of Tableau data visualizations to draw insights for various questions I wondered about the delays of US domestic flights in 2015. For each dashboard or visualization, I will summarize my main findings, explain my design choices and include the resources I used.

### **The first insight: Which states/ cities have the highest amount of arrival delays (in terms of minutes)?**

Link:

[https://public.tableau.com/views/sumofarrivaldelaysperstatesorcities/Dashboard1?:language=en-US&publish=yes&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/sumofarrivaldelaysperstatesorcities/Dashboard1?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link)

Summary:

I have created a dashboard, which includes a map, a bubble chart and a bar chart to look into answers for this question.

The map in top left corner shows the number distribution of arrival delays across different states. We can see that Texas has the highest number of arrival delays (215,553 minutes), followed by California (163,218 minutes) and Illinois (149,246 minutes).

The bubble chart in the top right corner shows a narrower view of arrival delays as it shows the top 10 cities with the highest number of arrival delays. It shows most arrival delays occurred in the city of Chicago (145,336 minutes), followed by Dallas-Fort Worth (90,378 minutes) and Houston (73,976 minutes).

Linking the map to the bubble chart, we can see most arrival delays in Texas occurred in their cities of Dallas-Fort Worth and Houston, whereas in California it mostly occurred in Los Angeles, Las Vegas and San Francisco.

In the bottom of the dashboard, I have included a bar chart, which allows us to compare the amount of arrival delays by cities.

I have included a filter where the user can see how the number of arrival delays vary across months for states/ cities. The insights I drew above were based on selecting all 12 months.

Design:

My goal is to allow users to draw insights from a board to a more specific view, therefore my dashboard starts with a map which showcases all the states, then it narrows down to the data of cities in a bubble chart. Including the state abbreviations in the bubble chart facilitates my analysis of the data. Finally, I choose to include a single-colour bar chart to display only data of

the top 10 cities with the highest amount of arrival delays in a descending order, as the users can compare the amount of arrival delays between cities in an easy way, without the distraction of different sizes or colours.

In terms of color of the dashboard, I choose not to use red-green colours as they are not suitable for color blindness.

Resources:

Udacity Knowledge Forum

### **The second insight: How did the arrival delays (in terms of minutes) change over months in 2015?**

Link:

[https://public.tableau.com/views/theamountofarrivaldelayacrossmonthsin2015/Sheet3?:language=en-US&publish=yes&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/theamountofarrivaldelayacrossmonthsin2015/Sheet3?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link)

Summary:

The line graph shows how the amount of arrival delays in terms of minutes changed over 12 months in 2015. We can see the highest amount of arrival delays occurred in June (248,586 minutes). The second highest point occurred in February (179,573 minutes). It also shows there was downward trend from February to April and June to September, which is the lowest point (-17,496 minutes). It is a negative number, which means the flights in that month arrived earlier than the scheduled time.

I also included a filter where users can see how the number of arrival delays vary across airlines. The insights I drew above are based on selecting all the airlines.

Design:

I choose to use the line graph as it showcases the changes of a variable (sum of arrival delays in minutes) over time in a clear manner. It also helps users look for trends.

I also use only one color for the line, as there is no extra information needs to be highlighted by different colors. I avoid using red-green colour as it is not suitable for colour-blindness.

Resources:

Udacity Knowledge Forum

### **The third insight: What are the main reasons of delays at each airport?**

Link:

[https://public.tableau.com/views/theamountofeachdelaytypeperairport/Sheet1?:language=en-US&publish=yes&:display\\_count=n&:origin=viz\\_share\\_link](https://public.tableau.com/views/theamountofeachdelaytypeperairport/Sheet1?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link)

Summary:

According to the post (<https://knowledge.udacity.com/questions/316109>) on Udacity forum, there are 5 types of delay, including air system delay, airline delay, late aircraft delay, security delay and weather delay.

The stacked column chart I created shows the reasons behind delay at top 10 airports (in terms of the total of these 5 types of delays). For example, the top 3 reasons for causing the delay at Chicago O'Hare International Airport are late aircraft delay, airline delay and air system delay. They are also the main delay types at the other 9 airports.

We can see weather delay and security delay are not the main reasons behind the delay, compared to the other 3 delay types.

I also included a filter where users can see how the amount of various delay types vary across airlines. The insights I drew above are based on selecting all the airlines.

Design:

Stacked column chart is chosen because my focus is to compare the totals of various delay types and amount of the individual delay type across different airports. I also arrange the chart in a descending order of the totals of 5 delay types, which allows users to compare the numbers between airports more easily.

I choose 5 different colours to depict different delay types as delay types is important information need to be conveyed in an explicit way. Same as the other two visualizations, I avoid red-green colours, as they are not suitable for colour blindness.

Resources:

Udacity Knowledge Forum