

# Create a Tableau Story

Darren Chiu

## Links to Story

First Version:

[https://public.tableau.com/profile/darren.chiu#/vizhome/Flights\\_Delay\\_Analysis\\_v1/Story\\_1](https://public.tableau.com/profile/darren.chiu#/vizhome/Flights_Delay_Analysis_v1/Story_1)

Final Version:

[https://public.tableau.com/profile/darren.chiu#/vizhome/Flights\\_Delay\\_Analysis\\_v3/Flight\\_DelayAnalysis](https://public.tableau.com/profile/darren.chiu#/vizhome/Flights_Delay_Analysis_v3/Flight_DelayAnalysis)

## Summary

The story aimed at analyzing the relationship between travel period and delay time, i.e. does a specific period in the year having more flight delays. After some brief investigations, we found that there are significant more delay for flights in December and we thus further analyse the data and found out the delay is mainly due to weather instead of a higher travel frequency in December.

We also identified that coast city airports are more impacted by the winter weather compare to inland city airports.

We also come up with some holiday season travel tips for readers on dates and airlines to avoid.

## Design

The story was designed to follow the investigation thought process to facilitate understanding of readers. Each visualization was following the logic of the previous one.

A line graph was plotted at the beginning to demonstrate the change of delay time throughout the year. Line graph was chosen because it can show the variation of a variable across time and allow readers to see the changes of delay time easily. Both Arrival and Departure delay were plotted with dual axis. Dual axis was chosen because it can help readers find out if arrival and departure delays varies together.

Next, a bar chart of number of flights across the year is used as bar charts are good for comparing values and in understanding the distribution of data. They can easily summaries large data sets in a single visual.

To understand the breakdown of delay reasons, 5 line graphs are plotted in parallel for comparing the variation of different delay reason across the year. We didn't plot all of them on the same graph using technique like dual axis because 5 lines might make the graph too complicated for easy comparison. To compensate this disadvantage, we plotted them in parallel, which allow some level of comparison without sacrificing the clarity of graphs.

A dual axis line graph was also plotted for actual and CRS Elapsed time. Dual axis line graph is again chosen for its property of easy comparison as we were aiming at finding out the difference between actual and CRS elapsed time.

To further investigate the reason of the delay, the delay was plotted on a map, encoded by color and size. Comparing to using chart of airport code against delay time, this allow us to pinpoint the relationship between geographical locations and the delay time. As a result, we could conclude that coast city airports are more impacted by weather. The arrival and departure delay time are encoded in different marks (bubble size and color) to allow displaying both information together on the same graph at the same time. To enable readers to do some investigations on their own, we provided some filters so readers can choose the range of delay they want to look at. For example, they can focus on 90<sup>th</sup> percentile instead of 50<sup>th</sup> percentile if they want. By dragging the filters around, they can also see the variation of location and delay time.

Two more dual axis line graph of arrival and departure delay is plotted to help providing suggestions on day of month and hours suggestions for readers. Line graph is also chosen again for its ability to show variation across time so readers can easily find out the best timing for taking a flight. Month filters is provided as a tool to investigate best timing for different months, as we only investigated one of the three delay peaks throughout the year.

At last, horizontal bar charts of airline delay time was plotted, and sorted based on arrival delay. Horizontal bar charts are good in this use case as it allow easy comparisons. The graph is sorted to allow even easier in descending order to help spotting the best and worst airlines.

After feedback from the readers, a few changes were also made including adding a timing (hours) analysis page, add filtering for average Dep Delay time on the map and better title and labelling. These are further discussed in the next section.

## Feedback

Several feedbacks were received from my friends:

1. Add timing selection page  
Feedback mentioned that it would be nice if I can add a page for analyzing the delay time distribution across the time within a day. It would be a nice addition to the airline and dayOf Month analysis
2. Add filtering for average Dep Delay on the map  
On the map, there are no way to filter the Dep Delay time so the conclusion that coast city airports are more prone to weather impact are not well-argued.
3. Better title and labeling  
It would be nice to add more labelling or title on the visualization, e.g. showing the airport code on the map.

#### 4. More interactive features

It was also suggested by Udacity reviewer that more filters and interactivity would help readers do their own investigations more effectively. So, some filters of destination and origin filter was added on the three suggestions graphs, which aiming at making the three pages a best flight searcher tool for reader instead of just a data analytic tool.

#### Resources

##### Airlines Code

<https://www.tylon.com/resources/airlinecodes.htm>

##### Tableau Usage

[http://onlinehelp.tableau.com/current/pro/desktop/en-us/maps\\_geographicroles.html](http://onlinehelp.tableau.com/current/pro/desktop/en-us/maps_geographicroles.html)

<https://www.interworks.com/blog/ccapitula/2015/04/15/tableau-essentials-calculated-fields-date-functions>