Udacity Data Analyst Term 2 Tableau Project
Z. McLaughlin

Main file can be found:

- https://public.tableau.com/profile/z.mcl#!/vizhome/ad_4_1/FlightDelays?publish=yes

For additional information on the project, see below

Links to the first version of your story (e.g. story_v1) and final version of your story

- First Version: https://public.tableau.com/profile/z.mcl#!/vizhome/ad_test_2/UAFlights?publish=yes
- Final Version: https://public.tableau.com/profile/z.mcl#!/vizhome/ad_4_1/FlightDelays?publish=yes

Summary: brief description of the visualization and the main story or findings conveyed

Looking at a 20 year span from 1988-2008 using the data from:
http://stat-computing.org/dataexpo/2009/the-data.html
demonstrated that the number of flights an airline has does not influence how often the flight arrival is delayed. In fact the number of delays are regular consistent to the point that they seem to be standard operating procedure.

Design: explain any design choices you made including changes to the visualization after collecting feedback

Because of the size of the data involved I could not look at all flights over the 20 year period. So I selected three years 1988, 1998, and 2008 to give a nice span of data. This data was still too large, so originally I selected to look at just United Airlines data. I frequently flew United during this period of time in and out of its hub at SFO (San Francisco Airport.) After getting feedback, I found that people got very bogged down by the fact that United has far fewer flights in 2008 than it did in 1998, and didn't find my findings very interesting. I decided to change the analysis to include the top 4 airlines in North America based on Wikipedia (https://en.wikipedia.org/wiki/List_of_largest_airlines_in_North_America):

- American Airlines
- Delta Airlines
- South West Airlines
- United Airlines

to show some of the other information over the 20 year period such as the fact that there was up and coming airline, South West, that was taking over some of the airline travel. I also decided to add in a few more visuals regarding delays by year per airline and delay times.

I also added a little bit of information in the conclusion regarding some of the bigger events that may have caused the downturn in the number of flights for the big airlines. It could have been interesting to dive into airline travel around 9/11, but I wasn't particularly interested pursuing that explration. If I had been able to use a larger dataset with Tableau it could have been interesting to look at the data year by year, but I was limited to 15 million rows of data.

Feedback: feedback received from others from the first sketch to the final visualization

Feedback # 1
I saw the visualization. I am surprised.

1. There is a decline in the number of flights from the '80s to '08s.
2. Delays are not related to the number of flights. It does appear that delays are seasonal namely summer and winter.

I am curious as to why there would be a decline in the number of flights. I think it is due to increase in flight efficiency and creation of major airport hubs. The question I have is: How does it compare to the big 3? i.e. Delta and AA.

It would be interesting to find out what happened immediately after 9/11 with respect to 2008. There has been a sharp increase from 88 to 98. If there was data from 2000 and post 9/11, I wonder if it is possible to attribute the sharp decline to post 9/11 concerns.

The thing that really surprises me is that ORD is a United Hub and the sharp decline in traffic was observed at ORD.

Feedback # 2 (This was a discussion so below is a summary)

There's no way that United has less flights in 2008 than it did in 1998, why would that be?
Delays aren't surprising.
Maybe you need to do more research on why United has less flights.

Resources: list of Web sites, books, forums, blog posts, GitHub repositories etc that you referred to or used in this submission (Add N/A if you did not use such resources).

Used python jupyter notebook to pre-clean the data down to what I was interested in.
Used Tableau documentation
Used https://stackoverflow.com/