Marketing Analytics Nanodegree Program

Project 4
Build Data Dashboards

By Audrey Tan
10/25/2019

Visualization I – 2015 On-time Performance Trends of Domestic Flights at 17 Major US Airports: Part 1

Introduction
In this visualization, I analyzed a Kaggle dataset tracking the on-time performance of US domestic flights in year 2015 to find out delay trends at 17 major US airports: ATL,CLT, DEN, DFW, IAH, JFK, LAS, LAX, MCO, MIA, MSP, ORD, PDX, PHL, PHX, SEA and SFO.

Analysis Approach
I looked at the 27,4964 records in the dataset to find the on-time performance trends of at the 17 airports as these airports accounted for 50.39% of the total delays in the dataset.
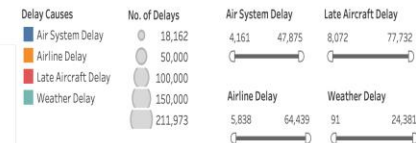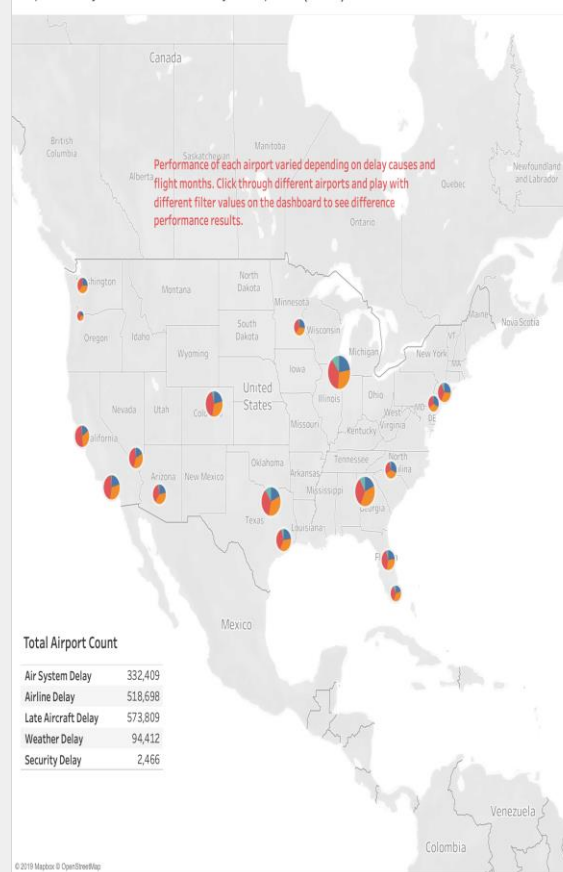
Design Choices
I built a map (encoded by size and color) of delayed flight counts with duration > 0 at the 17 major airports broken into 4 major delay causes, and a time-series chart (encoded by color) of 12-month trends on 4 major delay causes. Linked filters were provisioned on the 2 charts, to support animated view of different delay trend perspectives.
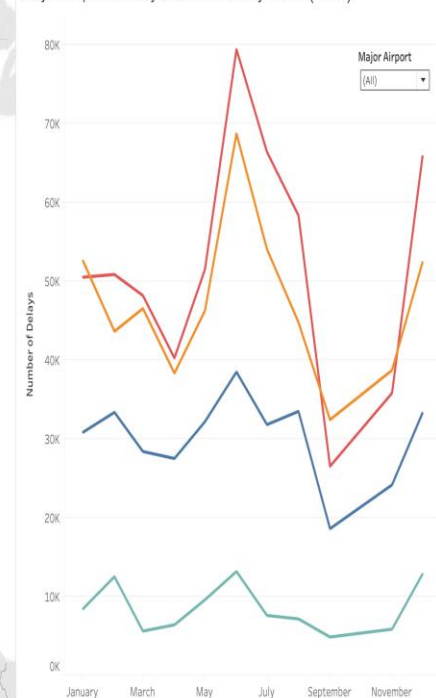
Insights
The visualization shows the on-time performance of the 17 major airports was impacted by 4 major delay causes - Air System, Airline, Late Aircraft and Weather Delays. Of the 4 causes, Late Aircraft and Airline delays were the top two culprits for the delays. The delay causes were less pronounced in April and September, and otherwise in June and December (this is in line with the seasonal effect of domestic air travels in the US).

Map of Delays and Causes at Major Airports (2015)

Performance of each airport varied depending on delay causes and flight months. Click through different airports and play with different filter values on the dashboard to see difference performance results.

Delay Causes
- Air System Delay
- Airline Delay
- Late Aircraft Delay
- Weather Delay

No. of Delays
- 18,162
- 50,000
- 100,000
- 150,000
- 211,973

Air System Delay		Late Aircraft Delay	
4,161	47,875	8,072	77,732

Airline Delay		Weather Delay	
5,838	64,439	91	24,381

Total Airport Count

Air System Delay	332,409
Airline Delay	518,698
Late Aircraft Delay	573,809
Weather Delay	94,412
Security Delay	2,466

© 2019 Mapbox © OpenStreetMap

Major Airports Delay Causes Monthly Trend (2015)

Major Airport (All)

(Number of Delays chart, January–November)

Sample Use Case Scenarios of the Linked Filter Effect
Below are 2 scenario examples on how to use the linked filters on the map and time series charts:

*EX. 1 – Filter on a Delay Cause*.
Click the orange "Airline Delay" pie segment of ORD airport on the Map to see the "Airline Delay" frequency of 47,875 at ORD airport, as well as the matching 12-month trends of delay causes on the Time-Series Line chart.

EX. 2 – Filter on Major Airport.
Select Chicago O'Hare International Airport on the "Major Airport" filter box on the Time-Series Line chart, to see the 12-month trends of the major delay causes at ORD airport, as well as the frequency counts of the 4 causes and their visual representation on the pie chart at ORD airport on the Map.

Tableau Public Links
Vis I Dashboard
Vis I Map worksheet
Vis I Table Chart worksheet
Vis Time-Series Line chart worksheet
Vis Storyboard
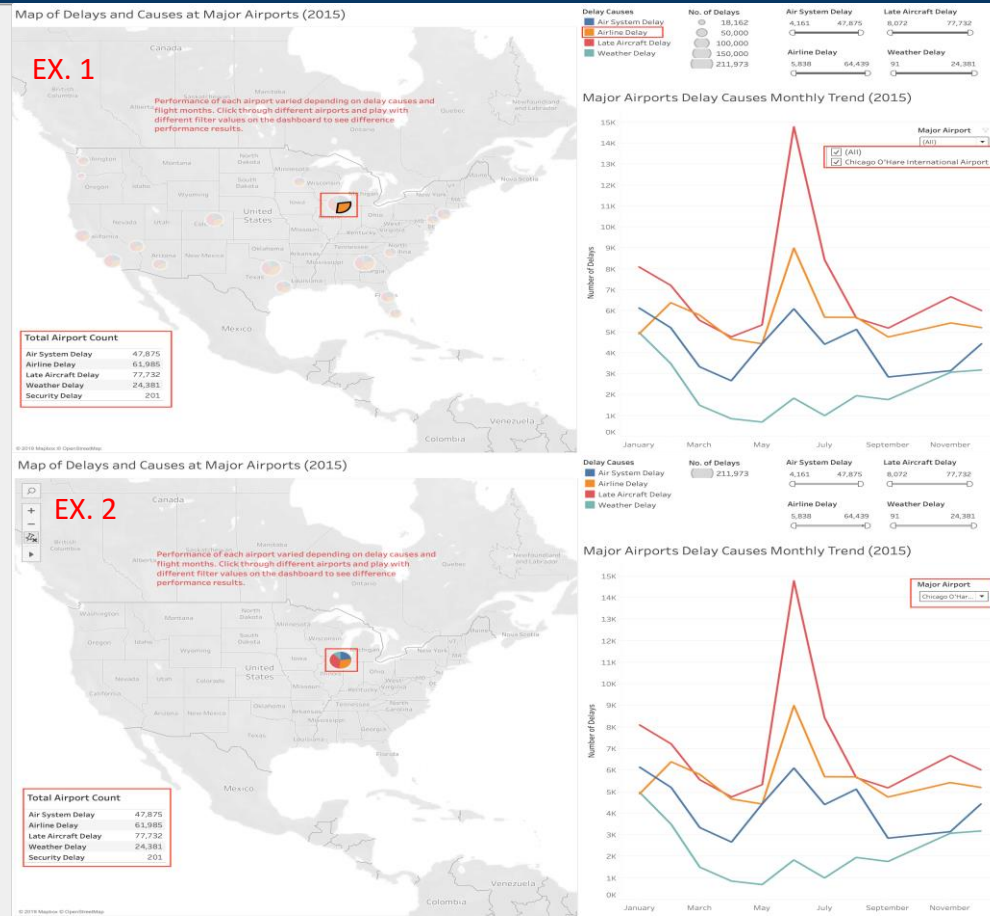
Introduction
In this visualization, I looked at the domestic flight performance data of the 17 major US airports, to find out the correlations between departure performance and the major causes driving the delays.

Analysis Approach
I examined the delay counts of the major delay causes from the dataset to see their effects on the departure performance and what factor, if any, might exacerbate these effects at the 17 airports.
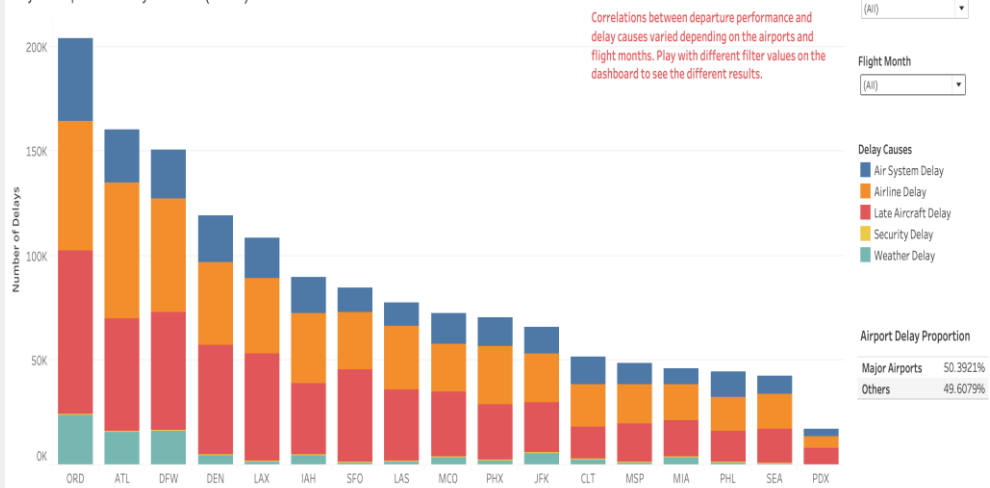
Design Choices
I created a stacked bar chart to show the counts of major delay causes (encoded by color) at the 17 airports. A scatter plot was also added to reveal the correlations between departure delay (duration > 0) and the major delay causes. Linked filters were added to the 2 charts, to allow animated view of varying correlation trends based on flight month and/or airport location value filtering. You can also click on a stacked bar on the bar chart to see the count of a delay cause at an airport, along with a matching scatter plot on delay duration.
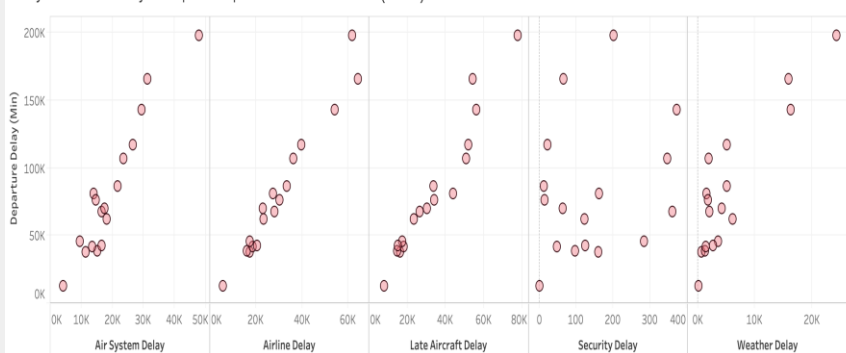
Insights
The visualization shows Airline and Late Aircraft delays had the most impact on departure performance, both had a clear positive correlation with departure delay duration. So was Weather delay but less clear for Security which had the least impact. Some Flight months (e.g. February, April, June, December) seemed to amplify the effects of the delay causes.

Major Airports Delay Causes (2015)

Correlations between departure performance and delay causes varied depending on the airports and flight months. Play with different filter values on the dashboard to see the different results.

Origin Airport
(All)

Flight Month
(All)

Delay Causes
- Air System Delay
- Airline Delay
- Late Aircraft Delay
- Security Delay
- Weather Delay

Airport Delay Proportion
| Major Airports | 50.3921% |
| Others | 49.6079% |

Delay Causes on Major Airport Departure Performance (2015)

Visualization II – Correlations between domestic flight departure performance and underlying delay causes at 17 Major US Airports: Part 2

Sample Use Case Scenarios of the Linked Filter Effect
Below are 3 scenario examples on how to use the linked filters on the stacked bar chart and scatter plot:

EX. 1 – Filter on Origin Airport.
Select one or more airport location in the Origin Airport filter, to see count distributions of delay causes of the filtered airport(s) on the stacked bar chart, and the matching scatter plot on delay duration and causes.

EX. 2 – Filter on Flight Month.
Select one or more month in the Flight Month filter, to see a different view of count distributions of delay causes of the filtered flight month(s) and the associated scatter plot on delay duration and causes.

Ex. 3 – Filter on a delay cause or a scatter plot spot.
Click on a stacked bar of an airport on the bar chart or a scatter plot spot to see a different view of the filtered criteria.

Tableau Public Links
[Vis II Dashboard](#)
[Vis II Stacked Bar Chart worksheet](#)
[Vis II Table Chart worksheet](#)
[Vis II Scatter Plot worksheet](#)
[Vis Storyboard](#)

Introduction
In this visualization, I compared the departure delay duration distribution of the 17 major US airports, to give a compact, quickly assimilated summary of the data.
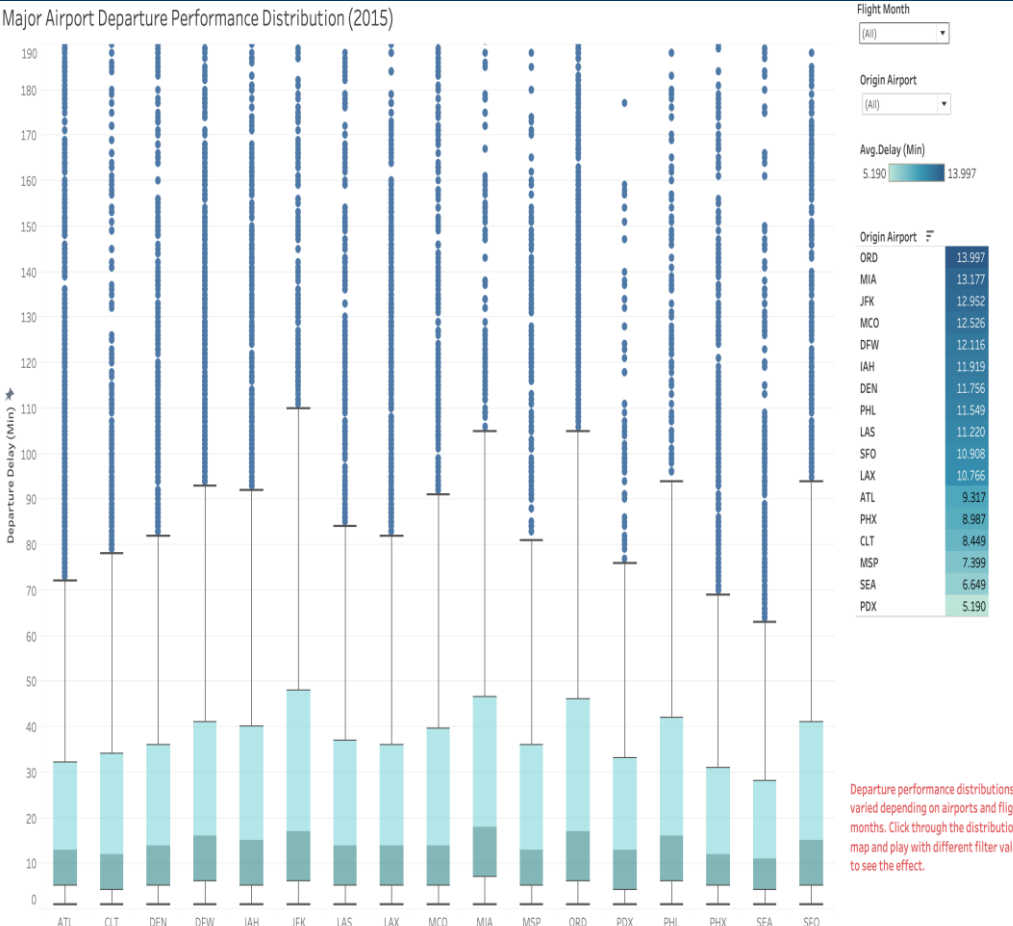
Analysis Approach
I inspected the shape of departure delay data distribution, its central values and variability among the airports, then compared the findings to draw an association between the delay duration and the airports.

Design Choices
A box plot is the natural choice for inspecting the shape of departure delay (duration > 0) , its central values and spread among the airports. I also included the average and standard deviation departure delay duration of the airports in a table, as another potential measure of spread. Linked filters on flight month and origin airport were added to support animated view of different distribution perspectives based on filtered values.

Insights
The visualization shows a right skewed distribution, with median < mean delay duration. The median, IQR and range values are reasonably similar, especially among same tiered airports (e.g. JFK & ORD & MIA & SFO, DFW & IAH, PHX & SEA, etc) across 12-month period. Filtering on certain busy flight months (e.g. June, December) shows an even closer similarity in centers and spreads among the same tiered airports. This suggests a consistent association between the airports and departure delays, with same tiered airports having similar departure delay durations.

Major Airport Departure Performance Distribution (2015)

Flight Month
(All)

Origin Airport
(All)

Avg.Delay (Min)
5.190 — 13.997

Origin Airport	
ORD	13.997
MIA	13.177
JFK	12.952
MCO	12.526
DFW	12.116
IAH	11.919
DEN	11.756
PHL	11.549
LAS	11.220
SFO	10.908
LAX	10.766
ATL	9.317
PHX	8.987
CLT	8.449
MSP	7.399
SEA	6.649
PDX	5.190

Departure performance distributions varied depending on airports and flight months. Click through the distribution map and play with different filter values to see the effect.

Sample Use Case Scenarios of the Linked Filter Effect
Below are 2 scenario examples on how to use of the linked filters on the box plot and table chart:

EX. 1 – Filter on Flight Month.
Select one or more month in the Flight Month filter, to see a different view of departure delay distribution on the linked box plot and table chart based on the filtered flight month(s).

EX. 2 – Filter on Origin Airport.
Select one or more airport location in the Origin Airport filter, to see a different view of departure delay distribution on the linked box plot and table chart based on the filtered origin airport(s).
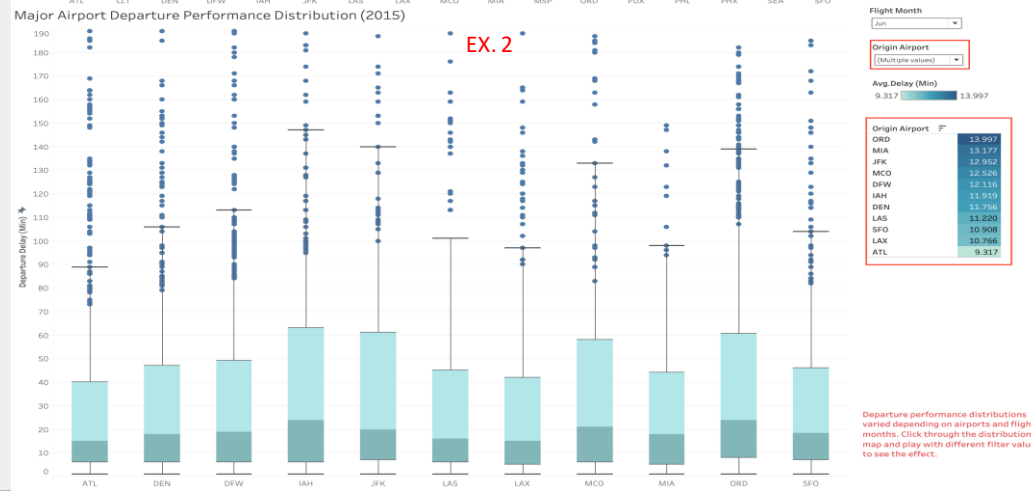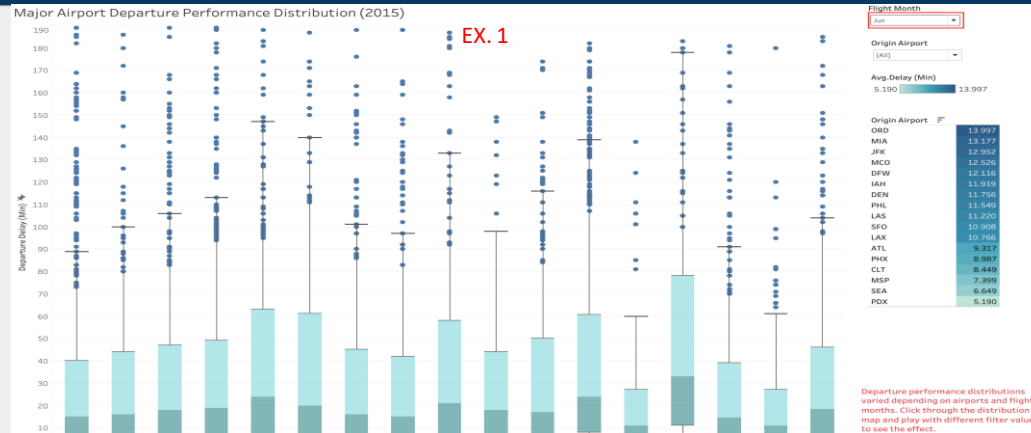
Tableau Public Links
Vis III Dashboard
Vis III Box Plot worksheet
Vis III Table Chart worksheet
Vis Storyboard

References

MAND Data Visualization

Lesson 4 Data Visualizations in Tableau
- Video: Marks & Filters

Lesson 5 Make Dashboards and Stories in Tableau
- Video-Building Dashboards and Stories with Trina
- My First Tableau Story

Tableau Public Gallery
- Time Series Line Chart
- Symbol Map with Pie Marks

Tableau Desktop and Web Authoring Help
- Build a Box Plot
- Set Action