Kyle Wilbert
Udacity Data Analyst Nanodegree
25 January 2019

Visualizations

Version 1

workbook wouldn't load on Tableau Public anymore and I didn't have a version locally, so I deleted it off my profile

Version 2

https://public.tableau.com/views/2008RITAFlightsDataAnalysisv2/TotalFlightsArrivinginNYCArea?:embed=y&:display_count=yes

Version 3

https://public.tableau.com/shared/Q6YCG7N5F?:display_count=yes

Final

https://public.tableau.com/views/OverviewofNYCairportsin2008Final/OverviewofNYCareaairportsin2008?:embed=y&:display_count=yes&publish=yes

Summary

The visualization is an overview of performance at the NYC area airports (Newark, John F Kennedy, LaGuardia) in 2008. It explores flight volumes for each airport, delays by day of week over time (mostly on Thursdays & Fridays and Summer), carrier competition, and destinations served (220 domestic destinations in 2008).

The dataset underlying this visualization is from RITA (Research and Innovative Technology Administration, within the US Department of Transportation). As the original dataset covering 2008 had more than 7 million rows, I extracted only the records where NYC airports were either the origin or the destination, resulting in a dataset of 752,855 records.

Design

Version 1

- Originally I was going to look at performance of the top 10 busiest airports by total number of flights
- Struggled to find a focus for the story
- After creating many somewhat related visualizations in one workbook, I was unable to complete a save to Tableau Public, and my desktop version crashed. I lost all my work
- At that point, I decided to look through the posters on the ASA site for this dataset to think about questions to ask and focus my story more

Version 2

- After Version 1, I established the backbone of the Final version here
- I abandoned sheet 9. I think I had planned on making an interactive map with flight paths. The data wasn't quite in the right configuration to make this work well.
- Left the project in worksheets and dashboards

Version 3

- Gathered the worksheets and dashboards into a story, adding captions to help establish a narrative and indicate insights gained from each visualization
- Settled on 1-2 visualizations per page, many with shared filters for interaction

Final

- Added airports.csv and airports.csv1 to data source; needed one source for Origin airports and the other for Destination airports
- Adding the source enabled me to replace airport codes with airport names throughout story
- Added a static stacked bar chart beside the interactive, filtered bubble chart, to enable better understanding of part-to-whole relationship between origin and destination airports (Card 4)
- Also added tooltip with airport destination name to bubble chart

Overall, I'm happy that I completed this project. It's fairly easy to get a large number of cool, good-looking visualizations with Tableau, but improving the data-to-ink ratio is fairly complicated and time-consuming. As a result, I compromised on many design choices simply because I didn't have the time or energy to dig really deep into customization.

In the first slide of my story, for example, I would have preferred to label the lines directly (airport code directly next to line, in corresponding color) so that I wouldn't need a legend. In looking through the forums for this fix, it appeared that the solution was complicated and kludgy, so I deferred making the change.

In the second slide, I thought side-by-side calendar heatmaps were an effective visualization showing an overview of flight volume and delays over time. The addition of the filter allows the user to look at all the airports together as well as each airport individually. I wish I had been able to synchronize the sliders on the visualization as that would have been more user-friendly than having to move them independently.

Next, the third slide shows carrier competition at each airport. I think the having both visualizations provides a more complete understanding. The line graph gives the user the big picture and the heatmap with data gives a more detailed understanding. Also, the ability to click each airline name, creates a highlight, which certainly helps get through all the colors.

I think the 4th slide simply is a neat way to show the variety of destinations as well as the number of flights and the originating airport. To me it seems to have a high data-to-ink ratio. Right off the bat, we can see that LGA has the four destinations with the highest amount of volume, despite the fact that it is the second busiest airport in the region. As noted above, after receiving feedback, I added a static stacked bar chart to show more clearly the part-to-whole relationship between each individual airport and number of destinations.

For the final slide, I wanted to tell a brief story about the large drop in number of destinations served that occurs in the Summer of 2008. Again I made the same compromise, keeping the legend rather than labeling the lines directly due to the complex solution required in Tableau.

Feedback

Vikram I. Mentor (Student Hub, 2019-01-25, 11:12 PM)

Hey @KyleW,
Excellent work with the exploration. I can already see that your project as it stands before this feedback is really awesome. :)

Let me suggest a few points that can help you get your project accepted:

- Avoid using abbreviations. For example, EWR, JFK, etc. As a viewer who has no prior knowledge of the dataset whatsoever, I expect the visualisations (charts, legends, title, descriptions) to provide all the information that I need
- You have also done a wonderful job of drawing insights and using them as descriptions. I suggest you add details about how using charts and their aesthetics helped you derive that inference in the write-up document under the design section.
- Finally, a general point and not related to the visualisation. As a code reviewer, I often see students not adding important findings in the summary section, do add that.
- Stacked bar charts are still a better tool when measuring quantitative stuff. That being said, the bubble chart is really interactive, I loved how the bubbles took the center stage when a filter was used :). However, I personally might choose a bar chart.

Stephanie A (Slack, 2019-01-25, 3:18 PM)

Nice story @Kyle Wilbert! You've got some really good charts, and it's a great look at the airports' activities. Just a couple things that popped to mind.
1. For the bubble chart it would be great to put the actual destination name in the tooltips as the 3-letter code is not intuitive. I think it's possible to do that, but I'm a bit rusty on Tableau!
2. It would be nice to have the full airport name in the colour legend for the airports.
3. It's interesting how the airports have different kinds of traffic.

The thing that I had difficulty with on my story was finding a main point and focusing on it; there is so much there!

All the best, Kyle!

Resources

Data Dictionary: http://stat-computing.org/dataexpo/2009/the-data.html
Supplemental Data Sources: http://stat-computing.org/dataexpo/2009/supplemental-data.html
Inspiration: http://stat-computing.org/dataexpo/2009/posters/

Create Date: https://onlinehelp.tableau.com/current/pro/desktop/en-us/data_dateparse.htm
Calendar Heat Map: https://www.edupristine.com/blog/how-to-create-calendar-heat-maps-in-tableau
Airlines We Lost in 2008: https://crankyflier.com/2008/12/31/airlines-we-lost-in-2008/