Create a Tableau Story

Tableau Public Workbook version 1 :
https://public.tableau.com/profile/manon.philippot#!/vizhome/Udacity-CreateaTableauStory1/2015Flight

Tableau Public Workbook final version :
https://public.tableau.com/profile/manon.philippot#!/vizhome/Udacity-CreateaTableauStoryFinal/2015FlightDelaysandCancellations

Dataset : https://www.kaggle.com/usdot/flight-delays#flights.csv

I. Summary

The U.S. Department of Transportation's (DOT) Bureau of Transportation Statistics tracks the on-time performance of domestic flights operated by large air carriers.

They created a dataset and published a summarizing information from 2015 in the USA on the number of on-time, delayed, canceled, and diverted flights.

This project intends to show some findings about this dataset : the most common origin and destination state or airport and some really cool information about flight delays and cancellations.

The findings that were shown are :

- the origin and destination states with the most amount of flights are California and Texas,
- the airports with the most amount of flights departing and arriving are ATL (Atlanta), ORD (Chicago) and DFW (Dallas-Forth Worth),
- Spirit Air Lines seems to be the worst airline in terms of delays, while Alaska Airlines seems to be pretty good arriving on average in advance compared to the arrival time,
- between September and October flights seem to arrive on average before the arrival time,
- the companies that cancelled the most amount of flights are Southwest Airlines, Atlantic Southeast Airlines and American Eagle Airlines.
- the main reason for flight cancellations is the weather, followed by the airline/carrier, the national air system and the security,
- flights are more likely to be cancelled on monday.

II. Design

Reasons behind chart types :

- map chart : our data contains locations and we study the distribution of variables across the different locations so the map will help the reader to visualize the data more clearly.
- bar chart: we want to understand the distribution of the data and compare the values / frequency for each category. We have more than 10 categories to display so it is

easier to read in the form of a bar graph. The name of the categories can be long sometimes so it is better if it is displayed horizontally.
- pie chart : we want to understand the distribution of the data and compare the values. There is not a big numbers of categories to display so it it can be more straightforward to display it as a pie chart.
- line plot: we want to view a quantitative variable over time in order to analyze trends in the dataset.

Detailed explanation of the visuals :

I decided to show the distribution of the origin and destination states with a map as it would be more easy to comprehend. I displayed the amount of flights for each state with a shade of blue. I also added the possibility to interact with the map (display the destination states on the second map corresponding to the origin state selected on the first map, and vice-versa) so that it is more engaging for the reader.

Concerning the top 20 origin and destination airports, I visualized them with a simple bar chart accompanied by a map. Clicking on a bar center the map on the localisation of the corresponding airport. At first there were colors on the bar but after the feedback I decided to remove them as it was a little bit disturbing for the reader.

Concerning the average delay per airline, I visualized the departure and arrival delay on two separate bar chart one next to the other. I added a color encoding for each airline so the reader could easily compare the position of a particular airline on both charts. According to the feedback, the colors were a little too intense so I made them less vibrant by reducing the opacity.

Concerning the top 20 average delay per airport, I used bar charts that I placed one next to the other. I wanted the reader to easily compare each bar chart and notice some redundancy in the airports shown in the different top 20.

Concerning the average delay by month, I used two line charts (for departure delay and arrival delay) that I put on top of each other for the same reasons as previously.

Concerning the cancellations and reasons, I wanted the page to be as complete and interactive as possible. I first used a bar chart to display the distribution of the cancellations over the airlines and defined it as a filter. I then added a table to display the distribution of the cancellation reasons for the filtered airline(s). I also added a pie chart to display the percentage of cancellations per day, for the filtered airline(s). I finally added a map displaying for the filtered airline(s) the distribution of the cancellations for each state, and for each state the distribution of the cancellation reasons with a pie chart. The goal is that by filtering the bar chart the reader can easily explore the data for more insights.

For every visualization, I tried to give the rights informations in the tooltip so that the reader can simply have all the information he needs to understand the visualization.

After collecting the feedback (cf next section), I corrected my Tableau Visualization :

- I corrected the typos

- I removed the colors on the third and fourth page so that the viewer can concentrate on the relevant part only
- I made the colors that were a little bit too intense a little bit less vibrant
- The font size of the explanatory text was set to size 11 instead of 9
- I displayed the text in a different way so it is easier to read
- I made the texts a little bit shorter when needed in order to keep the attention of the reader on the visualizations
- I added the missing unit on page 6 and renamed the label fields that were possibly not easy to understand

III. Feedback

I shared my Tableau Story in person with a coworker and this is her feedback:

- Some typos in the introduction and conclusion : “"intends", ”some”, “which”.
- The colors on the third and fourth page (Origin and Destination Airports) are not really useful in this situation and can distract the reader.
- The colors on the fifth page (Average Delay Per Airline) and on the pie chart of the eighth page (Cancellations and Reasons) are useful but seem to be a little too intense, maybe make them a little bit less intense.
- The font should be bigger, it would make it easier to understand.
- Sometimes you could make the text shorter (for instance the exact values that are already shown on the visualizations do not need to always be said).
- The text on pages 2 to 7 could be displayed in a different way and not in a big block.
- On the 6th page specify the unit of the y axis
- Rename some fields on some visualizations
- Overall good explanation and we manage to follow the presentation without any problem.

IV. Resources

https://onlinehelp.tableau.com/current/pro/desktop/en-us/buildexamples_pie.htm

https://www.youtube.com/watch?v=v7AsxBaBgyM

https://kb.tableau.com/articles/howto/renaming-dimension-column-row-headers