Build Data Dashboards

Data Visualization Project

2015 Flight Delays and Cancellations

Dataset

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. Summary information on the number of on-time, delayed, canceled, and diverted flights are published in DOT's monthly Air Travel Consumer Report and in this dataset of 2015 flight delays and cancellations.

Dataset link: https://www.kaggle.com/datasets/usdot/flight-delays

Insight #1: What were the number of flights and the number of canceled flights per month in 2015?

[Link to dashboard]

Summary

Here, we want to find out the number of flights in each month of 2015 and how many of these flights were canceled. We also want to find out the reasons for cancellation for the canceled flights. From the first visualization, we find out that July is the month with the most number of flights (26,810 flights) followed by August (26,313 flights). This visualization has a filter for Airlines where we can select an Airline to visualize the number of flights by this airline during each month of 2015. In the second visualization, we find out that the month with the most canceled flights was February where Weather conditions contributed to the most cancellations this month (782 flights). January is the month with the second most canceled flights where Weather was again the reason behind most canceled flights (375 flights).

Design

Vertical bar charts are used to compare different categorical or discrete variables (Months). The stacked bar chart is used to help simultaneously compare the counts of Cancelled Flights and notice sharp changes at the Cancellation Reason level that are likely to have the most contribution to the category totals such as the case with Weather.

Review #1: The line chart or the area chart is the perfect visualization to explore the numerical data over a time series, so it's better to change the bar charts to line or area charts.

Action: Charts were changed to a Line chart (Number of Flights Per Month) and an Area chart (Reasons for Cancelled Flights Per Month).

Resources

N/A

Insight #2: What were the airlines with the most delayed flights in 2015?

[Link to dashboard]

Summary

Here, we want to find out which airlines had the most delayed flights in 2015. From the first visualization (Pie chart), we find out that Southwest Airlines Co. had the highest share of delayed flights in 2015 (19.86% of all delayed flights). The airline with the second most delayed flights was American Airlines Inc. (12.45%), followed by Delta Airlines (10.98%). The second visualization confirms this by showing the number of delayed flights per airline sorted from the highest to lowest. Here we can see that the highest number of delayed flights was for Southwest Airlines Co (608,158 flights), followed by American Airlines Inc. (381,086 flights), then Delta Airlines (350,718 flights). Hovering on each bar shows the number of delayed flights for each airline per cause of delay. For example, the most common cause of delay for Southwest Airlines Co. is “Late Aircraft Delay” with 316,320 flight delays. This visualization can also be filtered by any continuous duration of months.

Design

The pie chart helps organize and show the number of delayed flights of each airline as a percentage of the total number of delayed flights. A horizontal bar chart is a good option for long category names such as Airline names because there is more space on the left-hand side of the chart for axis labels to be placed and horizontally oriented.

Review #1: Change the color palettes in the mentioned visualizations in section #5, try avoiding Red & Green colors in the used palettes is always a good practice to make your visualizations easily interpretable by the colorblind.

Action: The color palette was updated accordingly.

Resources

N/A

Insight #3: What were the states with the most canceled and diverted flights in 2015 based on the destination airport?

[Link to dashboard]

Summary

Here, we want to find out the states with the most canceled and diverted flights in 2015 based on the destination airport's state. The first visualization shows the number of canceled flights based on the state of the destination airport. In this visualization, we can see that Texas is the highest with 668 canceled flights. Illinois and California come second and third with 563 and 408 canceled flights, respectively. The second visualization shows the number of delayed flights based on the state of the destination airport. In this visualization, we can see that Texas is the highest with 130 delayed flights. Illinois and Florida are in a tie for second place with 65 delayed flights each.

Design

Map charts are used to compare values of canceled and delayed flights and show categories across geographical regions (states).

Resources

N/A