2015 Flight Delays

Bay Area Airport Delays
and Cancellations

Evette Trevino
July 31, 20233
WGU – C939
Data Visualization
Mr. Kozakoff

Overview

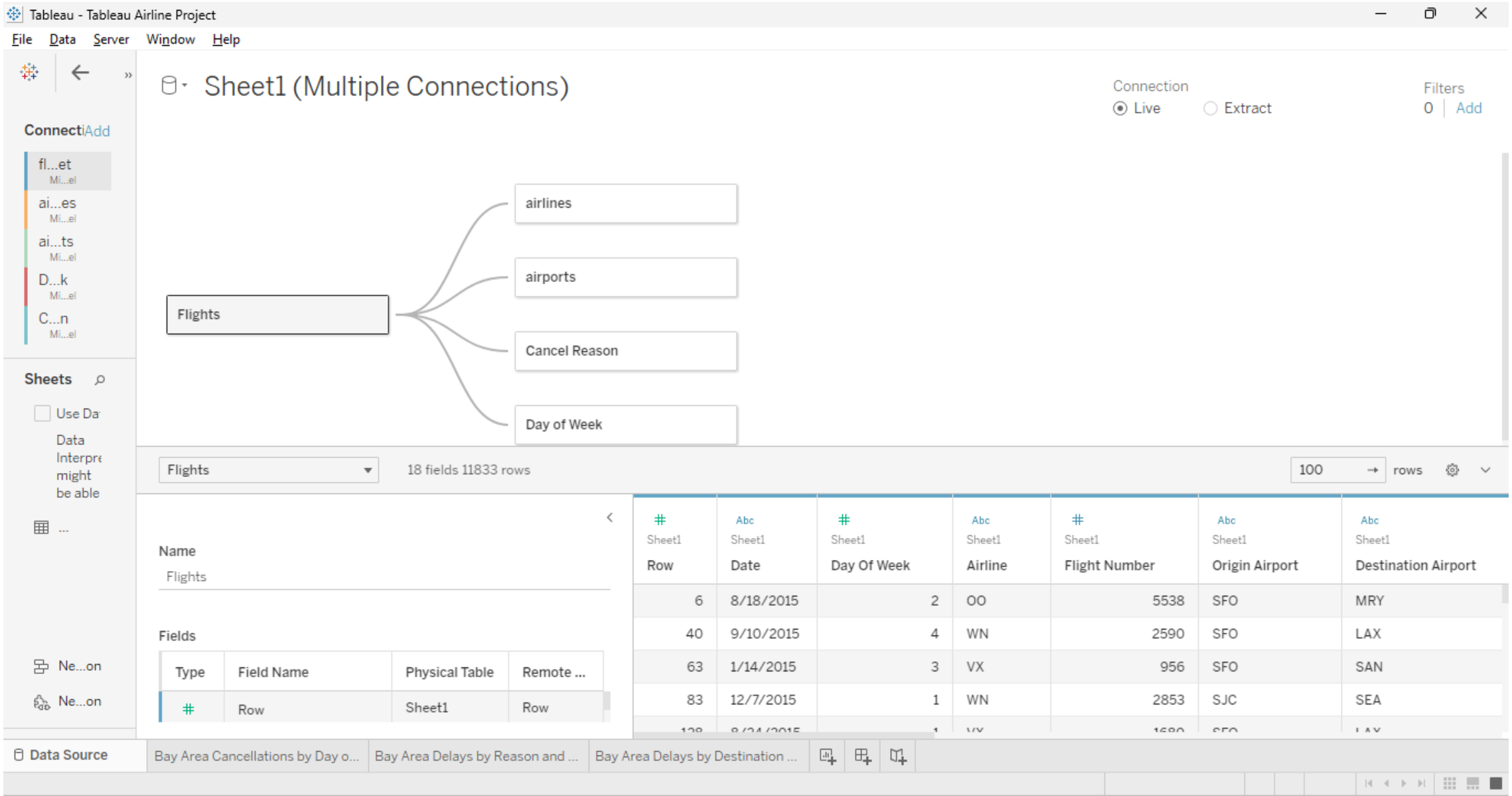
In this project, I'll create visualizations to reveal insights from the *Flights*, *Airports* and *Airlines* data sets using Tableau Desktop version 2023-1. I'll create visualizations to support the findings from the datasets. To examine insights, we'll talk about some of the reasons flight were delayed or cancelled and narrow data to the location of the Bay Area airport origins. We'll also join all three datasets in Tableau and see the use of a filter.

Connections

- Flights was the main data source used along with and connected with
  - Airlines joined on LATA code
    - This provides a 2-character alphabetic code for each airline which flies in the United States.
  - Airports joined on LATA Code
    - This file provides the airport name and their corresponding 3-digit LATA code.
  - Cancel Reason joined on Cancel letter
    - This file contains the numeric code for a cancel reason and the corresponding description for the cancel.
  - Day of the Week joined on Week number
    - This file contains the numeric code for the day of the week along with the corresponding day.

Connections

- Datasets used and how they are connected.



Bay Area Cancellations by Day of Week

First Insight

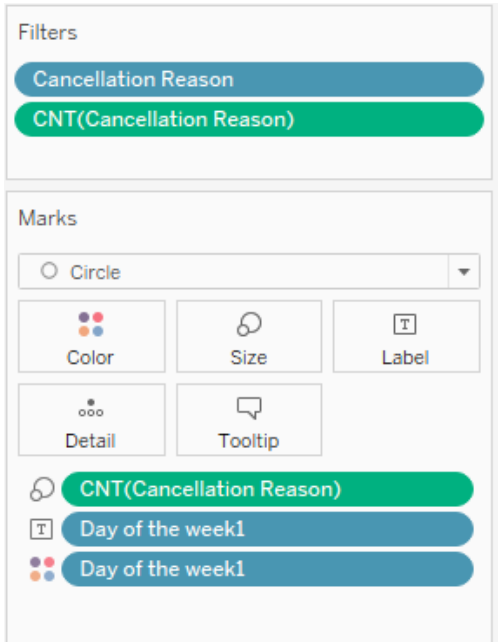
- Summary:
  This circle plot on the next slide displays the total cancellations in 2015 for the three major airports in the Bay Area.
    1. OAK – Oakland International
    2. SFO – San Francisco International
    3. SJC – Mineta San Jose International

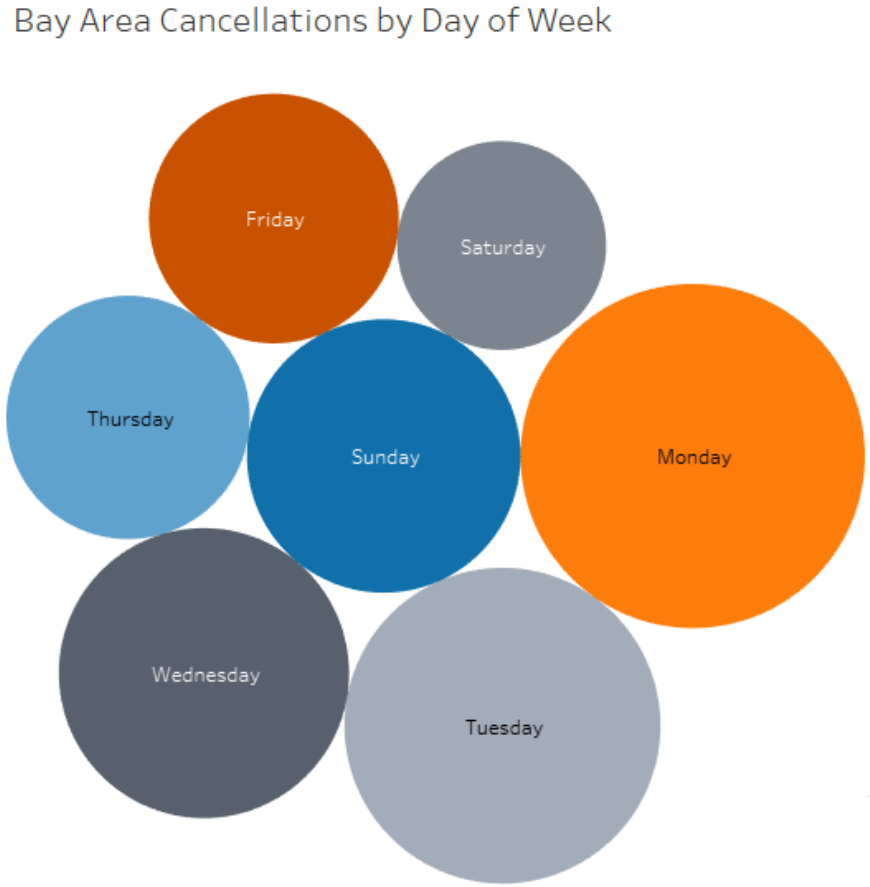
  Monday has the largest number of cancels than any other day of the week. It is the largest circle in the dashboard. The insight we can conclude is that Monday has the largest cancellations as it has the largest in shape. The line graph shows Monday has 38 cancellations, above any other day.

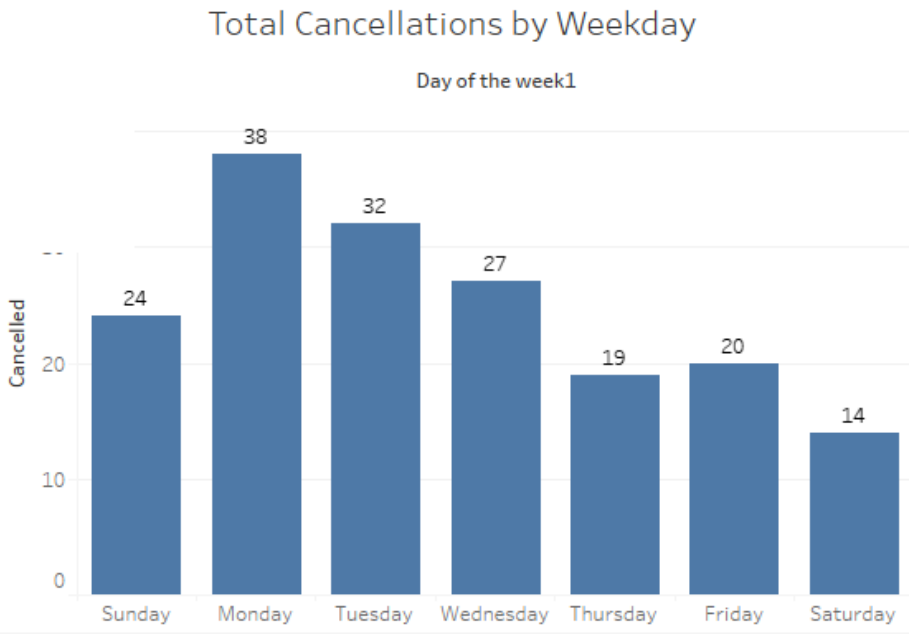
- Design Content: I selected a bubble chart as it was easy to view the data. Monday is the largest circle in size. I selected a bar graph to accompany the findings and help specify total cancellation by day of week. I also use the colorblind palette instead of the standard and brighter colors.

- Resources: N/A

Bay Area Cancellations by Day of Week







Bay Area Delays by Reason and Local Airport

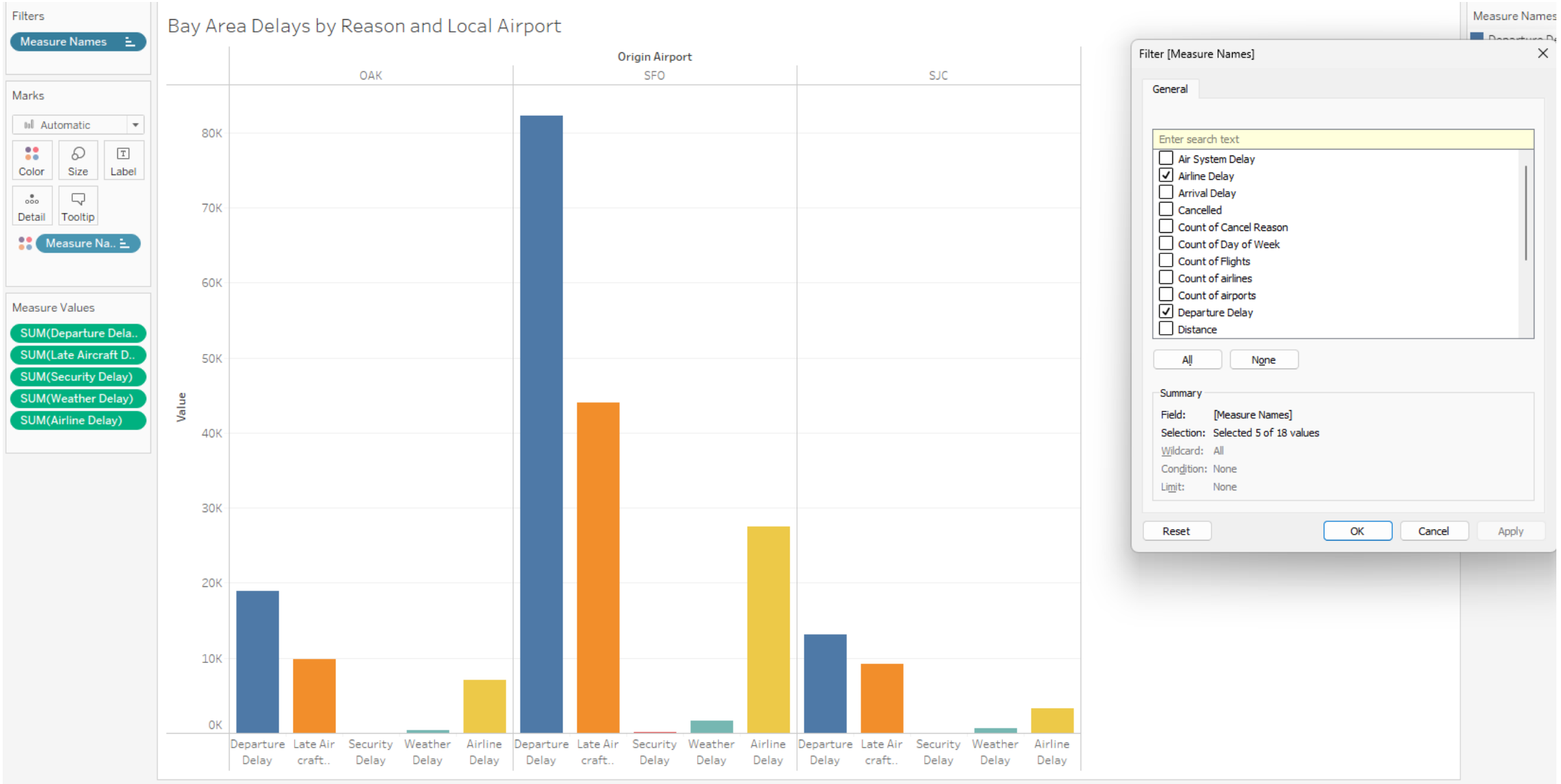
Second Insight

- Summary:
  This next slide is a bar graph and displays delay reasons for Oakland International Airport (OAK), San Francisco International Airport (SFO), and Mineta San Jose International Airport (SJC). San Francisco International has the highest departure delays in most categories. A filter was used to add in Departure, Late Aircraft, Security, Weather, and Airline delay reasons. For accuracy and using the help of the tooltip, SFO has 82,314 delays in total minutes, due to departure delays. SFO also leads in late aircraft an airline delay reasons. The insight we can conclude from this visualization is SFO has the highest chance for delay compared to two nearby airports.

- Design Content:
  I chose to use a bar graph for this visualization to clearly show the difference between all three airports.

- Resources:
  - N/A

Bay Area Delays by Reason and Local Airport



Bay Area Highest Delays by Destination

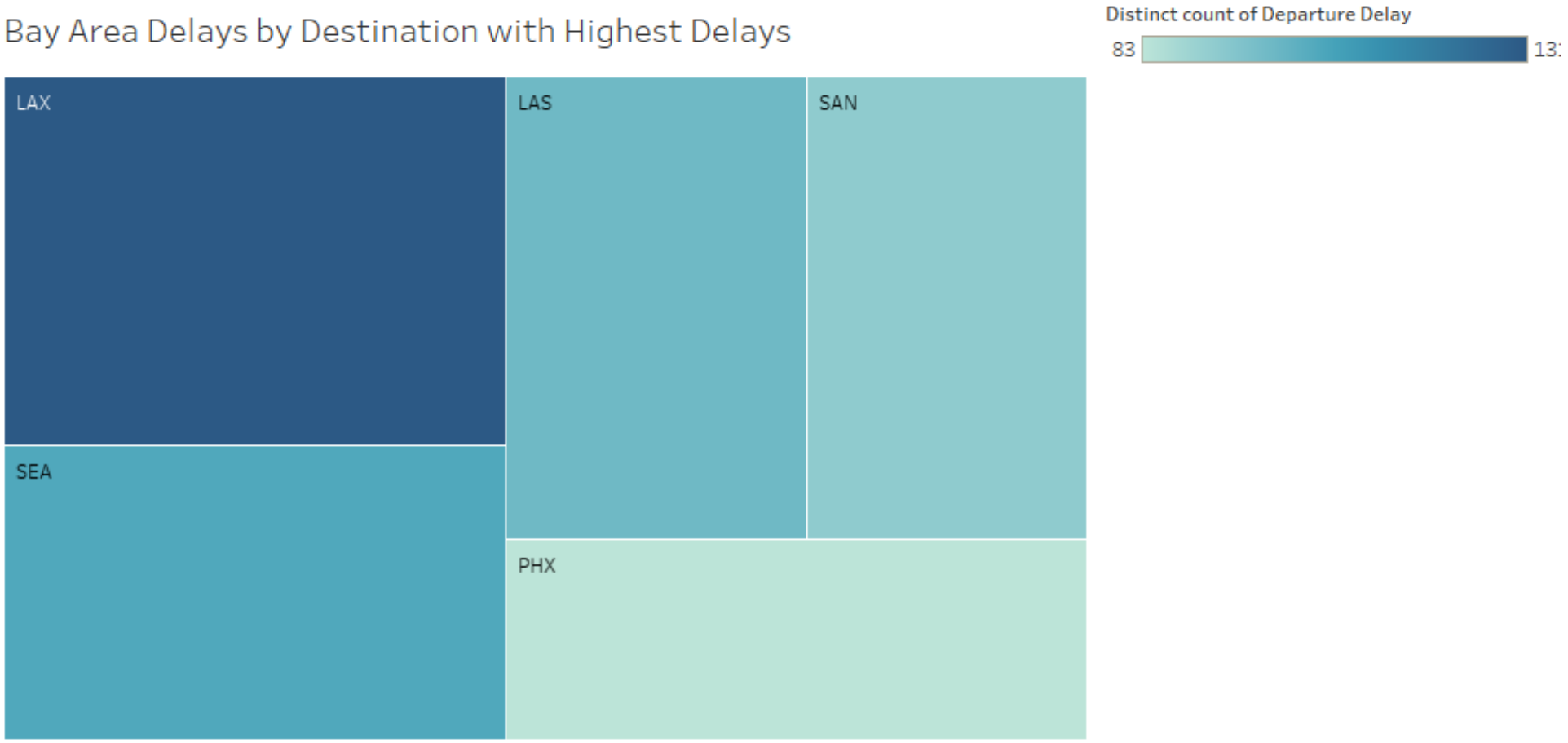
Third Insight

- Summary:
  The tree map below shows the destinations with this highest delays where flights originated from the Bay Area airports. LAX had the highest number of delays with 83. We are able to conclude this from looking at the tree map as the density of the color for LAX is the darkest in color. Related to this data is the bar graph on the next slide. The bar graph shows the airport origination and highest delays in minutes beginning at San Francisco International airport.

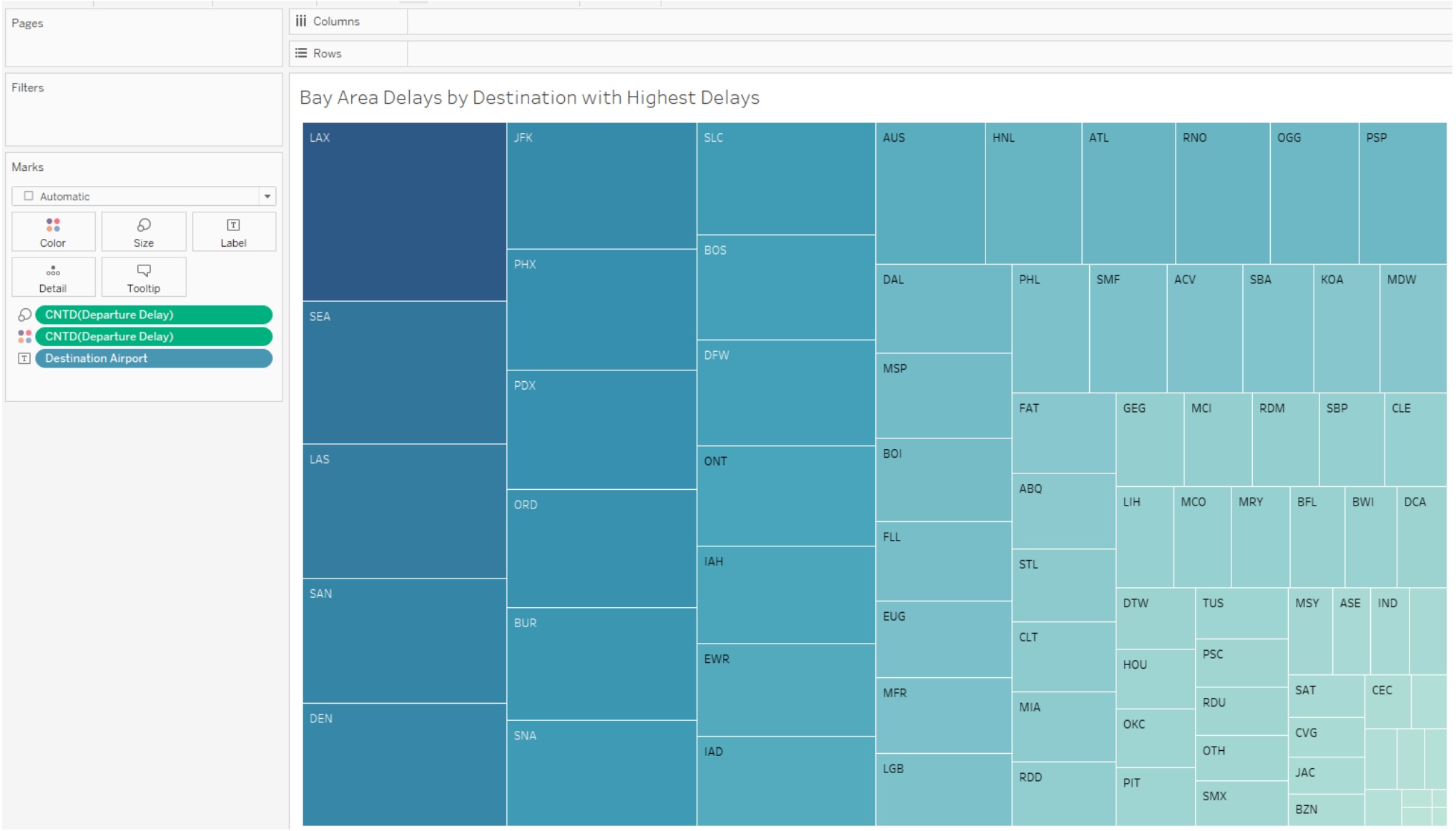
- Design Content:
  I chose a tree map as this was a more modern and simple way to show the number of delays by destination airport. I chose a bar graph to emphasize the origination location.

- Resources:
  N/A

Bay Area Highest Delays by Destination



Bay Area Highest Delays by Destination Without Filter



This is the same visualization as the previous slide. The difference is the location (CNT – Departure Delay) was removed. Now what is represented are all airports across the country, rather than only 3 major airports in the Bay Area.