

YEAR

2023

PRESENTERS

Dania Abdelmonem 2112241

Salma Mohammed 2110956

Lamar Omar 2115727

Air Quality In USA

Data Visualization Course Project

CLASS

CCDS-311



THE DATASET

Introduction

The dataset is made up of air quality data that was collected in 2020, across the United States by outdoor sensors.

The dataset offers a clear and in-depth assessment of the regional variations in ozone concentrations in the United States.

Since every entry in the dataset includes data regarding air quality, it is an valuable tool for studying and examining ozone level trends as well as environmental conditions.

NEXT

Our Questions

01

Question:

Which monitoring state's average AQI results are the highest?

02

Question:

What is the amount of outlier observations?

03

Question:

What connection exists between ozone and air quality?

04

Question:

Analyzing what is the difference in ozone levels among cities?

NEXT

Fix data types, convert to measure or dimension, and/or create hierarchies. Perform at least two such tasks.

Task 1

CONVERT TO MEASURE

Convert Parameter Code and Site Num
to measure

- # Observation Count
- # Observation Percent
- # Parameter Code
- # POC
- # Site Num
- # *daily_ozone_2020.csv (C...*
- 🌐 Latitude (generated)
- 🌐 Longitude (generated)
- # Measure Values

- ▼ ⚙ Country
 - # County Code
 - 🌐 County Name
- 📅 Date Local
- 📅 Date of Last Change
- Abc Datum
- Abc Event Type
- Abc Local Site Name
- ▼ ⚙ Location
 - Abc Address
 - 🌐 State Name
 - # State Code

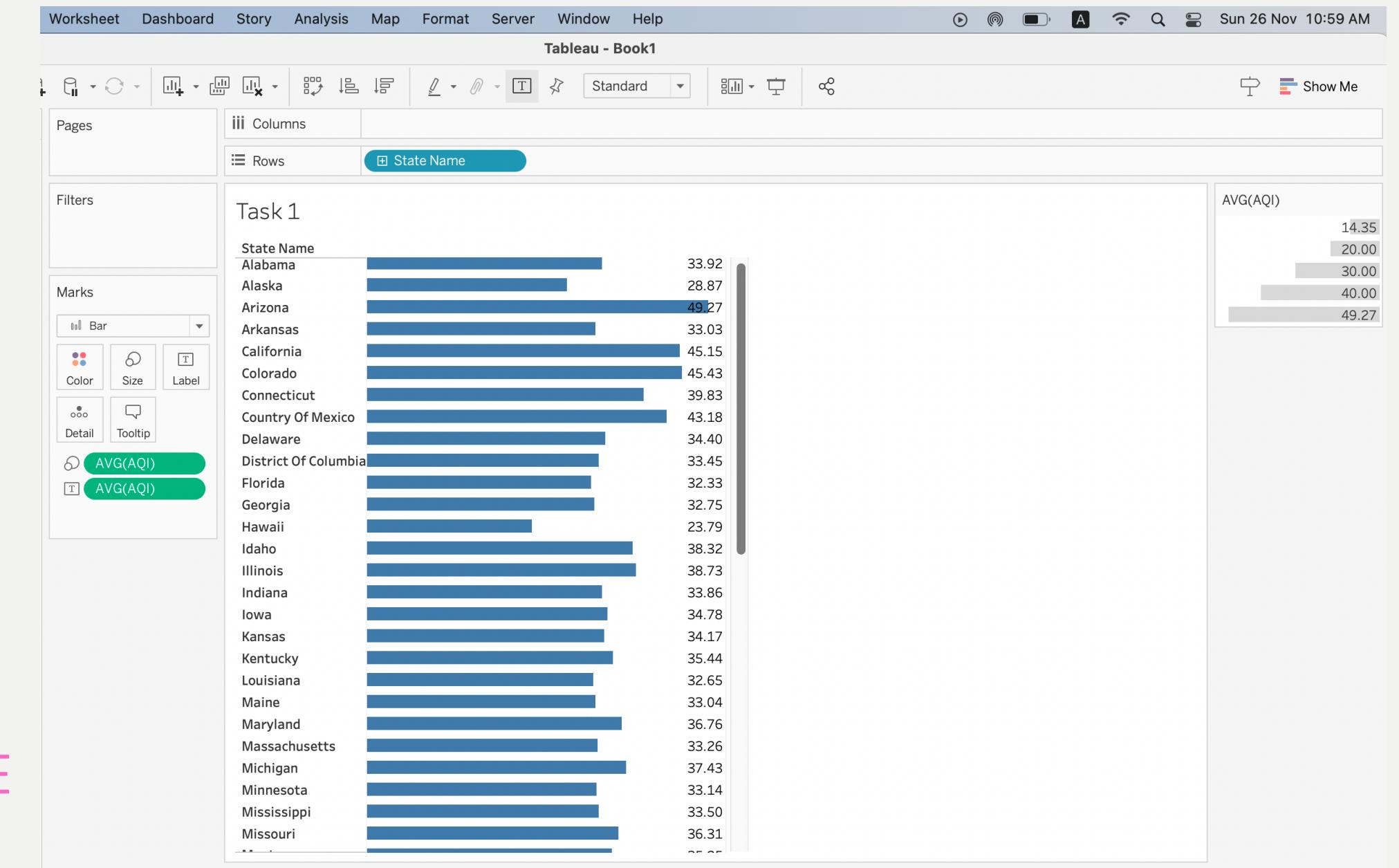
CREATE HIERARCHY

Hierarchy of Country w/ Country Name
& Country Code

Hierarchy of Location w/ State Code,
State Name, and Address

NEXT

Visualize single variables (at least one single categorical and one single continuous) relevant to your questions



Task 2

Part 1

WHICH MONITORING STATE'S AVERAGE AQI RESULTS ARE THE HIGHEST?

State Name (categorical)

We detected that the highest AQI value belonged to Arizona

NEXT

Visualize single variables (at least one single categorical and one single continuous) relevant to your questions

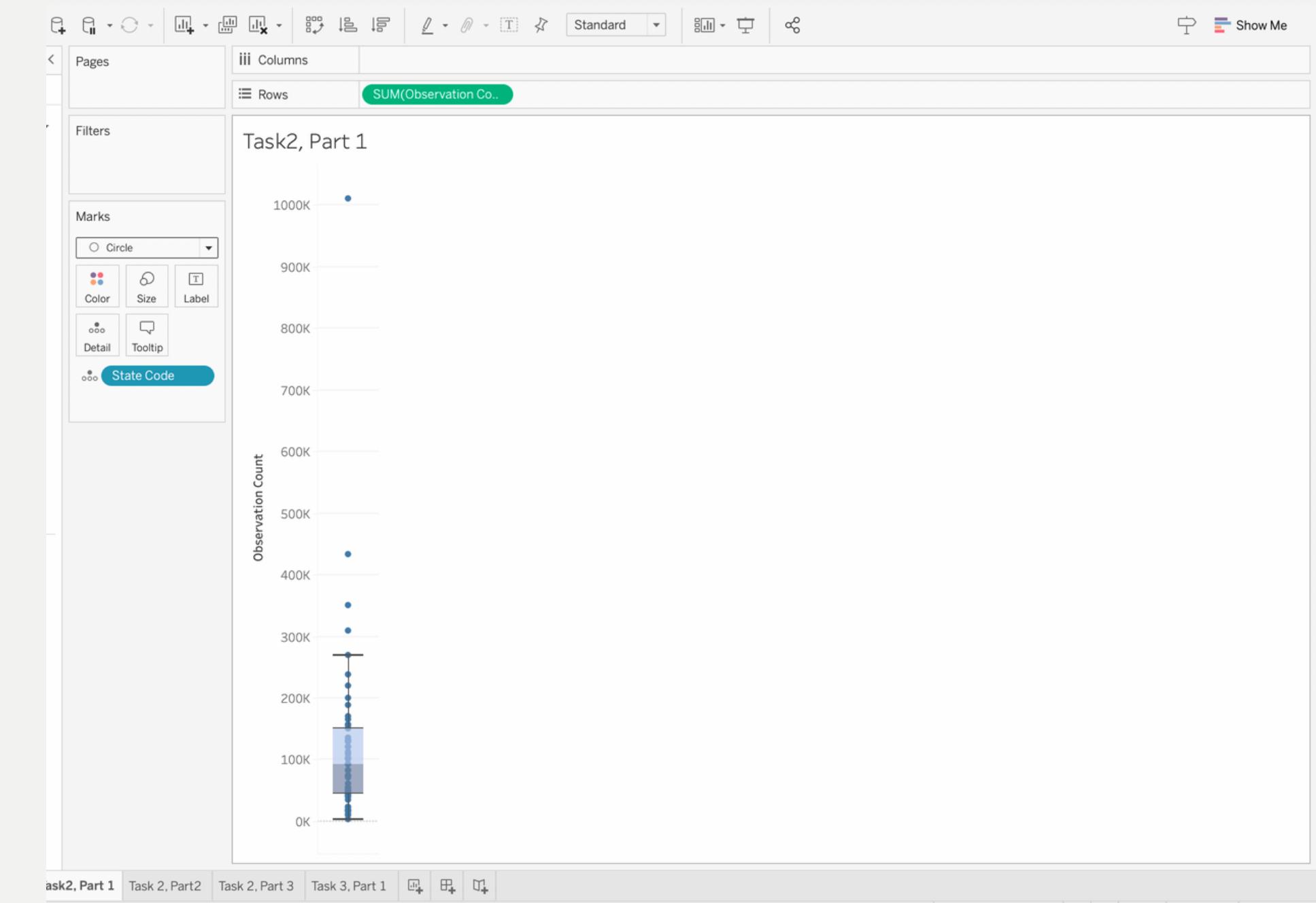
Task 2

Part 1 continued

WHAT IS THE AMOUNT OF OUTLIER OBSERVATIONS?

Observation Percent (continuous)

There appear to be four outliers in the Sum of Observations for the four states, which could indicate a mistake entry in the data.



NEXT

Visualize two continuous variables versus each other (at least one visual required) relevant to your questions

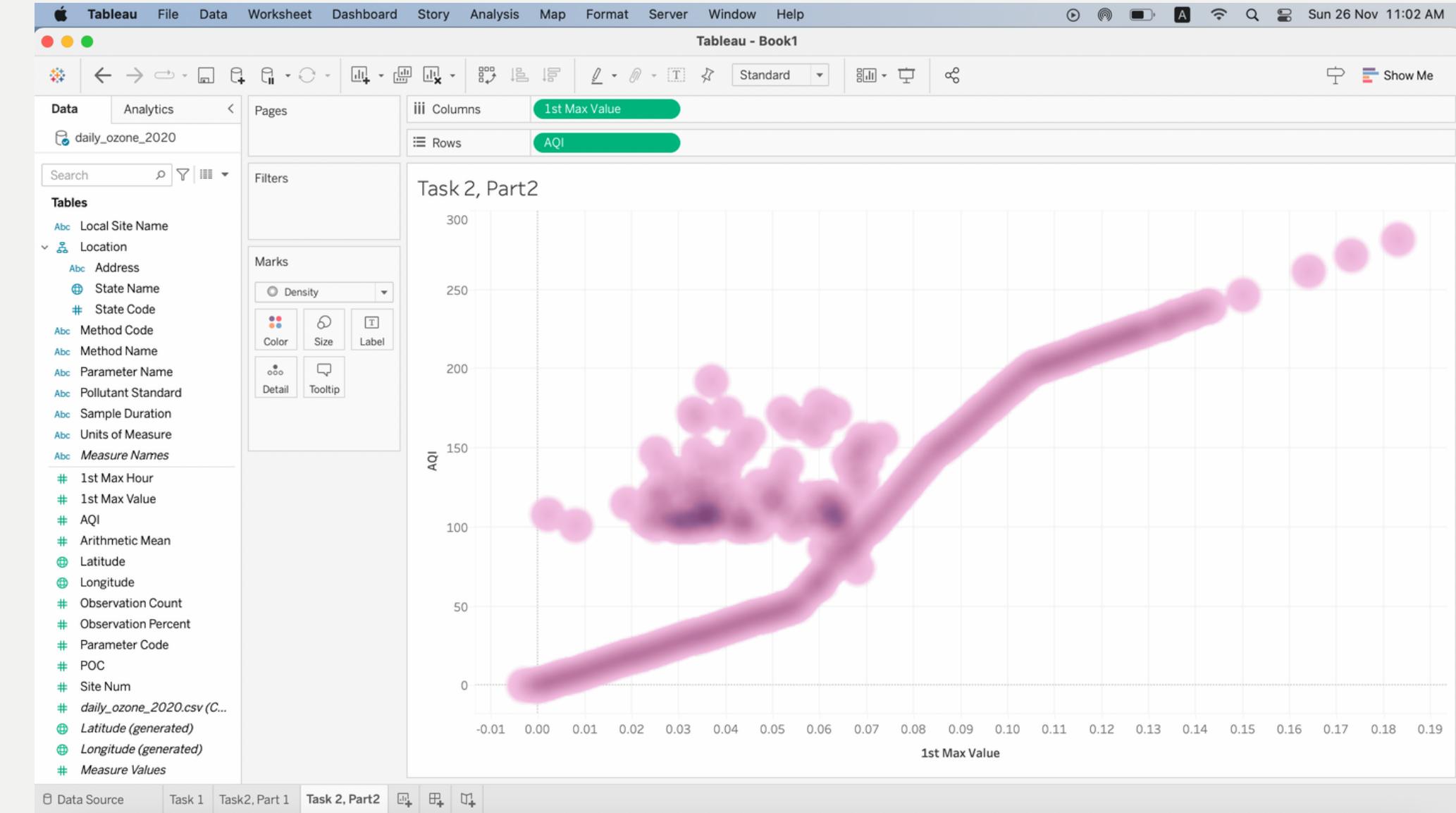
Task 2

Part 2

WHAT CONNECTION EXISTS BETWEEN OZONE AND AIR QUALITY?

1st Max Value (continuous) & AQI (continuous)

We can see that when the max value of ozone increases, the AQI (Air Quality Index) is also likely to go up with it.



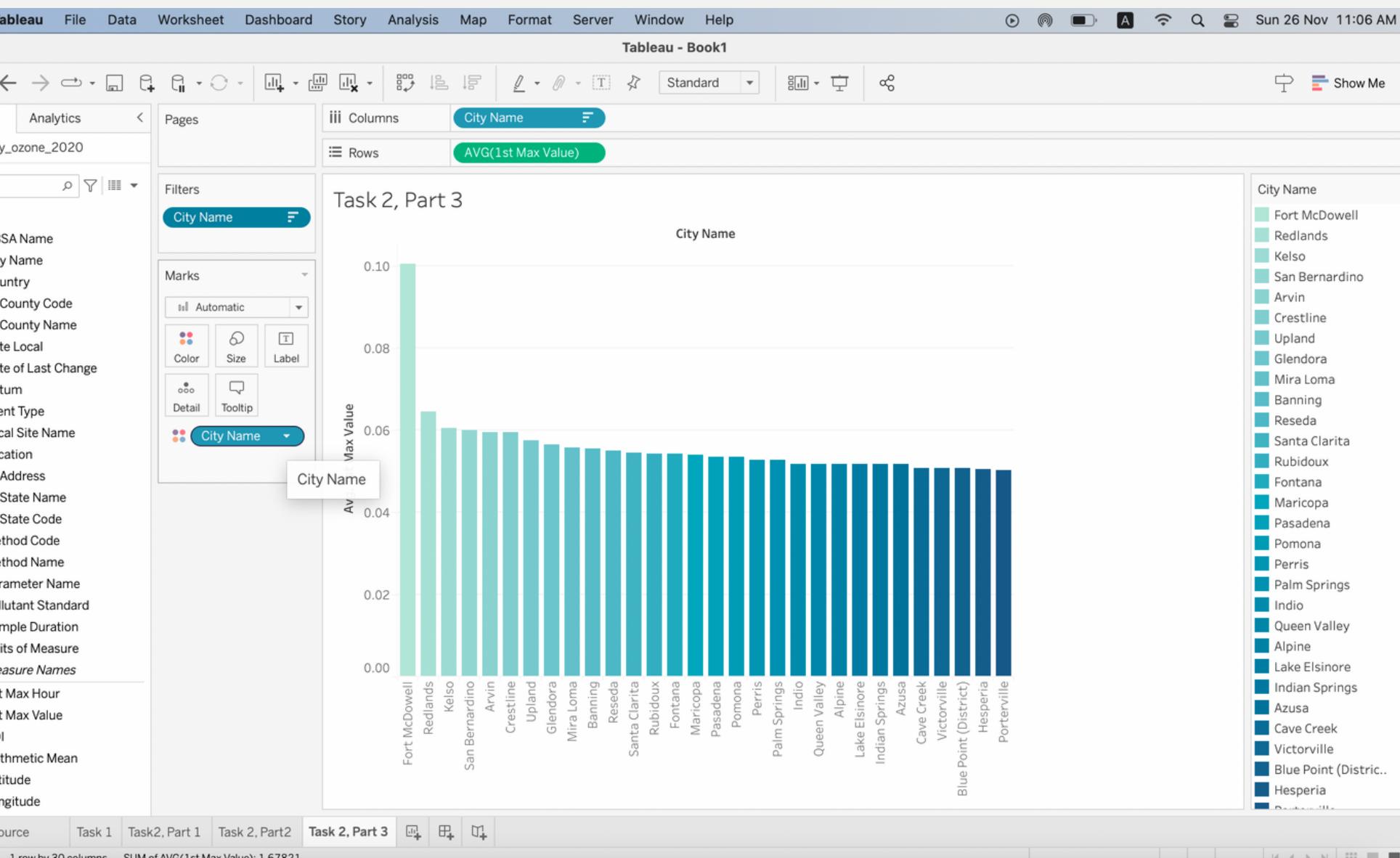
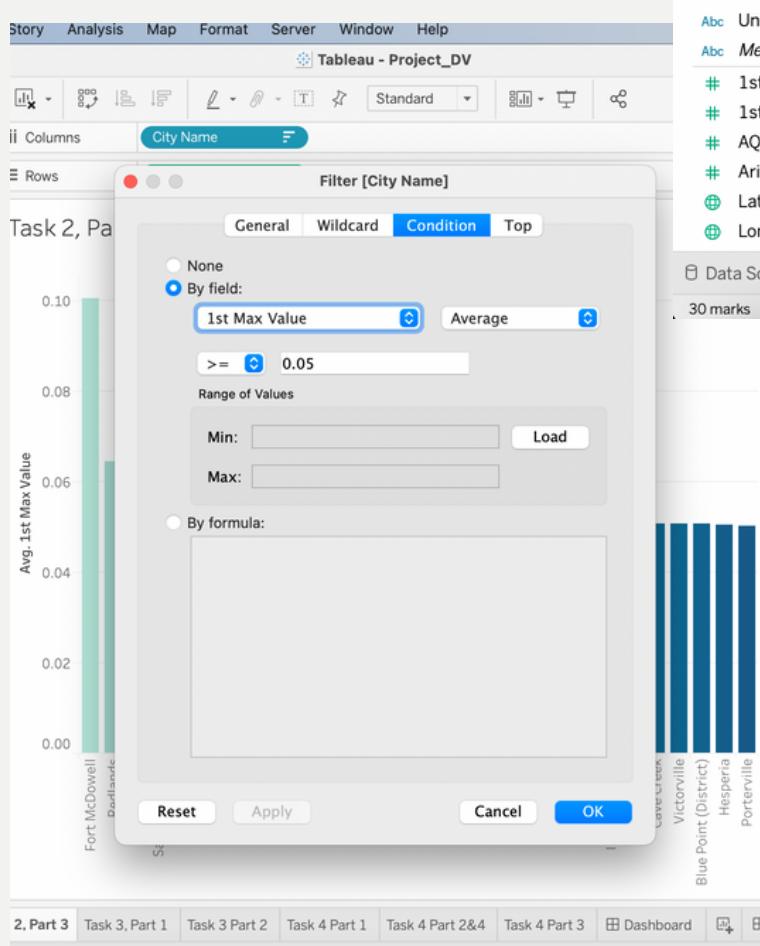
NEXT

Visualize two variables of different types versus each other (at least one visual required) relevant to your questions.

Use at least one filter in your visualizations

Task 2

Part 3



ANALYZING, WHAT IS THE DIFFERENCE IN OZONE LEVELS AMONG CITIES?

City Name (categorical) & 1st Max Value (continuous)

We realized that between cities, there are different ozone levels, then we added a condition filter(task 4 part 4) to retrieve cities that have more than or equal to an 0.05 ozone level.

NEXT

Use at least 2 different geospatial visuals to answer your questions, include at least three different types of maps

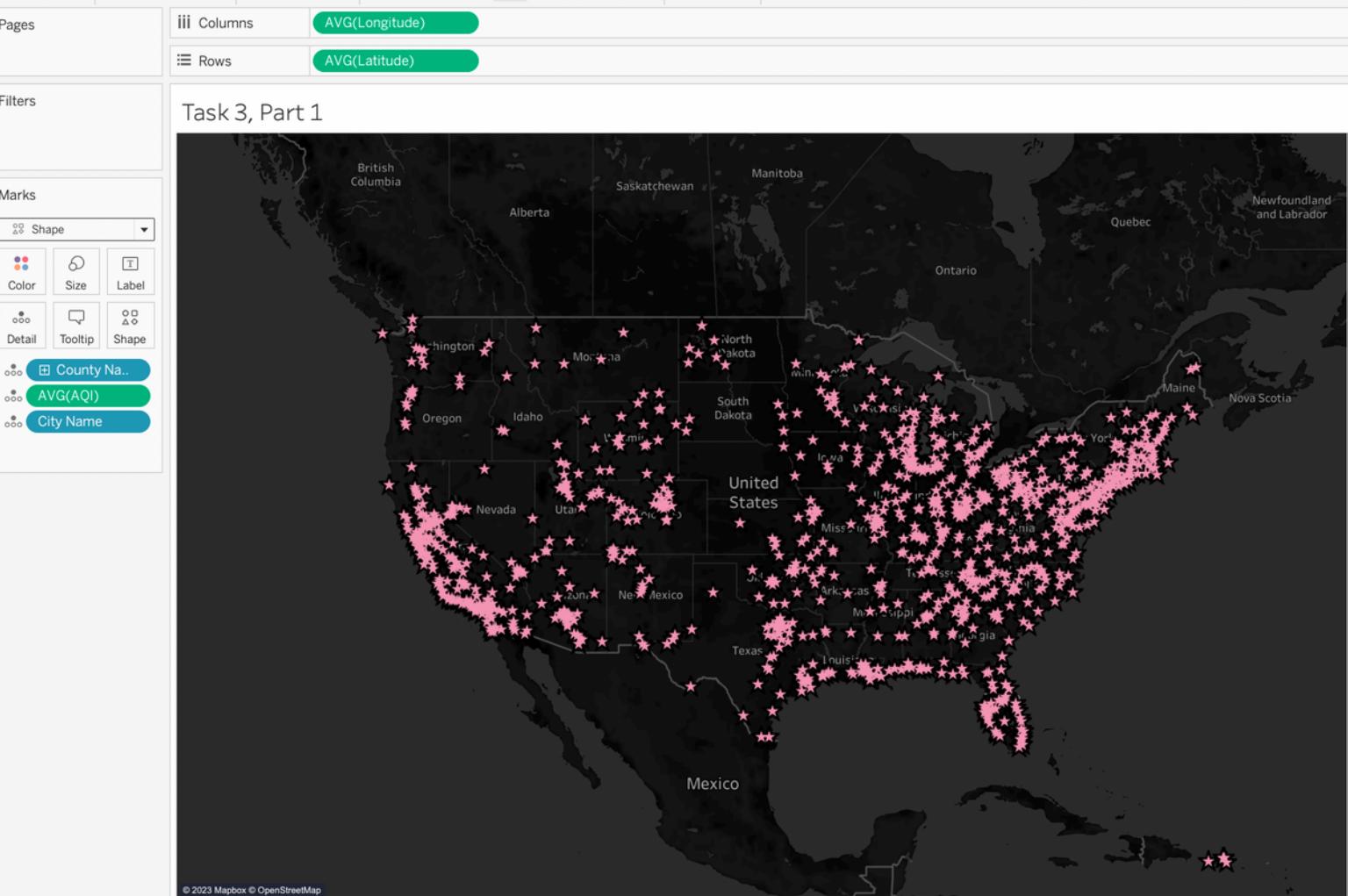
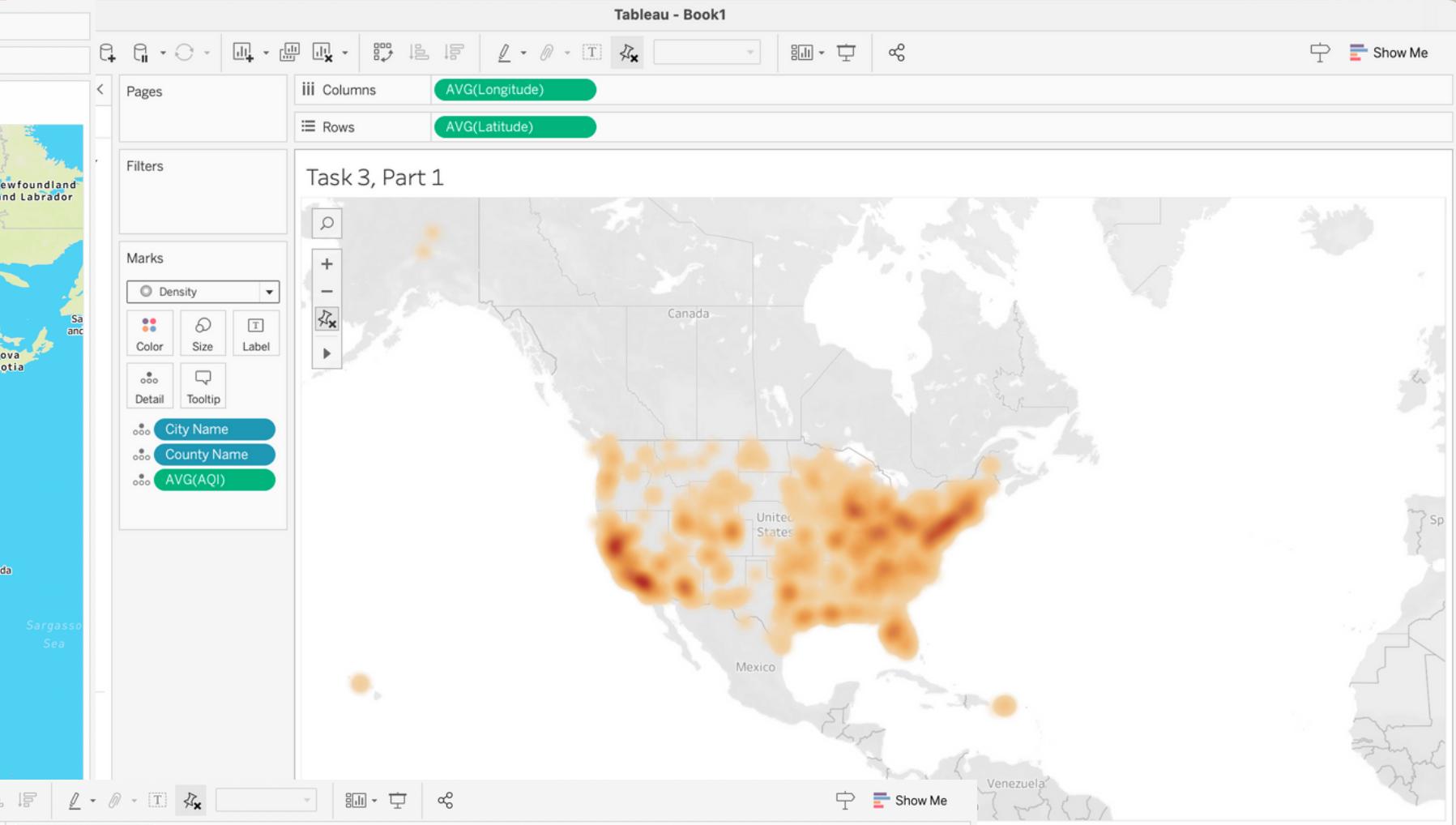
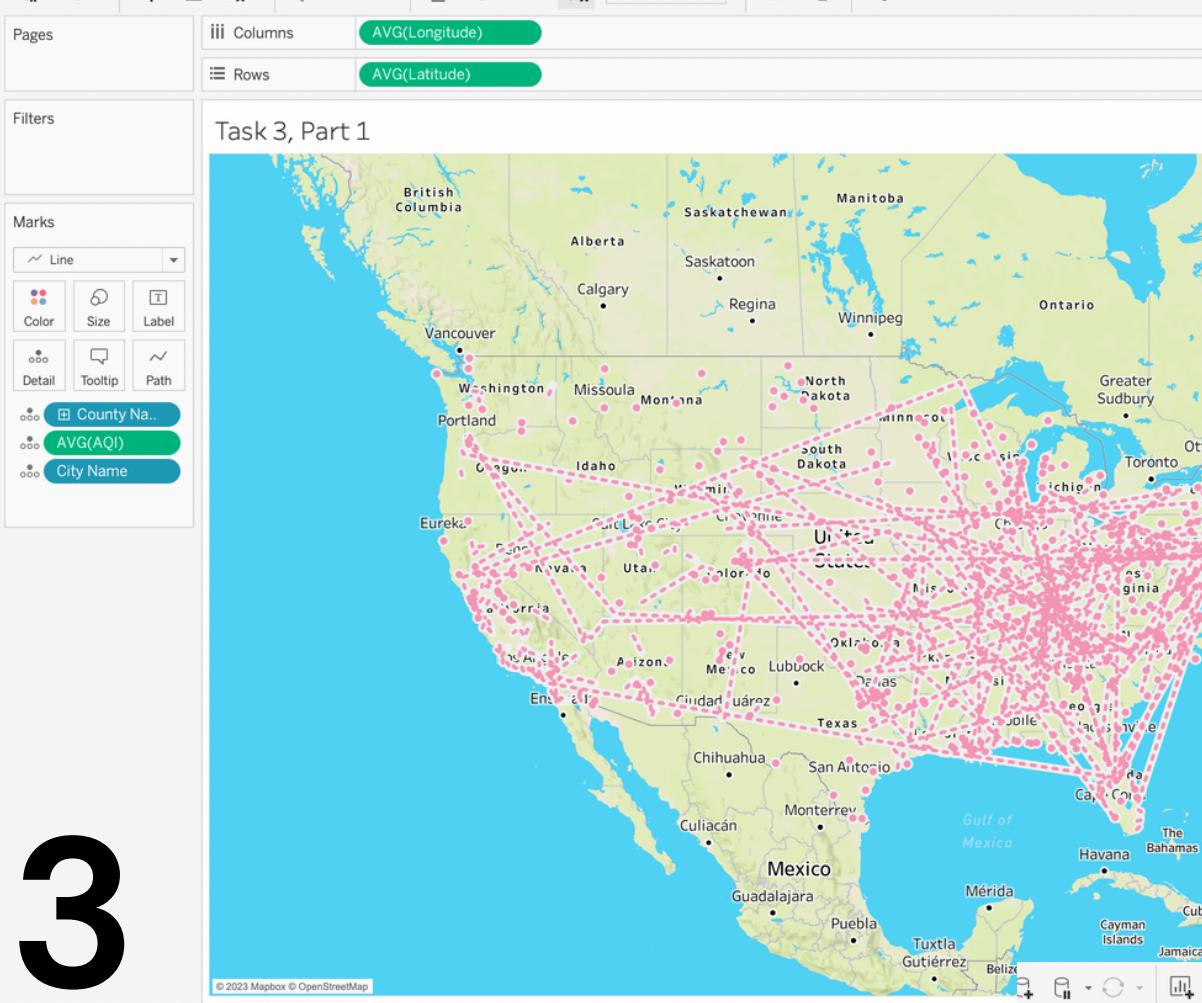
Task 3

Part 1

ANALYZING THE AVERAGE AQI BY COUNTRY AND CITY NAME

Using Filled map, Density map, and Line map.

Using these maps helped us visualize and see that the East had a higher AQI.



NEXT

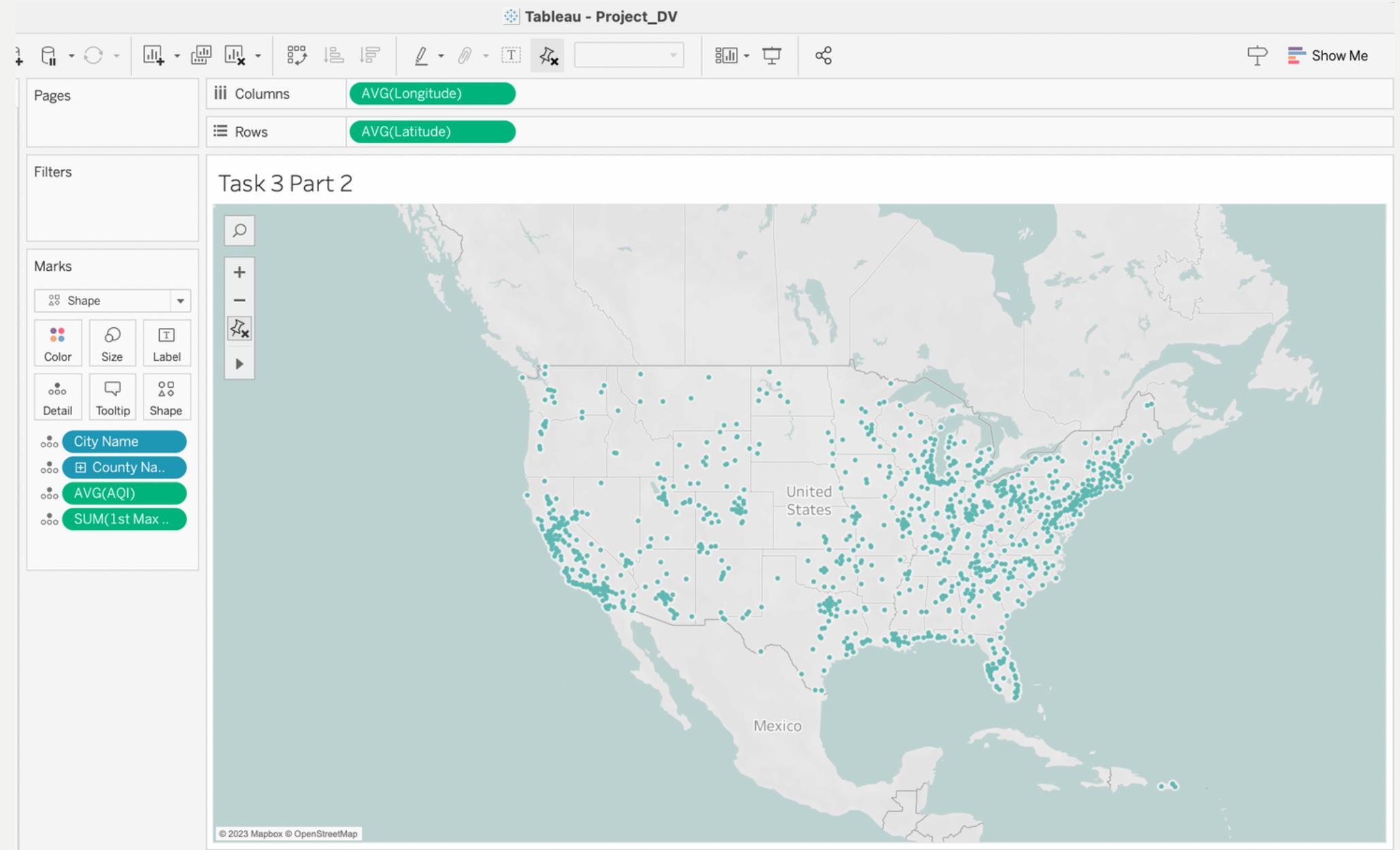
Four visual attributes should be used in these visualizations

Task 3

Part 2

USING 4 VISUAL ATTRIBUTES

1. City Name
2. Country Name
3. AQI
4. 1st Max Hour



NEXT

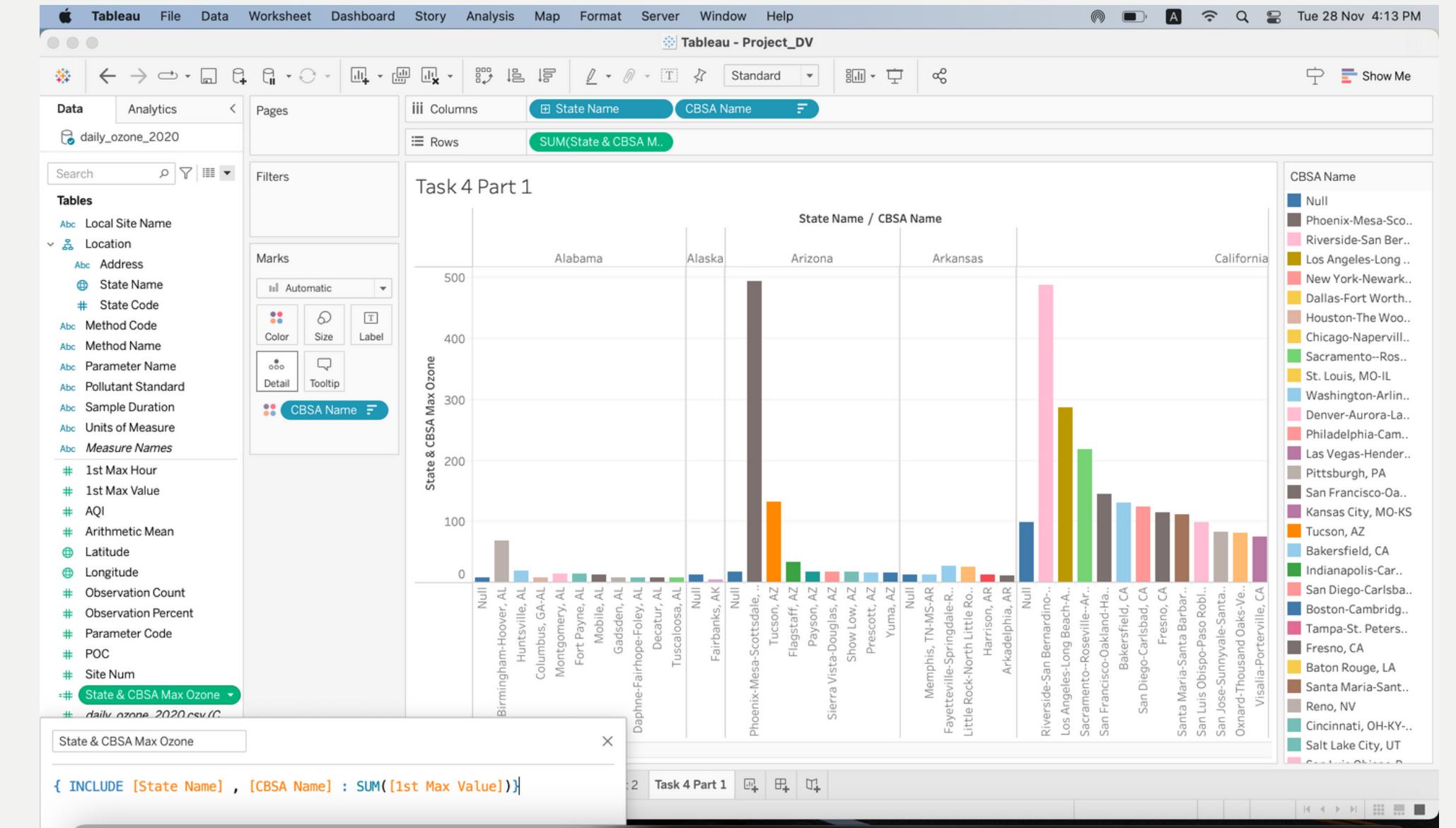
Use at least one LOD expression

Task 4

Part 1

LOD EXPRESSION

We used INCLUDE LOD Expression for calculating the sum of Max Ozone Value in each State and CBSA.



NEXT

Use at least one calculated field

Task 4

Part 2

FLAGGING EACH DAY THATS
CONSIDERED A “HIGH POLLUTION
DAY” BASED ON THE THRESHOLD AQI VALUE

This showed us most days had normal pollution.

Calculated fields allow you to create new data from data that already exists in your data source.

[NEXT](#)

The screenshot shows a Tableau worksheet titled "Task 4 Part 2". In the top right corner, there is a modal dialog box for a calculated field named "High Pollution Day". The dialog contains the following code:

```
IF [AQI] > 150 THEN 'High Pollution' ELSE 'Normal' END
```

Below the code, a message says "The calculation is valid." There are "Apply" and "OK" buttons at the bottom right of the dialog. The main workspace shows a table with columns "High Pollut.." and "State Name", and rows for various US states with their corresponding AQI values. The calculated field "High Pollution Day" is applied to the rows.

High Pollut..	State Name	AQI
High Pollution	Arizona	188.7
	California	174.2
	Colorado	156.0
	Connecticut	163.4
	Illinois	159.1
	Indiana	152.5
	Kansas	169.0
	Michigan	161.5
	New Mexico	156.0
	Puerto Rico	154.0
	Rhode Island	154.0
	Texas	163.6
	Utah	154.0
	Wisconsin	167.1
	Normal	Alabama
Alaska		28.9
Arizona		47.6
Arkansas		33.0
California		43.5
Colorado		45.3
Connecticut		39.6
Country Of Mexico		43.2
Delaware		34.4
District Of Columbia		33.4
Florida		32.3
Georgia		32.8
Hawaii		23.8

Use at least one table calculation

Task 4

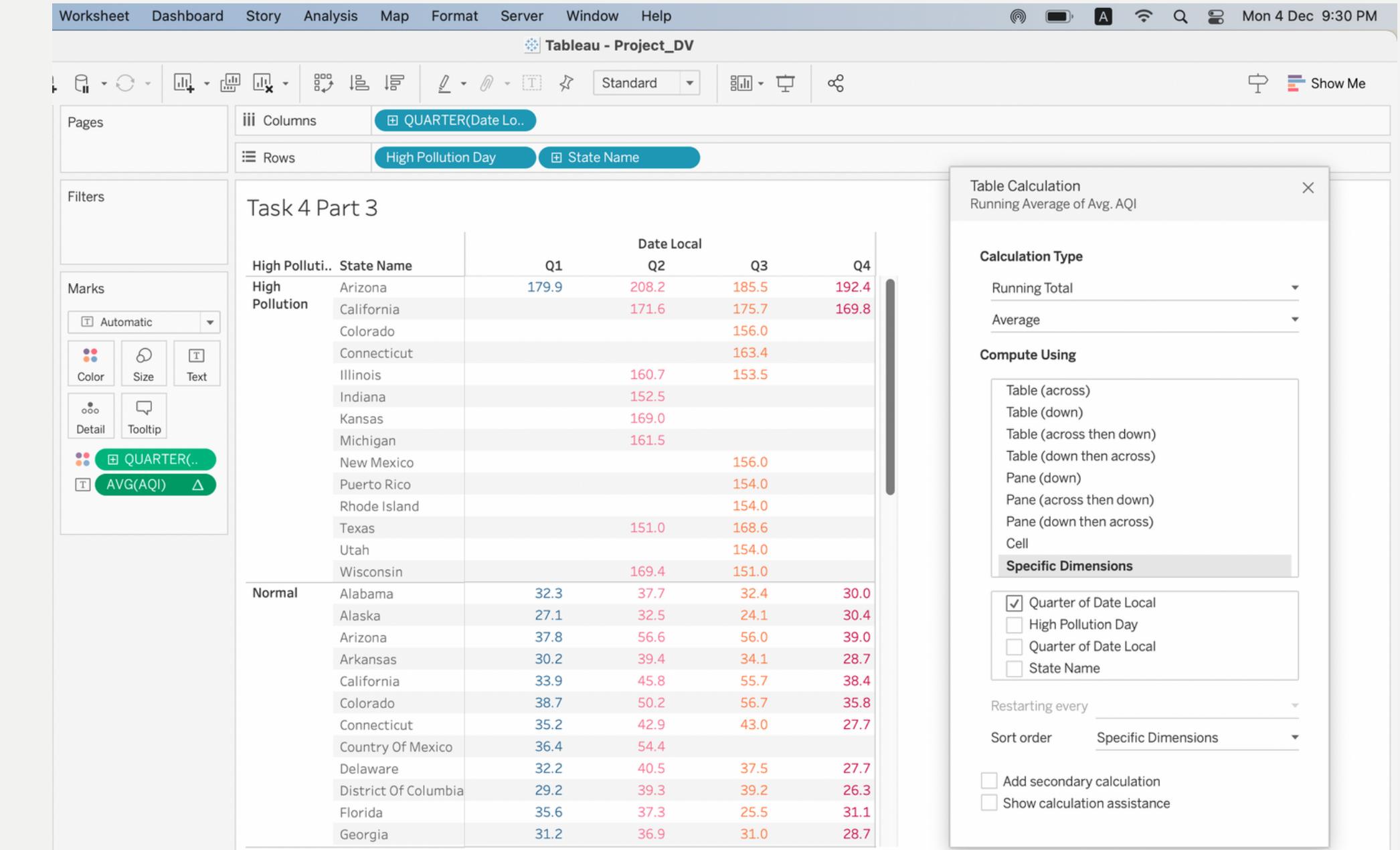
Part 3

CALCULATING THE AVERAGE OF
THE 4 QUARTERS IN THE YEAR

Using Table Calculation

filtered out of the visualization. You can use table calculations for a variety of purposes, including: Transforming values to rankings.

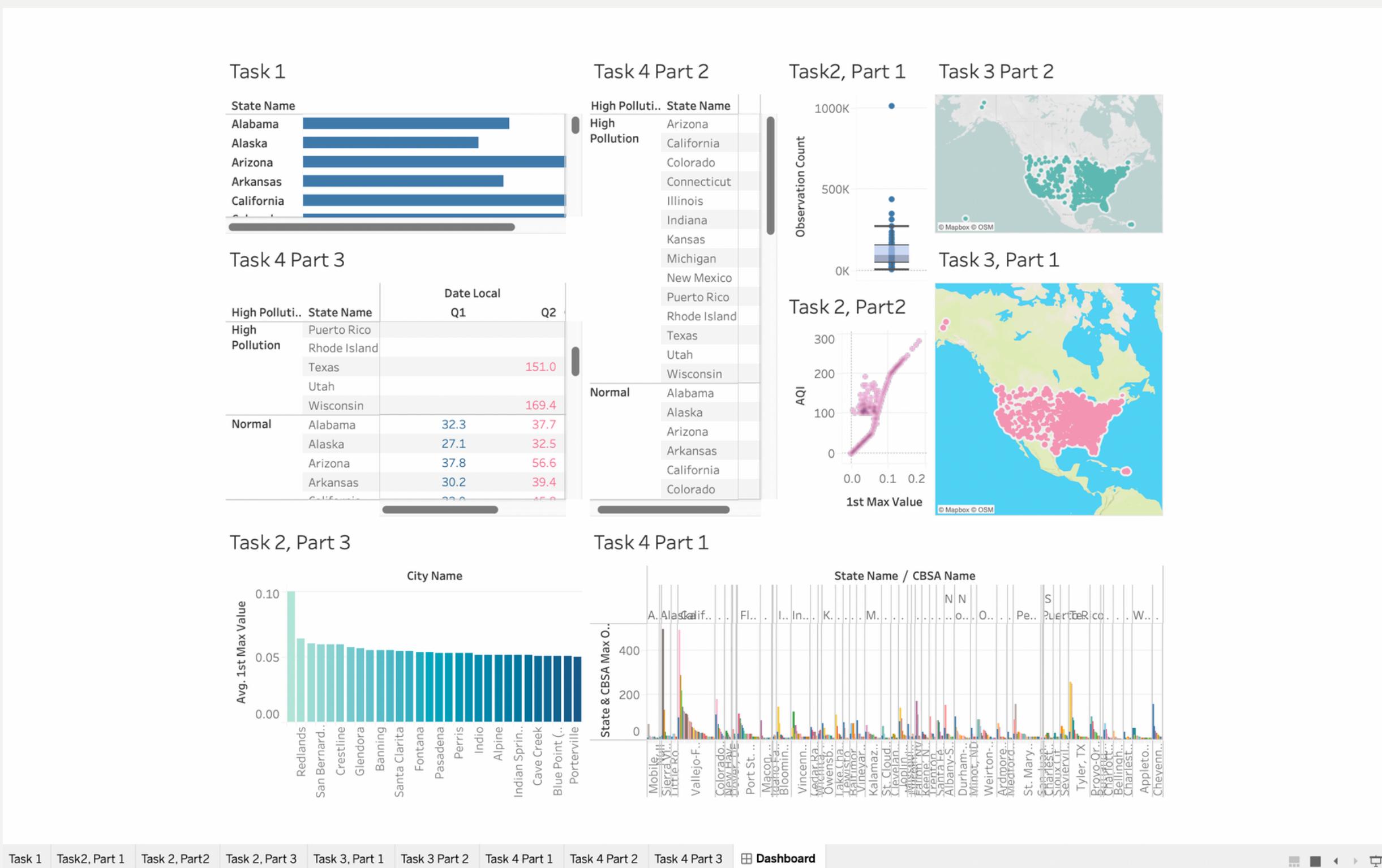
[NEXT](#)



Task 4

Part 5

DASHBOARD



NEXT

Thank You. ☺