

# DAILY OZONER ANALYSIS

---

Group members:

- |                   |         |
|-------------------|---------|
| • Amira Saleh     | 2006782 |
| • Enas Khan       | 2007143 |
| • Marwah Alhadi   | 2007881 |
| • Zain Abdulwahab | 1911102 |

# INTRODUCTION

## ABOUT THE DATA

- The data comes from EPA's Air Quality System (AQS). All the collected reports are aggregated and summarized by the EPA system. All the data were compressed into a single Comma-Separated Value.
- the data includes the pollution standards and the air quality standards.
- the numbers in the dataset we have were calculated are An 8-hour average that is calculated for every clock hour
- The dataset Included 52 States, 517 counties, and 674 cities in the USA.
- The data were calculated from 2020 to 2021.

# Variables & Hierarchy

*Select the Variables for Visualization:*

---

- 1st Max Value
- AQI
- Arithmetic mean
- latitude
- longitude
- State Name
- Observation Count
- Event type

*Hierarchy :*

---

- Location
  - State Name
  - County Name
  - City Name

# Questions

**Q1 | WHICH STATE HAD THE HIGHEST '1ST MAX VAULE' ?**

**Q2 | ANALYZE THE AVERAGE AQI BASED ON STATE NAME**

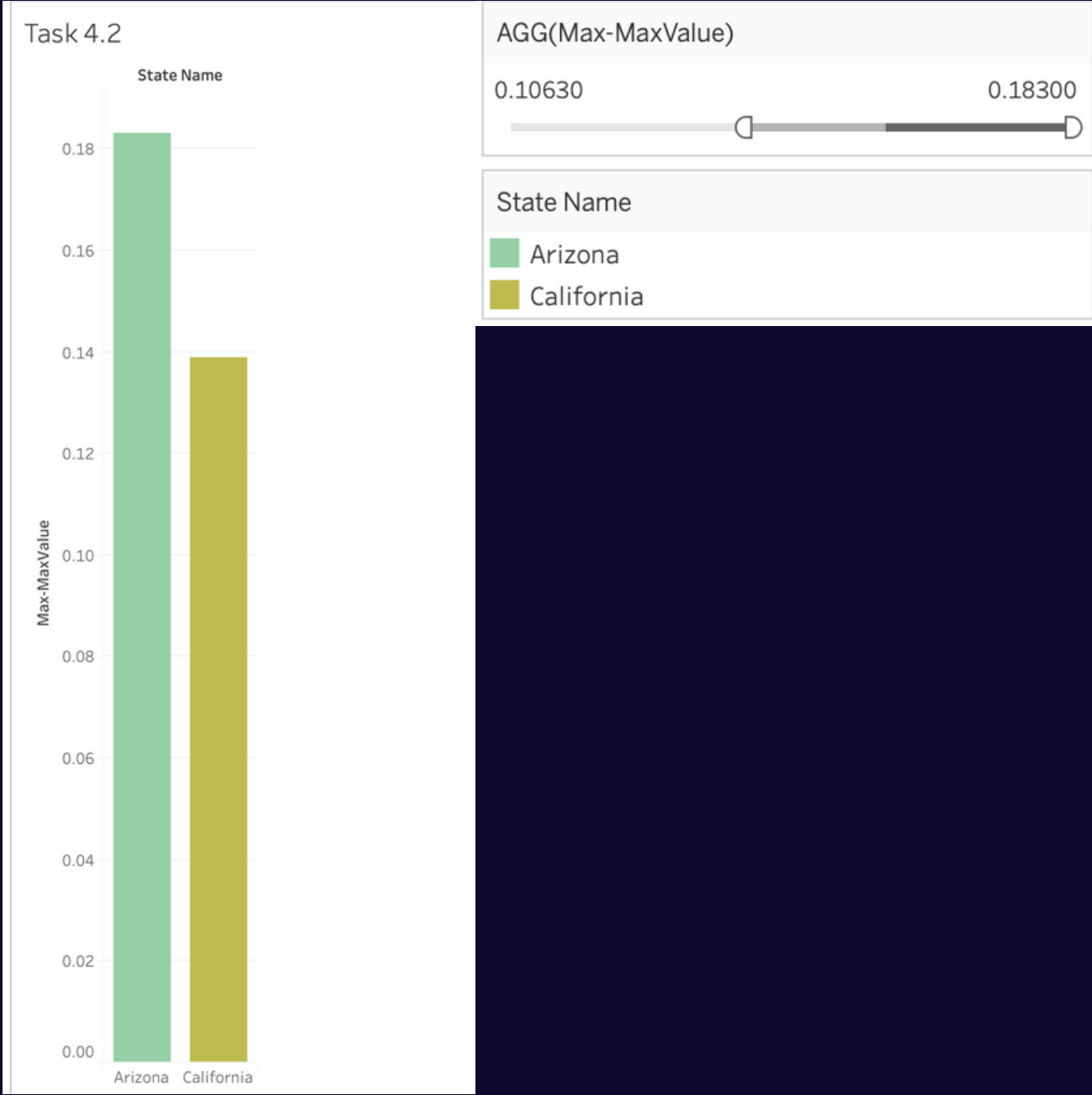
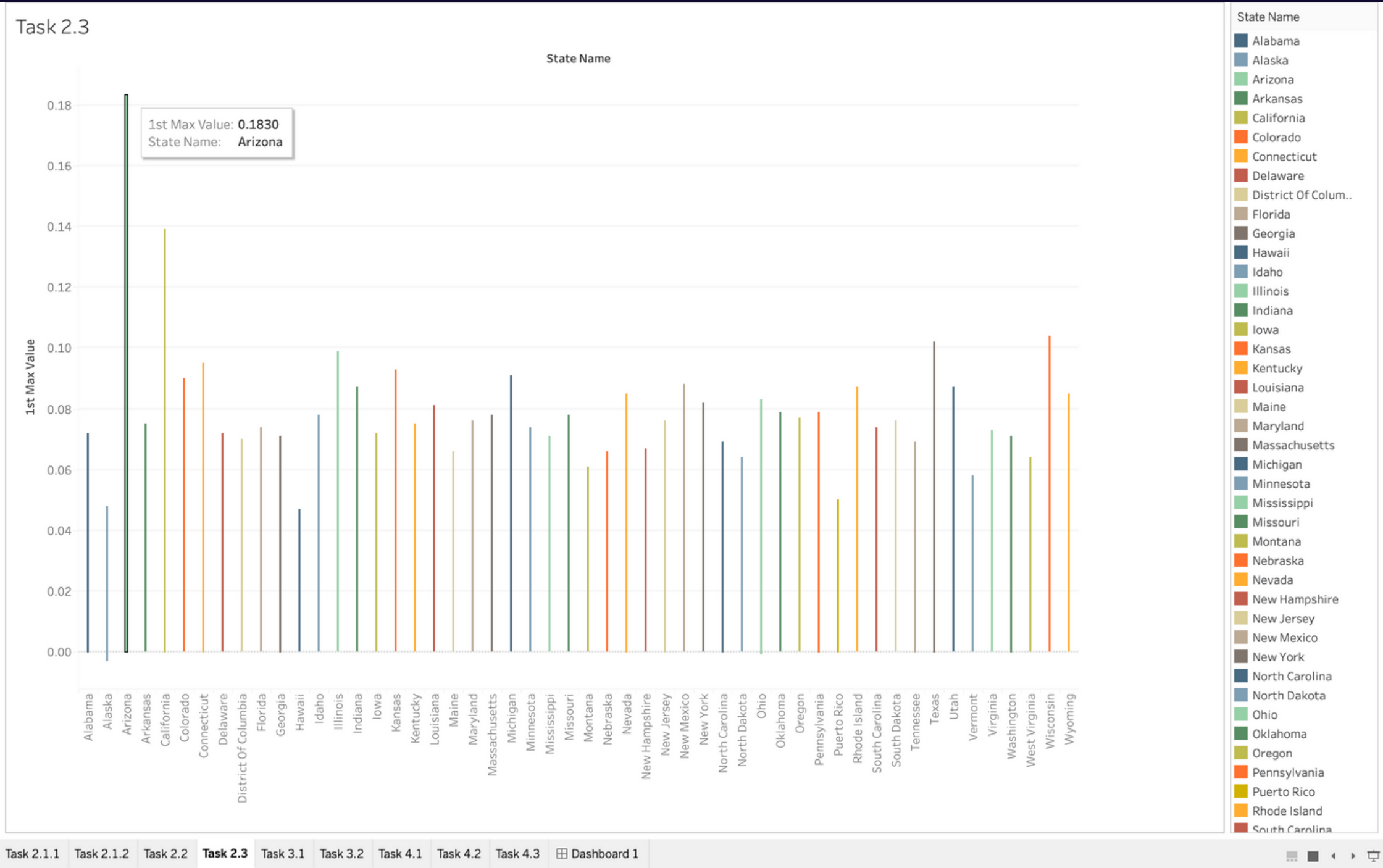
**Q3 | ANALYZE THE OBSERVATION COUNT AND ARTHMETIC MEAN BY STATE NAME**

**Q4 | ANALYZE THE EFFECT OF EVENT TYPE ON AQI?**



# Q1 | WHICH STATE HAD THE HIGHEST '1ST MAX VAULE' ?

TWO VARIABLES OF TWO DIFFERENT TYPES – BAR GRAPH (1)  
CACLCUALTION FIELD & FILTER – BAR GRAPH (2)

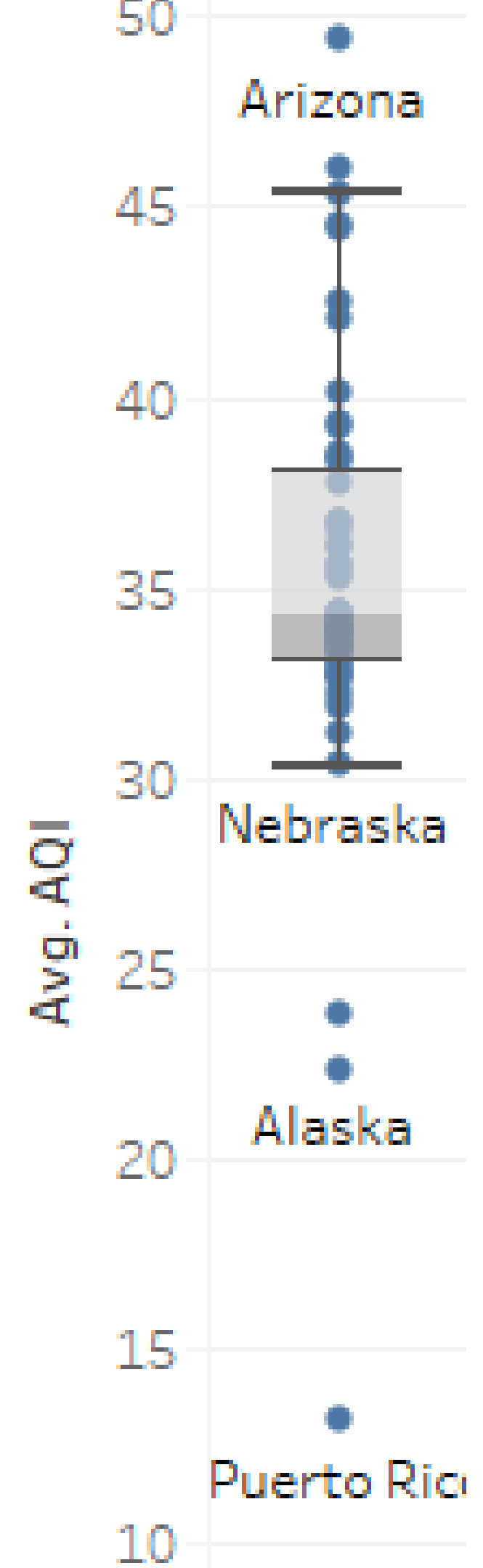


The question here is about the first max value of particle pollution that the measurement instrument caught .  
As shown above, Arizona and California has the average of max-value of pollution. That might be Due some geographical features and area of the state.

## Q2 | ANALYZE THE AVERAGE AQI BASED ON STATE NAME

### SINGLE VARIABLE ( CONTINUOUS) – BOXPLOT

At first, we tried to only represent the AQI as a single variable and used the state name as a label but as it is shown in the graph. it seems like Arizona and Alaska are outliers here, their values have passed the max and min lines so the box plot might not be the best way to analyze it.



## Q2 | ANALYZE THE AVERAGE AQI BASED ON STATE NAME

### TABLE CALCULATION

the table is showing the average AQI for the four quarters. In general, the fourth quarter has the highest numbers of all the quarters. But it is easier to analyze after sorting the results descending or ascending.

the first three states that have the highest average overall in the four quarters, are Arizona, Colorado, and Nevada.

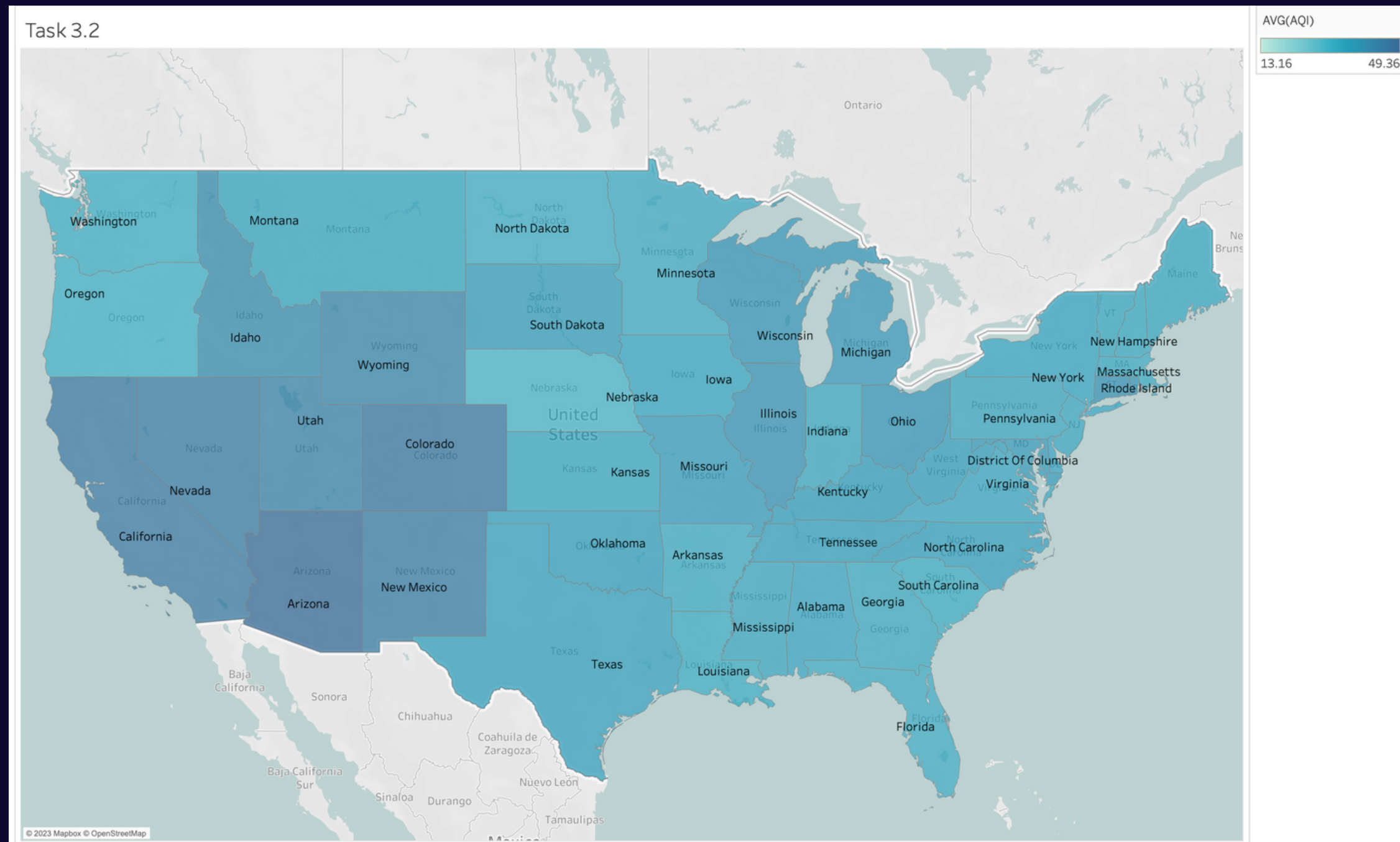
### Task 4.3

State Name	Date Local			
	Q1	Q2	Q3	Q4
Arizona	39.2	97.0	156.3	196.3
Colorado	38.2	89.0	148.7	184.1
Nevada	36.0	84.4	142.3	175.4
New Mexico	37.1	88.1	142.1	178.1
California	34.0	81.3	141.2	180.6
Utah	35.4	82.8	139.8	170.4
Wyoming	40.0	83.9	132.3	167.7
Idaho	34.0	75.2	122.6	151.8
Connecticut	34.8	77.7	121.0	148.7
Illinois	30.9	76.1	117.5	144.9
Rhode Island	37.7	78.1	117.3	146.0
South Dakota	34.0	76.3	116.9	146.0
Michigan	32.6	76.9	116.6	144.0
Ohio	32.0	74.3	116.6	143.4
Maryland	34.1	74.8	113.8	141.2
Wisconsin	31.6	75.2	112.6	140.7
Oklahoma	32.4	74.4	112.5	143.0
Missouri	32.0	73.8	110.6	138.3
Delaware	32.2	72.7	110.1	137.8
Indiana	28.4	70.9	110.1	135.3
Kentucky	32.2	73.2	110.0	137.3
Montana	34.8	73.0	109.9	136.0
New York	32.0	71.7	109.8	136.5
Pennsylvania	29.9	69.6	109.1	135.6
Texas	30.0	72.4	109.1	141.4
West Virginia	33.0	71.1	108.3	134.3
Kansas	29.3	70.9	108.1	136.4
North Carolina	35.1	74.3	107.9	138.3
Tennessee	32.8	72.7	107.7	137.5
District Of Columbia	29.2	68.6	107.7	134.0
New Jersey	30.5	70.0	107.6	133.3



## Q2 | ANALYZE THE AVERAGE AQI BASED ON STATE NAME

### GEOSPATIAL MAP



We can see that Arizona has the darkest shade which means the highest (AQI). and the higher the number the more polluted the air quality can be.

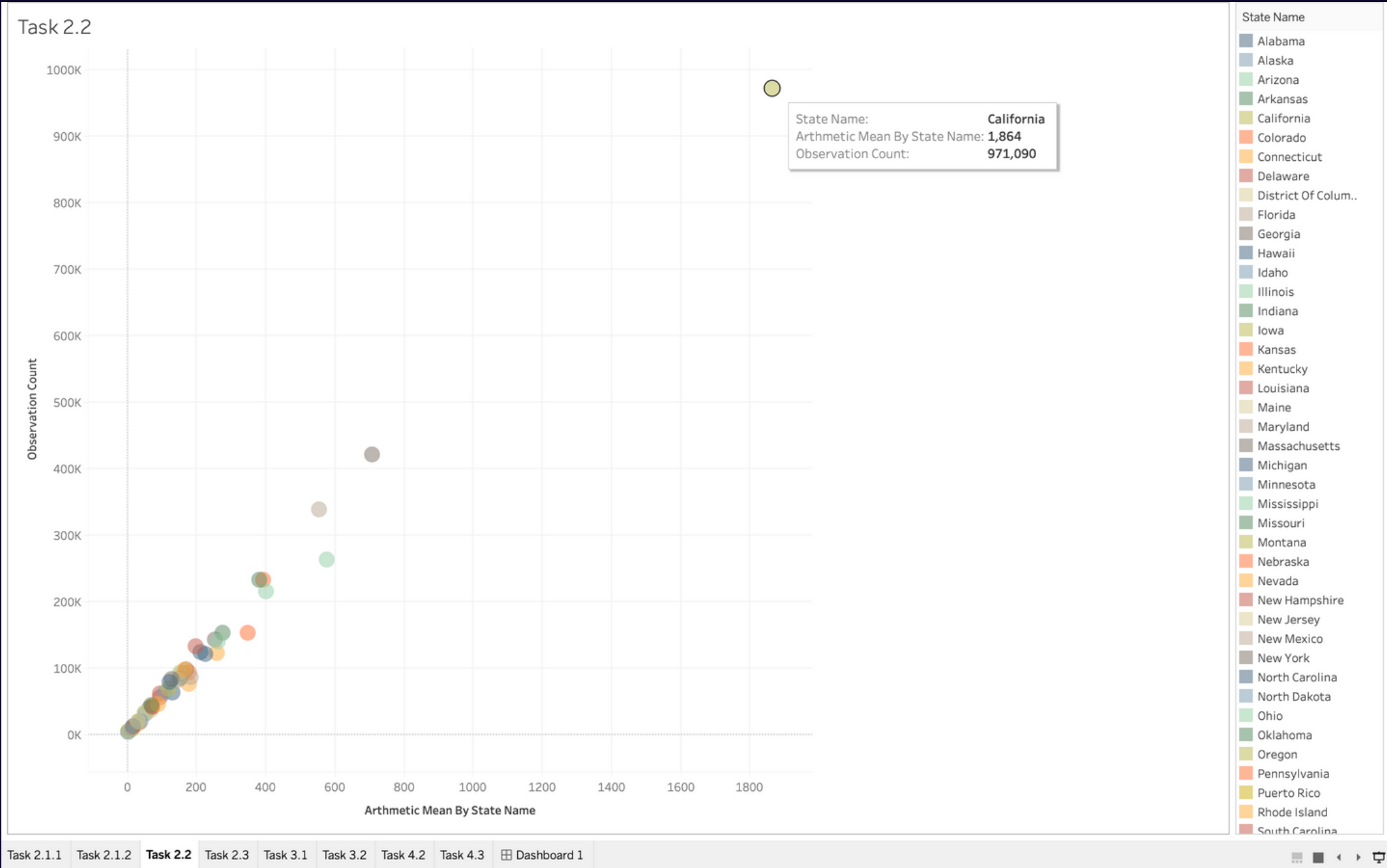
but it became harder to differentiate the other states due to the color similarity.

the AQI can be seen to associated with the area of the state too.



# Q3 | ANALYZE THE OBSERVATION COUNT AND ARTHEMETIC MEAN BY STATE NAME

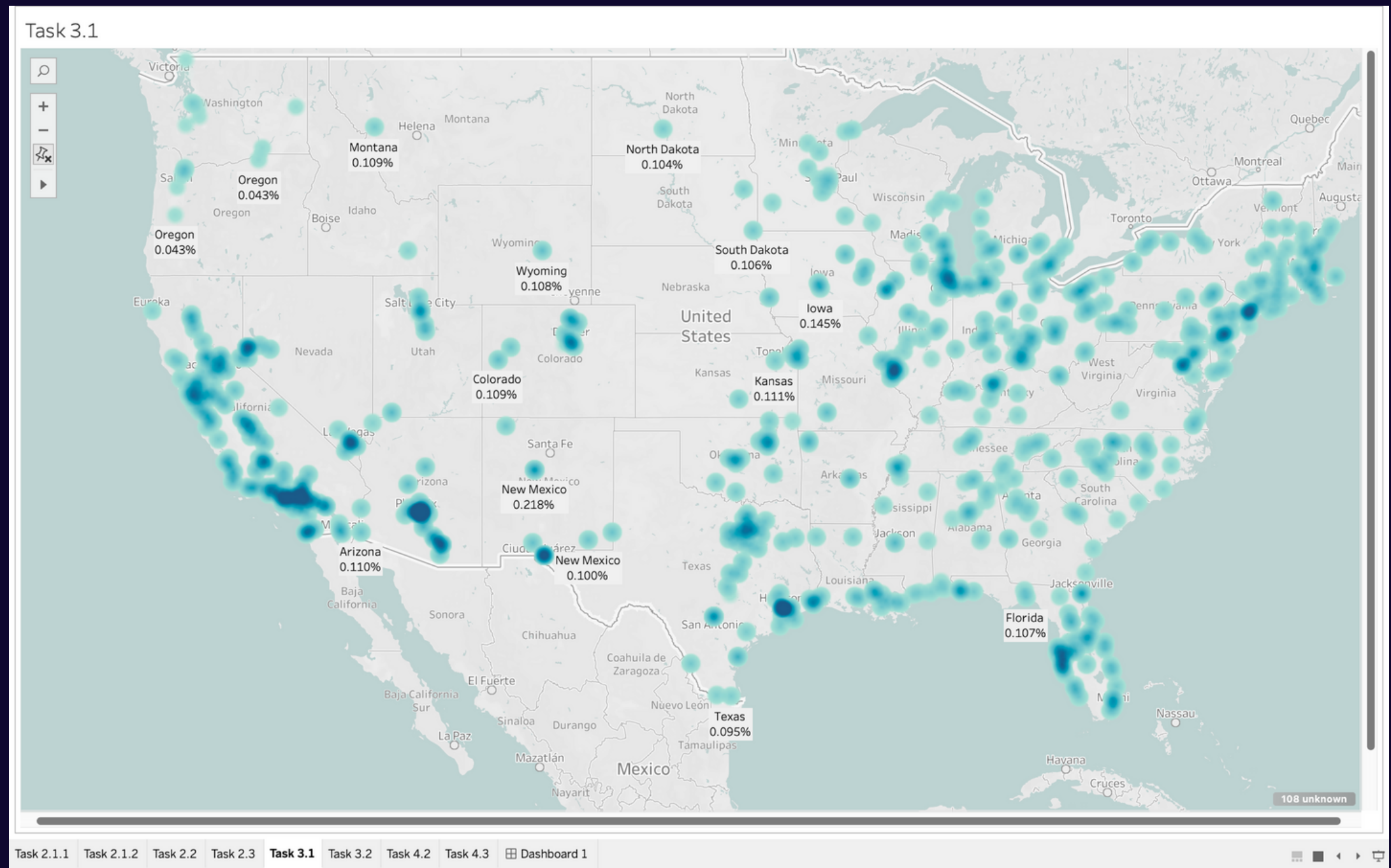
## TWO CONTINUOUS VARIABLES – SCATTER PLOT & LOD EXPRESSION



This scatter plot here is visualizing how the arithmetic mean is correlated with number of observations. We can see that most of the records are strong positively correlated. the arithmetic mean is the aggregation of results of the standards.

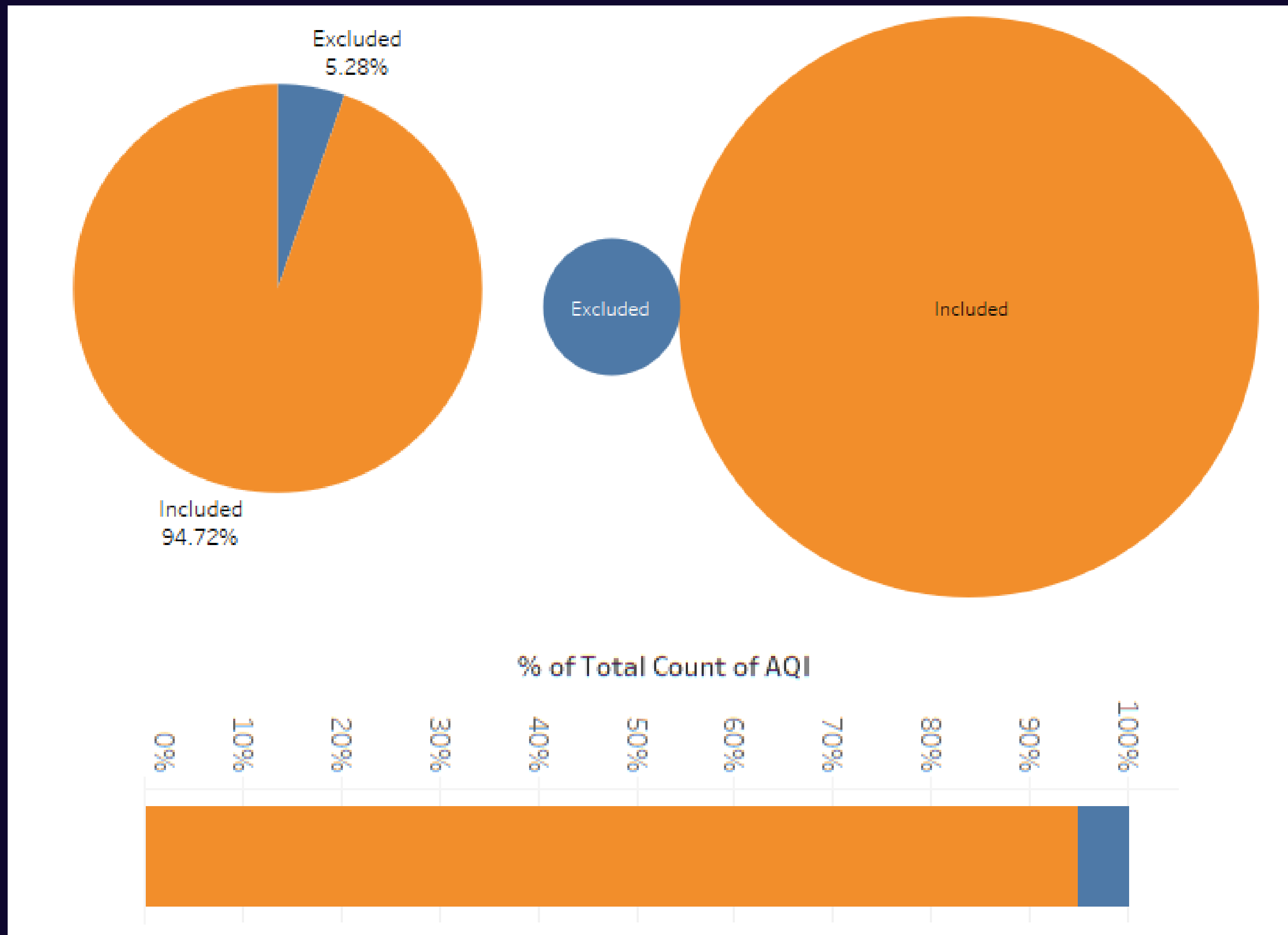
# Q3 | ANALYZE THE OBSERVATION COUNT AND ARTHEMETIC MEAN BY STATE NAME

## GEOSPATIAL MAP



In the geospatial map. We can see the distribution of the observations over the 52 states. The percentage of observations count are shown as text and average arithmetic is visualized as color brightness. The darker the shade is the higher the arithmetic mean. even though Arizona had less number of observations but still got a high arithmetic mean

## Q4 | ANALYZE THE EFFECT OF EVENT TYPE ON AQI?



These visualizations show us the effect of the occurrence of an event on the AQI ratio, we clearly note the impact of the occurrence of events on air pollution

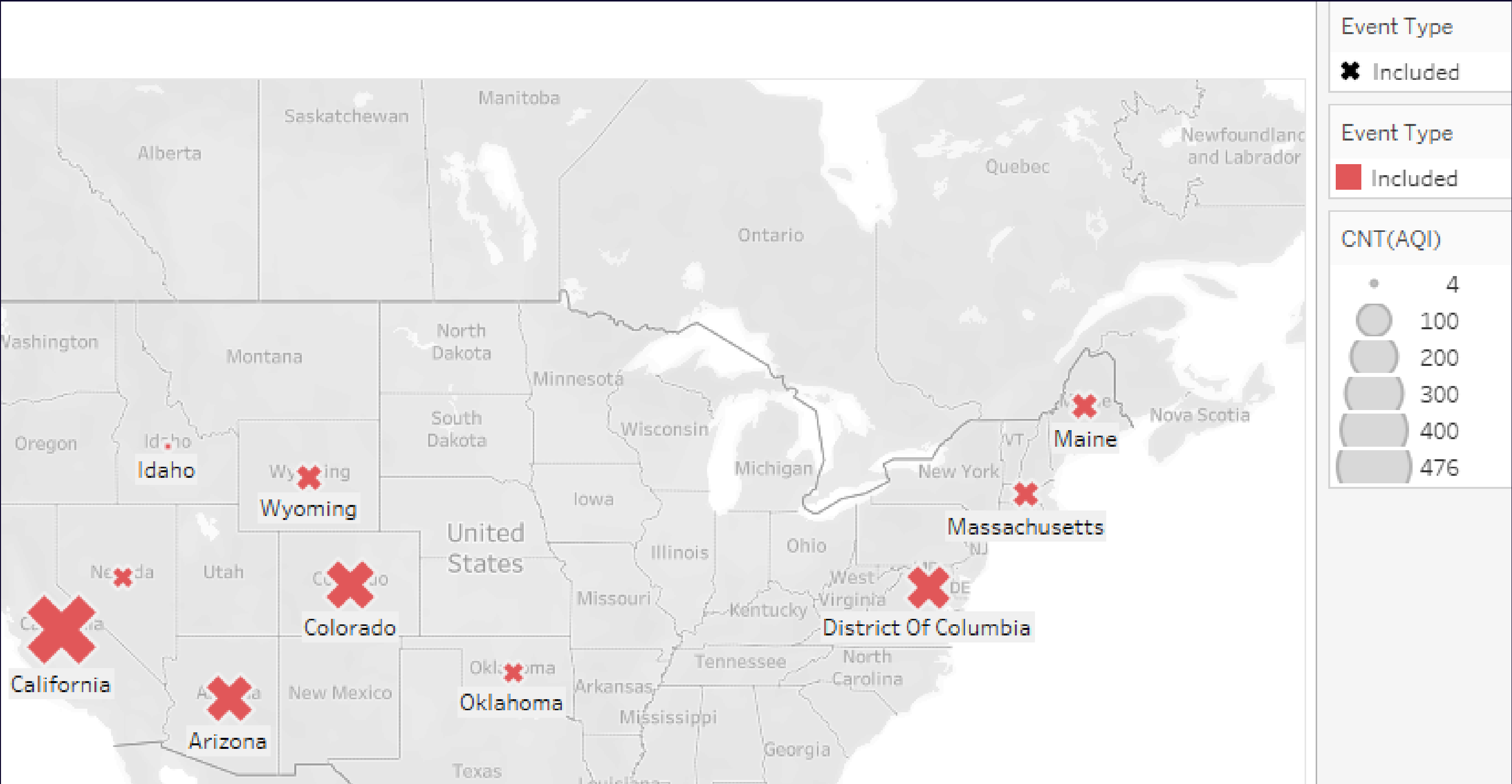
The representation was done in three different ways to convey the meaning

- bubbles
- Pie
- Fixed bars

In this case, we choose bubbles because it communicates meaning and feeling more, especially its association with air, which will make it the most appropriate choice

Q4 | ANALYZE THE EFFECT OF EVENT TYPE ON AQI?

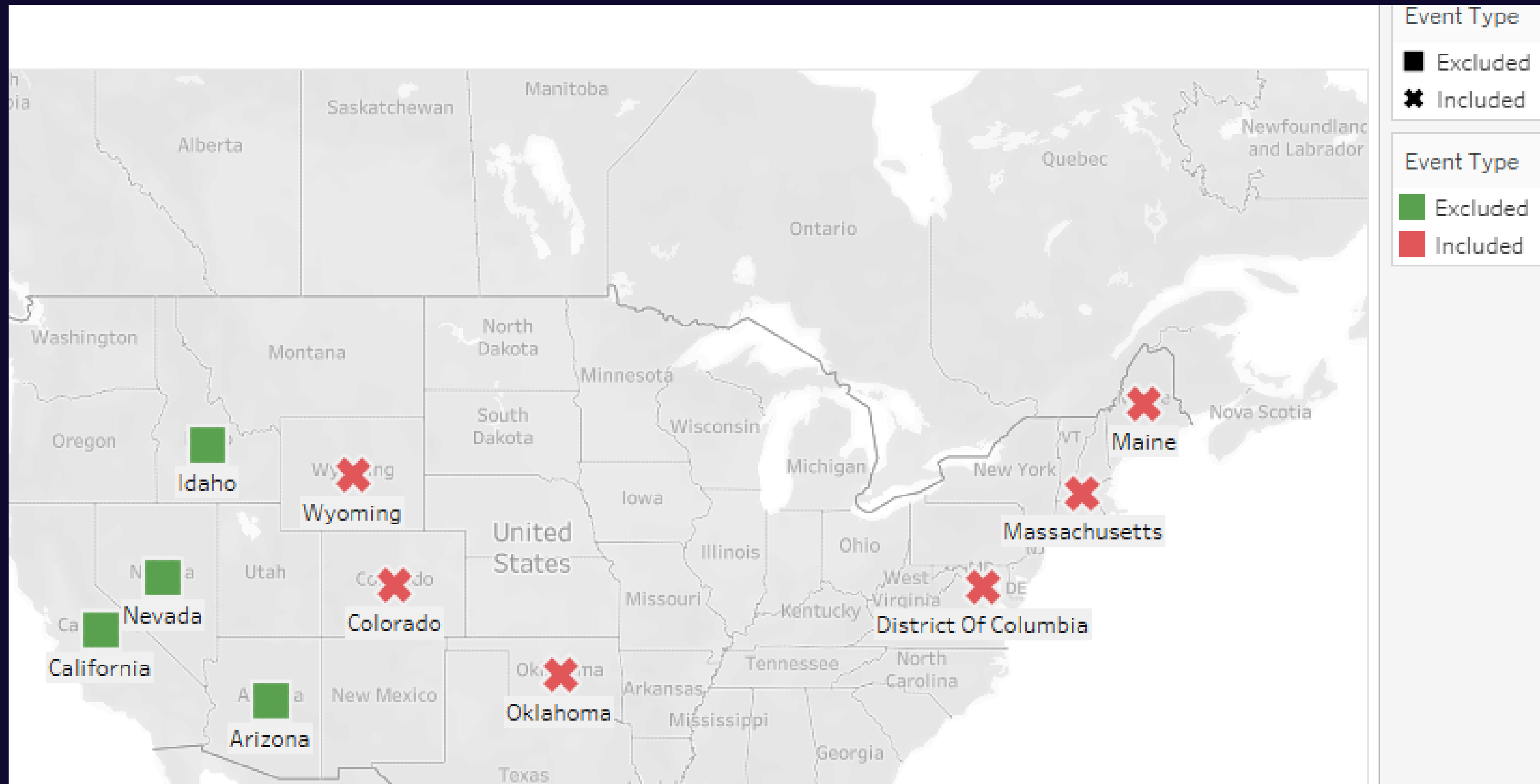
This visualization was used to display the aqi difference in the locations where the events occurred during the month (June and August).



Size has been used to represent it correctly

## Q4 | ANALYZE THE EFFECT OF EVENT TYPE ON AQI?

actually, this specifically answers the question: What are the places where events occurred during the months (of June and August)?



Colors and shapes are used for representation in this map

## CONCLUSION

After we visualized the data from different angles, we found that Arizona and California have the most 1st max vaule

We also represented AQI on a map and how it varies across states, along with what factors are likely to cause it.

We also showed a linear relationship between the mean and the number of observations taken, which may be because taking a lot of observations may be one of the reasons why the mean values are observed more.

On the other hand, we discussed the relationship of occurrence of events with pollution and how it caused an increase in air pollution, specifically the areas prone to events in specific months.





**THANK YOU  
FOR LISTENING**