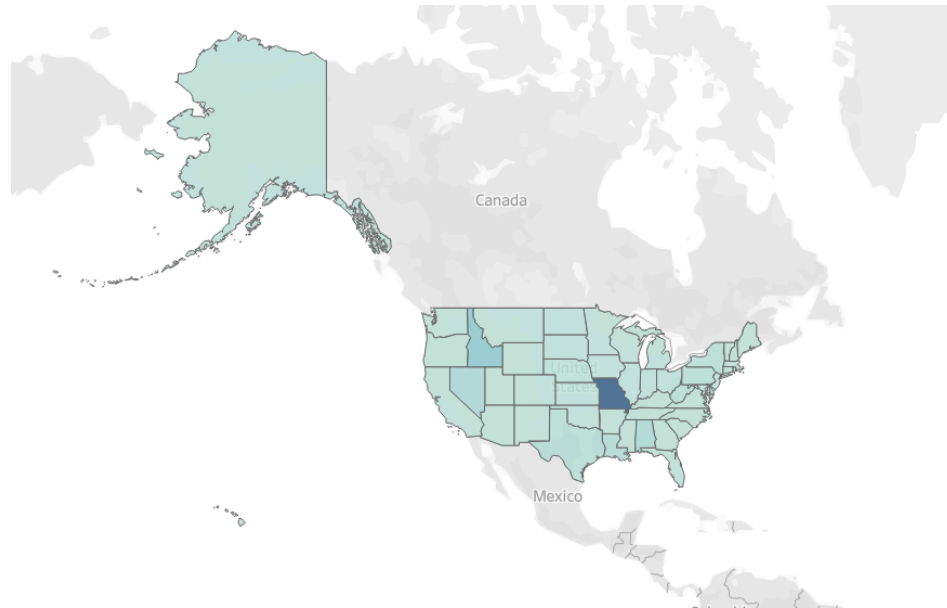


Prompts

In this white paper:

1. Repeat the context for your project, why did your group choose this topic?
 - a. We chose this topic for the project because we all were invested in the idea of the environment and how pollution impacts it. Coming at it from a data side based on states within the US seemed the most feasible given the amount of pollution data available to the public. Per state pollution in the United States is easily accessible data when compared to other kinds, such as other countries' pollution data.
2. Repeat the initial research question
 - a. How does water pollution differ between states based on landmass?
 - i. From there we decided to make our topic about pollution per size of state, water, land, etc. and tried to see if there were any correlations, positive or negative.
3. Include a written discussion of your model results.
 - a. What (if any) modifications to your data were done
 - i. Modifications that were done to our data include cleaning through the deletion of unimportant rows (Non-States) as well as unimportant columns (specific pollutant types). We focused on making our data have a column of US States only under the State column and had a few other columns that described land area as well as total pollutant loadings.
 - b. Interpret in clear everyday language that most people can understand, what the findings of the analysis
 - i. The findings of the analysis was that the bigger a state is, the more it tends to pollute. There were a few exceptions to this such as Missouri, but for the most part this trend (linear regression) held true.
4. Produce visual illustrations of your model results



- a.
 - i. An analysis of this graph is that based on total area and total pollutants, missouri pollutes the most per sq ft it has. Idaho comes in second as seen by the coloration.
5. In this section, you can very carefully begin to draw potential implications of your findings and propose possible explanations for your findings
 - a. Possible implications of my findings include trying to find out why bigger states tend to pollute more, see if it is possibly population based. I propose a possible explanation for my findings as bigger states tend to have more people, which means that there will be more pollution as a byproduct of each person in said state.
6. What remaining questions arise from your analysis?
 - a. Given what you found, what questions are still unanswered? What might be the “next phase” of the project?
 - i. Given what we found, some unanswered questions I have are how company density affects the amount of pollution in a state as well as how the pollution gets measured. I believe the next phase of the project would be looking at company density in all 50 states and DC, what companies pollute the most, and other company specific questions.
 - b. What limitations does your data/model have that would have been helpful in better understanding your research question?
 - i. Some limitations that our data/model has that would have been helpful in better understanding our research questions would be knowing the limitations of how the depth and breadth of pollution testing was done. Along with this understanding how it was measured to better understand the methodology would have been vastly helpful.