DSE 241 – Data Visualization

Submitted by Alejandro Hohmann

Exercise 3 Report: *Timeline of US State Populations and CO2*

Report Contents

[The Viz 2](#_Toc126272518)

[Motivation 2](#_Toc126272519)

[Data augmentation (if any) 2](#_Toc126272520)

[Tasks 2](#_Toc126272521)

[Expressiveness of design 2](#_Toc126272522)

[Effectiveness of the solution 2](#_Toc126272523)

[Interaction 3](#_Toc126272524)

[Conclusions 3](#_Toc126272525)

# The Viz

Aggregated to total USA

Diagram, map

Description automatically generated

Example of hovering over a state - TEXASDiagram

Description automatically generated with low confidence

Example of hovering over a state – CALIFORNIA

Diagram

Description automatically generated

# Motivation

This visualization is intended to quickly edify users on:

* which sectors create the most emissions
* which states create the most emissions
* what is the relationship between population and CO2 emissions (by sector)
* how these relationships have trended over time

# Data augmentation (if any)

Joining two excel sheets within Tableau

* State = State
* Year = Year

Changing datatype

* Tableau reads in *Year* as an integer, but in some cases it is a categorial variable (was helpful while initially brainstorming visualizations). Duplicated the *Year* measure, tight-clicked the pill from both sheets and converted to string.

Filtering data

* Population sheet includes a “USA” total under the state column – excluded this total due to it being redundant with Tableau total/subtotal functionality.

Renaming fields

* Renamed “*value*” from CO2 Emissions sheet to “*MMTCO2*” aka *Million Metric Tons CO2*.
* Renamed “*value*” from Populations sheet to “*Population*”.

# Tasks

The viewer is required to spend some time orienting themselves to the viz as there is a lot of information encoded. This is more intuitive when interacting with the viz, especially by hovering over different states, though not required.

# Expressiveness of design

Combining the scatter chart with the state map does a really nice job for an interactive viz.

# Effectiveness of the solution

There are two different color channels:

* Scatterplot color channel by sector
  + The color channel in the scatter plot does a good job of partitioning the industry for the user, it is easy to differentiate which industry is which. Especially because there are only five sectors.
* Avg MMTCO2 color channel by state
  + Does a decent job at demonstrating differences by state. At least for the larger emitting states of Texas and California

In the scatter plot, there is a size encoding for year. In the case of this viz, this does a decent job of showing the measurements for each year because there is not a lot of overlap. So it almost gives the appearance of a trendline. This size encoding would not work well if the data was more bunched together. However, in this case it does provide more context and the user can pretty clearly pick up on the timeline for a particular industry’s CO2 and Population. If there is some ambiguity, the scatter points can be hovered/clicked to identify the year using the tooltip.

The scatterplot also effectively shows the relationship between Population. MMTCO2 is on the y-axis to help the user hone in on this being the dependent variable and it especially works because population has consistently increased with time.

# Interaction

The user can take the viz as is or can hover over individual states. Looking at individual states is telling for two reasons:

* Identify largest sector of CO2 emissions for an individual state
* Identify if a particular sector in a particular state is increasing/decreasing emission over time, and whether this is due to population change or if there is an outside factor

The user can also scroll across the state map to view more states and zoom in on the map to more effectively interact with the smaller states. The tooltip helps the user identify which state they are viewing and the chart title at the top of the viz updates with each state as well.

# Conclusions

This viz is effective at conveying rank, trend, and correlation for CO2 and population across multiple dimensions (time, states, and sectors). The viewer can see that for the US, **Electric Power** is the largest CO2 emitter in every year in the dataset. Also, despite the US population increasing, emissions from **Electric Power**, **Transportation**, and **Industrial** sectors have come down in recent years after steadily increasing. The conclusions for individual states can be different.