PROJECT 5

SELECTED DATA: US Census Demographic Data

TOOL USED: Tableau Public

SUMMARY OF DATASET: This documentation contains some insights drawn from the US Census Demographic Data which contains the census data for all States in the United States, in 2015. The data is a pretty large dataset, having 37 columns and 3220 rows, as well as the 2015 demographic data for roughly 50 states and over 1,800 counties.

INSIGHTS

Question 1: Which State is the most populated?

Solution Approach: I filtered the data by state, with the sum of the total population in the colour and size marks. I then included the hierarchy made up of the County and State in the text mark and visualized the data with a **Treemap Plot**. This enables the viewer assess the total population by state, and county, if should the need arise. For texts that could not be contained in the treemap plot, the data can be seen by simply hovering around the plot.

Finding(s): From the visualization, it could be seen that California was the most populated State in 2015.

Link: US Commute Visualization (Project 5) | Tableau Public

Question 2: Does employment status/rate have any effect on the commute rate?

Solution Approach: I gained some insight into this by plotting the sum of employed and unemployed against the mean commute and filtered by State in a **Scatterplot.**

Finding(s): Unemployment correlated positively with the mean commute and so did the employed. Employed however has a medium correlation strength. It seems unemployed persons commute just as much as employed persons.

Link: US Commute Visualization (Project 5) | Tableau Public

N.B: *The worksheet containing the plot is appended above the page as a tab.*

Question 3: Which state does the most commuting occur?

Solution Approach: I gained some insight into this by plotting the State,County hierarchy against sum of the mean commute and filtered by State in **Bar Plot.**

Finding(s): From the plot, Texas has the highest rate of commuting.

Link: US Commute Visualization (Project 5) | Tableau Public

N.B: *The worksheet containing the plot is appended above the page as a tab*

I also created a dashboard and a story to further present my visualization and they can be viewed via this link (***N.B:*** *The dash board and story are appended above the page as a tab).*

CONCLUSION

This project contains a link via which you can view all plots which include a treemap, scatter plot and bar plot, as well as a dashboard and story. All these are appended as tabs above the published page.

A treemap was first used because it properly captured the answer to the question at a glance. With the size showing the population and the colour intensity showing the density, the visualization can be clearly understood. Also, the scatter plot was useful in showing the correlation between the variables, while the bar plot gave insight on the mean commute by state (and county if need be).

For my colour, I kept it blue as the plots already clearly visualized insight drawn (by length, height, size and correlation) and did not need variation of colours.

RESOURCES: Project Walkthrough by Josh