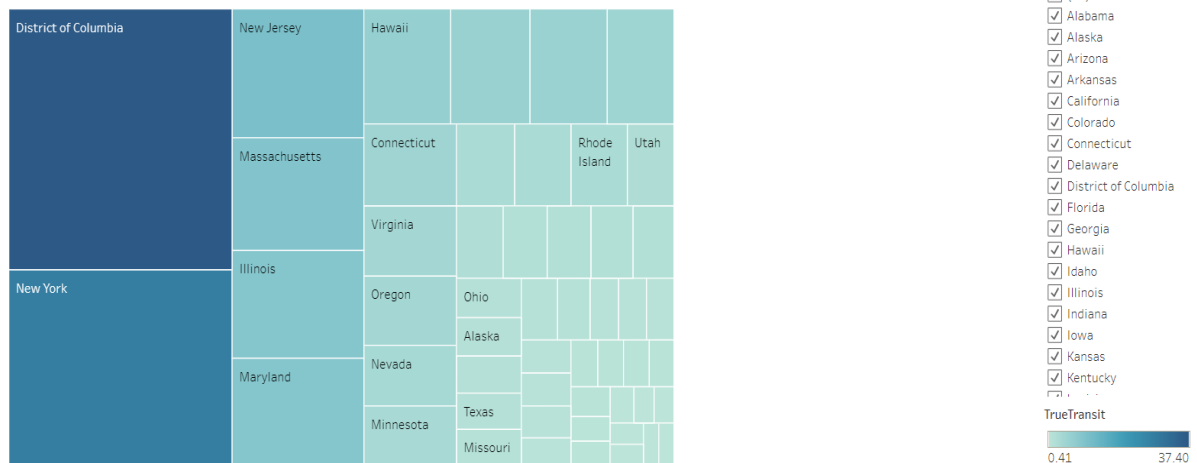# Report on Visualization of US Census Demographic Data.

In this report we intend to summarize the various findings from our visualization of US Census Demographic Data. This report is structured in a question-visualization-discussion. This is briefly shown below.

- **Which states have the best transportation?**
  The visualization that answers this question is shown below.

States in the USA with the best Transportation

The link to view this visualization is
https://public.tableau.com/app/profile/fidel.imasuen/viz/project5sheet1/Sheet3?publish=yes

Discussion

In order to find out which states have the best transportation we first need to understand the criteria for determining the effectiveness of transportation in a state. Some of the options to choose from are the mean commute time and the proportion of people in a state using the public transportation infrastructure. The option chosen is the proportion of people in a state using the public transportation infrastructure because the best way to find out the effectiveness of a system is by popular consensus.

This means that the higher the percentage of people using public transportation infrastructure in a state the more efficient the transportation infrastructure must be because if the infrastructure is not efficient then a lot of people won't want to use it. At first glance the mean commute time seems like a good way to determine states with very good public transportation infrastructure however this option was not used because the mean commute time is highly dependent on the size of the state i.e. a smaller state will have a smaller mean commute time compared to a larger state. Therefore using the mean

commute time as a metric is not an accurate way of determining the effectiveness of a state's public transportation infrastructure.
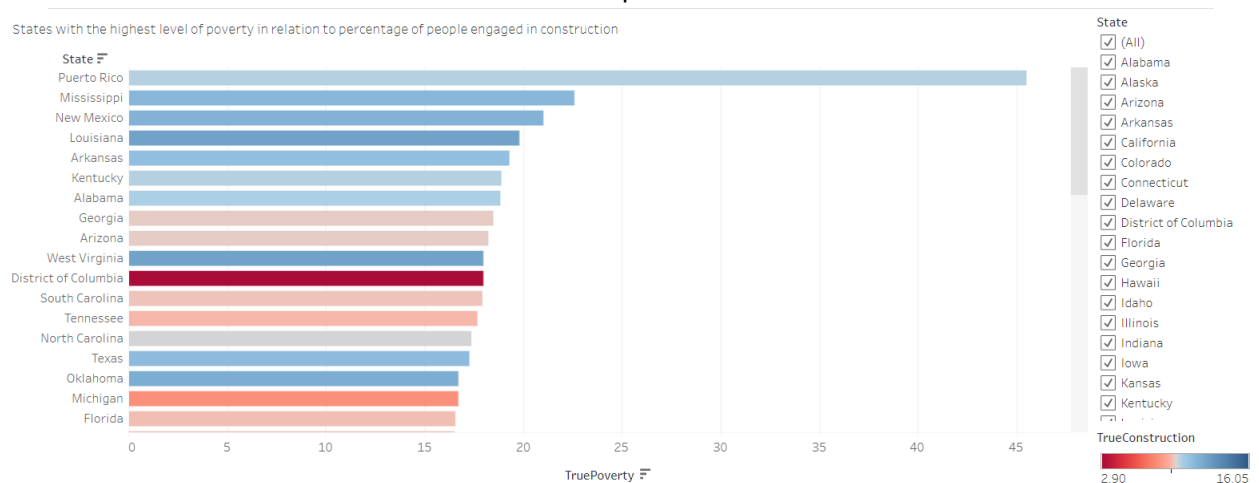
In order to visualize the states with the best transportation based on the percentage of people in a state using the public transportation infrastructure we need to have an accurate aggregation of the number of people using the public transportation from the county level all the way up to the state level. Using the sum function in Tableau on the transit column tends to give inaccurate results so I created a new calculated field called TrueTransit which is based on the formula

$$sum([Transit]*[Total\ population])/sum([Total\ population]).$$

Aggregating over this newly created column tends to give an accurate result. The visualization chosen to answer this question was the heat map which was designed to convey information based on color intensity and size. Hence from the picture shown above it is easy to see that District of Columbia is the state with the best transportation.

- **Is there a relationship between the level of poverty in a state with the percentage of people engaged in construction?**
  A screenshot of the visualization that answers this question is shown below.

The link to this visualization is given by
https://public.tableau.com/app/profile/fidel.imasuen/viz/project5sheet2/Sheet2?publish=yes

Discussion

In order to answer this question the visualization of choice used was a color coded bar chart. The length of the bar chart corresponds to the percentage of people living in poverty in a particular state while the diverging colors of the bar chart corresponds to the percentage of people involved in construction in a particular state.

In order to make this visualization accurate, aggregation issues for the poverty and construction column had to be resolved. This is because using the sum function over these columns tends to give errors when summing from the county level all the way up to the state level. Hence calculated fields called Truepoverty and TrueConstruction were created to solve this issues. The formula these calculated fields were based on is shown below.

For TruePoverty,

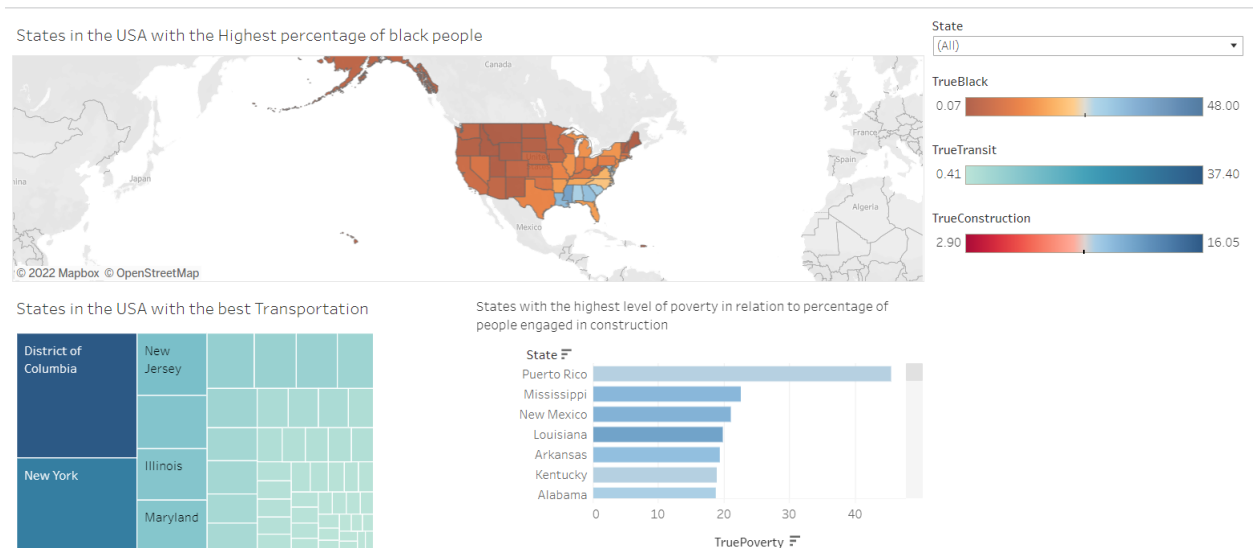$$\text{sum}([\text{Poverty}]*[\text{Total population}])/\text{sum}([\text{Total population}]).$$

For TrueConstruction

$$\text{sum}([\text{Construction}]*[\text{Total population}])/\text{sum}([\text{Total population}]).$$

Aggregating over these newly created columns tends to give an accurate result. From the visualization above we can see that state with the highest level of poverty is Puerto Rico. However, there does not seem to be a relationship between the percentage of people involved in construction and poverty levels in a State. For example Georgia and West Virginia have almost identical poverty levels yet they have widely divergent percentage of people working in construction.

- **Which States have the highest percentage of black people and what are some of the characteristics of these states?**
  The visualization to answer this question is shown below.

The link to this visualization is given by
https://public.tableau.com/app/profile/fidel.imasuen/viz/project5dashboards/Dashboard1?publish=yes

Discussion:

In order to answer the question posed, the visualization used is a dashboard. The dashboard contains three sheets and is a combination of a geographical map, heat map and coloured bar chart. In order to get an accurate percentage of black people in a state, aggregation issues when summing from the county level all the way up to the state level had to be resolved. This was done by creating a new calculated field called TrueBlack. The formula this calculated field is based on is shown below.

$$\text{sum}([\text{Black}]*[\text{Total population}])/\text{sum}([\text{Total population}]).$$

Aggregating over this newly created column tends to give an accurate result. Interactivity is incorporated in the dashboard, clicking on any state in the geographical map changes the heatmap and the barchart i.e we get to have closer look at the characteristics of a particular state. There are interesting insights that can be inferred from the dashboards for example we see that states close to the coast tend to have a higher percentage of black people and this gradually decreases as you move away from the coast.

Resources:
No external material was consulted in the making of this project