Dataset used: US Census Demographic Data

Insight no. 1

Link:
https://public.tableau.com/views/UScensusvizfinalproject/Publictransportationvizproject?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

Summary: My project is dedicated to analysis of US public transportation system on a state level. Our first dashboard could be called “paradox of New Jersey”. It contains three visualizations. First one is a map plot which is distributing whole states by median of mean commute time in minutes by color from blue (low) to orange (high). The text table on the right as well as filter above play a supportive role for the map plot and allow us to conclude, that New Jersey residents spend more time to get to work then their counterparts from other states, the fact, that doesn’t characterize NJ public transportation system in a good light. But, if we will take top 5 states in terms of median percentage of transit use New Jersey will appear on the first place again, which is a very curious insight! Also, I have to explain, why data from county level to the state is aggregated by using median. I think, that it seems to be much more ‘fair’ on a state level, since big cities with a high population density almost always appear as clear outliers, so using average aggregation affects the outcome and doesn’t represent the situation in the whole state appropriately. In other words – we can imagine the dead desert around the New York city main boroughs and transit average will still be relatively high. For the same reason, District of Columbia was excluded from the top 5, since it is counted as the whole state, but appears to be an outlier as being just a one big city of a special status without that “geographical diversity” regular states usually demonstrate. In a summary for the next dashboard we will discuss fairness of my aggregation’s choice in more details.

Design: Map plot and text table are working together here in more clear visual conveying of situation with a mean commute time in different states. The neutral colors appropriate for color blind people were chosen. The simple and clear bar plot below the map is made as interactive filter for another two visualizations.

Resources: N/A

Insight no. 2

Link:

https://public.tableau.com/views/UScensusvizfinalproject/Publictransportationvizproject?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

Summary: The second dashboard is further elaborating the message from the first one by % of transit use visual analysis on county level. As you could notice in previous dashboard all of the top 5 states in terms of public transport use are located in the northeast of the country. They were including in a county level map plot as well as New Your state and District of Columbia for better emphasizing the meaning of median aggregation method use. In a top 10 counties by % of transit use text table on the right you can find out, that 6 from 10 counties are located in New York State, which is not even in our top 5 from the first dashboard, which proves that they are actually outliers! The similar case appears with regard to District of Colombia, which is counted as county and state at the same time.

Design: Similar color palette as in the first dashboard with dark background was used for the map plot. Both text table and map are fully interactive. For better visual clarity, transportation means distribution pie chart was added to the tooltip.

Resources: N/A

Insight no. 3

Link:

https://public.tableau.com/views/UScensusvizfinalproject/Publictransportationvizproject?:language=en-US&publish=yes&:display_count=n&:origin=viz_share_link

Summary: Comparing the usage of public transport to driving your own vehicle, it is crucial to realize the two different concepts, laying behind those two options. In short, the key difference is that your car is your property that you should buy whereas public transportation is a service, meant to be affordable for everyone. In the final dashboard I was trying to find out, is there any correlation between the poverty level and % of commuting in public transport. For the same reasons as before, District of Colombia was excluded from the analysis.

The surprising insight is that no positive correlation was found. In fact, the scatter plot demonstrated the weak negative correlation, as poorer states appear to use public transportation in a smaller extent than their richer counterparts. Another unexpected insight is that top 5 states by average income exactly matches the top 5 states with a highest median % of transportation use from the first dashboard! In other words, in some sense, relatively rich folks prefer buses. Is it a matter of environmentalist's movement influence or simple the high quality of public transportation service or absence of unconscious connection between social status and private vehicle's usage? Only further investigations will show.

Design: The scatter plot was used as a best choice to represent potential correlation between two variables. The same color-blind friendly blue-orange palette was used to distribute states by average income.

Resources: N/A