Udacity: Data Visualization Project

Project Overview

Objective of this project to apply Data analysis skills along with data visualization and prepare/publish analysis into dashboard or story format. US Census Demographic Data is being used for Data analysis and visualization to showcase various findings and pictorial presentation.

Source of Dataset

Downloaded dataset from Udacity.

Tools Used

Tableau desktop & Tableau public.

Visualization Principles

Visualization can be used for Exploratory and Explanatory analysis, so we have performed both types of analysis in our case study.

- During exploratory analysis, we have identified set of dimensions, measures and relate them with the help of various worksheets.
- During explanatory analysis, we narrated a story to showcase analysis with the help of worksheet in defined sequence.

Area of Analysis.

Transportation System (Visualization #1)

1. **Tableau Public Link**

https://public.tableau.com/views/DS_Tableau_Project_Story1/Story1?:language=en-US&:display_count=n&:origin=viz_share_link

2. **Summary**

Certainly, this is very subjective questions and need to assess lots of factors to conclude but in Dataset, we have following columns related with transportation so we will use them to make some analysis.

- Drive: Percent commuting alone in a car, van, or truck.
- Carpool: Percent carpooling in a car, van, or truck.
- Transit (Public Transport): Percent commuting on public transportation
- Walk: Percent walking to work
- Other Transport: Percent commuting via other means

There is another column "MeanCommute Mean commute time (minutes)", but there is no direct relationship of its value with other columns.

In general, Self-drive, carpool, walk or other transports are not mode of transport to say State/County has best transport system, rather its Public transport (Transit) which is less-expensive, frequent, ensure connectivity & reachability, eco-friendly, flexible ticketing options.

Here we will analyse data to see which states has good Transit system, which counties of states contributed more into transit system, what is % usage of other modes of transport on those states/counties. We will also see modes of transport usage by people in high populated and less populated states.

3. Design comments

Following design principles considered:

1. Use of maps or symbols map with Geo based columns.
2. Use of Labels wherever it is significant.
3. Use of appropriate aggregate function to prepare correct information out of data.
4. Use of Dual-Bar chart to showcase comparative analysis of measures.
5. Use of appropriate filter, so end-user can explore various possibilities.

Population density & Diversity (Visualization #2)

1. Tableau Public Link

https://public.tableau.com/shared/76HCG8ZHN?:display_count=n&:origin=viz_share_link

2. Summary

In US census data, we have data about population in each states/county along with total number of Men & Women. We also have % population of various community like Asian, black, Hispanic, Native, White, pacific. Some analysis around this data can be useful for various purposes like:

1. Launching new business targeting to specific community.
2. Organize festive events on States/Counties based on % population.

3. Evacuation operations by Home country.
4. Migrants preferred State/County by specific region of people.

There could be many more business case of such data analysis.

3. Design comments

Following design principles considered:

1. Use of maps or symbols map with Geo based columns and use of Grey scale color to highlight information.
2. Use of appropriate color to match with generic saying Like "Women prefer Pink, and Men prefer Blue".
3. Use of aggregate function with "Size" marks to showcase density. Big circle is always considered high and small is low.
4. Use of Dual-axis chart with same scale for both the measures to showcase comparative analysis.
5. Use of appropriate stepped colours to showcase only useful range.

GITHUB Public Reference

https://github.com/ankjain004/Data_Visualization_Project.git