U.S. Income Analysis:Delving into State and City Economic Structures

Overview

The objective of analyzing the USA income dataset is to gain insights into the distribution of income across different states, counties, and cities. By examining mean, median, and standard deviation of income, we can identify trends and disparities. The dataset contains columns for latitude (Lat) and longitude (Lon).which is used to understand the geographic and demographic factors related to income can inform policy decisions and resource allocation.

Objective

1. Income Distribution Analysis
2. Geographical Income Analysis.
3. Comparison by State and City
4. Demographic Insights
5. Area Analysis

Key Objectives:

❖ Income Distribution Analysis

- ➢ Investigate how income varies across different states, counties, and cities.
- ➢ Identify regions with higher or lower income levels.
- ➢ At 88,657.64, New Jersey had the highest Average of Mean and was 50.26% higher than Michigan, which had the lowest Average of Mean at 59,001.12.
- ➢ Average of Mean and total Average of Median are positively correlated with each other.
- ➢ Across all 10 State_Name, Average of Mean ranged from 59,001 to 88,657.64, Average of Median ranged from 70,040.05 to 1,25,685.83, and Average of Stdev ranged from 42,550 to 57,617.

❖ Geographical Income Analysis

- ➢ Understand the central tendency and spread of income data.
- ➢ A map effectively displays geographical distribution, with color saturation representing income levels in different regions, allowing for easy identification of areas with higher or lower incomes.
- ➢ California had the highest Sum of Mean and was 4,957.89% higher than Wyoming, which had the lowest Sum of Mean at 5066450.
- ➢ California accounted for 11.81% of Sum of Mean.
- ➢ Across all 52 State_Name, Sum of Mean ranged from 5066450 to 2562557.

❖ Comparison by State and City

- ➢ At area of 256255700, California had the highest Sum of Mean and was 4,957.89% higher than Wyoming, which had the lowest Sum of Mean at area of 5066450.
- ➢ Sum of Mean and total Sum of Median are positively correlated with each other.
- ➢ California accounted for 11.81% of Sum of Mean.
- ➢ Sum of Median and Sum of Mean diverged the most when the State_Name was California, when Sum of Median were 73652384 higher than Sum of Mean.

❖ Demographic Insights

- ➢ Examine land and water area (ALand, AWater) in relation to income.
- ➢ The area analysis, a scatter plot or bubble chart is appropriate for visualizing the relationship between two variables. Placing population density on the x-axis and income levels on the y-axis allows for analysis of how income distribution varies with population density.
- ➢ Across all 5 Zip_Code, Sum of sum_w ranged from 62,475 to 6,12,241 .
- ➢ At 6,12,241.92, ZIP_PIN 60608 had the highest Sum of sum_w and was 879.97% higher than ZIP_PIN 93304, which had the lowest Sum of sum_w at 62,475.41.
- ➢ 60608 accounted for 55.51% of Sum of sum_w

5. Area Analysis

- ➢ Aleutians East Borough had the highest Average of ALand (9,63,64,04,567.75) and Average of AWater (1,51,68,95,958.06).

CONCLUSION

- The USA income dataset provides valuable insights into income distribution and geographic disparities.
- States like California and New York exhibit higher mean incomes, while others may have lower averages.
- Income variation exists within counties and cities, impacting overall economic well-being.
- Policymakers can use this information to address income inequality and allocate resources effectively.

BY AKHYATA MOHARANA

PD21_015