

## **Project**

### **US Census Demographic Data**

Following data analysis gives an insight into the US Census Demographic Data dataset in 2015. This analysis is looking more deeply at the following issues across America:

- What transport methods are used for commuting
- Poverty, income and unemployment correlations
- Self-employment correlations to unemployment, working at home and unpaid family work

#### **Type of transport used for commuting in US states in 2015**

Visualization: [https://public.tableau.com/profile/milja#!/vizhome/USCensusDemographicData\\_7/Dashboard1](https://public.tableau.com/profile/milja#!/vizhome/USCensusDemographicData_7/Dashboard1)

In the analysis of the transportation used for commuting to and from work shows that across the US the most popular method is commuting alone in a car, van or truck. In most of the states driving to work is used 70%-85% of the time. There are only two exceptions Alaska and District of Columbia where driving (47% and 34% respectively) to work stays under 50% of the total transport used. In District of Columbia people go to work and come from work mostly by using public transport (37%). Alaska is the state that has the highest percentage of walking (24%) and of carpooling (12%) to and from work. All percentages are given as the percentage of the total transport used commuting to and from work in a state.

For visualization I have used stacked bars as I wanted to visualize the whole transportation used rather than focus on the individual categories.

#### **Median income per capita across the US in 2015**

Visualization: [https://public.tableau.com/profile/milja#!/vizhome/USCensusDemographicData\\_7/Dashboard2](https://public.tableau.com/profile/milja#!/vizhome/USCensusDemographicData_7/Dashboard2)

Here I give an overview of the median income per capita in different states and counties. Figures express annual income per capita. Dashboard shows two maps where it is possible to see comparisons for both median income per capita in states and in counties.

On the first map there is shown the median income per capita in each of 52 states. Among the states the highest median income is in District of Columbia (\$47,675), Rhode Island (\$36,851) and Connecticut (\$35,529). The lowest median income is in Puerto Rico (\$9,217) and Mississippi (\$18,968).

On the second map it is shown the median income across the counties. Here the median income has been divided into categories. For this a calculated field has been created.

#### **How poverty correlates to income and unemployment rate in the US states in 2015?**

Visualization: [https://public.tableau.com/profile/milja#!/vizhome/USCensusDemographicData\\_7/Dashboard3](https://public.tableau.com/profile/milja#!/vizhome/USCensusDemographicData_7/Dashboard3)

According to our dataset, the average percentage of living under poverty level is 17,5%. Official statistics say that in 2015 it was 13,5%<sup>1</sup>.

These data may differ in few reasons. Firstly, in the current dataset the poverty rate is given in percentages and as calculations do not base on the figures of source data, which may be the cause of different results.

---

<sup>1</sup> Source: Income and Poverty in the United States: 2015: <https://www.census.gov/library/publications/2016/demo/p60-256.html>

Secondly, different statistics may regard the poverty conditions differently: if it is income that shows poverty or are there same living conditions included etc.

The same difference we can see in relation to the results of child poverty. Based on the current dataset, the average of child poverty is 24% of the population, but official statistics say that in 2015 it was 19,7%<sup>2</sup>.

Irrespective of the accuracy of the average poverty rate, we still could discover some correlations among our dataset. The visualization shows the correlation between poverty and income and between poverty and unemployment.

Correlations to occupations, private and public work has been regarded, but there was not any significant correlation shown.

The first graph shows a negative correlation between median income per capita and poverty rate. It means that the higher is the mean of income the lower is the poverty level in a state. From this scatter graph it is seen that the marks of District of Columbia on the one edge and Puerto Rico on the other edge are locating further from the others. This correlation is not surprising and although any causation cannot be made within these data, it seems quite logical result that the less you earn the more financial difficulties you may experience in your daily life.

The second graph shows the positive correlation between unemployment rate and poverty level. The higher is the unemployment rate in a state the bigger is probability that people are living under poverty level. Again, any causation cannot be brought out, but also these data seem showing a quite logical result. If you are unemployed there is quite high probability to experience financial difficulties.

These graphs also compare the level of the general poverty rate to child poverty rate. There are not any significant differences, only the percentages of children living under poverty level tends to be a bit bigger than general poverty level rates in almost every state. Unfortunately, it shows that there are more children living under poverty level than adults.

## **Self-employment in the different US states and some interesting correlations in 2015**

Visualization: [https://public.tableau.com/profile/milja#!/vizhome/USCensusDemographicData\\_7/Dashboard4](https://public.tableau.com/profile/milja#!/vizhome/USCensusDemographicData_7/Dashboard4)

Last insight gives an overview of the rate of self-employment and how it is correlating with other data points.

Visualization shows self-employment in the different US states and some correlations in 2015. Self-employment rate is expressed as the percentage of the total employment (16 and older who are employed). Average self-employment rate is 8% of total employment.

Self-employment rate is the highest in North-Dakota (15,4%), Montana (14,2%), South-Dakota (14,1%) and Nebraska (13,9%). Interestingly the lowest self-employment rate is in District of Columbia where the income is the highest in the US.

Trying to see if there are any interesting correlations between data points and self-employment, three interesting correlations came out. This is the reason to use scatter plots as they are one of the best graphs for visualizing correlations.

There is quite strong positive correlation between self-employment and unpaid family work. In some statistics family work is also regarded as part of self-employment so the strong correlation is an expected result.

---

<sup>2</sup> Source: Income and Poverty in the United States: 2015: <https://www.census.gov/library/publications/2016/demo/p60-256.html>

The middle scatter graph shows that unemployment and self-employment are negatively correlated. The higher is the unemployment rate the lower is self-employment activity in a state. For example, North Dakota with the lowest unemployment rate has the highest self-employment rate. Puerto Rico, that is in many cases exception to comparing to average data, can be also seen as an exception with the having very high unemployment rate, but average self-employment rate.

The upper scatter graph shows a quite strong positive correlation between working at home and self-employment. The four state that have the highest self-employment rates, have also the highest percentage of working at home. It is again not surprising correlation as self-employed people are working their own and very often they do not need special office space. They do their whole work or at least paper work at home office.

## **Conclusion**

It would be very interesting to compare this dataset with earlier or later periods to see changes in societies and to discover the trend lines.