Udacity Data Visualization Nano-Degree

Anca Filip
Vodafone

Project no.1 – Tableau visualization – 2015 US employment

1. The project can be found here:

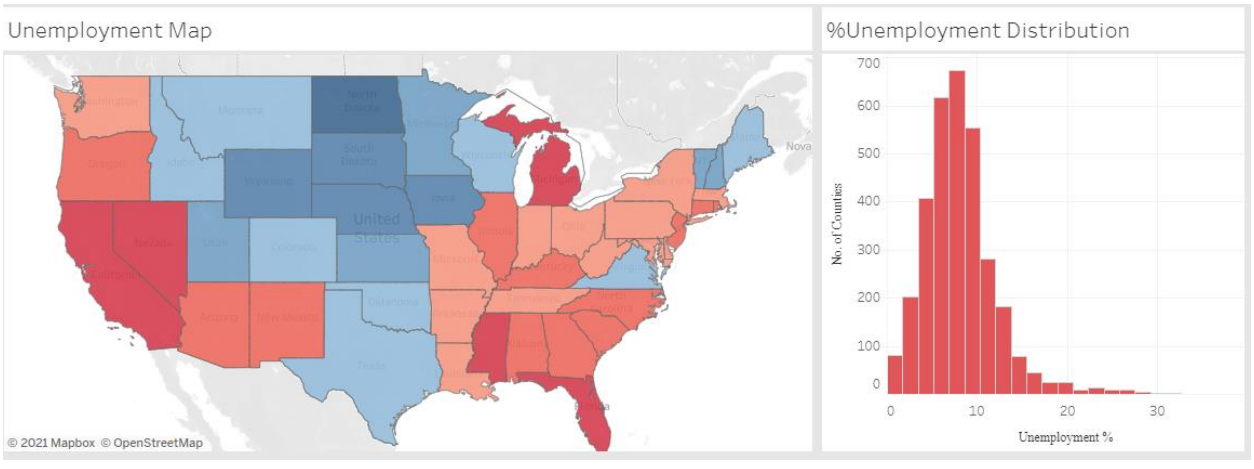
https://public.tableau.com/app/profile/anca3789/viz/AF_Udacity_Viz_Proj1_resub1/UnemploymentvsChildPovertybyCounty

2. **Summary**:

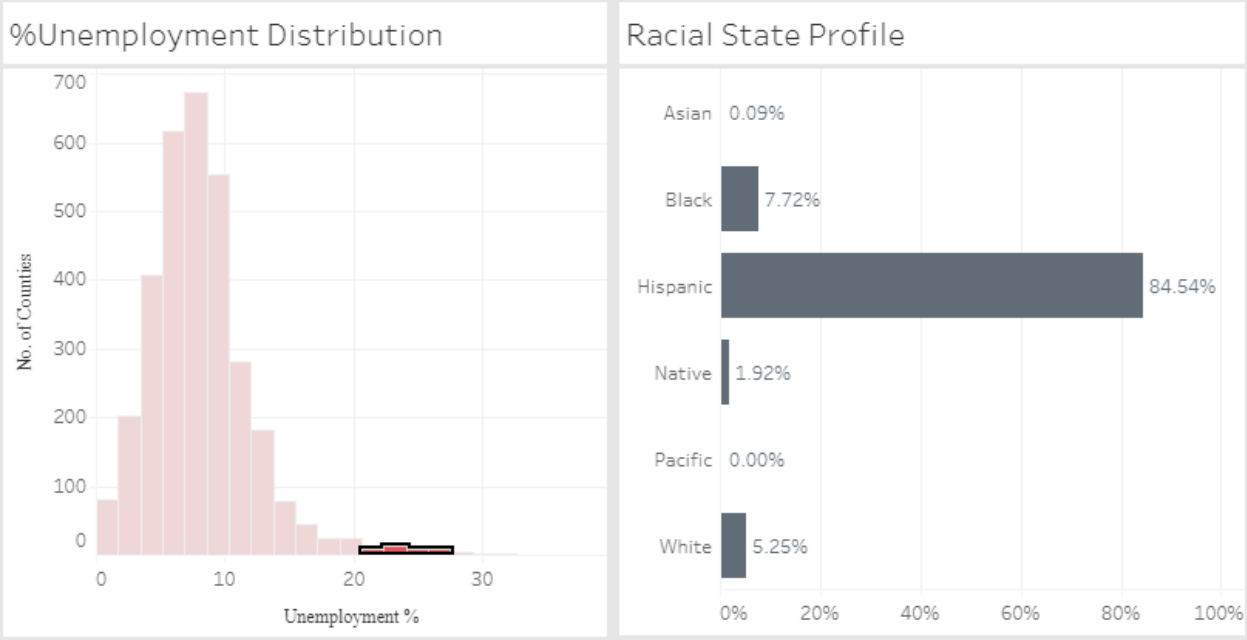
The dashboard's purpose is to show various dimensions of the employment status for US in 2015.
One of the key elements is the distribution of the unemployment rate, both in terms of value frequency and by state.
For the employed population I've included the view of industry, type of work and transportation mode.
I was also curious to see if the racial profile of the states is somehow correlated with the unemployment rate so I added the percentage of each race / state.

What I found out is that the **central and northern** parts of the US have a **lower unemployment** rate than the coastal states - which is obvious from the state heat map.

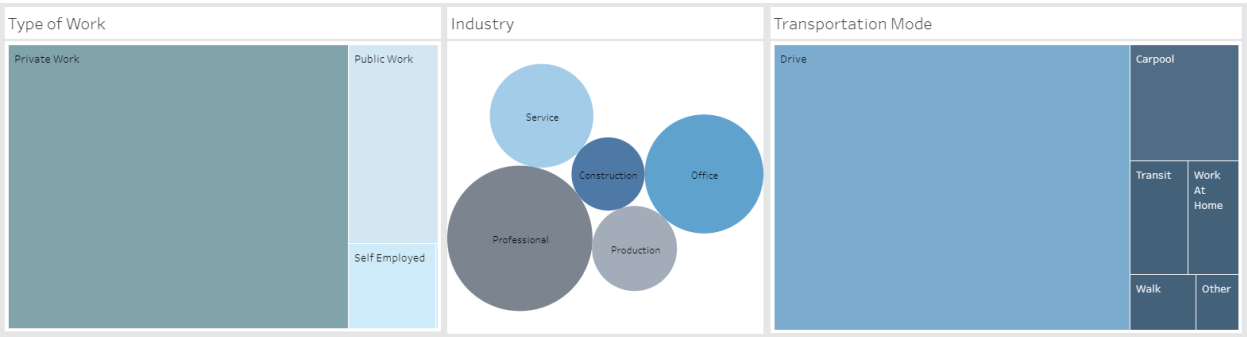
The **unemployment rate** seems to be almost normally distributed around a **median** of 7 % (almost **700 counties have a rate of ~7%**), with a slight **right skew** going **up to 30%** unemployment rate.



When I selected the **unemployment rates > than 20%** the racial composition of those specific counties showed a proportion of **85% of Hispanic** people.



In terms of the dimension of the employment, it appears that overall US, with some variations, the people are **Professionals (36%)** or **Office employees (24%),** working the the **Private sector (79%)** and they **Drive (77%)** to work.



3. **Design**:
   - 3.1. Dashboard:
     - **Rate of unemployment by state**: I decided to use a map to show the unemployment rate with a range of colors from red to blue: darker red showing the highest unemployment rate states and darker blue showing the lowest rates. The map was the obvious choice because it provides both efficient and effective display of more than 40 categories. Moreover, it's easy to use it as a filter clicking on each state, as opposed to scrolling through a drop down list – the map controls the entire dashboard. The colors seemed appropriate considering the fact that

unemployment rate is a negative aspect, therefore red would flag where the problems are.
To add extra depth to the information provided by the colors, the tooltip shows all population counts by race, total population and the unemployment rate.
Resub1: as pointed out by the reviewer I added the color legend for the chart.

- **Unemployment rate value distribution:** I used a histogram to display the unemployment rate by county: sometimes the central tendency measures (median or average) hide the overall picture. Even the states that are very red on the map have ~10% unemployment rate, so not very far from the median of 7%, however, when looking at the histogram we can see there are some counties with 20-30% unemployment rates. To keep consistency with the colors used for the map, the unemployment histogram bars are also red.
  The histogram was set to work as a filter because it's providing essential insight especially when we look at how the racial profile changes.

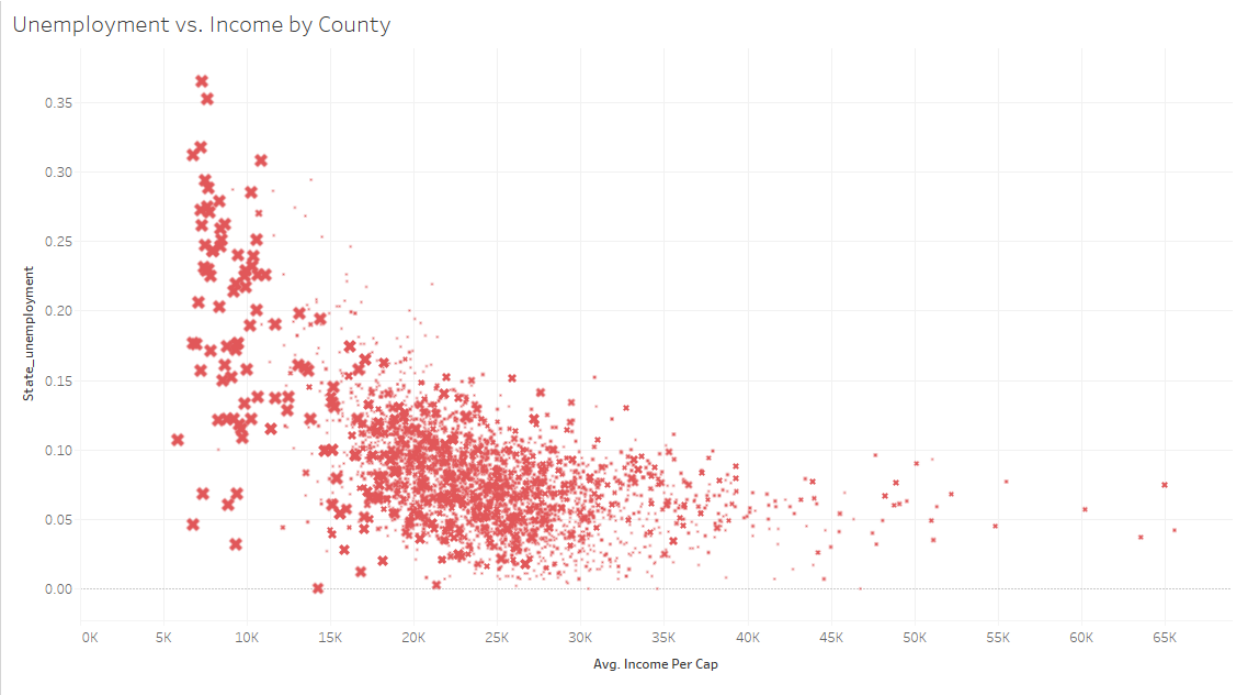
- **Racial State Profile**: I used a simple bar chart to show the percentage of each race – these percentages change when a state or specific unemployment range are selected in the map or the histogram, respectively. Since it not a measure of employment or unemployment I used a neutral grey.
- **Type of work, Industry and Transportation mode** – these measures show composition, so I used bubble and tree maps with blue hues to match the intuitive approach I started with in the map (unemployment = red, employment = blue). The main aim of these charts is to give the viewer a feel of how big/small each component is compared with the rest; however, since they are not very effective at conveying values or percentages, if you want to know such details, the tool tip will show both the absolute values and the percentage.
  These 3 charts are controlled by the map, so you can see for each state the status.

Resub1: As I noticed from the dashboard that the counties where the Hispanic population has a higher percentage also have a higher rate of unemployment, I decided to do a deep dive into the subject in order to create the 2 distinct visuals besides the dashboard.

3.2. Unemployment vs. Child Poverty by County and
3.3. Unemployment vs. Income by County

The most suitable chart to show the correlation between two numerical variables is a scatterplot; because I wanted to show not just the correlation between the unemployment and child poverty rates, but also the correlation of these with the Hispanic population percentage, in order to show the third measure I encoded it in the size of the scatter points.
I kept the color red for both of these charts as I intended the view as alarm signal, changing the symbol to differentiate the 2 views.

I kept the legend that explains the size of the symbols, however the main goal was to point out that indeed the counties with higher Hispanic % suffer on all standard of living points of view – unemployment, child poverty and income per capita – the bigger symbols are obviously clustered in the areas of the chart that show high unemployment / high child poverty, and also high unemployment / low income per capita.



4. **Resources**: US Flag Image: https://images.app.goo.gl/rpKjvHCDpDJyC38D9

Thank you!

Anca