

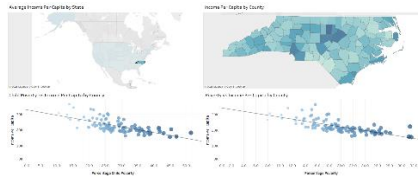
Build a Data Visualization Project

2015 US Census County Data Analysis

Insight 1:

What does Poverty look like for Income per Capita?

<https://public.tableau.com/app/profile/rodney.robinson/viz/IncomePerCapitaandPoverty-Dashboard/IncomePerCapitaandPoverty>



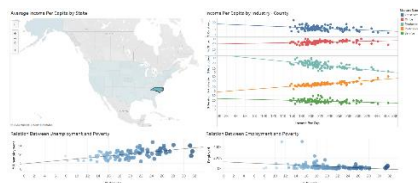
There is an inverse, or negative, correlation between income per capita and poverty. In both of the charts for 'Child Poverty' and 'Poverty' it is clear that there is a trend that shows as 'Income Per Capita' decreases the percentage of poverty increases.

(A fill map is used to allow the user to click on a state. When selected the user gets an enlarged view of the selected state with shading to depict the counties. The county's shading indicates its 'Income Per Capita' on a graduated scale – the darker the shading the greater the 'Income Per Capita'. Details can be obtained from a tooltip when hovering the mouse over a county of interest. Scatterplots are used to visualize the relationship between 2 variables. The bubble size represents the size of the percentage with larger percentages represented by larger bubbles. Details can be obtained from a tooltip when hovering the mouse over a bubble of interest.)

Insight 2:

What is the distribution of Industry and how does it relate to Income per Capita. What is the relationship between Employment and Poverty?

<https://public.tableau.com/app/profile/rodney.robinson/viz/IncomePerCapitaandEmployment-Dashboard/IncomePerCapitaandEmployment>



The data analysis shows a consistent trend for the relationship between 'Income Per Capita' and 'Industry'. It appears that for the 'Office' and 'Professional' industries there is a positive correlation between the percentage of individuals working in those industries and the income per capita for those industries. For instance, in North Carolina, Robeson county has an 'Income Per Capita' of \$16K in the 'Professional' industry with 25.1% of the workforce in that industry. Orange County has 55.1% of its workforce in the 'Professional' industry with an 'Income Per Capita' of \$36K. Increasing percentages of the workers have increasing 'Income Per Capita'. In contrast, the industries of 'Construction', 'Production' and 'Service' tend to have negative correlations between 'Income Per Capita' and

'Industry'. There is a trend showing decreasing percentages of workers in these industries have higher 'Income Per Capita'. The trends in these correlations suggest that certain industries provide better earnings potential to a greater percentage of the population than other industries.

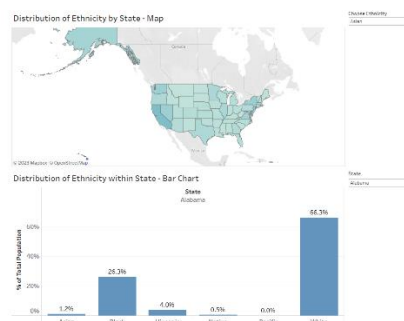
There is a positive correlation between unemployment and poverty. As 'Percent Unemployment' increases so does 'Percent Poverty'. And in contrast, there is an inverse correlation between employment and poverty. As the count of 'Employed' decreases the 'Percent Poverty' increases.

(A fill map is used to allow the user to click on a state. When selected the user gets a view of 3 scatterplots with data related to the selected state. Scatterplots are used to visualize the relationship between 2 variables. The stacked scatterplot allows the user to view data from 5 different categories of 'Industry' in one visual facilitating easier comparison of the industries. The bubble size represents the size of the percentage with larger percentages represented by larger bubbles. Details can be obtained from a tooltip when hovering the mouse over a bubble of interest.)

Insight 3:

What is the Ethnic distribution by State?

<https://public.tableau.com/app/profile/rodney.robinson/viz/DistributionofEthnicitybyState-Dashboard/DistributionofEthnicitybyState>



The 'Distribution of Ethnicity by State' dashboard provides valuable insights. For instance, the map shows the states with the highest percentage of a particular ethnicity. It can be seen that the states with the highest percentages of 'Black' are in the Southeastern states from Texas, Louisiana, and Mississippi eastward through South Carolina and up the east coast to Maryland. There are high percentages of 'Hispanic' in the Southwest, Florida, and Puerto Rico. It appears the highest percentages of 'Native' are in Alaska and the Midwest which maps closely to the locations of the largest "U.S. Domestic Sovereign Nations: Land Areas of Federally-Recognized Tribes" found on the map at <https://biamaps.doi.gov/indianlands/>.

The bar chart provides the opportunity to see 'Ethnicity' percentages broken down within each state. From this visual it can be seen that there are 4 states with percentages of 'White' greater than 92% (New Hampshire 91.4%, West Virginia 92.5%, Vermont 93.6%, and Maine 93.9%). The percentage of 'Black' in District of Columbia is 48%, which is the largest ethnicity in District of Columbia. Puerto Rico has a 'Hispanic' percentage of 99%.

(A fill map is used to highlight the states and the percentage ethnicity on a gradient shading scale. The user can filter through different ethnicity choices to see a graduated distribution of the selected ethnicity on the state level. An independent bar chart allows the user to filter through different states viewing a breakdown of percentages of each ethnicity in the selected state.)

Insight 4:

What is the Ethnic distribution by County?

<https://public.tableau.com/app/profile/rodney.robinson/viz/DistributionofEthnicitybyCounty/DistributionofEthnicitybyCounty>



This visualization provides a more granular analysis of ethnic distribution at the county level. As with the 'Distribution of Ethnicity by State' map certain trends stand out. For instance, with the nationwide view showing all of the counties throughout the country the distribution of 'Black' is evident in the Southeast. When filtering on 'Hispanic' at the nationwide view the distribution is Puerto Rico, South Florida, and from the southernmost tip of Texas up through the southwest and eastward to California and northwestward toward the Pacific Northwest. And again, as with the 'Distribution of Ethnicity by State' map 'Native' distribution tracks closely to federally-recognized tribes (see note and link above).

This visualization also provides the opportunity to filter at the state-view level providing a more granular breakdown with the ability to cycle through the different ethnicities for a more insightful view of distribution. For instance, when filtering on North Carolina much insight can be gained into the percentage distribution of 'Black'. I will elaborate more on this in the addendum that follows.

(A fill map is used to visualize percentage ethnicity on the county level. This very granular map allows the user to filter the view geographically – 'All' for nationwide view of counties or by choosing a state in the 'State' filter for a state-level view of counties. The user can then filter either geographical view by ethnicity.)

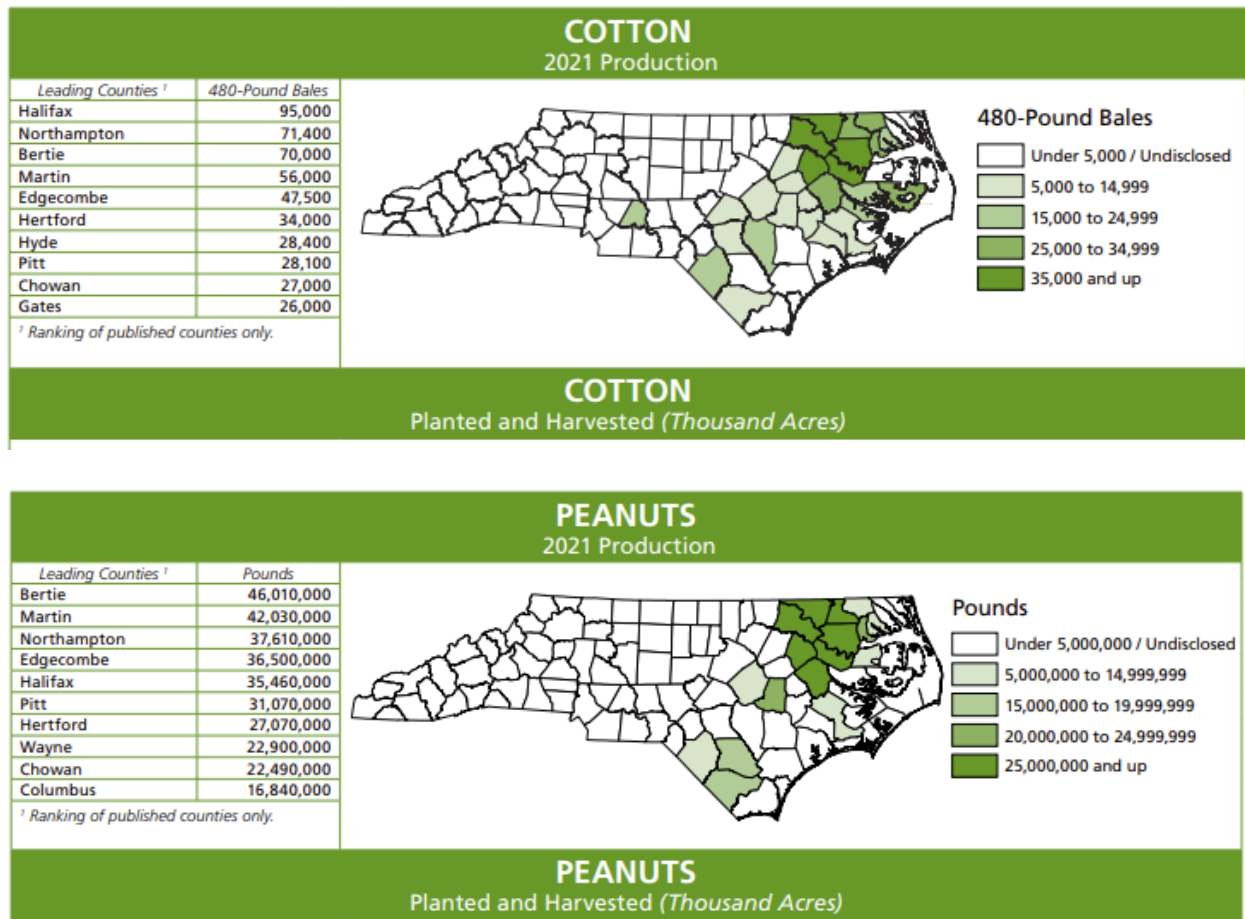
Addendum

Additional Insight:

What additional insight might be gained that could provide historical context to ethnic distribution throughout the country or a particular state?

When viewing the 'Distribution of Ethnicity by County' map and filtering on the state of 'North Carolina' and the ethnicity of 'Black' I was curious whether or not there were any historical contexts that could explain the high distribution of 'Black' in certain counties, especially when a large proportion of those 'high-distribution' counties were more rural. My curiosity led me to seek additional data. I first searched for North Carolina agricultural statistics and found a .pdf document entitled "2022 North Carolina Agricultural Statistics" at

https://www.nass.usda.gov/Statistics_by_State/North_Carolina/Publications/Annual_Statistical_Bulletin/AgStat/NCaGStatBook.pdf. This document provided two very insightful maps for certain crops. I have included them below:



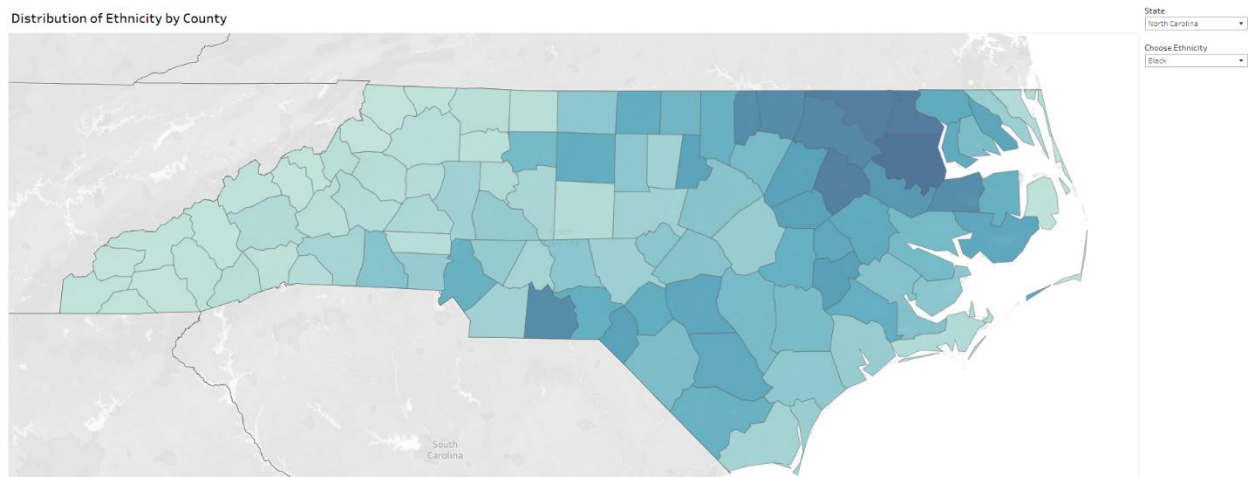
These 2 maps show counties with the highest production for these particular crops. These maps track very well with the 'Distribution of Ethnicity by County' map for North Carolina filtered for 'Black' ethnicity. There is a strong correlation, but can any insight be gained within a historical context? I

decided to look further for any historical context. Since these are historically ‘cash crops’ I looked for geographic cross references and found a map that displayed ‘Percent of population enslaved, 1860’ at <https://www.ncpedia.org/media/map/north-carolina-enslaved>, which I have included below:

North Carolina enslaved population by county, 1860



Comparing this map to the ‘Distribution of Ethnicity by County’ map (from my “2015 US Census County Data Analysis”) for North Carolina filtered for ‘Black’ ethnicity reveals a striking similarity (I have included a snapshot below for a quick comparison):



Does this suggest that the roots of this ‘Black’ distribution originated in slavery and has persisted for centuries? If so, why? What socio-economic, cultural or political factors, if any, have contributed to this pattern? These are obviously questions that can only be sufficiently addressed through more rigorous analysis of additional data.

Final Observations:

What insights cannot be gained from the “2015 US Census County Dataset”?

This data set in its current form presents challenges for certain data analyses. There are separate measures (counts) for ‘TotalPop’, ‘Men’, and ‘Women’. However, many of the other measures are percentages such as for ethnicity, poverty, industry, mode of transportation, unemployment. It is simple enough to execute ‘reverse’ math to convert those percentages into counts. This allows, for instance, to aggregate counts at the county level if necessary. Those counts can then be used to aggregate at the state level, if needed. But breaking down gender within different ethnicities or industries cannot be accomplished. Also, there are no age measures therefore preventing any analysis of age distribution by gender, ethnicity, etc. Furthermore, major industries have been omitted from the dataset – agriculture, healthcare, technology. And certain Industries, such as ‘Professional’, may have been broken down into more granular levels. These missing measures could have enabled a more aggressive and granular mining of the data to possibly reveal more and deeper insights.

**Personal note: This was a very interesting project. I learned a great deal and feel as though I have gained valuable skills that will help me continue to grow as a business/data analyst.*