Information Visualization

-HW1

Members and Team Contribution:

SONG TANG-SOT13

JIALE SUN-JIS96

QIUYANG XU-QIX21

	task a)	task b)	task c)	data clean	report
song tang			√	√	√
jiale sun		√		√	√
qiuyang xu	√			√	√

Goals

Generate visualizations to compare the poverty data by states, by races, and by time.

Data Pre-processing

We did several steps to aggregate row data. First, there are two records of 2013. According to footnotes, the methodology changed and they split-tested that year. So we decided to use (18) to represent 2013. And then we set year range as 1990 to 2015, and select Year, Number, Percent and Race as attributes of visualization. Finally, we merge row data to five races, by following rules:

White: White Alone and White

White not Hispanic: White Alone not Hispanic and White not Hispanic

Black: Black Alone or in Combination and Black

Asian: Asian Alone or in Combination and Asian & Pacific Islander

Hispanic: Hispanic (any race)

There was a category change in 2002. Although categories aren't exactly the same and different categories seems to have overlap. For instance, Asian Alone or in Combination contains Asian Alone. Our mapping strategy is choosing categories that cover the most in each race.

Application

1. Use dataset D1. To compare data across races and over time, generate a time-series plot to show how the poverty by race change from 1990 to 2015. You need to show at least three time points in the plot: 1995, 2005, and 2015. Omit race categories with "NA" values.

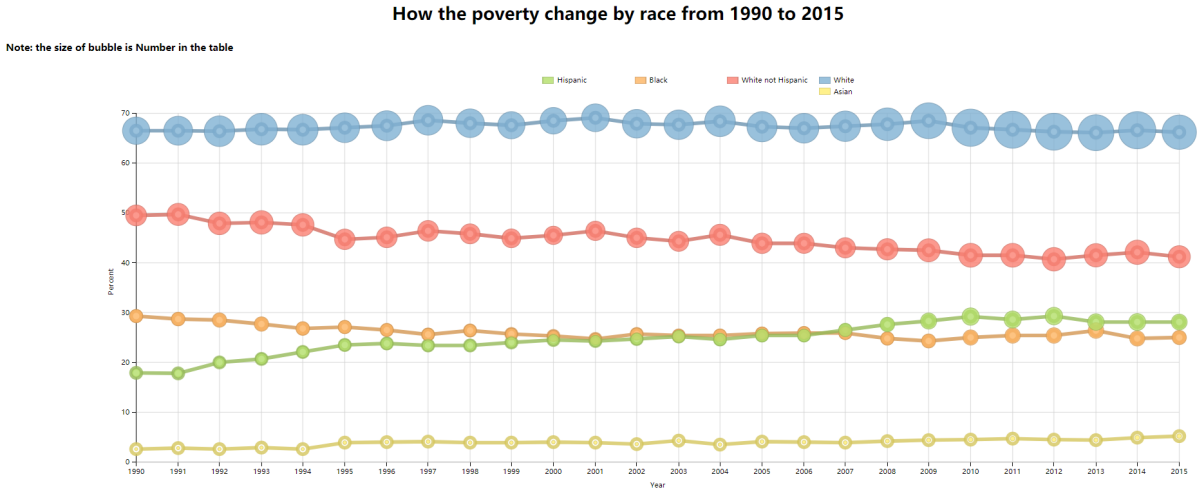


Figure 1

1.1 what you intend to show in the visualization

For dataset, we select Year, Number, Percent and Race.

This vis shows trends of five races in 1990 to 2015. Each node contains race, year, poverty number and rate. Vis maps the time to X, percent to Y. . The size means the population in poverty.

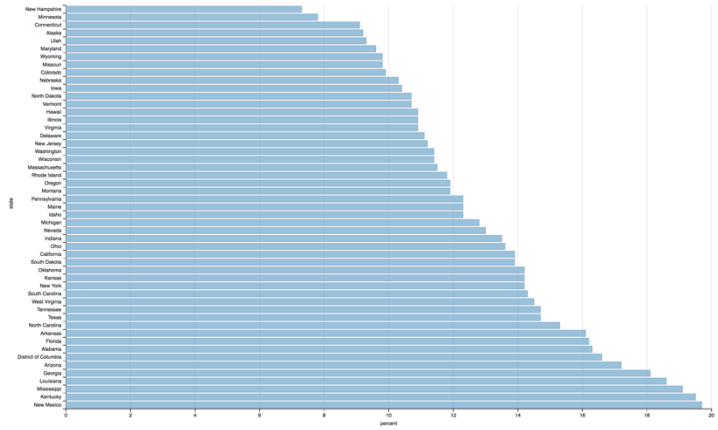
1.2 the rationale for the design (why is it an effective representation for the things you intend to show?)

Lines in colors indicate different races, and can clearly shows changing during a quarter of a century. Visualization also dynamically shows the number change over time, using areas.

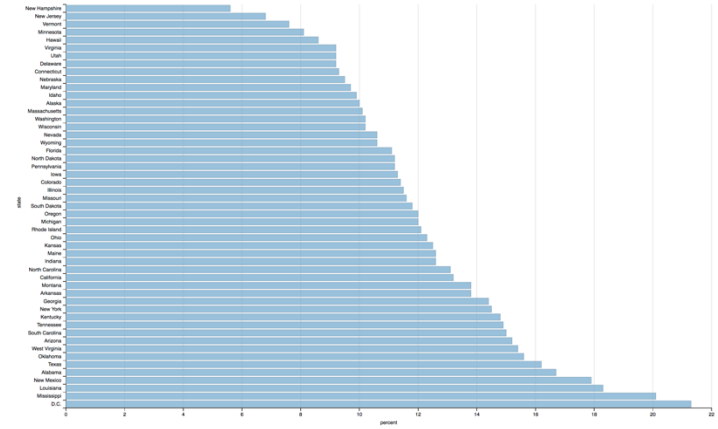
In this way, vis clearly separate categories. And also corresponding the time to the rate fairly intuitive.

2. Use dataset D2. To compare data across states and over time, generate three bar charts corresponding to poverty of the states in 1995, 2005, and 2015 (one chart per year), where each bar represents a state.

2015 state and poverty bar chart



2005 state and poverty bar chart



1995 state and poverty bar chart

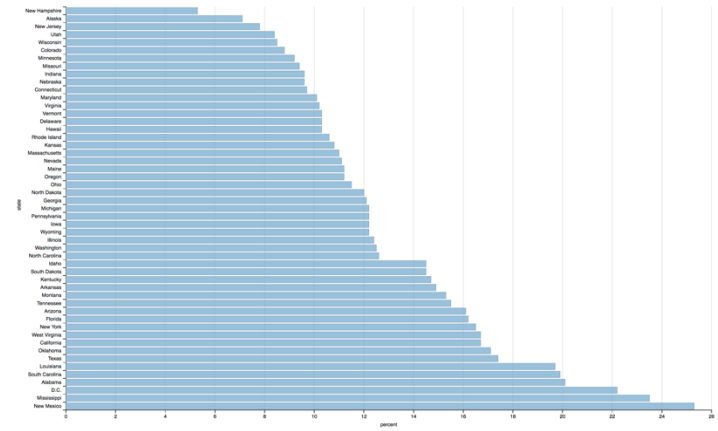


Figure 2

2.1 what you intend to show in the visualization

We want to show the distribution of poverty percent by state in three years: 2015, 2000, and 1995. Each bar in barchart present the percent and ranked in order from low to high poverty situation.

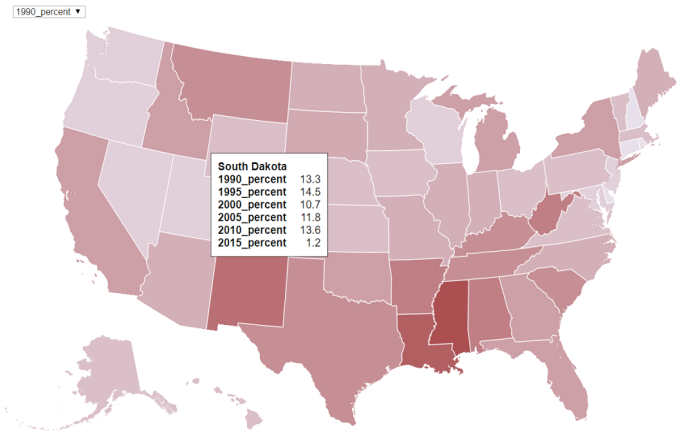
2.2 the rationale for the design (why is it an effective representation for the things you intend to show?)

Bar chart can illustrate the quantitative comparison of each variables. Though those three years are distributed separately, the plot shows the distribution of each state's poverty percent. Meanwhile, the bar chart is read horizontally for easier read by state.

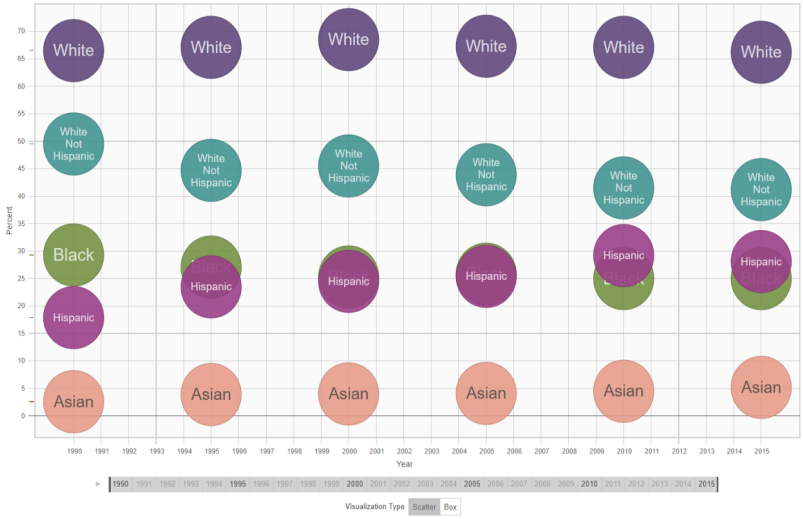
3.Use datasets D1 and D2. Use your creativity to design and generate a composite visualization to show how the data change between 1990 and 2015, by state and by race. The visualization can include multiple charts or mix different types of charts.

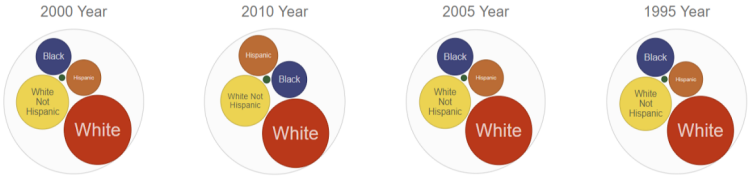
How the poverty change from 1990 to 2015 by states

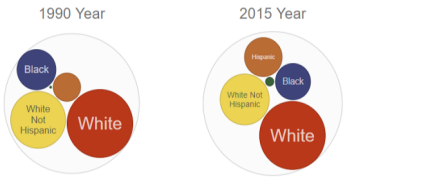
note: the lighter color, the lower percent



How the poverty change from 1990 to 2015 by Race







note: Above chart is for each year, every race density distribution

Figure 3

3.1 what you intend to show in the visualization

To show how the data change between 1990 and 2015 by state and by race, the visualization includes multiple charts or mix different types of charts. We divide it into three parts:

In first part, we use national geomap to show the poverty distribution by each state in U.S. The color degree of each state represent its poverty percent at certain year. We select six years: 1990, 1995, 2000, 2005, 2010, and 2015 in time period to compare the change of poverty percent, with each year's situation can show separately in dropdown list.

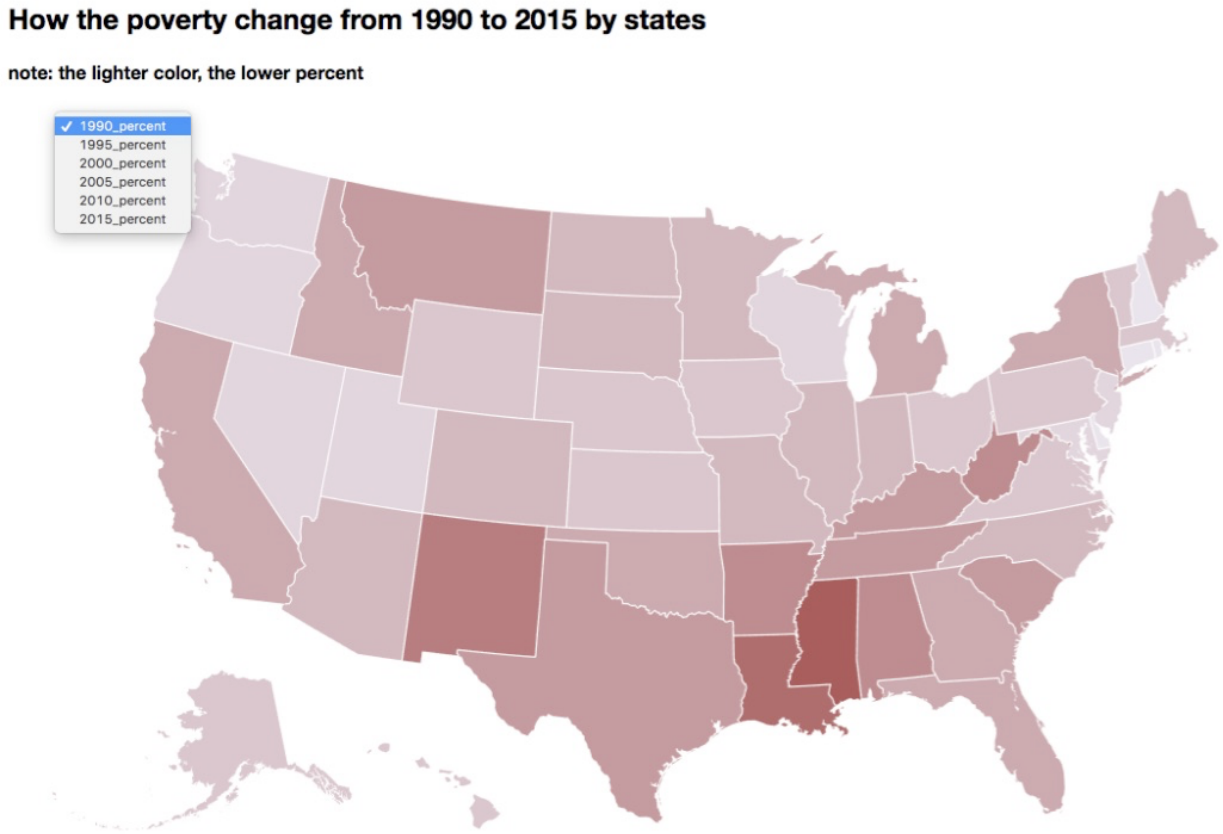


Figure 4

In second part, we use scatter plot and box plot to illustrate the poverty distribution by race. The diagram type can be chosen under the year bar.

In scatter plot, the x-axis represents the year and y-axis represents the poverty percent, then the race type is defined by each color. The position of each node clearly demonstrates the poverty percent of each race by year. In box plot, the x-axis is by year and y-axis is by poverty percent as well. It gives an overview of the change in each year. Each box provides more details in statistics: the interquartile range is based on the total data range that year.

How the poverty change from 1990 to 2015 by Race

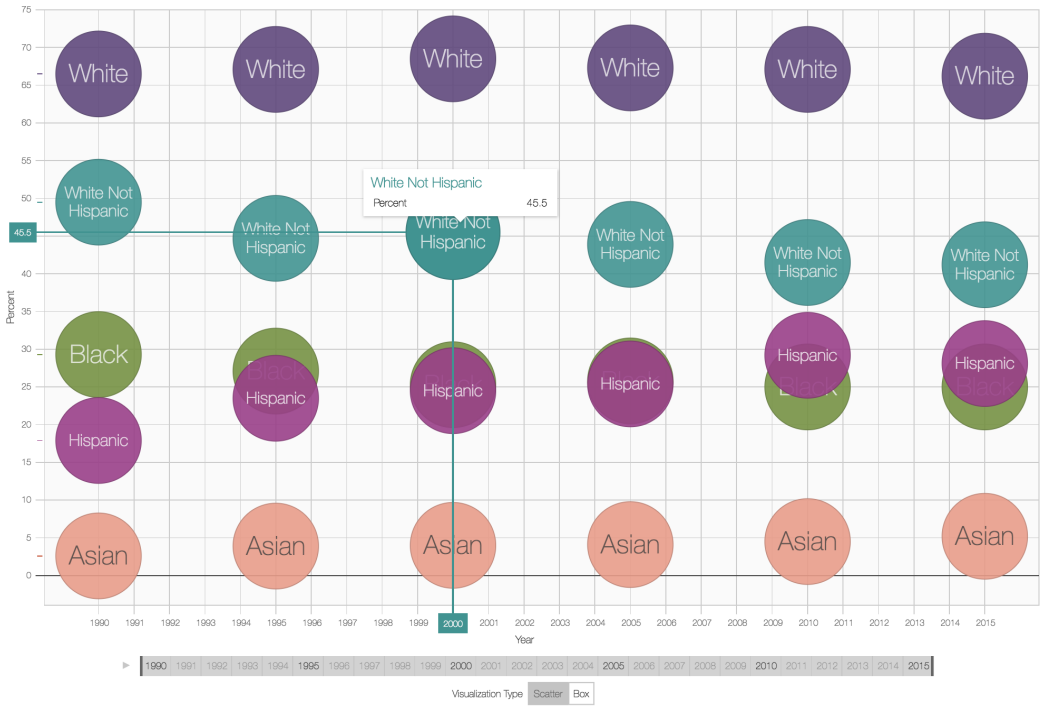


Figure 5

How the poverty change from 1990 to 2015 by Race

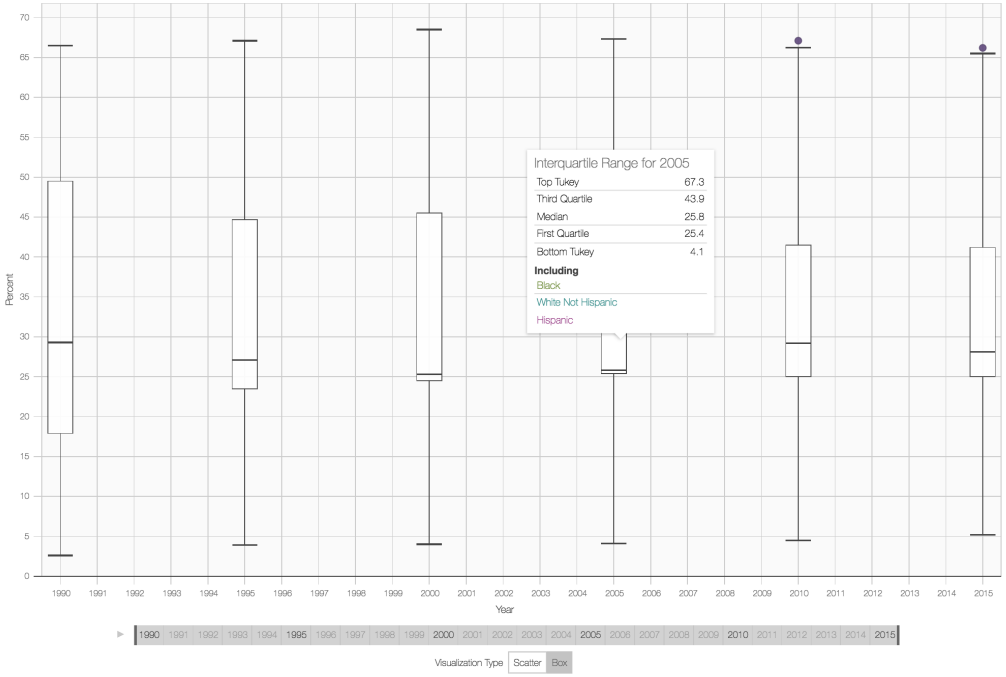


Figure 6

In third part, the bubble chart compare the poverty distribution among five races in every year, which helps to better understand the relationship among each race's poverty degree. The size of each bubble is the poverty population percent and the color distinguishes each race.

Figure 7

3.2 the rationale for the design (why is it an effective representation for the things you intend to show?)

First, the geomap visually shows the distribution of each state. Integrated with the color, the poverty degree of each state is clearly represented: the darker the color is, the poor the state is. When changing year in dropdown list, the map is updated with this year's poverty percent.

Second, in the scatter plot, each node shows this race people's poverty percent at that year. By horizontal comparison, the plot shows the poverty changes during years; By vertical comparison, the plot show the comparison of each race. In box plot, it gives an overview statistics in total poverty status comparison for further understanding.

Finally, the races in bubble chart are defined as five color, the larger the bubble size is, the poor the race is. Five races in each year are collected into one group, which gives more clearly visual for comparison by race.

Reference:

http://dimplejs.org/

https://github.com/topojson/topojson