

# Homework 1

## Task Description

1. Read the description about Educational Attainment at [Census.gov](https://www.census.gov).
2. Generate three visualizations to compare the data by states, by categories, and by time.
  - a) To compare data across categories and over time, generate a time-series plot to show how the (overall US) educational attainment of the three categories change from 1990 to 2009. You need to show at least three time points in the plot: 1990, 2000, and 2009.
  - b) To compare data across states and categories, generate three bar charts corresponding to educational attainment of the three categories in 2009, where each bar represents a state.
  - c) Use your creativity to design and generate a visualization to show how the data change between 2000 and 2009, across states and categories. The visualization can include multiple charts or mix different types of charts.
3. All visualizations need to be properly annotated and labeled to help clearly explain what the visualizations show.
4. Explain in your report:
  - what you intend to show in the visualization
  - the rationale for each design (why is it an effective representation for the things you intend to show?)
  - in the beginning of the report, describe your team contribution, including (a) names and PITT IDs of all members, and (b) one paragraph to briefly describe the contribution of each member

## Team Member and contribution

### Team Member:

Jingru Zhang (JIZ122)  
Yuxian Gong (YUG28)  
Ziyi Lan (ZIL21)  
Varanasi Anand (SKV9)

### Team Contribution:

We discuss and decide together the designs of all three visualization tasks, and then distribute the implementation of designs to team members:

Yuxian Gong focus on task a),  
Varanasi Anand focus on task b)  
Jingru Zhang and Ziyi Lan focus on task c)

## Visualization Design

### Task a)

In the visualization of task a), we intend to show how the overall educational attainment of the three categories change with time: Does each one of them increase or decrease from 1990 to 2009? How do the three categories compare with each other regarding the changing trend?

Time here is discrete, point-based, linear and ordered, so we decide to map it to 1D. The data in this task are quantitative and univariate. Considering the characteristics of both time and data, we choose line chart to plot educational attainment for each category from 1990 to 2009.

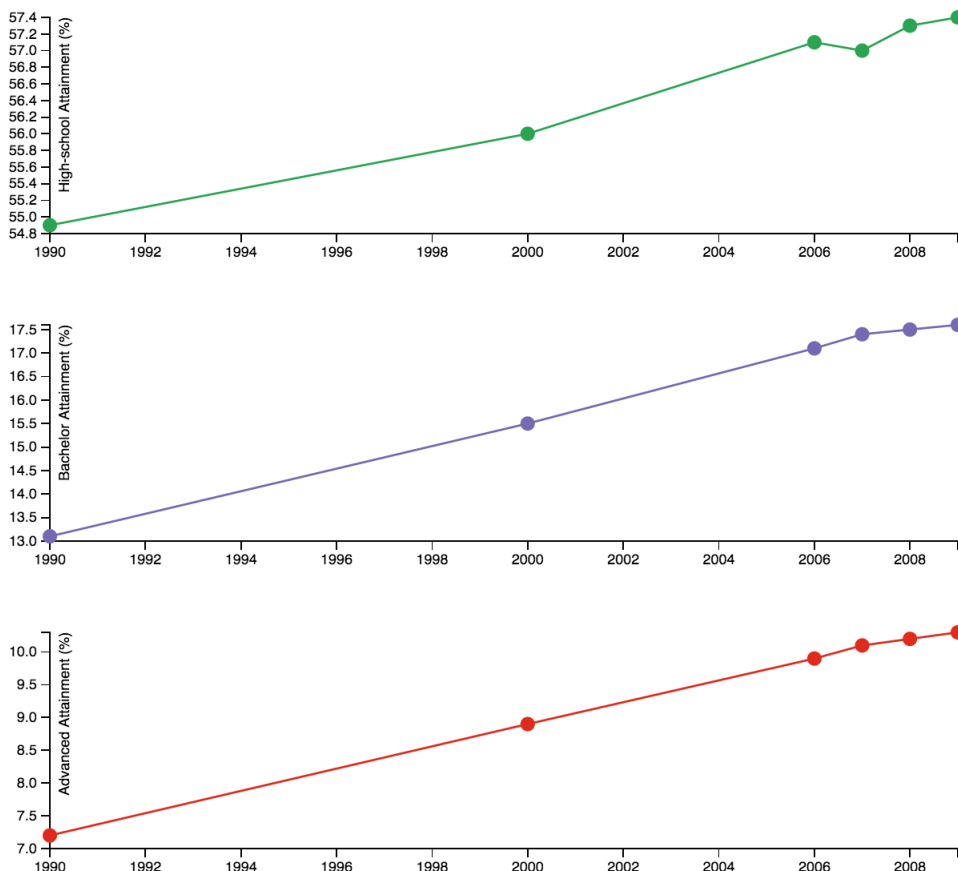
To better show the data and how they change with time, we plot two line charts: one plots each category in a separate chart, in which we can see the change trend of each category more clearly: the other plots all three categories in the same chart to better compare between categories.

#### 1. Line charts plot each category separately

##### Question 1

**Time-series plot to show how the (overall US) educational attainment of the three categories change from 1990 to 2009.**

**Line Chart 1 [ Separated ]**

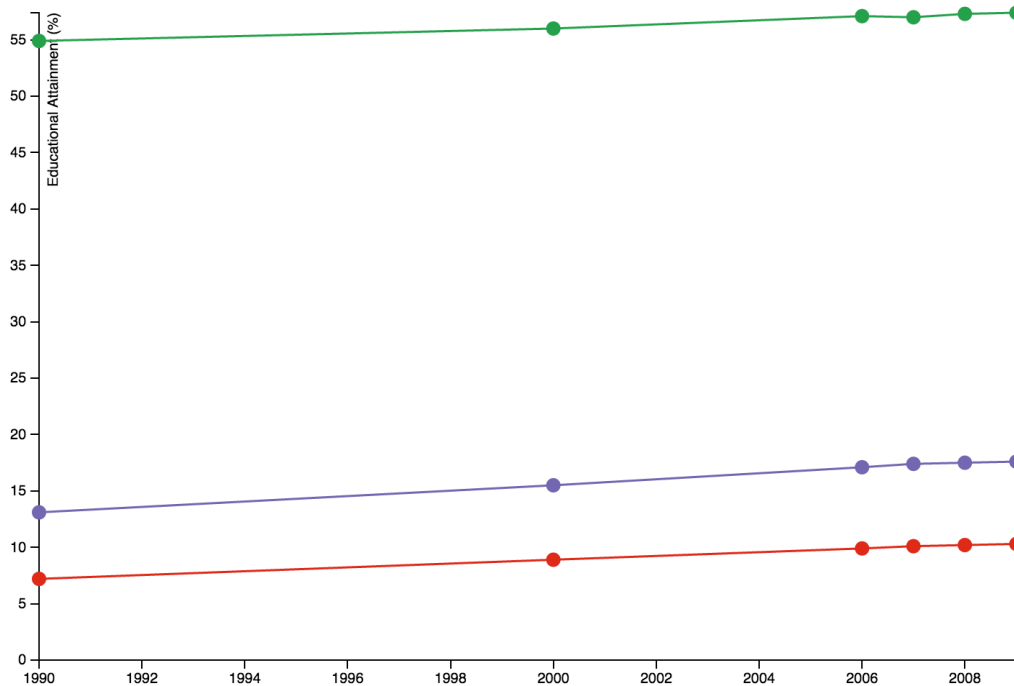


## 2. Line chart for all three categories together

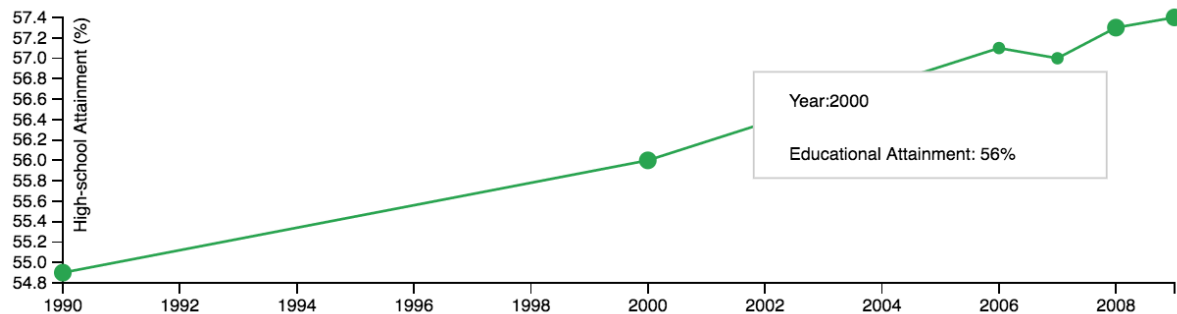
### Task a

Time-series plot to show how the (overall US) educational attainment of the three categories change from 1990 to 2009.

Line Chart 2 [ Together]



We use the “mouseover” and “mouseout” function to show detailed information on each point:



And once you use this “mouseover” to see detailed information and then move away, the point that you have just checked will become smaller. See the points for 2006 and 2007 is smaller then other ones on the above screenshot. In this way, you can know which points are checked and which not.

In both line charts:

- Green lines represent High School education attainment.
- Purple lines represent Bachelor's degree attainment.
- Red lines represent Advanced degree attainment.

### Pros:

- In line chart 1, you can see clearly how the three categories change with time; while you can compare them in line chart 2.

- You can check data for each point to get more detailed information and better understanding of the chart.

### Limitation:

The data in each category are very close to each other, so the three lines in line chart 2 seem like flat lines, which makes it difficult to tell how they change with time. So we plot separated line charts for each category, but the y axis does not begin from 0, which may be misleading.

### Code Source:

<http://bl.ocks.org/mbostock/3883245>

## Task b)

To compare data across states and categories, generate three bar charts corresponding to educational attainment of the three categories in 2009, where each bar represents a state.

Note: Data of year 2009 only is taken for these below charts.

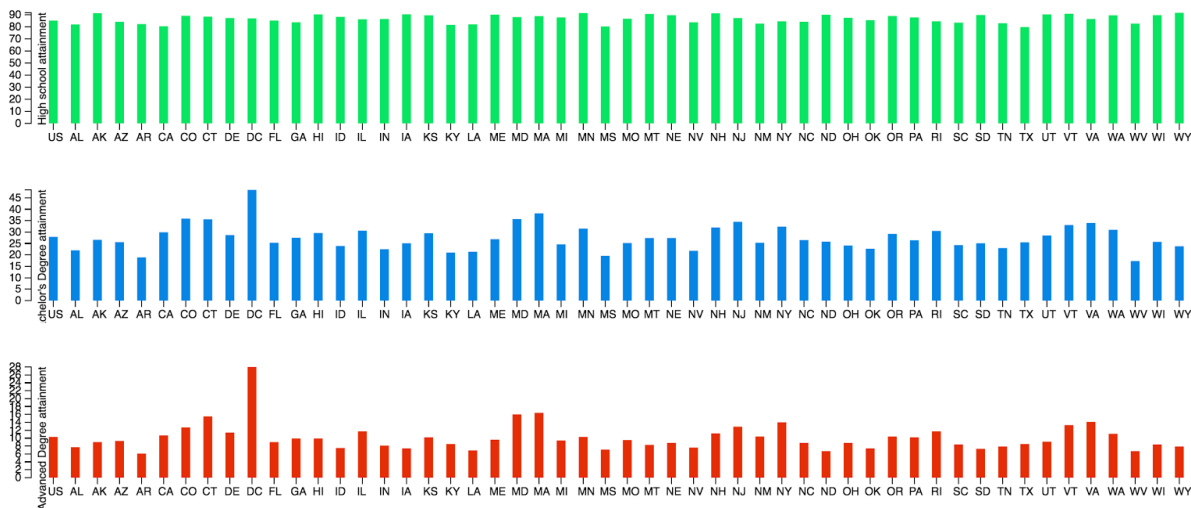
In this assignment we used 3 kinds of barcharts:

### 1. Simple Barchart

#### Question 2

Comparing data across states and categories.

Bar Chart 1 [ Simple Bar Chart ]



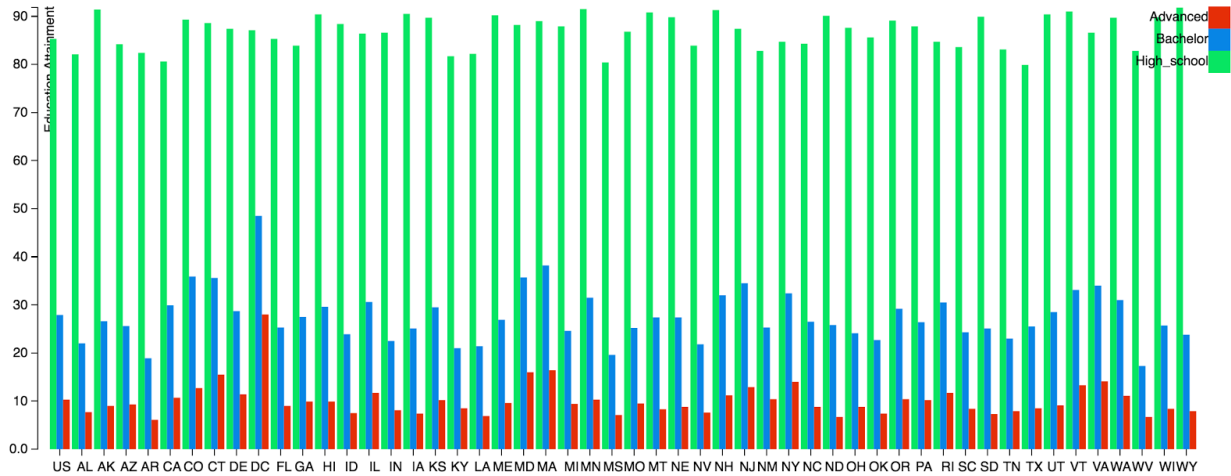
[ Source : <http://bl.ocks.org/mbostock/3885304> ]

### 2. Grouped Barchart

## Question 2

Comparing data across states and categories.

Bar Chart 2 [ Grouped Bar Chart ]



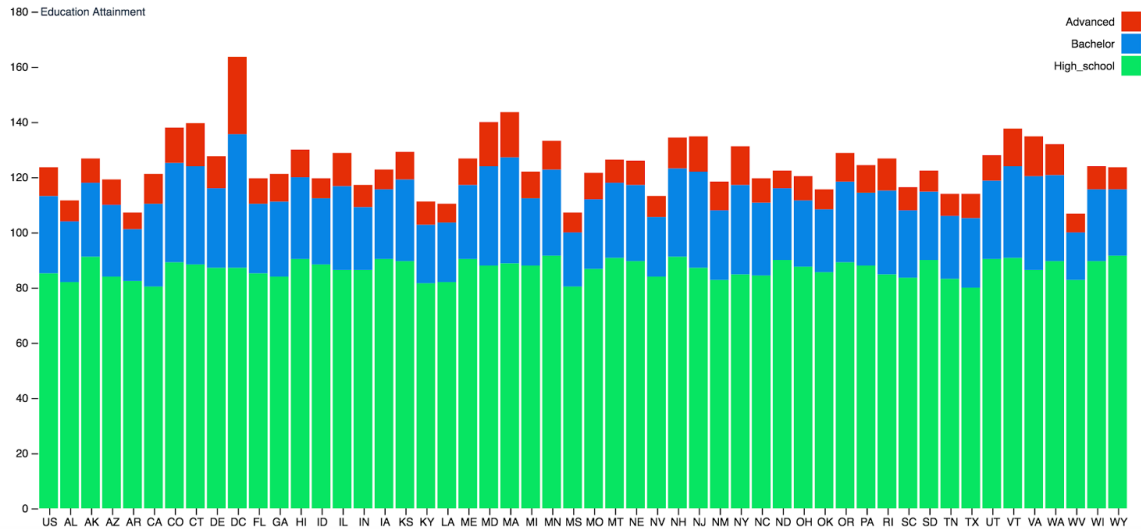
[ Source : <http://bl.ocks.org/mbostock/3885304> ]

## 3. Stacked Barchart

### Question 2

Comparing data across states and categories.

Bar Chart 3 [ Stacked Bar Chart ]



[ Source : <http://bl.ocks.org/mbostock/3885304> ]

In all the three barcharts :

- Green color represents High School education attainment.
- Blue color represents Bachelor's degree attainment.
- Red color represents Advanced degree attainment

In Bar chart 1, We can see 3 separate simple bar charts representing each education attainment category i.e., High School education attainment (Green), Bachelor's degree attainment (Blue), Advanced degree attainment (Red).

In Bar chart 2, all the mentioned 3 categories are group for each state.

In Bar chart 3, all the mentioned 3 categories are stacked together for each state.

## Pro :

Bar chart 1 is conveys the 3 categories education attainment rates for each state in a very simple way but Bar chart 2 is even more effective as we can compare attainment rates between the 3 categories. Bar chart 3 is very effective in getting to know the over attainment rates in each state. In Bar chart 3 we can see the overall attainment rate is high in DC ( District of Columbia ) and least attainment rate is in WY ( Wyoming )

## Limitations:

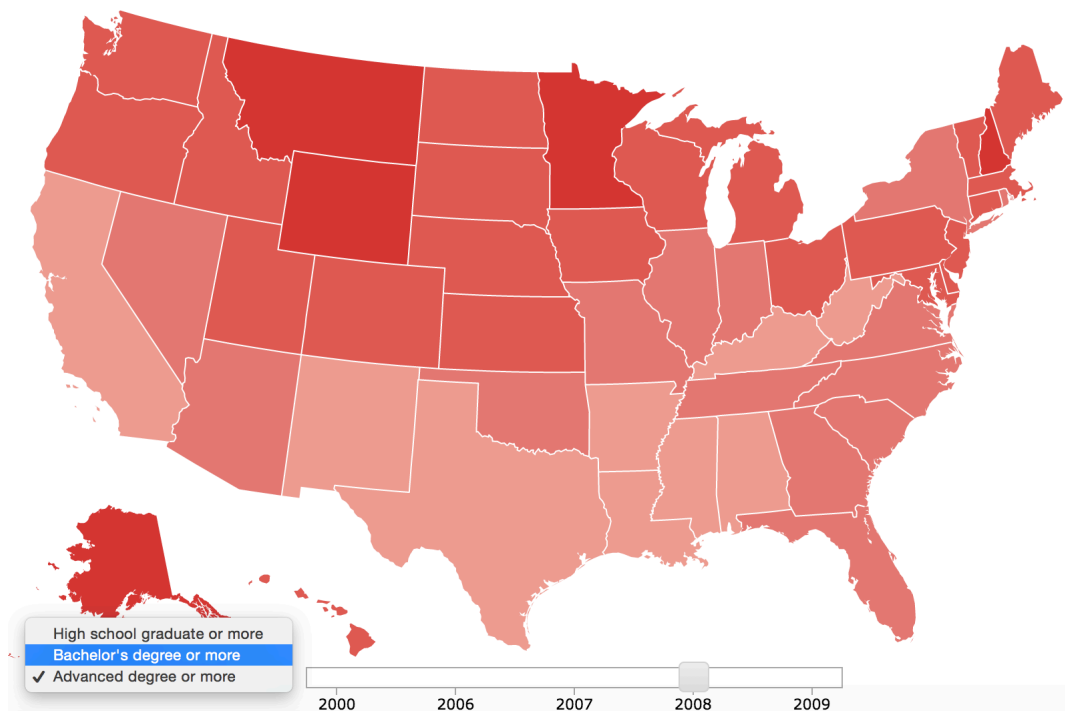
All the 3 barcharts are not interactive. This is in the sense, we didn't include functionalities like seeing the attainment rates upon hovering on each bar.

## Task c)

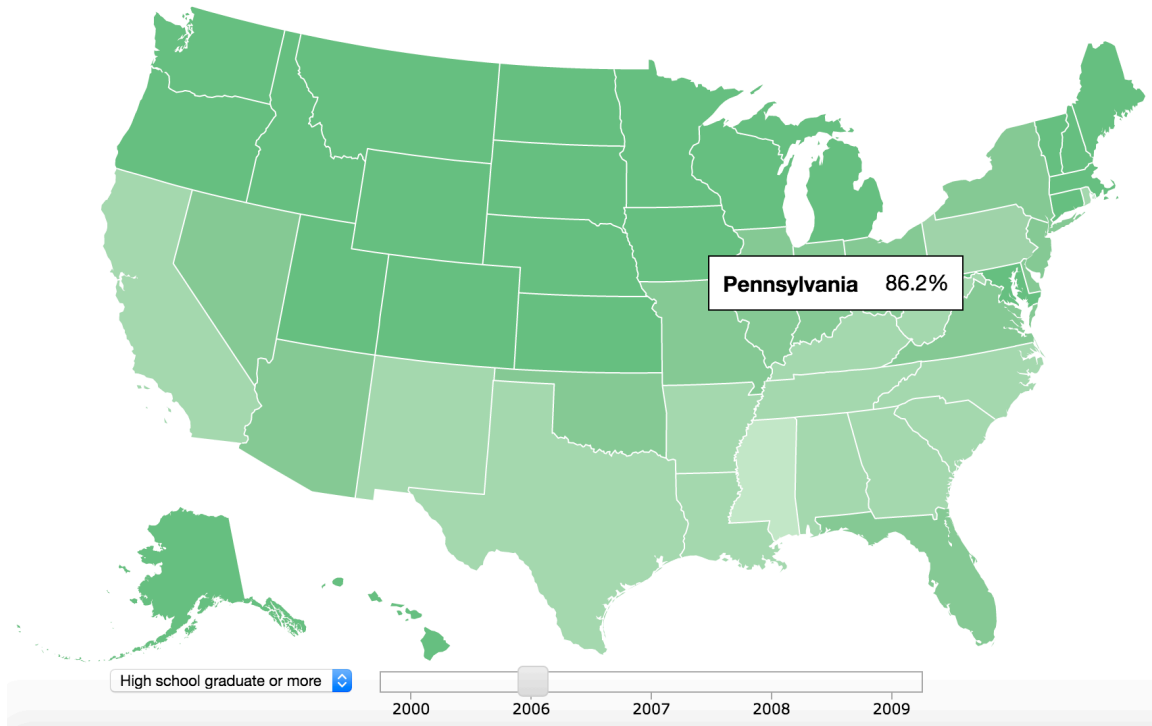
Design and generate a visualization to show how the data change between 2000 and 2009, across states and categories. The visualization can include multiple charts or mix different types of charts.

Data include geographic information, so we decide to use a map to show this information. Time is presented in slider. Category can be selected in a drop-down list ("High school graduate or more", "Bachelor's degree or more" and "Advanced degree or more"). Once you select a time point and a category, the map will show educational attainments (by intensity of color) in all states.

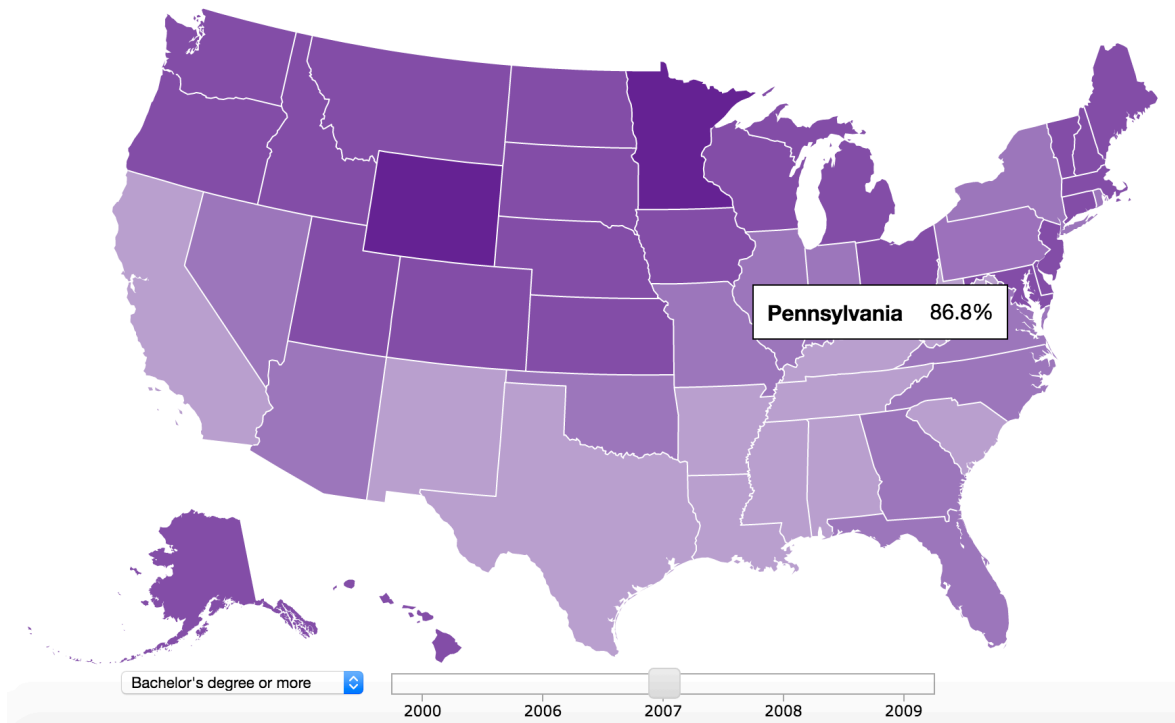
When mouse points to the location of specific state, the tip will show the name and the attainment rate of the chose year and category.



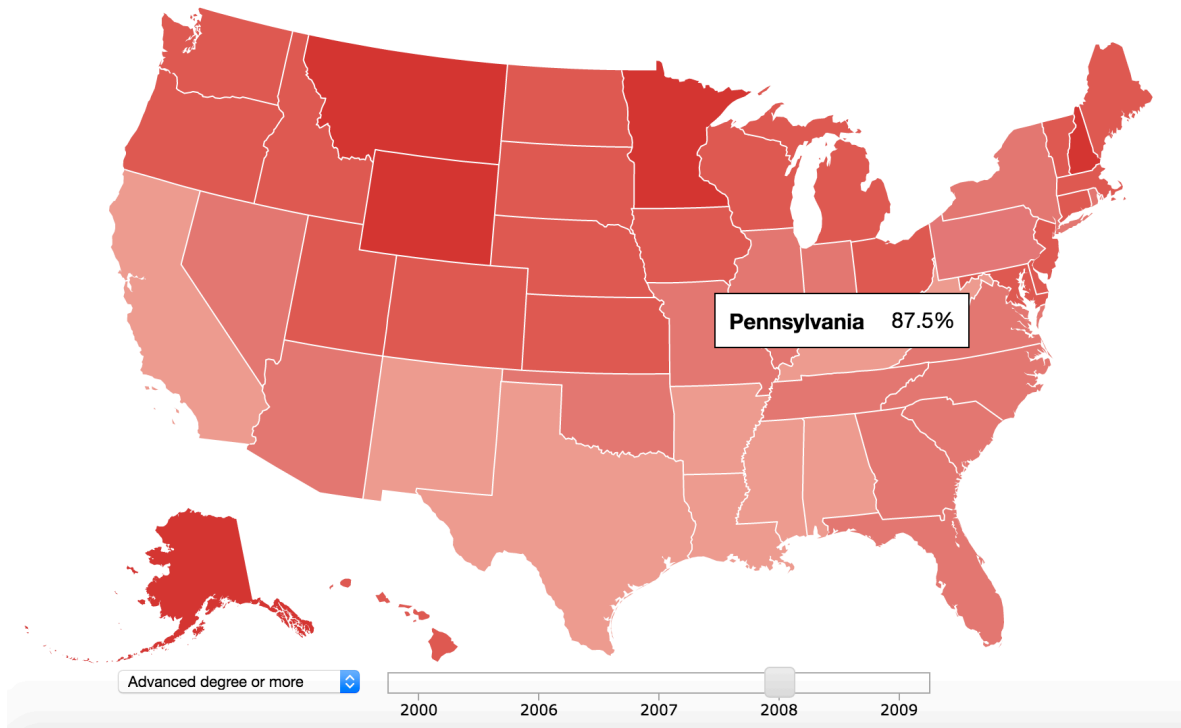
## Educational Attainment data from 2000 to 2009



## Educational Attainment data from 2000 to 2009



## Educational Attainment data from 2000 to 2009



On the map:

- Green filled map represents data in High School degree attainment.
- Purple filled map represents data in Bachelor's degree attainment.
- Red filled map represents data in Advancedr's degree attainment.

### Pros

- The map gives a very straightforward and effective way to see how data change across states and two years.
- You can also get detailed data for every state.

### Limitations

Comparison of data across states are direct, you can distinguish by the intensity of color. However, the comparison on several years and categories is indirect.

### Code Source:

<http://bl.ocks.org/cmdoptesc/fc0e318ce7992bed7ca8>  
<https://vida.io/gists/vh63p7sCnXuSctA7Y>  
<http://bl.ocks.org/mbostock/4699541>