# Visualization of Crime Survey Data

Akshay Gaur, Alekhya Kammadanam, Sudhindra Gopal Krishna
University of Oklahoma,
Department of Computer Science

***Abstract*** **Pew Research Center conducted a telephonic, survey between 07 and 11 January, 2015. The respondents were geographically diverse and represented all of USA. The survey had various questions touching different aspects of respondent's day-to-day life, like: "Improving education system", "Reducing crime" etc. The answers were in the form of a priority scale, the highest being "Top Priority" and the lowest being "Should not be done". The visualization shown above has been prepared ranking the various aspects based on the number of respondents giving it "Top priority". We chose to visually analyze statistics on the results obtained for the question on "Reducing crime".**

***Index Terms: Visual Analytics, Reducing Crimes, D3, Tableau.***

## INTRODUCTION

Surveys are conducted to know the trend and the changes that are affecting the day to day life of each and every individual. There are different kinds of surveys, like population survey (census), voter's survey and so on. There are different organizations to conduct these surveys. Pew Research Center is one such organization, who conducts survey all over United States of America. Pew Research Center, was started in the year 1990, with the name Times Mirror Center. In 2004 it was named as Pew Research Center. This research center in one of the reputed organization in USA. Every year, the survey will be conducted between 07 and 11 of January. Their survey includes the people's opinion on certain issue. They conduct a telephonic survey, and they ask the people to rate the question as either top priority or important but lower priority or not so important or should not be done. Few callers will not be knowing about that particular issue, so they have also included a section called don't know (DK). This survey had many questions like, reducing the crime rates, improving education, improving job situation and so on. These questions are directly related to president's ability to rule. For example, due to the terror attack at world trade center in 2001, people concern about the protection against terrorism increased, while conducting the survey in 2002. But the people's priority for defending the future terrorist attack stands top in every year. So visualizing these survey data will give a fine overview of the trend from past two decades. Visualization includes the survey from January 1995 to January 2015. Crime rate is an important factor in nation's development. Lower the crime rate, much safer is the country. This also affects the president's role in controlling such factors. So we decided to visualize the crime survey data, through which we will be able to establish the relation between the crime rates and people's concern towards reducing the crime.

Since this data is related to people's perspective, they are prone to change from one part of the region to other. So to get the exact data, we have also chosen to use the data which are provided by FBI. So with this we will

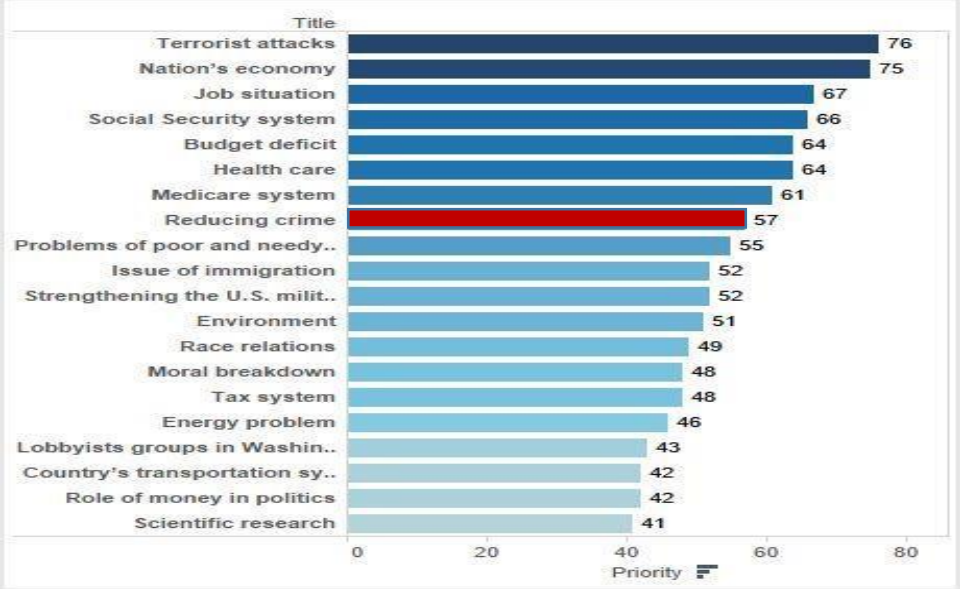



Fig1 : overall survey data of top priority.

be able to analyze whether the people concern about reducing the crime have changed or not. Figure 1 shows the top priority of all the survey conducted and they are arranged in an ascending order. From which we can infer that priority is given to control terrorism in a higher rate.

## 2 APPROACH

To visualize these crime data, we first thought of using a single tool, that is Tableau. Then after we processed the project plan, we used two different tools to visualize the data, so that we will be able to compare the function of two tools as well. The two tools that we chose to visualize the data are, *Tableau* and *D3*. The function about those two tools and the visualization obtained from those two tools are explained in the coming sections.

## 3 TABLEAU

In our project we are making use of tool by name Tableau. This is one of the visualization tool. Tableau expects the visualizer to just insert the data file which The main disadvantage of using Tableau is that whenever we need share the visualization, either we need to publish the visualization in the Tableau website. Otherwise we need to convert it into image or a PDF and share it with others, which makes way to the loss of interaction. The above visualization was created using **Tableau® Desktop Application** based on the responses received asking people to prioritize crime reduction. The exact question asked was: "I'd like to ask you about priorities for President Obama and Congress this year. As I read from a list, tell me if you think each should be a top priority, important but lower priority, not too important or should it not be done. (First,) should reducing crime be a top priority, important but lower priority, not too important, or should it not be done?".
is a csv or excel format file. To do that first we need to convert the rawdata from source and convert it into the Tableau dependent form that is ".csv". This software comes in handy for the user to interact. Tableau helps the user to see and understand the data and even analyze and visualize the information. To obtain such a visualization we start with choosing a file, after choosing a file Tableau automatically recognizes the data separates the attributes with the Dimensions and Measures. Dimensions are categorical fields that describe our data and Measures are numerical and can be used for mathematical purpose. Tableau reads the data dimensions and automatically helps the visualizer to choose the type of the visualization. After choosing the visualization we can edit the visualization to our need and look. To build the interaction we use the Filter option that is present in thesoftware. We can also make use of drag and drop option which Tableau creates its own visualization according to the data provided. Likewise, we can create as many visualizations that is required and then by clicking on the Dashboard Button we can place our visualization and we can even publish the same in Web, so the viewers can view our visualization. Tableau also connects social media for sharing our project. Figure 2 shows the stacked bar graph obtained using Tableau, the data set are taken from 19952015 of January. The x-axis represents the years and y-axis represents value.

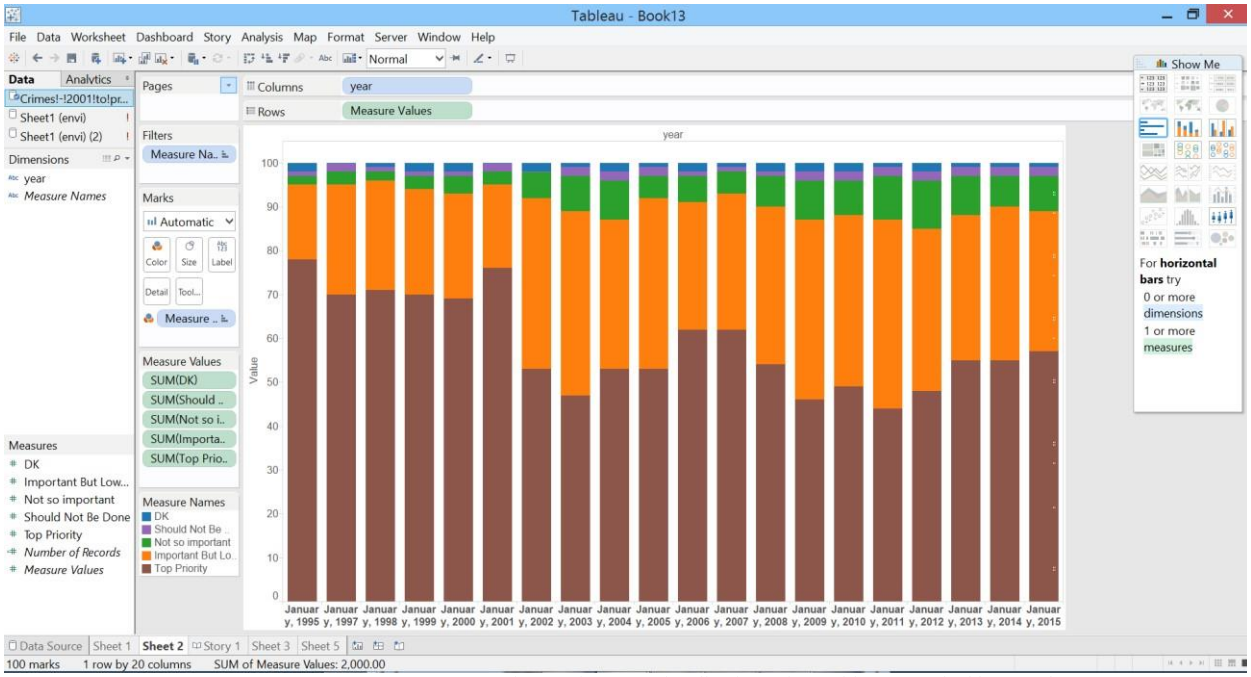


Fig 2: Crime rate visualization using Tableau's stacked bar graph

## 4 D3.js

D3.js is a JavaScript library for manipulating documents based on data." "D3's emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization components and a data-driven approach to DOM.

Learning D3 and implementing it was as expected and required about 100 hours of learning. However, once learnt, it is a pretty straightforward programming language. Most of the heavy lifting is done by the library itself. We were able to learn many visualization techniques and their implementation techniques. We have used Firefox as our choice of browser for implementing and testing our D3 visualizations. We are using HTML5 pages, CSS stylings and JSON as our choice of data source. We have successfully implemented basic visualization styles with animation as shared in the link above which might not necessarily make a lot of sense now as this is just a proof of concept of what can be achieved with D3. We had to create everything from scratch starting from the pages to the JSON file itself. We estimate that creating the JSON files will require 20% of the effort while creating the visualizations with the animation will require 80% of the effort.

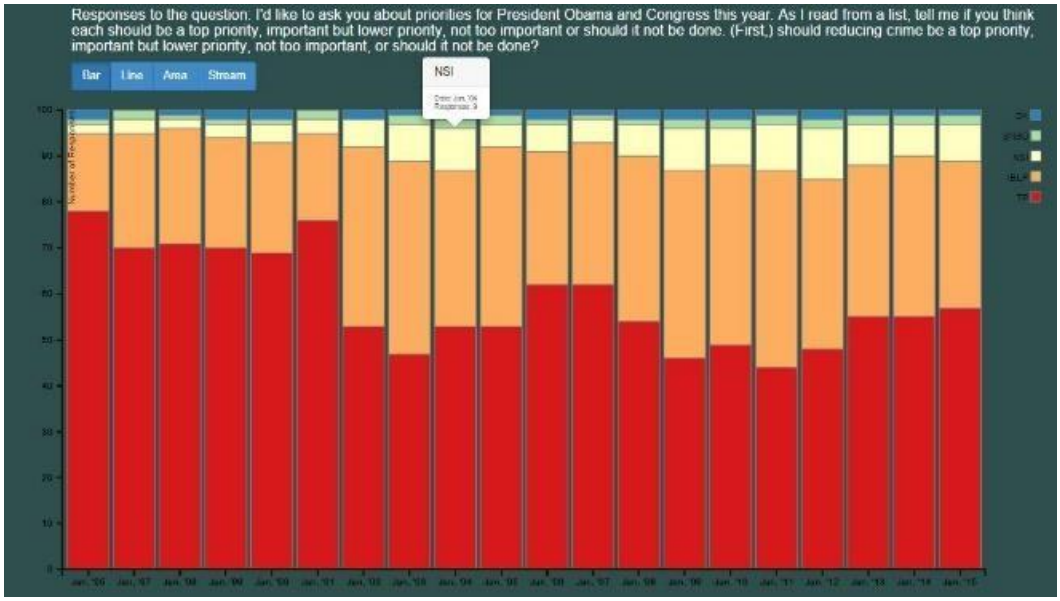



Fig 3a: Stacked Bar Graph using D3.js

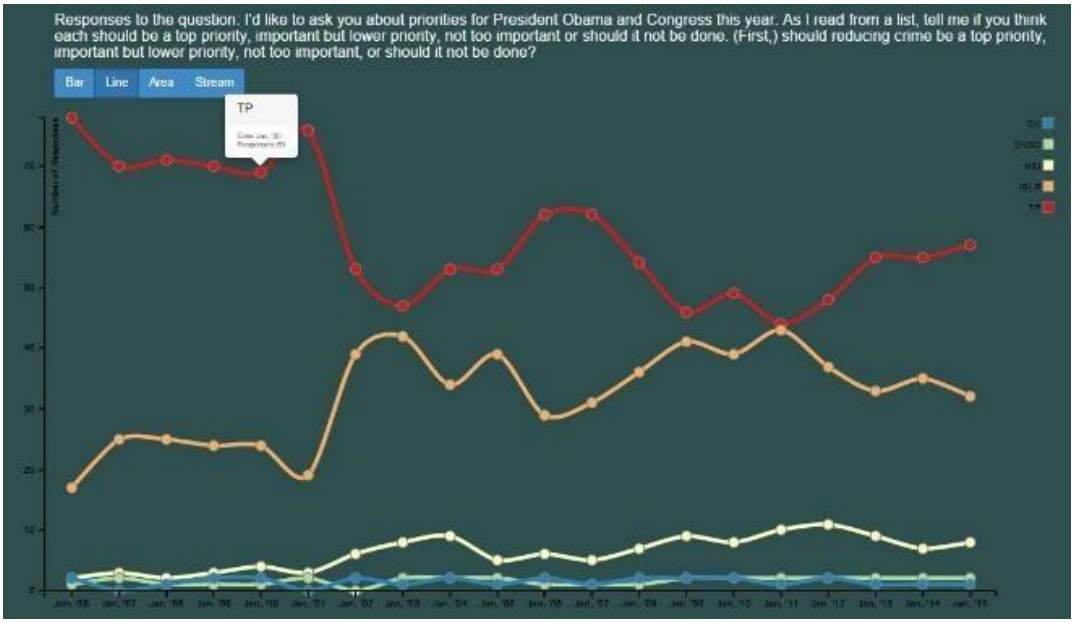



Fig 3b: Line Graph using D3.js

Figures 3a, 3b shown below are the visualization created using D3.js.

## 5 DATA SET

As mentioned in the previous sections, the extracted data in its raw form had many questions and each questions contained a detail survey. So from the pool of the questions we considered the most trending and happening issue that is, reducing crime, according to the overall statistics nearly more than 50% of people concern about reducing crime. Figure 5a shows the sample of the data set collected through the survey by the pew research center.

Since it's the perception of the individuals, the data provided cannot be concluded correct unless if it matches the original data. So as to compare the survey data with the original data, we extracted the data from FBI. So the sample dataset collected from FBI is shown in Figure 5b.

## 6 TABLEAU VS. D3.js

Tableau as said early it is very much easy to use and it gives different visualization for differentit. Say for example, whatever may be the data, the programmer must understand what data sets of data, which is by analyzing what visualization suits better for a particular data set. For example, if there is a dataset consisting of state names of USA and some data related to it, the Tableau automatically selects *maps* visualization. While that of D3.js, since it is a Java script it require a lot of programming knowledge, while working on it is and he should create relative visualization to it.

| | Top Priority | Important But | Not so imp | should not | DK |
|---|---|---|---|---|---|
| January, 2015 | 57 | 32 | 8 | 2 | 1 |
| January, 2014 | 55 | 35 | 7 | 2 | 1 |
| January, 2013 | 55 | 33 | 9 | 2 | 1 |
| January, 2012 | 48 | 37 | 11 | 2 | 2 |
| January, 2011 | 44 | 43 | 10 | 2 | 1 |
| January, 2010 | 49 | 39 | 8 | 2 | 2 |
| January, 2009 | 46 | 41 | 9 | 2 | 2 |
| January, 2008 | 54 | 36 | 7 | 1 | 2 |
| January, 2007 | 62 | 31 | 5 | 1 | 1 |
| January, 2006 | 62 | 29 | 6 | 1 | 2 |
| January, 2005 | 53 | 39 | 5 | 2 | 1 |

Fig 5a: Sample data on crime rate

| Year | Population | Violent crime total |
|---|---|---|
| 1995 | 262803276 | 1798792 |
| 1996 | 265228572 | 1688540 |
| 1997 | 267783607 | 1636096 |
| 1998 | 270248003 | 1533887 |
| 1999 | 272690813 | 1426044 |
| 2000 | 281421906 | 1425486 |
| 2001 | 285317559 | 1439480 |
| 2002 | 287973924 | 1423677 |
| 2003 | 290788976 | 1383676 |
| 2004 | 293656842 | 1360088 |
| 2005 | 296507061 | 1390745 |

Fig 5b: sample FBI crime data.

Since Tableau is that easy, its most use is found with the business analyst, where the business analyst will not have time to program and obtain the visualization. Since his main job is just to analyze data and create a report. But whereas with D3.js it is possible to obtain a good pixel perfect dashboard. D3 is often used for very specific UI/UX requirements or branding guidelines.

Tableau has got its own restricted environment, meaning only certain chart types can be used for a set of data and also dashboards must be shared via Tableau Server, and they are limited to the formatting options included within Tableau. While D3 can do which Tableau cannot. Licensing is required to access Tableau software, which is the greatest border for the few programmer. However, the student version is free for a year. But D3 is an open source visualization tool. With which at any point of time one can share the visualization, through website. Where the interactivity is kept alive. This gives D3 a greater advantage over Tableau.

## 7 FBI CRIME DATA SURVEY

Figure 6 depicts the crime rate from 1995-2014 which was collected by FBI. These data are represented that, as the

year goes on there is a steady increase in the population. But we can also infer from the graph that the crime rate is also decreasing steadily. The red line represents the decrease in crime rate, whereas the other line shows the increase in the population. As we have obtained this visualization, it is now easy for us to compare the visualization from the survey datasets, where people concern about the reducing crime is really needed or not. The visualization is obtained as shown in Figure 6b.

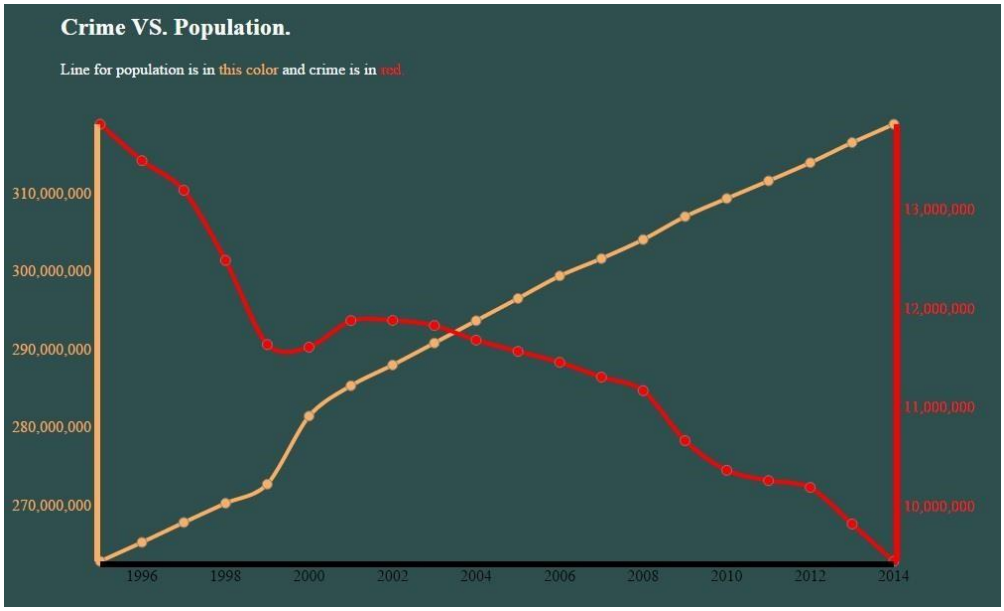



Fig 6a : FBI crime survey visualization

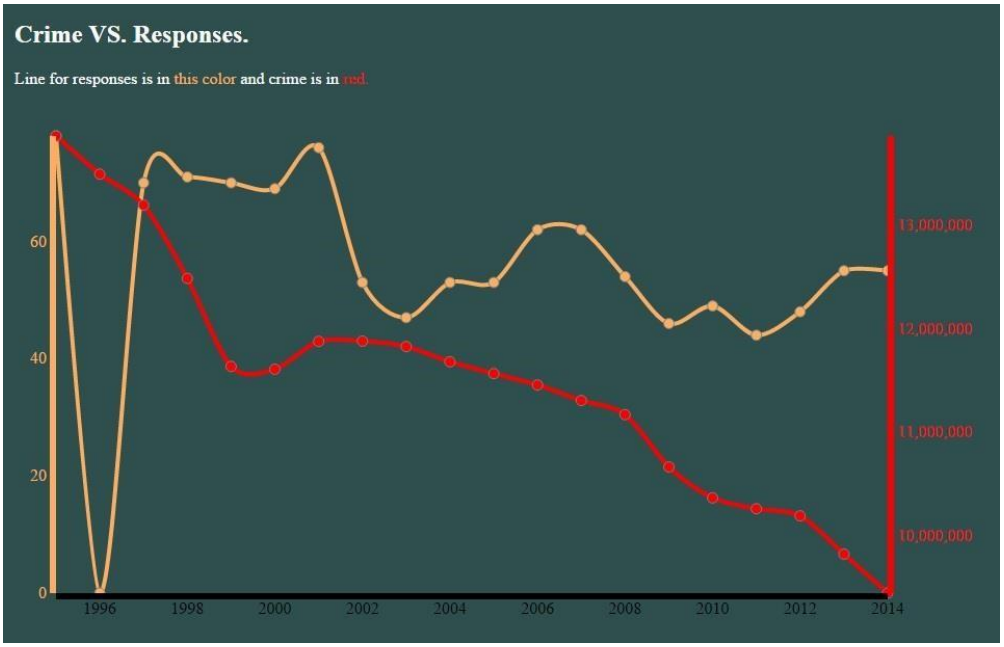
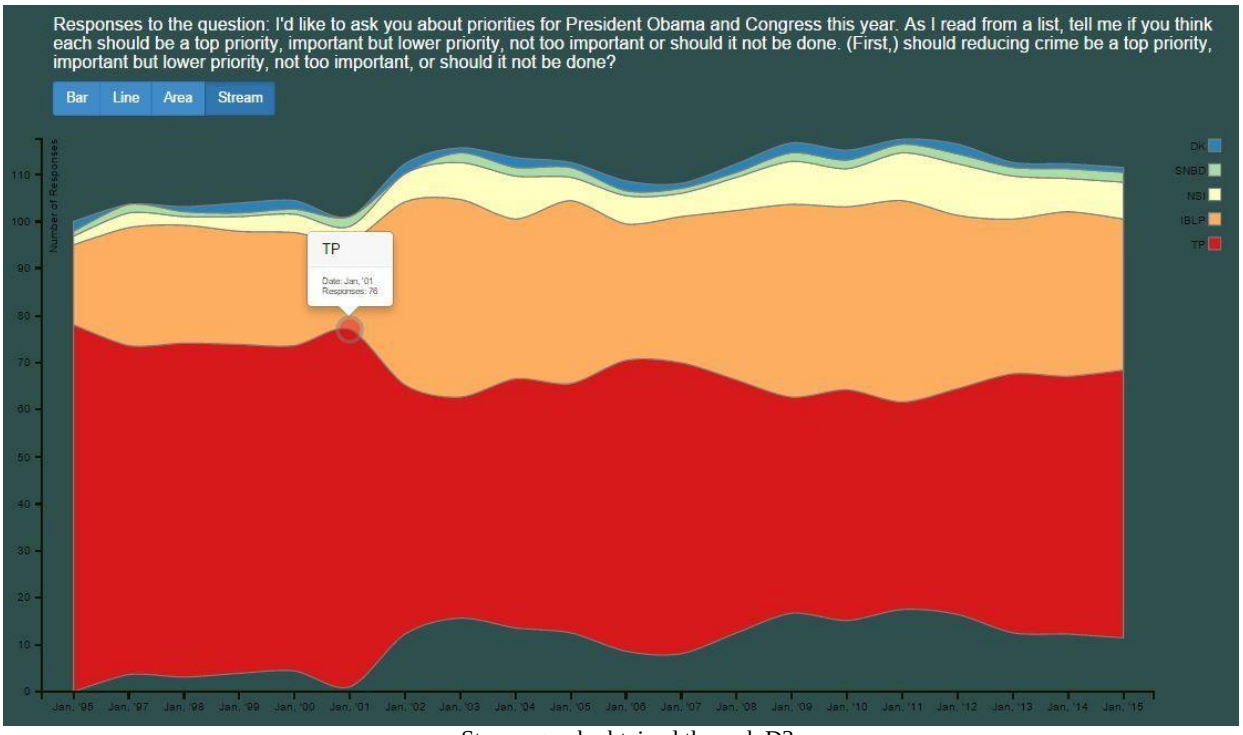



Fig 6b: People survey crime data

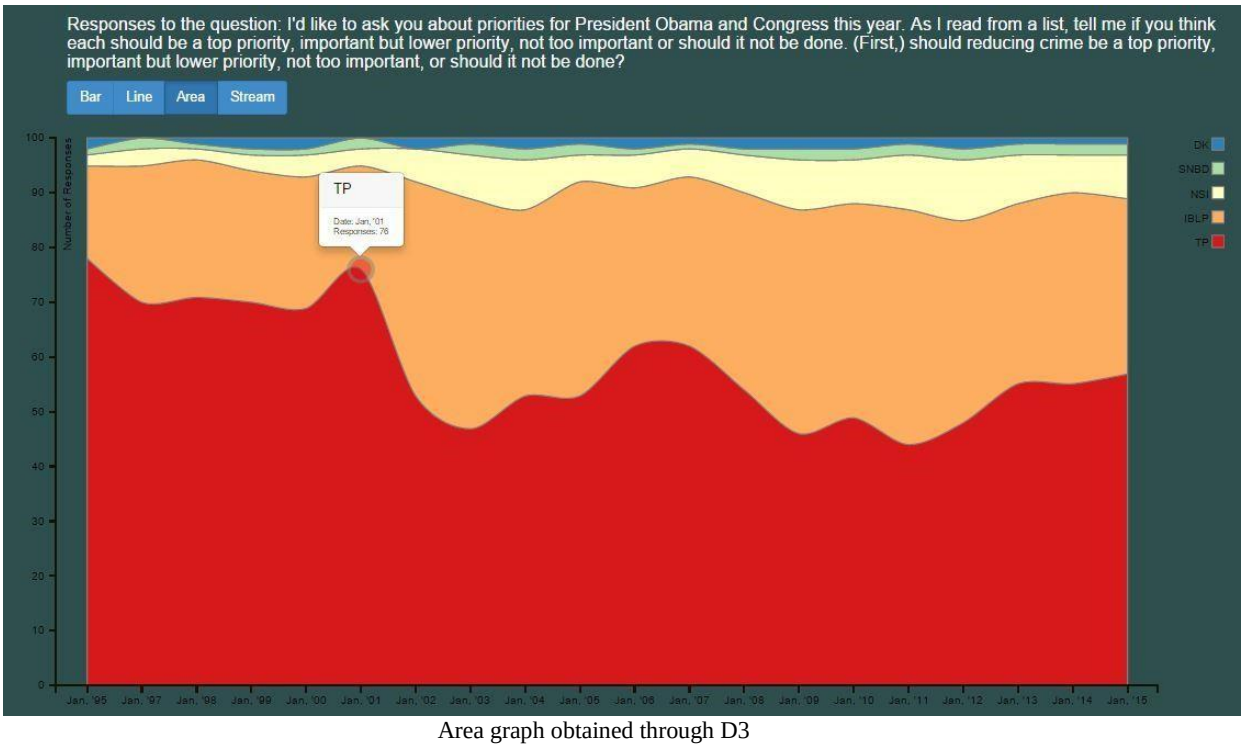
## 8 ANALYSIS and CONCLUSION

From the above two visualization, we can clearly differentiate between the FBI crime data with the people's opinion. In the FBI dataset, it is clear that, though there is increase in the population, the crime rate low. Whereas from the survey we have more people who are more concerned about reducing the crime rates, though there is no much crimes happening. From this we can clearly infer that the data are from the perception of people, which are sometimes not reliable.

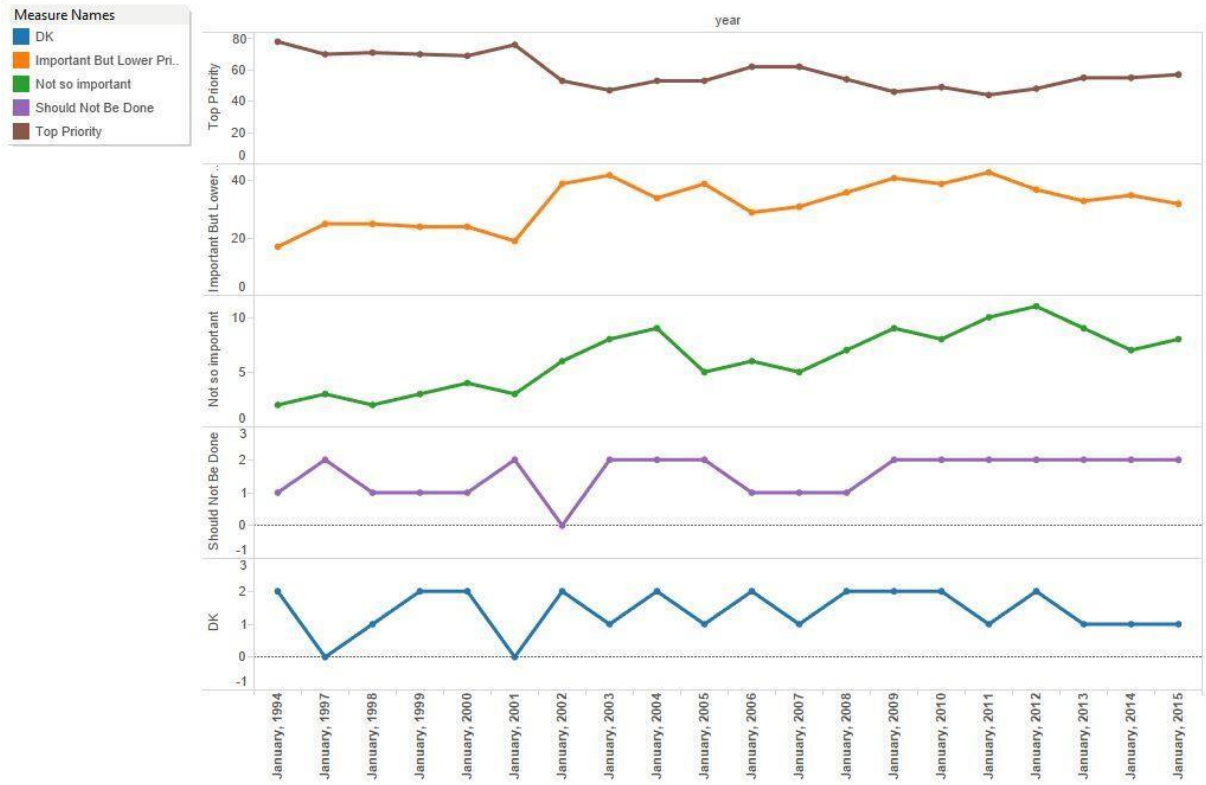
Here are some of the other visualizations obtained from both D3.js and Tableau.




Stream graph obtained through D3




Area graph obtained through D3




line graph obtained through Tableau

REFERENCES

[1] Pew Research Center: http://www.people-press.org/
[2] Tableau: www.tableau.com
[3] D3: www.d3js.org
[4] FBI: https://www.fbi.gov/about-us/cjis/ucr
[5] ColorBrewer: http://colorbrewer2.org/