CRIME TRENDS IN US STATES (1965 – 2012)

CS573 FINAL PROJECT – PROCESS BOOK

By: Amogh Raghunath, Arun Vadivel and Harsh Vardhan

CRIME TRENDS IN US STATES (1965 – 2012)

CS573

FINAL

PROJECT –

PROCESS

BOOK

CONTENTS

1. Overview and Motivation 2
2. Week 1 3
3. Week 2 9
4. Week 3 14
5. Week 4 20

OVERVIEW AND MOTIVATION

Crimes and terrorism are threatening to the world, taking away the peacefulness and happiness from the society. Coming to US for our Master's program, we were curious to know how US tackles its crime problems. We wanted to make an interactive visualization to understand the crimes types and their trends taking place in different parts of America.

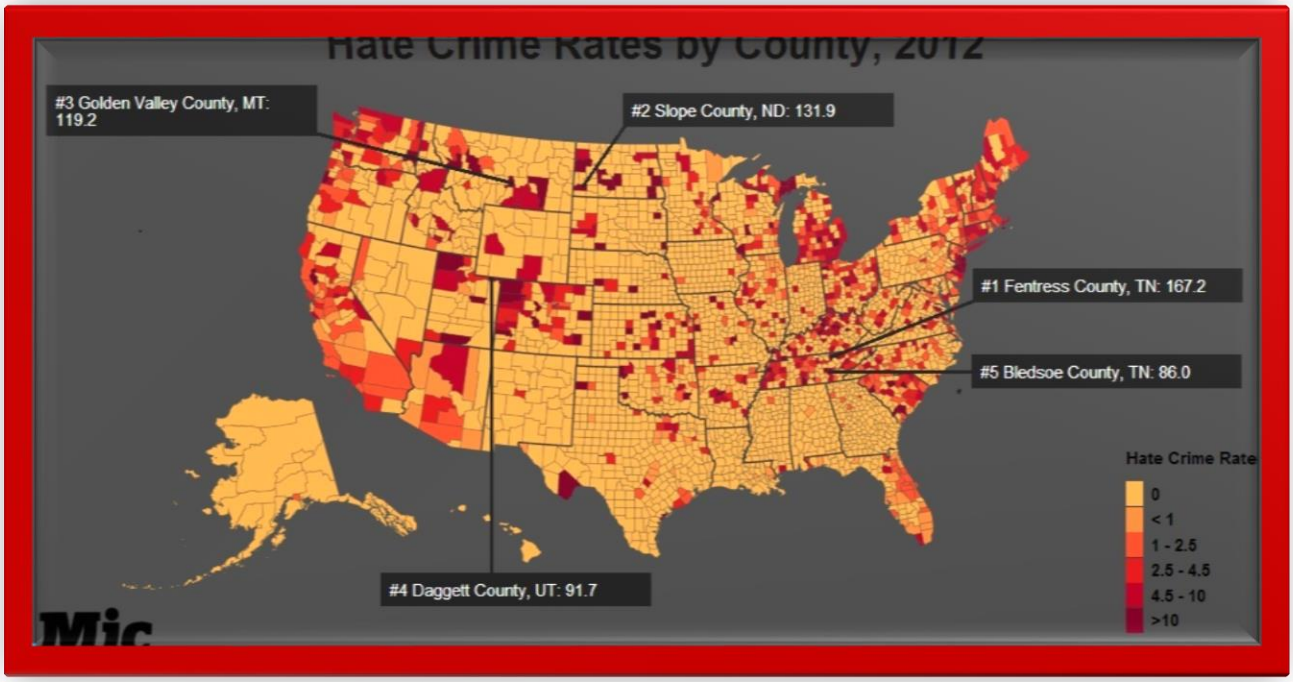
The objective of our project is to understand and explore what are the different categories of crime committed across the states of US. There were no good visualization available on crimes in US, which compared states or showed trends over the years. We want to compare every category of crime like rape, murder, etc. committed across the states of US. Over the years, are the crime tends to be decreasing or increasing or is it stable? We want to understand which states are worst or safest in terms of the data provided. We would like to learn which states have improved or become worst over the years.

WEEK - 1

1. Related Work

Before finalizing this visualization, we visited many websites that provides a visualization on crime reports. Only few visualizations impressed us and answered only few of the questions. One such visualization was restricted to the crime data for the year 2012. So, we took this one as a model and planned to extend it for many years to provide a better insight. Using crime data from the FBI's website, we planned to extend this visualization for the years 1960 to 2012.

Some of the visualizations that inspired us can be seen below:



2. Question

In this visualization we are trying to answer the following questions:

- How can we visualize the different types of crimes in every state?
- What are the crime trends in every state for different year from 1960-2012?
- What are the safest states and dangerous states in America to live?
- Which states have consistent crime control throughout these years?
- Which states have got worse compared to the past?

3. Dataset

Source - Crime report data from U.S. Department of Justice (FBI)

Link:
http://www.ucrdatatool.gov/Search/Crime/State/StatebyState.cfm?NoVariables=Y&CFID=120725516&CFTOKEN=bd96e08ed9febd35-29F3C038-DA0F-AE2A-333418C82FDB6803

The data is available in a .csv file for each state. Fields for each .csv file - Year, Population, Violent Crime Total, Murder, Rape, Robbery, Assault, Property Crime total, Burglary, Larceny - Theft, Motor Vehicle – Theft. The data is available from the year 1960 – 2012.

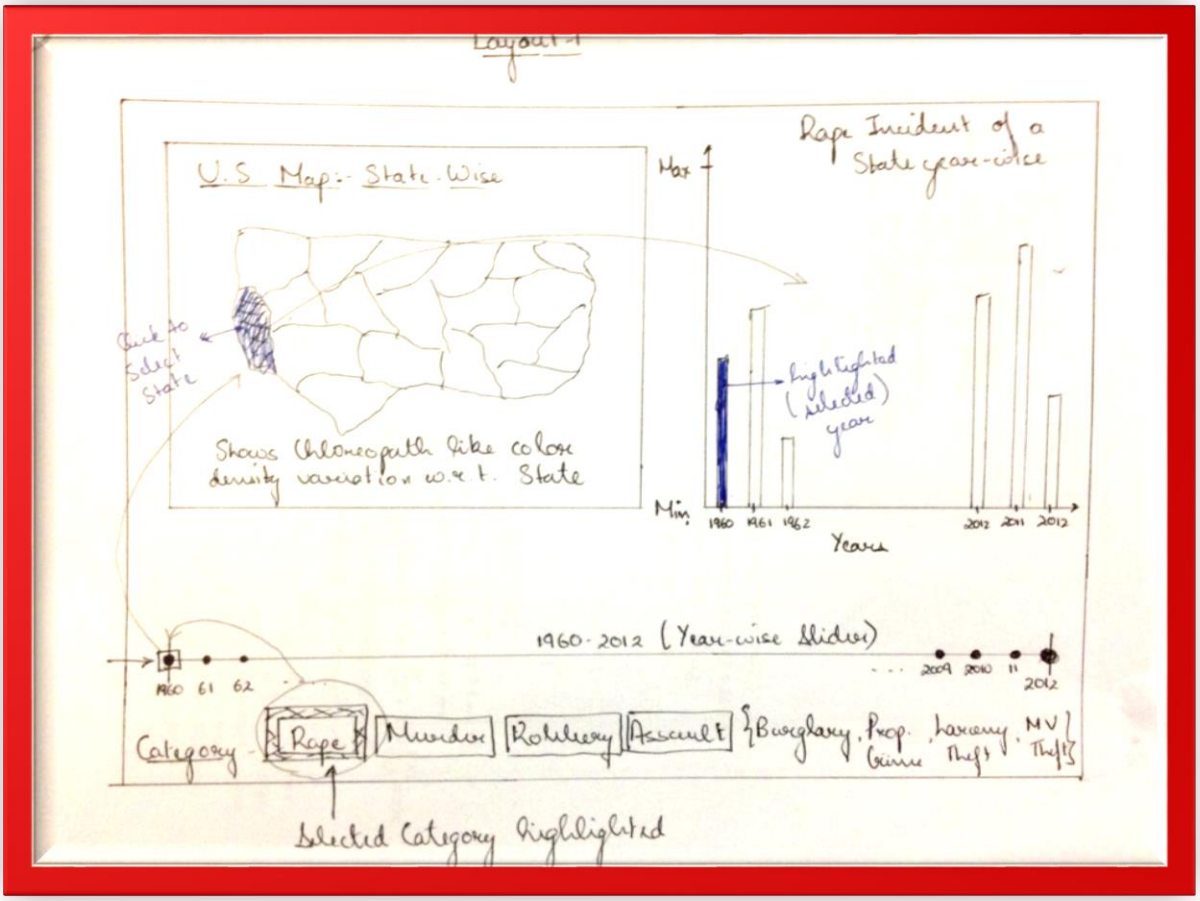
The data is clean. The only problem is that the data is separate for every state and for different crime. So, we have to consolidate all the state's data for every year and put it in a master CSV file. In order to do this we'll need to use Python to append all the data into a single master file.

4. Exploratory Data Analysis

Plotting rough sketches of visualizations on the gathered data is a good way to understand the insights of data. We tried to plot the data using Tableau as Scatterplot, Bar chart and Pie chart. We found that Bar chart would suit better compared to the other two when comparing the crime rates of each states for the years 1960 to 2012 by using the experiments performed on many users and by validating the scores. The mentioned Bar chart was planned to work on based on the map whenever the user hovers over a specific state.

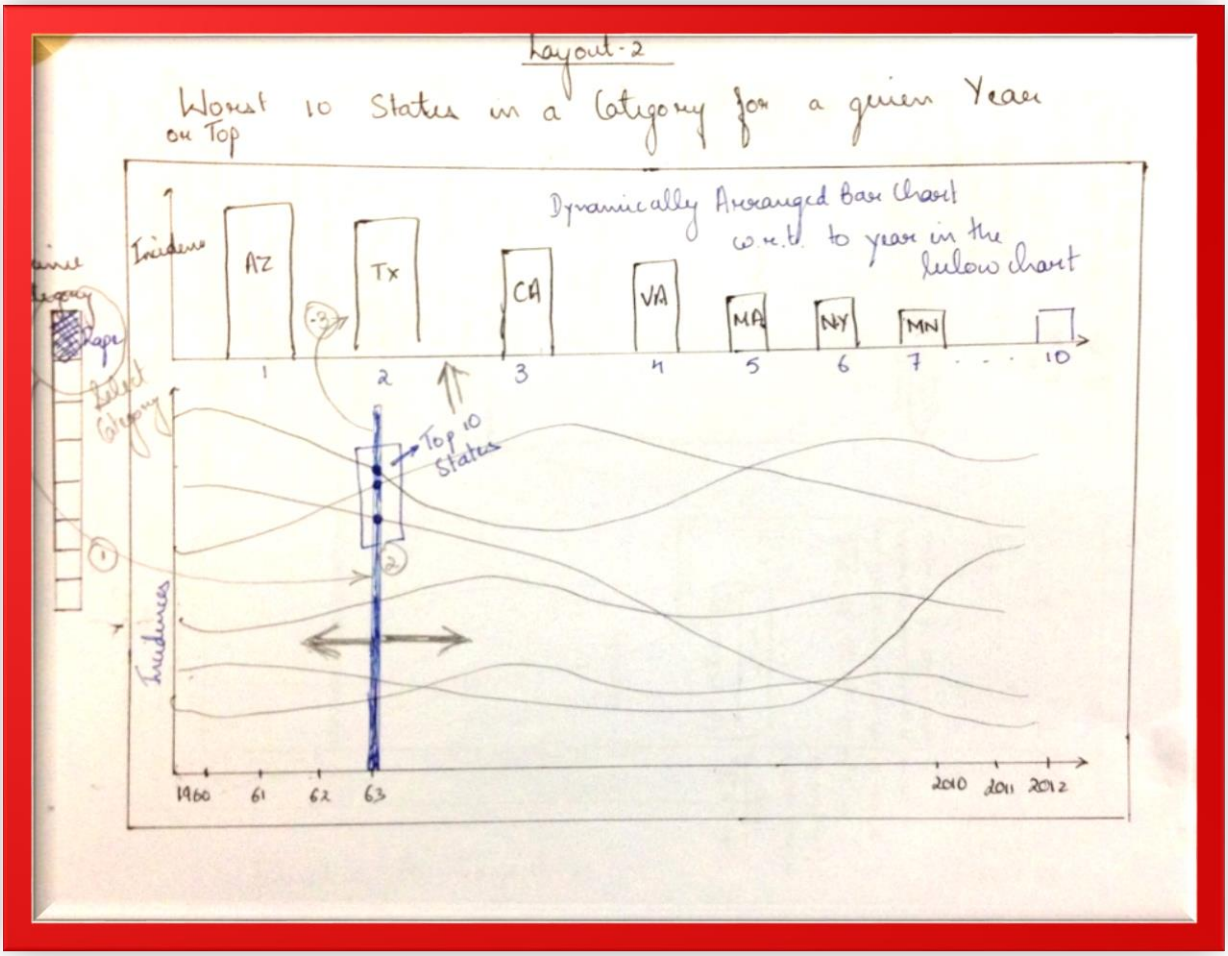
5. Design Evolution

Before starting with visualization and coding, we planned layouts that would help us to understand what are the best type of visualization we can use is for the data. For this exercise we used sketching as the basic strategy to figure out possible views, interaction, charts types and orientation. Below are the sketches and a brief description of what we are trying to achieve.



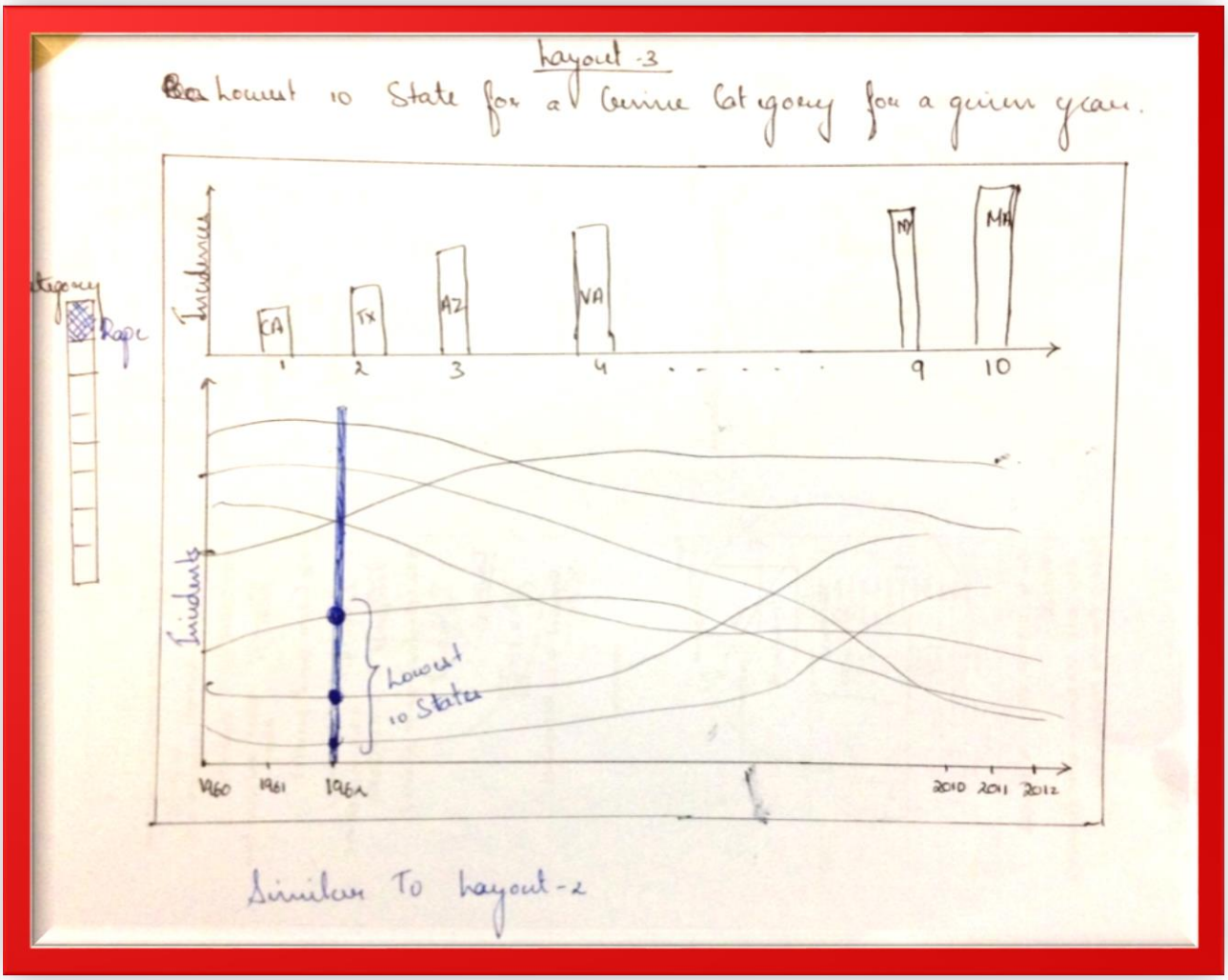
Layout 1

An interactive map of US, which shows variations in color when particular year and a crime category is selected. Also, a state selection will alter the bar chart and show all the crime rate for a particular crime category from 1960- 2012



Layout 2

It shows the trends in the form of line chart for all the states in the lower half for a given crime category. The user can hover over the line chart and for a particular year ranking based bar chart in the top will appear. We are planning to show only the worst 10 states in this visualization.



Layout 3

It is an exact copy of the layout 2, but instead it shows the top 10 states having the lowest crime rates for a given crime category.

WEEK – 2

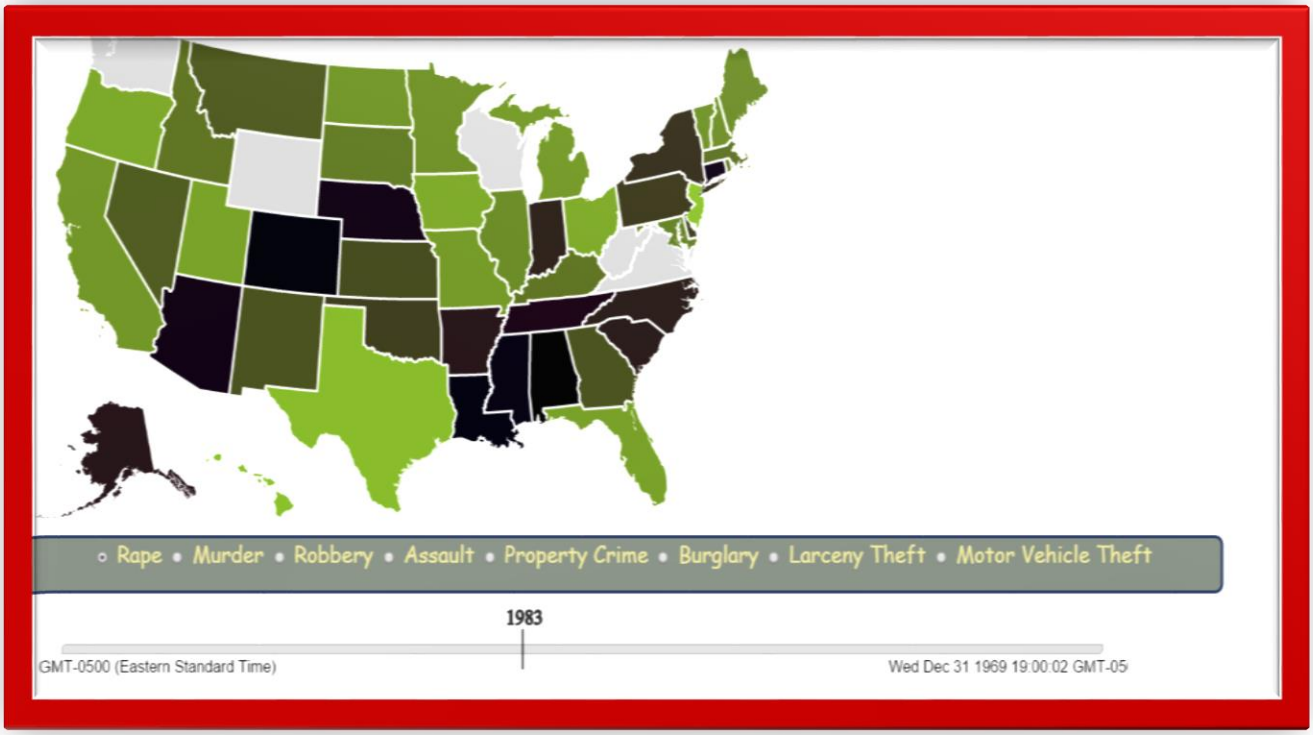
1. Data Processing

As we started with the code, we found key data missing for the state of “New York” from the year 1960-65. Since, we wanted to focus on all the states and our map would turn out to be blank for this state; we decided to shift the data range from 1965 – 2012. Also, we added a new column “id” to the dataset to identify and draw US states.

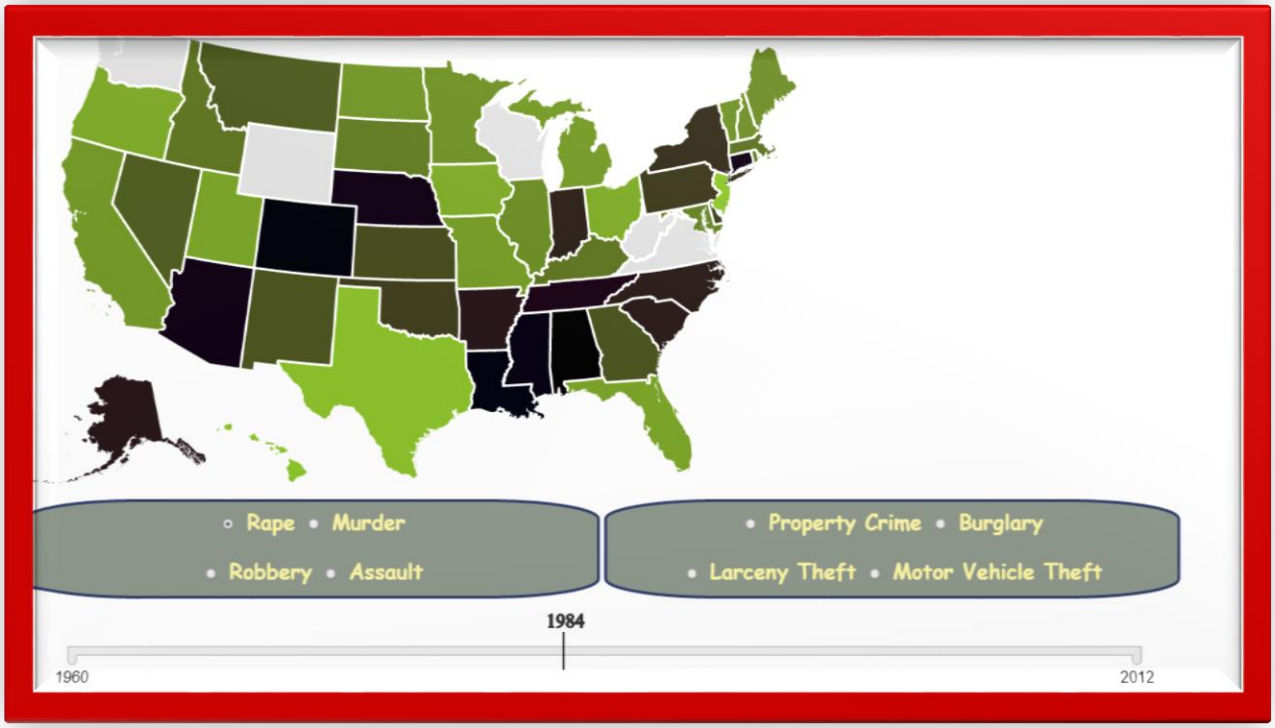
State	year	Violent	MNMR	rape_ra	robbery	assault	propert	burglar	larceny	motor_	hicle_theft
Massachu	1966	115.2	39	42	20	37	24	22	38	2	
Massachu	1967	127.6	39	38	24	38	26	26	39	2	
Massachu	1968	164	36	36	22	36	21	20	36	2	
Massachu	1969	187.9	37	39	19	36	19	16	35	2	
Massachu	1970	202.9	39	34	20	36	19	15	36	2	
Massachu	1971	266	36	39	14	34	17	14	33	2	
Massachu	1972	295.2	39	40	12	30	17	14	36	1	
Massachu	1973	351.9	37	34	11	30	15	13	34	1	
Massachu	1974	388.7	38	38	11	30	12	13	33	1	
Massachu	1975	442.6	40	29	10	27	13	14	35	1	
Massachu	1976	399.2	41	37	12	23	13	12	37	1	
Massachu	1977	425.3	44	32	13	20	17	14	38	1	
Massachu	1978	462	41	32	14	20	18	13	39	1	
Massachu	1979	531.3	42	33	13	17	16	12	37	1	
Massachu	1980	601.3	40	32	10	14	18	14	40	1	
Massachu	1981	628.6	41	31	9	16	22	18	41	1	
Massachu	1982	571.4	38	32	11	14	22	20	41	1	
Massachu	1983	576.8	41	28	10	11	23	21	42	1	
Massachu	1984	523.7	37	26	13	13	25	26	42	2	
Massachu	1985	538.2	40	24	13	15	26	25	41	1	
Massachu	1986	556.9	39	27	15	18	28	29	44	2	

2. Implementation

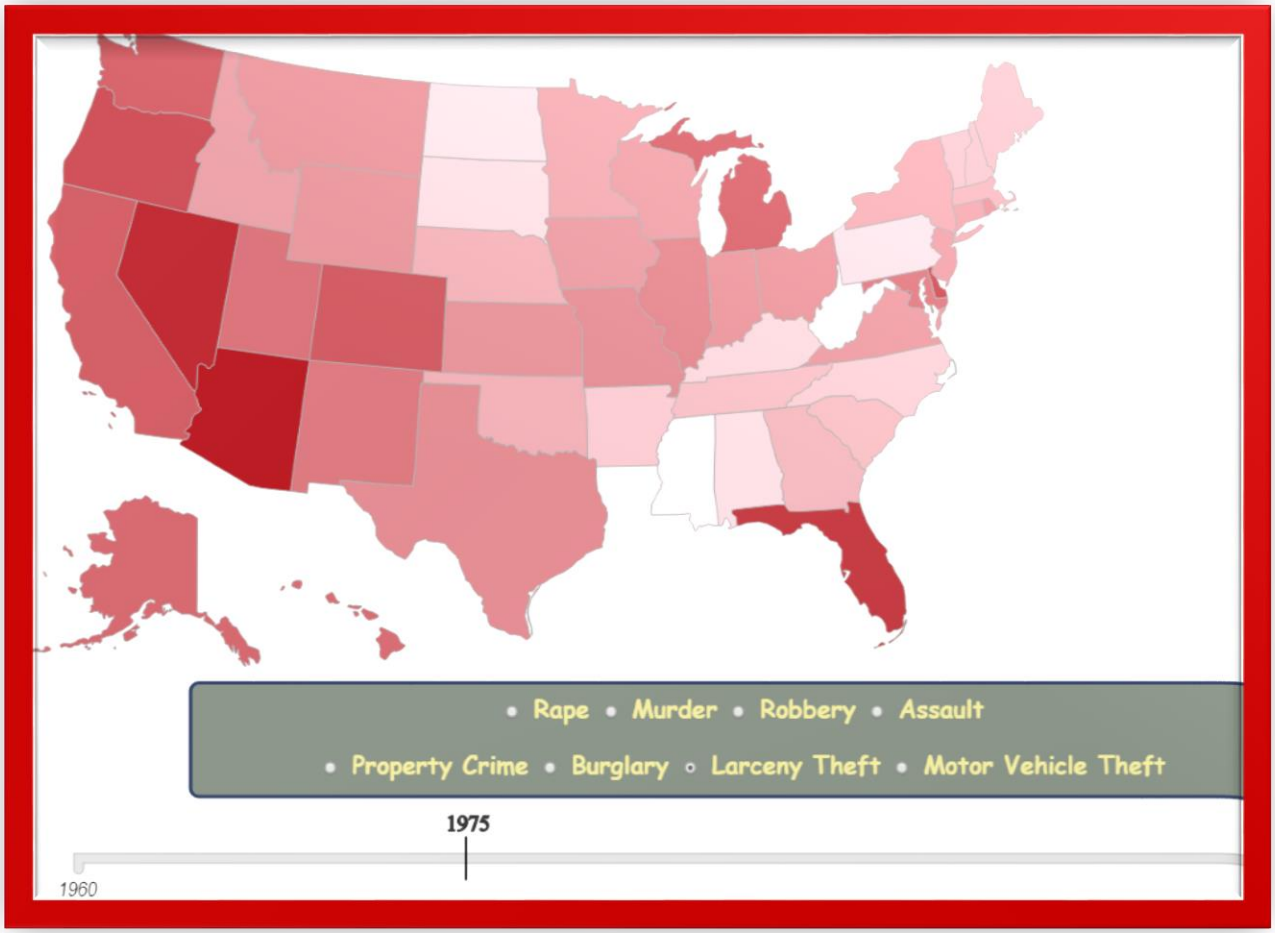
For the implementation part we started out with Layout -1. On the data side, we had gathered in the form of a .csv file having all the categories, state ids and crime rates. At start, first we tried to get the data only for the rape category and upload the US map for the year 1983 as the start of the visualization. Slider and a category selector in the form of the radio button is added.



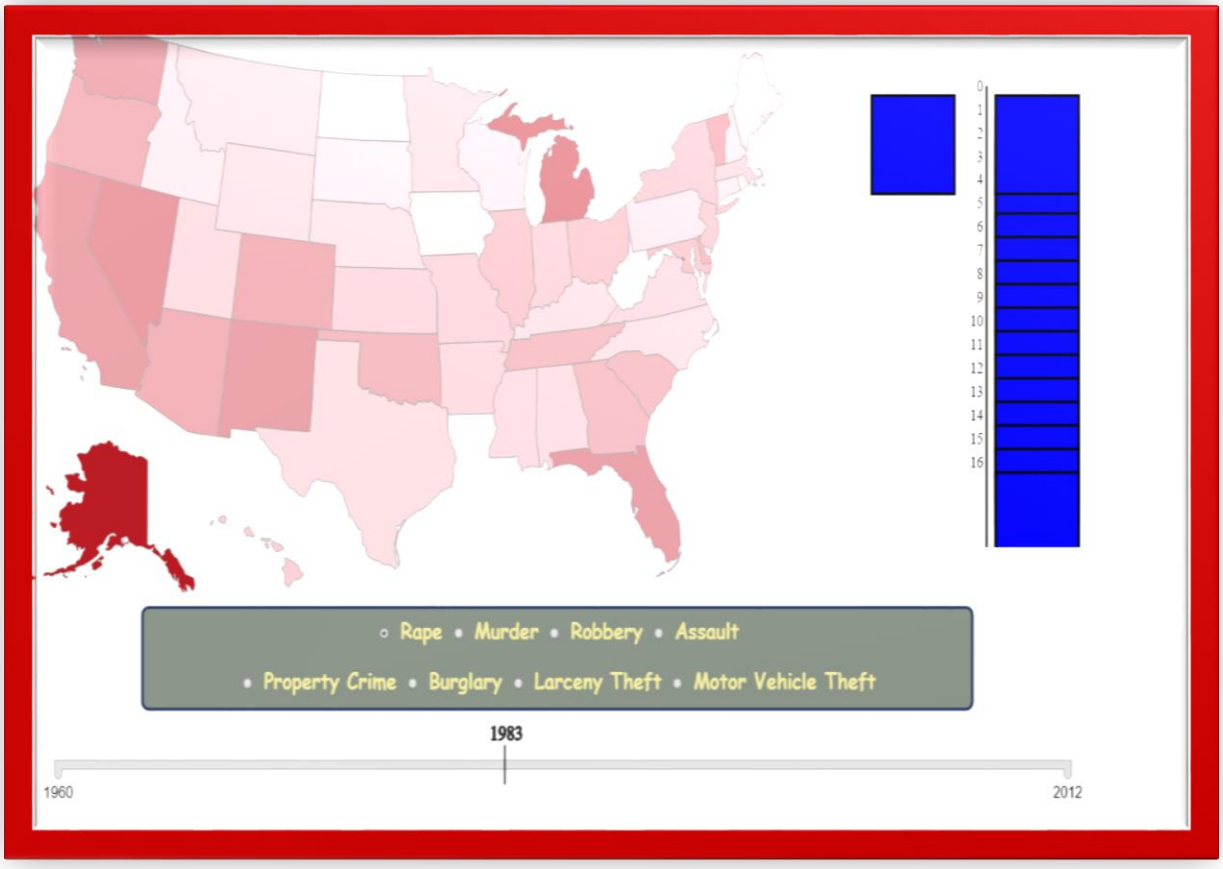
Next, we wanted to implement the slider to be working on the year. As we move the slider, the US map should change accordingly.



In the above visualization US states, the color used to fill covered a wide range of colors. We tried to implement a suitable color range. Also, we implemented the working of the radio buttons for various crime categories.



Next, we tried to implement the last part of the layout on the right side of the viz. We wanted a bar chart to be visible whenever the user hover over the state on the US map. The bar chart shows the bars for the selected crime category over the entire range 1965 – 2012. The below visualization shows just the implementation to get the bar chart on mouse over.



WEEK – 3

1. Feedback from Project Prototype

The feedback we received from the project presentation was insightful. The following changes were proposed:

- To integrate the time data into trends by simple chart type such as line chart.
- To drop one of the 4 dimension in the data to make it simple and having trends over the entire time range of 1965 – 2012.

To incorporate these changes in our viz, we decided to go for another layout which will show data across the years from 1965 – 2012 for different states of US. Also, we wanted to implement a viz in which we could compare one US state with other or all states by user input. So, going forward we went ahead with a line chart showing crime rates on y-axis and years on x-axis; category and state being controlled by user.

2. Design

Instead of implementing a bar chart, we decided to change the layout 1 to show the rank of a particular state for all the crime category. This change in the design will lead to removal of layout 2 & 3. The basic idea for the layout 2 & 3 is to find the top 10 and worst 10 state over the period of 1965 – 2012. But, we can implement this feature in the layout 1 itself by making a table to show the ranking

and that also for all the category which is an improvement over the previous layout. We have also to add the legends for the US map and a comment for describing how the data is in its raw state and how the user should interpret the US map.

Since, there will be a lot of interactivity in our visualization of layout 1, we decided to go for an introduction page prior to starting the visualization and explaining the idea and motivation behind the viz. This will also contain a link to the process book as well as the screen cast to our viz. Screen cast will help the user to understand the implementation and interactivity.

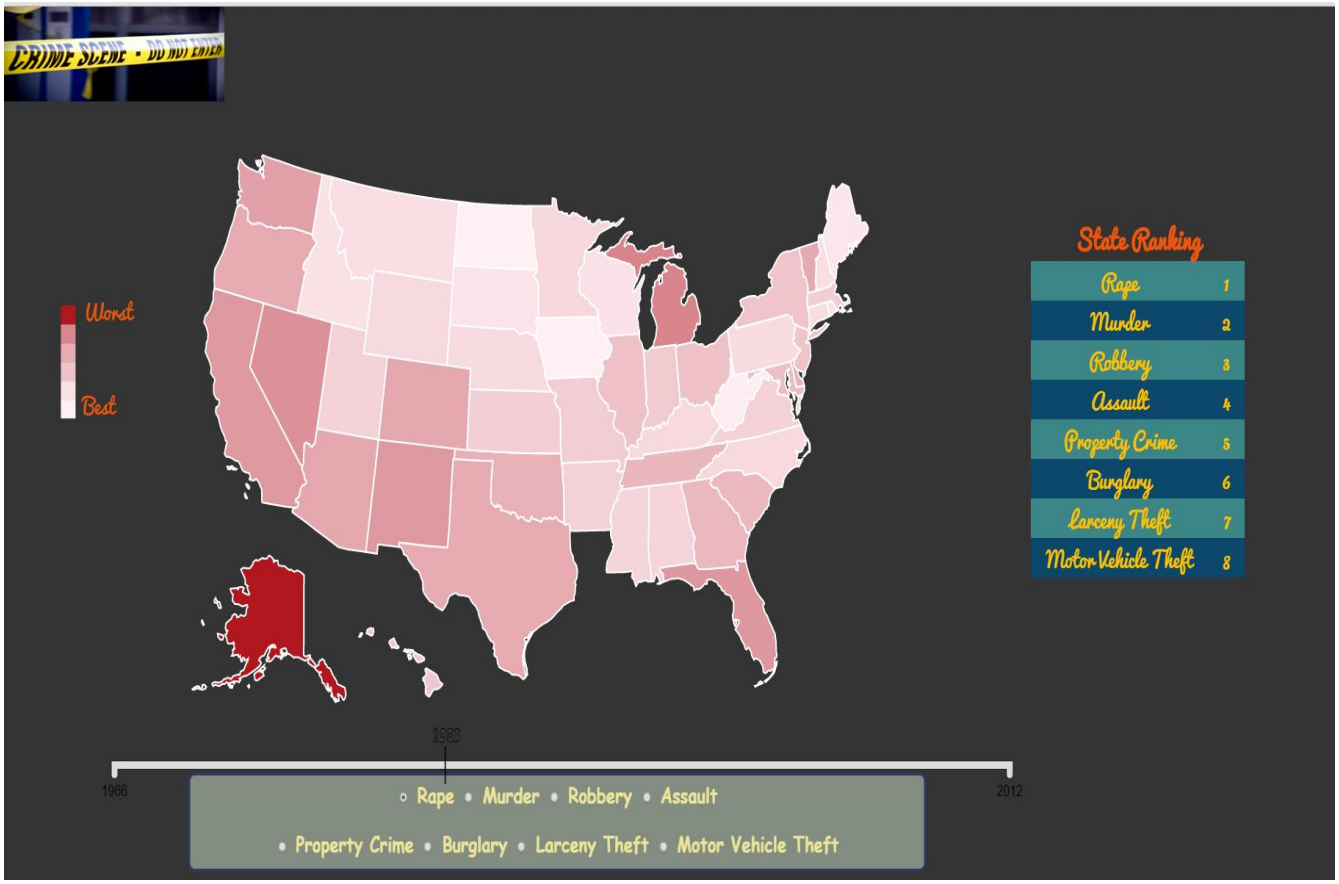
The radio button were to be replaced by images as a click images to select the category as it will make our viz more colorful and engaging.

3. Data Processing

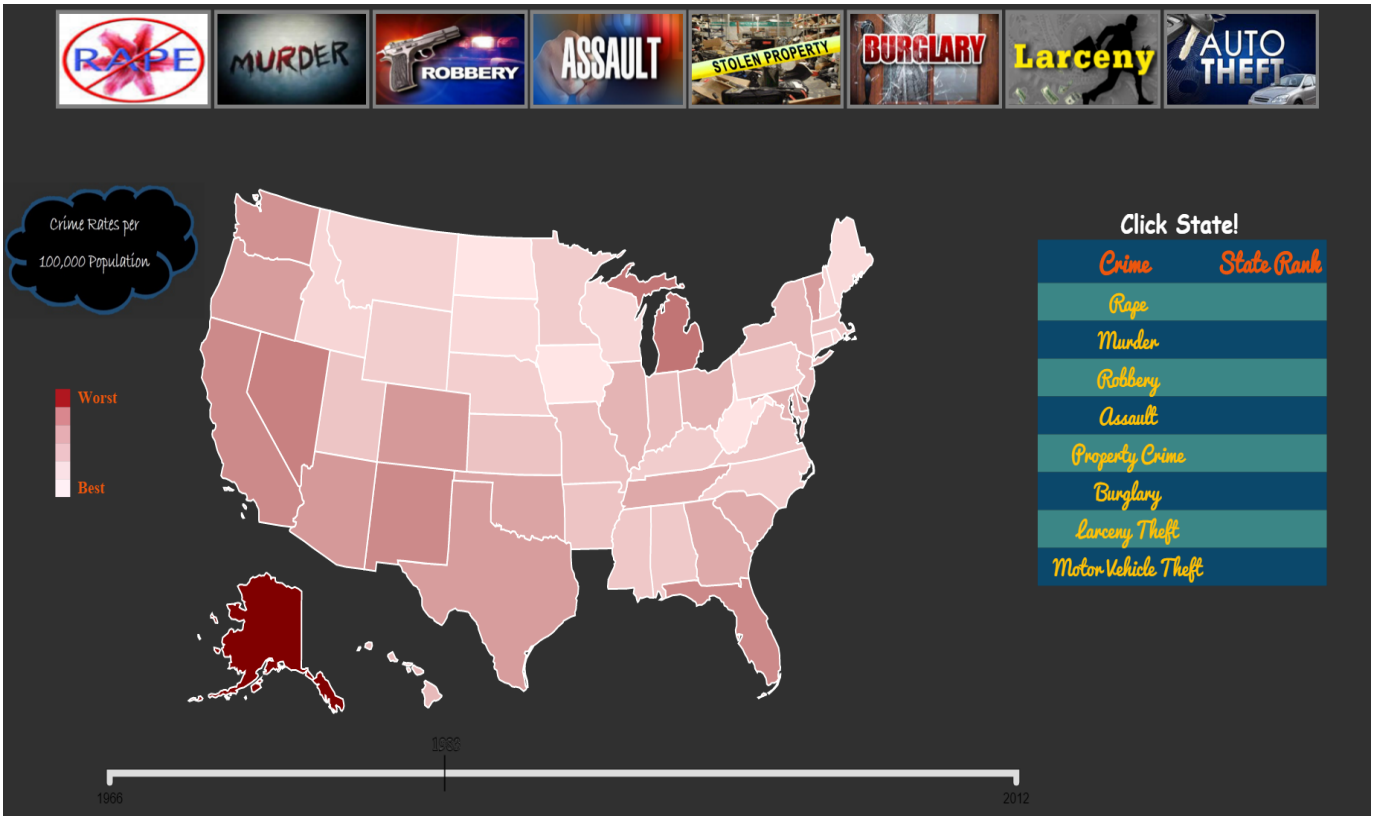
To implement the change in the design of the layout, we had to introduce rank into our dataset. This can be achieved by making a new column in our dataset for each crime category. The new ranking column will be showing the state rank 1 – 51 for a particular year and a crime category. This required grouping of data by year and then sorting it by crime rate. We used R to implement and data processing so that we can go forward with our idea and introduced new column in the crime dataset. The new columns are - rape_rank, MNMR_rank, robbery_rate, assault_rate, property_crime_rate, burglary_rate, larceny_theft_rate and motor_vehicle_theft_rate_rank.

4. Implementation

To make the new design change we wanted to add an html table in our viz. We have added legends, converted radio button to button with images and also added a cloud which explain the crime rates is defined as the crime per 100,000 population. In this week we have tried to implement lot of changes in our design, styling and over all appearance of the viz. Also, we have worked on improving the interactivity.

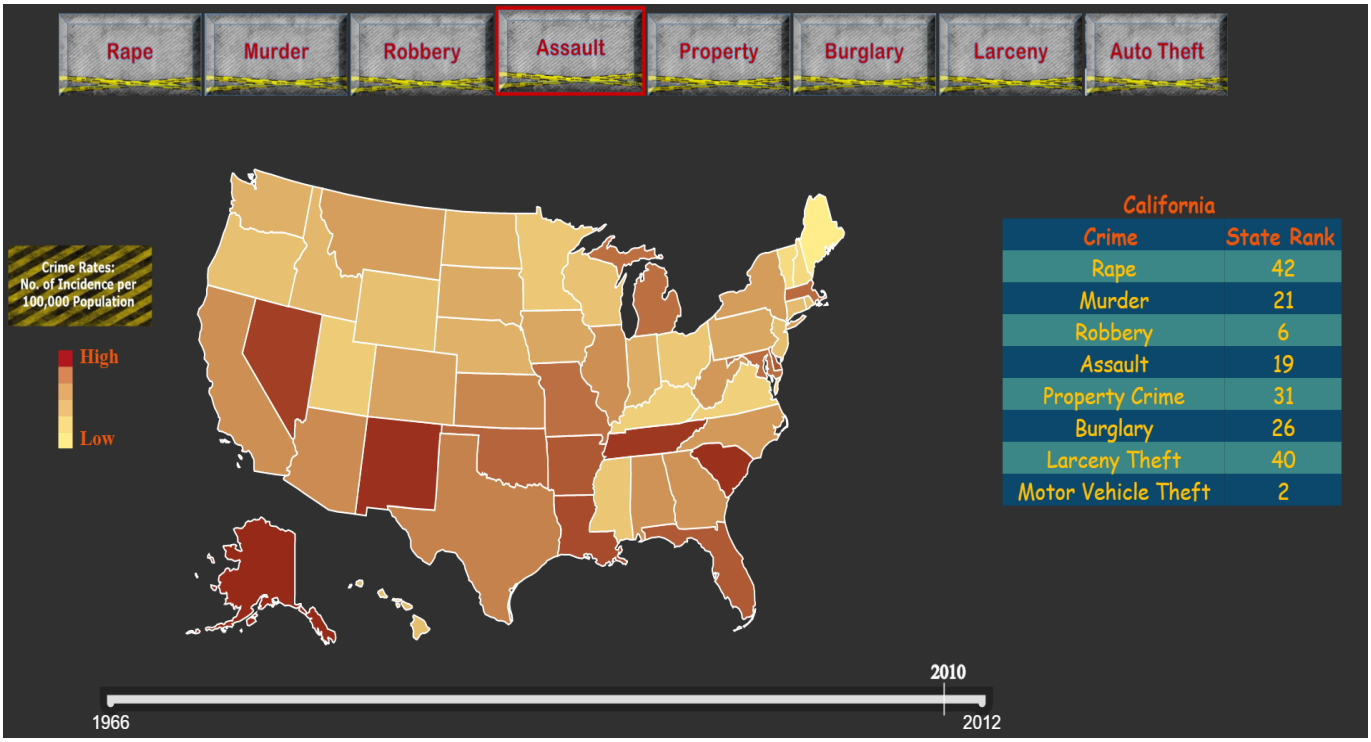


The white text “Click State!” instructs the user to select a state for the table to be loaded with the rank. Once a user clicks a crime category the border of the category images change to red showing the category selection.



Our team felt that, even though we had replaced simple radio button with images; the effect was not appropriate and gave it a very flashy and distractive look to the overall theme. We decided to change these images, the font family of the text in the table to make it less stylish and more readable; and lastly the comment about the crime rate definition. We implemented all these changes in the below snapshot. The slider also needed changes in its background, color of the

text, cross hair and selector; since it was not able to read the domain of the year and the selected year.



We also implemented the introduction page of our viz. It contains the textual description, the link to the process book, screen cast and viz.

THE CRIMINAL STATES OF AMERICA

Crime Report

Violent Crimes - Murder, Rape, Robbery and Assault *Property Crimes - Burglary, Larceny and Vehicle Theft*

The visualization shows crime rates i.e. crimes incidents per 100,000 population across the US States from 1965 - 2012.

Lets try to answer which states have improved in tackling crimes? States which has turned from bad to worse?

This visualization provides state crime ranking from 1 to 51 showing decreasing order of crime rates. Find out trends, state crime rank and compare.

Click on crime category, slide across the years and select states to unravel insights and trends. (Demo: Screen-Cast)

Screen-Cast

Process Book

Viz!

WEEK-4

1. Plan Based on Feedback

Based on the feedback we received during the prototype presentation, we planned to create a new layout which lets the user to compare different states. Another challenge we had is to eliminate one dimension from the data and present the four dimensional data as a three dimensional data. So we decided to combine the different categories as one violent crime data.

The new layout we finalized to implement was using multiple line chart. Since the layout's primary objective is comparison, line chart can depict the story better than other visualizations. The chart will have the years from 1965-2012 in the x-axis and the number of crime incidents in the y-axis. The design in a very important part of this layout since the data needs to be communicated to the user in the best possible way and as interactive as possible.

2. Design

There are plenty of ways to represent a multiple line chart. It is up to the designer to decide best the data can be communicated. We had to consider the theme of the story before building the layout. Since it is about crimes in USA, we wanted to present the layout dark but also keep it simple. Also the user needs to feel that it is a well visualized layout considering the advances in web technologies today.

When choosing the design for this layout, we wanted to compare the different states' crime rates from the years 1966-2012. We planned to achieve this by using multiple line charts and assigning different colors to the lines. This helps the user to analyze and compare between different data. But having plenty of lines in the chart will be too crowded to compare different states. So we came up with a tool tip that can display the currently selected states. This helps the user to visualize and analyze the trends in crime between different states over the years.

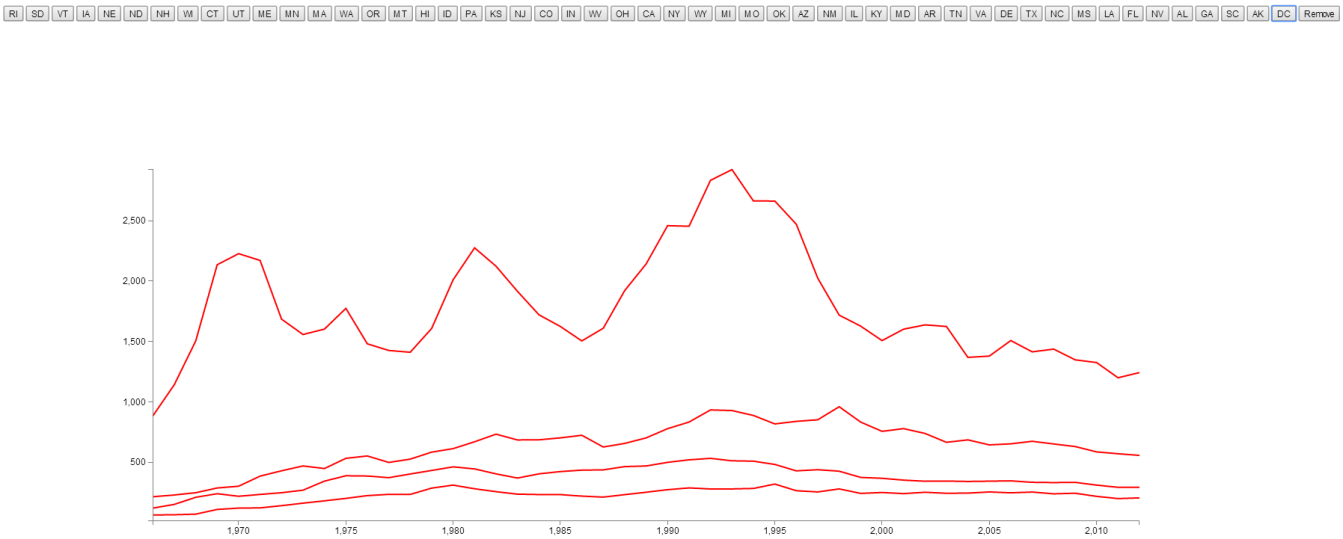
The primary consideration we had to take from the feedback is to convert the four dimensional data to three dimensional. So, we had to eliminate one dimension without hindering the story behind this visualization. Since the major objective of this visualization is to compare different states over the years, we decided to eliminate the different categories of crime and combine it into one 'Violent Crime Rate', since the different categories were already presented in the previous layout.

The layout will basically have a line chart, a check box and a back button to go towards the previous layout with an animated heading for the visualization. We represented the complete layout in a black background to showcase the darkness in the theme of crimes. Also, black layout will represent the different line colors in the visualization better than any other background color. For an eye-catching scroll check box with hover over functionality, we used a standard package from jQuery. For the BACK button and the animated heading transition, we made use of the CSS.

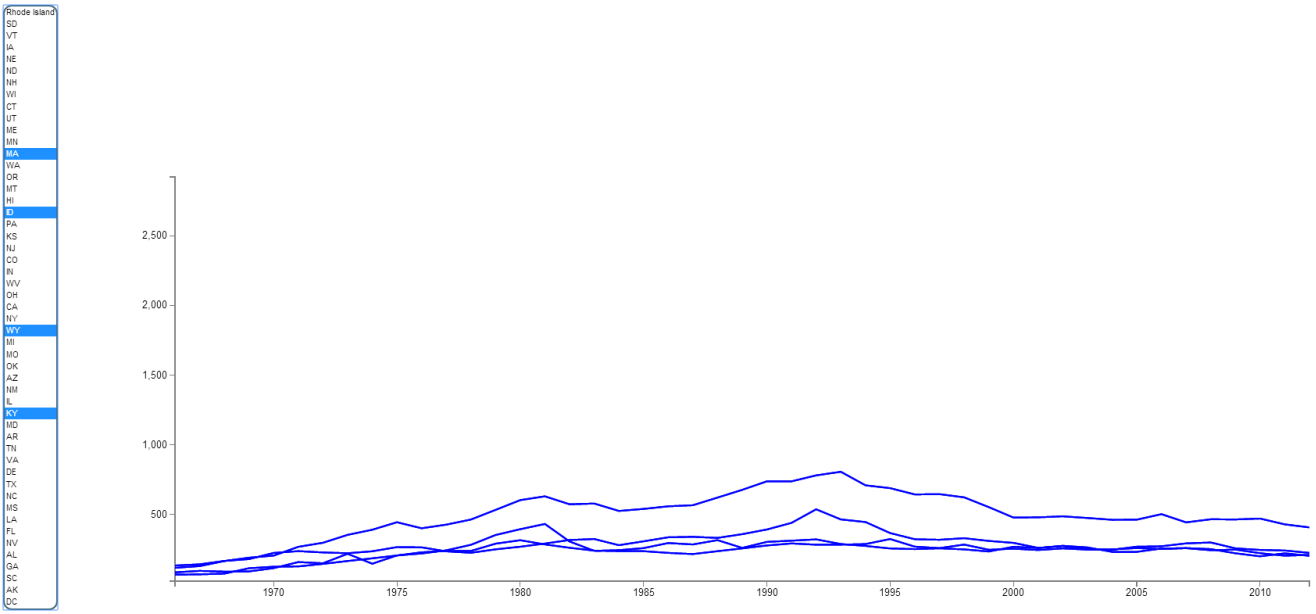
3. Implementation

The initial development of the proposed layout was to bring out a working prototype. So, we implemented a basic line chart which performs the required

functionality of comparing different states. For the first revision of the layout, we opted for selecting different states using multiple buttons. We were able to bring up a working prototype before polishing the layout to final visualization.

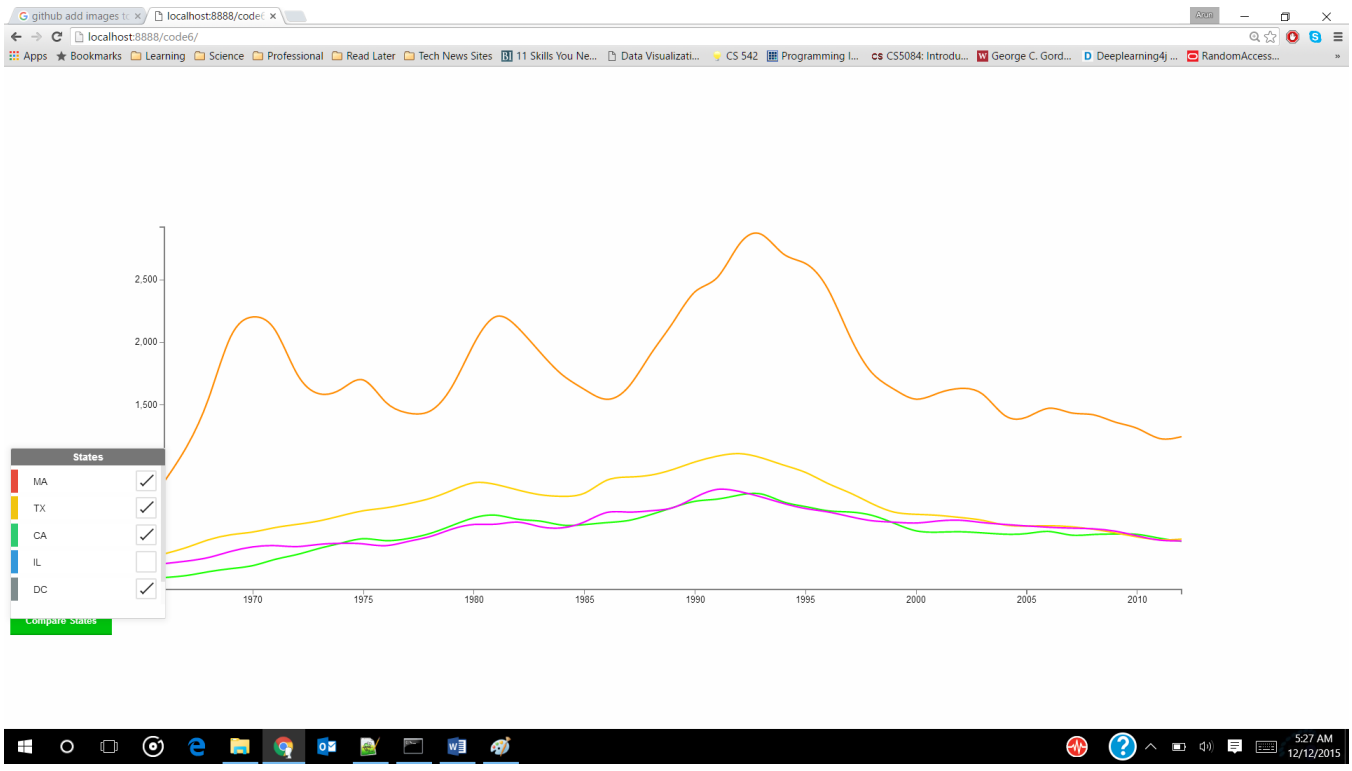


The next move towards our goal was to eliminate the redundant buttons for every states. We came up with a list box. But the problem we faced with this method was that user has to hold the control key in the keyboard for multiple selection of states.



Since we had a working prototype, we had to plan and finalize the best way of representing this visualization to communicate the story of the violent crime data from 1965-2012. We used the web to search different stylish objects. By navigating to some good websites, we got some good ideas that we can use to reflect on our visualization.

We used jQuery and used an package for hover over button with checkboxes ‘hover-dropdown-box.jquery’. This helped us make the visualization appear really good.



To match the theme of the story, we preferred a black background. A black background also represents the different colors of lines in the line chart better that any other background color.

For the line chart, we used thick axes in grey and small dotted grid lines in white. With a black background, this was the perfect combination we could come up with. For the checkbox, we used jQuery to add an inbuilt package ‘jquery-hover-dropdown-box’ which was in white and was a perfect fit for our layout. It provides a hover over functionality over a button and checkboxes for different states. We also added a BACK button to go to the previous layout incase the user wants to. One special effect we used was in the heading. It briefs the story in few words about what this layout is for and animation was used every words with the help of CSS.