Interactive Media Design CA 1

Matthew Roche – X00102929

**Contents**

[**CA Summary** 1](#_Toc433006291)

[**Background** 3](#_Toc433006292)

[**Seven Stages** 3](#_Toc433006293)

[Acquire 3](#_Toc433006294)

[Parse 3](#_Toc433006295)

[Filter 3](#_Toc433006296)

[Mine 3](#_Toc433006297)

[Represent 3](#_Toc433006298)

[Refine 3](#_Toc433006299)

[Interact 3](#_Toc433006300)

[**Problems & Solutions** 4](#_Toc433006301)

[GDP per State 4](#_Toc433006302)

[Pie Chart Issues: 5](#_Toc433006303)

[**Conclusions** 6](#_Toc433006304)

[**References** 7](#_Toc433006305)

# **CA Summary**

This goal of this CA is to provide visualisations regarding various violent crime statistics in the United States for recent years. An area in which this CA aims to provide some level of insight into, is violent crime over time and current trends. In order to achieve this goal, this CA will warrant the use of crime statistics over a number of years in order to demonstrate change and current trends. This CA will also highlight crime as it pertains to different states of the Contiguous United States, for which data showing crime statistics as they relate to specific states will have to be utilized.

Currently Gun Control and Race Equality are large concerns in the United States, for this reason this CA will highlight these concerns as they relate to Violent Crime. The CA will attempt to highlight different weapons types and the frequency with which they are used to commit crimes. It will also look at Violent Crime in relation to race by examining the rate at which crimes are committed by each race and the rate at which each race is a victim of violent crime.

These visualisations will ideally clear up misconceptions and provide an accurate insights into Crime in the US. Through these insights this CA is intended to provide a broader understanding of Violent Crime and the demographics surrounding it.

# **Background**

Violent Crime in the United States has been a very contentious issue over the last decade, according to Forbes mass shootings have been on sharply rising over the last decade, as of the 18th of October 2015 there have been 305 mass shootings in the US.

While this CA takes a much broader approach to Violent Crime than Mass Shootings, they are a globally publicised event, for this reason I felt a look at Crime in the United States would be a very interesting subject to study and attempt to visualise.

When considering this topic over others, I was heavily influenced by the impressive collection of statistics made available to the public by the FBI through FBI.gov. These statistics cover a wide range of topics relating to crime, such as crime by State, County, City etc. as well as crime by different demographics such as race, gender, sexual orientation, religion and more. Interestingly, the statistics provided by the FBI do not just numbers regarding numbers crimes, they also provide statistics such as weapons used, motivations and more.

The ready availability of interesting data sets is what lead me to choose this topic over other choices, I feel that with this information this CA will be informative and relevant to the stated goals in the summary.

There have been various data visualisations done in relation to this topic, many of which utilize the same statistics I opted to make use of. Given the very relevant and highly publicised nature of the topic I have chosen all of what I will demonstrate has at some point been demonstrated before. In order to separate this CA from what has been done before, I will aim to provide interactivity and multiple visualizations which support one another to provide better context.

# **Seven Stages**

## Acquire

I considered a number of different data sources when preparing for this CA, prior to choosing my topic I looked at information made publically available by Amazon through Amazon Web Services’ Datasets Website. This website has a large number of interesting datasets, a number of which I would like to work with at some point.

Another consideration is the publically available datasets published by the Irish Government at data.gov.ie. This was a major consideration as the information available pertains to Ireland and therefore the discoveries I make would have a larger relevance to me personally.

Ultimately I decided to go with the FBI Crime statistics as I found it be a comprehensive and informative set of information which could be worked with easily, as well as for the reasons outlined in the Background.

## Parse

The FBI Dataset is quite extensive, there are a large number of statistics available and from a great number of years. My methodology for categorising this data to be used, involved focusing on one particular area, for example Crime Rate by Race then searching the FBI database for relevant statistics. I would then choose the most recent data for use in the chart. The two main areas I focused on were Violent Crime and Race, as a result I choose datasets relating to the Crime Rate since 1994, Crime Rates in relation to states and two datasets involving offenders and victims of violent crime.

I also had to separately find and tie states and races together by population as this statistic was not available in the FBI datasets, doing so enabled me to get the Crime Rate per Capita statistics. I also looked into using Google’s “Google Refine” software, however I found that the FBI Datasets were already cleaned up and ready for use.

## Filter

The majority of the datasets I utilized included information which was not relevant to the topic I had chosen. The dataset highlighting Murders by State and Types of weapons for example included states which were not part of the contiguous United States (Alaska, Hawaii, etc.) which I removed and recalculated the totals to reflect.

A large amount of filtering went into the Racial Crime Statistics, the original datasets I utilized included many other demographics, such as Religion, Gender, and Sexual Orientation etc. In order to focus upon race as a topic I filtered other demographics out, this also required a recalculation of totals.

The Racial Crime statistics also require filtering to remove property based crime from the dataset as this CA’s focus would be upon violent crime. As a result of this I restructured the data and recalculated the totals to accurately reflect upon violent crime only.

I separated the weapons statistics out from the states in order to be able to generate a Bubble Chart to focus entirely upon them.

## Mine

As I mentioned in the Filtering area a lot of the charts had to be formatted to reflect new totals as data was removed. To do this I added formulas to the excel tables to automatically sum up new totals as changes were made. This was useful several times, particularly in the case of the Racial Datasets which went through many iterations to remove statistics from groups such as Native Americans/Alaskans, Pacific Islanders and multiple/other races.

I also utilized some data mining in the representation stage of the Crime Since ’94 chart, I used Tableau to generate a trend line based on the average rate of crime over the years. Similarly I used Tableau calculated fields to generate crime statistics per 100,000 people where relevant.

## Represent

## Refine

Improve the basic representation to make it more visually engaging

## Interact

Add methods for manipulating the data / controlling what features are available

# **Problems & Solutions**

## GDP per State

One of the areas in I ran into problems was attempting to contrast the murder rate per state versus the GDP per state. To do this I created a map diagram and used each state’s GDP as a colour highlight, this can be seen here:

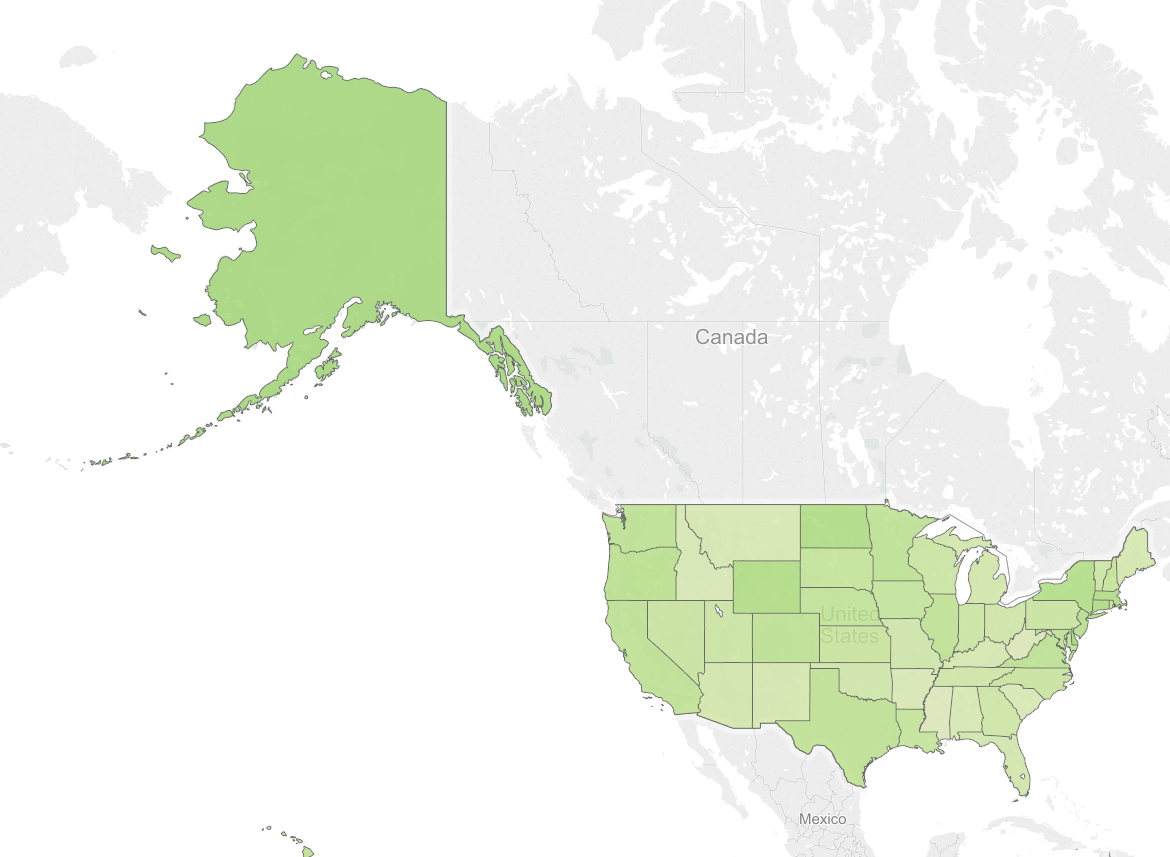


Figure 1: GDP per State

I opted not to use this visualization however, as I did not feel there was a strong correlation between GDP and murder rate and as such I did not think this visualization fit with the other visualizations I was presenting.

Instead of looking at wealth inequality as it pertains to crime in the United States, I opted to look at the race issues, which I felt fit with the other data I am displaying. Luckily there is a large amount of data available on this topic.

## Pie Chart Issues:

Another issue I encountered was the use of a pie chart to represent types of weapons used in crimes for the year 2012. I found that this chart was unsuitable as the labels would not display correctly, it also wasn’t entirely clear which slices represented which data. This can be seen here:

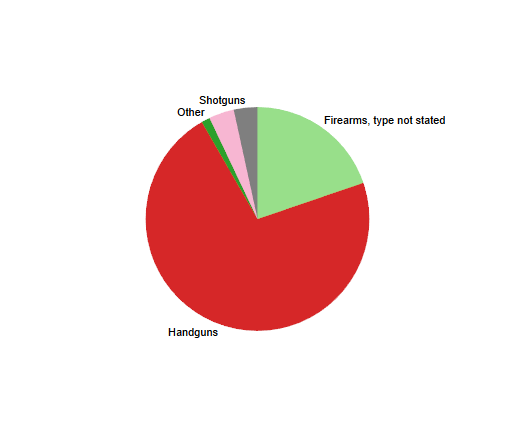
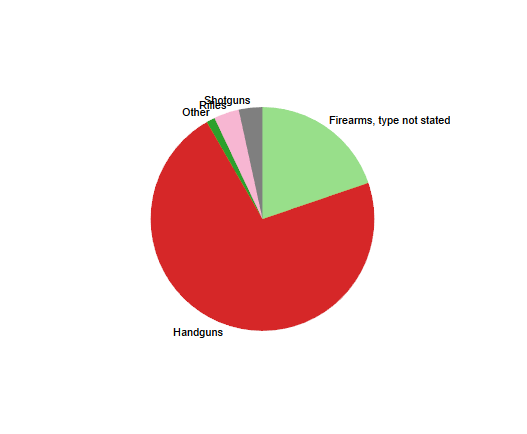


Figure 2: Pie Chart Label Omission

Figure 3: Pie Chart Label overlapping

This can be worked around by allowing Labels to overlap one another as demonstrated in figure 3, however I felt this clutters up the diagram making it difficult to read at a glance.

Given this, I decided that the pie chart would not be a suitable visualization to display this data, in the final version of the CA I opted to us a Bubble Chart. The Bubble Chart has the advantage of displaying difference through size, while also allowing labels to display accurately.

# **Conclusions**

# **References**

Diamond, D., 2015. *Mass Shootings Are Rising. Here's How To Stop Them..* [Online]   
Available at: http://www.forbes.com/sites/dandiamond/2015/06/18/charleston-deaths-are-an-american-tragedy-mass-shootings-are-rising/  
[Accessed 12 October 2015].

Mass Shooting Tracker, 2015. *Mass Shooting Tracker.* [Online]   
Available at: http://shootingtracker.com/wiki/Mass\_Shootings\_in\_2015  
[Accessed 18 October 2015].

[Figure 1: GDP per State 3](#_Toc433006070)

[Figure 2: Pie Chart Label Omission 4](#_Toc433006071)

[Figure 3: Pie Chart Label overlapping 4](file:///C:\Users\Matthew\Desktop\EAD_Labs\IMD\Data\Document\Interactive%20Media%20Design%20CA%201.docx#_Toc433006072)