
STEVENS
INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

# Informatics for Engineering Management

EM 624 Fall 2016 – Final Project

# **Law Enforcement Deaths Analysis**

Authors

Chay Pimmanrojnagool

Date

December 3 ,2016

## **Research Questions:**

This project is to analyze the Statistical Data of Law enforcement death and whether it has a correlation with any of related factors or not. Consequently, we have to look for the second set of data which might be an attribute of the deaths. We choose 2 factor for the analysis. First, the economic factor which indicated by Unemployment Rate and GDP Growth Rate, this factor represents the burden of the American citizen and decay of the society. The second is Crime Factor, which indicated by Crime Rate, which can identify the risk of the officer at that time. Thus, we have stated the questions below in order to answer the question.

1. What the Police Deaths Data Told us from 1900?
2. Is there any relationship between Deaths of the law enforcement and Crime Rate?
3. Is there any relationship between Deaths of the law enforcement and GDP Growth?
4. Is there any relationship between Deaths of the law enforcement and Unemployment? Rate

## **Research Motivation**:

Currently, the controversy about the civil right and the police brutality are one of the society main topic. There are a lot of the debate about the right of the victim and the safety of law enforcement. Consequently, 2 organization occurred, Black Life Matter and Blue Life Matter. The Blue Life Matter support the police action because they concern about the safety of the police. For the civilian, it is hard to understand the situation of them and we might curious about which are the factor that might affect the police safety. Thus this project was created to analyze some of the "might-be-related" factor whether they have an effect to the law enforcement safety or not. The result will clarify some information to the police about their safety.

## **Dataset Description:**

We have obtained 4 Data set

1. Police Officer Deaths
   The Dataset is a record of police officer deaths in the US between 1970 -2016. The Dataset recorded Name, department, working shift, course of death, location, and time.
   URL: https://www.kaggle.com/five-thirty-eight/police-officer-deaths-in-the-us

2. US. Unemployment Rate
   The Dataset of National Unemployment Rate between 1950-2016 from the Bureau of Labor and Statistic
   URL: http://data.bls.gov/pdq/SurveyOutputServlet

3. US. GDP Growth Rate
   The Dataset of United States' GDP Growth Rate between 1960-2016 from The World Bank ,Data bank
   URL: http://databank.worldbank.org/data/reports.aspx?source=world-development-indicators&Type=TABLE&preview=on

4. National Crime Rate
   The Dataset of National Crime Rate between 1950-2012 from FBI Uniform Crime Reporting Statistic
   URL: https://www.ucrdatatool.gov/Search/Crime/State/RunCrimeStatebyState.cfm

Due to there is no a single dataset that combine all of required data thus we have to down load them from the difference sources which the most burden is to find the dataset that have been recording for a long period and genuine. Fortunately, those Datasets are considered as reliable due to all of the record came from government organization.

**Data Preparation:**

The Data Sources allow the user to download the file in a XLMS or CSV format except Unemployment Rate data which display the CSV format, thus the user have to copy the content to the text file. Another problem is the dataset have the difference recording period, so we to perform an inner join before doing an analysis. Cleaning data is necessary because all of the data have a difference data structure and irrelevant content. Most of the cleaning have been done using python but deleting small vignette or some description text have been done manually

Police Officer Deaths Data has the longest recording period (1791 - 2016) and very clean but there are columns that no need for calculation and it has been recording by occurrence (each row for each death), thus we have to cut out the irrelevant columns and when we want to categorize, count() is necessary. Also we will reduce the data period into 1900 - 2015

After copy the Unemployment rate data from the webpage to text file, we can import text file to DataFrame directly but it has been recording in monthly thus we have to find an average for each year before analyzing.

Furthermore, US GDP Growth Rate is the most unstructured data, it was created in column dimension and the content was a mess, thus we needed to transvers data and performed data cleaning. In the others hand, National Crime Rate is the most structured and cleanest Data.

Finally, in order to analyze the data, we need to use Merge function to make the 2 dataset in the same recording period and 2016 data will be cut out due to incompletion, thus the questions will be analyzed in 1900-2015 ,1950-2012 ,1960-2015 and 1950-2015 relatively

**Methodology**:

We use python for doing Data Cleaning, Data analysis and Data Visualization. First, we look at the Police Deaths Data, the Data range is start from 1791 but too old record might give fault information. Moreover, the changes in modern era when American Pop culture was flourish seem to be more reasonable to be focused, thus we choose to extract the Data after 1900.

First, we look for some basic information from the police deaths data. In Figure 1, The graph displays the Number of police death in each year which we can see that there are 2 peaks point. According to additional research, the first increasing, start around 1920 and end 1933, was effected by the Prohibition Era, which Mafia rose the power through its success in the illicit liquor trade, they used all the power and made a war on cops to take control of this business. Another interesting piece of information is, in October 29, 1929, the United State faced the biggest market fall down in the history, and create the Great Depression from 1929 to 1939, but the number of death fall down after the end of Prohibition Era, that might indicate that the economy factor might not affect the deaths.

The second rising was around 1968 to 1975, The one of the historical explanation is, in October 27, 1970, Congress passes the Comprehensive Drug Abuse Prevention and Control Act of 1970. Richard Nixon declared war on drug which become public United State Public enemy number 1.

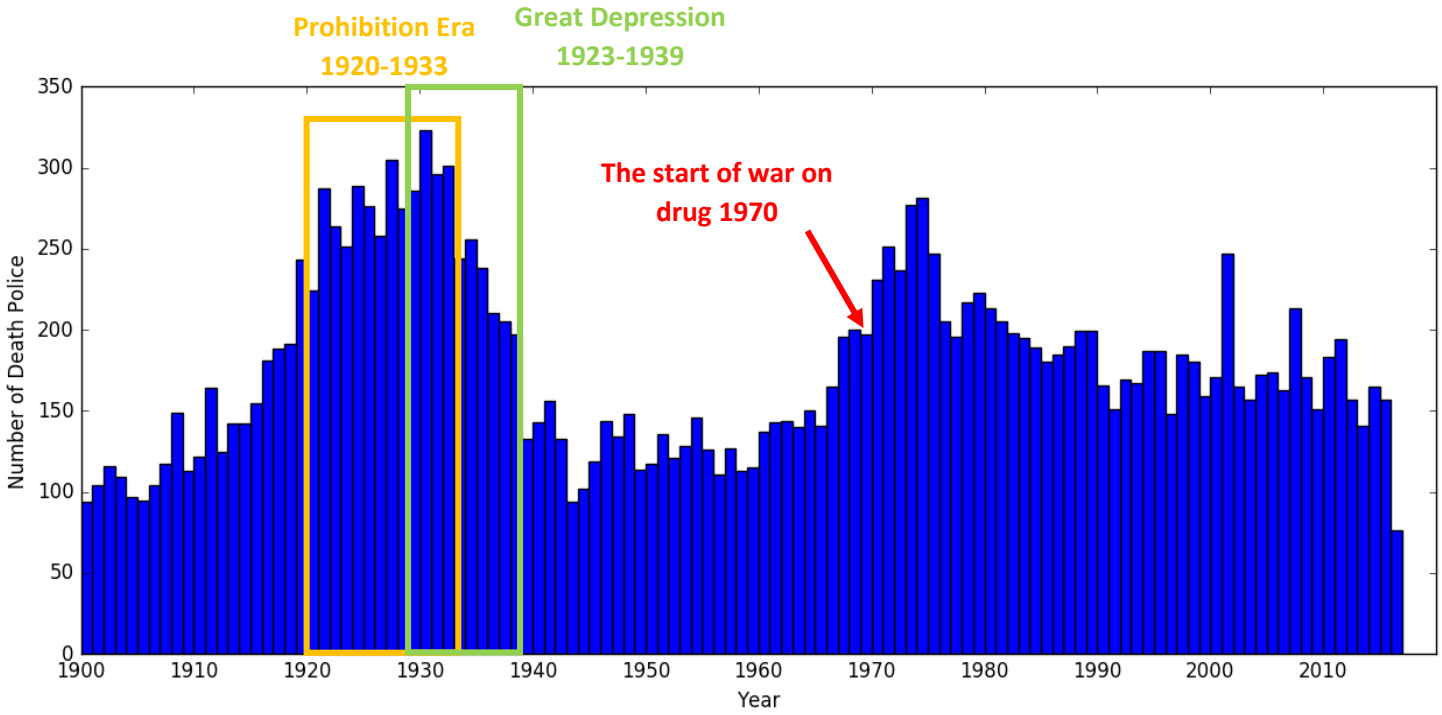



**Figure 1**

Next, we plotted the category. In figure 2, we categorized number of police deaths by cause of death, the graph indicates that Gunfire is main cause of the deaths.

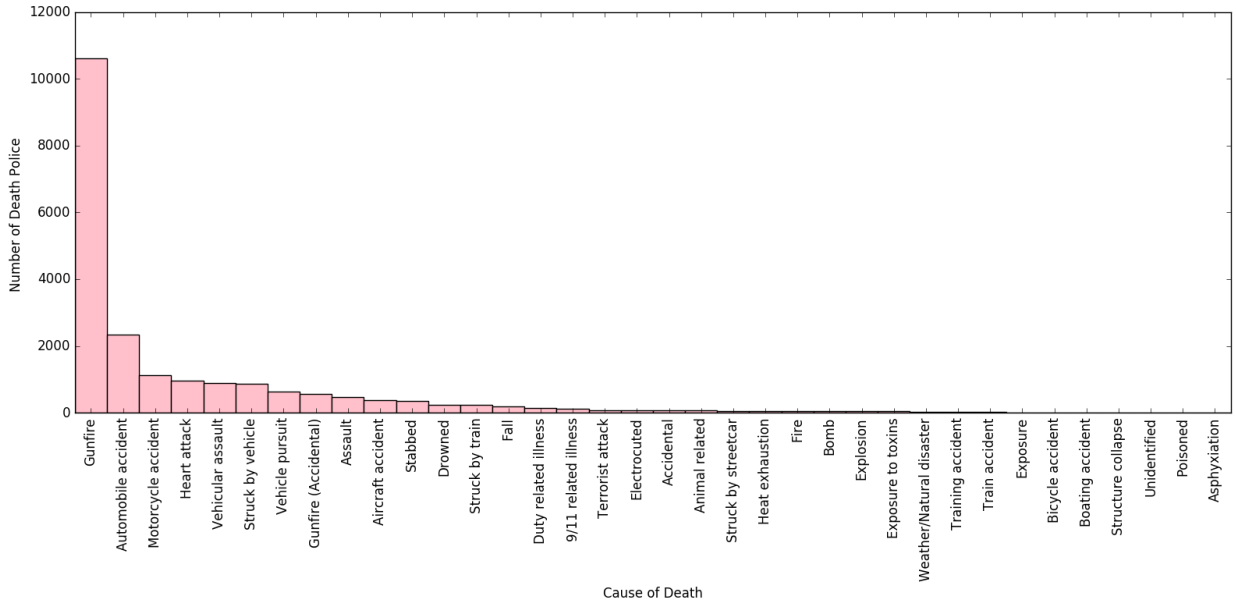
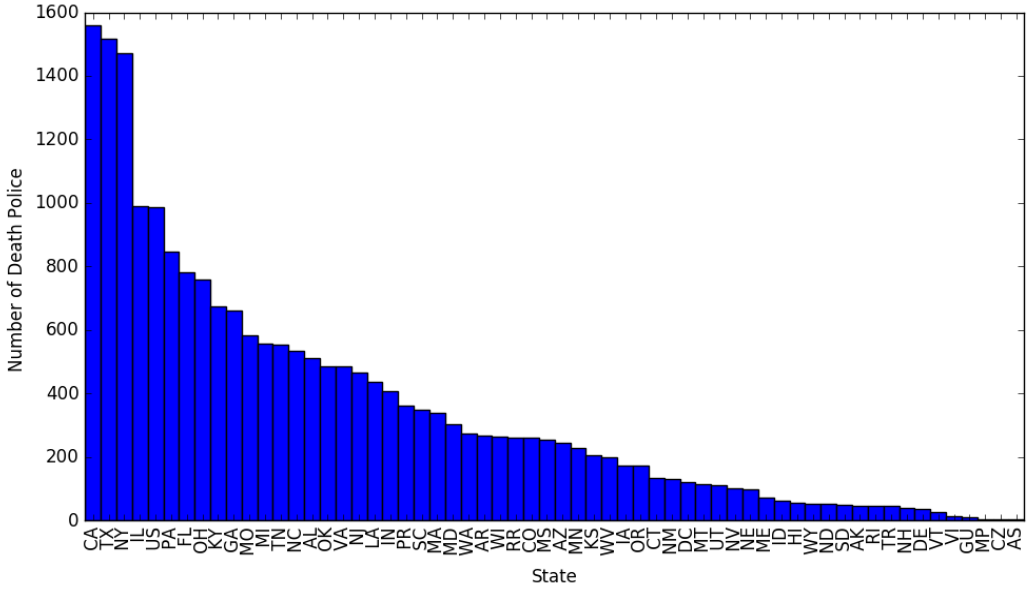



**Figure 2**

In figure 3, we plotted the deaths by state and it shows that California, Texas and New York has the highest number of police deaths.




**Figure**

Then, in order to find the relationship between the number of death and crime rate, we use US. Crime Rate Data from FBI Database. The crime rate can be categorized into 2 section, Violent Crime Rate and Property Crime Rate. Although, the obtained data has range only between 1960 to 2012. Thus, in order to analyze the data, we have to extract the police deaths data in the same period and merge it to Crime Data.

From the graph, now we can see that the crime rate started to rise since middle of 1960 and the rate fall in the mid 90's. According to historical research, there were a lot of change and events between this period (1965 – 1995). First, there was a conflict between race, the civil right become the main issue, African-American Civil Rights Movement occur in 1960-1970, Civil Rights Act of 1964. Moreover, America was in the middle of Vietnam war (1954 –1975) and facing the Cocaine Outbreak. The drug dealer in Los Angeles converted the powder of cocaine into "crack," a solid smokeable form of cocaine, that could be sold in smaller quantities, to more people. It was cheap, simple to produce.

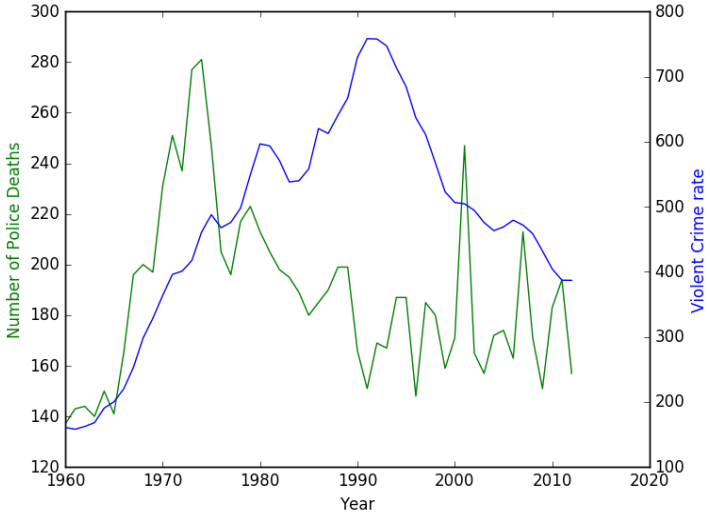



**Figure 4**

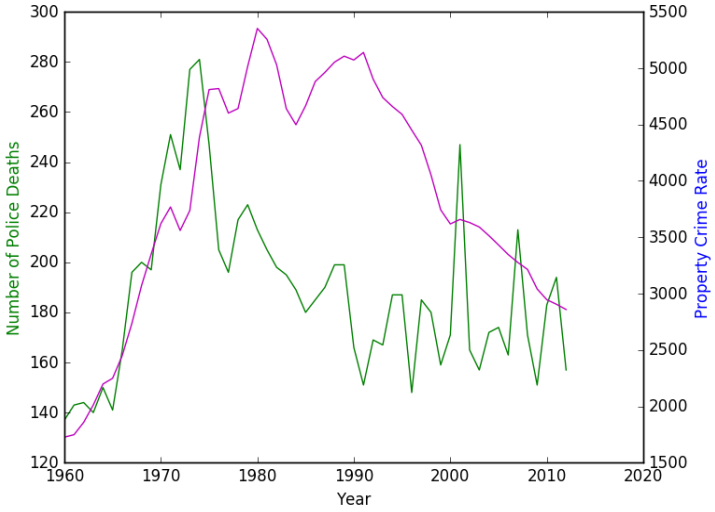



**Figure 5**

Even though, according to the figure 4,5 and 6, the number of police death don't move in the same direction with the violent crime rate nor property crime rate nor all-crime rate. They upsurge at the same time in the mid 60's but after 1970 the number of police death started to fall down despite the crime rate which was still high. Thus, statistically there is no relationship between Crime Rate and The number of police Death in the big picture, but, according to the rose in the mid 60's to the mid 70's (Figure 6) when the war on drug occurred, the number of police deaths might need additional factors such as government policy to make it move in the same way as crime rate.

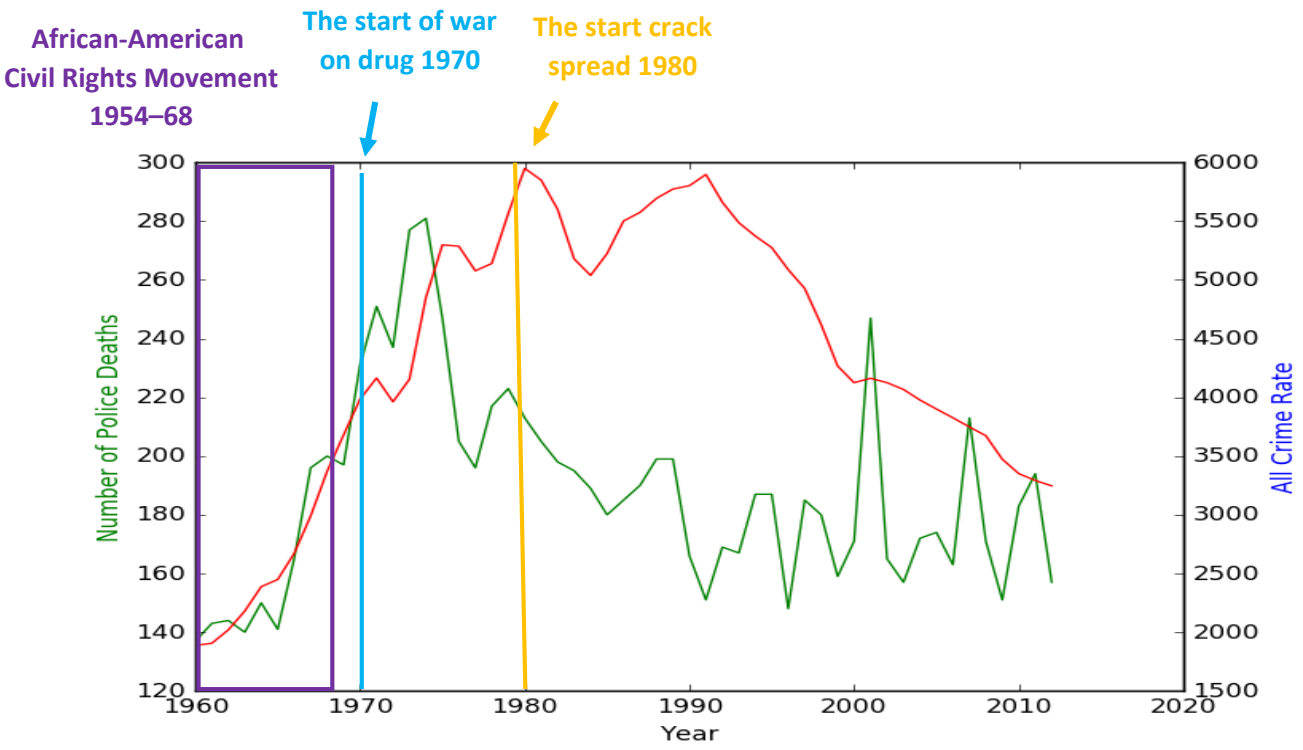



**Figure 6**

The next attribute that might affect the police death is economy. We choose 2 economic indicators to do the analysis, first is the GDP Growth Rate and the second is US. Unemployment Rate. The GDP Growth Rate recorded by The World Bank was calculated from the difference of the GDP in that year compare to the prior year in percentage, and the range of the Data is between 1950 -2015. Thus we need to extract the Police Deaths data in that period of time, and calculate the difference of the deaths in every year in percentage in order to analyze the correlation the changes between GDP Growth Rate and The number of Police Deaths.

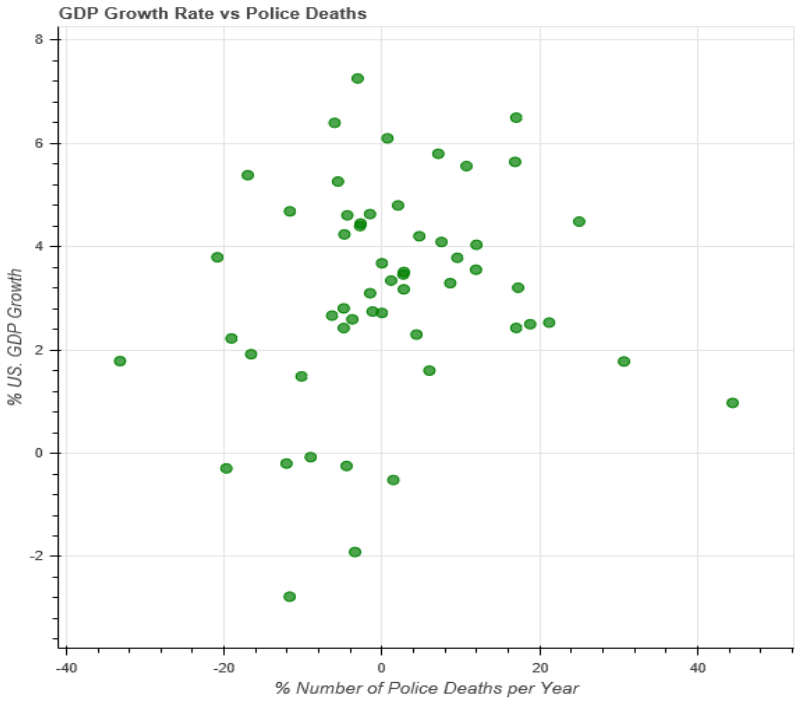



**Figure 7**

According the figure 7 scatter plot, the plot is very disperse and do not form in particular pattern, thus there is no correlation between GDP Growth Rate and The number of Police Deaths.

The second economic indicator is Unemployment Rate. The data given by the Bureau of Labor and Statistic has a range is between 1950 – 2015 and it has been recording monthly. Thus we have to find the annual average before finding the correlation with the number of police deaths.

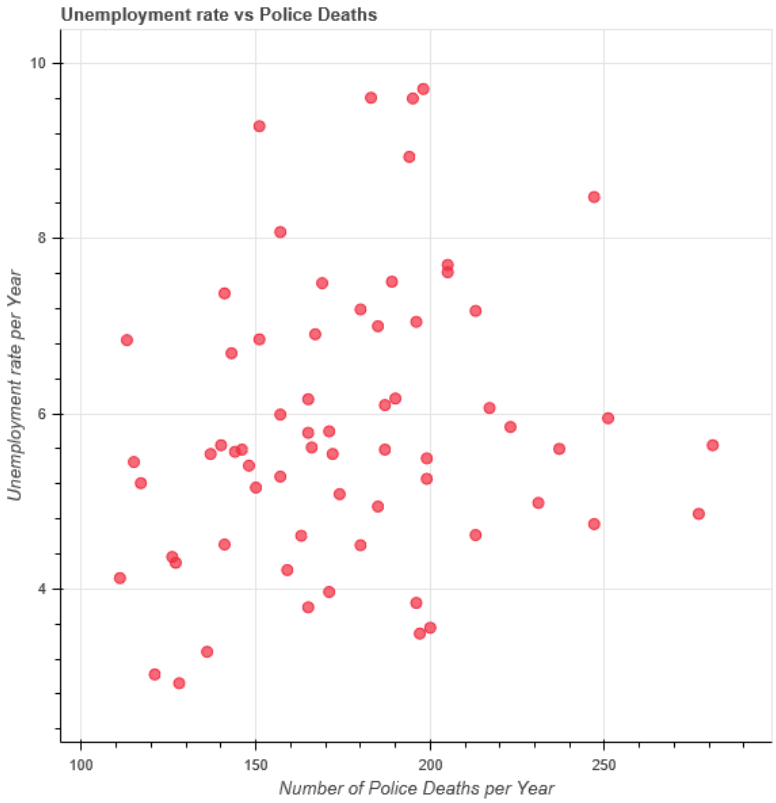



**Figure 8**

According the figure 8 scatter plot, the plot is very disperse and do not form in particular pattern, thus there is no correlation between unemployment Rate and The number of Police Deaths.

**Result**:

Throughout the analysis, now we can answer the developed questions.

1. What the Police Deaths Data Told us from 1900?

According to the analysis for the basic information, the upsurge of the police deaths was occurred two time. The first time was around 1920 and end around 1935 when The Prohibition Era took place and the Mafia Association made war against police to keep the illicit liquor trade running. The second time was between the late 60's to 1975 when US. Declare war on drug.
Moreover, most of the police were killed by gunshot and, California, Texas and New York have the significantly highest number of police deaths respectively

2. Is there any relationship between Deaths of the law enforcement and Crime Rate?

According to figure 4, 5 and 6, There is same movement between late 60's to 1975 but, in the big picture, there is no relationship between the number of police deaths and all crime rate, property crime rate and violent crime rate.

3. Is there any relationship between Deaths of the law enforcement and GDP Growth?

There is no correlation between the number of police deaths and United State GDP Growth rate due to there is no specific pattern in scatter plot (figure 7)

4. Is there any relationship between Deaths of the law enforcement and Unemployment rate?

There is no correlation between the number of police deaths and unemployment rate due to there is particular pattern in scatter plot (figure 8)

Consequently, according to the given datasets, the economy which is indicated by unemployment rate and GDP growth rate are not the factor that affect the number of police deaths. Also the crime rate has specific relationship the death which mean that there is no risk for the law enforcer to work in high-crime-rate circumstance. In the other hand, the historical event or the government policy which create a crucial time such as Prohibition Law and War on Drug seem to be the reason that might affect the deaths.

**Reference**:


http://bluelivesmatter.blue/organization/

http://www.history.com/topics/prohibition

https://en.wikipedia.org/wiki/Prohibition_in_the_United_States

https://en.wikipedia.org/wiki/Crack_epidemic

http://www.drugfreeworld.org/drugfacts/crackcocaine/a-short-history.html

http://iipdigital.usembassy.gov/st/english/publication/2008/04/20080407123655eaifas0.7868769.html#axzz4RoLwdZD4

https://en.wikipedia.org/wiki/1960s

http://www.history.com/topics/vietnam-war/vietnam-war-history

https://en.wikipedia.org/wiki/War_on_Drugs

http://www.drugpolicy.org/facts/new-solutions-drug-policy/brief-history-drug-war-0

https://web.stanford.edu/class/e297c/poverty_prejudice/paradox/htele.html

http://www.history.com/topics/mafia-in-the-united-states