

# DSC520 Final Project – Andrea Fox - 23 Oct 2019

## Section 1 – Week 9 – Getting Started

1. Provide an introduction that explains the problem statement you are addressing. Why would someone be interested in this?
  - a) Watching the news every morning in Kansas City, Missouri it seems like there is always a story regarding gun violence. Citizens are becoming concerned and calling for action . The new mayor, Quentin Lucas, ran on the platform that he would work on decreasing violence. The real issue is that crime in general has increased in the city over the past three years, but the media is using gun violence stories because it is a hot topic nationwide.
2. Draft 5-10 Research questions that focus on the problem statement.
  - a) What is the frequency of gun violence happening in comparison to other crimes reported?
  - b) Are there specific areas that are more heavily targeted?
  - c) Are the crimes reported only involve suspects or does it also include victims?
  - d) Is crime really increasing or just seems like it because of heavier reporting from media?
  - e) Does crime increase at a specific time of day?
  - f) Is crime heavier during specific parts of the year?
3. Provide a concise explanation of how you plan to address this problem statement.
  - a) I plan on addressing this topic by looking at the previous three years of crime reported by Kansas City Police Department. I will look heavily at the crime description and whether the firearm flag has an indication of yes. After gathering, cleaning, and modeling the data I plan to bring this information to the local news so that the city is receiving accurate reporting, and to KCPD so that they can target areas that may need additional patrols.
4. Discuss how your proposed approach will address (fully or partially) this problem.
  - a) If my hypothesis is accurate than it can bring accurate awareness to the actual crime taking place in the city. This allows citizens to be more informed and perhaps feel somewhat safer knowing that not everything happening crime-wise is gun related. A secondary benefit of this study is to bring to KCPD so that they can see what areas may have a heavier burden of crime in comparison to other parts of the city. This will allow them to adjust patrols to areas in need as well as having better predictions on time crime is most likely to happen.
5. Do some digging on a dataset that you can use to address the issue. Original source where the data was obtained is cited and, if possible, hyperlinked. Source data is thoroughly explained (i.e. what was the original purpose of the data, when was it collected, how many variables did the original have, explain any peculiarities of the source data such as how missing values are recorded, or how data was imputed, etc.).

### a) *KCPD Crime Data 2018*

- <https://data.kcmo.org/Crime/KCPD-Crime-Data-2018/dmjw-d28i>
- **Description:** This is the most current information as of the date of upload. This provides the user the ability to view the most current crime information within Kansas City, Missouri. The displayed information is the most current information from the data source as of the date of upload. The data source is dynamic and therefore constantly changing. Changes to the information may occur, as incident information is refined.

While the Board of Police Commissioners of Kansas City, Missouri (Board) makes every effort to maintain and distribute accurate information, no warranties and/or representations of any kind are made regarding information, data or services provided. The Board is not responsible for misinterpretation of this information and makes no inference or judgment as to the relative safety to any particular area or neighborhood.

- **Date Created:** 12-18-2017
- **Last Updated:** 04-19-2019

b) KCPD Crime Data 2017

- <https://data.kcmo.org/Crime/KCPD-Crime-Data-2017/98is-shjt>
- **Description:** *This is the most current information as of the date of upload. This provides the user the ability to view the most current crime information within Kansas City, Missouri. The displayed information is the most current information from the data source as of the date of upload. The data source is dynamic and therefore constantly changing. Changes to the information may occur, as incident information is refined. While the Board of Police Commissioners of Kansas City, Missouri (Board) makes every effort to maintain and distribute accurate information, no warranties and/or representations of any kind are made regarding information, data or services provided. The Board is not responsible for misinterpretation of this information and makes no inference or judgment as to the relative safety to any particular area or neighborhood.*
- **Date Created:** 01-03-2017
- **Last Updated:** 03-05-2018

c) KCPD Crime Data 2016

- <https://data.kcmo.org/Crime/KCPD-Crime-Data-2016/wbz8-pdv7>
- **Description:** *This is the most current information as of the date of upload. This provides the user the ability to view the most current crime information within Kansas City, Missouri. The displayed information is the most current information from the data source as of the date of upload. The data source is dynamic and therefore constantly changing. Changes to the information may occur, as incident information is refined. While the Board of Police Commissioners of Kansas City, Missouri (Board) makes every effort to maintain and distribute accurate information, no warranties and/or representations of any kind are made regarding information, data or services provided. The Board is not responsible for misinterpretation of this information and makes no inference or judgment as to the relative safety to any particular area or neighborhood.*
- **Date Created:** 01-08-2016
- **Last Updated:** 04-19-2019

6. Identify the Packages that are needed for your project.

- a) ggplot2
- b) readr
- c) class

- d) ggm
  - e) pastec
7. What types of plots and tables will help you to illustrate the findings to your research questions?
- a) Scatterplots to get an idea of where my data might be trending. I plan to look at gun violence by year and gun violence in comparison to other crimes on a year to year basis.
  - b) Histogram showing gun violence by month and year
  - c) Histogram showing all crime for on a year by year basis
  - d) Histogram showing crime by area. (haven't decided if I want to look at each year separately or together or both)
  - e) Histogram showing crime rates by time of day
  - f) Histogram of all crimes where a firearm is involved (this may not necessarily be a shooting).
  - g) Chart of predictions based on crime type, firearm involved, time of year, time of day, and area in KC.
8. What do you not know how to do right now that you need to learn to answer your research questions?
- a) If cleaning data is involved how is that done?
  - b) Will I need to split the data and set up train/test like week 8?
  - c) What methods are needed?
  - d) Will I need to set up a few different models and cross validate to verify the accuracy of my hypothesis?
  - e) Is there bias in the data?