

Name: Andrea Fox

Date: 4 November 2019

Title: Section 3 – Week 11 – Final Project

Section 3 – Week 11 – Starting Your Write-ups

- ☐ Discuss how you plan to uncover new information in the data that is not self-evident.
- ☐ What are different ways you could look at this data to answer the questions you want to answer?
 1. First I plan to look at each year's crime and see if the fire arm flag indicates yes. I want to see if on a yearly basis the amount of crime involving firearms is higher than "regular" crime.
 2. I then plan on combining my yearly data into 1 and look at the overall crime involving a firearm to find out whether my problem statement was accurate or not.
 3. I'm considering looking at age and sex to see if those are relevant factors into my data, but someone pointed out in a reply from last week that it might enforce a stereotype bias, which I don't want to do. I'd love more feedback on this one.
 4. I want to look at number of crimes based on time of day. I think I'll do this by looking at everything from 7am-7pm and vice versa or something along those lines. I'm making an assumption that most crime happens at night, but want to verify with the data.
 5. I also want to look at number of crimes by month across each year and then combine into one graph. Are there months with more crime vs others?
- ☐ Do you plan to slice and dice the data in different ways, create new variables, or join separate data frames to create new summary information? Explain.
- ☐ How could you summarize your data to answer key questions?
- ☐ What types of plots and tables will help you to illustrate the findings to your questions? Ensure that all graph plots have axis titles, legend if necessary, scales are appropriate, appropriate geoms used, etc.).
- ☐ What do you not know how to do right now that you need to learn to answer your questions?
 1. How do I fix my margins when plotting crime against zipcode?
- ☐ Do you plan on incorporating any machine learning techniques to answer your research questions? Explain.
 1. I don't think so at this time. Mostly because while I understand the concepts of machine learning I'm not confident enough in trying to incorporate that into my project. I think on this I'd need more practice to be able to make using machine learning more effective.

Suggestion from the course professor: Some additional questions you may want to consider asking yourself as you work through this section of the project:

- ☐ What features could you filter on?
- ☐ How could arranging your data in different ways help?
- ☐ Can you reduce your data by selecting only certain variables?
- ☐ Could creating new variables add new insights?
- ☐ Could summary statistics at different categorical levels tell you more?

- How can you incorporate the pipe (`%>%`) operator to make your code more efficient?