Assignment 8.1

Scott Breitbach

7/23/2020

## Section 1

### Introduction:

Lincoln, Nebraska is the home of the Cornhuskers and every fall, Husker football is a huge draw to the city of Lincoln. When there are home football games, Memorial Stadium, which has a seating capacity of 85,458 and has sold out the last 375 games, essentially becomes the third largest city in Nebraska.

In a city of just under 300,000 people, I would like to see what impact, if any, this huge influx of people has on public safety, including crime and traffic incidents.

### Research questions:

1. Does crime increase on Husker game weekends compared to non-game weekends?
2. Do traffic incidents increase on game weekends compared to non-game weekends?

If yes to 1 or 2:

* Which day of the week has the largest change from baseline?
* Does game time have an impact (morning or evening game)?
* Is there a difference in before-game vs after-game?
* Does the opposing team have an impact?
* Is there an opposite effect on away-game weeks when many people travel to watch the game?

### Approach:

I plan to look at several years of crime and traffic data (at least 2017-2019) as well as Husker football game date and time information to generate crime and traffic incident baselines for non-Husker home game weeks and compare the baselines to Husker home game weeks to evaluate for any effect.

### How your approach addresses (fully or partially) the problem:

This approach will be exploratory in nature, in an effort to determine whether there is a corresponding increase in crime and traffic incidences alongside the increase in individuals to the city.

### Data:

All of my datasets were obtained through the [Open Data and Performance Management](http://opendata.lincoln.ne.gov/) page for the city of Lincoln.

Files:

* [LPD\_2013\_2020\_Arrests\_and\_Citations\_De\_Coded.csv](https://opendata.arcgis.com/datasets/69363105cc3f4f73a3318cafed030dfa_0.csv)
* [LPD\_2017\_2020\_Incident\_Reports.csv](https://opendata.arcgis.com/datasets/dc814856aa6645879c3a0aa7e7d527e0_0.csv?outSR=%7B%22latestWkid%22%3A4326%2C%22wkid%22%3A4326%7D)
* [LPD\_Traffic\_Crashes\_2013\_2020.csv](https://opendata.arcgis.com/datasets/fd7f05be61da45b3b14be8780f7685b2_0.csv)

I may also look at Traffic Stops, though those .csv files are compiled by individual year.

All data was created by the Lincoln Police Department and compiled via the LPD Records Management System. The Incident Report data has 25 variables and 116,263 observations. When loading this dataset, there were 50 parsing failures, so I will have to investigate and see if I can solve the issue, or just exclude those data points if I can’t. The Traffic Crashes data has 18 variables and 24,500 observations, and the Arrests and Citations data has 366,000 observations of 18 variables.

Many of the variables are meaningless to me, but I will primarily be looking at total incidences and the dates and times on which they occurred. I will also need to find or generate my own dataset with the dates and times of the Cornhusker football games, whether they were home or away, and who they played. I will also need to convert to a common date time format between files in order to perform the analysis.

### Required packages:

I will need the following packages

* readr
* ggplot2
* lubridate
* dplyr
* hmisc
* car

I will probably add more as I need them.

### Plots and table needs:

I will need a chart of incidents over time (perhaps by week) could be useful to illustrate whether there are spikes on game weeks. If there is an increase in incidents on game weeks, bar charts showing the average increase by opposing team, day of the week, pre- or post-game, or any other correlations that reveal themselves during analysis.

### Questions for future steps:

I probably have all of the tools I need, though I’m sure I will need a refresher on some of the steps along the way, especially when it comes to converting dates to a common format as well as a way to compare the date ranges between datasets.

One thing that I would like to learn more about is creating my own functions, so perhaps I will get a chance to do that within the scope of this project.