Visualization for County crime against Population and Staffing

Shivani, William, Chaitra

11/6/2017

Scatter plot for Crime in County's of Nebraska against Population of the County

Load the Required Libraries

```
library(ggplot2)

library(ggrepel)
```

Load the dataset

```
county_crime <- read.csv('County_crime_and_population_combined.csv')
lawCrime <- read.csv('lawStaff_vs_Crime.csv')
```

Create a subset for counties with comparatively low population

```
county_crime_with_low_population <- county_crime[county_crime$Population <= 10000 & county_crime$Popula
```

Plot the scatter plot for the county's in the above subset

```
p <- ggplot(county_crime_with_low_population, aes(Population, Total, color = County))

p + geom_point(shape=21, fill="White", size=3,stroke=1.5) + theme(legend.position = "none") + geom_text
```
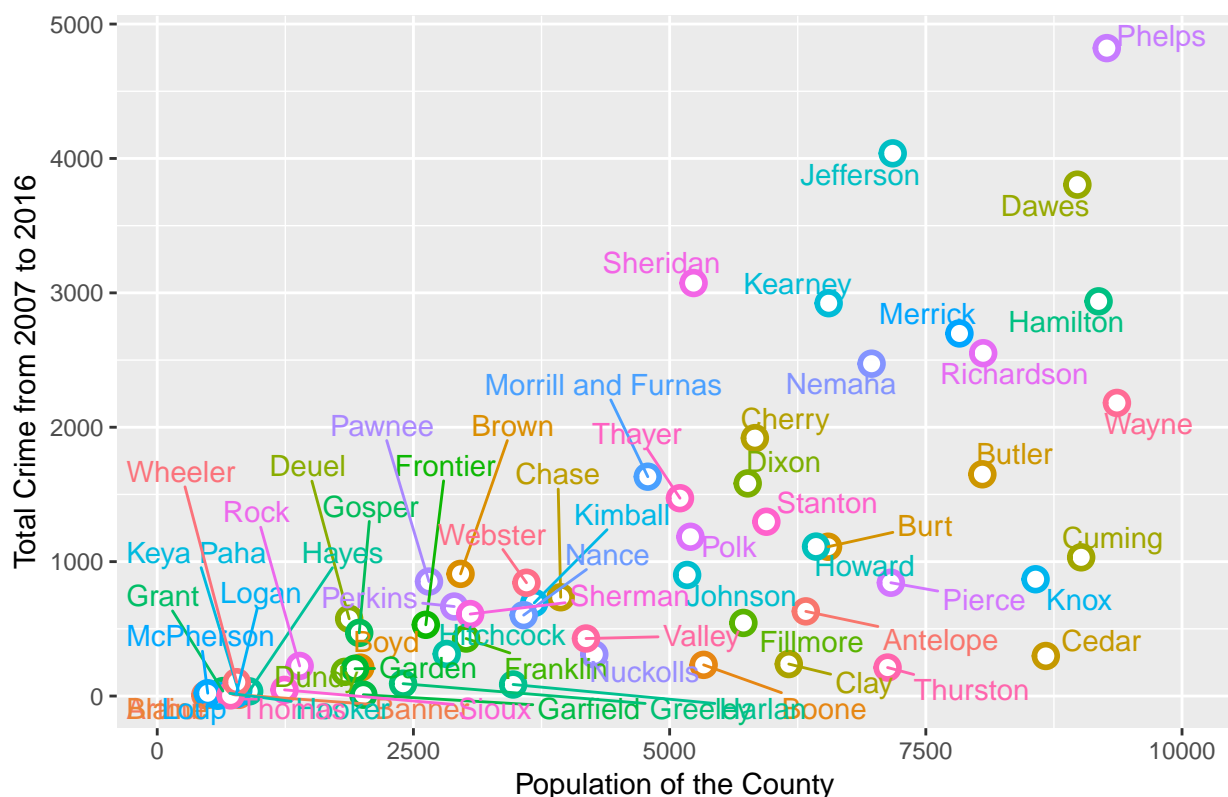
Scatterplot for Total crimes in the Nebraska County against Population
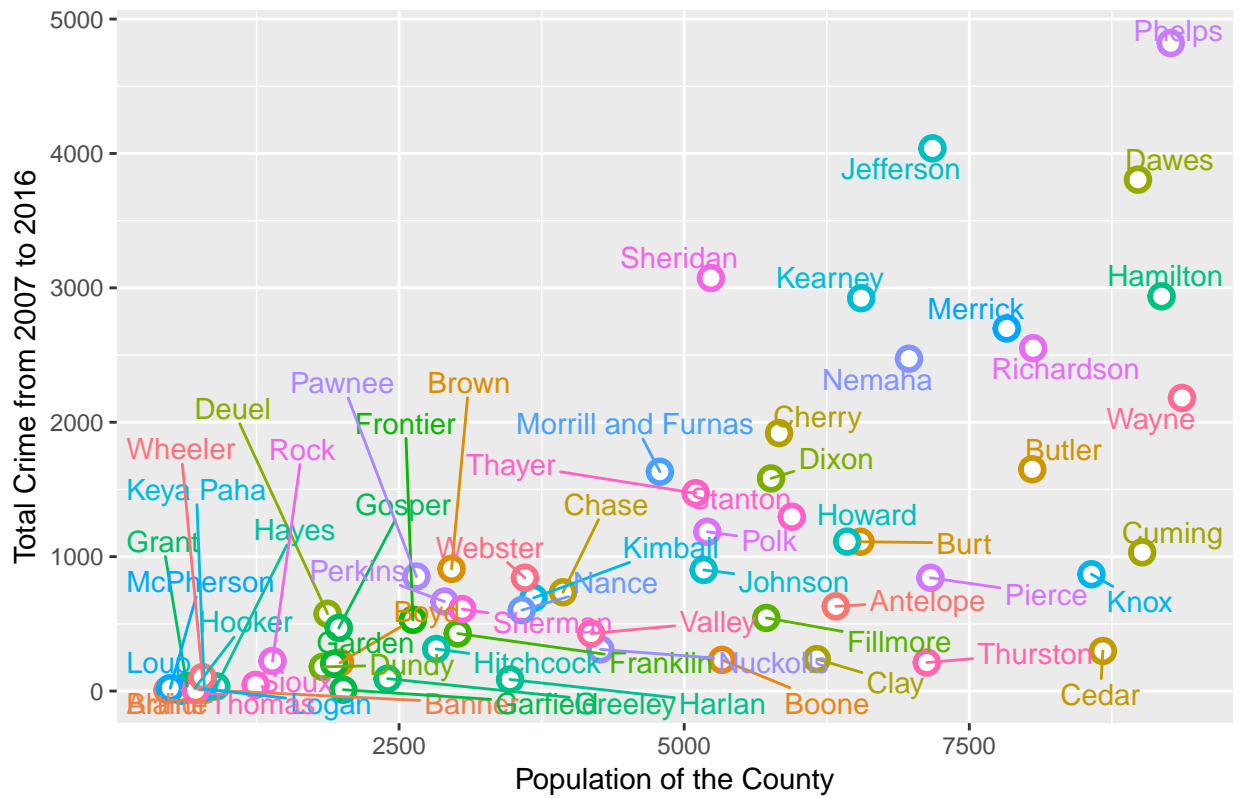
Final Plot with proper title and axis label

```
p + geom_point(shape=21, fill="White", size=3,stroke=1.5) + theme(legend.position = "none") + geom_text
```

Scatterplot for Total crimes in the Nebraska County against Population

Plot for Crime against population of the counties (with population less than 50000) shows many variation. As some of the counties like "Phelps" for an instance represents general idea of higher population and higher crime but then there are counties like Cedar and Knox which show the exact opposite with higher population and very less crime. This provides us an insight to explore other factors than population which might be a major factor on number of crimes in the counties.

Using a smaller set of the data set gives more specific details from the visualization perspective and hence it was better to group a set of Counties and visualize crime data for that particular group.

Crime_vs_lawStaff Part

Data Preparation

See "lawStaff_vs_Crime_Plot" R script ### Set working Directory

```r
lawCrime <- read.csv('lawStaff_vs_Crime.csv')

# Replace null values with 0, from inconsistent data cleaning
columnsToReplace <- lawCrime[,c("us_so_ft", "us_so_pt", "num_res_so",
                                "us_c_ft", "us_c_pt")]
columnsToReplace[is.na(columnsToReplace)] <- 0
lawCrime[,c("us_so_ft", "us_so_pt", "num_res_so",
            "us_c_ft", "us_c_pt")] <- columnsToReplace

# Create a total full time officer column,
# combining male and female records
```

```r
lawCrime$total_so <- (lawCrime$ft_so_m + lawCrime$ft_so_f)
# Same for full time civilians
lawCrime$total_c <- lawCrime$ft_c_m + lawCrime$ft_c_f


#################
### Map Prep ###
###############

# Mapping county zip codes to names
postalCodes <- read.csv('us_postal_codes.csv')
colnames(postalCodes)[5] <- "county"

# Just keep zip codes and county names, and states
postalCodes <- postalCodes[,c(-2,-4,-6,-7,-8)]

# Rows in dataset matching counties in lawCrime and located in NE
indexes <- which(postalCodes$State == 'Nebraska' & postalCodes$county %in% lawCrime$county) #& ])
# Keep only matching rows in the zipcode file
postalCodes <- postalCodes[indexes,]
# Keep only one zip for each County
postalCodes <- postalCodes[!duplicated(postalCodes$county),]


# Create a new dataframe merging LawCrime with county zip codes
lawCrimeMapping <- merge(lawCrime, postalCodes, by = "county", type = "inner")

# Remove state and department columns
lawCrimeMapping <- lawCrimeMapping[,-c(21, 3)]


#!!!!! Map only 2016 data !!!!!#
lawCrimeMapping <- lawCrimeMapping[lawCrimeMapping$year == 2016,]
# Aggregate a number of columns in the event you want to map different things
lawCrimeMapping$county <- as.character(lawCrimeMapping$county)

lawCrimeMapping$pop_covered <- as.numeric(lawCrimeMapping$pop_covered)

lawCrimeMapping <- aggregate(lawCrimeMapping[,c(3,4,5,6,7,8,9,10,12,13,14,17,18)],
                             by=list(county=lawCrimeMapping$county),
                             FUN=sum)
```

Mapping the Data

  See "lawStaff_vs_Crime_Plot" R script

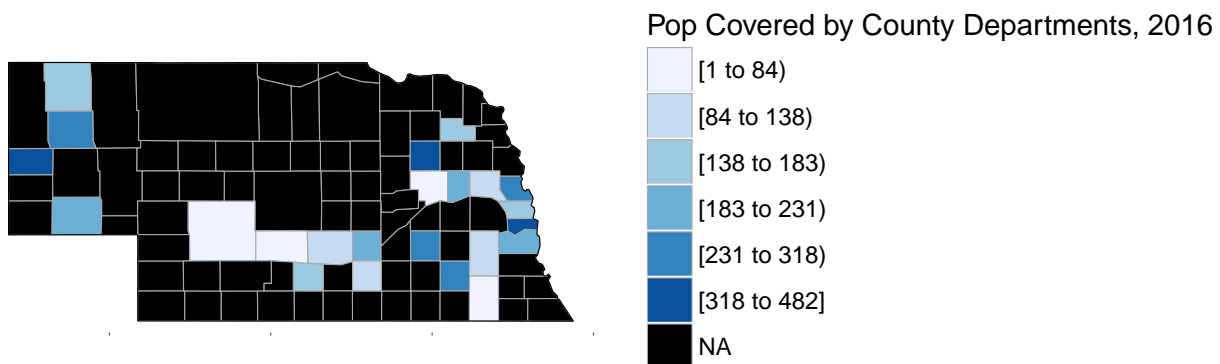**Looking at Number of Sworn Officers by County**

  Pretty interesting to see there are many counties which have no police department, our policing data does not reflect how resources are spread out for each one amongst neighboring counties. Generally, there is one police department per county with 1 exception. In the middle of the state, we see that the cluster of departments are responsible for patrolling the entire North to South region of the state.

Number of Sworn Officers, 2016
- [7 to 14)
- [14 to 16)
- [16 to 33)
- [33 to 43)
- [43 to 140)
- [140 to 848]
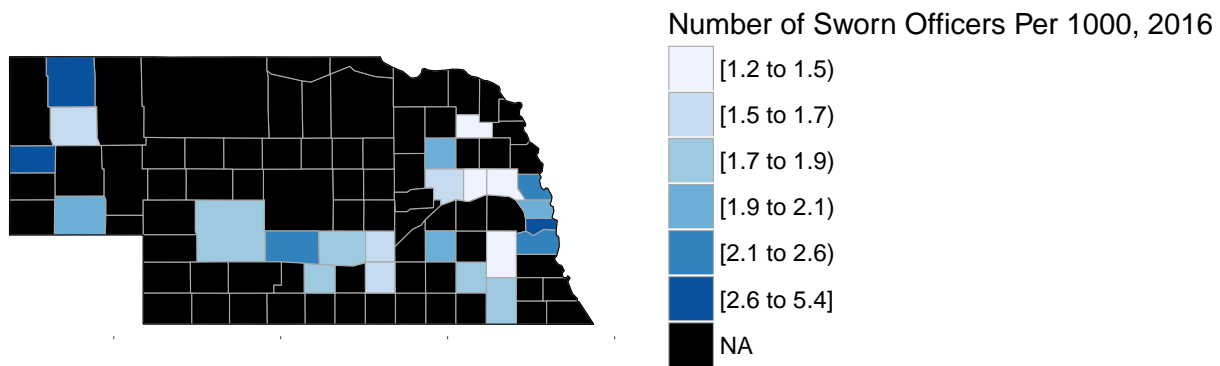- NA
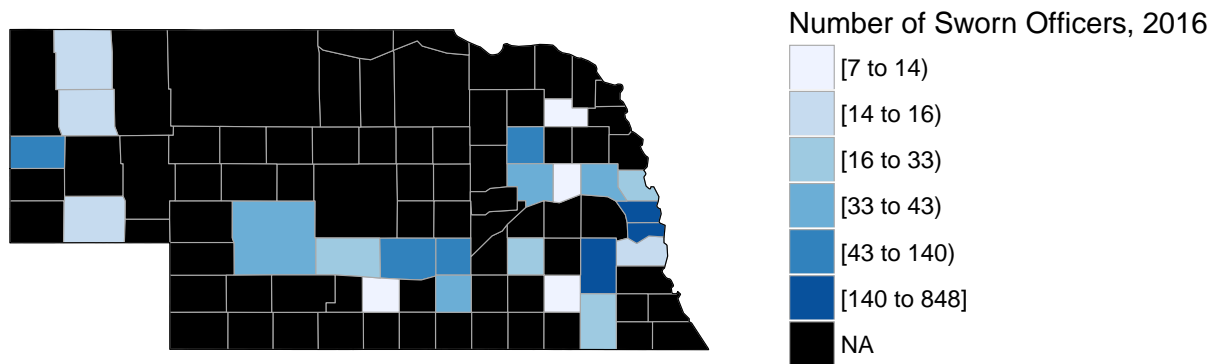
Population covered by County

While we mapped by county, each county represents a department. Here we can see the general population densities of Nebraska, as well as the additional burden placed on those central agencies - as they have a relatively high amount of population to cover given the size of the geographic area they operate in.

county_mapPop

Pop Covered by County Departments, 2016
- [1 to 84)
- [84 to 138)
- [138 to 183)
- [183 to 231)
- [231 to 318)
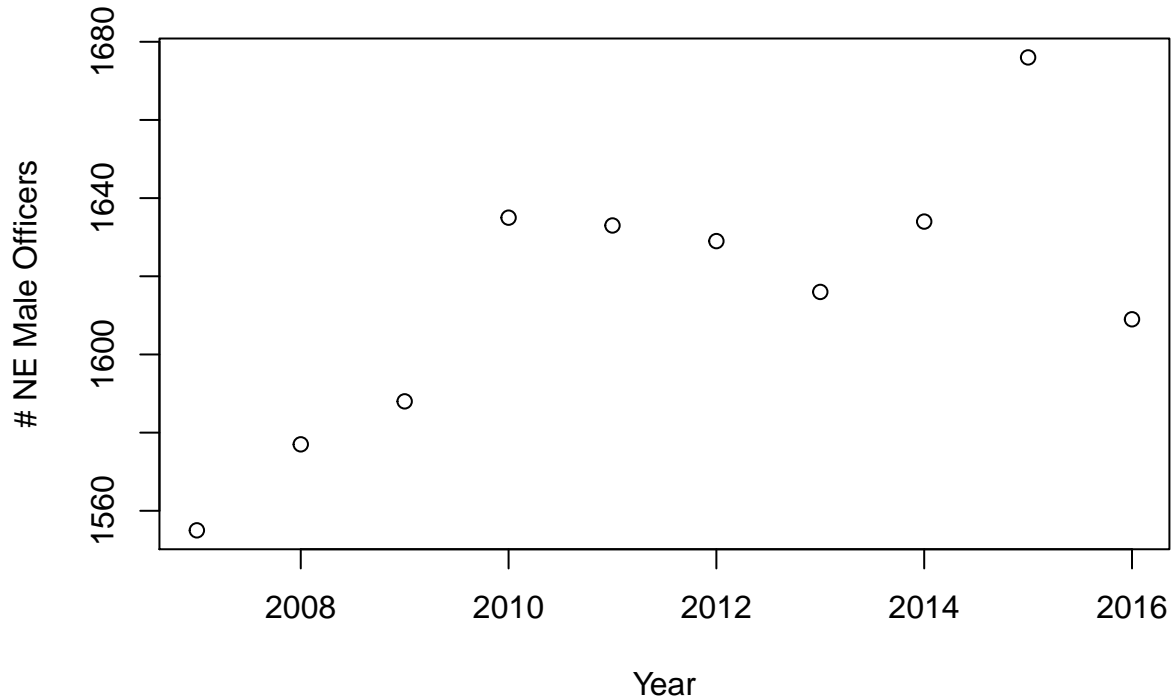- [318 to 482]
- NA

SO per 1000 People by County

This map does not translate well, as obviously departments, while not located in many counties have to go and investigate crimes in counties without a dedicated police department. It is interesting there appears to be lower police presence on the panhandle portion of the state, while the central region enjoys the best officer to civilian ratio.

Number of Sworn Officers Per 1000, 2016
- [1.2 to 1.5)
- [1.5 to 1.7)
- [1.7 to 1.9)
- [1.9 to 2.1)
- [2.1 to 2.6)
- [2.6 to 5.4]
- NA

Annual SO Hiring Trends by Gender

Hiring trends appear consistent statewide by gender, although only 5-10 female officers are being hired or leaving each year. Due to the relative involatile nature of this visualization, I'm not sure it's meaningful in our research.

Male SO Staffing Trends

Female SO Staffing Trends