W205 - Final Project - Streetlights and Crime

Andrew Burrage

April 28, 2016

```
## Warning: package 'ggplot2' was built under R version 3.1.3
```

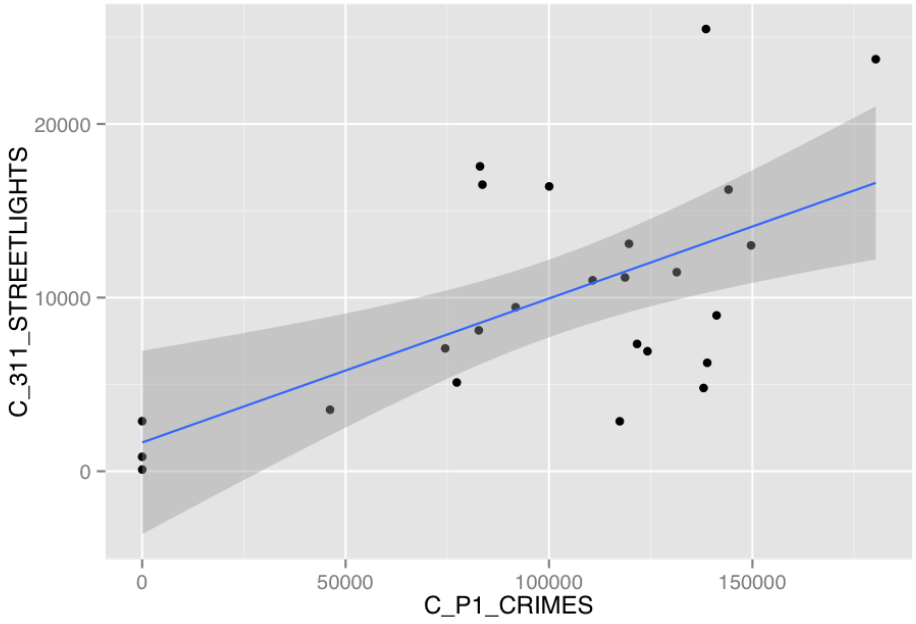
```
## Warning: package 'stargazer' was built under R version 3.1.3
```

```
##
## Please cite as:
##
##  Hlavac, Marek (2015). stargazer: Well-Formatted Regression and Summary Statistics Tables.
##  R package version 5.2. http://CRAN.R-project.org/package=stargazer
```

```
## Warning: package 'reshape2' was built under R version 3.1.2
```

The purpose of this project is to expand upon prior research regarding the relationship between crime and streetlight outages in Chicago. I downloaded crime and streetlight outage data from Chicago and San Francisco open data websites. This R script will hopefully shed some light on where crime and streetlight outages are happening within Chicago and San Francisco and how closure time of a streetlight outage request is affected by the crime rate within the requesting district.

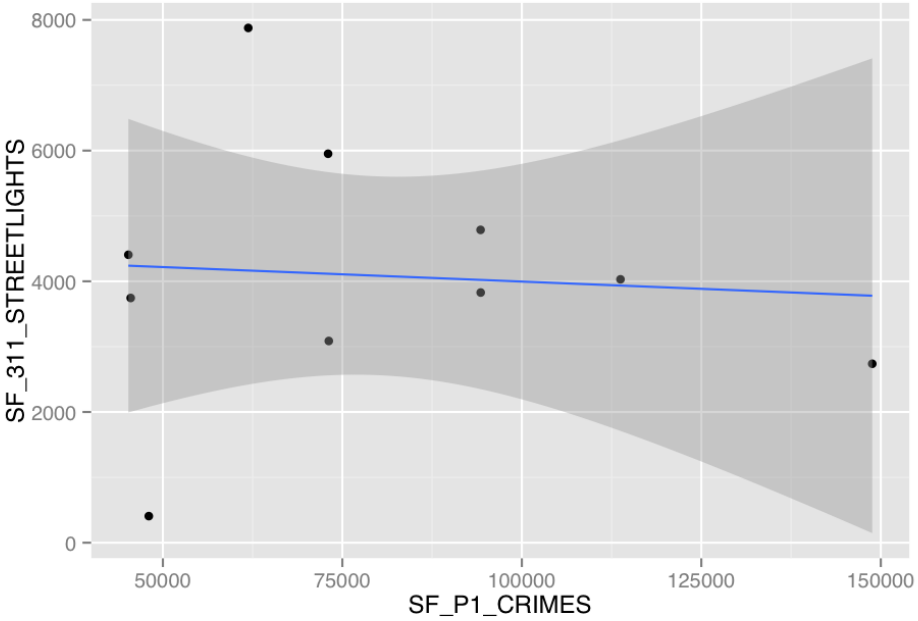
The below scatterplot shows the total number of Part I crimes (Homicide, Rape, Robbery, Aggravated Assault, Burglary, Larceny-Theft and Vehicle Theft) and total number of 311 streetlight service requests.



As you can see in the graph there seems to be a positive relationship between Part I crimes and 311 streetlight service requests in Chicago. To get a better understanding of this relationship I will run a regression to measure the statistical relationship.

```
##
## Call:
## lm(formula = C_P1_CRIMES ~ C_311_STREETLIGHTS, data = chi)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -69581 -28417   5566  32700  60088
##
## Coefficients:
##                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)        57025.914  14387.978   3.963 0.000616 ***
## C_311_STREETLIGHTS     4.355      1.208   3.605 0.001490 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 39150 on 23 degrees of freedom
## Multiple R-squared:  0.3611, Adjusted R-squared:  0.3333
## F-statistic:    13 on 1 and 23 DF,  p-value: 0.00149
```

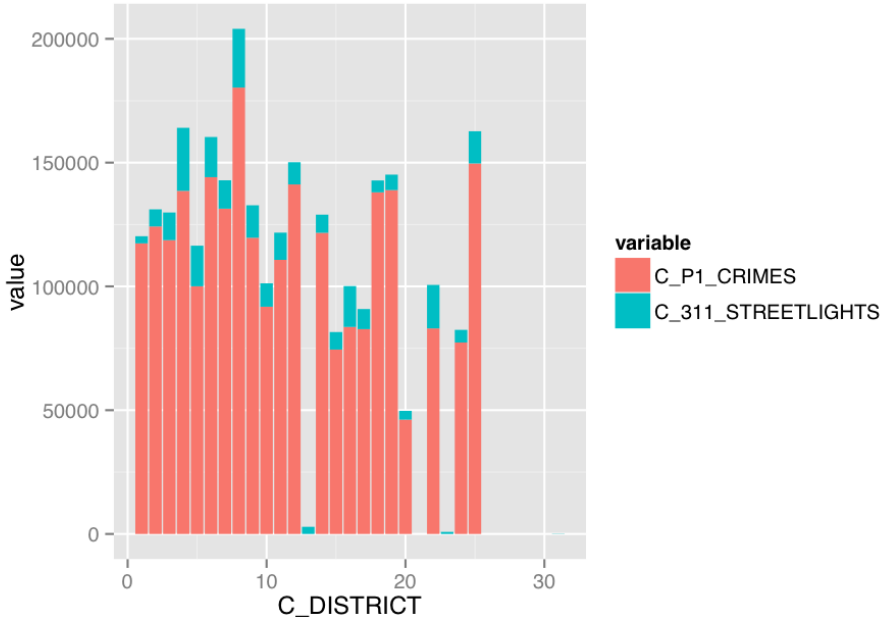
The results show there is a statistically significant relationship between the number of Part I crimes and 311 streetlight service requests. To get a better understanding and see if there may be a relationship in other big cities I will look at San Francisco's Part I crimes and 311 streetlight service requests.



##

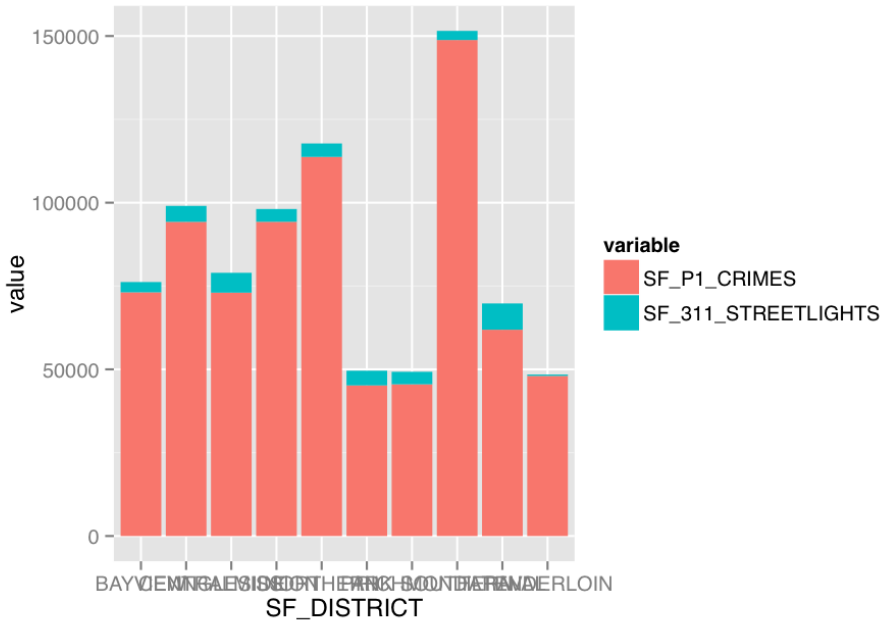
```
## Call:
## lm(formula = SF_P1_CRIMES ~ SF_311_STREETLIGHTS, data = sf)
##
## Residuals:
##    Min     1Q Median     3Q    Max
## -36470 -28891  -6164  15066  67294
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)         85029.711  26950.609   3.155   0.0135 *
## SF_311_STREETLIGHTS    -1.284      5.998  -0.214   0.8358
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 35470 on 8 degrees of freedom
## Multiple R-squared:  0.005699,   Adjusted R-squared:  -0.1186
## F-statistic: 0.04585 on 1 and 8 DF,  p-value: 0.8358
```

The scatterplot does not show a relationship between Part I crimes and the number of 311 streetlight service requests in San Francisco. The results of the regression show there is not a statistical relationship between Part I crimes and the number of 311 streetlight service requests in San Francisco. There is a significant difference between the data from San Francisco and Chicago.

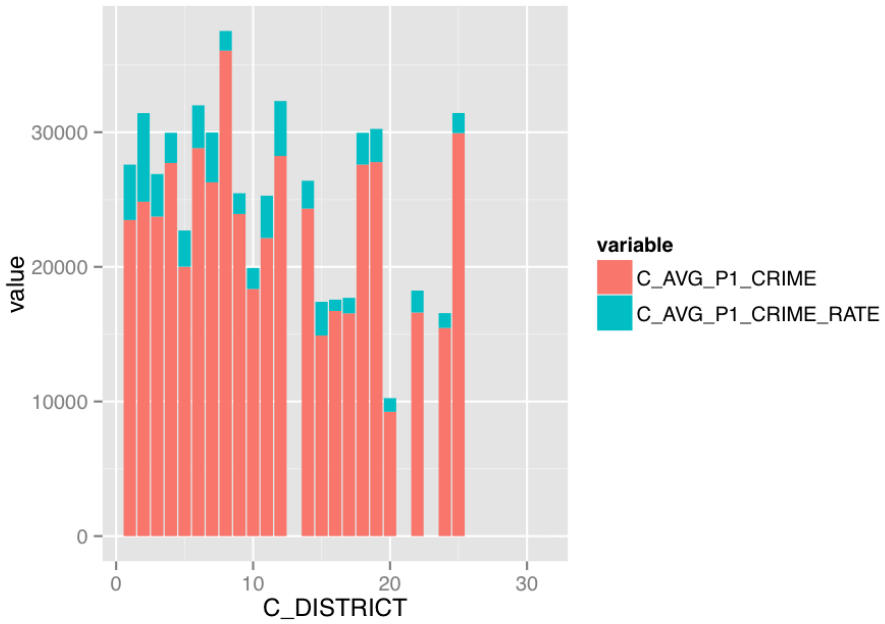


The above stacked bar graph shows the number of Part I crimes and number of 311 streetlight service requests for Chicago. You can see district 8 has the largest number of total Part I crimes and one of the largest number of 311 streetlight service requests. It's important to note this graph does not account for population within each district. Without accounting for population the total number of crimes could make a heavily populated district seem unsafe. To recitify this, I will calculate the crime rate within each district in Chicago and San Francisco using the poopulation figures of each district.

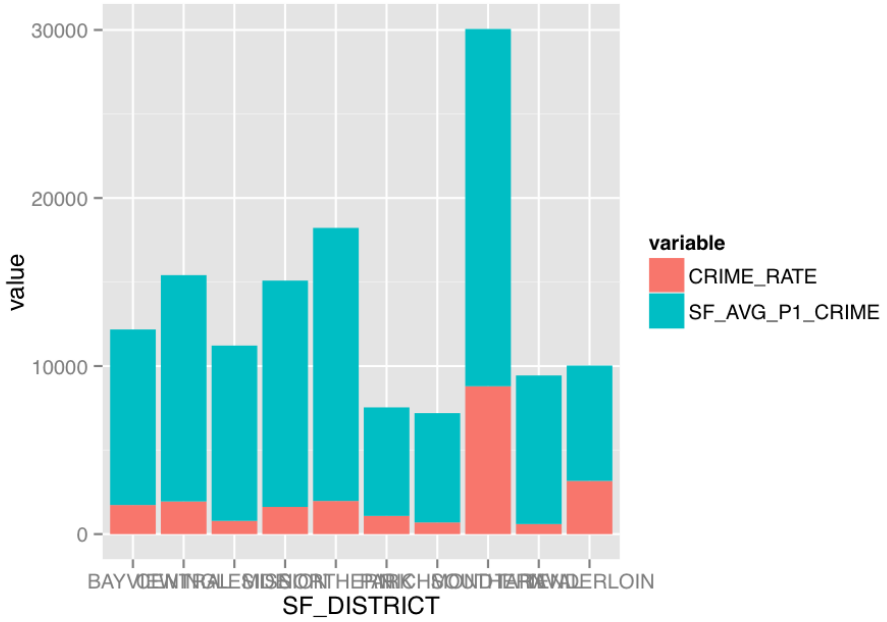
Similarly, below is the same stacked bar graph showing Part I crimes and number of 311 streetlight service requests for San Francisco. The Southern district had the greatest number of Part I crimes.



As previously mentioned, to get an accurate picture of what crime is like in these police districts population needs to be accounted for. Below is the stacked bar graph for Chicago showing the average annual crime by district for the past 5 years with the average annual crime rate per 10,000 people for the past 5 years. Unlike the total crime count, you can see districts 2 and 12 have the largest crime rates per 10,000 people in Chicago.



Below is the stacked bar graph showing the average annual crime by district for the past 7 years with the average annual crime rate per 10,000 people for the past 7 years in San Francisco. The Southern district has the largest average annual crime rate per 10,000 people and total average annual crime count.

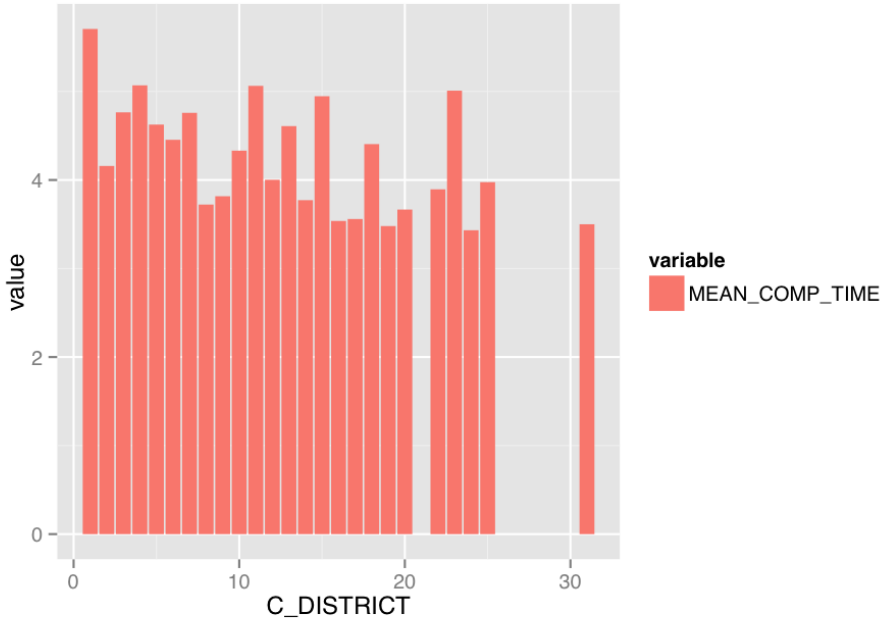


Below are the summary statistics for Chicago and San Francisco, respectively, that show the time (in days) it took to close a 311 streetlight service request.

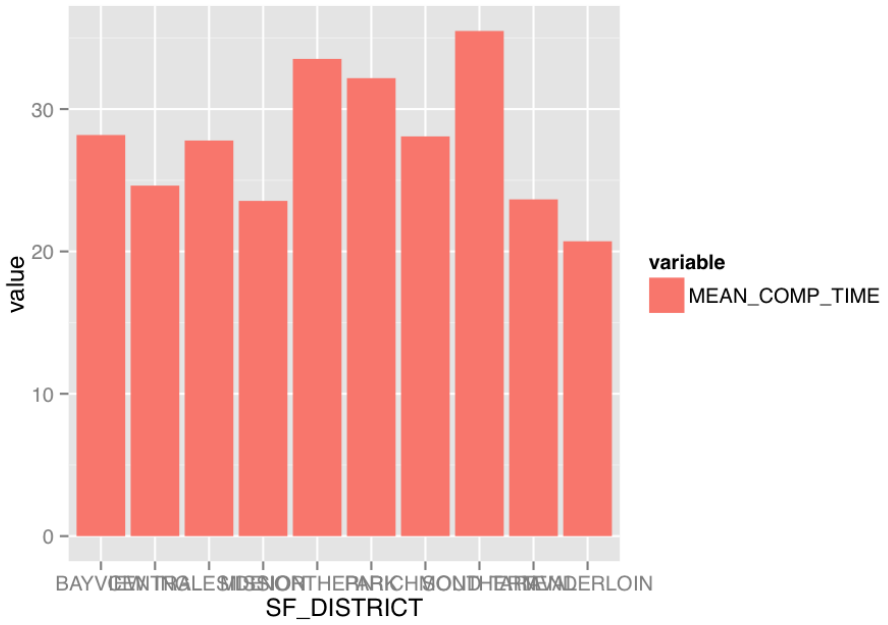
```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
##    0.000    2.000    3.000    4.213    5.000 1053.000        8
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
##   -5.00    2.00    5.00   26.66   15.00 1482.00       2
```

Now that I roughly know which districts have the highest crime rates, I will explore the streetlight outage request closure times within such districts. The below graph shows the average closure time (in days) for streetlight outages within Chicago. As a whole Chicago's average closure time is clustered around 3 to 5 days. District 1 has the longest average closure time of approximately 5 days. Districts 2 and 12 had the highest crime rates but do not have the longest or shortest closure times. At first glance, it does not seem crime rate has a big impact on streetlight outage request closure time in Chicago.



On the other hand, the Southern district within San Francisco had the largest crime rate and has the longest streetlight outage request closure time in San Francisco, with an average of approximately 37 days. It is important to note it takes San Francisco a significantly longer period of time to complete streetlight outage requests than Chicago. The Tenderloin district had the shortest average closure times, approximately 20 days.



Although I found a statistical correlation between Part I crimes and streetlight outage requests in Chicago but no correlation in San Francisco does not mean it is a spurious correlation. It should be used to gather additional data within multiple cities to conduct in depth research on the topic. I was surprised to see the Southern district at the top of the list for both crime rate and closure time. One thing is for certain, there is a big difference between Chicago's and San Francisco's streetlight average closure times. If the relationship between streetlights and crime is real then San Francisco would benefit from expediting their request closure process. Furthermore, all 311 service request data should be analyzed to ensure our cities are responding to and fixing legitimate problems within our city in a timely manner.