City Crime and Safety

Problem Statement

Introduction

My City has hired us, Team 5556 (better known as the the Dream Team), to determine their safety rating. A large set of crime data including date, description of crime, general location, and beat number (the specific zone of patrol) has been provided for analysis. The city is divided into 22 districts with a total population of 2.8 million people, with each district having comparable populations. The city has also asked us to write a two page non-technical report for the Mayor to describe our findings. We propose several targeted options and suggestions in this report as well as an overall improvement plan to the mayor to reduce crime.

Objectives

Our main objective was to develop two reliable models using the crime data set provided. We first developed a weighted model in which some crimes had greater severity and counted for more than others, and then a second unweighted model that used arrest rates to determine safety. The safety rating is the success rate of criminals getting arrested.

Assumptions

- Some crimes are more severe than others, making the effect of certain crimes greater on the overall safety rating of My City.

- The 22 districts of My City have equal populations. Chicago, a city with very similar statistics and the same number of districts, had approximately equal populations in each district as indicated by the population map in Appendix A.

- The first one or two digits of each beat number indicate which district the beat belongs to, which was determined using the Chicago Police Department standard for beat numbers.

- All of the crimes included in the provided data set are worthy of arrest.

- The two weeks of crime data provided are proportional to the crime of My City during an entire year.

- There are no extreme outliers represented in the two weeks of crime data provided. No one category of crime is drastically higher or lower than it normally would be.

- Districts 13, 21, and 23 do not exist. The provided beat numbers do not indicate any of these three districts.

- There are no additional police beats. Beats that are not listed do not exist.

- All arrests are carried out successfully by My City police.

- All crimes in my city are reported.

- The police can handle any number of arrests.

- The city can hire more personnel if the Mayor determines it is necessary.

- The data provided is 100% accurate.

Solution

Model 1: Weighted Model

In our first model, we assigned a severity rating to each crime.

Out[178]=

Crime	Severity
* Other Offense	1
Criminal Damage	3
Theft	4
* Criminal Trespass	2
Battery	8
Assault	7
Narcotics	2
* Prostitution	2
Robbery	9
Weapons Violation	2
Motor Vehicle Theft	8
Homicide	10
Interference With Public Officer	3
* Public Peace Violation	5
* Gambling	1
Deceptive Practice	3
Burglary	8
Crim Sexual Assault	9
Sex Offense	8
Offense Involving Children	8
Arson	5
* Intimidation	4
Kidnapping	9
Stalking	8
Liquor Law Violation	3
Concealed Carry License Violation	3
Other Narcotic Violation	2

estimated severity based on similar crimes

These severities are based off of guidelines created by the Georgia Board of Pardons and Paroles (see appendix). Certain severities had to be estimated by our team because predefined measurements did not exist. The severities estimated by our team were based off of crimes we viewed as similar; these crimes had official ratings.

The number of times each type of crime occured within the entire city was then calculated using a *Mathematica* program (see appendix).

Out[191]=

| Crime	Occurences
Other Offense	643
Criminal Damage	1093
Theft	2658
Criminal Trespass	310
Battery	2040
Assault	712
Narcotics	1179
Prostitution	92
Robbery	403
Weapons Violation	142
Motor Vehicle Theft	382
Homicide	20
Interference With Public Officer	65
Public Peace Violation	119
Gambling	24
Deceptive Practice	509
Burglary	566
Crim Sexual Assault	43
Sex Offense	23
Offense Involving Children	69
Arson	24
Intimidation	5
Kidnapping	11
Stalking	7
Liquor Law Violation	21
Concealed Carry License Violation	1
Other Narcotic Violation	1

The number of times each crime occured was then multiplied by the severity and divided by the population of *My City*, 2,800,000, and finally multiplied by 100,000 in order to get the crime occured per 100,000 people.

$$\frac{\text{occurences}\,(\text{severity})}{2\,800\,000}\,(100\,000) = \text{Threat per } 100\,000 \text{ people}$$

Crimes by Type per 100,000 Population With Severity
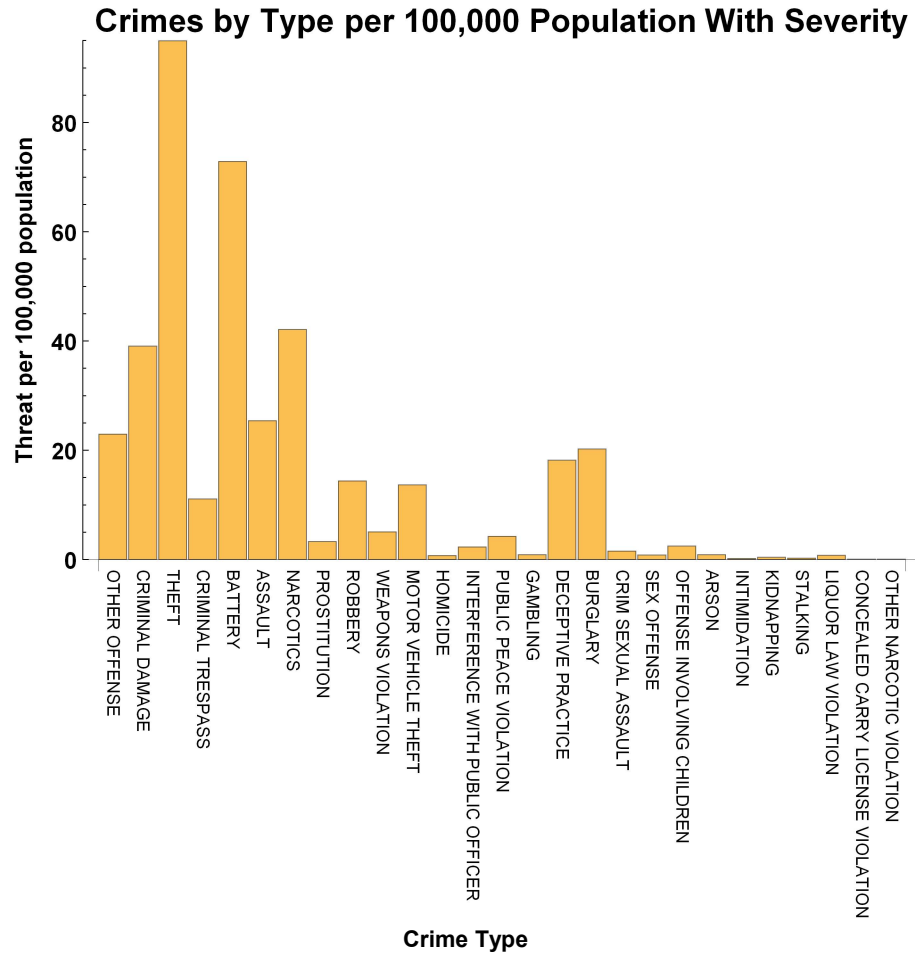
Next, the number of crimes per district was calculated. By using the Chicago Police Department method of creating beats, we used a program coded in Python to extract the number of crimes per district from the given data set. The number of successful arrests per district was also calculated. The following graph illustrates the number of crimes per district and the number of crimes that had successful arrests.

Crimes And Arrests In Each District

Out[95]=
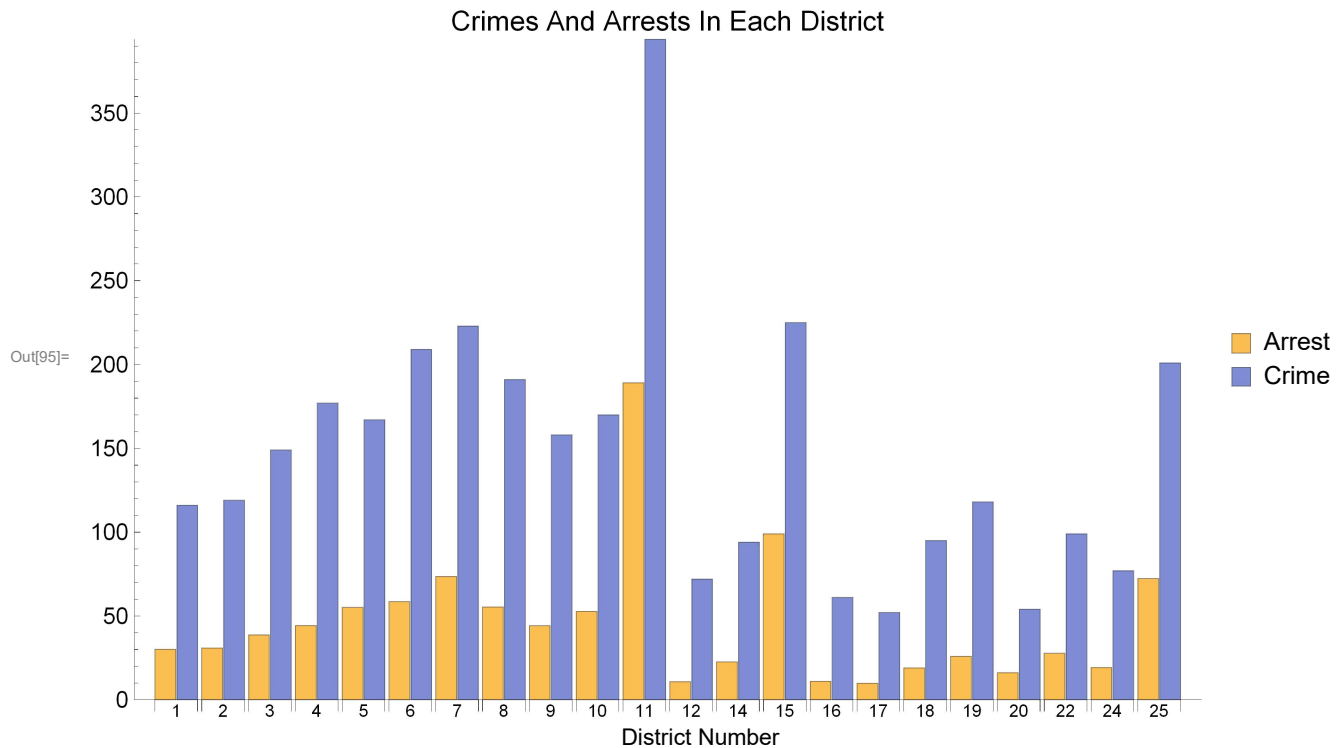
Finally, we created a safety rating for our first model. We did this by getting the severity rating for each crime, and multiplying by 0 if there was no arrest, and 1 if there was an arrest. This way, we can get the weighted value of the arrested crimes. Then, we divided by the sum of all the severity ratings for all crimes so that the success rate would be a percent out of 100.

$$\frac{\sum \text{severityRating} * \text{arrest} \{0, 1\}}{\sum \text{severity}} = \text{Success Rate}$$

Strengths of Model 1

- Accounts for certain crimes more than others
- Accounts for arrest rate
- The majority of crime severity is based on official guidelines

Weaknesses of Model 1

- Some severities had to be estimated by our team
- Crime is crime, and some people may feel severity is of no concern

Model 2: Unweighted Model

In our second model, we first found the number of crimes in each district using a Python program (see Appendix A). The number of successful arrests were once again compared to the total number of crime per district.

Crimes And Arrests In Each District
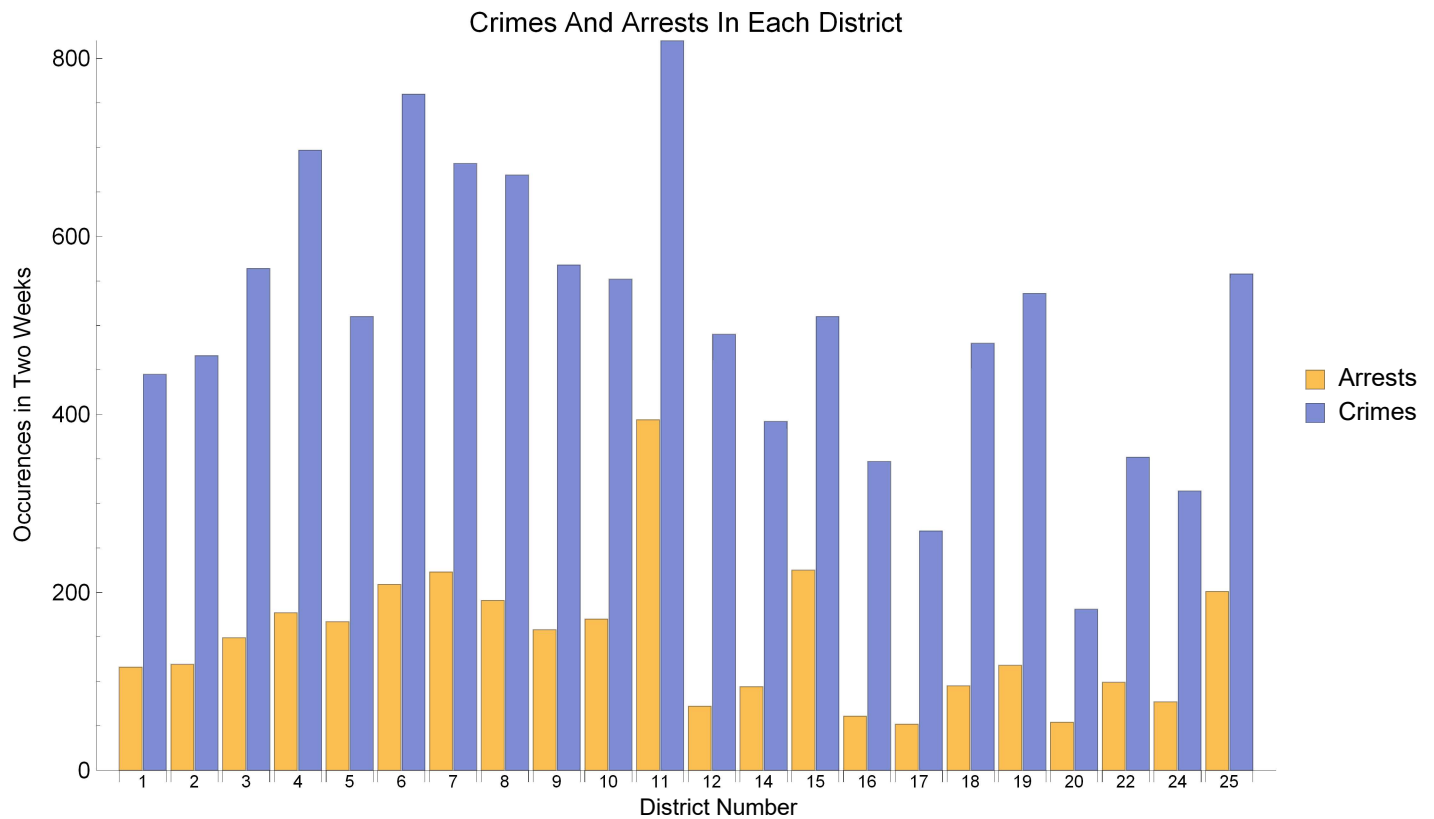
We developed a sequence of ratios between arrests and crimes for each district that accurately represents success in thwarting crime. A ratio of 1 implies infallible law enforcement whereas a ratio of 0 and other lower numbers would imply a lack of successful law enforcement.

$$totalArrests \, / \, totalCrimes = successRate$$

Strengths of Model II

- More concrete, contains no opinionated measure of severity.
- The model is very effective, but also very simple.

Weaknesses of Model II

- Less accurate because a petty crime is considered equal to a felony such as murder

Extensions

In a future model, we would like to find districts that have unusual distributions of types of crime. We could use this data to create a more targeted solution for reducing crime and improving city safeness. Also, we would like to analyze what types of crime is common on everyday of the week and find a potential pattern. We know crime overall is more apparent on Fridays; however, we would not know for sure what type of crime causes this spike. We would also like to use secondary descriptions and specific locations to create a detailed crime map.

Conclusions

In conclusion, the city as whole isn't very safe with a safety rating of 0.28 and 0.22 respectively for each of the models. Because a rating of 1 would imply perfect law enforcement and safety, these low ratings would suggest that only about a fourth of the city's crime is successfully thwarted. Thus, we can conclude that the city is not a safe place to live in general.

Report to the Mayor

November 11, 2015

To the Mayor of My City,

The safety of citizens and property is essential to the quality of living in any city. Constantly fighting crime and working towards creating a safer society should always be a priority. However, without proper data analysis, it is nearly impossible to make informed decisions on how to improve a city. In the case of My City, extensive amounts of crime data have already been recorded, and analyzing this data gave us valuable information.

Extracting data from two weeks of police records in July of 2014, we represented the important information through a chart of crimes and arrests per district, and a chart of crime types per population. We noticed a few important discrepancies between the 22 districts.

First, we noticed that theft and battery were by far the most common crimes, far outranking any other crime at 94 and 72 occurrences per 100,000 people, respectively. Many crimes such as intimidation, kidnapping and stalking were very rare, barely visible on our charts.

Secondly, we made a crime severity ranking system based on average jail sentence. Factoring the severity system into our crime per population, we were able to take into account how different crimes pose different threat levels to civilians. In addition, more severe crimes correspond to longer jail sentences, and create a greater burden on the jail and criminal justice system. Thus, the graph could show which crimes were causing the most danger to civilians and costing the most for the jail system. On this graph, battery and theft led once again at approximately 600 and 400 arbitrary danger units.

In order to combat the high levels of battery, we propose a citywide outreach program in order to educate the public on the risk of battery they are currently at. It would serve to make more people aware of their surroundings and minimize the chances of persons getting battered.

In order to combat the high levels of theft, we propose an increased presence of police officers in and around popular shopping districts and residential areas at night. This would discourage possible theft because there would be a greater chance of arrest.

Fourth, we observed district 12 has the most crime. This is mainly because district 12 had the lowest arrest rate of the city districts. Increased the presence of police officers all over district 12 would be a clear warning to potential criminals the city will not tolerate these illegal activities.

Crimes by Day of Week
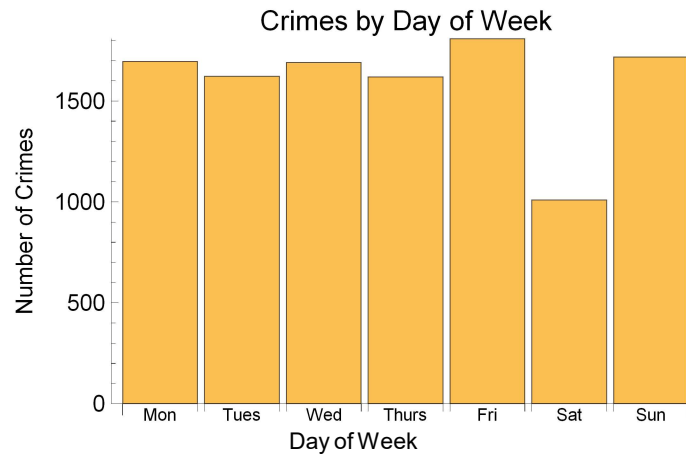
Finally, we noticed Fridays was the most popular day for crime. We propose businesses and shopping centers close earlier than than normal until this rate has gone down.

Thanks,
Team 5556

Works Cited

Know Your District. (n.d.). Retrieved November 11, 2015, from
        https://portal.chicagopolice.org/portal/page/portal/ClearPath/Get Involved

Calculating Crime (n.d.). Retrieved November 11, 2015, from
        http://oag.ca.gov/sites/all/files/agweb/pdfs/cjsc/prof10/formulas.pdf

Chicago Beats (n.d.). Retrieved November 11, 2015, from
        http://gis.chicagopolice.org/pdfs/district_beat.pdf

Crime Rankings (n.d.). Retrieved November 11, 2015, from
        https://pap.georgia.gov/sites/pap.georgia.gov/files/CSL-s_Post_1-1-2006_considerations.pdf

Appendix

A)
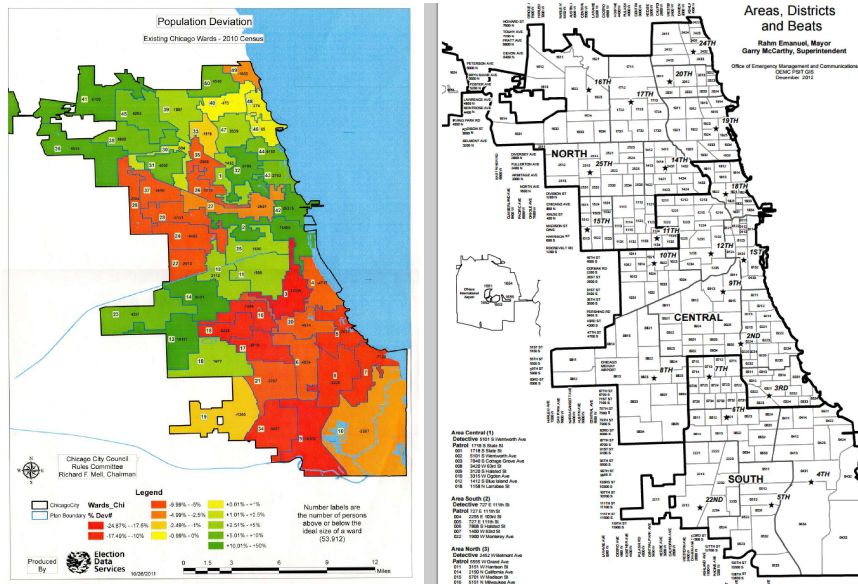
B)

```
############################################################################################
# Imports ##################################################################################
############################################################################################


import xlrd,itertools,xlsxwriter,matplotlib.pyplot as plt,math


############################################################################################
# Open Excel File ##########################################################################
############################################################################################


cases=xlrd.open_workbook('data.xlsx')
sheet=cases.sheet_by_index(0)


############################################################################################
# Severity Ranking #########################################################################
############################################################################################


severity={'KIDNAPPING':9,'CONCEALED CARRY LICENSE VIOLATION':3,'PUBLIC PEACE VIOLATION':5
```

```
################################################################################
# Old Data ######################################################################
################################################################################


dist=[math.floor(sheet.cell_value(i,7)/100.0) for i in range(1,sheet.nrows)]
vals=[sheet.cell_value(i,2) for i in range(1,sheet.nrows)]


data=zip(dist,vals)


################################################################################
# New Data ######################################################################
################################################################################


distNums=[]
avrgNums=[]


for area in set(dist):# calculates severity of crimes in particular region


points=[severity[v] for b,v in data if b==area]



summed=float(sum(points)) # find the total severity
length=float(len(points)) # find the total offenses


avrg=summed/length


distNums.append(area) # unique area code
avrgNums.append(avrg) # average severity


################################################################################
# Show Plot #####################################################################
################################################################################


plt.scatter(distNums,avrgNums)
plt.show()


################################################################################
################################################################################
################################################################################
```