Problem B: Curbing Crime Violence

Team 2796

Contents

1 Introduction

As urban develop throughout the world continues to expand, there is a natural contingency for crime. The weak socioeconomical status of parts of the city may make the regions conducive to problems with gangs and violence. For regulation of city crimes, resources are allocated to provide these deprived communities with opportunities for jobs and education. However, other cities put emphasis on enforcement of the crimes via an increased officers and investment police department resources. Regardless, crime is an issue that concerns any populated metropolitan area. In this study, we look to determine the source of crime.

2 General Assumptions and Justifications

Assumption: Similar populated sized cities have similar areas, amount of schools, and layout.
Justification: Population of a city is limited to the size of that city; a smaller sized city cannot grow and be as populated as a larger area city. The population affects the number of schools and other infrastructure; more people require more schools and a more infrastructure. Similarly populated cites, therefore, would have a similarly sized region and a similar amount of schools.

Assumption: The city can be modeled using a pure square grid system.
Justification: When observing the two cities that best matched the demographics of the given city, we noticed that these cities mostly followed a grid system. We generalized that this city can be modeled with a pure square grid, therefore, to make the problem more solvable.

Assumption: Schools are placed where there is a population need.
Justification: Schools are placed where there is need. If a region in not populated, then there would not be a need for a school there.

Assumption: People are not predisposed to commit crime.
Justification: If people inherently commit crime, then under any circumstance, they will still commit crime, making the problem impossible to solve. Assuming solvability of the issues is necessary to solve it.

Assumption: National-level crimes involving state and federal officials will not be committed.
Justification: Few crimes become national headlines and the chance that it will happen in a specific city is even smaller.

Assumption: Unemployment statistics do not include teenagers.
Justification: The U.S unemployment rate is calculated by the number of people who collect unemployment checks. Teens do not formally collect unemployment checks and are not counted in the unemployment numbers.

Assumption: Discouraged workers are not significant in this model.
Justification: Discouraged workers are sample of the population that is not getting unemployment checks. It is nearly impossible to accurately calculate the number of discouraged workers in the American economy.

Assumption: The total city population does not include the number of people in prison.
Justification: There are more people in prison then there are people in the city. The only way to justify this fact is if the prison population is not include in the city population.

Assumption: Population density decreases constantly as the distance from a center increases.
Justification: The center of the city is defined as the area with the highest concentration of commerce and people. As we move away from the city, population density tends to decrease as we can see in cities such as Phoenix, AZ and the two cities that best match the given citys profile: Springfield, MA, and Salinas, CA (Rex, 2000).

3 Objectives

4 Task 1: Creating the Model

4.1 Profiling the City

Given the population statistics, we searched for cities with similar population and located in similarly populated county. We searched within ±5% tolerance of the 2008 population of this city. Looking through the Census Bureau 2008 July estimates, we found 22 cities that had similar population sizes. After running a Q test, to determine whether the population of the 24 cities matched the given city, we determined that 2 of these cities were outliers, and threw them out. Afterwards, we profiled the amount of people in the county that each respective city was located in, and performed a Q test to find which counties were similar to the given citys county. We found that only 2 of these cities matched the profile with a ±10% tolerance: Springfield, Massachusetts and Salinas, California.

After choosing these two cities, we took the average of their area to determine the area of our city, which we calculated to be 28 square miles. Looking at the two cities, we determined that these cities can be modeled on a Cartesian coordinate plane. We calculated the average of high school between the two cities and determined there were 5 high schools.

4.2 Surveying the Problem

In order to identify the strategy for improving the city, we surveyed different aspects of the sample city distribution given. The primary method for performing these surveys was multiple regression analysis. The multiple regression analysis was performed on the parameters with regards to time. The multiple regressions can perform regression for multiple independent-variable parameters (X) and a single dependant-variable result (Y). β_n is analogous to the slope of the Pearson linear regression of the independent parameter and the dependant-variable. The value is the y-intercept analogy of the Pearson linear correlation. The $\beta_{standardized-weight}$ gives the relative influence of the parameter on the dependant variable. Once we establish which independent variables are correlated with the dependant variables, we can proceed by approaching possible solutions.

4.3 Correlated Parameters with Violence Incidents

A multiple regression analysis was run on city population (X_1), county population (X_2), the inverse of high school enrollment (X_3), the rate of student failure to graduate (X_4), the high school drop outs (X_5), and unemployed population (X_6) with the overall incidence of violence (Y). The multiple regressions were done as a function of the time (in years). All statistical tests were run using VassarStats. The regression was given in the following general form, where X_n is a parameter and Y is the incidence of violence:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + +\beta_6 X_6 + \beta_7 X_7 \qquad (1)$$

$$\begin{aligned} Y \approx &- 18433.2981 + (-0.0522)X_1 + 0.0463X_2 + 7.586 * 10^7 X_3 \\ &+ 463.1819X_4 + (-1.5634)X_5 + (-0.0293)X_6 \end{aligned} \qquad (2)$$

We observed a relatively significant correlation between all the parameters and the incidence of violence ($R^2 = 0.804$, $df = 6$). The elimination of any individual parameter $|n| - 1$ reduced the strength of the correlation where $|n|$ is the number of parameters and $p(X)$ is the p-value of statistical significance $p(X_{|n|-1}) > p(X_{|n|})$. The selected parameter provided the optimal statistical significance between the parameters and the incidence of violence. Correlation of all permutations of parameters and incidents of violence are shown in the following multiple regression correlation matrix:

	X_1	X_2	X_3	X_4	X_5	X_6	Y
X_1	1	0.369	-0.219	-0.302	-0.375	-0.119	-0.489
X_2	0.369	1	-0.969	-0.8	-0.587	0.055	-0.542
X_3	-0.219	-0.969	1	0.859	0.668	0.005	0.572
X_4	-0.302	-0.8	0.859	1	0.798	-0.054	0.744
X_5	-0.375	-0.587	0.668	0.798	1	-0.156	0.625
X_6	-0.119	-0.055	0.005	-0.054	-0.156	1	0.107
Y	-0.489	-0.542	0.572	0.744	0.625	0.107	1

Weak Positive Correlation (0<R<0.5)
Strong Positive Correaltion (0.5<R<1)
Absolute Positive Correaltion (R=1)
Weak Negative Correaltion (-0.5<R<0)
Strong Negative Correaltion (-1<R<-0..5)

Table 1

A partial negative correlation was observed between X_1 and X_2 with Y. These parameters are attributed to the characteristics of the sample population. This suggests that an increase in the population yields a lower incidence of violence. Others studies (Wirth, 1938; Sampson, 2006) have observed density-dependant urban crime, contradicting the results of this data. Due to the multiversity of the influence of population change on the function of the society, we assume that the correlation is a response to a societal response from population growth not given in the given parameters. For the purpose of this study, we will consider density (D) and incidence of violence (V) to be directly proportional ($D = kV$). The $\beta_{standardized-weight}$ of each parameter is given as the following: X_1 (-2.3327), X_2 (8.6558), X_3 (9.4401), X_4 (0.279), X_5 (-1.8489), and X_6 (-0.962). The $\beta_{standardized-weight}$ values show that among the parameters enrollment in high school was most significantly correlated.

4.4 The Effect of Education on Unemployment

Although other studies (Ashenfelter & Ham, 1979; Barry & Hannan, 1997) have pointed to the inverse correlation between education and unemployment, we decided to determine whether there is a statistically significant correlation between the inverse of high school enrollment (S_1), the rate of students that fail to graduate (S_2), the number of high school drop outs (S_3), and the population that is unemployment (U) for this particular city. All calculations and data were acquired using the same method as the previous section (Correlated of Parameters with Violence Incidents). A significant correlation was not observed between any educational parameters and unemployment ($R^2 = 0.0457$, $df = 3$, $p = 0.0319$). The following is the correlation matrix of all permutations of this multiple regression:

	S_1	S_2	S_3	U
S_1	1	0.859	0.668	0.005
S_2	0.859	1	0.798	-0.054
S_3	0.668	0.798	1	-0.156
U	0.005	-0.054	-0.156	1

Table 2

4.5 The Effect of Education on the Juvenile Jail Population

Since we observed that the educational aspects of the community did to lead to unemployment, we decided to identify any effect of education on the number of juvenile inmates. A multiple regression analysis was done as shown previously testing the influence of the inverse of high school enrollment (H_1), the rate of student failure to graduate (H_2), and the number of high school drop outs (H_3) on juvenile inmates (J). A highly significant correlation was observed between the educational parameters and the juvenile jail population ($R^2 = 0.9804$, $df = 3$, $p < 0.01$). This yielded the following functionalized regression formula:

$$J = \alpha + \beta_1 H_1 + \beta_2 H_2 + \beta_3 H_3 \tag{3}$$

$$J \approx -54362.8309 + (5.54)10^8 H_1 + (-4920.2935)H_2 + (-8.7266)H_3 \tag{4}$$

The $\beta_{standardized-weight}$ of each parameter is given as the following: H_1 (0.7617), H_2 (-0.0617), and H_3 (-0.3384). The most attributable parameter is high school enrollment, followed by the failure to graduate and the number of high school drop outs. The following is the correlation matrix of all permutations of this multiple regression:

	H_1	H_2	H_3	J
H_1	1	0.025	-0.516	0.935
H_2	0.025	1	0.438	-0.191
H_3	-0.516	0.438	1	-0.758
J	0.935	-0.191	-0.758	1

Table 3

The Effect of Education on the Jail Population The significant correlation between juvenile jail population and education parameters suggested that the there may be a correlation between the effect of education on the general jail population. The factors that were used in this multiple regression analysis were the influence of the inverse of high school enrollment (E_1), the rate of student failure to graduate (E_2), and the number of high school drop outs (E_3). These were used to determine any potential influence on the total prison population (P). A moderately significant correlation was established between the parameters and the total prison population (R^2 = 0.618, df = 3). This resulted in the following multiple regression analysis function:

$$P = \alpha + \beta_1 E_1 + \beta_2 E_2 + \beta_3 E_3 \tag{5}$$

$$P \approx 258481.4447 + (-9.41 * 10^8)E_1 + (-46732.4669)E_2 + 58.1383E_3 \tag{6}$$

The reduced significant of correlation compared to the highly significant correlation established between the education parameters and the juvenile jail population. This may be attributed to dangerous behavior committed following the completion of high school. The standardized-weight of each parameter is given as the following: E_1 (-0.876), E_2 (-0.2106), and E_3 (0.5143). The most attributable parameter is the number of high school drop outs, followed by the failure to graduate and high school enrollment. The following is the correlation matrix of all permutations of this multiple regression:

	E_1	E_2	E_3	P
E_1	1	0.859	0.668	-0.713
E_2	0.859	1	0.798	-0.553
E_3	0.668	0.798	1	-0.238
P	-0.713	-0.553	-0.238	1

Table 4

4.6 The Effect of Unemployment on the Jail Population

In addition to education, we looked at unemployment as a source for the high jail population. A Students t-test was performed on the population unemployed and the prison population. The F-test revealed a significant difference in variance between the two samples (F = 16.57, df = 8, p = 0.000319). A statistically significant difference was observed between the two samples (t = -67.64, df = 16, $p < 0.0001$). This statistical rejects the idea that unemployment and prison population are related for the crime present in this given city.

4.7 The Effect of Behavior on Jail

We define behavior through the inverse of high school enrollment (B_1), the rate of student failure to graduate (B_2), the population of high school graduates

(B_3), and the population unemployed (B_4). These behavioral factors were used observe any possible correlations to the prison population (R). We observed a relatively significant correlation ($R^2 = 0.7893$, $df = 4$). This correlation yielded the following equation of multiple regressions:

$$R = \alpha + \beta_1 B_1 + \beta_2 B_2 + \beta_3 B_3 + \beta_4 B_4 \tag{7}$$

$$\begin{aligned} R \approx & 280436.9018 + (-8.62 * 10^8) B_1 + (-43210.6989) B_2 \\ & + 43.7365 B_3 + (-1.7211) B_4 \end{aligned} \tag{8}$$

The $\beta_{standardized-weight}$ of each parameter is given as the following: B_1 (-0.8026), B_2 (-0.1947), B_3 (0.3869), and B_3 (-0.4228). The most attributable parameter is high school enrollment, followed by the failure to graduate and the number of high school drop outs. The following is the correlation matrix of all permutations of this multiple regression:

	B1	B2	B3	B4	R
B1	1	0.859	0.668	0.005	-0.713
B2	0.859	1	0.798	-0.054	-0.553
B3	0.668	0.798	1	-0.156	-0.238
B4	-0.005	-0.054	-0.156	1	-0.476
R	-0.713	-0.553	-0.238	-0.476	1

Table 5

4.8 Effect of Education on Parole Violations

To observe the potential relationship between education and the effect on post-high school behavior, we studied the correlation between three educational parameters [the inverse of high school enrollment (F_1), the number of high school drop outs (F_2), and the rate of failure to graduate (F_3)] and unemployment (F_4) with the number of parole violations per year (L). The relationship, although moderately significant ($R^2 = 0.6865$, $df = 4$), incorporated multiple variables, allowing for a reasonable extend for variation of the coefficient of determination (R^2). This provided the following equation and correlation matrix:

$$L = \alpha + \beta_1 F_1 + \beta_2 F_2 + \beta_3 F_3 + \beta_4 F_4 \tag{9}$$

$$L \approx 1.49 * 10^5 + -4.9 * 10^8 Z_1 + 109.25 Z_2 + -356.065 Z_3 + -1.498 Z_4 \tag{10}$$

	F_1	F_2	F_3	F_4	L
F_1	1	0.668	0.37	0.005	0.108
F_2	0.668	1	0.51	-0.156	0.613
F_3	0.37	0.51	1	-0.17	0.179
F_4	0.005	-0.156	-0.17	1	-0.486
L	0.108	0.613	0.179	-0.486	1

Table 6

4.9 Results of the Population Characteristics Survey

The following flow chart gives an overview of the survey results:

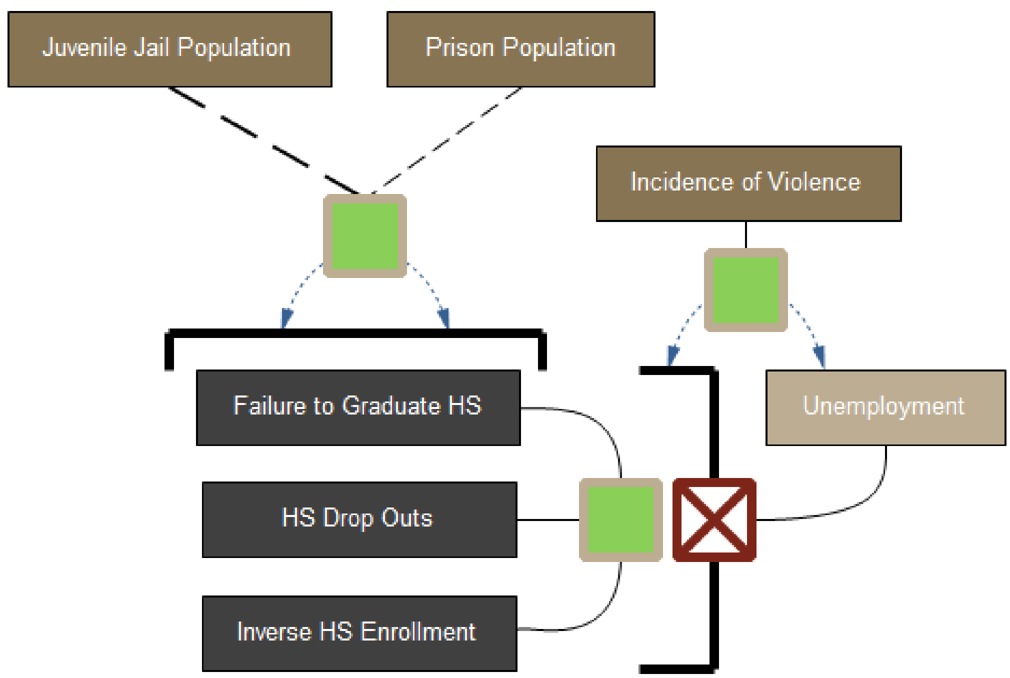

Figure 1

All correlations were done as sets, so brackets were used to show the sets that were used for correlation. For example, the education-related parameter set was used in multiple tests for comparison to different dependant-crime associated factors. Unemployment had no correlation to the education parameter set. However, a correlation was established within the high school parameter set. Incidence of violence showed a correlation with the education parameter set and unemployment. Additionally, juvenile jail population and prison population both showed a correlation with the education parameters, although the juvenile jail population had a stronger correlation.

For these statistical tests, we can determine significant parameters. Through these tests, we can see that the rate of student failure to graduate, the high school drop outs, high school enrollment, and unemployed population are significant factors. Furthermore, we looked at the interconnection between multiple parameters and observed the complexity interaction of between parameters and their effects. The goals of this solution model will be to address these parameters.

5 Task 2: Mapping the Problem

5.1 Objectives

Our question was to model the incidence of violence and to give a solution to reduce violence in the city. Our first objective in finding a solution to reduce violence is to identify various changes that can be made in the city in order to improve factors such as high school, enrollment, high school graduation, unemployment, etc. Our next objective is to find how these changes will specifically (percentage wise) affect the before-mentioned factors. By finding these values we will be able to find a direct effect of the change not only on immediate factors but also the effect on the overall change in the prevalence of incidence of crime and in turn number of incarcerated persons within the city. The final objective is to find the overall improvement in the crime rate of the city through all of the changes we made to see the effectiveness of our proposed improvements. Another side objective in this section is to model the city graphically in order to obtain a visual representation of our changes.

5.2 Mapping Overview

A computer-based program that can model a simple city such as the one in question is determined to be extremely important to facilitate both quantitative and qualitative simulation of our solutions. Using MATLAB, a program was written to model population distribution using random simulation, estimate percentages of populations going to each school, and determine the effect of a new school in conditions similar to that described in the prompt. After running the simulation several times, it was determined that introducing a school towards the center would result in 10

To model population density of the city, it was first assumed that the city was in the confines of a square. It was additionally assumed that the population was most heavily concentrated in some region in the center of the city, and that the distance from this central location was inversely proportional with population density. The program, however, did simulate some variance in where the center of the city really is, as is common in most mid-size and large cities. Additionally, the population density was not a perfect Gaussian distribution, but rather closely approximated it. Thus, applications of the model could expand to most cities, but not just the perfect city.

To model high school locations and the volume of students, it was assumed that a population in a certain region was proportional with regular participants of school, and thus density determined on a map generated of a city could be representative on a grayscale image, with white demarcating the highest populated regions and black the lower populated region. The four high schools were also assumed to enroll equally-sized populations of students, though not necessarily exactly equally-populated regions. These existing high schools would then be present in the center of one of the four equally-sized quadrants that could be made from the city. The new high school, on the other hand, would be built near or around the center of the city, and though enrolling boundaries were confined to a smaller region than the other schools, the population of students would remain more or less the same as in the other four schools. The new high school would enroll from a diamond-shaped portion of the city in order to optimize travel distance for all four schools.

5.3 Gaussian Distribution

To complete the model of the city we needed to model how the population density varied over the span of the city. To do this we used a Gaussian distribution to show the population density as a function of the distance from the mean density. Gaussian (normal) distribution is utilized in statistics as a form of probability distribution. It is generally used to represent the spread of reality based random variables that cluster about a mean value. The function to represent a normal distribution in two dimensions can be modeled by:

$$f(x) = \frac{1}{\sqrt{(2\pi\sigma^2)}} e^{\frac{-(x-\mu)^2}{(2(\sigma)^2)}}$$

The function above is defined by two constants (μ and σ^2). These constants represent certain statistical values:

$$\mu \in \mathbb{R}$$

is the mean value and $\sigma^2 > 0$ is the variance. For the purposes of creating a model that would satisfy the parameters that we set up (a mean value μ=7,8,9,or10 and a maximum of 32) we had to manipulate the function into the one below:

$$f(x) = 32e^{\frac{\pi(x-\mu)^2}{128}}$$

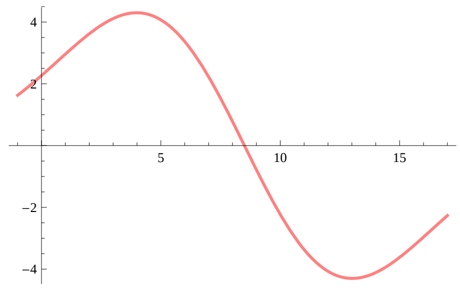

Figure 2

The graph above was calculated with a =8.5 for the purpose of showing a general solution in the middle of our data. The graph below shows the equation graphed with the =7,8,9,and 10 overlaying each other. The reason for choosing those four mean values instead of just the average is to make it more realistic. In real life the center of a city is not located in the geographical center, but is in the general vicinity. Since we divided our city into a 16x16 matrix, we chose the population center to be randomly picked within the center 4x4 matrix.

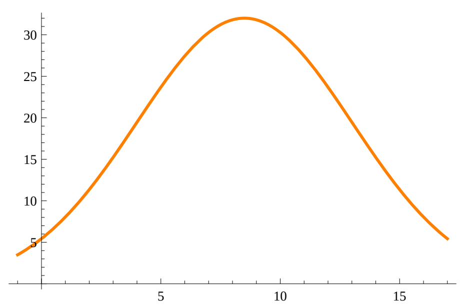

Figure 3

Another measure that was taken to make the map of the city more realistic was to not model the density strictly off of a normal distribution, but to let it vary a little around the normal distribution. To put this into our particular model, we use the derivative of the normal distribution function above in order to find the rate at which the density changes in order to calculate how much the grayscale value of each concentric perimeter from the randomized center would decrease from the previous perimeter. The equation modeling that is seen below:

$$f'(x) = -\frac{\pi}{2}e^{\frac{\pi(x-\mu)^2}{128}}(x-\mu$$

)

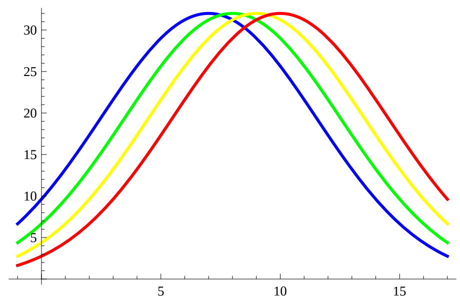

Figure 4

Then to add the realistic variance we allowed the program to shade the density of each square by following method: Take the rounded value of the normal distribution function for that square Take the rounded value of the derivative of the distribution function for that square Create a range defined by: $f(x) - 1/2f'(x)$, $f(x) + 1/2f'(x)$ Pick a random number within that range as the grayscale value

By following this method we were able to create a population density model that follows a normal distribution, but also allows for variance that would normally be seen in the real world, such as the center and rate of change.

5.4 Results

A series of simulations revealed that the four existing schools would, on average, have their student bodies cut by 9.21%. Correlation

6 Task 3: Finding a Solution

6.1 Overview

From the survey, we recognize that 3 of the 4 main factors to address are school related: high school graduation, high school dropout, and high school enrollment. Any solution to the crime rates must address the school system. We assume the following definition of each of these parameters: high school graduation is the number of people who graduated from high school, high school dropout is the number of people who dropped out of high school since the beginning of the school year, and high school enrollment is the number of people who are still registered as students in the school.

6.2 Proposal 1: Creation of a City Wide After-School Program for Failing Students

Justification: Statistically, the most significant issue is high school graduation, which is purely influenced by the academic performance of the students. In order to address this, after-school programs will be created in all schools for the purpose of educated students. Each program will be headed by at least one or two qualified teachers from the school. The rest of the teachers would be a mix of part and part time facility. Students would receive individual attention in the subject areas that needed to be taught the most. According to a study on California after school programs, the inclusion of after-school programs in themselves reduce high school dropout rate by 27% and an increase in graduation rate by 21% (Brown et al, 2002). Taking a fixed proportion of this to fit to our data, we can model effectively change in dropout rate to be 13.5% and the change in graduation rate to be 10.5%. This is because of the increased amount of people served, and therefore, the reduced effect in treating each individual student. The Department of Justice (DOJ) through the Office of Justice Programs (OJP) sponsors 4.2 billion dollars worth of after-school programs meant to keep students out of trouble (Maurer, 2010). Tapping into several of these funds we could fund this after school help sessions. According to The Cost and Benefits of After School Program, after school programs cost 10,038 dollars to implement, but reduces other related cost in the future, on the low end, about 89,500 dollars for a net gain of 79,500 dollars (Brown et al, 2002). The program would have to be externally funded for about 5 or 6 years before the benefits of the program start to outweigh the external cost.

6.3 Proposal 2: Create a New School

Justification: Various studies have shown that the school size is correlated with the number of crime incidences (Wilson et al, 2001). This is primarily because smaller schools have lower student to teacher ratio, and can give each student more individualized focus. Smaller schools have an 8% increase in enrollment, Studies have shown that between a school of 300 and 1000 students, the average crime rate goes up by 56%. Newer schools tend to have a lower dropout than the dropout rate around the school (4.6% vs 12.9%) . The creation of a new school, therefore will directly reduce the dropout rate (7%), and increase the enrollment (8%) (Chicago 2010). In addition, in the creation of the new school, we will generate a number of permanent teaching jobs and temporary construction jobs, reducing the unemployment rate. Springfield Public High School has about 148 teachers in its similarly sized high school. We will instead increase the number of teachers in our new high school to 160 to reduce the student to teacher ratio. With the addition on non teaching jobs, we can expect the school to employ about 200 people. With construction jobs, assuming the school takes a year to build, and the cost of the contractor fees to the US average (Reed Construction), and the average wage of the worker is 842 dollars per week (BLS) we can expect to create 60 temporary construction jobs (Reed 2008) (BLS, 2010). According to Reed Construction, the average school cost 17 million to create. Funding can be secured by borrowing bonds against the saving that we plan to make in the next two proposals.

7 Task 4: Model Regression for Predicted Solution Implementation

A multiple regression was run to develop an equation using the following parameters: the inverse of high school enrollment (Z_1), high school drop outs (Z_2), the rate of student failure to graduate (Z_3), and unemployment (Z_4). The dependant variable (V) used for modeling the relationship between these variables and crime was incidence of crime. Although a regression was established between all parameters and the dependant variable, a multiple regression analysis was run on this particular subset of parameters due to the influence of the proposed actions on these specific parameters. A moderate significant was observed for the multiple regressions of these factors ($R^2 = 0.7154$, $df = 4$). The primary implication of using correlation as source to characterize this model is the ambiguity in whether the correlation is due to causation. For the purpose of simple implementation of this model and the diversity of sources suggesting causation, all correlations established by this regression will be acknowledged as the result of causation. The regression provided the specific equation and correlation matrix as follows:

$$V = \alpha + \beta_1 Z_1 + \beta_2 Z_2 + \beta_3 Z_3 + \beta_4 Z_4 \tag{11}$$

$$V = 385.1969 + 1618665.7845 Z_1 + 0.1943 Z_2 + 7.4918 Z_3 + 0.0073 Z_4 \tag{12}$$

	Z_1	Z_2	Z_3	Z_4	V
Z_1	1	0.668	0.37	0.005	0.572
Z_2	0.668	1	0.51	-0.156	0.625
Z_3	0.37	0.51	1	-0.17	0.736
Z_4	0.005	-0.156	-0.17	1	0.107
V	0.572	0.625	0.736	0.107	1

Figure 5

Using the given multiple regression correlation, we can substitute predicted values for the inverse of school enrollment, the number of high school drop outs, the rate of student failure to graduate, and the population unemployed in order determine the effects of the previously suggested (Task 3) solutions to the crime problem. This equation will provide a prediction of the number of incidents that will occur in the city after the completion of this crime-curving reform. The changes in the parameter values for the equation will be based on known variation attributed to the parameter due to the known effect of the proposed solution shown in existing societal studies. These changes will all be based relative to the 2008 city education parameters and unemployment data provided. For parameters influenced my multiple of the proposed solutions, we will take the influence (

$$V = 385.1969 + 1618665.7845(1.046 * 10^{-4} * 0.92)) + 0.1943(147 * 0.715) + 7.4918(0.11 * 0.895) + 0.0073(17805 \tag{13}$$

$$V = 667.41 \tag{14}$$

As shown through this equation, the V value was around 69 incidents less than the cities 2008 value (736). This is an approximately 10% reduction in the incidence of violence. By performing the proposed goals to improve the citys education and employment characteristics, we will be capable of making this improvement.

8 Discussion

8.1 Pros/Cons of the Model

The pros to this model is that it address the long term goals of improved high school education with different levels of approach, initially looking at the vocational, after-school education, and then an increased number of schools. Our model served to help deal with the problems of unemployment and education. However, we neglect to account for the needs of enforcement of the community. In addition, we neglect to consider the factors of real estate on the development of a new school. One positive for this model is that we account for density in our MATLAB model of the optimal school placement. However, the number of schools and the density characteristics that we obtained were based on only two cities of similar country and city population. If more cities were used the accuracy of the model would have been better. The ability of this model to be visualized is another positive attribute, particularly when conveying the model to the public. Furthermore, the analysis of the problem through multiple forms of multiple regression analyses and Students t-tests provided support for the models and establishes the legitimacy of our model. However, additional tests could have been used to give a deeper insight into the condition of the city.

8.2 Topics for Future Study

Future studies should look into incorporating all parameters into the model. For example, we could model the distribution of employment throughout the city and observed any method for improving the unemployment problem. Another characteristic to observe are the distributions of law enforcement by looking at the ability of the police to maintain the stability of regions of concentrated crime, primarily in densely population regions. Education parameters such as the teacher to student ratio and quality of the school program should be areas for future considerations. It would be particularly significant to see the distribution of the parameters with regards to its location in the city. Further statistical tests must to be performed to establish the best form of correlation. One possible way to approach these statistical tests is the parameterization and testing of generated decision trees using data mining through the WEKA 3 data mining software. Other factors and larger sample sizes are important factors to study.

9 Letter to the Mayor

My fellow residents of Honeycomb, right now we live in a time of great strife. The city roads are poured with the screams and cries of the innocent being taken advantage of. We have conducted various statistical analysis and models and we have determined the 3 most important factors that determine crime rates, high school dropout rates, high school enrollment rates, and high school graduation rate. What this tells us is that committing crime is not controlled by a single point in time, it tells us that long term factors, many coming from adolescence age, drives the formation of the criminal mind. While this may not deter crime right away, it tells us that crime can be tackled early on, and people can be moved away from committing crime. We propose three programs to address these factors and in the long term reduce crime rates. First, create a system of afterschool programs, both educational and recreational. Our research has shown that keeping kids in school and keeping them involved after school, is the best way to deter crime. Secondly, we will create a system of vocational learning within our schools. Some of our students are more likely to be interested in being electricians, plumbers, and carpenters, jobs that are quickly fleeing our cities. Instead of being made to go through the core curriculum, a separate system will be made to teach our youth the skills they need to stay off the streets and become productive members of society. Lastly, and most ambitiously, I propose creating a new high school in the center of the city. This school will lower the student to teacher ratio of all four schools, giving the youth more attention they need, especially the ones at risk for crimes. New schools are also less likely to have violent reputations and established culture, allowing for us and the students to create a safe, working environment for these schools. According to the model, we would reduce the incidence of violent crimes by 70 crimes.

10 References

Alsapugh, John W. (2003). The Relationship between School, Size, Student Teacher Ratio and School Efficiency. Education 114(4), 593-601. Retrieved from

Amos, Jason (2008). Dropouts, Diplomas, and Dollars: U.S High Schools and the Nations Economy. Washington D.C: Alliance for Excellent Education. Retrieved November 14, 2010 from http://www.all4ed.org/publication_material/reports/Econ2008

Ashenfelter, Orley, Ham, John (1979). Education, Unemployment, and Earnings. Journal of Political Economy 87(5), 99-116. Retrieved from http://www.jstor.org/stable/1829910

Barry, F., Hanna, A (1997). Education, Deprivation, Hysteresis, Unemployment. Retrived from College Dublin, Department of Political Economy Web site. http://logec.repec.org/scripts/paperstat.pf?h=RePEc:fth:dublec:97/24

Brown, William O., Frates, Steven, B., Rudge, Ian, S., Tradewell,Richard L (2002). The Cost and Benefits of After School Programs. Retrieved from Claremont-Mckenna College, The Rose Institute Web site: www.claremontmckenna.edu/rose/publications/pdf/after_school.pdf

Bureau of Labor Statistics. (2010). Construction. Retrieved November 15, 2010 from http://www.bls.gov/oco/cg/cgs003.htm#nature

Chicago Public Schools (2010). Small Schools: Research Retrieved November 15, 2010 from http://smallschools.cps.k12.il.us/research.html

Hughes, Katherine L., Bailey Thomas, Mechur, Melinda (2001). School-to-Work: Making a Difference in Education. Retrieved from Columbia University, Institute on Education and the Economy Web site: http://www.tc.columbia.edu/iee/stwrelease.htm

Hyperphysics (2010). Gaussian Distribution Function. Retrieved November 15, 2010 fromhttp://hyperphysics.phy-astr.gsu.edu/hbase/math/gaufcn.html

Raphael, Steven, Stoll, Michael (2004). The Effect of Prison release on Regional Crime Rates. Papers on Urban Affairs 204-255. doi.10.1353/urb.2004.0011

Reed Construction Data (2008). Construction Cost Estimate: High School Construction Costs. Retrieved November 15, 2010 from http://www.reedconstructiondata.com/rsmeans/models/high-school/

Sampson, Robert. J. (1983), Structural Density and Criminal Victimization. Criminology, 21(2), 276293. doi:10.1111/j.1745-9125.1983.tb00262.x

South Dakota State University (2010). Higher Earnings for College Graduates. Retrieved November 15, 2010 from http://www.sdstate.edu/news/college-earnings-commentator.cfm

Stern, David, Dayton, Charles, Paik, Il-Woo, Weisberg, Alan (1989). Benefits and Cost of Dropout Prevention in a High School Program Combining Academic and Vocational Education: Third-Year Results from Replications of the California Academies. Educational Evaluations and Policy Analysis 11(4), 405-416. Retrieved from http://www.jstor.org/stable/1163992

Swanson, Gordon I (1978). Vocational Education: Fact and Fantasy. The Phi Delta Kapan 60(2), 87-90. Retrieved from http://www.jstor.org/stable/20299230

Tung, Rosann, Ouimette Monique, Feldman, Jay (2004). How Are Boston Pilot School Students Faring? Boston: Center for Collaborative Education. Retrieved November 15, 2010 from http://www.ccebos.org/pilots.faring.2004.pdf

Tyler, John H, Lofstrom, Magnus (2009). Finishing High School: Alternative Pathways and Dropout Recovery. The Future of the Children 19(1), 77-103. doi:10.1353/foc.0.0019

U.S Bureau of the Census (2008). Annual Estimates of Resident Population Counties: April 1, 2008 to July 1,2008. Retrieved November 14, 2010 from http://www.census.gov/popest/counties/CO-EST2008-01.html

U.S Bureau of the Census (2008). Population Estimates: Incorporated Places and Minor Civil Division. Retrieved November 14, 2010 from http://www.census.gov/popest/cities/SUB-EST2009-4.html

U.S Department of Education (2010). Implications and Conclusion. Retrieved November 15, 2010 from http://www2.ed.gov/pubs/VoEd/Chapter3/Part8.html

United States Government Accountability Office. (2010). Oversight of DOJ Funds for Recreational Activities (GAO publication No.10-661R). Washington D.C: Retrieved from www.gao.gov/new.items/d10661r.pdf

Wilson, David B, Gottfredson Denise, C., Najak, Stacy, S. (2001). School-Based Prevention of Problem Behavior: A Meta-Analysis. Journal of Quantitative Criminology 17(3), 247-272.

Writh, Louis (1938). Urbanism as a Way of Life. American Journal of Sociology 44(1), 1-24. Retrieved from http://www.jstor.org/stable/2768119

11 Appendix

LATEX Code