HIGH SCHOOL MATHEMATICAL CONTEST IN MODELING

OUTSTANDING PAPER

HiMCM

The contest offers students the opportunity to compete in a team setting using applied mathematics in the solving of real-world problems.

November 2015

Additional support provided by the National Council of Teachers of Mathematics (NCTM),
the Mathematical Association of America (MAA),
and the Institute for Operations Research and Management Sciences (INFORMS).

Editor's Comments

This is our eighteenth HiMCM special issue. Since space does not permit printing all eight Outstanding papers, this special section includes one of the eight, minus the appendices. We emphasize that the selection of this paper does not imply that it is superior to the other outstanding papers. We also wish to emphasize that the papers were not written with publication in mind. Given the 36 hours that teams have to work on the problems and prepare their papers, it is remarkable how much they accomplished and how well written many of the papers are. Visit www.mathmodels.org to view all the outstanding papers.

HiMCM Director's Article

Dr. William P. Fox
Department of Defense Analysis
Naval Postgraduate School
Monterey, CA 93943
wpfox@nps.edu

Dr. Kathleen Snook
Mathematics Education Consultant
Consortium for Mathematics and Its Applications, COMAP
Bedford, MA 01730
kgsnook@gmail.com

After eighteen years running the HiMCM, I am handing over the contest directorship to Dr. Kathleen Snook. I thank COMAP for their continued organizational support of this important contest, and I especially thank all the advisors and students that have made running this contest so fun!

As we complete the eighteenth year of the High School Mathematical Contest in Modeling (HiMCM), we celebrate that this contest is, and continues to be, a fantastic endeavor for students, advisors, schools, and judges. The mathematical modeling ability of participating students and their faculty advisors is very evident in the problem solutions and professional submissions we receive. In efforts to have their teams improve and compete at a higher level, some schools have asked us to visit and discuss the modeling process. We are happy and excited to do so, and we hope this trend continues.

The teams truly accomplished the vision of our founders by providing *unique* and *creative mathematical* solutions to complex open-ended real-world problems. This year students chose from two problems, both representing "real-world issues." (See the Judges' Commentary that follows for the problem statements.)

The HiMCM is still moving ahead, growing with a "positive first derivative!" We hope this growth continues as more high schools engage their students in mathematical modeling and participate in the contest. Figure 1 shows a

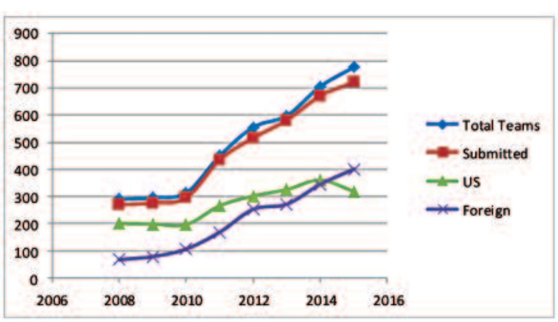

Figure 1: *Number of HiMCM teams vs. contest year from 2008-2015*

plot of the growth over time. The trend continues to reflect an increasing engagement of high school students in mathematical modeling.

The 2015 contest had 776 registered teams resulting in 719 total submissions (92.07%), a registration increase of about 10% over last year. In total, 2784 students competed in the contest. This represents an increase of 3.5%. Teams from the United States represented 24 states and numbered 318. This number was down a considerable 12% from last year. We had 401 foreign teams, representing a 16.23% growth. China represented about 87% of the foreign entries.

Of the 2784 student participants this year, 1030 or about 36.99% were female, 1702 were male, and 52 did not specify gender. Since the start of HiMCM, 8,261 (36.5%) of 12,606 total participants have been female. We hope that all competing students will continue to pursue further STEM education.

Summary of Facts from the 18th Annual Contest

- A wide range of schools/teams competed including teams from the United States, Turkey, United Arab Emirates, Philippines, Singapore, Hong Kong SAR, and China.
- The 776 registered teams from U.S. and International institutions represent over a 10 % increase in participation.
- There were 2784 student participants, 1754 (63%) male and 1030 (37%) female.
- Schools from twenty-four states in the U.S. participated.

Commendation
We congratulate all students and advisors for their varied and creative mathematical efforts. We continue to encourage all teams to submit their solutions to experience the learning impact and satisfaction of fully completing this challenging contest. Of the 719 submissions, 273 completed Problem A and 446 Problem B. The 36 continuous hours to work on the problem provided for high quality papers. We commend all teams for the overall quality of their work. Judges frequently comment about the amazing amount of mathematical thought and supporting material students are able to put together and present in their paper within the 36 hours.

Teams again and again prove to the judges that they have "fun" with their chosen problem, demonstrating research initiative and creativity in their solutions. This year's effort was again a success!

Judging
We ran three Regional Judging sites in December 2015 that

Judging Results

Problem	Outstanding	%	National Finalist	%	Finalist	%	Meritorious	%	Honorable Mention	%	Successful Participant	%	Total
A	4	1%	7	3%	22	8%	71	26%	96	35%	73	27%	**273**
B	4	1%	7	2%	30	7%	117	26%	146	32%	142	32%	**446**
Total	8	1%	14	2%	52	7%	188	26%	242	34%	215	30%	**719**

Figure 2: *Eighteenth Annual HiMCM Statistics*

also included remote judges. The regional sites were:

Naval Postgraduate School in Monterey, CA

Francis Marion University in Florence, SC

Carroll College in Helena, MN.

Each site judged papers for problems A and B. The papers judged at each regional site may or may not have been from their respective region. All judging is blind with respect to any identifying information about the participants or their schools. Judges ranked papers as Finalist, Meritorious, Honorable Mention, and Successful Participant. All finalist papers from the regional sites were sent to the National Judging in Seattle, Washington. This year, ten judges from both academia (high school and college) and industry participated in the National Judging. At this level, judges first select the best of the Finalist papers as National Finalists, and then, from this set, choose the "*best of the best*" as Outstanding papers. This year there were fifty-two papers in the final discussions. Eight papers earned the Outstanding award and fourteen were deemed National Finalists. The National judges commended the regional judges for their efforts in selecting the high quality Finalist papers forwarded to Seattle. We feel that this Regional and National Judging arrangement provides a good structure for the future as the contest grows.

Outstanding Teams

- Illinois Mathematics and Science Academy, Aurora, IL, USA (6117 & 6129)
  6117 Advisor: Steven Condie
  Team Members: Jason Chen, Noor Michael, Paul Nebres, Kushagra Gupta
  6129 Advisor: Steven Condie
  Team Members: Angad Garg, Joy Qiu, Brice Wang, Daniel Sohn
- Malibu High School, Los Angeles, CA, USA (5967)
  Advisor: Henry Wadsworth
  Team Members: Gannon Earhart, Harry Putterman, Sam Burton, Izzy Putterman
- Mills Godwin High School, Henrico, VA, USA (6018)
  Advisor: Todd Phillips
  Team Members: Linh Bui, Nidhi Kumar, Uday Patil, David Qin
- North Carolina School of Science and Mathematics, Durham, NC, USA (6080 & 6083)
  6080 Advisor: Daniel Teague
  Team Members: Tejal Patwardhan, Vibha Puri, James Chapman, Jonathan Kuo
  6083 Advisor: Daniel Teague
  Team Members: Jeffrey He, Simon Marland, Justin Zhang, Edward Zhuang
- Shenzhen Middle School, Shenzhen, China (5473)
  Advisor: Wentao Zhang
  Team Members: Yuepeng Yang, Pinxuan Yu, Pengcheng Zhang, Fengyu Xie
- The Affiliated High School of Shanxi University, Taiyuan, China (6057)
  Advisor: Lu Liu
  Team Members: Alexander Wu, Yuchen Zhang, Letian Leng, Xinyu Liu

National Finalist Teams

- Changle NO.1 Middle School, Shandong, China (5979)
- China Welfare Institute, Shanghai, China (5679)
- Evanston Township High School, IL, USA (5791)
- Friends' Central School, PA, USA (5372)
- Glenbrook North High School, IL, USA (5955)
- Holy Ghost Preparatory School, PA, USA (5941)
- Kapolei High School, HI, USA (6094)
- NC School of Science and Mathematics, NC, USA (6082 & 6084)
- Oregon Episcopal School, OR, USA (5773)
- Shenzhen Middle School, Guangdong, China (5483)
- Stanford University Online High School, CA, USA (6182)
- The High School Affiliated to Renmin University of China, Beijing, China (5813)
- The Mississippi School for Mathematics and Science, MS, USA (6157)

Common Core State Standards

The director and the judges asked that we add this paragraph. Mathematical modeling is one of the six conceptual categories of the Common Core State Standards for Mathematics in High School. We clearly realize the mapping of this contest to the concept of modeling through the high school curriculum. This contest provides a vehicle for using mathematics to build models to represent and to understand real world behavior in a quantitative way. It enables student teams to look for patterns and think logically about mathematics and its role in our lives. In a future *Consortium* article we will analyze one of our HiMCM problems by mapping it to applicable mathematics required by the Common Core Standards.

General Judging Comments

The Judges' Commentary that follows provides specific comments on the solutions to each problem. In this section, we provide the following as general comments about the solutions from a judge's point of view.

Participants must ensure their papers follow the contest rules posted on the contest site. Papers need to be coherent, concise, clear, and well written. Teams should present their solution and analysis in fewer than 30 pages using at least 12-point font. It is not the number of pages, but the ability to complete all contests requirement and communicate the solution in a succinct and clear fashion that will merit recognition. Students should use spelling and grammar checkers before submitting a paper. Foreign papers should insure that all symbols in tables and graphs are in English.

Judges are best able to analyze a paper when students restate the problem in their own words and clearly preview the focus of their paper. Papers recognized at the National level articulate the problem, present assumptions with justifications, explain the development of the model and its solutions, and then support the results mathematically and communicate them clearly. Modeling assumptions should include only those that come to bear on the solution (this can be part of simplifying the model). Laundry lists of assumptions that do not play in the context of model development are not considered relevant and deter from a paper's quality. Papers need to show clear development of the mathematical model(s), and define all variables. Presenting more than one model, without explaining and choosing the most appropriate model to answer the questions, is detrimental to your paper's success. Teams that merely present a model without showing the development of that model do not generally do well.

Judges also value thinking "outside of the box." This varies from problem to problem, but usually includes model extensions or sensitivity analysis of the solution to the team's inputs. Students need to attempt to validate their model, even if by numerical example or intuition. A clear conclusion and answers to specific scenario questions are also key components. The strengths and weakness section is where the team can reflect on their solution and comment on the model's strengths and weaknesses. Attention to detail and proofreading the paper prior to final submission are also important since the judges look for clarity and style.

Citations are very important within the paper as well as either a reference or bibliography page at the end. Teams that use existing models should reference them during the exposition and not just include a reference/bibliography in the back of the paper. This is true for all graphs and tables taken from the literature. Use in line documentation, footnotes or endnotes to give proper credit to outside sources. We encourage citations within the paper in sections that deal directly with data and figures, graphs, or tables. Most papers did not use references within the paper, yet we saw many tables or graphs that obviously came from websites. We have noticed an increase in use of Wikipedia. Teams need to realize that, although useful, information from Wikipedia might not be accurate. Teams need to acknowledge this fact.

The Future

The contest, which attempts to give all high school students, and especially the under-represented students, an opportunity to compete and achieve success in mathematics, appears well on its way in meeting this important goal. We continue to strive to improve the contest, and we want the contest to grow. Any school/team can enter, as there are no restrictions on the number of schools or the numbers of teams from a school. A regional judging structure offers flexibility to accommodate the number of teams.

These are exciting times for our high school students. Mathematics continues to be more than just learning skills and operations. Mathematics is a language that involves our daily lives. Applying the mathematical principles and concepts that one learns is key to individual and societal future success. The ability to recognize problems, formulate a mathematical model, use technology, and communicate and reflect on one's work is key to success. Students gain confidence by tackling ill-defined problems and working together to generate a solution. Applying mathematics is a team sport!

Advisors need only be motivators and facilitators to encourage students to be creative and imaginative. Success is not only about the procedural technique used, but the conceptual understanding that discovers how assumptions drive those techniques to a valid solution and conclusion. I encourage all high school mathematics faculty to get involved, encourage your students to be problem solvers, make mathematics relevant, and open the doors to future success.

Mathematical modeling is an art and a science. Teach your students through modeling to think critically, communicate effectively, and be confident, competent problem solvers.

International Flavor of the Contest
The award format has changed slightly as the contest continues to grow internationally.

The HiMCM papers are designated as:
Successful Participant
Honorable Mention
Meritorious
Finalist
National Finalist
Outstanding

Contest Dates
Mark your calendars the next HiMCM will be held in November 2016. Registration for the 2016 HiMCM will open in September. Teams will choose a consecutive 36-hour block within the contest window to complete the problem and electronically submit a solution. Teams can learn more and register online at www.himcmcontest.com.

MathModels.org
Powered by COMAP content, Mathmodels.org has been reimagined as a new resource to make math modeling a year round activity. Teachers and students may use the materials found on this site to enrich their classes and help prepare students for mathematical modeling competitions. www.mathmodels.org

The International Mathematical Modeling Challenge (IM^{2}C)
We are pleased to announce the establishment of a new international secondary school mathematical modeling competition. The purpose of the IM^{2}C is to promote the teaching of mathematical modeling and applications at all educational levels for all students. It is based on the firm belief that students and teachers need to experience the power of mathematics to help better understand, analyze and solve real world problems outside of mathematics itself – and to do so in realistic contexts. The Challenge is being launched in the spirit of promoting educational change.

All USA teams that successfully compete in the HiMCM contest and are awarded a designation of Meritorius or above (Meritorius Finalist, National Finalist, or Outstanding) will be invited to compete in The International Mathematical Modeling Challenge, IM^{2}C. From these participants we will select the two top teams to move on and represent the USA in the IM^{2}C international round.

The second annual IM^{2}C is set to take place March 16th, 2016 through April 22, 2016.

To learn more about the IM^{2}C visit www.immchallenge.org

Judges' Commentary

Problem A: Preventing Road Rage

Have you ever been traveling down a highway when you see one or both of these signs?

In some cases, the road simply loses a lane as it enters a less traveled section. In other cases, the road narrows because of construction or roadwork. When this happens on a fairly busy highway, it may be a recipe for road rage*.

Assume we're on a busy two-lane road reducing to a one-lane road. Once a driver sees a sign indicating the right lane is going to end, he or she has a choice: drive in the right lane or drive in the left lane. Often, many drivers move into the left lane fairly quickly. This causes the left lane to become congested and slow down. Some drivers remain in the right lane and are able to maintain their speed (and pass the vehicles in the left lane which are now moving slower).

When a lane closure is approaching, there will be signs indicating the distance of the lane closure as it nears (e.g. 1 mile, ½ mile, 500 ft.). As the end of the right lane approaches, some of the right lane drivers merge into the left lane, while others continue heading toward the lane closure. It's at this point that the probability of road rage rapidly increases. Horns blast and gestures are made. Some cars will pull half way into the right lane in an attempt to prevent any cars from passing only to have another vehicle swerve around them and continue toward the lane closure point. At times, a vehicle will stay in the right lane directly next to, and driving at the same speed as, a chosen vehicle in the left lane in an effort to make the travel toward the lane closure "fair."

Some states have done research on this phenomenon, but there is no consensus around best practices. Your team is tasked to provide a fresh analysis of this issue or add value to any existing analysis. Teams should not simply mimic previous research.

Part I: Analyze the various driver actions and their implications in lane closure situations on a major highway.

Part II: From your analysis in Part I, address and support fair and efficient driver actions:

a) In the case of two lanes merging to one, what is the "fairest" way for drivers to behave as they approach a lane closure? What is the most "efficient" way for drivers to behave as they approach a lane closure? Is there a difference between "fair" and "efficient?"

b) Use your analysis to address and support what drivers should do if a three-lane highway is reducing to two lanes? How should driver behavior change, if at all, if the three-lane highway is reducing to one lane?

c) Discuss and support any differences in fair and efficient driver behavior for lane closures on a secondary road with a 35 mile per hour speed limit versus a highway with a 65 mile per hour speed limit.

Part III: Policies and Practices

a) From your analysis in Parts I and II, develop guidelines for inclusion in the Department of Motor Vehicles driver education materials and signage used by the Department of Highway Safety to encourage appropriate driver actions when approaching a lane closure.

b) In addition to the HiMCM contest format, write a one-page cover letter in support of your guidelines and signage to the Director of the Department of Transportation, urging adoption of your recommendations.

* Road rage is aggressive or angry behavior by a driver of an automobile or other road vehicle. Such behavior might include rude gestures, verbal insults, deliberately driving in an unsafe or threatening manner, or making threats. Road rage can lead to altercations, assaults, and collisions that result in injuries and even deaths. It can be thought of as an extreme case of aggressive driving. (http://en.wikipedia.org/wiki/Road_rage)

Judge's Comments
William P. Fox, Naval Postgraduate School and Kathleen Snook, COMAP

Problem Author: Kathleen Snook, COMAP

Teams addressed this problem through various mathematical models. We envisioned that students might use simulation modeling for traffic flow and vary parameters to measure the effect. Successful teams clearly defined the concepts of both "fair" and "efficient." They then created a model for fairness and efficiency. We saw many mathematical forms. We found student approaches both rich and robust. A number of teams, however, strictly found and then used the "zipper method" from online sites. Those teams had to develop some value added to be successful in the modeling approach category. Some teams used cellular automata or simulated the process using some Monte Carlo Method. We remind those teams and future teams that judges do not read the computer codes. Teams are required to supply either a flow chart or algorithm for the code they develop and include that in the body of the solution. Judges felt that sensitivity analysis would be very useful discriminator in this problem.

Problem B: City Crime and Safety

What can we make of the massive amount of crime statistics collected in major cities? Beyond just reporting numbers, how can we use these data to determine the safeness of a city?

Assume that you and your modeling team live in My City, a large international hub of commerce, technology, finance and travel, with a current population of 2.8 million people impacted by a metropolitan area of an additional approximately 6 million people.

The data set provided (My_City_Crime_Data.xlsx) shows two weeks from police reports in My City and includes crimes listed by case number, date of occurrence, primary and secondary crime descriptions, crime location, whether an arrest was made, whether or not this was domestic crime, and the beat number of the police route.

Part I: Using mathematical modeling, analyze the data. Create a safety rating for My City. Use your safety rating to specify a measure of how safe My City is.

Part II: In addition to the HiMCM contest format, prepare a 1-2 page non-technical report for the Mayor of My City to describe your findings.

Judge's Comments
William P. Fox, Naval Postgraduate School and Kathleen Snook, COMAP

Problem Author: William P. Fox, Naval Postgraduate School

Most teams started off well by using basic statistics to organize and perform some initial analysis of the data provided. Additional exploratory data analysis with charts, plots, and displays helped to see relationships. Teams had to define "safety" and then create a metric to determine the safety of the city. To be successful, teams needed to show the possible range of values and whether large values meant safe or unsafe. A good strategy was to normalize to a 100-point scale. Teams then had to search Internet sites to find data from other cities with which to first test their model and then to compare with My City so they would know whether My City was really safe or unsafe.

General Judge Comments

The judges were excited about the amount of effort put into the modeling. Some papers, however, were unnecessarily long, over 70 pages, without appendices or computer codes. Researching the problem topic on the Internet and then putting everything one finds into a solution paper is not necessary or helpful. We understand that teams will initially conduct research on the topic and will develop several different approaches and models prior to deciding their final solution strategy. Teams should be brief in their explanations of preliminary models that they built upon for their final model. The lengths of the papers are sometimes artificial because teams include every possible model form, even those not used, in their discussion. We encourage teams to include only what they determined useful in addressing the posed problem and answering its questions.

Summaries

The executive summaries are getting better, but there is room for improvement. This has been an ongoing issue since the contest began. We recommend faculty advisors spend some time with their teams and advise them to write a complete and well-written summary. Many summaries appear to be written before the teams start and only state how they will solve the problem. Teams should write their summaries last. Summaries should include the particular questions addressed and their answers. Teams should consider a brief three paragraph approach: a restatement of the problem in their own words, a short description of their method and solution to the problem (without giving any mathematical expressions), and the conclusions providing the numerical answers in context. The executive summary should entice the reader, in our case the judge, to read the paper. If a summary is not strong and with "good" results, management may never read the paper.

These comments also apply to the non-technical papers or "memos" required in most problems. These memos may briefly describe the model in text form, not mathematically, but the focus is on why the model and its results are applicable and important to the reader. The key is to interpret and communicate the results.

Assumptions with Justifications

All assumptions should have some justification of their importance in the modeling process. Assumptions help to simplify the problem so that it can be solved. Facts presented in the problem statement are not assumptions. Teams should not list assumptions that do not impact on the model used or developed. Some teams included new assumptions as they proceeded in their solution. It is better to list all assumptions together at the start of the paper. For example, a team could state assumption #7 and note that it is used in model #4.

Variables and Units

Teams must define their variables and provide units for each of them. Teams should define and list variables when the model is developed. When variables are listed all together at the start of the paper, judges must go back and forth to understand each model development.

Model

Teams need to show a clear link between the assumptions they listed and the building of their model or models. Too often models and/or equations appeared without any model building effort. Teams must reference equations taken from other sources. Teams should show how the model was built and why it is the model chosen. If teams present several models, each should build on to the previous. Presenting several model forms hoping to WOW the judges does not work. We prefer to see sound modeling based on good reasoning.

For example, in the car-merging problem many students immediately presented the zipper merge model without any discussion as to why it was appropriate. Often the zipper model, graphs from the Internet and background material was lifted from sources and not referenced. Anything not created by a team must be referenced.

Model Testing

Model testing is not the same as testing arithmetic. Teams need to compare results or attempt to verify (even with common sense) their results. Teams that use a computer simulation must provide a clear step-by-step algorithm. Lots of trials and related analysis are required when using a simulation. Sensitivity analysis should be done in order to see how sensitive the simulation is to the model's key parameters. We compliment those teams that related their models to real data.

For example, consider the crime problem where we provided information about many types of crimes. For example, assume there were 15 murders in 1,000 crimes. Is the true rate 15/1,000? What are the chances that more murders were committed and not found yet? Thus, analysis should be performed on rates greater than 15/1,000 to measure that impact. This is sensitivity analysis. Both problems were open for good sensitivity analysis (also deemed "what if?" analysis), but the judges found few papers doing any sensitivity analysis.

Computer-generated solutions

Many papers used extensive computer code. Computer code used to implement mathematical expressions can be a good modeling tool. However, judges expect to see an algorithm or flow chart from which the code was developed. Successful teams provided some explanation or guide to their algorithm(s)–a step-by-step procedure for the judges to follow. Code is not read for papers that reach the final rounds unless the code is accompanied by an algorithm that is clearly and logically explained. A complete algorithm, if code is used, is expected in the body of the paper with the code in an appendix.

Graphs
Teams can very effectively use graphs if presented well. All graphs need a verbal explanation of what the team expects the reader (judge) to gain (or see) from the graph. Legends, labels, and points of interest need to be clearly visible and understandable, even if hand written. We realize that teams may use graphs, charts and diagrams found on the Internet. When doing so, teams must clearly attribute the source of the illustration. For example, in Problem A, many teams used an identical diagram of cars merging without a source reference. In the final judging, we use both printed and electronic copies of submissions. Thus, references to color in a graph were not helpful when using a non-color printed copy. Teams should restrict their graphs to various gray-scales and perhaps differing markings (line, dashes, dots, etc.).

Strengths and Weaknesses
Teams should be open and honest here. What could the team have done better? How could the model be better? In this section, teams can review the assumptions they made and how they impacted strength or weakness of the model.

Conclusions
This section deals with more than just results. Conclusions might also include speculations, extensions, and generalizations. This is where all scenario- specific questions should be answered. Teams should ask themselves what other questions would be interesting if they had more time and then tell the judges about their ideas.

References
Teams may use references to assist in their modeling, but they must also reference the source of their assistance. References should be cited where used and not just listed in the back of the paper. Teams should also have a reference list or bibliography in the back of the paper. We request that teams use some in-line reference to graphs, figures, and direct explanation of materials taken from other sources.

Adherence to Rules
Teams are reminded that detailed rules and regulations are posted on the COMAP website. We remind teams that only inanimate resources may be used. Teams cannot call upon real estate agents, bankers, hotel managers, or any other real person to obtain information related to the problem. Additionally, the 36-hour time limit is a consecutive 36 hours.

2015

18th Annual High School Mathematical Contest in Modeling (HiMCM) Summary Sheet

Team Control Number: 5967

Problem Chosen: B

Summary Sheet:

Crime is a problem shared by essentially all major metropolitan areas. It is often a priority of local and city governments to reduce the spread of criminal activity within their jurisdictions, as a city's safety is often an indicator of its economic success and the happiness of its citizens. As such, it is important to be able to gauge a city's safety with respect to crime in comparison to others, in order to determine if a city is effectively combatting the crime within its borders. This is the purpose behind our mathematical and computational model, which considers many different factors contributing to the safety of a city.

Our model generates a singular, easily interpretable 'crime rating' that can be compared to that of other major cities' in order to judge the overall safety of citizens. To do this, our model was constructed keeping in mind that we would need to be able to evaluate My City as accurately as possible utilizing the massive amount of information provided to us. At the same time, however, we would also need to be able to compare the safety of My City to that of other real-life cities which may not have as readily available data. As such, the model we designed is not only accurate, but flexible.

We constructed our model around the following parameters that would contribute to a city's safety: the number of crimes taking place over a given period of time, the number of people affected by these crimes, the severity of different crimes, the efficiency of police in arresting criminals, the proximity of different crimes to residential areas, and the relative distribution of crime throughout the city. In order to take into account all of the crimes reported and listed in the given spreadsheet, we utilized Microsoft Excel and Java programming to quickly analyze the data and yield results in a timely manner. We were also able to use far less comprehensive statistics from online sources in the same model to make comparisons between My City and real-life cities in order to determine safety.

The simplest version of our model generates a crime rating of 4.4 for the general metropolitan area of My City, and when additional factors are considered, the crime rating reaches 17.47. Downtown Detroit, an area historically known for its crime, received a basic score of 37.7, and received a score of 149.6 when additional factors were projected. Minneapolis, a city deemed the safest major city in the United States by Forbes magazine, received a basic score of 2.6, and received a score of 10.32 when additional factors were projected. The advanced scores required projections for the real-life cities because sufficient data was not available to use the full model, which requires data on the distribution of crime and the rate of arrests. For context, Seattle's greater metropolitan area received a basic score of 2.8 and Los Angeles county received a score of 3.1. This shows that My City is not nearly as unsafe as downtown Detroit but is still lagging behind some major cities in terms of safety.

Dear Mayor:

We understand you were interested in a professional evaluation of your city's safety. We understand the nature of your request, as such a measure would not only provide insight to you about future city planning, but would also be a compelling statistic to aid prospective citizens with their real estate decisions. In our model we took into account many of the factors contributing to public safety that we could derive from the data you provided to us in the Excel spreadsheet. Utilizing Excel and our own programming, we were able to create a model that would generate a 'crime rating' for your city that could be compared with other major cities' to determine the safety of your citizens.

We used numerous other cities as points of reference with which to compare yours. In order to do this we had to limit the types of data we could include in our model, because it was difficult to find detailed statistics on all crime for additional cities. Initially we considered only total population, the amount of each type of crime, the severity of different types of crimes, and the number of people victimized. Using these factors alone, our model yielded a crime rating of 4.4 for your city. To put this in context, Minneapolis, one of the safest cities in the country according to Forbes magazine, received a rating of 2.6, and downtown Detroit, a city historically known for its criminal activity, received a 37.7. We also conducted additional investigations on the Seattle metropolitan area and Los Angeles county, which received comparable scores of 2.8 and 3.1 respectively. Thus, while your city isn't ranked quite as safe as the safest in the nation, it is still far safer than downtown Detroit. This isn't to say that your city is without crime, however. In the process of calculating the overall crime rating we also found that over 10,000 crimes took place in your city in just a short two week period, with an average of 714 per day. By comparison, Minneapolis has an average of 64 crimes per day- if you account for the population difference, if Minneapolis had the same population as your city, there would be approximately 149 crimes per day. This shows the degree to which improvements could be made for a city the size of yours. While My City is clearly in good hands, we would still recommend that you pursue further measures to combat crime in your city.

Although these numbers alone can provide you a good understanding of the situation of crime in your city, we expanded upon our results with more factors to provide you with a more accurate rating. As a consequence of this consideration of additional factors, however, it becomes more difficult to make reliable comparisons to other cities. These new factors included the uniformity of crime distribution, the percent of people arrested for their crimes, and the locations of the crimes with respect to residential areas. The safety of your city, we figured, would also depend on these factors. With these new metrics we calculated a new crime rating of 17.4 for your city. This value cannot be directly compared to the aforementioned safety rankings of Minneapolis and Detroit because the number of crimes considered and the new metrics make it distinguishable. We did provide theoretical projections you could use for other cities, which include 10.3 for Minneapolis, and 149.6 for Detroit, but these reference points are somewhat inaccurate because they aren't based on real data from these cities. Ideally, you should contact mayors of the other cities in order to procure more accurate crime statistics, which we can then use to provide better and more accurate reference points with which we can compare your city. This will allow you to get a better sense of how safe your city is in the context of the safety of others.

We hope that you find the flexibility and efficacy of our model compelling, and adopt it for future application; we also suggest asking other cities for data on their crimes to help improve the accuracy of our model.

Thank you for contacting us to create this model for you,
-Team 5967

1 - Introduction

1.1 Background

Over a period of time, based on individual police reports and other methods of data entry, it is possible to collect and organize large amounts of information about the abundance and types of criminal offenses that take place throughout a city. These statistics are important in how they can be used to judge the overall safety of a city, which is an important factor that contributes to economic sustainability, population growth, and overall happiness, among other things. My City is a hypothetical city which has a downtown population of 2.8 million people and an additional 6 million people living in the surrounding metropolitan area, yielding a total population of 8.8 million people. The mayor of My City has provided a spreadsheet detailing all of the reported crimes within My City over a two-week period, in hopes that we would be able to quantify and evaluate the safety of My City as a whole.

1.2 Restatement of Problem

The inherent challenge in providing a concise analysis of My City lies in the massive amount of information that must be analyzed to do so. Our group decided that in order to create an applicable ‘safety rating,’ a mathematical model would be necessary. This model would have to take into account all of the factors that contribute to a city’s safety, which may or may not have been present in the given spreadsheet. The safety rating must be easily interpretable by the mayor of My City and must also be comparable to the safety rating of other well-known cities. The accuracy of our model as a whole can be gauged by its ability to accurately rate existing cities and by the number and relevance of factors that are considered in judging the city.

2 - Definitions, Assumptions, Justifications, and Variables

2.1 Definitions

Here we will define terms that will be used throughout the paper:

Crime rating: This is the singular value that we are attempting to calculate which will describe the impact of crime on a city. It will be used to determine which cities are safer than others. It is represented in our equations as C.

Beat: For all intents and purposes in this paper a beat is a region within a city within which crime takes place.

Victimization: Victimization occurs when one is profoundly affected by the consequences of a crime, even if one is not explicitly targeted. An example of this is a close relative of a murder victim, whose lifestyle and well-being is drastically changed as a result of the crime.

2.2 Assumptions and Justifications

Below we list some assumptions that were necessary to make in order to expediently and effectively construct our model, along with justifications for each assumption:

Assumption 1: Beats that are not referred to in the given spreadsheet do not exist.
Justification: The mean number of crimes in any given beat was calculated to be 40, and the lowest number of crimes for any given beat was 5. As a result, if beats did exist in the city where crime did not take place, there would likely be very few.

Assumption 2: The data given is for both the main city and the surrounding area described in the problem. All 8.8 million people are included.
Justification: There was not enough information to determine which crimes took place in or out of the city, and as such we assumed the crime report included both crimes within the city and in the surrounding area.

Assumption 3: The percentage of crimes taking place in residential areas contributes to the danger/safety rating of a city, but the percentage of crimes which are related to domestic/personal affairs does not.

Justification: If two cities have the same amount and severity of crime, the city in which less crime takes place in the home should be considered safer. However, the risk of victimization of a domestic crime is independent of one's location.

Assumption 4: One single person will not be the victim of multiple crimes during the two week period.
Justification: My City has a population of 8 million people, and according to our model, about 2,000 are victimized by crime in one day. As such, there is a 0.024% chance of being victimized any given day. Multiplying this number by 14 for the two week period and squaring the result yields .001%, which is the likelihood of being affected twice by crime in the same two week period. We considered this likelihood negligible because of its low value and because of the complications that would arise from needing to account for these people.

Assumption 5: A crime in somebody's home is worse than a crime in a public place.
Justification: If two cities have the same abundance and severity of crime, then the city in which more crime takes place outside the home is safer, because it is possible to make the conscious decision to stay at home and lower the likelihood of victimization.

Assumption 6: If a crime is listed with a "no" under the arrest column, the perpetrator escaped.
Justification: Had the perpetrator not escaped, an arrest would have been made.

2.3 Variables

The following variables will be used within our model:

n = number of instances of a particular crime
n_h = number of instances of a particular crime occurring in a home or private residence
n_a = number of instances of a particular crime in which the perpetrator was arrested
n_{total}= total number of crimes taking place in a city over a given period
S = the severity of a particular crime
P = the population of the city
P_v = the number of people victimized by a particular crime
d = the number of days over which the surveyed information was taken
C = Crime Rating
C_{ind} = Crime Rating for each individual type of crime
$C_{with\ deviation}$ = Crime Rating with variation between beats factored in

3 - Model Description

3.1 Utilizing Abundance of Crime Incidents and City Population

Before we began to analyze the massive amounts of information given in the spreadsheet for 'My City,' we figured it was important to first consider what factors would play into our determination of which cities are safer than others. The simplest of these and probably the most obvious is the number of crimes that takes place every day within the city. The easiest way to begin, as such, was to create a 'crime rating' value based on this number:

$$C = \frac{n_{total}}{d}$$

Where C is our hypothetical 'crime rating' value, n_{total} is the total number of crimes in the data given, and d is the number of days surveyed.

This equation is a good place to begin, because a city with more individual crimes every day will have a higher crime rating than a city with fewer crimes every day. Next, it is important to consider the size of the city in order to accurately compare cities of different sizes. According to the previous model, given two cities with the same number of crimes per day, the crime rating of both would be the same, even if one city was far larger. A smaller city with the same number of crimes as a larger city should instead have a larger crime rating, because crime affects a larger percentage of the city's population. Adjusting our model yields the following:

$$C = \frac{n_{crimes}}{dP}$$

Where P is the population of the city surveyed.

3.2 Incorporating Chance of Victimization

After this we had a good scale for judging the overall crime rating of a city, independent of its size. We then decided to utilize the percent likelihood of becoming a victim of a crime as a factor in determining our crime rating. To do this requires knowledge of how many people are victimized by crime as a fraction of the total population:

$$C = \frac{n_{crimes}}{d} \times \frac{E}{P}$$

Where E is the number of people victimized. To find E, it is necessary to find the number of people victimized by each crime that has taken place. E is thus the total result of the sum of each individual crime multiplied by the number of people it affected. This step is based on the idea that multiple people can be victimized by crime even if they are not explicitly targeted. As such, it is incorrect to assume that the number of people victimized by crime explicitly equals the number of crimes committed. To account for this, we assigned a hypothetical average number of people who are victimized to each type of crime:

Type of Crime	Number of People Victimized
Theft	2
Criminal Damage	2
Criminal Trespassing	3
Battery	4
Assault	4
Narcotics	1
Prostitution	2
Robbery	1
Weapons Violation	2
Motor Vehicle Theft	4
Homicide	6
Public Peace Violation	3
Interference with a Public Officer	4
Gambling	1
Criminal Sex Assult	3
Burglary	4
Sex Offense	3
Offense Involving Children	4
Arson	5

Deceptive Practice	1
Kidnapping	6
Liquor Law Violation	3
Other Offense	2
Stalking	3
Other Narcotics Violation	1
Concealed Carry Licence Violation	1
Intimidation	2

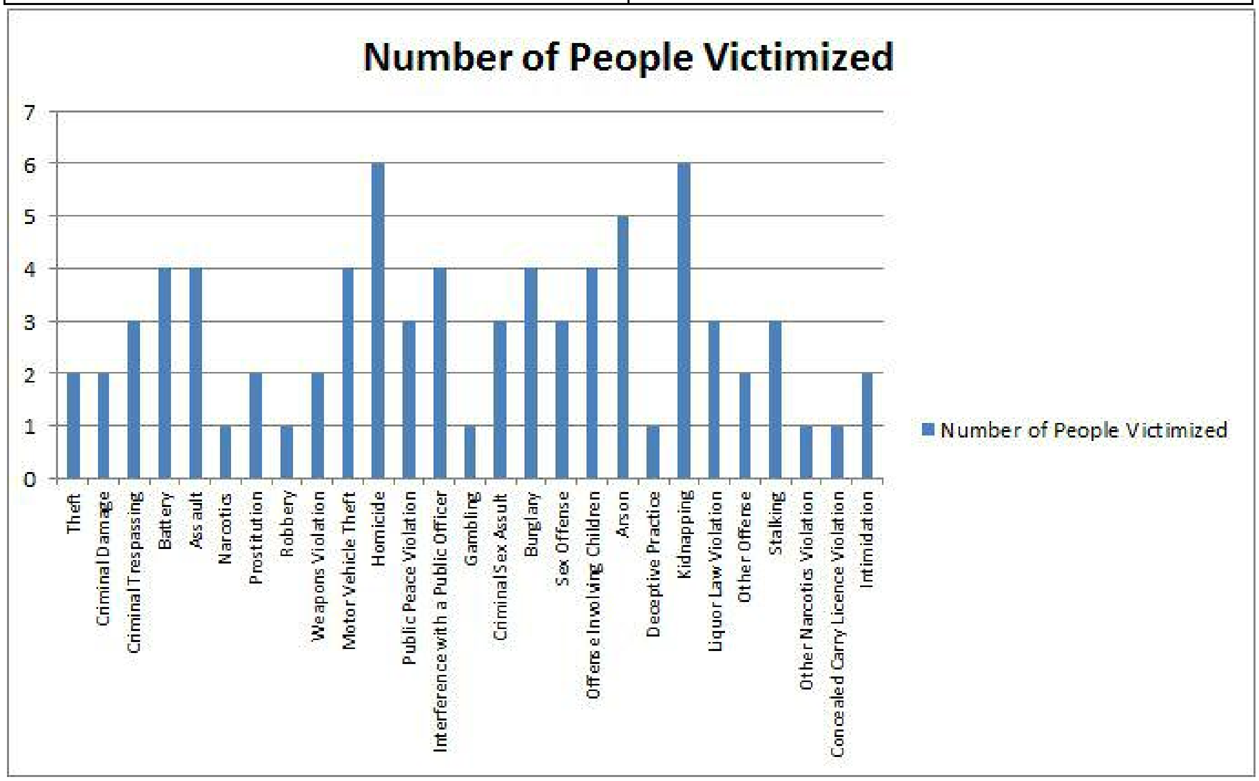

Figure 1a & b: This is the relation between each crime and our estimate on how many people are victimized.

These values can be modified to yield different results.

A mathematical expression for E can be written as such:

$$E = \sum nP_v$$

Where n is the number of instances of a specific type of crime and P_v is the corresponding number of people victimized. The summation will thus add together the products of n and P_v for each type of crime. Since there are 27 different types of crimes in the report from My City, writing a full expression would be impractical, and as such it is easier to use a program to perform this calculation.

3.3 Incorporating Severity of Crimes

After we had accounted for the chance of victimization, we decided that it would also be ideal to include the severity of different crimes in our calculation. If two cities have the same amount of crime and have the same chance of victimization, the city with more severe crimes should be considered less safe. To account for this, we performed research to determine how real-life city governments determine the severity of certain types of crime. Using this research we developed a 1 to 10 scale of severity with which each type of crime could be ranked. This was incorporated into our equation in a manner similar to the way in which we incorporated the number of people affected by a type of crime. It is also important to note that we opted to solve for the crime rating based on individual types of crime rather than the crime rating as a whole:

$$C_{ind} = \frac{(nSP_v)}{dP}$$

Where S is the severity of the crime, and C_{ind} is the crime rating for each individual type of crime; we will later add each individual crime rating to create one overall crime rating. Therefore the use of E is not necessary for calculating the overall crime rating of the city, but we later discovered that E has practical applications in determining the likelihood of being victimized by a crime.

The following table shows the severity we assigned to each type of crime:

Type of Crime	Severity
Theft	1
Criminal Damage	1
Criminal Trespassing	2
Battery	4
Assault	6

Narcotics	2
Prostitution	3
Robbery	7
Weapons Violation	3
Motor Vehicle Theft	4
Homicide	10
Public Peace Violation	2
Interference with a Public Officer	3
Gambling	1
Criminal Sex Assult	9
Burglary	5
Sex Offense	8
Offense Involving Children	8
Arson	6
Deceptive Practice	2
Kidnapping	10
Liquor Law Violation	2
Other Offense	4
Stalking	6
Other Narcotics Violation	2
Concealed Carry Licence Violation	3
Intimidation	5

Figure 2a: This is our relation of each crime to its severity, 10 being the most and 1 being the least severe.

Graph of Above Table:

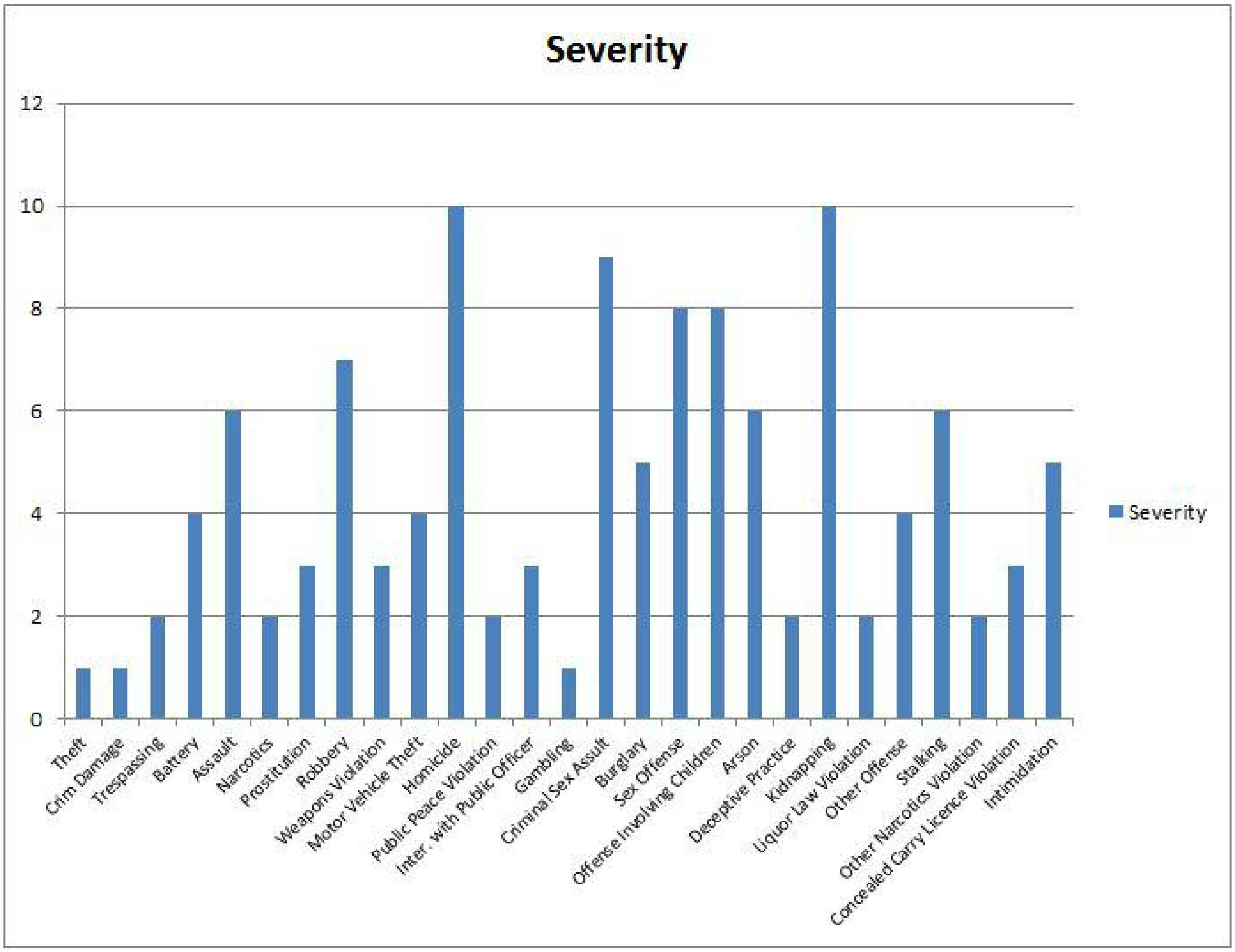

Figure 2b: This is our relation of each crime to its severity, 10 being the most and 1 being the least severe.

The severity values assigned are based upon extensive research of ranking systems utilized by real life city and local governments in determining the severity of different types of crime. Like the values we assigned in Figure 1, these values can be modified to yield different results based on how severe certain crimes must be considered in different situations.

3.4 Incorporating Police Efficiency and Residential Safety

At this point, our model was already very comprehensive, incorporating not only the number of crimes taking place in a given city, but also the severity of each crime, the number of people affected, and how large the city is.

In order to further improve our model's representation of overall safety in our city, we added a factor representative of how many of the perpetrators were arrested. This factor will cause the crime rating to increase if a city's police force is not as effective at apprehending criminals. This is based on the idea that given two cities with equal crime abundance and severity, the city where more perpetrators are arrested should be considered safer. The following equation will give us a number between one and two; if all criminals are arrested after committing a crime, the value will be .75, decreasing the original crime rating. If no criminals are arrested, the value will be 1.25, causing the crime rate to increase.

$$A = 0.75 + \frac{n - n_a}{2n}$$

Where n_a is the number of people arrested for each crime. Hence:

$$C_{ind} = \frac{(A \times n \times S \times P_v)}{dP}$$

To make our model even more accurate we created a new factor that would incorporate the percentage of crimes that took place in residential areas. This is based on the idea that given two cities with equal abundance and severity of crime, the city in which less crime takes place within or near the home should be considered safer. In a similar manner to how *A* was incorporated, we created a new factor *H* whose value will also lie between .75 and 1.25. If no crimes take place in the home, then the value of *H* will be .75, lessening the crime rating. If all of the crimes in a city take place in the home, the value of *H* will be 1.25, increasing the crime rate:

$$H = 0.75 + \frac{n_h}{2n}$$

Where n_h is the number of each crime that took place in somebody's home. Hence:

$$C_{ind} = \frac{(H \times A \times n \times S \times P_v)}{dP}$$

When we evaluate the equation above for each type of crime, and summate them, we find the crime rating for our city:

$$C = \sum C_{ind} = \sum \frac{(H \times A \times n \times S \times P_v)}{dP}$$

When writing the full equation we find:

$$C = \sum \frac{(0.75+\frac{n_h}{2n})\times(0.75+\frac{n-na}{2n})\times n\times S\times P_v)}{dP}$$

For quick reference,

n = number of instances of a particular crime
n_h = number of instances of a particular crime occurring in a home or private residence
n_a = number of instances of a particular crime in which the perpetrator was arrested
n_{total}= total number of crimes taking place in a city over a given period
S = the severity of a particular crime
P = the population of the city
P_v = the number of people victimized by a particular crime
d = the number of days over which the surveyed information was taken
C = Crime Rating
C_{ind} = Crime Rating for each individual type of crime
$C_{with\ deviation}$ = Crime Rating with variation between beats factored in

Since our current model gives us answers all less than 0.0001 we introduced a scaling factor of 10,000 to our equation, simply for ease of understanding for the viewer:

$$C = 10^4 \times \sum \frac{(0.75+\frac{n_h}{2n})\times(0.75+\frac{n-na}{2n})\times n\times S\times P_v)}{dP}$$

This constant of 10^4 causes the value of C to generally lie between 0 and 50 for most cities, as will be explained in our analysis.

3.5 Extension: Incorporating Uniformity of City

When looking at the beat numbers, we realized that it was not possible to draw conclusions about the locations of different beats relative to one another. Initially this impeded our ability to utilize the relative dispersal of crime throughout the city within our model- however, we later realized that using individual statistics for each beat could allow us to draw conclusions about the distribution of criminal activity. To do this, we decided to utilize the standard deviation of the sums of the products of severity, the number of instances of crimes, and number of people affected for each crime in a beat. A larger standard deviation would indicate that, on average, the abundance and severity of crime differed by greater amounts from area to area within a city. A smaller standard deviation, on the other hand, would indicate that the majority of areas in the city had the same amount of crime and severity.

$$x_b = \sum \frac{nSP_v}{d}$$

The above equation expresses a simple way to represent the impact of crime on a certain beat. Ideally, it would be best to observe the standard deviation of crime rate *C* for every beat, but the spreadsheet does not indicate the population of each beat. We also could potentially include the aforementioned factors *A* and *H* for each beat to have a more accurate representation of the impact of crime in each beat, but we ruled that the simplified version of our model without these factors would be adequate for determining the standard deviation of the impact of crime in each area of the city.

$$\sigma = \sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2+\ldots}{n_b}}$$

The above equation takes the standard deviation for all of the beats in a given city, where x_1, x_2, x_3, etc. represent the value for x_b evaluated for each beat in the city. The symbol n_b represents the number of beats in the city.

We discussed the degree to which the standard deviation should affect the overall crime rating, and decided that when factoring the standard deviation into our formula, we should divide it by the mean of the crime ratings. This is called the coefficient of variation. This is very likely to be a value less than 1, and will still remain relatively small even in extreme circumstances with high variability. It is also unitless, which allows it more versatility in being incorporated into our model.

We also discussed whether or not the uniformity of a city would contribute to its safety or danger. Ultimately, we concluded that for safer cities, uniformity should contribute to the city's overall safety, and thus lower the crime rating. This is because cities trying to optimize their safety should avoid harboring individual pockets of crime and relegating it to certain areas. However, for more dangerous cities, we concluded, uniformity should contribute to a city's crime rating. This is because prospective visitors or residents of relatively unsafe cities would value the presence of areas with higher safety. In order to model this, we used the average of the simplified crime ratings for Detroit and Minneapolis, two cities that represent typical perceptions of unsafe and safe cities, respectively. If the crime rating, when calculated without factoring in standard deviation, yields a value less than 25.30, then the standard deviation factor should increase the crime rating. If a value less than 25.30 results from calculating crime rating without factoring in standard deviation, then the standard deviation factor should decrease the crime rating. Keeping this in mind, and adjusting constants in order to keep the uniformity factor from drastically changing the value, the following rules were devised:

$$C_{\text{ with deviation}} = C\left(\frac{1+\frac{\sigma}{mean}}{4}+1\right) \text{ if } C < 25.30$$

$$C_{\text{ with deviation}} = C/\left(\frac{1+\frac{\sigma}{mean}}{4}+1\right) \text{ if } C > 25.30$$

This factor requires a detailed knowledge of the number and type of crimes taking place in each beat within a city, and thus can rarely be applied without a comprehensive report from the city.

3.6 Final Model and Application Guidelines

The final, completed model, which assumes a comprehensive knowledge akin to that provided to us within the spreadsheet for My City, can be written as follows:

$$C = 10^4 \times \sum \frac{(0.75+\frac{n_h}{2n})\times(0.75+\frac{n-na}{2n})\times n\times S\times P_v)}{dP}$$

To factor in the uniformity of crime distribution throughout the city, utilize the following rules as dictated in Section 3.5:

$$C_{\text{ with deviation}} = C\left(\frac{1+\frac{\sigma}{mean}}{4}+1\right) \text{ if } C < 25.30$$

$$C_{\text{ with deviation}} = C/\left(\frac{1+\frac{\sigma}{mean}}{4}+1\right) \text{ if } C > 25.30$$

It is important to keep in mind that *in order* to compare cities using our model, the same information must be utilized in the model for both calculations. As such, to compare crime ratings between two cities where n_a and/or n_h are unknown requires that one utilize the abbreviated version of the model, shown here:

$$C = 10^4 \times \sum \frac{nSP_v}{dP}$$

Because finding detailed information about the arrest rate for different types of crimes and the frequency of crimes occurring in residential areas is exceedingly difficult without a comprehensive report, using the above model is far more convenient for comparing different real-life cities.

Additionally, when using the above model, it is important to note that accurate comparisons between cities can only be made if the *same* types of crimes are used to evaluate for C. For example, if data are not present for prostitution violations in one city, but are present in another, then summing $\frac{nSP_v}{dP}$ for prostitution violations when calculating C for both would cause the second city to appear more dangerous than the first simply because the second city had more data on reported incidents. As such, when comparing two cities' C values, it is important to only summate the terms for each crime that are represented by data for both cities.

3.7 Strengths and Weaknesses of our Model

Strengths:

-Our model accounts for a large number of factors that could potentially contribute to the impact of crime on a city, including the number of incidents, the severity of the crimes, the size of the city, the number of perpetrators that escape arrest, the proximity of crimes to residential areas, and even the distribution of crime throughout the city

-Our model is flexible- if certain information is not available one can simply remove factors and create an abbreviated version of our model that still produces viable crime ratings

-One can effectively compare different cities even if the amount of data available is inconsistent for the two cities

-The model, by default, assigns reasonable and researched weights to the severity of crimes and the importance of arrests that reasonably scales real-life cities based on their data

-Our model was devised around the performance of key cities such as Minneapolis, Detroit, and Los Angeles in order to allow a reasonable placement of other cities and to provide an accurate scale for our results

Weaknesses:

-The model requires at a minimum the knowledge of a city's population, the number of crimes taking place over a certain period of time, and what different types of crime took place and how many of each took place

-Our model assumes a fixed weight and number of persons victimized per crime; to be incredibly accurate, it would be possible to assign a more accurate weight and number of persons to each individual infraction, but this requires careful analysis of very large amounts of data, which could be possible with more time

-The weights assigned to factors A, H, and the standard deviation, as well as the weights of each type of crime were assigned arbitrarily based on extensive research and calculations, and were designed to make reasonable changes in the C value based on our interpretations. Ultimately, the severity of crimes relative to one another and the importance of other factors are entirely subjective and thus cannot be reliably quantified by any standard
-To gather many of the values necessary for computing C, computer programs are necessary to avoid painstakingly large amounts of manual computation and data entry

4 - Analysis

4.1 Spreadsheet Analysis

Our model was designed to rate a city based on how safe it is, given crime reports over a two week period. By taking all of the relevant information for a given city, we could then calculate its crime rating. To gather all of the necessary information and to perform the necessary summation, we decided to use Excel. We believe it is best to utilize multiple models to affirm our results, so in addition to our Excel analysis, we also created a computer program which would parse the individual rows in the original spreadsheet and thus scan through all of the reported crimes over the entire two-week period.

In our Excel analysis, we assigned rows to each type of crime, since our model is essentially based on a sum of calculations for each individual classification of crime. Within each row we used COUNTIF() functions to determine the number of crimes that took place by having Excel search through the original spreadsheet to find all of the crimes that were labeled a certain type and count the number of each. We used similar Excel functions to determine how many arrests were made for each separate type of crime and how many of them took place in locations that would imply they were in residential areas. Once we had these values, we could easily perform the calculation for C in our model, though this wouldn't include the standard deviation factor.

Excel was similarly used to calculate the crime rating using the abbreviated model for other cities using data from online sources about the number and types of crimes.

4.2 Programming Analysis

To calculate the standard deviation we needed to iterate through all of the data and for each beat create a data entry which represented the sum of the severity of each crime times the number of times that crime happened times the number of people that crime affected. These values were taken from the tables that we attached earlier. To store the data that represents the net crime of a beat we used an array where the dimension represents the beat. To add the excel data to our IDE we exported the excel data in a tab delimited format. We looped through each line of the text document such that at the location in the array where the beats where the value equaled the sum of all of the crime data that occurred in that beat. From that point we calculated the standard deviation.

To make sure that the program was effective we used a small subsection of the data set that was provided to test the model for standard deviation/mean. Because the data set was quite small we were able to manually affirm that the program worked.

Below is our flow chart for the program we wrote:

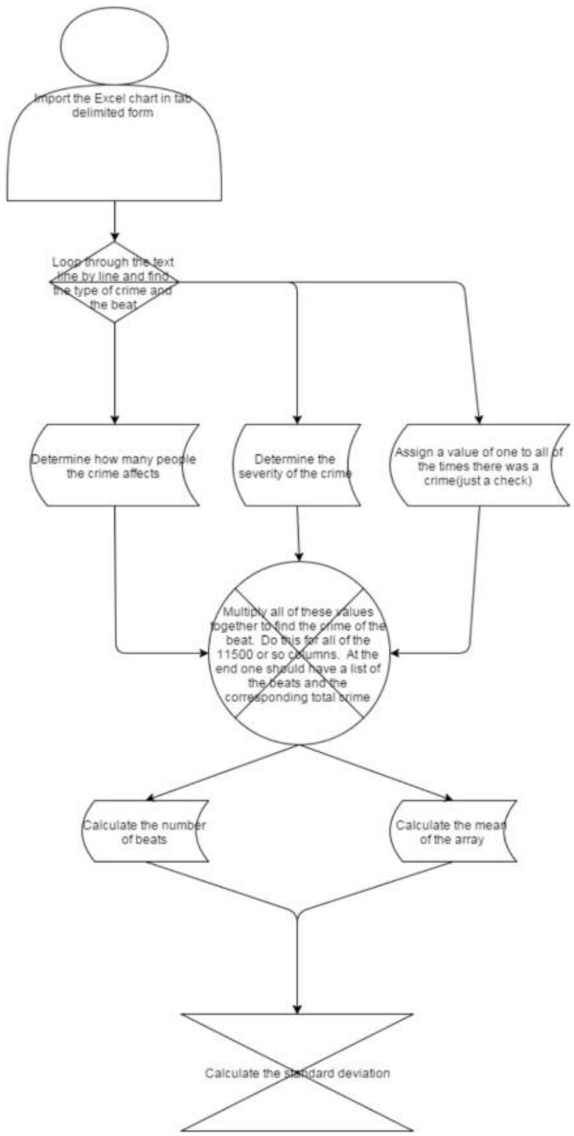

Figure 3: A flow-chart outlining the central strategies we used in our program to calculate the standard deviation.

To test the efficacy of our model we used two real world cities to provide a basis of comparison for the results we got for My City. In comparing My City with other real world cities the essential principle is using the same amount of data from both cities. Often times it is hard to find data concerning a city that is of the same resolution as the data that we were provided. Thus using only data that we have a point of comparison to makes the most sense, in essence this ensures that the d value is the same for the both cities.

4.3 Sensitivity Analysis:

Varying the severity of the crimes:
By doubling the severity of each crime, we also double the Crime Rating. They are directly proportional.
Varying the population of the city:
When all else is constant the higher the population; the lower the Crime Rating.
Varying the efficacy of the police:
When the police department is one hundred percent efficient the result of the other calculations is multiplied by .75; when it is zero percent efficient the result is multiplied by 1.25.
Varying the number of crimes which take place in a residential area:
When the rate of crimes in homes is one hundred percent the result of the other calculations is multiplied by 1.25; when it is zero percent the result is multiplied by .75.
Varying the amount of variation between beats:
When there is no variation the result is multiplied by 0.75; when the standard error is double the mean the result is multiplied by 1.25.

5 - Results and Conclusion

By using a simplified version of the our model and only considering the types of crimes that were documented by all cities:

$$C = 10^4 \times \sum \frac{(n \times S \times P_v)}{dP}$$

We concluded that 'My City' had an overall Crime Rating of 4.4. To put this in context, Minneapolis, the safest city, received a score of 2.6, and Detroit, the most dangerous city, received a score of 37.7.

By using the equation for the percent chance of being victimized we were able to conclude that on a given day there is a 0.023% chance of being victimized, and in a given year there is an 8.83% chance of being victimized.

The following table on factors in the types of crime that all of them shared

	My City (Greater metro area)	Minneapolis (greater metro area)	Detroit (city)	Seattle (Greater metro area)	Los Angeles (county)
Total Grief Per Day	3910	1021	2352	1005	3119
Crime Rating	4.4	2.6	37.7	2.8	3.1

Figure 4: The relation between major cities in the US and My City and their crime ratings.

Given information about the arrest rate, location, and distribution of crime throughout each city, we would be able to add resolution to our results by using our full model- however, since such statistics were only available for My City, the following comparisons were made by projecting the crime ratings for Detroit and Minneapolis linearly with respect to the growth of My City's crime rating. The following table shows how the crime rating for My City changes when more and more of these factors are considered, and how it would compare with that of Minneapolis' and Detroits' crime ratings:

	Not factoring in whether or not the crime was in a home. Not factoring in whether or not the criminal escaped.	Not factoring in whether or not the crime was in a home.	Factoring in both homes and escapes.	Factoring in variance in crime per beat..
Crime Rating For My City	8.606	9.763	13.976	17.47
Projected Rating for Minneapolis	5.085	5.768	8.257	10.321
Projected Rating of Detroit	73.737	83.650	119.747	149.683

Figure 5b: A more in depth relationship between My City, Minneapolis, and Detroit.

Projections are purely hypothetical, but provide some context for understanding our results. Here is a graph of our function and projections.

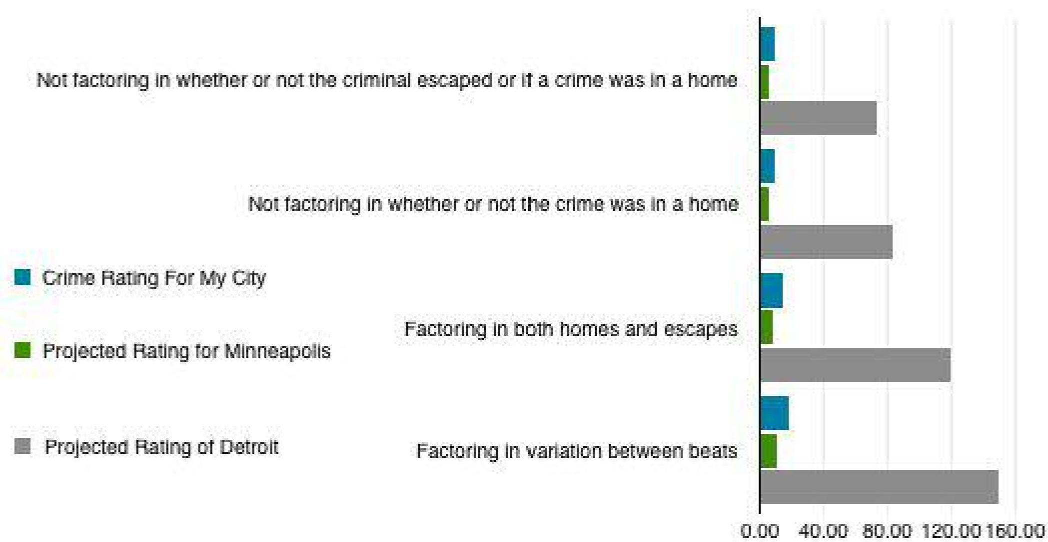

Figure 5b: This is the graph for the table above in Fig. 5.

Our model shows that the city provided is a more dangerous place to live than Minneapolis by a small factor; and is a safer place to live that Detroit by a large factor. This trend continues when we use our full model and the complementary projections. Ultimately, our model shows that My City is safe by comparison to downtown Detroit, and thus is not in danger of being historically remembered for its crime, but steps must be taken in order for the overall safety of the city to match that of most other major cities, such as Minneapolis, Seattle, and Los Angeles.

Earlier in the process we calculated that the overall percent chance that a person will be affected by *any* crime to be 8.8%. This number closely resembles the 12% provided by the Florida Correctional Department, therefore we can conclude that our model is, yet again, accurate.