Towards the Development of Urban Crime Metrics that Match Perception

Abstract This project involves using a data-driven approach to develop crime models and metrics that better match perceived safety indicators. The often-used metrics for studying urban crime, such as murder rate, typically do not match the perceived safety threat that comes from these crimes occurring. This causes a disconnect between analysts studying crime and the residents of the cities who are potential crime targets. Although much work on crime data has been done in the past, much of it involved predicting actual cases of crime. The research into matching urban metrics to perception has been very limited, with only a few previous studies that aim to rank relative safety rates of regions. Work related to urban crime metrics largely focused on scaling metrics that normalize values in proportion to population. This project would allow for creation of novel metrics that aim to better match the perceived reality and thereby providing better information to policy makers as to how to improve urban safety in a meaningful manner. The project would involve using data from the Chicago data portal, which contains both crime incidents and perceived safety scores over time. The data would be used to calculate values of several crime metrics and compared against neighborhoods. This would include a proposed new metric that is based on the assumption that any crime affects all people almost equally within the area where it occurs. This is different than the commonly quoted crime rates, which divide crimes among all persons equally at a global level. The trends in these different metrics would then be compared against safety scores to determine which ones produce better correlation. The most correlated metric would then be compared against a predictive model derived from machine learning, and tested for accuracy. It is hypothesized that the proposed metric would offer a better match to perceived perception, while being a simple and easily interpretable metric. However, the machine learned models might offer better accuracy, which will be tested. Students would be tasked with validating this hypothesis and experimenting with the different models and metrics to determine possible improvements. Students working on this project will have to first familiarize themselves with current literature on the subject. They will then use Python and related data scientific packages to process the data, derive the values of the needed metrics, and use machine learning to produce and test predictive models. They can then proceed to experiment with other methods. Lastly, they would be responsible for documenting the processes and coming up with ideas for trying to improve results.

I. Introduction and Background

Public perception of crime has often times not been reflected in statistics shown in the news. In a recent article by NBC News, Ken Dilanian writes in a frustrating tone about how Americans are simply wrong about crime trends [1]. The author cites Gallup polls showing most Americans believe local crime has been increasing in 2023, while it has actually decreased [2]. This is framed as an irrational human behavior that is not grounded in reality. In both the NBC News and the Gallup articles, authors blame the perception on media coverage and indirectly, on the inability of humans to correctly assess the crime levels. However, the same Gallup polls that were cited in the articles show that levels of crime perception have actually been slightly lower in 2023 than 2022. This suggests that while the perceived value of crime levels may not be matching to the crime statistics, it seems to trend with it.

While humans may certainly have a hard time accurately assessing current or past crime levels based on their knowledge and media coverage may play a role in altering perception to some degree, the bigger issue might be the way crime statistics are computed. To correctly assess the impact of crime, one must take into account how crime affects people individually. With statistics like the murder rate, the number of crimes is divided equally by the whole population. A city with ten times the population of another, with the same number of crimes, will have a tenth of the crime rate. In

such a case, people who argue that human perception of crime is irrational would likely state that the proper perception of crime levels should also be a tenth of the smaller city. The argument is that, assuming crime happens at random, a higher population would linearly scale the chance of being a victim. People should thus feel a lot safer in a more populated city than a smaller one with

the total effect of crime on individuals. To see this, consider the scenario where a violent crime has recently occurred in front of a high-rise residential building. Since this crime is so close to people living in that building, it will likely affect each resident equally. Every resident may now be more careful about leaving the building and may also invest in protection and surveillance devices. Now, consider a different scenario where the same residents are now living in several smaller buildings, spread over a large area. If we still have a single crime occurring in front of one of the buildings, the residents of other buildings may not be affected as much. In both scenarios, the crime rate would have been the same, but the perceived safety threat would have been much lower on average in the

the circle formed by the dotted outline, we take the square root to estimate the distance value r, within constant of proportionality (

[13] H.W. Kang and H.B. Kang, Prediction of crime occurrence from multi-modal data using deep learning *PloS one*, *12*(4), e0176244. 2017.

[14] K. Jenga, C. Catal, and G. Kar, Machine learning in crime prediction *Journal of Ambient Intelligence and Humanized Computing*, *14*(3), 2887-2913. 2023.

Funding related to your proposed research: **none**

Criteria for student applicants: **some programming experience (esp. Python) required**

List of seminar topics you are willing to cover (Selecting from list below or suggesting your own):
- o Ethics in Research/Scholarship
- o Literature Search and Library Resources
- o Problem-Solving Skills and the Scientific Method
- o Presentation Skills
- o Technical Writing
- o Resume Writing and Marketing
- o Preparing for Graduate School
- o Interview Skills
- o Mock Presentation Supervisor (Practice for Symposium)
- o Other: **Data Analysis and Data Management**