Case Study 1
SYS6018

Erik Langenborg :: el4gf
Brady Fowler :: dbf5sd
Tianye Song :: ts7fx
Lev Zadvinskiy :: lz4uc

Identify the Problem

Increased data collection in recent decades has allowed researchers and public servants to better describe the occurrence of crime. The creation of more robust datasets has potential to reveal underlying behavioral patterns related to the rate of criminal offenses and elicit policy prescriptions tailored to specific, measurable factors. The goal driving these efforts is to build a long-lasting public safety environment and efficiently allocate resources intended to assuage crime.[1] Qualitative social theory can provide a launching point for quantitative analysis, but our goal is ultimately to better understand which *controllable* factors can be tweaked to minimize crime rates in a measurable way.

Transforming raw information into prioritized actions is a problem worth solving since crime impacts society at large, determines the planning of law enforcement, and the budgeting of politicians. Currently, police forces are adopting real-time crime prediction applications to dispatch officers to the most essential locations and re-route patrols as anticipated risks change. These algorithms may be improved by the addition of previously unmeasured factors uncovered in further research.

The surrounding events and factors that may contribute to crime include: demonstrations, sporting events, natural events/disasters, traffic levels, weather, holidays and many more. In the section titled "Understanding the State-of-the-Art" we review and bucket selected literature exploring public safety in more detail. Generally speaking, these studies seek to "explain" the factors that contribute to crime or "predict" where and when crime is most likely to occur next based on historical patterns.

Define Objectives & Metrics

Our goal in this case study is to examine occurrences of public safety threats, such as theft, homicide, rape, or assault. Identification of trends and predictions can assist in reduction of these undesirable events. As noted above, the correlation of problem areas, events, or other factors with crime incidents, public services can provide insights to the mechanisms that lie behind crime.

Measurements used to calculate the frequency and location of crimes are available from a variety of sources ranging from national datasets like the FBI's Uniform Crime Reporting Program to local law enforcement agencies like the Washington DC Police Department's Crime Incidents (ASAP) database. Ancillary datasets reflecting the factors that may be correlated to crime like demographic mix, marriage and employment rates and property values can be retrieved from think tanks like PEW Research and Crime Prevention Research Center or government bodies like the U.S. Census Bureau and Bureau of Labor Statistics.

Measuring these public issues and potentially correlated influences require quantifiable metrics. Rates of offenses in geographic areas present the most basic indicators, providing a view of 'what,' 'where,'

---

[1] Many analytics companies like IBM have stepped into the "public safety" space with the aim of uncovering some "hidden insight" that may be the key to public safety: http://www.ibmbigdatahub.com/presentation/what-data-driven-government-public-safety

and 'when'. Other side effects can be used to verify the success of police efforts, such as mean incomes, business infrastructure projects, or public satisfaction surveys.[2] The table below demonstrates a breakdown of metrics that past research suggests may be correlated with the "primary event" of crime:

Primary Event - Crime	Discrete Events	Temporal Events	Infrastructure Conditions	Social Conditions
Robbery / Theft	Demonstrations	Holidays	Strength of Economy	Demographic Mix
Homicide	Sports Events	Seasonality	Unemployment Rates	Gender Distribution
Hit & Run	Natural Events	Time of Week	Consumer Spending	Wealth Inequality
Vandalism / Graffiti	Rush Hour	Bar Closing Time	# Law Enforcement Agents	
Sexual Assault			# Floating Population	
Assault Patients in ER				

Understanding the State-of-the-Art

A survey of peer-reviewed literature reveals that research exploring factors related to criminality can generally be placed in one of four main domains: Discrete Events, Temporal Events, Infrastructure Conditions, and Social Conditions.

Discrete Events

Discrete Events are intermittent occurrences that may influence local public safety, such as weather systems, sporting events, political protests, or entertainment venues. These conditions are examined for the potential to influence crime rates during their occurrences and are typically organized into durations and locations.

Crime records also contain a temporal and geographic component, allowing the two events to be compared: one paper[3] in particular correlates crime rates against sporting events. Data is pulled from civil sources and sport times are collected from publishers such as the NFL and the NBA. The particular study used a least squares regression with a model of:

$$\text{Crimes} = \beta_0 + \beta_1 \text{Status} + \beta_2 \text{Month-Day-Hour} + \beta_3 \text{Year}$$

The influence of weather on crime is a more thoroughly explored version of discrete events, examined for this report in two papers[4, 5]. Information on climate conditions was collected from the NCDC, injury reports were pulled from hospital trauma registries, and crime statistics were collected from municipal authorities. Applying negative binomial regression models and controlling for individual factors, the paper found a strong relationship (IRR = 1.0128, p < 0.001) between rising heat and violent crime.

---

[2] "Predicting Crime: A Review of the Research." Stephen Schneider. 2002. This research review catalogues many factors discussed in "crime prediction" studies since 2000.
[3] "Entertainment as Crime Prevention: Evidence from Chicago Sports Games." Hannah Lacquer, Ryan Copus. April 25, 2014.
[4] "Investigating the relationship between weather and violence in Baltimore, Maryland, USA," Samuel Michel et al. Jan 2016.
[5] "Crime, Weather, and Climate Change." Matthew Ranson. May 2014.

Precipitation (IRR = 0.9804, p = 0.013) and snow (IRR = 0.9827, p < 0.001) provided a mild deterrent to violence and crime related injuries. No correlation was detected for wind speed.

The study used a precision of half-hour increments and concluded that crimes decreased in Chicago before, during, and after popular broadcast games, with varying degrees of intensity. The various types of crimes denoted react differently to each of the classes of discrete events (for example, NBA games are related to the reduction of violent crime, but had no statistically observable impact on drug crimes).

Both types of discrete events covered in the relevant papers left two obvious avenues of improvement: greater degrees of temporal and geographic precision. The examined sport events were broadcast nationwide, eliminating any ability to correlate location effects of local people attending a sports arena. Furthermore, the authors noted that crime statistics were often mislabeled in time or location by the police departments, hampering more detailed analysis.

Temporal Events[6]

In their paper "Even Criminals Take a Holiday: Instrumental and Expressive Crimes on Major and Minor Holidays", researchers Ellen Cohn and James Rotton investigated the idea that crime rates would vary around holidays.[7] Specifically, they hypothesized that violent crimes would increase in frequency while property crimes would decrease during holiday seasons. The researchers collected 911-call data from the Minneapolis Police Department and categorized the resulting offenses into expressive crimes (assaults, disorders and domestic violence) versus instrumental crime (theft, burglary and robbery), and coded holiday periods with binary dummy variables. Further data from the National Weather Service determined whether or not the holiday-effect was merely due to weather factors.

Since some factors (like disorderly conduct) exhibited significant skew in distributions, the researchers applied a logarithmic transformation or square root transformation to either "reduce the variable's coefficient" or create a normal distribution, respectively. The resulting regression models showed that only "major holidays" like Christmas and New Years exhibited significant relationships with crime rates. Since the study was restricted to Minneapolis in 1985-87, the findings may not be generalizable to other cities or countries. Furthermore, the study did not attempt to control for other geospatial factors like pre-existing crime rates or other important demographic factors which is a notable area for improvement.

Two Norwegian researchers, Ingeborg Rossow and Thor Norstrom explored the relationship between bar closing time and violent crime.[8] They collected data from 2000-2010 on changes in bar closing time and crime rates, then applied cross-sectional time series analysis to determine that changes in "last call" timing is associated with a crime reduction of nearly 16%. Their time-series regression model further accounted for city level dummy variables intended to remove any intercity covariation or unobserved differences. This study has implications for policymakers who may consider shortening

---

[6] Other temporal events like "moon phase" have been colloquially theorized to impact crime rates. At least one paper (http://www.sciencedirect.com/science/article/pii/S0047235210000589) explores this hypothesis.
[7] "Even criminals take a holiday: Instrumental and expressive crimes on major and minor holidays." Ellen Cohn, James Rotton. August 2003.
[8] "The impact of small changes in bar closing hours on violence.The Norwegian experience from 18 cities." Ingeborg Rossow, Thor Norstrom. June 15, 2011.

drinking hours as an attempt to curb crime spurts. Since the authors do not list all city factors accounted for, it is unclear whether or not they accounted for factors like mode of transportation to the bars in question or population density which could be highly related to after-closing time criminal activity. Furthermore, the method could be improved by increasing the relatively small number of observations (n=774) through additional data collection.

Infrastructure Conditions

A large proportion of published literature related to public safety involved the state of underlying infrastructure conditions and their effects on recorded crime. The focus of these papers was primarily on the relationship between the rates of criminal activity (mainly burglary and theft) as explained by factors like proximity to urban transportation, floating population (i.e. individuals with temporary residence like tourists, students or commuting workers) and the size of the local police force.

One paper exploring police force size as a factor impacting crime rates systematically surveyed 62 studies and 229 findings from papers published 1968 to 2013.[9] The goal was to prove the general assumption held by many policymakers: that the expansion of the police department would directly reduce crime occurrence. The authors concluded that the size of the police force had little impact on levels of crime, further arguing that *advancements* in policing practices would likely have more impact on crime than a mere expansion of force size.

A group of researchers in China examined the potential effect of roads urbanization, transport development and complexity of land use on robbery rates in a small Southern town in China.[10] The researchers, L. Weihong, W. Lei and C. Yebin compiled spatial temporal, or topographical, data from 2008-2012 and conducted cross-analyses on property crime hot spots to develop a forecasting model. The quantitative cross analysis utilized by the researchers combined daily activity logs, geographical location, traffic data and levels of crime activity; produced results showed significant correlation between observed variables. With accurate data, the forecasting model developed in this study could be used by police force as an effective tool to analyze any "non-passion" crime types and develop better patrol routes.

A third study using geospatial data by E. Groff, D. Weisburd, and S. Yang examined crime trends on a micro scale.[11] Police incident reports from Seattle, Washington from 1989-2004 were used to establish temporal crime trajectories (a roadmap of daily crime activity) across the city. Researchers observed variability in the levels of crime across various sections of Seattle, arguing that, in order to understand crime patterns one must examine **localized clusters** (i.e. streets and neighborhoods) rather than looking at the large area trends. The results of the analysis indicate strong variability between high crime rate and low crime rate streets, a relationship that would be *easily* omitted at a macro level. Researchers concluded that further examination of small area trends is warranted using more robust information system that are available today to better understand crime and delinquency.

---

[9] "Conclusions from the history of research into the effects of police force size on crime—1968 through 2013: a historical systematic review" YongJei Lee, John E. Eck, Nicholas Corsaro. August 19, 2016.

[10] "Spatial−temporal forecast research of property crime under the driven of urban traffic factors." Elizabeth R. Groff, David Weisburd, Sue-Ming Yang. March 30, 2016.

[11] "Is it Important to Examine Crime Trends at a Local "Micro" Level?: A Longitudinal Analysis of Street to Street Variability in Crime Trajectories" Elizabeth R. Groff, David Weisburd, Sue-Ming Yang. January 1, 2010.

Notably, few of the "infrastructure" centered studies included factors describing the demographic and social conditions of a region but most do note the necessity of deeper examination using additional descriptive datasets to isolate the effects of infrastructure and demographic factors.

Social Conditions

Qualitative social theories have long suggested that social "stability" may be a strong crime deterrent on the individual level. The strength of social structures like marriage, wealth distribution and neighborhood cohesion have been studied both on the individual and community level in relationship to delinquency rates.

Researchers Jeremy Porter and Christopher Purser combined national crime data from the FBI's UCR Database with US Census Bureau data to estimate the spatially bound correlation between "social disorganization" and criminal behavior.[12] Their paper builds on previous well-received literature; rather than modelling the propensity for committing crime on the individual level (i.e. a child is less likely to engage in criminal behavior if their parents were married) they combine measures of racial heterogeneity, family disruption, socio-economic status, and urbanization to proxy "social disorganization" and analyze the correlation against various crimes. They conclude that communities with lower levels of disorganization also experience lower levels of crime.

Other research investigating the social factors of inequality on crime rates intends to provide a more granular understanding of variations between community than a more macro-view than the "social disorganization" research described above. Alyssa Chamberlain and John Hipp used the National Neighborhood Crime Study to explore how crime rates vary across "resource deprived neighborhoods," accounting for city-wide factors like poverty levels, income levels, median home value, percent of homeowners versus renters, racial heterogeneity.[13] Their models exhibited a statistically significant relationship between higher levels of inequality *within* a geographic tract and higher levels of violent crime. Notably, the authors stressed the importance of accounting for the attributes of nearby neighborhoods and use "spatially lagged" variables to account for "spatial spillover" of both correlated driving factors and anti-crime efforts.

[INSERT REST OF REPORT HERE!]
[Rough idea is to build off the "floating population" and "bar closing time" papers to see if we can use Uber surge data (proxy for high number of floating populants in an area) being related to higher levels of "personal crimes" like assaults/muggings. We have attained geotagged DC crime data as well as Uber Surge data procured by Washington Post analysts.]

---

[12] "Social disorganization, marriage, and reported crime: A spatial econometrics examination of family formation and criminal offending." Jeremy Porter, Chistopher Purser. September 2010.

[13] "It's all relative: Concentrated disadvantage within and across neighborhoods and communities, and the consequences for neighborhood crime." Alyssa Chamberlain. John Hipp. November 1, 2015.

Appendix
Unfamiliar Items

Prof Gerber - We kept a running list of concepts and methods that we weren't familiar with. For your convenience we split them into those groups and ordered in terms of perceived increasing difficulty!

Concepts:

- Logging Variables that have significant skew versus Square Rooting Others to get a more normal distribution"
- The combining use of dummy variables and hierarchical regression
- Gini Coefficient
- Global Moran's I
- Heteroskedasticity-robust standard errors
- Poisson regression with semi-parametric variables

Methods:

- Kernel Density Functions
- Kernel Decay Functions
- Eigen Values
- Negative Binomial Regression
- Spatially Lagged Variables
- Autoregressive integrated moving average (ARIMA)
- Space-time autoregressive model
- First difference generalized method of moments
- Neural networks (GA-BP)
- Ripley's K-function and bivariate K-function
- Dasymetric mapping

Citations

"Conclusions from the history of research into the effects of police force size on crime—1968 through 2013: a historical systematic review" YongJei Lee, John E. Eck, Nicholas Corsaro. August 19, 2016. http://link.springer.com.proxy.its.virginia.edu/article/10.1007%2Fs11292-016-9269-8

"Crime, weather, and climate change" Matthew Ranson. May 2014. http://www.sciencedirect.com/science/article/pii/S0095069613001289

"Entertainment as Crime Prevention: Evidence from Chicago Sports Games." Hannah Lacquer, Ryan Copus. April 25, 2014. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2429551

"Even criminals take a holiday: Instrumental and expressive crimes on major and minor holidays." Ellen Cohn, James Rotton. August 2003. http://www.sciencedirect.com/science/article/pii/S0047235203000291

"Investigating the relationship between weather and violence in Baltimore, Maryland, USA" Samuel J. Michel, Han Wang, Shalini Selvarajah, Joseph K. Canner, Matthew Murrill, Albert Chi, David T. Efron, Eric B. Schneider. January 2016. http://www.sciencedirect.com/science/article/pii/S0020138315004180

"Is it Important to Examine Crime Trends at a Local "Micro" Level?: A Longitudinal Analysis of Street to Street Variability in Crime Trajectories" Elizabeth R. Groff, David Weisburd, Sue-Ming Yang. January 1, 2010. http://link.springer.com.proxy.its.virginia.edu/article/10.1007/s10940-009-9081-y

"It's all relative: Concentrated disadvantage within and across neighborhoods and communities, and the consequences for neighborhood crime." Alyssa Chamberlain. John Hipp. November 1, 2015. https://asu.pure.elsevier.com/en/publications/its-all-relative-concentrated-disadvantage-within-and-across-neig

"Predicting Crime: A Review of the Research." Stephen Schneider. 2002.
http://www.justice.gc.ca/eng/rp-pr/csj-sjc/jsp-sjp/rr02_7/rr02_7.pdf

"Spatial−temporal forecast research of property crime under the driven of urban traffic factors." Elizabeth R. Groff, David Weisburd, Sue-Ming Yang. March 30, 2016.
http://link.springer.com.proxy.its.virginia.edu/article/10.1007/s11042-016-3467-2

"The impact of small changes in bar closing hours on violence.The Norwegian experience from 18 cities." Ingeborg Rossow, Thor Norstrom. June 15, 2011.
http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3380552/pdf/add0107-0530.pdf