**Temporal Patterns in Intimate Partner Violence**

Kelley Stieler

1. **Project Problem and Hypothesis**

For the past six years, the National Resource Center on Domestic Violence has released a “Holiday Toolbox” to address the increased media attention around the holidays regarding the perception that incidents of domestic violence increase around the holidays. The “Holiday Toolbox” notes that to date there are no comprehensive national studies that support the very common perception that incidents of abuse increase around the holidays. In fact, a study of the number of calls to the National Domestic Violence Hotline over the holidays showed that the number of calls decreases during the holidays. However, this study is limited in that it only captures calls to the hotline and not actual reports of incidents to police. Nonetheless, the perception of the correlation between the holidays and domestic violence perpetuates.

The move towards open data sources for police departments provides an opportunity to improve efforts to understand, prevent and reduce instances of intimate partner violence (IPV). Examining police department crime incident reporting against date and time may reveal temporal patterns in reported incidents of IPV. By uncovering patterns and trends in incidents of IPV advocates and law enforcement can better align resources and education used to combat IPV.

For this project, I will track the fluctuation in number of reported incidents of IPV in numerous major cities based on the time of the report, day of the week, and date, specifically isolating holidays.

**Problem Statement:** Determine if there is an association between number of reports of IPV to police and the season or day of the week using data from Los Angeles, Oakland and San Francisco Police Departments.

**Hypothesis:**

H0: Number of reported incidents of IPV do not fluctuate based on the time of day, day of the week or time of the year.

H1: Number of reported incidents of IPV is correlated to the time of day, day of the week and time of the year.

1. **Datasets**

In California, three cities, Los Angeles, San Francisco, and Oakland, provide public access to crime incidents data. The datasets generally include, the date and time the crime occurred or was reported, the type of crime and the location of the crime. Below is a comparison of the relevant information contained in each dataset:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Police Department** | **API** | **Data Range** | **Date** | **Day of Week** | **Time** | **Description** |
| San Francisco | Yes | 01/01/2003 – 12/06/2016 | M/D/Yr | Yes | Yes | Domestic Violence |
| Battery Against Former Spouse |
| Restraining Order Violation |
| Inflict Injuree on Cohabitee |
| Los Angeles | Yes | 12/31/2010 - 08/04/2016 | M/D/Yr  Date Occurred and Reported | No | Yes | Spousal (Cohab) – Simple Assault |
| Spousal (Cohab) – Aggravated Assault |
| Violation of RO |
| Violation of TRO |
| Oakland | Yes | 2013 - 2015 | Yes | No | Yes | Domestic Violence |

A cursory review of Oakland’s crime data reveals numerous problems. There is a dataset for years 2013, 2014, and 2015. However, the 2015 data stops in mid-December, before the holidays, which are an important part of the analysis. Additionally, a large portion of the domestic violence reports in the 2013 dataset are dated in 2005-2012, with a few said to occur decades from now. As this is a large portion of the data, it is unclear if this could be remedied by omitting these from the data.

On initial review, the San Francisco and Los Angeles data appears far cleaner than Oakland’s data. A brief sample of one day of output from the San Francisco API after filtering out unnecessary variables:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Category | Descript | DayOfWeek | Date | Time |
| SECONDARY CODES | DOMESTIC VIOLENCE | Wednesday | 1/1/03 | 8:30 |
| SECONDARY CODES | DOMESTIC VIOLENCE | Wednesday | 1/1/03 | 20:10 |
| SECONDARY CODES | DOMESTIC VIOLENCE | Wednesday | 1/1/03 | 16:00 |
| SECONDARY CODES | DOMESTIC VIOLENCE | Wednesday | 1/1/03 | 21:00 |
| SECONDARY CODES | DOMESTIC VIOLENCE | Wednesday | 1/1/03 | 0:01 |

One substantial, though not insurmountable, difference between the Los Angeles and San Francisco data is that the Los Angeles data does not include the day of the week the crime occurred. Thus, I will have to create a new variable Day of the Week and figure out how to input the day of the week based on the date. Another concern is that San Francisco only lists a date and time for the crime without indicating if the date and time correspond to when the report was made or when the incident occurred. Los Angeles data includes both.

Overall, all of the data suffers from one similar deficiency. The data dictionaries do not include how domestic violence is defined or categorized. If an incident of IPV is reported as a simple assault or sexual assault, it will not be captured in the data.

Given the deficiencies in the Oakland dataset, I may include Chicago in the analysis. The City of Chicago also has an API for Crime Incident Reports. Unlike the California cities, Chicago’s data includes a variable of whether or not the crime qualified as an act of domestic violence under the State statute. Thus, it makes it easy to determine if, for example, stalking or a restraining order violation involved domestic violence. The data dictionary is far more robust than that of the other cities.

1. **Domain knowledge**

To date, no comprehensive national study exists linking specific days or times of the year with an increase in intimate partner violence. Anecdotal evidence and studies of specific populations do exist, each suggesting a relationship between certain times of the year and an increase in IPV. For example, a 2005 study examined incident reports of IPV in Idaho from 1995 to 2001.[[1]](#footnote-1) The study found a correlation between increased incidence of IPV and specific holidays and days of the week. The population studied, those in Idaho, may not be representative of the general population as this is a small State with a largely rural, rather homogenous, population. Additionally, the data examined in the study is now nearly two decades old. Nearly a decade later, in 2014, a study examined the incidents of IPV reported to the US Air Force’s Family Advocacy Program.[[2]](#footnote-2) Similarly, this study also found correlation between incidents of IPV and time of the year, including a marked increase on Super Bowl Sunday. Again, as in the Idaho study, this population studied is not representative.

Another oft-cited study came to a different conclusion, namely that reports of domestic violence do not increase over the holidays. Data provided by the National Domestic Violence Hotline, cited heavily by the Domestic Violence Resource Center, suggests that calls to their hotline decrease during the holidays. However, notably, reports to a hotline are not the same as reports to law enforcement.

1. **Project Concerns**

Overarching concern is that my model will be very basic given the limited number of features (date/time). The descriptive statistics alone may answer many of the questions. To counter this, I can break apart the date and time into numerous variables – such as time of day, day of week and by holiday. This will require a lot of pre-work before I can begin any analysis. Additionally, for the Los Angeles dataset, there will there will be the initial challenge of creating a day of the week variable using the date.

Finally, I have some questions about the type of model to use. My inclination is logistical regression model by transforming the target variable – number of incidents – into a bivariate variable (above or below mean incidents/day). I suggest this type of model because it will produce odds ratios, so I can compare the odds of increased incidents across the numerous variables.

Overall, this study has numerous limitations. The data I am analyzing only includes incidents of domestic violence reported to the police. Unreported incidents or incidents reported to hospital staff, hotlines, to shelters or counselors will go unreported. Also, I will only be examining data from cities. This is largely because more rural police departments do not have readily accessible data. Thus, a more nationwide and representative sample is not possible given the data available. Finally, as noted above, datasets do not contain detailed data dictionaries. Thus, it is unclear how crime incidents are classified. For example, a crime reported as stalking or a restraining order violation may or may not be domestic violence. Including or excluding this from my analysis will impact my results.

More generally speaking, creating a temporal framework to examine domestic violence has its own risks.[[3]](#footnote-3) In so doing, it creates the idea that domestic violence occurs in temporal patterns and that it is not about a power dynamic that is ever present. An over emphasis of any temporal patterns may take away from the broader concerns. On the other hand, if there are temporal patterns, they should not be ignored as understanding the patterns provides opportunities to implement impactful interventions.

1. **Outcomes**

I intend to have a separate set of outcomes for each city since the data ranges for each city only overlap for a couple years. Much of the output data will include descriptive statistics and graphical visualizations. For example, a listing of the mean number of incidents reported per day or season and bar charts comparing the mean number of incidents of IPV by holiday, season or day of the week. I could potentially use the census bureaus data to calculate the incidence rate. Other studies often calculate the incidence rate, which requires including the population data for the cities.

Ideally, I would like to create a logistical regression model in order to calculate odds ratios of holidays and non-holidays or weekday versus weekend, using a target result of above or below mean reported incidents per day. Additionally, it would be useful to be able to predict the number of incidents likely to be reported on a given date in the future.

1. Vazquez, S. P., Stohr, M. K., & Purkiss, M. (2005). *Intimate partner violence incidence and characteristics: Idaho NIBRS 1995 to 2001 data.* Criminal Justice Policy Review, 16, 99-114. [↑](#footnote-ref-1)
2. McCarthy, R. J., Rabenhorst, M. M., Milner, J. S., Travis, W. J., & Collins, P. S. (2014). What difference does a day make? Examining temporal variations in partner maltreatment. *Journal of Family Psychology*, 28, 421-428. [↑](#footnote-ref-2)
3. Jeltsen, M. (December 12, 2016) *Why It’s Dangerous To Claim Domestic Violence ‘Spikes’ Over The Holidays* Huffington Post. [↑](#footnote-ref-3)