# Washington, DC Violence Forecaster: A Case Study Comparing Weather and Crime

matthew d'anna

George Mason University

Spring 2014

## Introduction

This project explores the potential relationships between weather and crime. Specifically, this study seeks to answer a relatively simplistic question: what, if any, is the relationship between temperature and violent crime in Washington, DC? In an attempt to answer this question, this paper examines daily weather data correlated with daily crime counts, specifically murder and robbery, to identify and potentially predict amounts and locations of violent crime.

## Literature Review

As a social science, criminological research generally seeks to relate data to physical sciences, as a means to enhance comprehension and uncover new contextual meanings. Routine Activities Theory encourages a logical approach to understanding crime as a component of a potential victim's surrounding environment. A primary facet of this environment is the weather. As such, studying the effects of temperature on crime is a popular topic in criminology. One of the earliest studies was conducted by Anderson (1987), which found a greater prevalence of violent crime during hotter quarters of the year. Further, when controlling for social, demographic, and economic variables, differences in violent crime appear to be related to the hotness of a city. Gamble and Hess (2012) studied temperature and violent crime in Dallas, Texas from 1993-1999, noting a curvilinear relationship: violent crime increases as temperatures exceed 80º F, and then decrease at temperatures above 90º F. Field (2013) noted similar conclusions for both violent and property crime in the United Kingdom at aggregate monthly, quarterly, and annual levels. These increases were notably independent of season variations. Interestingly, no relationships between crime and levels of rainfall or sunshine were identified. While this is a small sample of available research, the theme is consistent: there is some degree of relationship between high temperatures and violent crime.

## Methodology

Answering the question of weather and crime involves a multi-faceted fusion process of six data elements: temperature, windspeed, assault, robbery, murder, and combined violent crimes. All data was

queried for approximately 5.5 years, from January 1, 2009-May 5, 2014. The National Climate Data Center (NCDC) has two weather stations servicing the greater District of Columbia metropolitan area, as noted in Figure 1. Daily averages for temperature and windspeed were calculated for both weather stations, resulting in a single figure for each variable, each day (Figures 2-4).

**Figure 1: Weather Station Map**

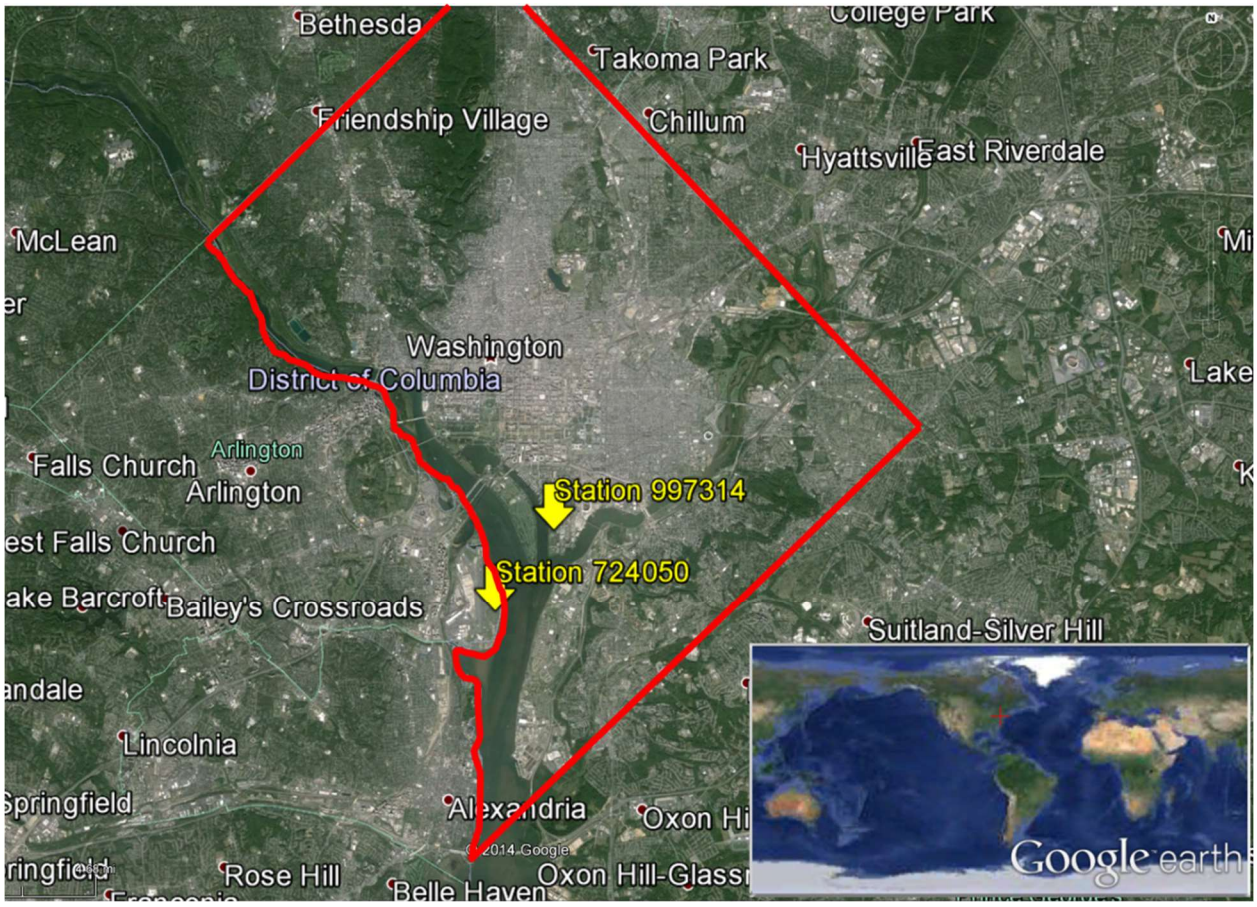



**Figure 2: Statistics for Temperature and Windspeed**

| | Temperature | Windspeed |
|---|---|---|
| N | 1946 | 1946 |
| **Measures of Central Tendency** | | |
| MEAN | 58.09 | 6.31 |
| MEDIAN | 58.03 | 5.80 |
| MODE | 41.70 | 5.25 |
| SD | 17.29 | 2.74 |
| SKEW | -0.12 | 1.01 |
| MEAN + SD | 75.38 | 9.05 |
| MEAN - SD | 40.81 | 3.57 |

**Figure 3: Washington, DC Average Daily Temperature**

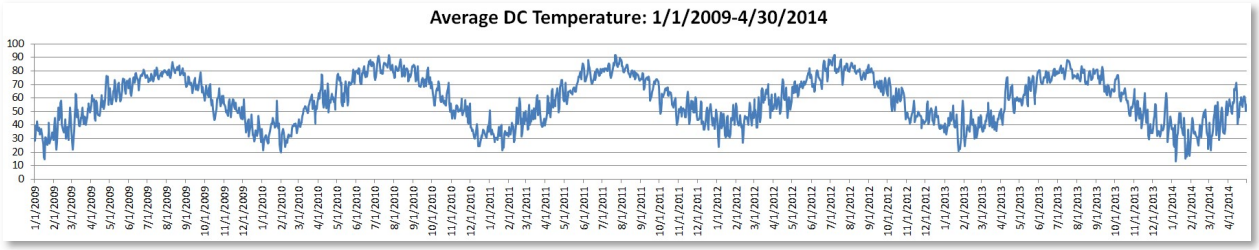
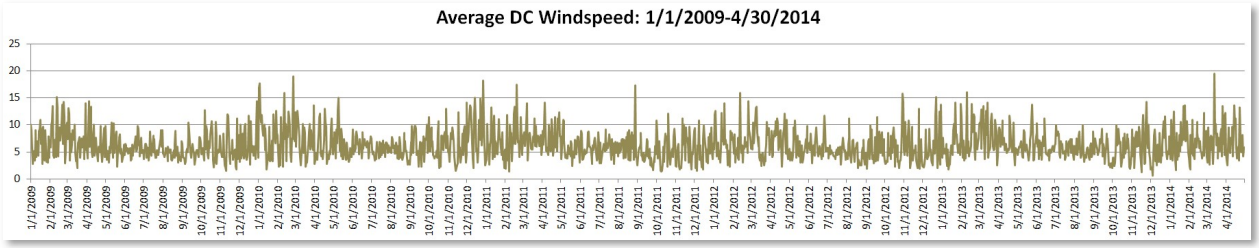



**Figure 4: Washington, DC Average Daily Windspeed**




The crimes analyzed were based on the universally accepted FBI Uniform Crime Report definitions for murder, robbery, and assault. Crime data was obtained from a RAIDS Online, a publicly available crime mapping website from BAIR Analytics. RAIDS Online does not provide comprehensive coverage of the entire country; however the jurisdictions included generally have reliable data over time. Daily counts for each crime, as well as a total violent crime count, were calculated, as described in Figures 5-9. Assaults were excluded from analysis after it was determined to be unreliably collected over time.

**Figure 5: Crime Statistics**

| | Total Violent | Murder | Robbery | Assault |
|---|---|---|---|---|
| N | 1799 | 1790 | 1799 | 1345 |
| % of total N | 92.45 | 91.98 | 92.45 | 69.12 |
| **Measures of Central Tendency** | | | | |
| MEAN | 15.83 | 0.31 | 10.43 | 6.81 |
| MEDIAN | 15.00 | 0.00 | 10.00 | 7.00 |
| MODE | 16.00 | 0.00 | 9.00 | 5.00 |
| SD | 6.38 | 0.58 | 4.35 | 3.22 |
| SKEW | 0.40 | 2.06 | 0.26 | 0.38 |
| MEAN + SD | 22.21 | 0.89 | 14.78 | 10.04 |
| MEAN - SD | 9.46 | -0.27 | 6.08 | 3.59 |
| AUTOCORRELATION | 0.52 | 0.06 | 0.38 | n/a |

**Figure 6: Washington, DC Average Daily Violent Crime**

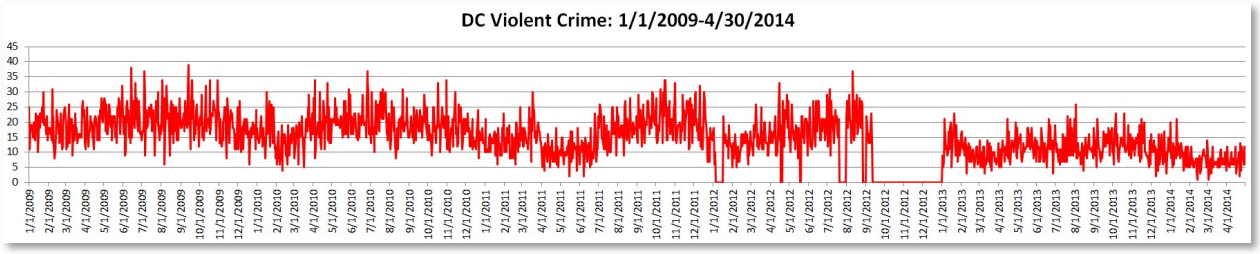



**Figure 7: Washington, DC Average Daily Murder**

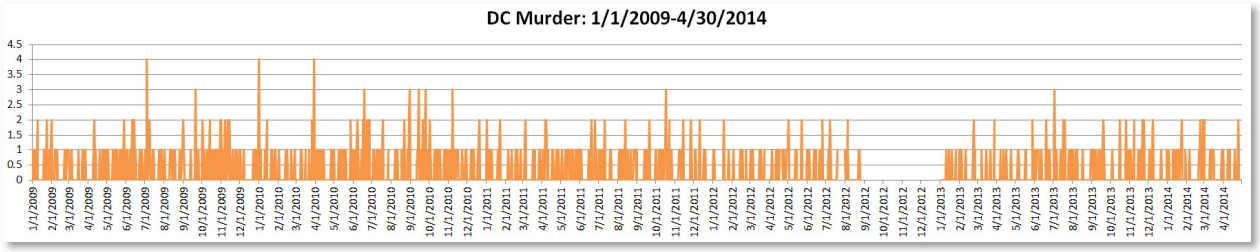



**Figure 8: Washington, DC Average Daily Robbery**

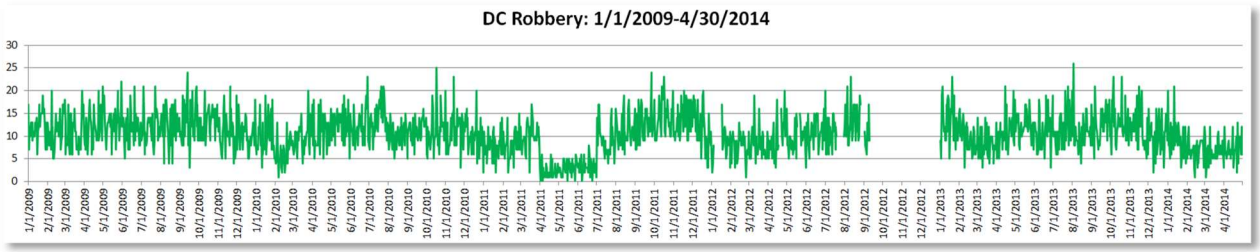
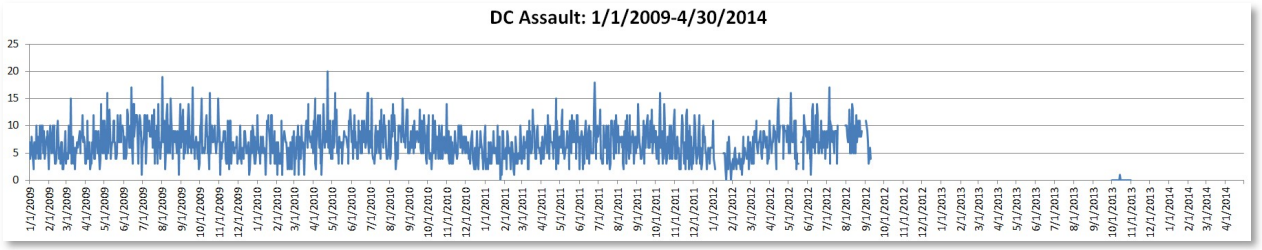



**Figure 9: Washington, DC Average Daily Assault**




Based on these data, a master table of temperature, windspeed, each violent crime, and total violent crime per day was created, as the basis for analysis. The intent was to a) create a model for determining future amounts of violent crime based on these variables, b) determine how similar a prediction is to previous dates, and c) identify common locations based on dates with similar crime counts.

## Analysis

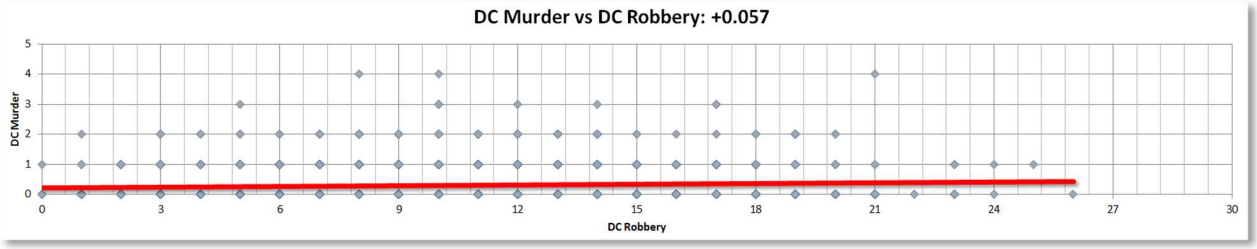
Each variable was correlated to identify any significant relationships of interest, as seen in Figures 10-20.

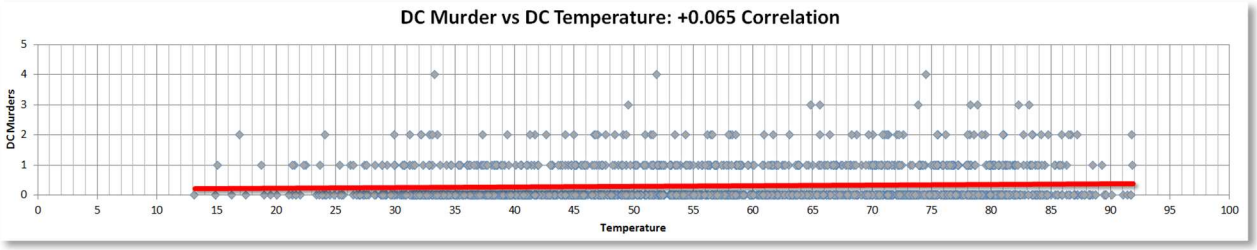
**Figure 10: Correlations**

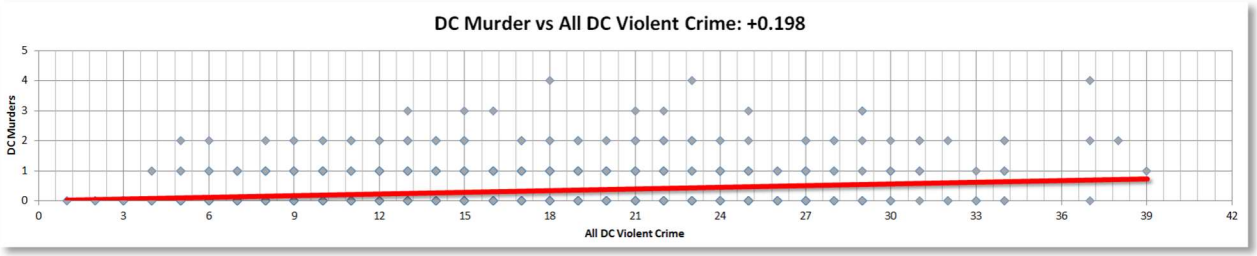
| Variables | Violent Crime | Murder | Robbery | Temperature | Windspeed |
|---|---|---|---|---|---|
| Violent Crime | *n/a* | 0.1977 | 0.7609 | 0.2873 | -0.1304 |
| Murder | 0.1977 | *n/a* | 0.0571 | 0.0648 | -0.0224 |
| Robbery | 0.7609 | 0.0571 | *n/a* | 0.1605 | -0.1245 |
| Temperature | 0.2873 | 0.0648 | 0.1605 | *n/a* | -0.2196 |
| Windspeed | -0.1304 | -0.0224 | -0.1245 | -0.2196 | *n/a* |

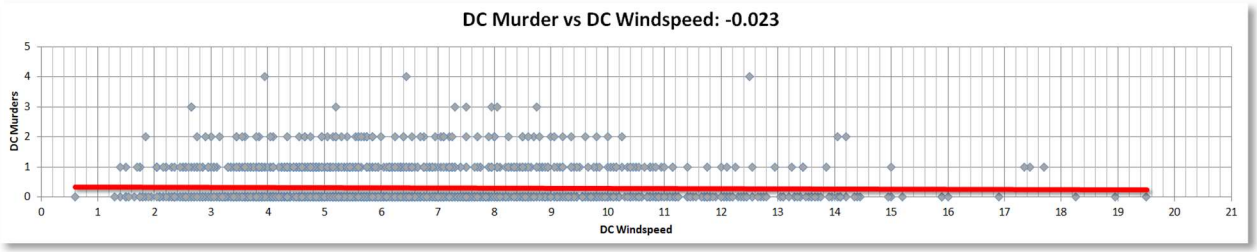
A strong positive relationship between robbery and violent crime is expected, given that the majority of violent crime in this study is robbery. Interestingly, temperature was a moderate positive relationship with all violent crime, with weaker relationships with robbery and murder, respectively. Overall, windspeed has negative correlations with all other variables; meaning as windspeed increases, crime decreases.

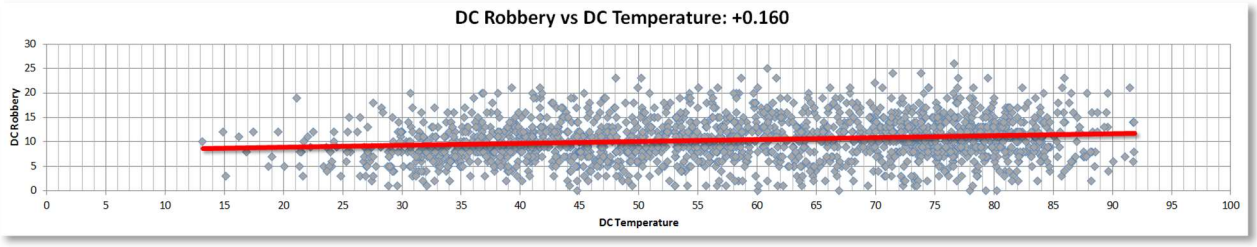
**Figures 11-20: Correlations**

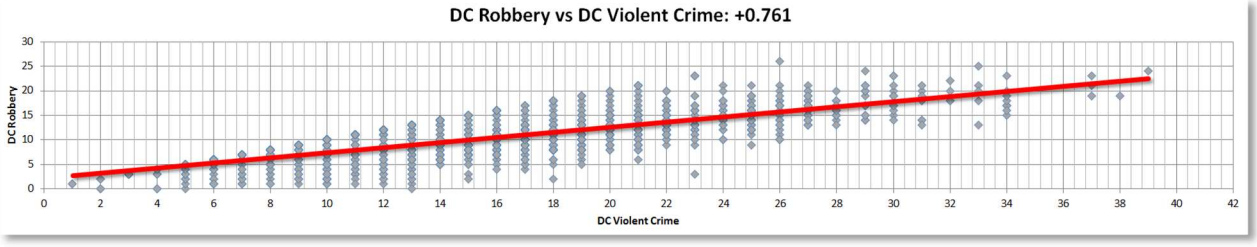

DC Murder vs DC Robbery: +0.057
DC Murder
DC Robbery

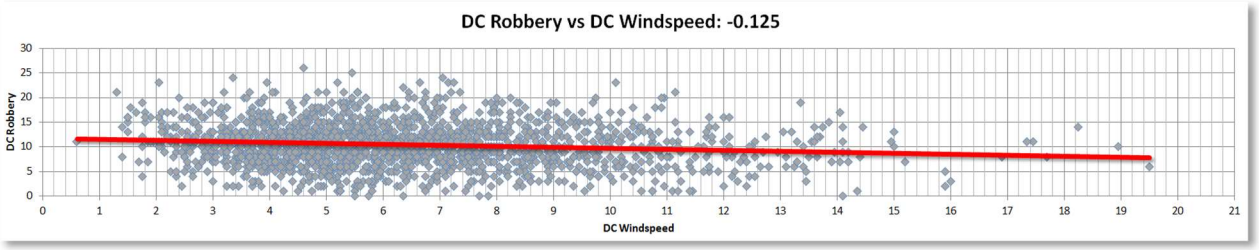

DC Murder vs DC Temperature: +0.065 Correlation
DC Murders
Temperature

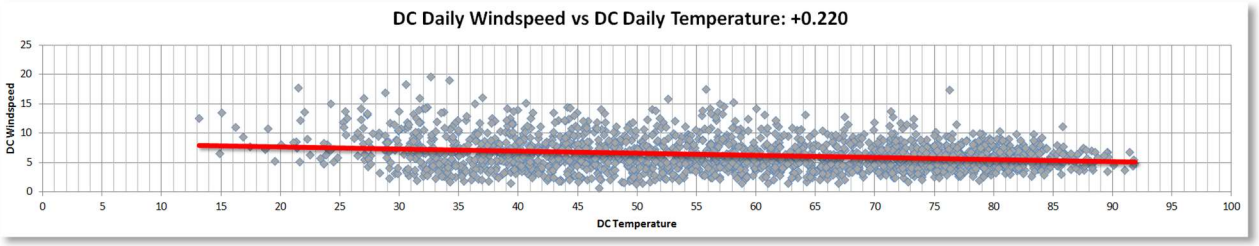

DC Murder vs All DC Violent Crime: +0.198
DC Murders
All DC Violent Crime

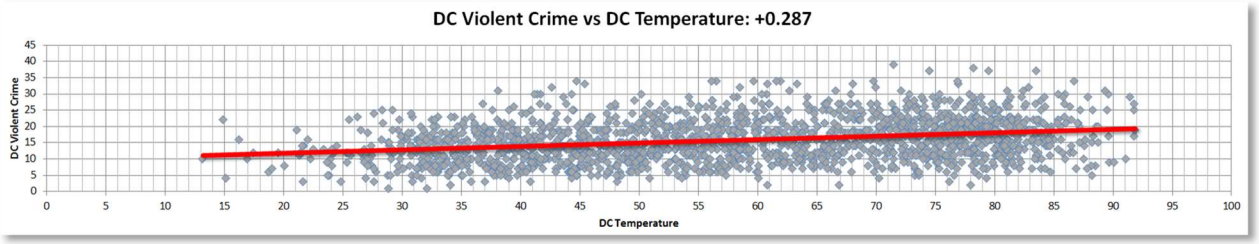

DC Murder vs DC Windspeed: -0.023
DC Murders
DC Windspeed

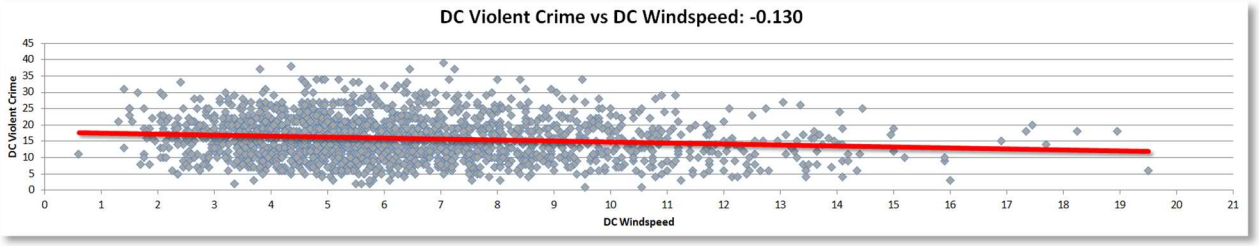

DC Robbery vs DC Temperature: +0.160
DC Robbery
DC Temperature


DC Robbery vs DC Violent Crime: +0.761
DC Robbery
DC Violent Crime




Given these relationships, a regression model was created. Violent crime was the dependent variable, and temperature, windspeed and the previous days' violent crime were independent variables. The prior day violent crime counts were deemed significant from autocorrelation calculations. This model had an $R^2$ of 0.542. The model was heavily driven by the Prior Day Crime variable; however, in models with this variable as the only independent variable, the $R^2$ was only 0.424. Thus, temperature and windspeed make meaningful contributions to this model. Further, using the residuals of the regression, high and low buffers were created for each predicted value (Figures 21-23).

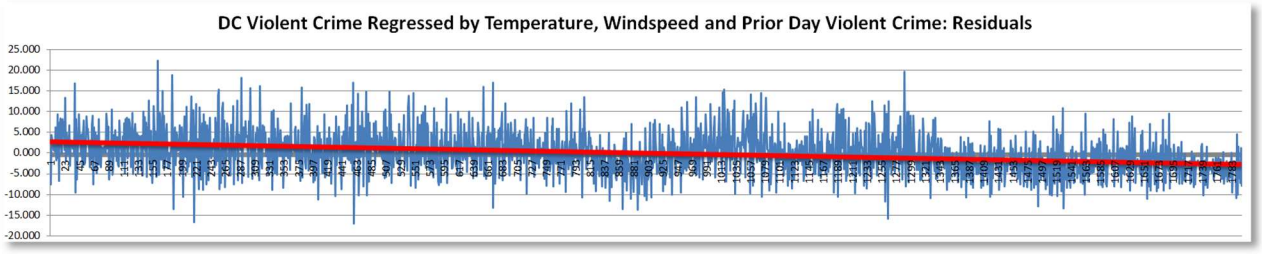
**Figures 21-23: Regression Equation and Residuals**

Coefficients[a]

| Model | | Unstandardized Coefficients | | Standardized Coefficients | t | Sig. | 95.0% Confidence Interval for B | |
|---|---|---|---|---|---|---|---|---|
| | | B | Std. Error | Beta | | | Lower Bound | Upper Bound |
| 1 | (Constant) | 5.914 | .637 | | 9.289 | .000 | 4.665 | 7.163 |
| | Temperature | .053 | .008 | .146 | 6.908 | .000 | .038 | .068 |
| | Windspeed | -.105 | .048 | -.045 | -2.208 | .027 | -.199 | -.012 |
| | PriorDayViolent | .473 | .021 | .473 | 22.835 | .000 | .432 | .514 |

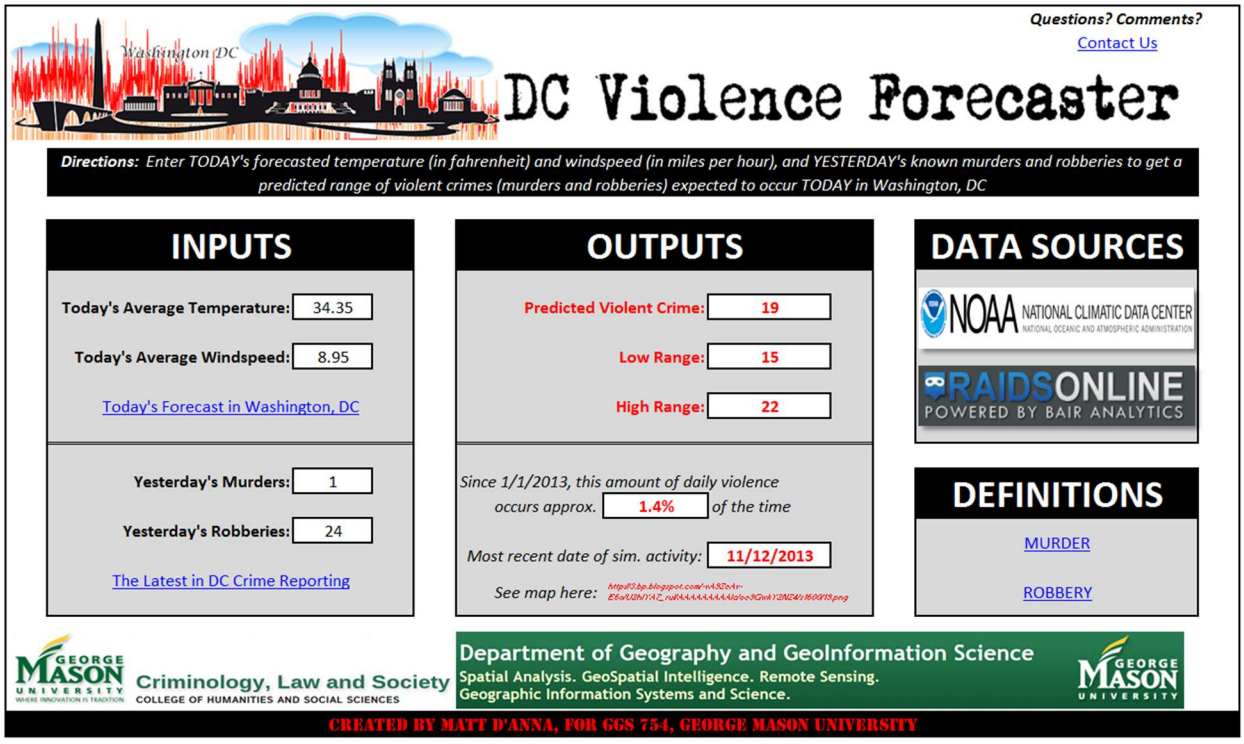
a. Dependent Variable: ViolentCrime

| Violent Crime Residuals | |
|---|---|
| MEAN | 4.23 |
| STANDARD DEVIATION | 3.29 |
| MEAN + STANDARD DEVIATION | 7.52 |
| LOW BUFFER | -3.76 |
| HIGH BUFFER | 3.76 |




Based on this equation, a user-defined input was created, as seen in Figure 24. This tools enables a user to input the day's temperature and windspeed, and yesterday's violent crime. If the user is unaware of the metrics, the hyperlinks enable them to find that information. Based on those inputs, a prediction, with high and low buffers, is made. Further, the prediction is compared to a recent, high-performing subset of records, January 1, 2013-April 30. 2014. This process identifies the percentage of cases matching the prediction, the date of the most recent match, and a corresponding map of the crime locations for that date. The data is made available through RAIDS Online, and is sanitized at the street/block level.

**Figure 24: User Input Screen for DC Violence Forecaster**




Once created, the five most days were analyzed in real-time to test the models performance. The results are seen in Figure 2 below.

**Figure 25: Real-time Crime Predictions**

| Date | Temperature | Windspeed | Prior Day Violent | Prediction | Actual | Difference |
|---|---|---|---|---|---|---|
| 5/1/2014 | 66.70 | 5.55 | 6 | 12 | 8 | 4 |
| 5/2/2014 | 64.40 | 6.90 | 8 | 12 | 8 | 4 |
| 5/3/2014 | 64.35 | 4.15 | 8 | 13 | 10 | 3 |
| 5/4/2014 | 66.45 | 7.35 | 10 | 13 | 7 | 6 |
| 5/5/2014 | 58.80 | 8.65 | 7 | 11 | 10 | 1 |

## Conclusions

A small, positive relationship exists between both temperature and windspeed with violent crime in Washington DC. However, the greatest predictor for the amount of violent crime in the city on any given day is the previous day's violent crime. A combination of these variables appears to yield the strongest predictability. Further research, incorporating additional environmental features, will likely strengthen models such as these.

## References

Anderson, C.A. (1987). Temperature and aggression: Effects on quarterly, yearly, and city rates of violent and nonviolent crime. *J Pers Social Psychology, (52)*6, 1161-1173.

Field, S. (2013). The effect of temperature on crime. *The British Journal of Criminology, (32)*3, 340-351.

Gamble, J.L., & Hess, J.J. (2012). Temperature and violent crime in Dallas, Texas: Relationships and implications of climate change. *West J Emerg Med, (13)*3, 239-246.