September 30, 2009

Insights Into Crime Trends in Washington DC

Introduction

As citizens of Washington DC you are probably concerned about your safety and the general crime statistics of your neighborhood. In this report, we will try to show some of the less known facts about criminal activities in the United States Capital.

Data

The DC Metropolitan Police Department provides details about individual crimes to the public. These include violent crimes (assault with a dangerous weapon, homicide, sexual assault) and property crimes (arson, burglary, robbery, stolen auto, theft, theft from auto). Overall, there were 127,600 crimes reported to occur between January 1, 2006 and September 9, 2009. Each reported crime has 23 attributes, including time (the date and police shift of the reported crime), location (latitude, longitude, block address, ward, police district), and crime categories (offense and method), along with other identifiers of the incident. The data is available at [http://data.octo.dc.gov/](http://data.octo.dc.gov/).

In The News

Criminals Take Vacations Too

We looked at Washington DC’s Criminal Reports using Spotfire from the beginning of 2006 until the end of August 2009. We wanted to see if there was any trend in criminal activity depending on the day of the week and roughly time of day, so we made a bar graph showing the number of criminal reports grouped by day of week. Each bar is broken down into the particular shift during which the report was made (MID – 11:00pm-07:00am; DAY – 07:00am-03:00pm; EVN – 03:00pm-11:00pm).

We found out that Washington DC’s criminals prefer to take a slight “break” two days earlier than most people. In other words, the days of Thursday and Friday have, on average, the least criminal activity compared to the rest of the week. Moreover, criminal activity starts back into full swing on Sunday and peaks during Monday when most of the population goes back to work.
We then wanted to see if there are any trends regarding criminal activity depending on the month of the year. For this purpose we loaded data starting with the year 2006 and ending in 2008 inclusive. We filtered out all the data from 2009 since it was incomplete and we did not want to add any bias in the graph for the first eight months of the year.

First, the winter holiday season does not have the highest criminal activity, as we expected. Instead, Washington DC’s wrongdoers prefer the summer months of May through July for their pursuits. But what probably strikes most just by glancing at the graph is the big dip in illegalities during the month of February. We also took into consideration that February is the month with the least number of days during the year, but even after that adjustment, the difference is obvious. We cannot seem to find any reasonable explanation for this finding, except that after a busy season, even the criminals may need a break.

**The Safest Ward May Not Be That Safe For You**

We looked at all the data again and this time we wanted to see if certain city blocks were more likely to have high criminal activity. We also grouped the city blocks by their ward number and we made the treemap visualization above. Immediately we noticed that Ward 3, the one with the least amount of crime in all DC, contains two blocks with a high number of crime reports, some of the highest in DC. Thus we wanted to see what their rank was with respect to all other city blocks.
We put together the two visualizations above, with the top bar graph showing the “Top 10 Hotspots” in Washington DC in terms of criminal activity. The highlighted elements represent the two most crime infested city blocks in Ward 3, which are ranked 2nd and 5th overall in all of Washington. Clearly, even though you might live in a peaceful neighborhood overall, danger can be just one block away.

Assaults with Deadly Weapons Fall for the 4th Year in a Row

We looked at the number of reported assaults with deadly weapons during each quarter since 2006 and we noticed that, despite the seasonality of the data, on average the number has been decreasing steadily for the last four years. We broke down the data by method used in the reported assault, and we saw that the sharpest decline comes from the “other method” category. But, reports of assaults with guns and knives are also enjoying a steep decline. Assaults using clubs declined the least, however they were not very numerous to begin with.

Critique

Our analysis was performed using Tibco Spotfire. Overall, Spotfire is a fantastic tool for data exploration. Users can pose questions about their dataset, and in most cases find the answers quickly and easily. However, we did find areas that could use improvement, both in terms of functionality and usability.

Main Critiques

We have two main critiques of Spotfire. First, filters are very powerful tools, but a couple enhancements could allow even more intuitive “slicing and dicing” of data. Especially for data with a large number of columns, a cluttered filters panel distracts from its power. Two suggestions we have for the filter panel are:

1. Users most commonly want to examine filters that affect the currently displayed data. To make these important filters easy to identify, a right-click option to “Show active filters” would be useful. Otherwise, a user must scroll through all available filters (23 in our dataset) and manually hide each unnecessary filter, or pre-arrange the filters into filter groups.

2. The filterable data attributes should mirror the attributes available within visualizations. Specifically, for date columns users can choose to display derived attributes like “Month” and
“Day of Week”, but these are not directly filterable fields. They can be added in a separate, single-column filter hierarchy (which is unintuitive, because you are creating a “hierarchy” of one column), but may be better represented as a separate filter type for date columns.

Second, Spotfire should give users granular control over missing data points in charts with multiple series. For example, we wanted to show the reduction in assault activity by method (e.g. “gun” or “knife”) and year. However, not all methods of assault (e.g. “club”) occur every year, but instead of a bar with height zero, nothing is displayed. Via the axis properties, Spotfire allows all-or-nothing options for displaying missing values, but what is really needed is a choice to show missing values for a selected set of columns.

**Minor critiques**

- It was difficult to discover how to plot aggregate data in a scatter plot. According to the documentation, this is “only available if the markers in the scatter plot have been aggregated by setting the ‘Marker By’ option to something other than (Row Number).” To facilitate ease-of-use, Spotfire should select a smart default setting for ‘Marker By’.
- Users should be able to directly use the data displayed in a Cross-Tab visualization as a new data source or data table.
- In places where correlated visualization settings come from different menus, it would be helpful to have a link between the two. For example, it would be helpful to have a link from visualization axis settings to alter sort order for a column, instead of exiting the visualization properties window to open the column properties window.

**References**