Final Project - Seattle Crimes Predictive Policing Project   
PowerView GeoSpacial Analysis



The Seattle PD has an opening for a Business Intelligence Analyst. A meeting with a recruiter discovered that they utilize a full-day interview including a hands-on analytical session (they are also know to provide 'Training day' experiences on the second interview, that is a subject for a later job prep assignment). An Internet scan discovered a [public database of police incidents](http://www.seattle.gov/police/crime/default.htm). You are asked to perform some pre-interview job preparation. You are asked to download a database and create a management report of recommendations related to one category of crime (ie, assault, car prowl) based on your original analysis. You can work alone or in groups of two, if you think you will be more productive.   
  
BTW - here is a [link to NYC crime databases](https://data.cityofnewyork.us/) - in case you want to compare crime trends.

**Steps**

1. The data has already been prepared for you and is [here](file:///\\cb-ot-srv03\CCB\Public\~mfeatherman\Common\Data\SeattlePD_PowerPivot.zip).   
   Analyze just the data for the last 5 years. The data is incomplete in 2008 so do not assume a large spike in crime has occured.   
     
   This linked data file is a zipped Excel file that has the data imported into PowerPivot (inside Excel). You probably have to be using a version of Excel 2010 or Excel 2013 that has the powerpivot add-in. You can skip ahead to #3 if you can successfully download, unzip and open the linked datafile.
2. This next portion just explains how the data got transformed and entered into first SSMS, then Excel.   
     
   The Seattle.Data.gov website was accessed and a .csv file was downloaded and the data was imported from the seattle.gov website, was trimmed and the columns renamed and imported into SSMS. Then the following query was used to add more date & time columns to improve the analysis. Notice this query below retrieves the table from SSMS adds the extra calculated columns and copies the data into a new table. This query below did the trick.

SELECT [IncidentID] ,[Offense] ,[Offense Detail] ,[Date Occurred],[District]  
,[Latitude],[Longitude] ,[Location] ,[Month]  
, DATENAME(month, [Date Occurred]) as [Month Name]   
,[Year],[YearMonth]  
, DATEPART(weekday, [Date Occurred]) as [Numeric Day of Week]  
, DATENAME(weekday, [Date Occurred]) as [Day of Week]  
, DATEPART(day, [Date Occurred]) as [Day of Month]  
, DATEPART(dayofyear, [Date Occurred]) as [Day of Year]  
, DATEPART(week, [Date Occurred]) as [Week#]  
, DATEPART(Hour, [Date Occurred]) as [Hour of day]   
INTO [Featherman\_Analytics].[dbo].[SeattlePDIncidents]  
FROM [Featherman\_Analytics].[dbo].[SPD]

Finally the completed datatable was imported into Excel using PowerPivot. Notice the location field, it contains both lattitude and logitude (a combination of the lattitude and longitude fields.) We need all three of these fields to use a map in power view. So we have to pay particular attention to these fields.   
   
If you want to take a look at the data table and its schema then connect to cb-ot-devst03.ad.wsu.edu, uid = mfstudent, pwd = BIAnalyst, database = featherman\_Analytics. Query the SeattlePDIncidents table and take a good look at the columns (in design view) and the data (you can just select the top 1000 rows).

1. **Note before you begin**:  
   You can jump into the analysis to get a feel for the dataset and a good idea of the crime category (or district) that you want to analyze, report on and improve. When you turn in the written report however, please be sure to include a cover page (adding a picture of the crime if possible is helpful), an executive summary, a table of contents, the detailed analysis (first with an overview, then sliced up using different criteria such as district, and date time dimensions). Next add maps, statistics, correlations, scattergrams to show relationships, anything you feel is necessary to investigate and examine the crime). Add all this analysis into a word processing document, and weave in your analysis and explanations. Finish with your recommendations to reduce the crime that you analyzed. Add a section that mentions the further data that you recommend be gathered and analyzed to further understand the crime.  
     
   **Start the analysis** by choosing a type of crime (assault vs. burglary, etc.) then running pivot tables and charts. After you use Powerpivot to get the data into Excel, use PowerPivot's pivot table/chart to insert several tabs each that contain charts and tables. You can also use Tableau to perform the analysis. Be sure to use local Seattle maps however. Only analyze data for the last 5 years (filter out any older data).  
     
   Use the day of week, week#, hour of day etc fields to draw many line charts and column charts. Be sure to show the data for each year. For example if you are analyzing day of week, then do not comingle all the years data into one chart. For example you would have 5 different line charts each representing one of the five prior years. You can either create 5 line charts each with one year of data or one chart with 5 data series.  
     
   Remember that each time you comingle data you are assuming (you know what happens when you assume something right?) that there is no variance in the data. For example if you show all the data in one chart for variance by day of month, then you are assuming that month or season do not introduce any variance into the data. Be careful as crime patterns are known to significantly change in the summer, around holidays, and when universities open and close. In this case you might want to provide a chart with 12 lines for the days of the month (one for each month of the year). That would be too many lines, so you can perhaps better create a concatenated field somehow and create one chart for each season, each that has 3 months of data that show days 1-30 or 31. The key takeaway is that comingling data hides information.  
     
   **Your job is to explain to the Police Chief what is going on with a particular type of crime**. Are there predictable patterns in the data? Are there spikes of activity based on day of week, day or month, month of year etc? Use the yearmonth column to draw charts that span several years (you can also concentrate of the last 4 years of data). Use a different tab to examine each different trends perhaps comparing day of month for each month of the year or day of month (for example March as compared to the prior March). Slice the data  different ways until you understand the data. Organize your analysis grouping similar analysis on each tab or powerview. Remember you will be submitting your excel file and an accompanying MS-Word document.  
     
   Count up the ID's (not sum) to get an idea of what incidents are common, and what district has a lot of incidents. You can right-click the Type field in the pivot table fields list and filter the number of values shown (for example only show the incident types that occurred  > 300 times, or just the top 10 incidents). Insert a slicer to the spreadsheet and examine the incidents by month, and year and other categories.
2. After you have a good feel for the trends of the data, then insert different powerview sheets. Use different ***maps*** to examine the data, looking at geographic districts in Seattle.
3. This is an assignment to let you practice importing data, and Powerview (however your professor will often provide the data more easily for you if you prefer). Read about the mapping features and other features of powerview using the book chapters listed in the PowerView Module. Use the mapping feature to take a closer look at the crime.
4. **This is a powerpivot & powerview (or tableau assignment)**, so please experiment with and create a workbook with several powerview reports. Please use maps and other features to really push your analysis. Try to use scattergrams over time or PowerMap as well.
5. ***Bring in 1 more data source to explain your analysis and recommendations. Also mention your ideas for the additional data that could be used to further your analysis of the crime.***
6. ***Your job is to:  
     
   a) analyze one category of the crime incidents (such as car-related incidents over time) or   
   b) all the incidents for one district over time.   
     
   Provide a report that can be given to the Police Chief that provides your recommendations.***
7. Your project will primarily be graded from your word processing document. If you have any animations or other analysis that can only be viewed in the Excel file Upload your spreadsheet with the power views reports and crosstabs, charts. etc (whatever Excel feature you need for your analysis and write-up). Add your word processing document that is addressed to the police chief. Better grades are given for more professional and insightful work.
8. Here are some design tips  
   a) crop the charts that you import  
   b) rescale the X and Y axis on charts to remove white space.  
   c)

The best projects will actually be submitted to Seargent Verhover at the Seattle Police Department so please turn in professional work. thank you.

Here are some example pivot charts to give you some ideas.



