Examining Spatio-temporal Crime Patterns and High Crime Areas in Raleigh, North Carolina
Emily Antonezyk, University of Pennsylvania, Spring 2014

VIS-STAMP outputs four data visualization tools: the Map Matrix, Self-Organizing Map (SOM), Parallel Coordinate Plot (PCP) and Spatial Temporal Matrix (S-T Matrix).

- Map Matrix: gives a visual depiction of clusters by areal designation over time, here it shows changes in crime clustering all over Raleigh, but one code - 27610 (blue cluster at bottom right in each map) - is consistently high crime over the 8 year period.
- SOM: shows the relationship clusters by color, with larger color circles indicating larger clusters. The background coloring in the SOM is also significant, while and lighter backgrounds appear more similar between adjacent clusters while dark grey or black backgrounds indicates more polarized clusters. This SOM reveals that blue clusters are relatively large and mostly set apart from adjacent neighbors.
- PCP: shows the nested means of each crime type for each cluster, in this case, the blue strands are neighbors.
- S-T Matrix: impacts each zip code and its accompanying clusters, enabling users to see how individual zip codes have changed over time (high crime to low, steady increase/decrease, stall). While some change / remain the same over time, the blue cluster in the 27610 zip code is still of interest.

Running analyses in VIS-STAMP helped to inform subsequent regressive analyses.

To further explore the crime data, SAS statistical software was used to run eight separate analyses for each crime type at the block group level for all of the block groups in Raleigh. Each analysis began using US variables, which were whittled down to seven or fewer using stepwise regressive tactics. Selecting significant predictors through stepwise regressions enabled a leaner analysis through OLS regression, the results of which are presented in Tables 1 & 2. Higher r-squared values (on a scale of 0-1, higher values indicating how much of the variation within the dependent variable, or crime type, is explained by the model) share models that are better explained by the predictive variables; in the batch of models shown to the above, the variables are best associated with the instance of simple assaults, and least likely to predict the instance of mental health commitments.

The 27610 zip code, consisting of 33 block groups, was identified as consistently high crime in the VIS-STAMP analyses. While there are not enough records (6 blocks groups) to conduct a more intensive stepwise / OLS regressive analysis, Pearson Correlation Matrices for each crime type give indication of potential predictive variables in each case. It is acknowledged that correlation does not indicate causation, but it is nonetheless interesting to see similar patterns emerge in the two analyses; in particular, that predictor variables seem to associate more strongly and consistently with more violent crimes: aggravated & simple assault, burglary, and associated consistently with higher crime crime areas across crime types.

Maps and charts give context to the study area. The first map shows crime density of the eight major crime types from the regressive analyses across Raleigh. Parcel maps in the center show the layout of the city in terms of current land use, and a close-up of the 27610 zip code. The final table shows statistics for the city of Raleigh gathered from the US Censuses as compared with statistics from the 27610 zip code.

To add: Adjust r squared values (for each US census model).