U of M Data Visualization and Analytics Bootcamp

Project 1

Team: Andrew Bartczak, Jon Simpson, Terrence Cummings

**Project description**

**Hypothesis:** Homebuyers are strongly influenced by crime statistics (negatively) and school ratings (positively) with respect to the amount they will pay for otherwise similar homes.

Identify the impacts of crime and school quality on home values in Minneapolis neighborhoods. Normalize for other factors with clear relationship to value such as square footage, number of baths/beds, etc. to isolate the impact of crime and schools.

**General process:**

1. Collect home value data by neighborhood. (**Terrence**)
   1. Segment homes by size and other variables correlated with home value.
   2. Plot gmaps heatmap to visualize home value variability by neighborhood.
2. Collect crime statistics by neighborhood. (**Jon**)
   1. Segment into categories of crime (e.g. violent vs. non-violent).
3. Collect school rating or other statistics correlated with school performance by neighborhood. (**Andrew**)
4. For each neighborhood we will then have:
   1. Profile of the average neighborhood home (i.e. size, bed/baths)
   2. Profile of the neighborhood (nearby) school(s).
   3. Profile of the neighborhood crime.
   4. Home values in the neighborhood.
5. Using multivariate regression we can then identify how much of the variance in home value is attributable to neighborhood crime and school quality (as opposed to simply the size or quality of the house itself).
6. Come up with a compelling way to visualize within a neighborhood the relationship between home value, crime, and schools. Plots, charts, regressions, maps.
7. Create PowerPoint.

**Data process:**

1. Get data
2. Clean data
3. Plots and observations
4. Save as csv to avoid having to rerun API calls repeatedly

**Potential weaknesses:**

1. Missing other variables which may have strong influence on home value (nearby shopping, recreation, etc.)
2. Importance of school and crime varies by demographic, i.e. young vs. old, family vs single/couple.

**Other notes:**

1. Google maps delineates neighborhoods so should be possible to map by neighborhoods.
2. Open Minnesota “X, Y Coordinates” in their data are not lat/long and the conversion to lat/long is complicated. Use Google Maps API to get lat/long when needed.