**Final Assignment - "Battle of the Neighborhoods"**

***Introduction/Business Problem***

In some of the previous labs and assignments, we had the opportunity to work with the location data of New York and Toronto to make clusters based on venue categories in neighborhoods. Also, in prior sections of the course, we were required to work with and analyze crime data of Chicago and San Francisco. Combining these two ideas together and applying it to my home city of Atlanta, Georgia, United States, I am going to attempt to determine whether we can find any correlation between the alcohol/weapon venue categories in a neighborhood and the number of crimes or types of crimes in the neighborhood. I will also perform exploratory data analysis as I work through the project and ideas come to mind. This project would be of particular interest to anyone who is curious about the relation between alcohol/weapon venues-types of a location and crime of that location. The police department, venue owners, or social workers may also be interested in the results of this study as it may provide insights relative to their fields. It will be interesting to have an "unbiased" look; for example, when we are in a particular location we automatically make assumptions based on the appearance of the buildings, infrastructure, people, etc. However, this study will not so much be looking at the perceived quality of the venues, but on the pattern of venue occurrence. I may also look in to venue ratings once clusters are made to explore whether a lower average rating of venues has any correlation to crime.

Previous studies have shown that higher density of bars/liquor stores have higher rates of violent crime, so I will investigate whether this holds true for Atlanta. <http://resources.prev.org/documents/alcoholviolencegruenewald.pdf>

***Data Description/Examples***

***Atlanta Data*** - I will use Wikipedia to get data concerning Atlanta neighborhoods and NPUs(Neighborhood Planning Unit, geographically based). Conveniently, the crime data I will use includes the latitude and longitude of the crime in addition to the neighborhood name and NPU, so I will take the mean lat/long for crimes in a given neighborhood to determine its coordinates(to be used with Foursquare API). Finally, I will use the open data csv files from the Atlanta Police Department to acquire crime data; for simplicity, we will only use the crime data for 2018.

<https://en.wikipedia.org/wiki/Neighborhood_planning_unit>

<http://opendata.atlantapd.org/Crimedata/Default.aspx>