**Disparities in Access to Facilities that Promote Health and Well-Being in Washington, DC Neighborhoods**

**Data**

To perform the analysis described above, I will make use of the Foursquare location database supplemented with neighborhood-level location information obtained from the on-line real estate service Zillow. Foursquare contains not only precise location data (street address and latitude/longitude) but also descriptions of each entity’s basic function/purpose, which is defined as one of several hundred “Categories” or sub-categories. Table 1 below provides examples of the categories that I will examine for this project. The presence of several types of entities within a neighborhood is likely to promote better health, convenience, satisfaction, and general well-being. At the same time, the presence of other types of entities within a neighborhood may work to the detriment of local residents, due to possible exposure to pollutants, unsafe conditions, excessive noise, or predatory economic activity. Finally, the presence of some entities may have positive effects in some cases, but negative effects in others or have impacts that are difficult to predict in isolation.

|  |  |  |
| --- | --- | --- |
| **Table 1**  **Illustrative Foursquare Entity Categories of Interest:**  **Presumed Impact on Human Health and Well-Being** | | |
| ***Positive*** | ***Neutral or Ambiguous*** | ***Negative*** |
| Supermarket | Convenience Store | Bar |
| Pharmacy/Drug Store | Military Base | Waste Facility |
| Bank | Warehouse | Check Cashing Facility |
| Spiritual Center (place of worship) | Gas Station | Prison |
| Athletic and Sports Facility | Recycling Facility |  |

One issue that must be resolved in conducting a project based on entity location is that of scale. Washington, DC is a medium-sized city bounded by somewhat arbitrarily drawn lines, which date from the city’s designation as the U.S. national capitol, on three sides and the Potomac River on the remaining (south) side. Washington has approximately 20 zip codes, many of which are of very irregular size and shape. Accordingly, an analysis performed on the basis of zip code might not yield sufficient granularity and consistency to support an analysis of the type of interest here. Rather, an analysis conducted at the community level, according to widely recognized neighborhoods found throughout the city, would be more likely to yield valid results.

The identity of specific neighborhoods is not a field contained within Foursquare’s data set. Fortunately, I have been able to obtain location data with this level of granularity from a different source. The real estate services firm Zillow has compiled data on neighborhoods in major U.S. cities and made them available for public use. I have located and downloaded this information, containing the latitude, longitude, and boundary locations of Washington, DC’s neighborhoods. The data may be viewed and downloaded at the following link: <https://public.opendatasoft.com/explore/dataset/zillow-neighborhoods/export/?q=DC&refine.state=DC>. According to the designations used in Zillow’s data set, the District of Columbia has 137 identified neighborhoods, yielding a level of granularity that should be sufficient to meaningfully analyze information on this city of approximately 700,000 residents.