Big Apple, Bigger Impact: Social Determinants and Hospital Rankings in New York City

Lawrence Chan | GIS 101 | Spring 2019

Tufts University | School of Arts and Sciences

The Factors

The social determinants of health are strong indicators of health outcomes of patients, and the paramount importance of health points to the importance of these social determinants as crucial to understanding differences in populations (Wilkinson, 2003). The significance of these health factors are shown in Figure 1 (Remington, 2015). But, do these determinants also translate to differences in hospital ratings? Since healthcare data are hard to parse by everyday patients, offices such as the U.S. Centers for Medicare & Medicaid Services (CMS) through the Department of Health and Human Services provide easy tools for patients to compare hospitals.

Health Factor	Percentage Makeup of Health
Health Behaviors	30%
Clinical Care	20%
Social & Economic Factors	40%
Physical Environment	10%

Figure 1. Social Determinants of Health

Figure 2. Hospital Compare

Called "Hospital Compare", this tool available at medicare.gov (Figure 2) shows patients a star-rating of hospitals, on a scale of 1 to 5. According to CMS, these ratings serve as a "summary of quality measures" that "reflect common conditions that hospitals treat" (DDHS, 2019). As the largest city in the United States, New York City and its healthcare system serve a tremendous patient population, so an analysis concerning this geographic region can provide useful insight into millions of lives. Due to the large percentage makeup of Social & Economic factors, these were the points of interest in this study. This research aims to show the difference in patient populations hospitals in the region serve and how these demographic factors affect the health outcome and thus rating of hospitals.

Data Selection

Data	Source	Type	Notes
Hospitals	NYU, CUNY	Raster	Location data
Census Blocks	U.S. Census Bureau	Vector	GEOID
Education Level	U.S. Census Bureau	Table	Count of Highest degree attainment of High School, Bachelors, Masters, Professional, and Doctorate
Median Household Income	U.S. Census Bureau	Table	Count of households who received Food Stamps / SNAP in the last 12 months
Race	U.S. Census Bureau	Table	Count of: Total, White alone, Black or African American alone, American Indian and Alaska Native alone, Asian alone, Native Hawaiian and Other Pacific Islander alone
Hospital Rating	U.S. Department of Health and Human Services	Manual	Star rating of hospitals

Sources

Remington, P. L., Catlin, B. B., & Gennuso, K. P. (2015). The County Health Rankings: Rationale and methods. Population Health Metrics,13(1). doi:10.1186/s12963-015-0044-2

U.S. Centers for Medicare & Medicaid Services. (n.d.). Hospital Compare. Retrieved May 7, 2019, from https://www.medicare.gov/hospitalcompare/search.html

Wilkinson, R. G., & Marmot, M. G. (2003). Social determinants of health the solid facts. Copenhagen: World Health Organization, Regional Office for Europe.

Methods

Data Preparation. Data tables with demographic information were downloaded from their respective sources and were edited in Microsoft Excel 2016 to change GEO.id2 column contents to text format.
Education Exception. Due to the format of the data and the desired "people who at a minimum earned a regular high school diploma" (HS), a new column was generated (BS+), evaluating to the sum of "Regular high school diploma", "Bachelor's degree", "Master's degree", "Professional school degree" and "Doctorate degree". Likewise, another column was generated to represent the "people at a minimum earned a bachelor's degree". Other than these two new column additions and GEOid conversion, no other data manipulation in Excel was conducted to contribute to the final analysis.
Spatial Queries. Each dataset was imported independently into new dataframes in ArcGIS Desktop 10.6.1. Hospital and Census Block were then imported into each dataframe. A join was conducted on the Census Block layer and on the corresponding table on the GEOid and GEO.id2. With the data now imported, spatial data was gathered. The first hospital was selected with a Select by Attribute with the query "OBJECTID = 1". A Select by Location was then conducted from on the Census Block from the select feature in the Hospital layer, with the method being "are within a distance of the source layer feature" of 1 mile from the selected hospital. Next, the attribute table of the Census Block was opened and Statistics was used to group all the blocks together near the selected hospital. See Figure 3 for a summary of data calculations. This data gathering was conducted for every hospital in the dataset.

Factor	Extracted Values
Food Stamp/SNAP Recipient	SUM of total, SUM of Recipient counts
Education Level	SUM of total, SUM of HS, SUM of BS+
Median Annual Household Income	MEAN of Median Income
Race	SUM of total, SUM of White alone, SUM of Black or African American alone, SUM of American Indian and Alaska Native alone, SUM of Asian alone, SUM of Native Hawaiian and Other Pacific Islander alone

Figure 3. Data Collection

Data transformation. Copying of SUM was used to account for differences in population density in areas. In Microsoft Excel 2016, SUMs of counts were divided by the SUMs of totals, to derive the true percentage in the geographic region. Eg. (SUM of HS was divided by SUM of total to derive the percentage of people in the region who attained at least a high school diploma). After these rates were generated, hospital ratings needed to be joined with their respective ratings. Medicare.gov Hospital Compare search was used--each hospital was queried and their rating was entered as an integer between 1 and 5 inclusive.
Data Analysis. Next, an analysis needed to be conducted to test the strength of association of demographic factors and hospital rating. Data was imported into STATA 15 and a pairwise correlation was conducted between the rating and each demographic indicator measured. Figure 4 details STATA commands entered.

```
pwcorr stars  race_w_p race_b_p race_ai_p race_a_p race_n_p, obs sig star(5)
pwcorr stars  hs_p bs_p, obs sig star(5)
pwcorr stars  hs_p bs_p, obs sig star(5)
```

Figure 4. STATA 15 Commands

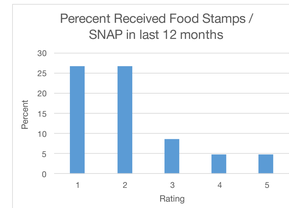
Visual Analysis. Data stored inside Microsoft Excel 2016 was used to generate graphics with the built-in graphics functions. The bar chart option, pie chart option, and scatter plot (with linear regression trendlines) were added.

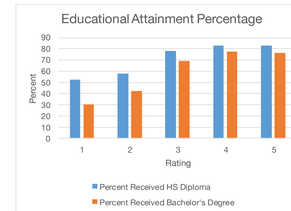
Data Flow

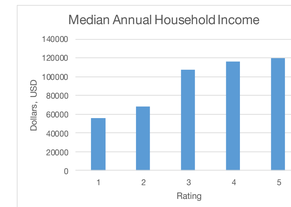
US Census Bureau → ArcGIS → Excel → STATA

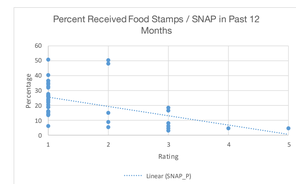
Results

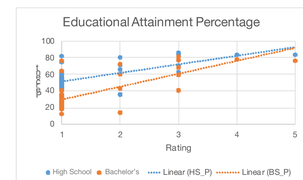
According to the graphs derived from the data, there are some general trends about the data. The figures below show that food stamp usage is correlated with lower ratings. Shows that the higher the rate of education in an area, the higher the rating. Shows that the higher income the population around the hospital is, the higher the rating. Finally, figures show racial trends from low ranking to high ranking and how racial diversity drops with higher-ranked hospitals.
These conclusions are not only supported by the visualizations but are also statistically significant. It was found that there is a strong negative correlation between percentage of recipients of Food Stamps / SNAP and rating (r_s = -0.5314, p = 0.0004). It was found that there is a strong positive correlation between percentage of high school graduates and rating (r_s = 0.5854, p = 0.0001). It was found that there is a strong positive correlation between percentage of college graduates and rating (r_s = 0.5821, p = 0.0001). It was found that there is a strong positive correlation between income and rating (r_s = 0.5686, p = 0.0001). It was found that there is a strong positive correlation between percentage of White population and rating (r_s = 0.5253, p = 0.0005). It was found that there is a strong negative correlation between percentage of Black / African American population and rating (r_s = -0.4119, p = 0.0083). Finally, it was found that there is a strong negative correlation between percentage of "other" race population and rating (r_s = -0.3993, p = 0.0107). Other racial properties were statistically insignificant, without hitting a 95% CI and the 0.313 Spearman's coefficient for n = 40 (the size of the dataset). In all, it was shown that multiple demographic factors of the population are statistically significantly tied results of the rating system.

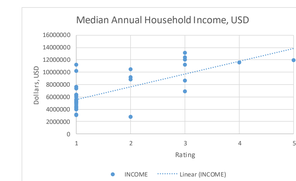


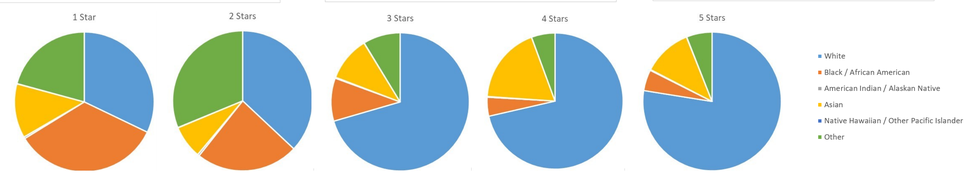












Demographic	Pearson's Coefficient	p value
Food Stamp / SNAP Usage*	-0.5314	0.0004
Education Attainment		
At Least High School*	0.5824	0.0001
At Least Bachelor's*	0.5854	0.0001
Income*	0.5686	0.0001
Race		
White*	0.5253	0.0005
Black / African American*	-0.4119	0.0083
American Indian / Alaskan Native	-0.3101	0.0515
Asian	0.1742	0.2823
Native Hawaiian / Other Pacific Islander	0.1217	0.4543
Other*	-0.3993	0.0107

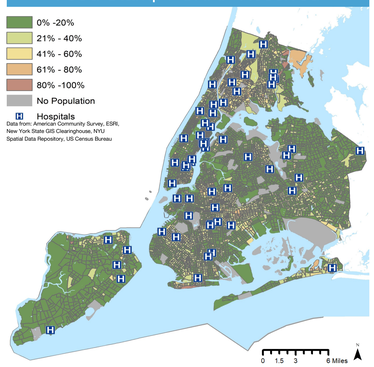
Figure 5. Summary Table

* p value ≤ 0.05

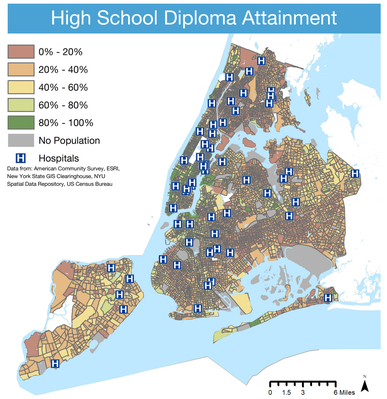
Conclusion

The results of this study (Figure 5) find that there are certain demographic factors of populations surrounding hospitals that can be linked to an association to the CMS rating system. These include food stamp usage (-0.5314), high school attainment (0.5824), bachelor's degree attainment (0.5854), income (0.5686), percent White population (0.5253), percent Black and African American population (-0.4119), and percent "other" population (-0.3993). These findings support the fact that traditionally known demographic factors are linked to health outcomes through the hospitals patients go to. These findings add to the large body of literature that assert these associations and strengthen the argument that social factors must be acknowledged with attempts to understand health outcomes. Unfortunate yet true, this also affirms the fact that many factors that determines one's health outcome are derived from things people can do little to nothing about. In all, there is an association with demographic factors and health ratings of hospitals.
These data should be used by health officials to understand the population makeup of hospitals in New York City. Decisions around budgeting, construction, and administration should be conducted with demographic factors in mind, because optimal health of the population is the ultimate goal. Further work should include hospitals from different areas around the United States and a better understanding of a hospital's service region through patient records. With these possible continuations and more, it is my hope that officials in health departments around the country take social determinants of health more heavily into consideration as these factors' association with health outcomes increases.

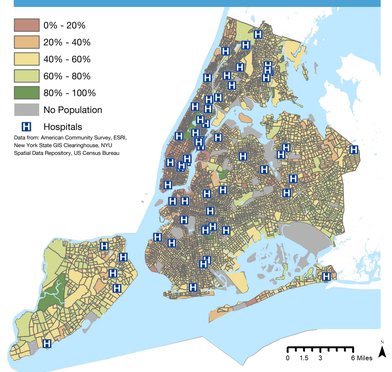
Food Stamp / SNAP Received



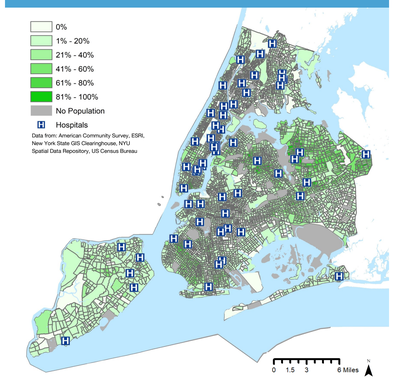
High School Diploma Attainment



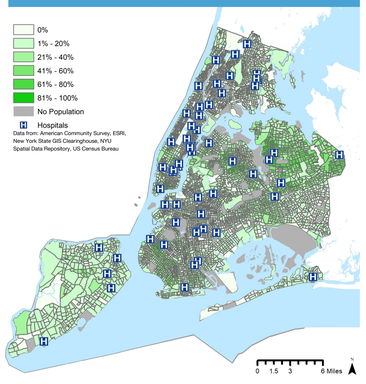
Bachelor's Degree Attainment



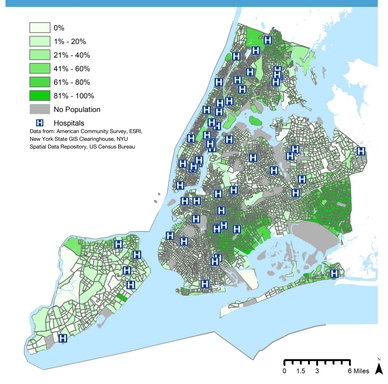
Race: American Indian / Alaskan Native



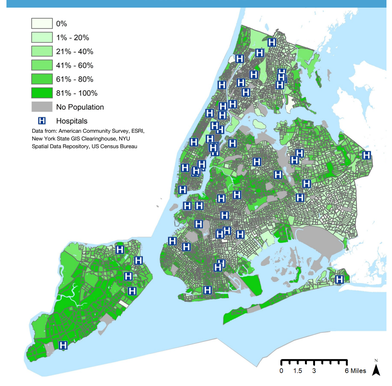
Race: Asian



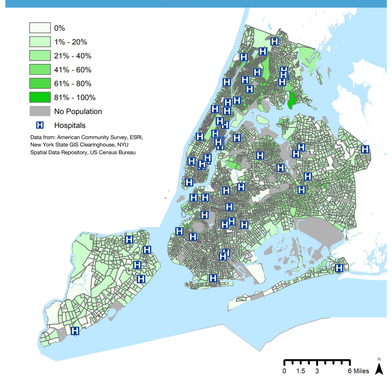
Race: Black / African American



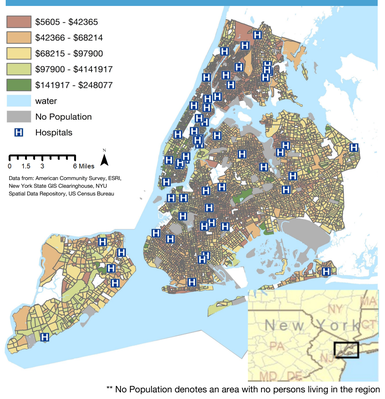
Race: White



Race: Other



Median Household Annual Income, USD



** No Population denotes an area with no persons living in the region