Hospital Expansion into Toronto

By Frederick R.

Date:5/26/2020

**Introduction/Business Problem:**

Using the Toronto datasets, the author set out to see if his employer (an United States HealthCare company) would benefit in expanding into Canada, specifically the Toronto region. The company had two questions it wanted answered: (1) could the expansion be profitable based on need and (2) were certain neighborhoods in more need of medical facilities. The concept of need is one that is present in many U.S. jurisdictions where medical facilities are based on population density, this is often called a *Certificate of Need.*

**Data:**

Data was collected via FourSquare’s API as well as geographic information from Geocoder/JSON and data pulls from the web. The web data pulls were conducted as part of the Week 3 Module for the Capstone project. Additionally, the concept of a *Certificate of Need* (*CoN*) was used in the analysis, while not normally used for business decisions, it was considered a valuable metric for decision-making. The certificate of need value does not have a set numeric value, so for the purposes of this analysis if was considered 0.75, or rather a ratio of 1:3.

A close up of a map

Description automatically generated

Figure 1: Concentration map of medical centers in Toronto

The above map shows a concentration of hospitals in the downtown and central areas of Toronto, the neighborhoods indicated in red all have a value over 0.75 as well as a high concentration of hospitals/medical facilities. There were no areas indicated that had a high concentration of hospitals with a metric under 0.75; the number associated with *CoN* as discussed previously. The neighborhoods in blue have a lesser degree of hospital concentration but still had a major presence. Overall, there will over 3300 different medical facilities identified in the area.

**Methodology:**

Upon data collection and processing two analyses were done. The first was a logistic regression to determine if the over-all expansion into the region would be potentially profitable. The second analysis used K-means clustering to see if there were certain Burroughs or neighborhoods that were more in need of additional medical facilities or hospitals.

**Results:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Algorithm** | **Jaccard** | **F1-score** | **LogLoss** |
| **Logistic Regression** | 0.54 | 0.56 | 0.57 |
| **KNN** | 0.67 | 0.63 |  |

Figure 2:- Summary of Results.

When the analysis for the over-all expansion was run a result of 0.55 was the overall average which on a scale of -1 to 1, indicates a positive association, but not a particularly strong one. This test was set up by dividing all the neighborhoods into regions and then taking the physically groupings and comparing hospital concentration.

While the second analysis was dependent of the first for business purposes, the analysis was run regardless. The K-means results were more promising then the logistic regression and observed that an even stronger concentration (r~0.65) was in fact present. Thus, the regions used in the overall analysis have a moderately strong association.

**Discussion:**

The data was presented to management at the hospital chain who would make the decisions reached in the conclusion. A common trend in this report is geographic concentration, which is important for the warranting of a new hospital facility. Should a hospital be made to close to another, they will take away business from the other and if the population is too small the hospital will not be profitable as well. Both are factors in the *CoN* that was mentioned previously. The author admits that the studies results were predictable. Upon graphing the concentrations, an apparent pattern emerged, additionally, the *CoN* was translated into a numeric value, which inherently creates room for interpretation.

**Conclusion:**

Ultimately, the expansion was not recommended. The heavier populated areas are already the most saturated with hospitals and medical facilities. Should an expansion occur it should be in the suburban areas of Toronto. Before expansion, additionally factors should be brought in for analysis such as population density and disease rates.