What neighborhood is the most 'Minnesota Nice'?

Michael O. McAnally
14 July 14, 2019

1. Introduction
    a. Background
    'Minnesota Nice' is a characteristic transcribed to people from the state of Minnesota in the Midwest of the United States. While it is hard to quantify, it is similar to pornography, where 'I know it when I see it'.

    The State Board of Tourism has collaborated with the cities of Minneapolis and St. Paul to promote tourism and relocation. They believe that publishing 'top ten' lists of most 'Minnesota Nice' neighborhoods in the Minneapolis and St. Paul (referred to as the Twin Cities) region will be highly successful.

    The McAnally Consulting Group (MCG) has been contracted by the Minnesota State Board of Tourism to perform a statistical study that creates top ten lists of 'Minnesota Nice' neighborhoods for the entire Twin Cities as well as neighborhood types.

2. Data Acquisition and Cleaning
    a. Data was collected across multiple sources. Data for this study was collected from:
        i. The Twin Cities have been releasing limited publicly available data derived from the 2010 Census and 2013 American Community Survey on mncompass.org.
        ii. Geolocation data from the Google API
        iii. In-depth venue data for parks, churches, schools, and broad general data across all venue types
        iv. Walking scores for locations in the Twin Cities from the Walkscore API
    b. Data was accessed from the Google Geocoding, Walkscore, and Foursquare APIs
    c. When data was not available for particular factors of interest (see: Median Household Income, Median Home Value) a median value of the factor for all the neighborhoods was imputed for the 'NaN'.
3. Exploratory Data Analysis
    a. Summary statistics were calculated and checked for reasonability across the factors of interest in determining 'Minnesota Nice'.
    b. Simple maps were created to visualize location of the 102 neighborhoods in the Twin Cities.

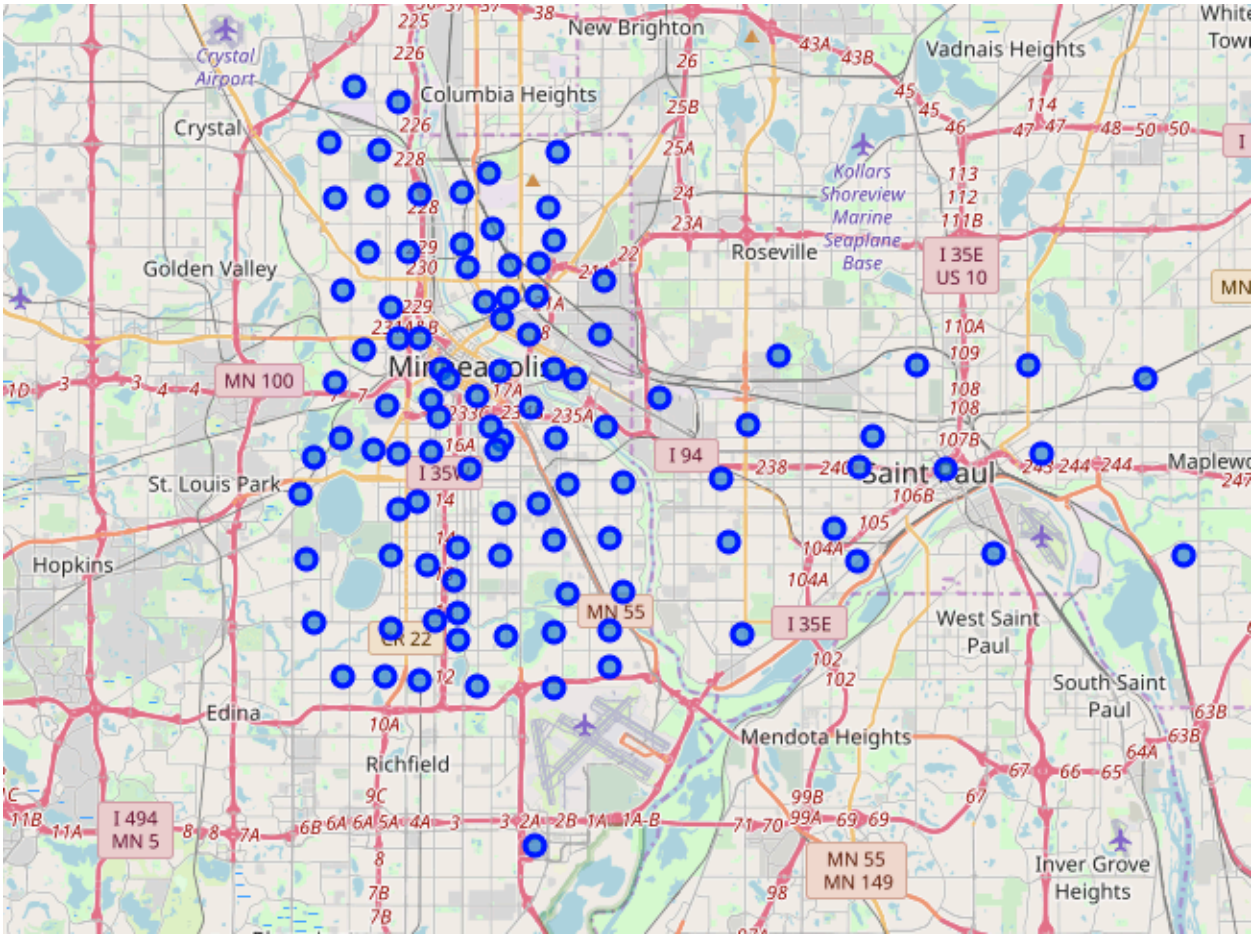

Figure 1: All 102 Neighborhoods in the Twin Cities

c. A stability analysis for K-means clustering was performed to examine the best number of clusters for neighborhood analysis.

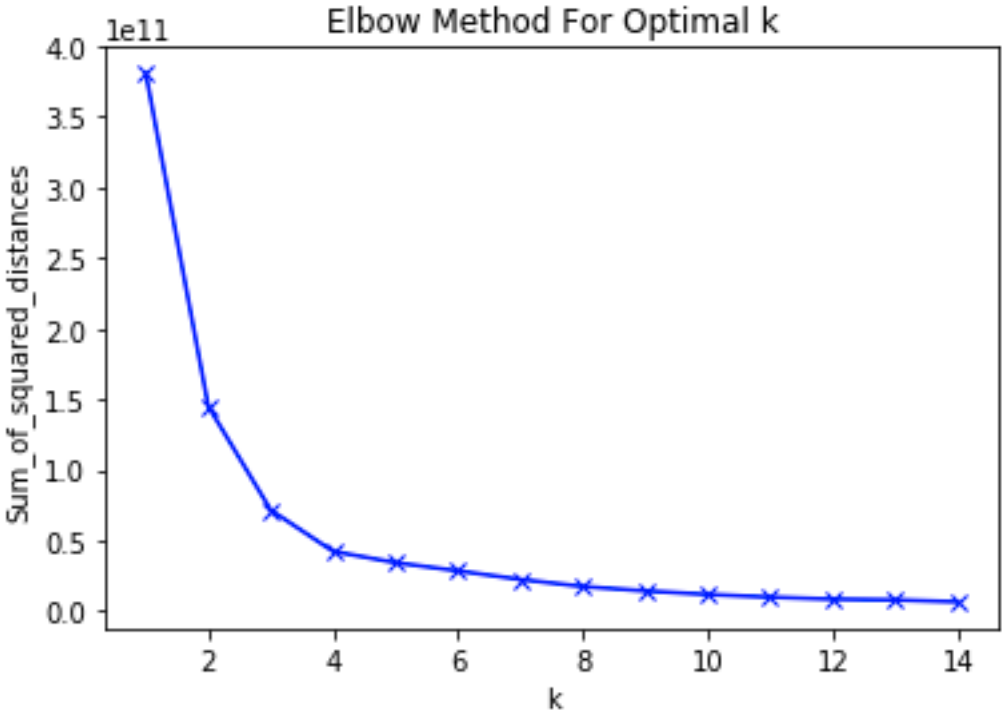

Figure 2: Analysis of potential cluster sizes in K-means clustering analysis

4. Predictive Modeling
    a. The main analyses in this work were:
        i. A mean of ranks in the factors of interest normalized for the number of neighborhoods to give a 'Minnesota Nice' factor in the interval of [0,1] with '1' being the most 'Minnesota Nice' neighborhood.
        ii. Top ten neighborhoods (with Minnesota Nice score):
            1. Nicollet Island-East Bank, Minneapolis 0.921569
            2. St. Anthony West, Minneapolis 0.859477
            3. Downtown East, Minneapolis 0.848856
            4. Sheridan, Minneapolis 0.826797
            5. Bottineau, Minneapolis 0.808824
            6. Field, Minneapolis 0.792484
            7. Page, Minneapolis 0.776144
            8. North Loop, Minneapolis 0.772059
            9. East Isles, Minneapolis 0.770425
            10. West Calhoun, Minneapolis 0.746732

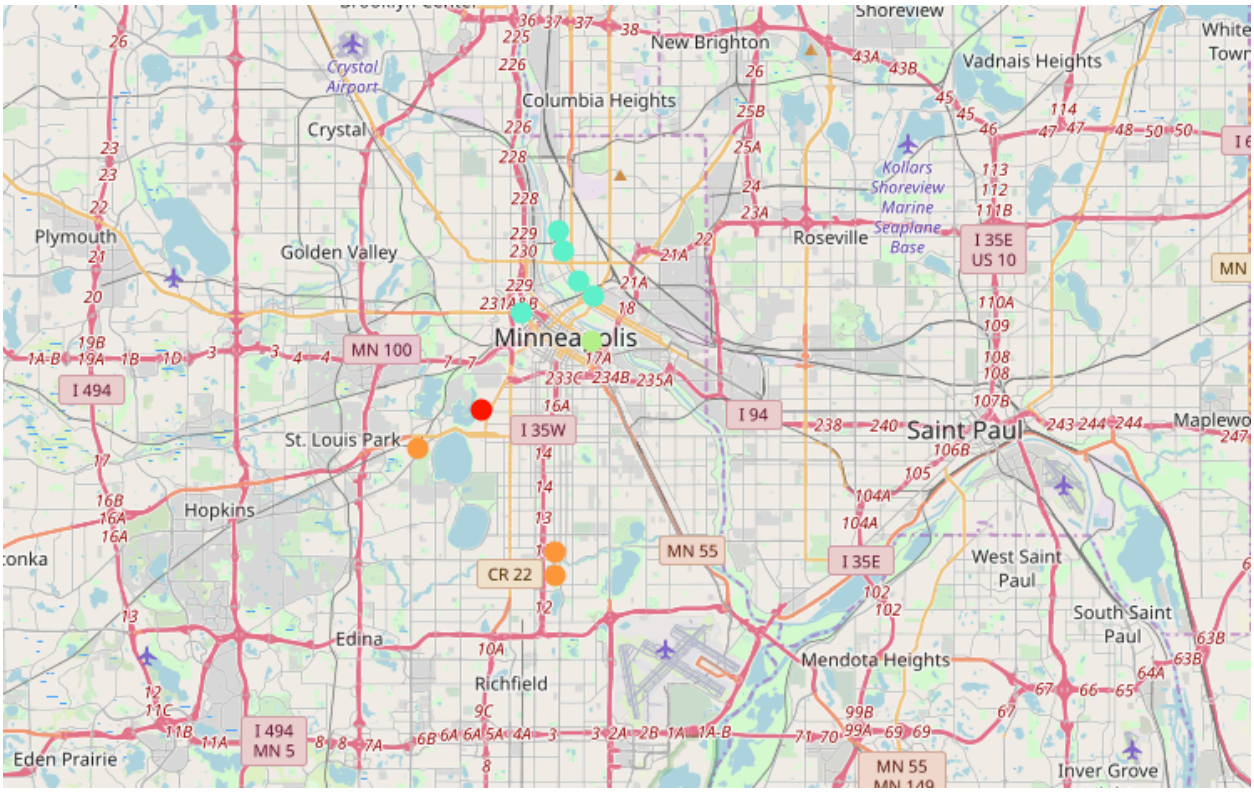

Figure 3: Locations of top ten 'Minnesota Nice' neighborhoods

iii. A cluster analysis of the neighborhoods with no geographic data attached. Using the top ten venue types per neighborhood, parks per capita, churches per capita, schools per capita, median household income, median home value, walk score, percentage of population under 18, total population, and average household size were used to cluster the neighborhoods in six clusters.

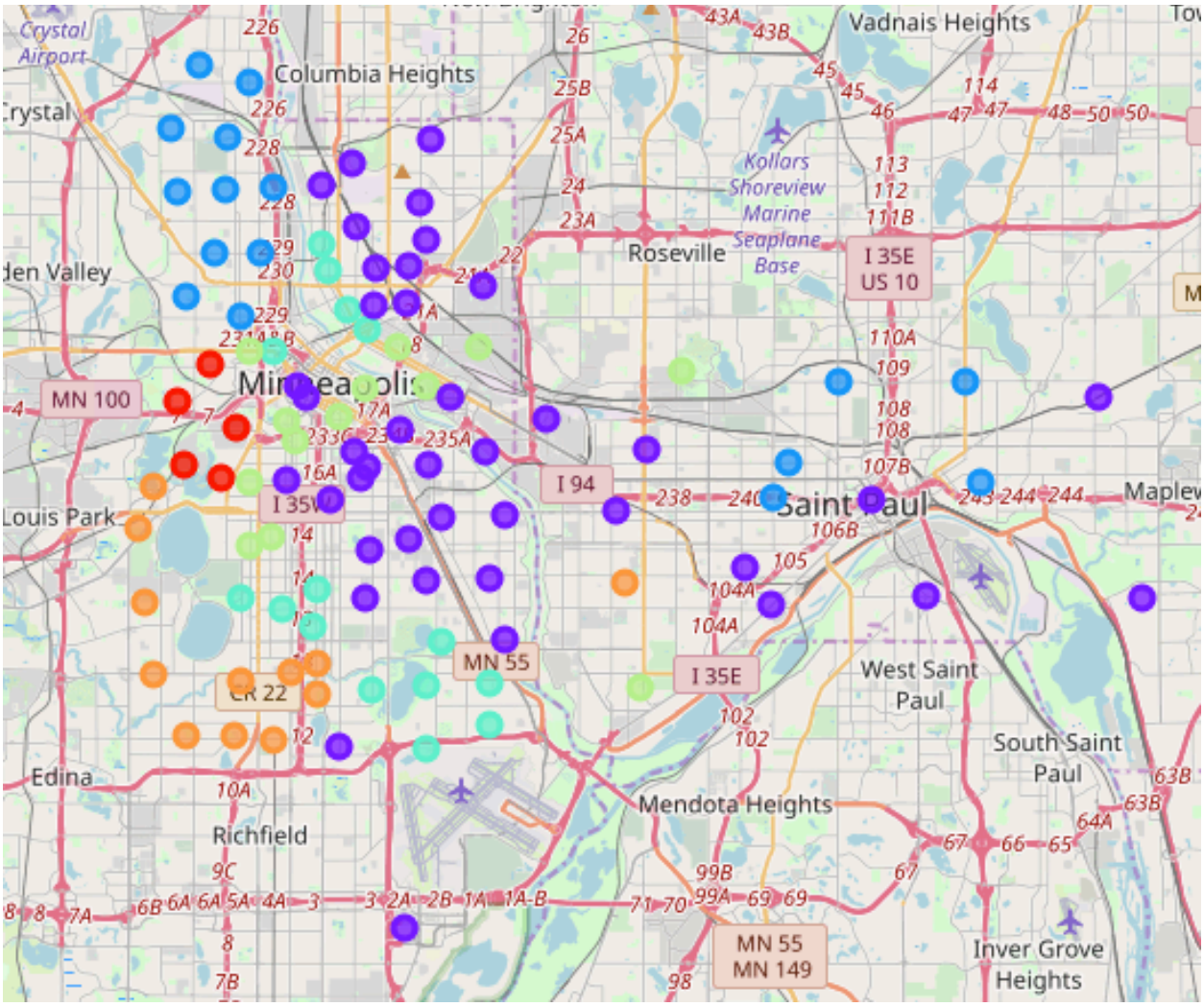

Figure 4: Location of neighborhoods colored by cluster assignment

5. Results and Conclusions
    a. High 'Minnesota Nice' neighborhoods were correlated with high income and home value.
    b. Even without geographic information in the clustering analysis, the clusters are localized geographically. This suggests that the *character* of the neighborhood is highly dependent on the neighborhoods surrounding it.
6. Future work
    a. Further work could be performed on the definition of the 'Minnesota Nice' factor.
    b. Statistical models could be developed to see what factors are the most influential in predicting median home value.
    c. Statistical models could be developed to analyze if a certain type of venue is influential in median home value.