A Recommendation Tool for Real-Estate Neighborhood Search

neighborhoodsearch.net
fyd113/team163

Morteza Maleki, Brad Beattie, Nivedita Sengupta, Poonam Dongare, Pankaj Kalan, Yidan Fu

College of Computing, Georgia Institute of Technology, Atlanta, Georgia

Introduction

Accurate information is essential to good decision-making. This is especially true when relocating or investing in real estate. We propose a recommendation tool to aid in selecting a neighborhood. When relocating to a new city, what tools are available to aid in selecting a neighborhood that suits personal preferences? How can one sort through the various attributes that could impact that decision and create a rational analysis in an unfamiliar place? This question and our project could potentially affect anyone involved in searching for a new place to live, investing in real estate, or researching these topics.

Currently available commercial tools (Zillow, Realtor, Trulia, Apartments.com) focus on available inventory and provide the user with some selection of limited attributes closely related to those properties but fail to consider these other essential decision attributes.

Methods

We created a novel tool offering a user three benefits a) it allows a user to select multiple attributes b) prioritize them and c) dynamically visualize their tailored results on a map. We created indices for each of the decision attributes then visualize the resulting weighted index based on the user's preferences.

Feature Value Calculation and Normalization:
Individual values are calculated for each feature based on the desired target value. The target value is taken as an input from the user (e.g. median home price of $300k to $500k) or predefined (crime rate should be minimum). The values are normalized to arrive at a scaled index in the range [0,1]. Table 2 depicts such normalization.

User Weights:
The user can select a preferred categorical weight of "Low", "Medium" or "High" for each feature. We define "Low" = 0, "Medium" = 0.5, and "High" = 1. Individual user weight values are normalized so that all weights sum to 1:

$$\omega_n = \frac{\omega_{selected}}{\sum \omega_{selected}} \tag{1}$$

User Weighted Average:
Our final recommendation index, I_{rec} displayed in the visualization is the weighted average of the individual feature indices,

$$I_{rec} = \omega_{user_1} I_{attr_1} + \omega_{user_2} I_{attr_2} + ... + \omega_{user_n} I_{attr_n} \tag{2}$$

where each I_{attr_n} is the unity-based normalized feature index and each ω_{user_n} is the user-specified weight for that feature where$(\sum \omega_{user_n} = 1)$

Datasets

Table 1. Data Used in Recommendation Tools

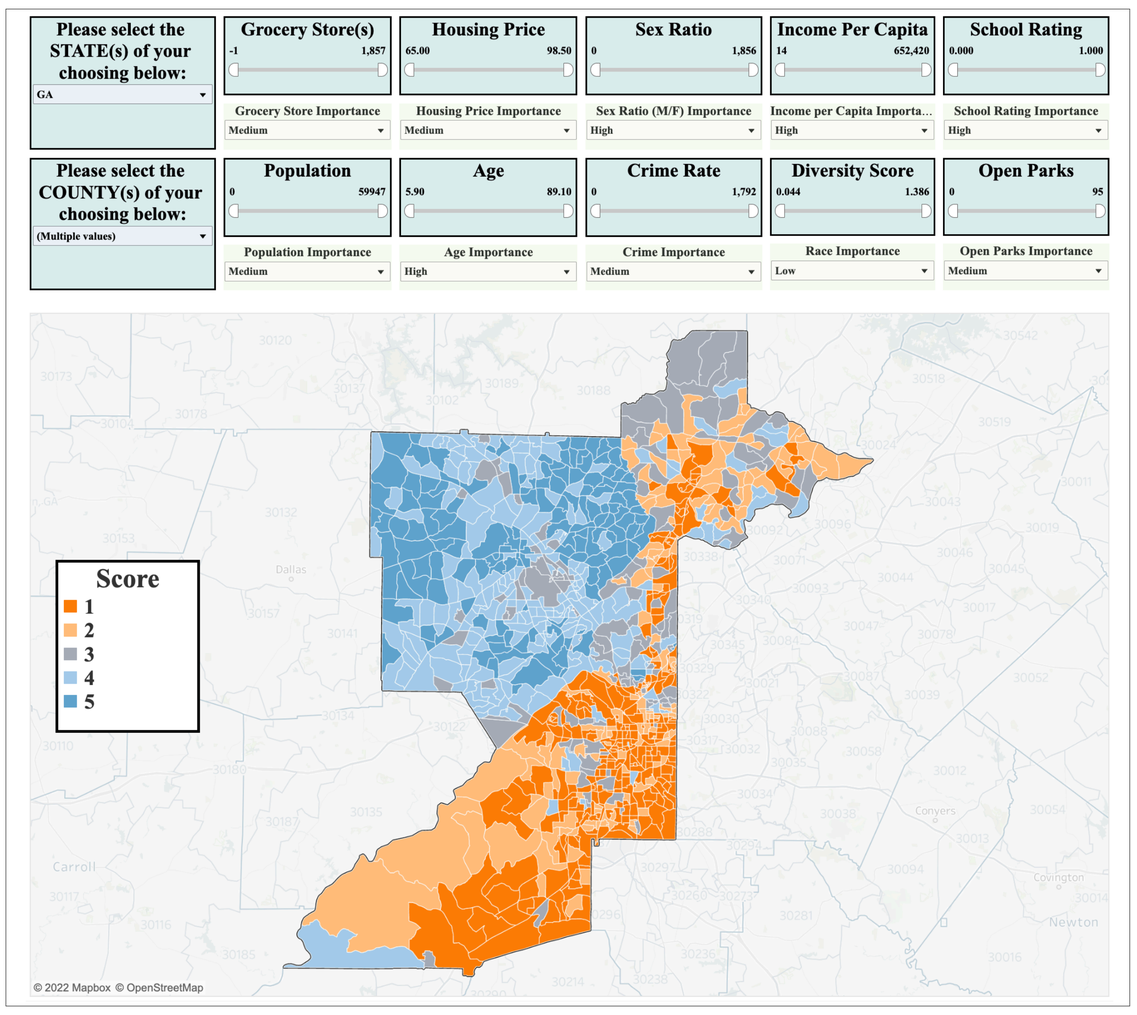
Data Attribute	Data Source	Number of Records	Number of Features	Year
Population Sex Race Home Price Median Age Income per Capita	Census	242,341	12	2020
Crime Rate	FBI	3,137	24	2021
Grocery Availability	Census & openIPSR	1,036,463	31	2017
Parks Availability	Census	73,058	11	2018
School Rating	Census/Niche	16,394/13,822	4/32	2020/2022

Feature Selection/Normalization

Table 2. Features used in Interactive Dashbaord

#	Feature	Desired Target	I_{Attr}
1	Population	User selects upper limit and lower limit from the	if $(I_{lower} < I < I_{upper})$
2	Sex Ratio	available range shown on a slider	then 1
3	Home Price		else 0
4	Median Age		
5	Income per Capita		
6	Racial Diversity	Racial diversity is predefined to be maximum from the available range. It is calculated as an inverse of the standard deviation of the population of four major races.	$I_{diversity} = \frac{1}{\sigma(I_{wh}, I_{bl}, I_{as}, I_{hi})}$ $I_{normalized} = \frac{I_{diversity}}{max(I_{diversity})}$
7	Grocery Availability	Predefined as maximum from the available range	$\frac{I - I_{min}}{I_{max} - I_{min}}$
8	Parks Availability		
9	School rating		
10	Crime Rate	Predefined as minimum from the available range	$1 - \frac{I - I_{min}}{I_{max} - I_{min}}$

Interactive Dashboard



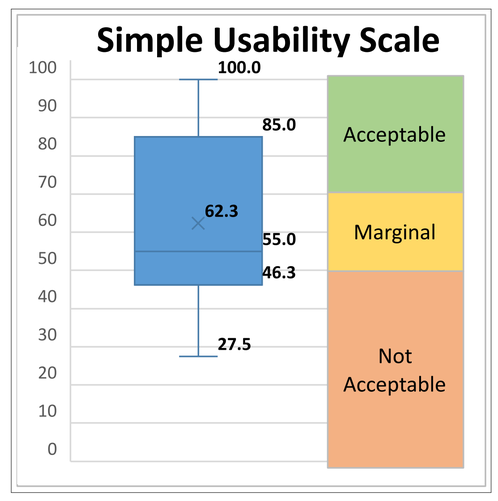
Evaluation

Metrics
We identified two criteria that could yield measurable results: user satisfaction and time to reach decision. Due to time and budge limit, our focus is on measuring user satisfaction.

User Satisfaction
We utilize Brooke's standardized usability assessment (SUS) [1] to provide an empirical measure of usability. At scale, we would want to recruit a large sample size and utilize multiple evaluations, but given the limitations of our scope it was only practical to sample a small number of family and friends.

We found that in our sample group, the SUS testing yielded scores with mean 62.3 (shown in first graph below). Based on the adjective ratings (shown in second graph below), we are within the marginal acceptability range.

Comparison with Other Methods
A more thorough approach for evaluation our method would split the recruited sample population in half and have one group evaluate our solution and the other group evaluate any solution they are able to find on the internet. In this case, it could measure the user satisfaction and time on task simultaneously.



Conclusion

Our work has resulted in an interesting exploratory tool that has the potential for further refinement especially on the number of dimensions considered. This tool could provide insights to a subset of renters and home buyers that none of the existing market tools possess. [2] Our tool provides a unique method for visualizing and understanding data that affects where we live. Ultimately, we feel that a more informed decision will produce better results for the user and the community.

References

[1] John Brooke et al. Sus-a quick and dirty usability scale. *Usability evaluation in industry*, 189(194):4–7, 1996.

[2] Radosław Cellmer, Aneta Cichulska, and Mirosław Bełej. Spatial analysis of housing prices and market activity with the geographically weighted regression. *ISPRS International Journal of Geo-Information*, 9(6):380, 2020.