Week 6: House Hunting, a mixed-method approach

Nate Bachmeier

TIM-7211: Research Design Methodologies

January 31, 2021

Northcentral University

# House Hunting, a mixed-method approach

Finding the perfect home within a densely populated metropolitan area like New York City or Tokyo is challenging. Within these cities, the price of homes fluctuates widely, as does its culture and amenities. The current pandemic also adds another layer of complexity, which imposes safety-regulations on home showings. Residents address these challenges by examining listing on real-estate web sites like Zillow and Redfin. However, these portals are full of erroneous information, limiting their usefulness. This work proposes an alternative solution, called the utility score, using a mixed-method mechanism.

# Problem Statement

Densely populated areas like the Greater New York City Metropolitan area have massive public transportation systems. These systems allow workers to live as far as Philadelphia or Connecticut and use rail-ways to commute into the city within ninety minutes. Meanwhile, other geographically closer areas might have indirect service routes, making a commute to Manhattan even longer. Property tax rates across this region range from 0.9 to 2.8%, making the net monthly cost fluctuate widely on same-priced houses. Further, the amenities like access to shopping, quality of schools, and crime-levels vary widely between housing blocks.

These distinctions make it challenging to compare two houses with traditional filters, such as housing prices. Homebuyers respond to this needle in a haystack problem by randomly searching real-estate sites like Zillow and Redfin. This process is highly inefficient and requires significant time investments. Alternatively, brokerage services and real-estate agents exist, but these services are expensive and require a time commitment. While some buyers are willing to wait for the perfect home (global maximum), others become discouraged or settle for what they can find (local maximum).

# Purpose Statement

Finding the global maximum is challenging because a system for making apples-to-apples comparisons does not exist. Without this capability, it is not possible to assess the available homes mechanically. However, a home is a profoundly personal purchase that makes this assessment opinionated. Consider the differences between a young couple and a family with children. While the family wants high-quality schools and quiet streets, the couple places a greater focus on the nightlife.

Since this problem contains elements of both quantitative and qualitative analysis, a mixed-method approach is most suitable. Mixed-method research draws on the strengths of both numerical and non-numeric information (Creswell, 2014). If the solution uses only numerical value, then it would not meet the personalization requirement. Similarly, a purely qualitative approach would miss-out on the statistical nature of comparing houses.

The strategy needs to begin with a qualitative process that identifies the individual buyer’s preferences (see Table 1). Next, these preferences become arguments into a *house* *utility scoring function.* This function enumerates the different housing areas within the commute distance and provides a match score. Finally, a quantitative study examines these scores to seek the best deals on fantastic homes.

Table 1: Example Preferences

|  |  |  |
| --- | --- | --- |
| Preference | Description | Value-type |
| Commute Information | How far will the person travel for work | Minutes and Location |
| Quality of Schools | The value placed on the public school system around the home | 1 to 5 stars |
| Availability of Nightlight | The value placed on easy access to bars, pubs, and late-night restaurants | 1 to 5 stars |
| Area Feel | The value placed on the community being dense | Rural to Urban |
| Apartment Feel | The value placed on the property is an apartment or condo | 1 to 5 stars |
| Family Home Feel | The value placed on the property is a family home | 1 to 5 stars |
| Home Age | The value placed on the property being recently built or renovated | 1 to 5 stars |
| Budget | The acceptable range of total monthly expenses | Range in Monthly Dollars |
| Crime Levels | The willingness to live in a higher risk area in exchange for a better deal | Low to High |

# Data Collection

There are fundamentally two artifacts that this research attempts to produce. First, identifying the influential preferences. Second, a function that ranks housing options across the commutable distance.

## Qualitative Information

The real-estate Multiple Listing Service (MLS) fed exposes hundreds of attributes about an available home. However, this information is incomplete and frequently lacks data about the specific community. It also contains numerous attributes (e.g., compliance requirements) that are irrelevant to the prospective buyer. Instead of reusing this schema verbatim for the preferences, an open-ended survey needs to occur with buyers at different phases of their journey. For instance, someone looking for a week has different requirements than another person that has completed the purchase.

## Quantitative Information

# Data Analysis