Enhancing The Home-buying Experience with Data Science

By George Pinto

No need to request edit permission. To make a copy for your presentation, use “File->Download”.

Abstract

1. In our Real Estate business we need to figure out how can use modern data science to drive sales into all types of consumer demographics while removing economic/social discrimination from our sales process and sell to people we may not have reached with a regular sales format and:

2. Use that same data science to enhance the consumer experience to the point that each home buyer feels that we were working just for them, working to find exactly the environment they were looking for at a price they could afford

3. Using the Google Geocoding API, median home price ranges by Neighborhood from Zillow's economic data for research and Sci-Kit Learn machine learning libraries, I have created a model in which I have clustered Chicago, IL neighborhoods by the things that make them fun, or interesting or dependable to people along with the price data in each location so they can find the things they're looking for in the environment surrounding their home, and their home at a price they can afford.

Motivation

Psychologically, people are driven by the environment they live in, a home is a place they can eat, sleep and co-exist in while their environment is what they strive for, they want better schools, better grocery stores, better neighbors with similar interests. They want a better life. That environment they want is very precious, very personal and a much stronger impulse for buying that simply just the home itself. Yes, I want a sales solution that will increase sales, but I want a solution for them, for people, to get a lot more from the homebuying experience than what they expected -- to find everything they want in terms of the physical space surrounding their lives at a price they can afford – not just a home.

Dataset(s)

1. From https://www.zillow.com/research/data/ I obtained my custom csv file as follows: I selected the Data Type as ZHVI (Zillow Home Value Index) for the current month and the Geography data arranged by Neighborhood, then proceeded to download the csv file.

2. For the type of locations surrounding each of the Neighborhoods the places API from Foursquare https://developer.foursquare.com/places-api by creating an account as a developer and obtaining credentials to use their API.

3. For matching locations with latitude and longitude data I used Google's geocoding API https://developers.google.com/maps/documentation/geocoding/start

4. I also pulled data from the City of Chicago's Geographical Information systems to obtain accurate coordinates from different Chicago Neighborhoods: https://www.chicago.gov/city/en/depts/doit/provdrs/gis.html