Relocation Neighborhood Recommendations

Author: Trinath Sahu
Version 1.0
Last updated date: 2nd November 2019

Contents

Revision History

Use this table to identify items that are new and/or changed in each release of this document.

Revision Number	Revision Date	Author	Summary of Changes	Changes marked
1.0	02-Nov-2019	Trinath Sahu	Initial draft version.	N

1. Introduction

1.1. Background

As per report, between 2012 and 2013, about 36 million Americans moved to a new home. That is about one in every 8 people move to a new location. The United States Census Bureau has a new report with a much bigger dataset about why people move. Top reasons of these movements include new job opportunity, Long term onshore travel, Own home, and job transfer. People who move within the same county are almost twice as likely to do so for housing-related reasons as those who move to a different county. The reasons people move to different counties are pretty evenly distributed among family, jobs and housing, whereas same-county movers do so disproportionately for housing-related reasons.

Advantages of relocation might include:

Add on to job security. Sometimes, relocating for work is the only way to keep your job. In many cases, it is strictly mentioned in the offer letter that wherever employee is put he/she has to go to that place and work. If you decide to make the move, your employer may feel you are loyal to the company. This could result in job security if your employer has to make cuts later on.

It can be an opportunity for career advancement. It is not uncommon that employees look for job change to get new opportunity in terms of roles, responsibilities, pay etc. In addition to job security, relocating for work can be a career accelerator, leading to promotions. It can give the opportunity to excel, take things to the next level and set the bar for others behind you.

It can increase your standard of living and quality of life. Kate Windleton is a relocations manager at Strong Move, a company that helps people relocate both domestically and internationally. She said the possibility of a better standard of living is an important consideration. Take into account the salary offered as well as the cost of living. You may be taking a pay raise or pay cut, but depending on the new cost of living, you can end up with a smaller or larger disposable income.

It is great for personal development and new experiences. Relocation is an opportunity to start fresh and reinvent yourself. Try the things you've always wanted to do but haven't made the time for.

You can make new friends. Moving is the perfect opportunity to welcome new, positive friendships and leave the toxic ones behind. You can use online apps, rec leagues and local meetups to find people who share your same interests.

You can move to a better climate. Whether you are currently living somewhere cold and you've always wanted to live in a tropical area or vice versa, this may be your chance to finally move. Moving to your desired climate may improve your quality of life.

1.2. Problem

While there are lot of advantages of relocation, at the same time there are lot of challenges too in the relocation. Some of them include:

You must find housing in an unfamiliar area. If you are unable to take advantage of corporate or temporary housing, finding a place to live in a new city can be a major challenge.

Right schooling and child care can be tough. If you have children, you will need to find housing that is within the boundaries of the school district you would like your children to attend. You will also have to find trustworthy child care, which can be worrisome if other family members have watched your children, shuttled them to and from school and extracurricular activities, etc. Utilize trusted internet forums and ask around for references.

Transportation system. You might not have your owne vehicle in the new place, so you will be dependent on public transportation system. So, you will be interested in a place where transportation service is nearby.

And many more…

If you are currently residing in a location where your needs are nearby, you might feel uncomfortable in moving to different location. Sometime, these challenges refrain you from taking up new good career opportunity and hence becomes road blocker for your career growth. So, it is very important as well as a difficult task to find a locality in unfamiliar place which should be similar to the one you are staying currently. The task becomes even more challenging when you relocate to a large city where it is very difficult to decide which factors to consider and which to factor out.

1.3. Interest

It will be of interest to all those people who relocates from Newyork to Toronto (which is a big crowd). If some automated system suggests them the recommended places where they can move to and live in, at least it will reduce their effort drastically by limiting scope of their search/exploration. All that you need is a place in the new neighborhood which is similar to your current neighborhood. If it can say neighborhood x, y and z in Toronto are similar to your current neighborhood 'a' in Newyork then you can imagine how much it will be helpful to them.

NOTE: Though this project is specifically to help people who is relocating from Newyork to Toronto, the underlying program can easily be modified to make it work for any other countries/locations.

2. Data acquisition and cleaning

2.1. Data Source:

In order to solve the problem, we will need location data (like venues around the Toronto locations), Postal code, Boroughs, neighborhoods in Toronto and New york, and their longitude & latitude.

For location data, we will use Fourthsquare location data service provider APIs;
For postal code, Boroughs and neighborhoods, we will scrap the Wikipedia page
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

We will get latitude and longitudes from a csv file called as Geospatial_Coordinates.csv.

2.2. Data cleansing:

Since there is no file/dataset present which has the Postal Code, Boroughs and neighborhoods and we are using Wikipedia website page for the same, website scrapping feature needs to be used to extract the required data. Luckily, thanks to outstanding Pandas read_html command which can search for any tables present in the page and stores in pandas dataframe directly. But, that does not end there. Following tasks need to be done in order to clean up the data:

➢ Many rows have Borough' column as 'Not assigned'. We will process the cells that have an assigned borough and Ignore cells with a borough that is Not assigned.

➢ It seems, more than one neighborhood can exist in one postal code area. For example, in the table on the Wikipedia page, you will notice that M5A is listed twice and has two neighborhoods: Harbourfront and Regent Park. These two rows will be combined into one row with the neighborhoods separated with a comma.

➢ If a cell has a borough but a Not assigned neighborhood, then the neighborhood will be the same as the borough. So, the value of the Borough and the Neighborhood columns will be Queen's Park.

2.3. How data will be used to solve the problem:

1. Cleanse the neighborhood data as mentioned above
2. Store the lattitudes and longitudes of all neighborhoods in dataframe
3. Call Foursquare APIs to explore the neighborhoods in Toronto city. I.e. get all nearby venues around all the neighborhoods in Toronto. Do the same for current neighborhood in Newyork.
4. Store all the venues around all Toronto neighborhoods.
5. Then add an additional entry with all the venues that are present nearby current neighborhood in Newyork.

6. Group the neighborhoods into Clusters by using clustering algorithm(such as Kmeans). Since we have added an entry for current neighborhood in Newyork, all the Toronto neighborhoods which are similar to the Newyork one will fall into same cluster. That's the trick of adding this additional row.
7. Visualize the data by highligting different clusters in different colors. So, all the Toronoto neighborhoods present in the same cluster as point 5 are similar to current Newyork neighborhood. Suppose cluster# for the current neighborhood is found to be 8 , then all the neighborhoods of Toronto which is assigned cluster no. as 8 will be the ones you are looking for. Those neighborhoods will be highlighted in special color with proper legend. The folio map will clearly show them which will help you to take decision on which locality to move into.

2.4. Feature Selection:

➢ Venue categories around neighborhoods of Toronto: We will call Foursquare APIs to get the venues in neighborhoods of Toronto. Then using Onehot coding, we will transform them into features.

➢ Top venue categories around current neighborhood in Newyork. This is required to compare new neighborhoods against current one.

Some examples of feature columns (which is venue category in this case):
Park, school, College Academic Building, College Arts Building, General College & University, Restaurant, pizza, Bank,Gym, Bus station, bus stop, cafe, dry cleaner, food court, Grocery Store, Health & Beauty Service   Health Food Store, High school, Movie Theater, Kids store, Laundry Service, Medical Center, Hospital, Pharmacy, Metro station, Shopping Mall etc.