Data

In this program, I will leverage the power of Foursquare, in which we can find detailed locations data. I will use the url that starts with https://api.foursquare.com/v2/venues/explore to explore nearby venues of the locations that the user input. After some categorizing, it ultimately generates a venues metric that lets me know whether or not the location has restaurants, shopping mall, bar… nearby. Combined with the ratings provided by the user, I am able to guess how much the user likes each venues. This can be done by calculating the dot result of two metrics, the venues metric and the rating metric. After having the resulting metric, I can see the user's preference, and use it to calculate the possibility that the user will like other locations.

Let's look at a simple example. Say Bob has been to Rouge Hill, Scarborough. Bob gave the place a rating of 10, which means he loved it. Now, the program can find data from Foursquare that tells us what are the most abundant venues in Rouge Hill, specifically, bars, Vietnamese restaurants, and coffee shops. With this information, I will boldly guess that Bob enjoys cities that have bars, Vietnamese restaurants, and coffee shops. Then, I will find cities that meet this criteria, and recommend it to Bob. Of course, this recommendation does not seem very convincing when it is based on only one rating. That's why I programmed it to ask for five. In fact, the more information we have about the user, the better recommendation we can give to him/her. This scale of this program can be expanded for more accurate predictions if needed in the future.