Rental Listing – Neighborhood Metrics

1. A description of the problem and a discussion of the background. (**15 marks**)
2. A description of the data and how it will be used to solve the problem. (**15 marks)**

Introduction/Business Problem

One of the biggest challenges while looking for a new home is to find the neighborhood information of a rental unit. Important information like how many schools, restaurants, playgrounds etc. could be very useful while looking at a listing in a rental website. This gives the potential tenant to decide the worthiness of a house based on his/her preferences. For example, families with young children would be interested in schools and other kid related information while young couples would be interested in dining and entertainment options nearby.

For this project, I will focus on the city of Toronto. The solution would provide scores to a rental listing based on 5 metrics, which are education (schools and other places of learning), entertainment (cinema halls, parks, etc.), healthcare (hospitals, clinics etc.), basics (restaurants, grocery stores, etc.) and shopping (shopping malls, stores, etc.). Scores would be calculated based on comparing the number of venues in a kilometer radius to the average number of venues in a kilometer radius in the borough.

The goal of this project is to give a easily interpretable scores that would highlight the location advantage for a rental listing, compared to other listings in the same borough, and aid in decision making.

Data and Approach

I will be using Foursquare APIs to retrieve nearby venues for a listing by categories that we are interested in. I will predetermine what venue categories will fall into a what metric. For example, ‘Supermarket’ and ‘Restaurants’ in venue category in Foursquare will be aligned with metric ‘Basics’. ‘Movie Theater’, ‘Museum’ and ‘Nightclub’ would be aligned with metric ‘Entertainment’

Publicly available information will be used to calculate density for each metrics (categories identified previously) in a borough. We will use geolocation, borough area and multiple Foursquare API calls to retrieve number of venues by the categories we are interested in and calculate density for each.

For each listing, scores ranging from 1 to 5 will be assigned for each metric. A rating of ‘3’ or ‘Good’ means the listing aligns closely with the borough average, ‘5’ or ‘Excellent’ rating means the listing has much more number of venues than the average and ‘1’ or ‘Poor’ has much less number of venues than the average.

Rating	Text rating	What it means
5	Excellent	Borough average + two standard deviation
4	Very Good	Borough average + one standard deviation
3	Good	Borough average
2	Fair	Borough average – one standard deviation
1	Poor	Borough average – two standard deviation