Spatial Analysis Using Toronto School and Neighbourhood Data

Proposal Report

Ziyuan Shen

November 6, 2024

1 Background/Introduction

The spatial distribution of educational settings plays a critical role in determining access to quality education, especially in a diverse urban environment such as Toronto. Understanding the distribution of different types of schools across neighbourhoods can provide valuable insights into potential disparities and help identify opportunities for a more equitable distribution of resources.

1.1 Research Question

How are different types of schools such as Elementary, Secondary and Adult Learning spatially distributed across Toronto's neighbourhoods, and are there notable spatial clusters or patterns in this distribution?

2 Data and Proposed Method

2.1 Data Description

- **TCDSB Schools Dataset:** This dataset contains information on the location of Toronto schools expressed as point data. The study area for this dataset is the entire City of Toronto, with a focus on educational facilities managed by the Toronto Catholic District School Board (TCDSB). The type of spatial data represented is point data, and the dataset contains 210 locations. Each point represents the geographic location of a school. Attributes in this dataset are shown in Table 1.

Table 1

Attribute	Description
_id	Unique row identifier for Open Data database.
OBJECTID	ESRI-generated ID for spatial data.
TCDSB_CODE	School unique code
SCH_NAME	School name
ADDRESS_FULL	School address
MUNICIPALITY	Municipality where the school is located
POSTAL_CODE	Postal code of the school
TCDSB_LEVEL	School type: E–Elementary, S–Secondary, AL–Adult Learning, ADM–Administration Building
STATUS	School status
geometry	Geospatial information for each school, represented as point data.

Note: Description of variables used in the Toronto school dataset.

Toronto Neighbourhood Shapefile: The Toronto Neighbourhood Shapefile provides polygonal data representing the boundaries of Toronto neighbourhoods. The study domain for this dataset is also the City of Toronto and the type of spatial data represented is polygonal. The dataset includes 157 neighbourhoods, each represented as a polygon that defines its geographic boundaries. Relevant attributes include:

1.Neighbourhood ID: a unique identifier for each neighbourhood.

2.Geographic Boundaries: spatial data representing the boundaries of each neighbourhood to assess the spatial relationship to school locations.

These datasets will be used to explore the spatial distribution of schools within the defined neighbourhoods and to assess potential clusters and differences between different types of schools.

2.2 Proposed Spatial Analysis Methods

2.2.1 Spatial Computations:

Buffers and Proximity Analysis: Create buffers around each school to determine its catchment areas and analyze the proximity between schools and neighbourhood boundaries. This may help me understand the spatial reach and accessibility of different school types.

Spatial Joins: Use spatial joins to combine neighbourhood boundary data with the school point data. This will help me to evaluate how the distribution of different school types correlates with neighbourhood geographic characteristics.

Distance Calculations: Calculate the distances between schools and the boundaries of their neighbourhoods to understand accessibility and spatial relationships better.

2.2.2 Spatial Statistical Analysis:

Point Pattern Analysis: Point pattern analysis will be applied to determine whether the distribution of schools in Toronto is random, homogeneous or clustered. This included calculating nearest neighbour distances and analyzing school clustering patterns.

Moran's I: Calculate Moran's I to assess spatial autocorrelation in the distribution of different school types. This will help determine whether the locations of specific types of schools exhibit clustering.

3 Summary Statistics

For further analysis, I split the geometry column into long(longitude) and lat(latitude) columns and then removed the geometry column from the TCDSB Schools Dataset.

I will perform an Exploratory Data Analysis (EDA) to summarize the spatial properties of the dataset.

Firstly, I created a preliminary visual map as shown in @map-1 below. that overlays the location of the school and the boundaries of the Toronto neighbourhoods on the map. The map helps to illustrate the spatial distribution of schools in different neighbourhoods . The map visualization includes the following features:

Neighbourhood Boundaries: Represented as polygons, showing the geographic extent of each neighbourhood.

Schools: Represented as dots, showing the location of different types of schools within the neighbourhood.

Click here to view the interactive map

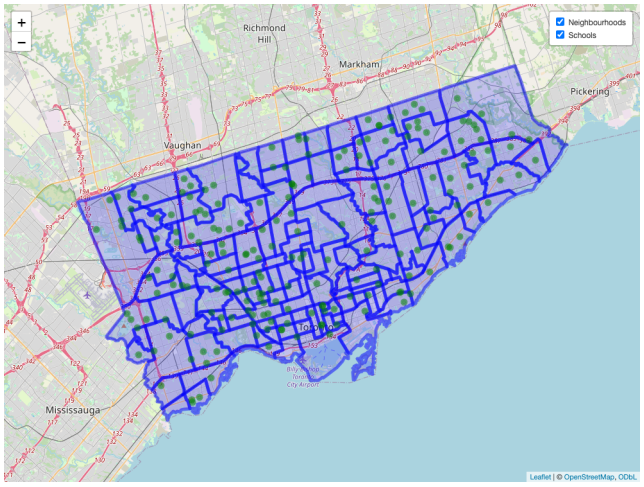

Figure 1: Leaflet Map of Toronto Schools

Then, I tried to summary the number of schools within each neighbourhood and their types in Table 2. There are four types of school. Calculating basic descriptive statistics to understand the distribution of different school types across neighbourhoods.

Table 2: Summary Statistics for School Types

School Type	Total Schools	Average Latitude	Average Longitude
ADM	5	43.76468	-79.37903
AL	10	43.71300	-79.39242
E	164	43.72274	-79.39980
S	31	43.71784	-79.40457