

1 Data

1.1 Delineating Neighborhoods

The city of Vancouver is broken up into 22 different neighborhoods, which are referred to as Local Areas by the city. An open data portal is also maintained by the city of Vancouver, which contains the following fields:

- MAPID: A unique 2 character identifier for each neighborhood.
- Name: The colloquial name for each neighborhood.
- Geom: The geometry of the boundaries between each neighborhood in the GeoJSON format. The coordinates are in latitude and longitude.
- geo_point_2d: The centroid of each neighborhood in latitude and longitude.

A truncated sample of the file is included for reference:

MAPID	Name	Geom	geo_point_2d
KC	Kensington-Cedar Cottage	{“type”:“Polygon”, “coordinates”: [[[-123.056587, 49.261981], ...]]}	49.246686, -123.072884
...
...

We also need a way to characterize different neighborhoods in all of the communities outside of Vancouver. For consistency’s sake, we will use the Forward Sortation Area portion of Canadian postal codes as a way to designate neighborhoods. In Canada, postal codes are six character long alphanumeric strings in the format A1A 1A1, where A is a letter and 1 is a digit. The first three characters in any postal code designate the Forward Sortation Area, which is a geographical region in which all postal codes start with the same three characters. Statistics Canada maintains boundary files, which are updated after each census is performed. The boundary files contain the following fields:

- CFSAUID: The unique ID of each forward sortation area. These are the first three characters of each postal code.
- PRUID: A unique 2 digit number that designates which Canadian province the forward sortation area falls in.
- PRNAME: The name of the province that the forward sortation area falls in.
- geometry: A polygon designating the boundary of each forward sortation area. The coordinates are in the Lambert Conformal Conic projection.

A truncated sample of the file is included for reference:

CFSAUID	PRUID	PRNAME	geometry
V3E	59	British Columbia / Colombie-Britannique	POLYGON ((4052781.049 1997949.749, ...))
...
...

1.2 Characterizing Neighborhoods

Once we have designated the different neighborhoods in our area of interest, we need a way to characterize each neighborhood. For this purpose, we will use Foursquare location data. Using the explore endpoint of the Places API, we are able to generate a list of venues that fall within a certain distance of a user supplied location. We can use this endpoint to generate a list of venues that fall within a certain neighborhood:

Neighborhood	Latitude	Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Kensington-Cedar Cottage	49.246686	-123.07288	Famous Foods	49.248356	-123.071471	Grocery Store
Kensington-Cedar Cottage	49.246686	-123.07288	Anytime Fitness	49.249640	-123.075450	Gym / Fitness Center
...
...

Once we have a list of all of the venues that are present in a given neighborhood, we can aggregate the venues based on their category. This will allow us to determine the frequency of each venue category in each neighborhood. The frequency of each venue category can then be used to characterize each neighborhood, allowing us to compare them to other neighborhoods.