

# Informe completo selección de vecindarios en Toronto

## Planteamiento del problema

Toronto es una de las ciudades más importantes de Canadá. Cuenta con entorno urbano multicultural. Es un lugar ideal para estudiantes de intercambio, ya que se comparte una riqueza cultural e idiomática. Por lo que, es importante que un estudiante pueda elegir el barrio para residir, que se acomode a su entorno cultural. Es así como surge el planteamiento del problema con ¿cuál es el vecindario ideal para un estudiante de intercambio?

## Datos a utilizar

Se utilizarán los datos geoespaciales de los vecindarios, así como la creación de una etiqueta según proximidad.

## Metodología

Para el proyecto, se utilizó las bibliotecas de Python Numpy, Pandas, Geopandas, Plotly, Scikit learn, Requests y Geopy.

Se aplicó la metodología de ciencia de datos, donde primer se realizó la limpieza de datos, eliminando los duplicados y depurando los valores missing. Luego, se procedió con la exploración de datos, aplicando al visualización y descripción estadística de datos.

También se aplicó machine learning, utilizando el algoritmo de clustering para agrupar a los vecinos en base a la proximidad según sus características. Por último, se analizan los resultados en base a visualizaciones.

## Resultados

Se realizó la limpieza de datos, donde se definió como variables el código postal, ciudad, vecindario, latitud y longitud.

|   | PostalCode | Borough          | Neighborhood                     | Latitude  | Longitude  |
|---|------------|------------------|----------------------------------|-----------|------------|
| 0 | M3A        | North York       | Parkwoods                        | 43.753259 | -79.329656 |
| 1 | M4A        | North York       | Victoria Village                 | 43.725882 | -79.315572 |
| 2 | M5A        | Downtown Toronto | Regent Park, Harbourfront        | 43.654260 | -79.360636 |
| 3 | M6A        | North York       | Lawrence Manor, Lawrence Heights | 43.718518 | -79.464763 |
| 4 | M7A        | Queen's Park     | Ontario Provincial Government    | 43.662301 | -79.389494 |

De esto, se identificó que la ciudad de Toronto cuenta con 100 vecindarios.



De estos vecindarios, se generó una etiqueta para la identificación de clusters según proximidad.

| Cluster Labels | PostalCode | Borough | Neighborhood                                      | Latitude  | Longitude            |
|----------------|------------|---------|---|---|----------------------|
| 0              | 4          | M3A     | North York  | Parkwoods   | 43.753259 -79.329656 |
| 1              | 4          | M4A     | North York  | Victoria Village                                    | 43.725882 -79.315572 |
| 2              | 2          | M5A     | Downtown Toronto                                  | Regent Park, Harbourfront                           | 43.654260 -79.380836 |
| 3              | 3          | M6A     | North York  | Lawrence Manor, Lawrence Heights                    | 43.719515 -79.464763 |
| 4              | 2          | M7A     | Queen's Park                                      | Ontario Provincial Government                       | 43.662301 -79.389494 |
| ...            | ...        | ...     | ...   | ...   | ...                  |
| 98             | 1          | M8X     | Etlolcote   | The Kingsway, Montgomery Road, Old Mill North       | 43.853854 -79.506944 |
| 99             | 2          | M4Y     | Downtown Toronto                                  | Church and Wellesley                                | 43.665860 -79.383160 |
| 100            | 4          | M7Y     | East TorontoBusiness reply mail Processing Cen... | Enclave of M4L                                      | 43.662744 -79.321558 |
| 101            | 1          | M8Y     | Etlolcote   | Old Mill South, King's Mill Park, Sunnyside, Hui... | 43.836258 -79.498509 |
| 102            | 1          | M8Z     | Etlolcote   | Mimico NW, The Queensway West, South of Bloo...     | 43.828841 -79.520999 |

### Conclusiones

Con ello, se identificaron 4 clusters, lo que podría ayudar a un estudiante de intercambio para mudarse en el cluster que se encuentre cerca a su centro de estudios. Considerando que los vecindarios más alejados de la costa, son los de menor costo de arrendamiento.

