

Battle of the Neighborhoods

Coffee Shop in Vancouver, BC, Canada

Author: Rafael Vidigal de Moraes

Date: April 25th, 2021

1. Data

In order to fetch the crime details of Vancouver it was used a real world dataset that can be found at Kaggle datasets, this dataset can be found [here](#). Although this dataset was filled with important information regarding the criminal activity, all coordinates data are offset and in some cases it was not disclosed in order to provide privacy protection, as declared by the City of Vancouver's website. In order to fetch the Boroughs it was necessary to do it from Wikipedia. Geocoder API and Foursquare API were used in order to fetch venues listed in the neighborhoods.

The population dataset was also fetched from Wikipedia, saved in a csv file and uploaded to GitHub..

Properties of the Crime Report

- TYPE - Crime Type
- YEAR - Recorded year (2018)
- Month - Recorded Month
- DAY - Recorded Day
- HOUR - Recorded Hour
- MINUTE - Recorded Minute
- HUNDRED_BLOCK - Recorded Block
- NEIGHBORHOOD - Recorded Neighborhood
- X - GPS longitude
- Y - GPS latitude

Wikipedia's data set didn't require web scraping. The page has information regarding Neighborhood and boroughs.

Open Cage API

- **Neighborhood:** Name of the neighborhood in the borough.
- **Borough:** Name of the borough.
- **Latitude:** Neighborhood Latitude.
- **Longitude:** Neighborhood Longitude.

The dataset from Kaggle was a heavy file and GitHub couldn't handle it. The Vancouver Crime Dataset had approximately 600.000 rows. Therefore the only crimes there were the crimes that took place in 2018, in order to reduce the dataset. It was created a csv file out of the dataset and it was uploaded to GitHub.

In order to provide privacy protection the coordinates were improperly encoded. This information was declared by the City of Vancouver and can be found in its website and in a disclose on Kaggle. Because of the misguided information in columns X and Y I decided to drop them along with columns Month, Hour, Day and Year.