Where to open the next Doghotel in Toronto

Capstone project | Karin Leisink | September 2019

Find the best location in Toronto for starting a dog hotel

Requirements

- There should be no dog sitting or dog boarding facilities in the neighbourhood yet and preferably neither in the surrounding area
- There are a substantial amount of registered dogs in the neighbourhood
- Dog owners need to have a sufficient income (at least $70,000 a year)

Stakeholders and interested parties

- Chelsea, who wants to start a dog hotel in Toronto
- Other people who want to start a dog hotel
- Future data scientist who look for interesting Data Science case studies such as this one
- People who are considering a data science career and who are interested in doing the IBM Coursera Data Science introductory course.

Datasets

Nr	Dataset	Source
1	List of Toronto postal codes with corresponding Neighbourhood and Borough names	Scraped from a wikipage
2	List of Geospatial Codes of Toronto neighbourhoods.	Supplied by Coursera
3	List of registered cats and dogs in Toronto (2019)	Toronto Open Data
4	List of neighbourhood profiles in Toronto from the 2016 census	Toronto Open Data
5	List consisting of postal codes and Neighbourhood names to link the neighbourhoods mentioned in the neighbourhood profiles list to the scraped Toronto postal codes	Self-constructed with the help of wordpostalcode.com
6	List of dog boarding and dog sitting services in Toronto.	Self-constructed with the help of the google search engine.

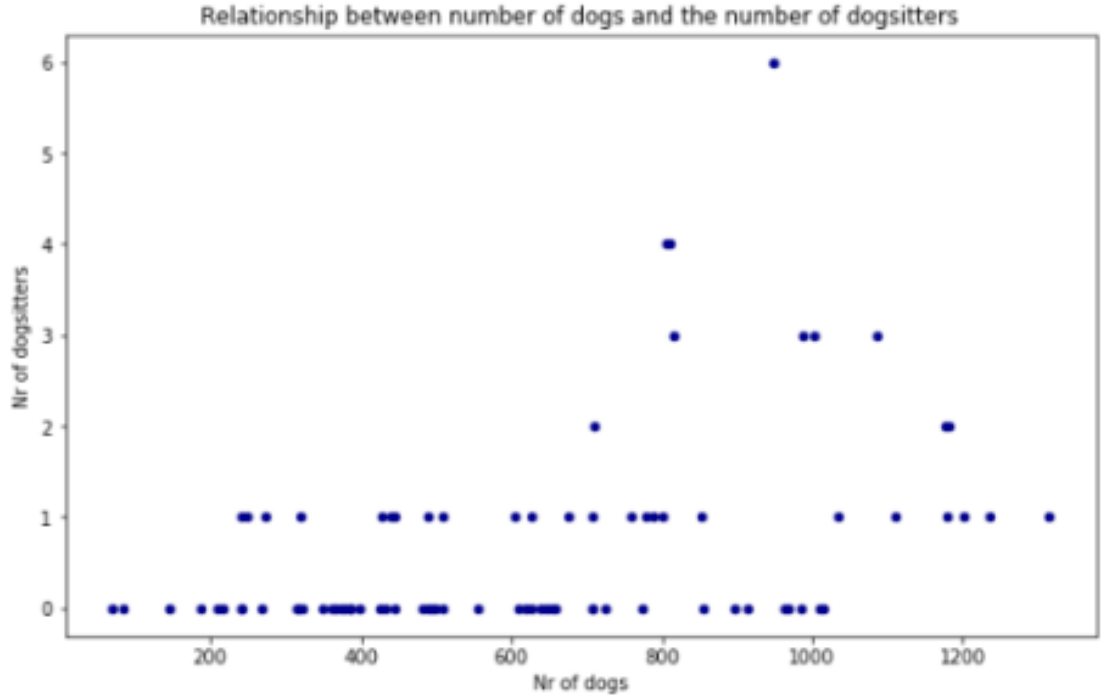
Most important columns after merging all datasets

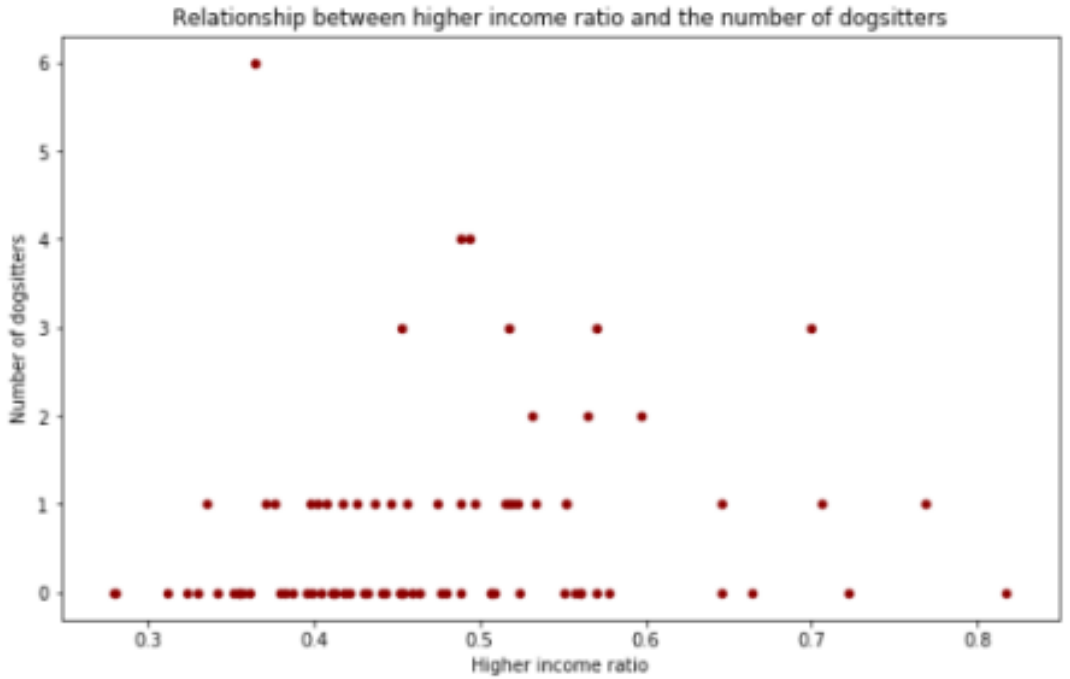
- The number of registered dogs per neighbourhood
- Percentage of households earning at least $70,000 a year per neighbourhood
- The number of dog boarding and dog sitting facilities per neighbourhood

Methodology

- Produce scatterplots to explore the relationship between the most important columns in the joined or base dataset
- segment the data using the K-means clustering technique
- Profile or name the clusters
- Visualize the obtained clusters on the map of Toronto
- Consider potential dog hotel locations

Scatterplots to investigate relations between data



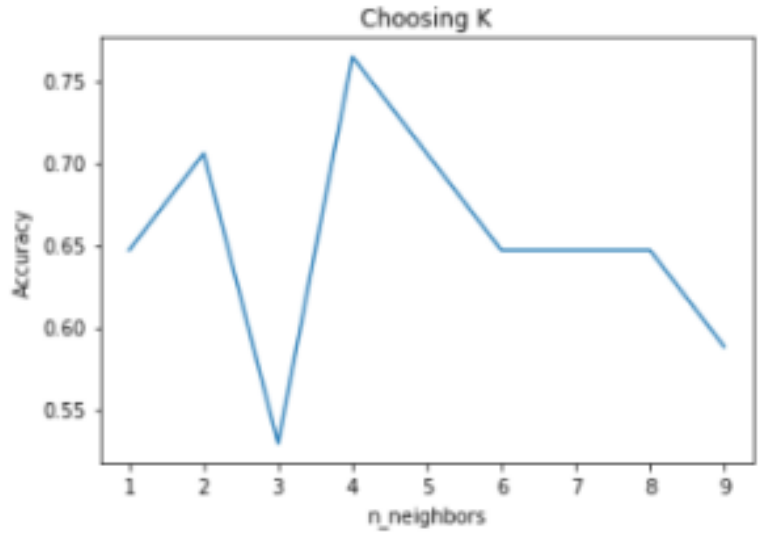


- Relationship between number of dogs and number of dog sitters identifiable
- Relationship between higher incomes and number of dog sitters not so strong

K-means clustering technique

- Optimal accuracy when k = 4
- 82 Toronto neighbourhoods divided into 4 clusters

	FSA	70000_plus_ratio	DOG	Dogsitters_Count	Neighbourhood
59	M6G	0.487944	812	4.0	Christie
61	M6J	0.493375	807	4.0	Little Portugal, Trinity
63	M6M	0.356486	620	0.0	Del Ray, Keelesdale, Mount Dennis, Silverthorn
66	M6R	0.334792	780	1.0	Parkdale, Roncesvalles
69	M8W	0.517753	815	3.0	Alderwood, Long Branch
71	M8Y	0.517052	708	1.0	Humber Bay, King's Mill Park, Kingsway Park So...
81	M9W	0.476506	639	0.0	Northwest

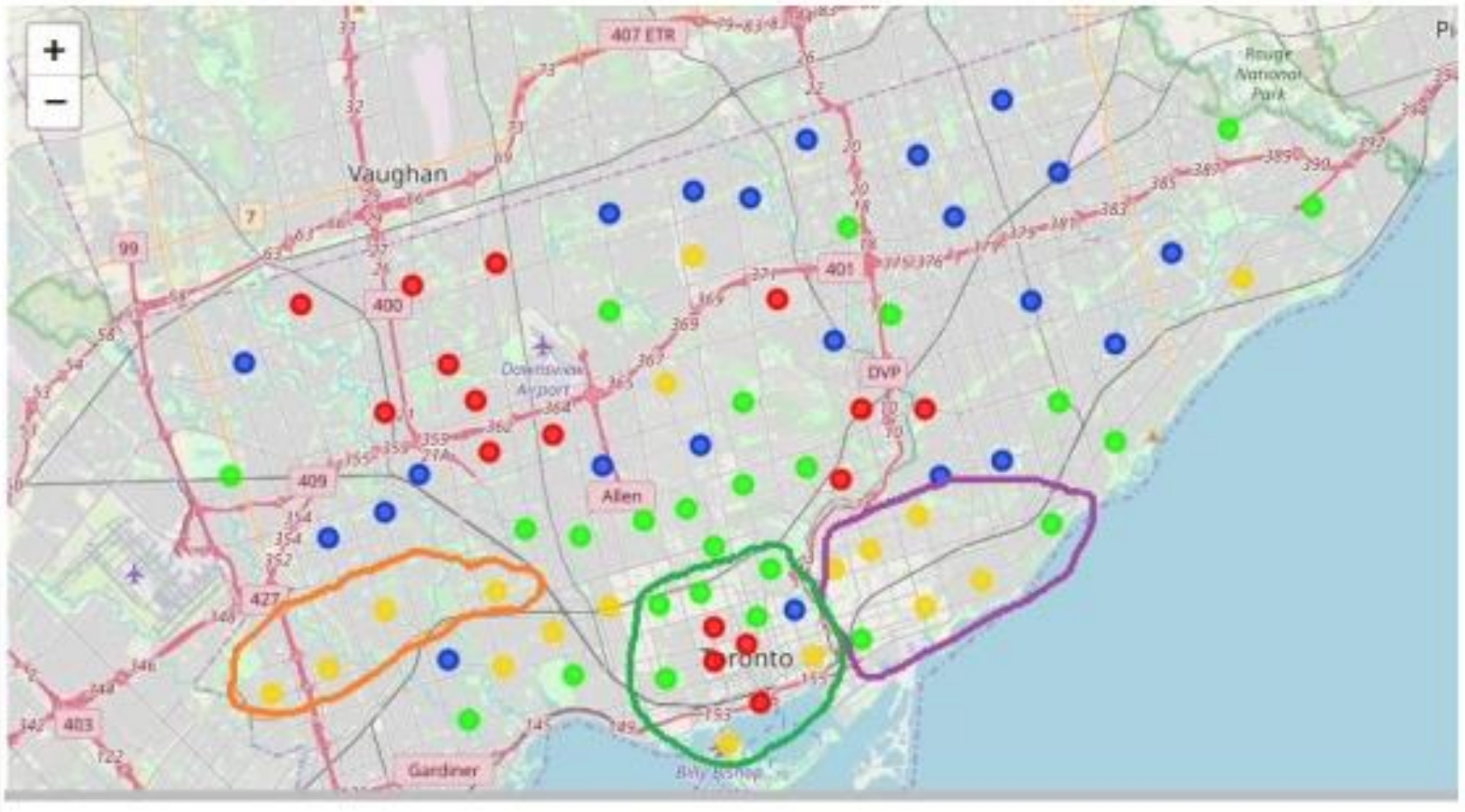


Name the clusters

- Cluster 0 (red): low dog count, just 40% of households earning $70,000 or more, very few dog sitting businesses
- Cluster 1 (green): medium dog count, 50% of households earning $70,000 or more, - a few neighbourhoods with 3 to 4 dog sitting businesses
- Cluster 2(blue): low to medium dog count,45% of households earning $70,000 or more, very few dog sitting businesses
- Cluster 2(yellow): high dog count,50% of households earning 70,000 or more, 2 in 3 neighbourhoods have dog sitting business

(The colours refer to the clustermarker on the map in the next slide)

Visualize the obtained clusters on Toronto map



Consider possible dog hotel locations - 1

- West of Toronto – area around Islington (orange encircled on main Map)

- No competing dog sitting facilities in neigbourhoods
- Proximity to main road in ‘Cloverdale Islington’

Consider possible dog hotel locations - 2

- East York (purple encircled on main Map)

- On average 1 competing dog sitting facility in every neighbourhood.
- Proximity to main road
- Main road is commuter car traffic route from East of Toronto towards City centre

Indications for 2 extra indicators by studying city centre

- Centre of Toronto – 3 neighbourhoods with 4 to 6 dog sitting facilities (green encircled on main Map)

- Proximity to main road: easy access to dog hotel
- Proximity of a commuter car traffic route: easy for dropping off and picking up of dog on road to and from work

Results

- 2 possible areas with several neighbourhoods to start dog hotel
- Initial requirements appear to have drawbacks:
    1. No strong relation between high incomes and presence of dog sitting facility in neighbourhood
    2. High dogcount in neighbourhood maybe not relevant - do customers wish a doghotel near home, near work or elsewhere?
- 2 new possible indicators discovered during analysis:
    1. proximity to main road
    2. proximity to commuter car traffic route

Recommendations

- Do research into the 2 newly discovered indicators
- Examine the spending habits of potential customers – are they willing to spend a considerable amount of their (not so high) income on dogs while saving money on other matters?
- Find out whether potential customers want a doghotel near their homes, their work or elsewhere