

---

# Recommendation System for Airbnb listings

Ioannis Petromichelakis

*Applicant for The Data Incubator fellowship program*



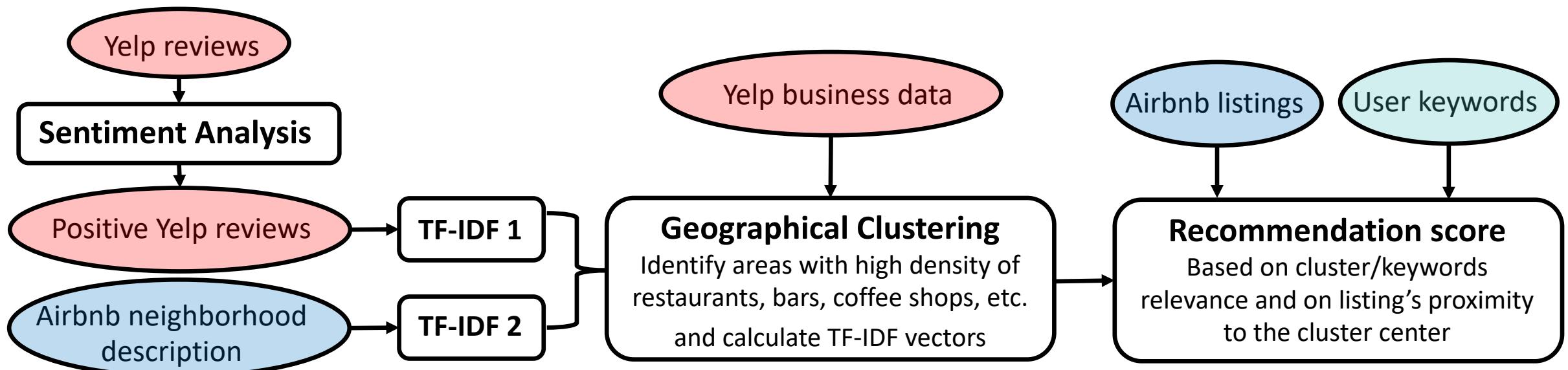
# Outline

## Assumptions and Objective

- The “**style**” of the neighborhood of settlement should much **traveler’s preferences**
- **Restaurants, bars, coffee shops, etc.** are a proxy for the “style” and “vibe”
- Reach those places in a **few steps** and feel **comfortable** spending time there
- Identify neighborhoods that much traveler’s preferences
- Choose Airbnb listings that ideally located within these neighborhoods

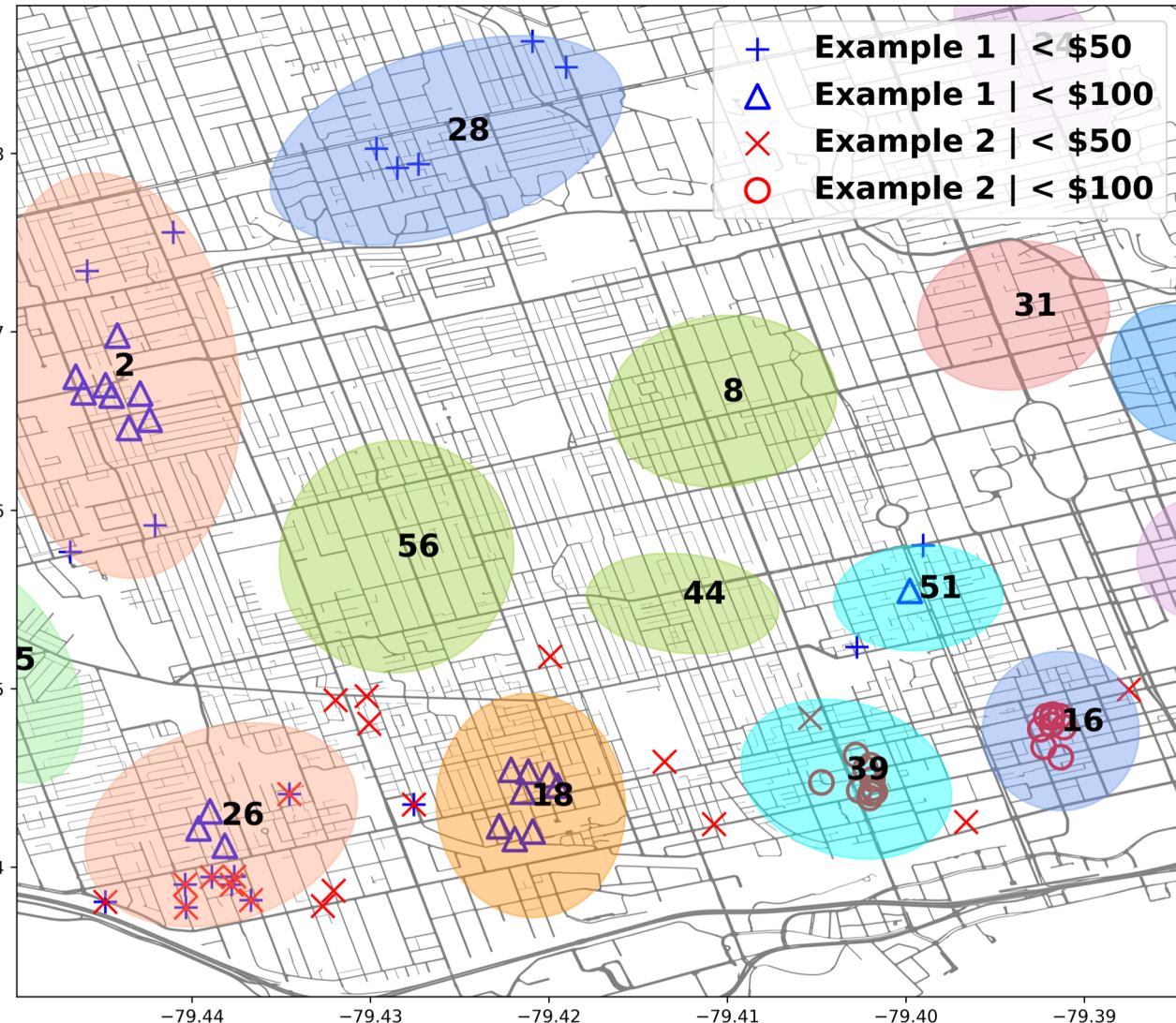
## Data:

- Yelp Challenge Dataset (5.5GB) from [yelp.com](http://yelp.com)
- Airbnb Listings in Toronto and Montreal (200MB) from [insideairbnb.com](http://insideairbnb.com)



# Results - Toronto

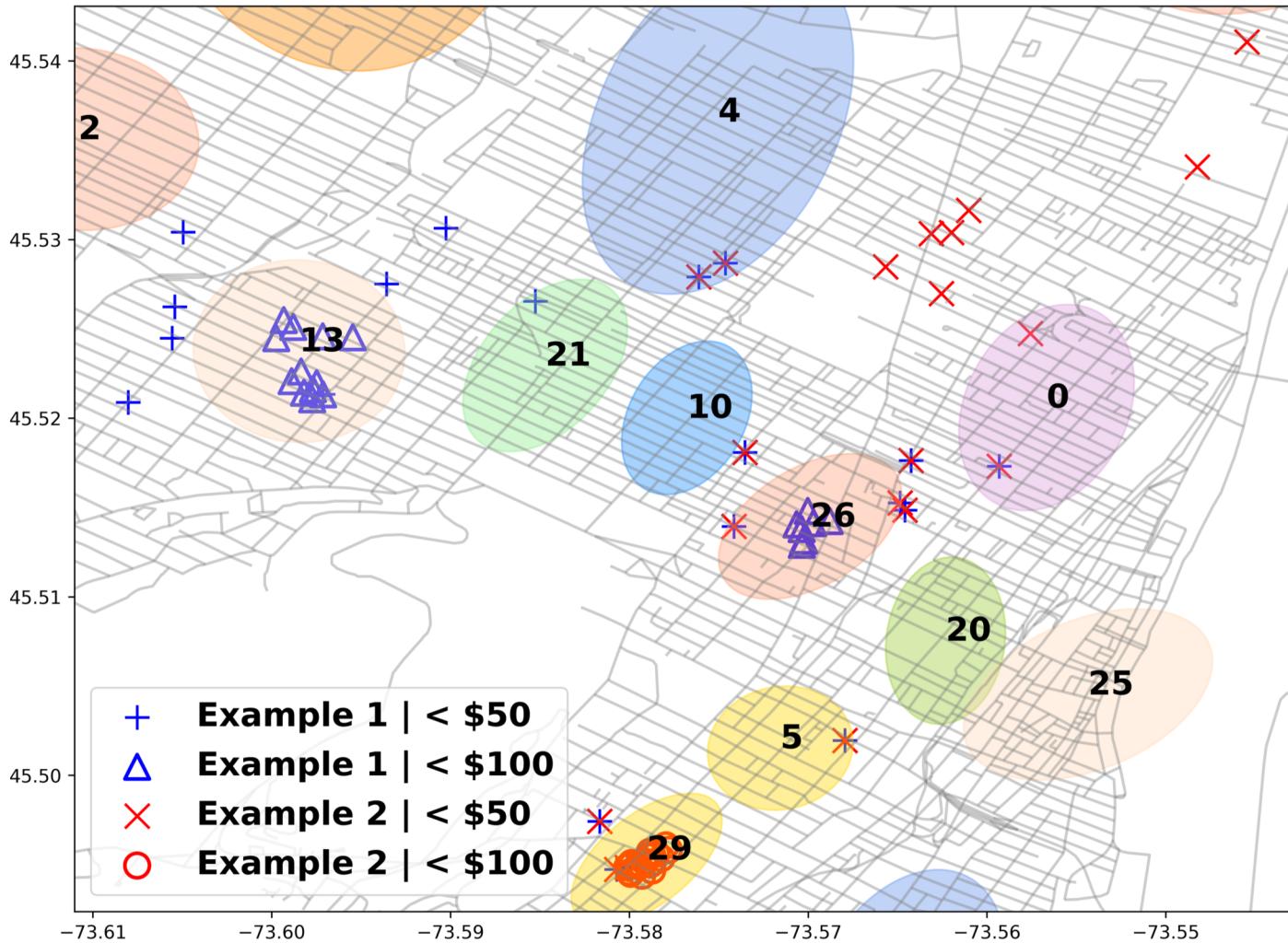
Example 1: ['hipster', 'hippy', 'art']   Example 2: ['clubs', 'nightlife']



Cluster	Scr 1	Scr 2	# listings	Av. Price	Neighborhoods
2	<a href="#">6.6</a>	<a href="#">2.3</a>	279	\$101.33	Dovercourt-Junction 69%
16	<a href="#">2.8</a>	<a href="#">5.0</a>	794	\$204.44	Corso Italia-Davenport 21%
18	<a href="#">6.8</a>	<a href="#">3.8</a>	487	\$177.35	Waterfront Communities 94%
					Little Portugal 37%
					Niagara 33%
					Trinity-Bellwoods 19%
26	<a href="#">6.1</a>	<a href="#">4.2</a>	319	\$125.96	Roncesvalles 40%
					South Parkdale 38%
					Little Portugal 18%
28	<a href="#">4.1</a>	<a href="#">1.6</a>	125	\$119.88	Wychwood 53%
					Humewood-Cedarvale 22%
39	<a href="#">3.5</a>	<a href="#">5.1</a>	769	\$199.43	Waterfront Communities 68%
					Niagara 20%
44	<a href="#">3.1</a>	<a href="#">3.7</a>	174	\$141.52	Trinity-Bellwoods 53%
					Palmerston-Little Italy 38%
51	<a href="#">6.4</a>	<a href="#">1.9</a>	250	\$122.03	Kensington-Chinatown 98%

# Results - Montreal

Example 1: ['hipster', 'hippy', 'art']   Example 2: ['clubs', 'nightlife']



Cluster	Scr 1	Scr 2	# listings	Av. Price	Neighborhoods
0	4.5	<u>7.4</u>	787	\$109.20	Ville-Marie 100%
5	5.8	<u>8.8</u>	325	\$186.20	Ville-Marie 76%
13	<u>10.3</u>	5.3	603	\$122.17	Le Plateau 24%
20	<u>8.0</u>	6.4	453	\$158.07	Le Plateau 77%
21	<u>6.6</u>	5.5	691	\$115.44	Outremont 14%
25	5.3	<u>7.0</u>	420	\$187.64	Ville-Marie 98%
26	<u>8.3</u>	<u>8.2</u>	823	\$159.81	Le Plateau 70%
29	<u>6.4</u>	<u>9.3</u>	728	\$125.02	Ville-Marie 30%
					Ville-Marie 99%

# Future directions

---

- Improve the Recommendation System
  - More sophisticated word embedding
  - Capture trends via temporal information
- Extend the System by utilizing more data
  - Include many cities by scrapping Yelp, Airbnb and possibly Instagram
  - Identify activities that are correlated with specific neighborhoods
- Deploy a Destination Recommendation component

Thank you!

# Outline

