FOOD RECOMMENDATION SYSTEM

DSI 16 Capstone Presentation

Dominic Ong

SINGAPORE

Restaurant no-shows on the rise as Singapore's F&B scene struggles through the pandemic

BY MAY SEAH | 29 JUL 2020 |

Many diners are making several bookings per night, choosing one and cancelling the rest at the last minute, say restaurants and chefs who are at the end of their tether.

Problem Statement

- Earlier this year, COVID-19 forced many countries into lockdowns to curb the spread of the virus.
- While staying at home, many people turned to online food delivery services for their daily meals.
- Many lesser known restaurants businesses have struggled during these times due to the lack of customers and orders.

Solution

- More people are starting to frequent restaurants again in Phase 2.
- Build an alternative restaurant recommendation system based on user reviews and sentiment analysis.
- Make people more aware of local restaurants and help keep their businesses alive.

Overview

Explanatory Data Analysis

- Pandas
- Seaborn
- Plotly

Natural Language Processing

- Sentiment Analysis
- Topic Modelling

Recommendation System

- Location-Based
- Collaborative-Filtering
- Content-Based

Yelp Toronto Dataset

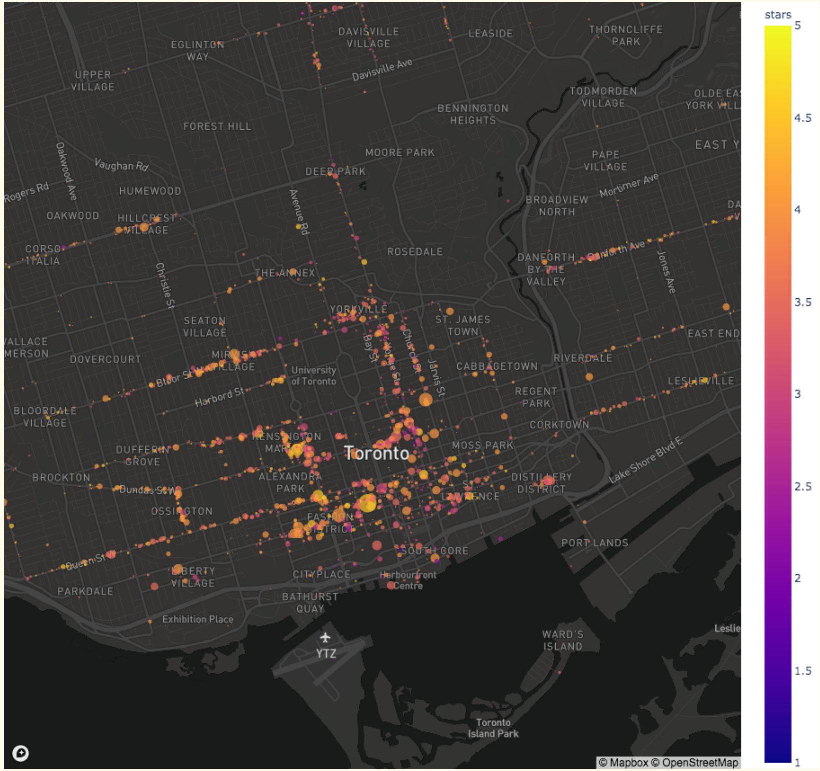
The Yelp dataset is a subset of Yelp's businesses, reviews, and user data that has been made publicly available for use for personal, educational, and academic purposes.

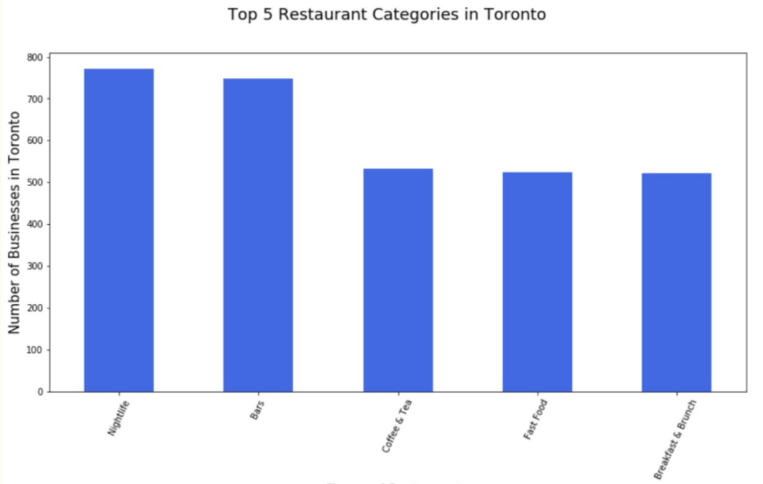
Initial: 44,023 Restaurants
Toronto: 5,471 Restaurants

Initial: 8,021,122 Reviews
Toronto: 253,050 Reviews

Initial: 1,047,892 Users
Toronto: 44,485 Users

EDA on Restaurant Data

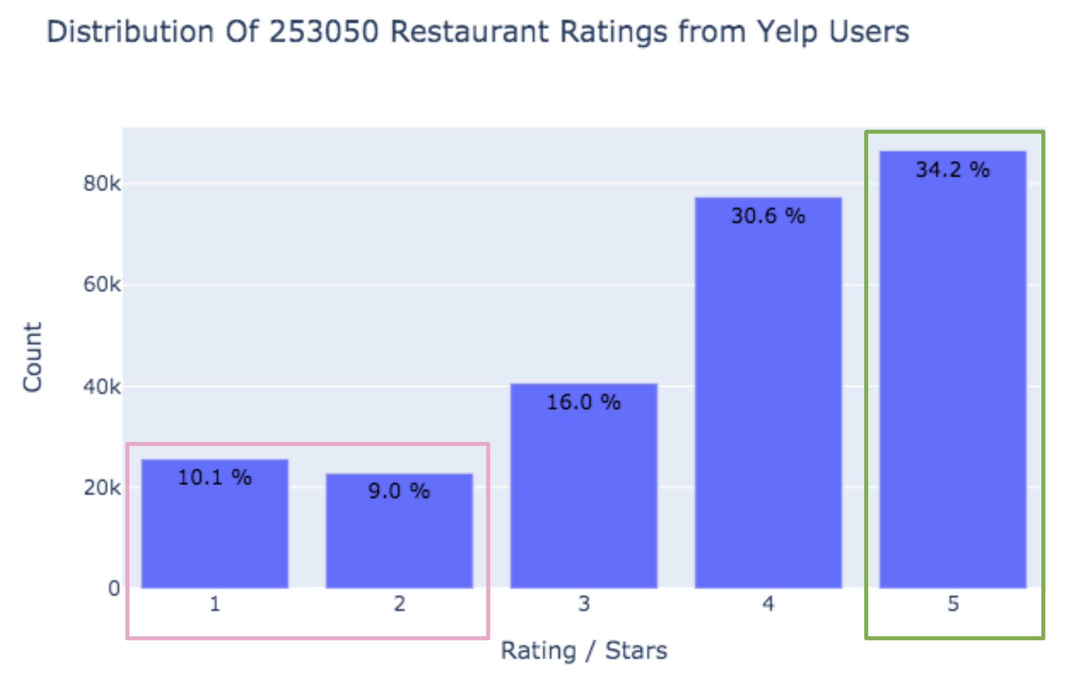




Top 5 Most Popular Restaurant Categories:

- **Nightlife** (772 Restaurants)
- **Bars** (748 Restaurants)
- **Coffee & Tea** (532 Restaurants)
- **Fast Food** (524 Restaurants)
- **Breakfast & Brunch** (522 Restaurants)

EDA on Reviews Data



86,543 Positive Reviews

- **Pai Northern Thai Kitchen**
- **Most # of Positive Reviews (2,758)**

48,332 Negative Reviews

- **Momofuku Noodle Bar**
- **Most # of Negative Reviews (1,010)**

EDA on Reviews Data

Negative Reviews

Positive Reviews

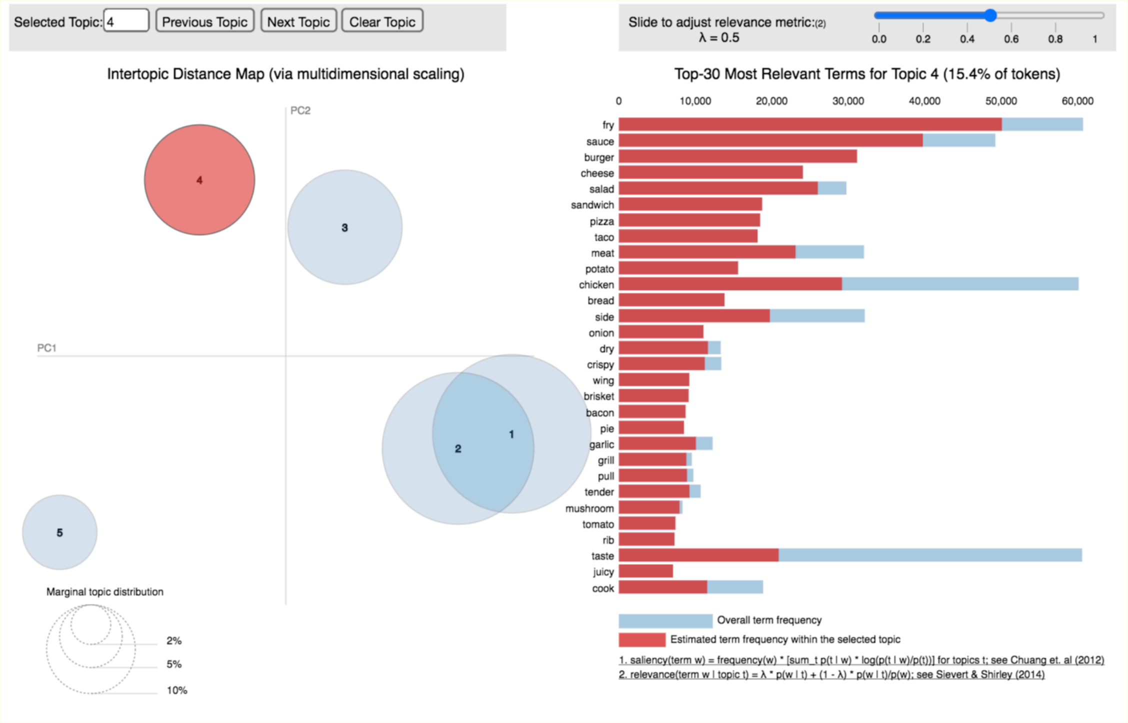
Sentiment Analysis

Yelp Rating (1-5) + Polarity Score [-1, 1] × Compound Score [-1, 1] = Super Score (0-6)

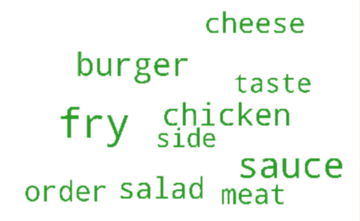
Polarity Score and Compound Score: Sentiment Analysis on Text Reviews

- Polarity Score: positive statement = 1, negative statement = -1
- Compound score: most positive = 1, most negative = -1

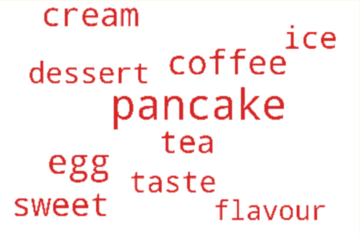
Topic Modelling



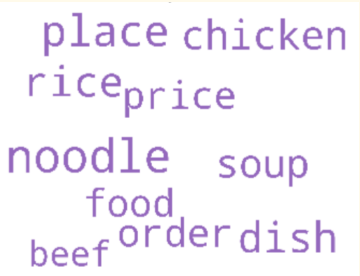
Western (Topic 3)



Dessert (Topic 4)



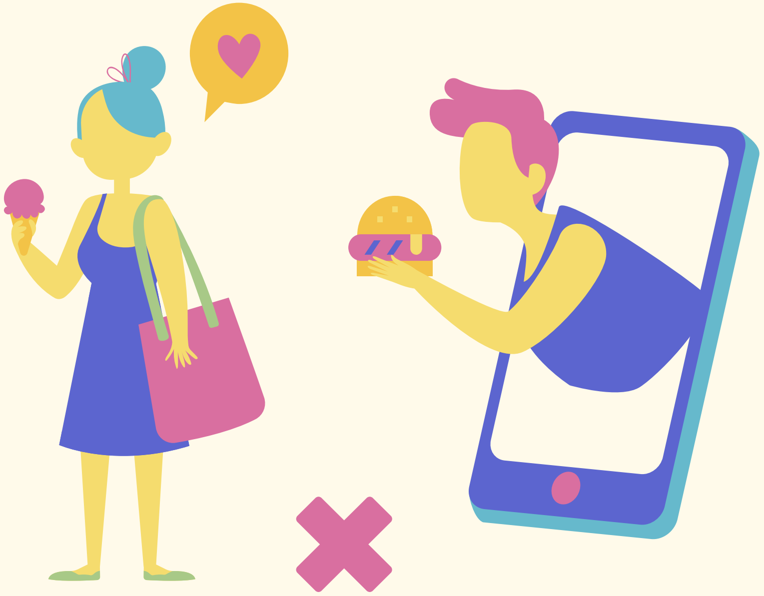
Asian (Topic 5)

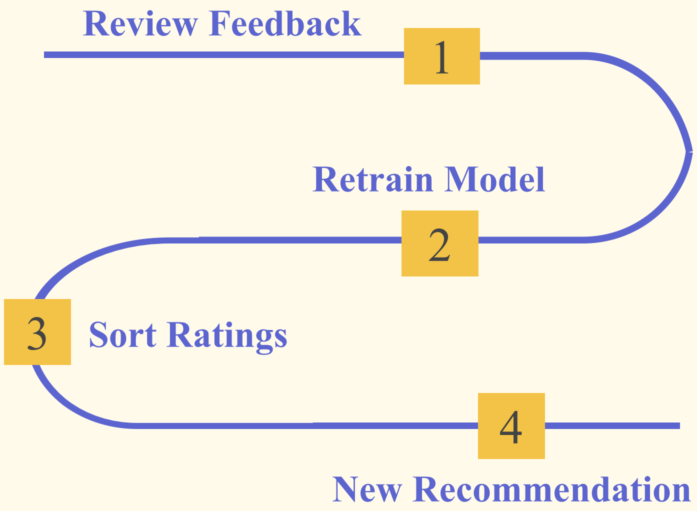


Building Our Recommendation Systems

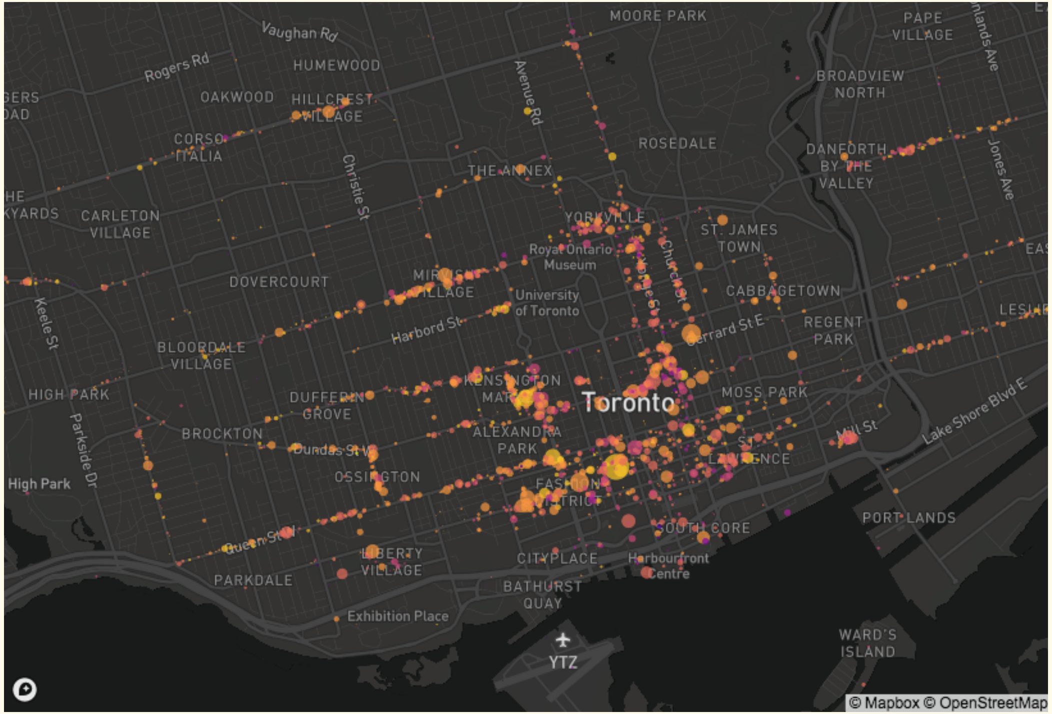
Cold-Start Problem

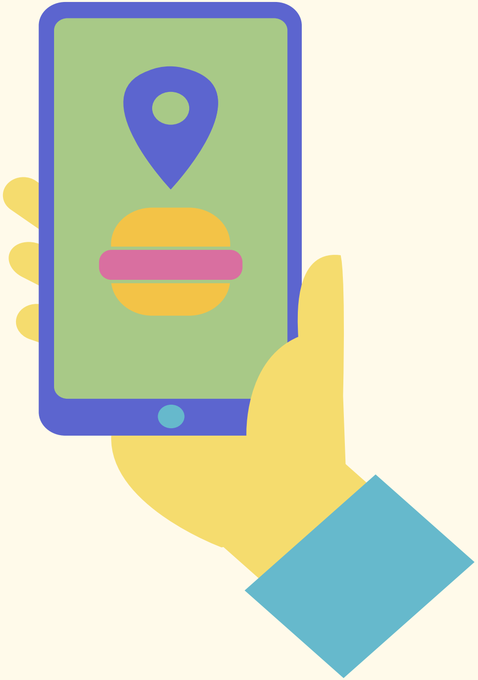
It is difficult to provide personalized recommendations to new users without any information about their prior preferences. We have to use a naïve solution to solve the cold-start problem.





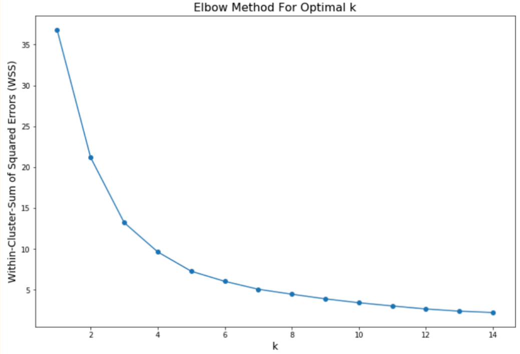
Location-Based Recommendation

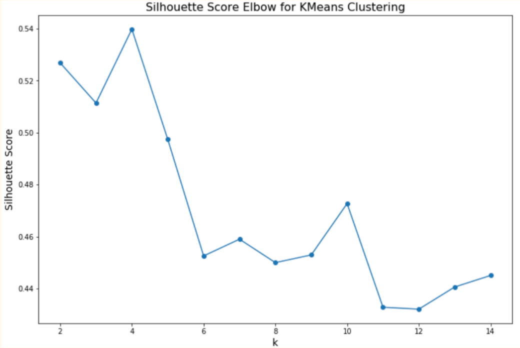


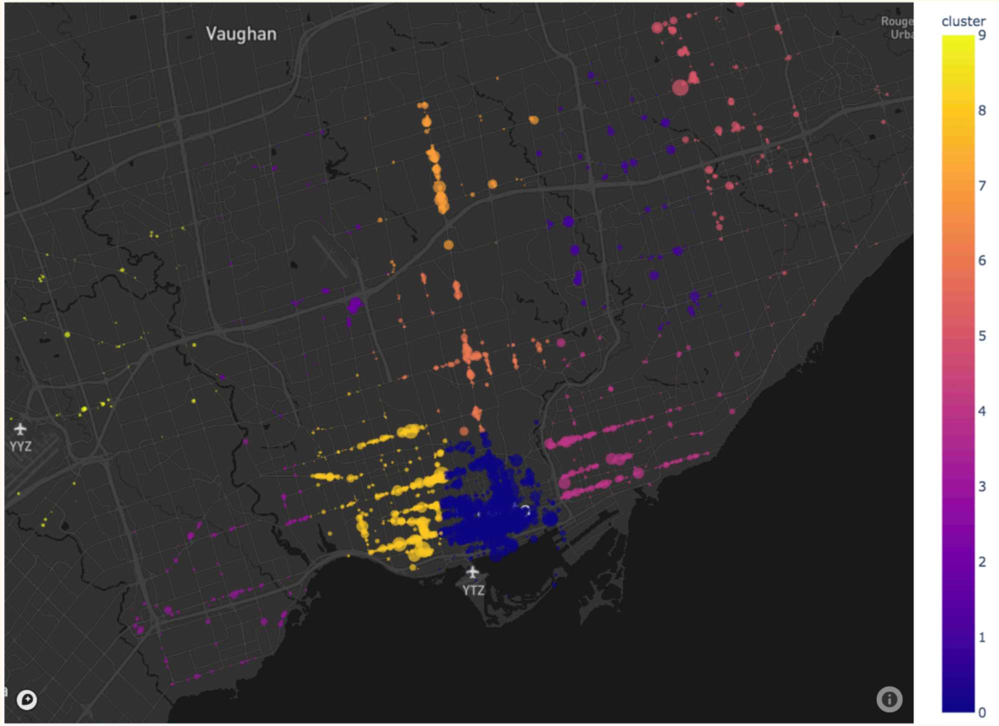


Location-Based Recommendation

Group Restaurants together based on geographical proximity using K-Means Clustering.



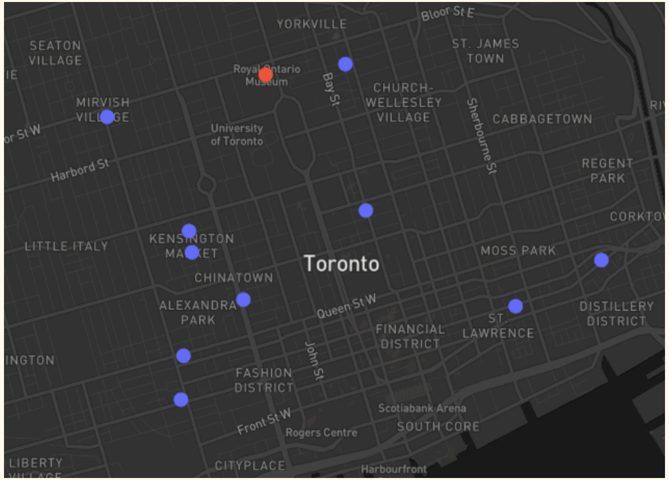




Location-Based Recommendation

Case Study 1: Royal Ontario Museum

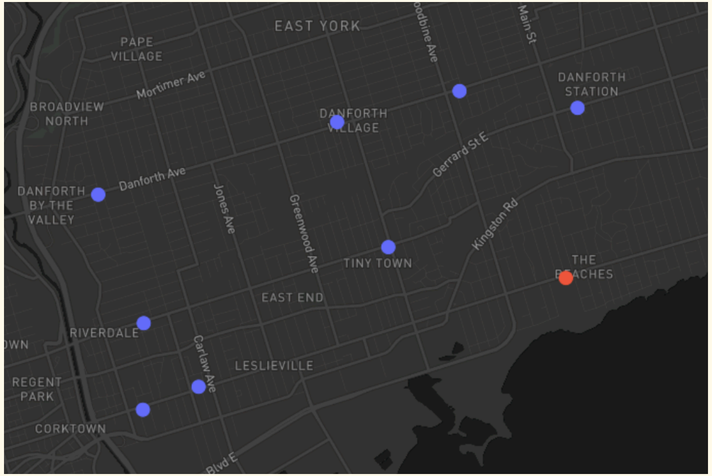
Shawn is a tourist in Toronto. After visiting the Royal Ontario Museum, he wants to have a meal but is unsure of what nice restaurants are around his location.



name	latitude	longitude	categories	stars	review_count	cluster
Big Trouble Pizza	43.651328	-79.397038	Pizza, Restaurants	5.0	48	4
Mallo	43.664631	-79.410738	Coffee & Tea, Food, Cafes, Nightlife, Sandwich...	5.0	36	4
Ooshee Mediterranean Oven	43.668469	-79.386778	Restaurants, Mediterranean, Bakeries, Lebanese...	5.0	20	4
The Vegan Extremist	43.656321	-79.402501	Indian, Vegan, Restaurants	5.0	15	4
Muncheez	43.644059	-79.403310	Juice Bars & Smoothies, Restaurants, Creperies...	5.0	12	4
Strange Love Coffee	43.647247	-79.403050	Coffee & Tea Supplies, Coffee Roasteries, Food...	5.0	11	4
Fast Fresh Foods	43.657823	-79.384715	Restaurants, Salad, Sandwiches, Food	5.0	8	4
Casamiento	43.654775	-79.402205	Latin American, Pop-up Shops, Salvadoran, Shop...	5.0	8	4
Old Town Bodega	43.654212	-79.361035	Cafes, Restaurants, Food, Coffee & Tea	5.0	8	4
YEE: Gettin' Cozy with Butter Chicken Roti	43.650860	-79.369668	Local Flavor, Restaurants, Indian, Yelp Events	5.0	8	4

Location-Based Recommendation

Case Study 2: Kew Gardens



Jeremy has just explored Kew Gardens. After a tiring walk, he is looking for a restaurant to sit down and have a relaxing meal.

name	latitude	longitude	categories	stars	review_count	cluster
I'll Be Seeing You	43.658993	-79.348074	Cocktail Bars, Bars, Comfort Food, Wine Bars, ...	5.0	49	6
Viva Shawarma	43.708528	-79.295665	Halal, Middle Eastern, Falafel, Mediterranean,...	5.0	30	6
Pomarosa Coffee Shop & Kitchen	43.683242	-79.325450	Cafes, Coffee & Tea, Restaurants, Sandwiches, ...	5.0	29	6
Elvy & Flo Cafe	43.666286	-79.347986	Restaurants, Coffee & Tea, Food, Cafes	5.0	15	6
Papyrus	43.677126	-79.353285	Vegan, Vegetarian, Restaurants, Middle Eastern...	5.0	11	6
The Dock On Queen	43.660927	-79.341583	Cafes, Sandwiches, Restaurants	5.0	10	6
YEE: The Breakfast Club: Lazy Daisy's Café Bis...	43.672696	-79.319486	Yelp Events, Restaurants, Local Flavor, Breakf...	5.0	10	6
JJ Bean	43.719810	-79.313194	Coffee & Tea, Cafes, Bakeries, Coffee Roasteri...	5.0	10	6
Gyoko Sushi & Bar	43.685862	-79.311198	Sushi Bars, Restaurants	5.0	9	6
Prologue Cafe	43.684435	-79.297439	Coffee & Tea, Restaurants, Cafes, Food	5.0	8	6

Collaborative-Filtering Recommendation

Collaborative Filtering finds a smaller set of users with tastes similar to a particular user. It looks at the restaurants they like and combines them to create a ranked list of suggestions.

Steps:

1. Pivot Table on Super Score Ratings
2. Truncated Singular Value Decomposition
3. Item-Item Matrix based on Cosine Similarity

Collaborative-Filtering Recommendation

Case Study: Pai Northern Thai Kitchen

- Most # of Yelp Reviews
- Thai Cuisine

1. Montchant (Café)
   Correlation: 0.975

2. Rainbow Food (Chinese)
   Correlation: 0.972

3. Masala King (Indian)
   Correlation: 0.942

Content-Based Recommendation

Content-Based Recommendation recommends restaurants based on similar categories and keywords.

Steps:

1. Topic Modelling
2. CountVectorizer
3. Cosine Similarity

Content-Based Recommendation

Case Study: Sidewalk Pizzeria

- Italian Cuisine
- Keywords: Italian, restaurants, pizza, great, place, food, …

1. Fusilli Restaurant

2. Tavolino

Limitations and Looking Forward

- More Reviews and Data are required to improve the Recommendation Systems.
- Use Bi-Grams and Tri-Grams in Sentiment Analysis and Topic Modelling.
- Deep Learning & Neural Network architectures for collaborative filtering can be explained in the future.
- Incorporate Graph Theory for location-based systems to optimize travelling and delivery routes.

Demo (in the Future)

When I figure out how to deploy on Heroku…

Thanks!

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**.

Please keep this slide for attribution.