

CAPSTONE PROJECT

SG Food Recommender

Prepared for: **Yelp Management**

Presenter: Choo Wende

Date: 08 Sep 2022

TABLE OF CONTENTS

01

Introduction

Background
& Problem Statement

02

Data Analysis

Key Findings
& Topic Modelling

03

Recommendation System

Methodology
& Evaluation

04

Recommendation & Conclusion

Recommendation
& Future Works

01

Introduction

Background
& Problem Statement



"Where you want to eat ah?"



Background

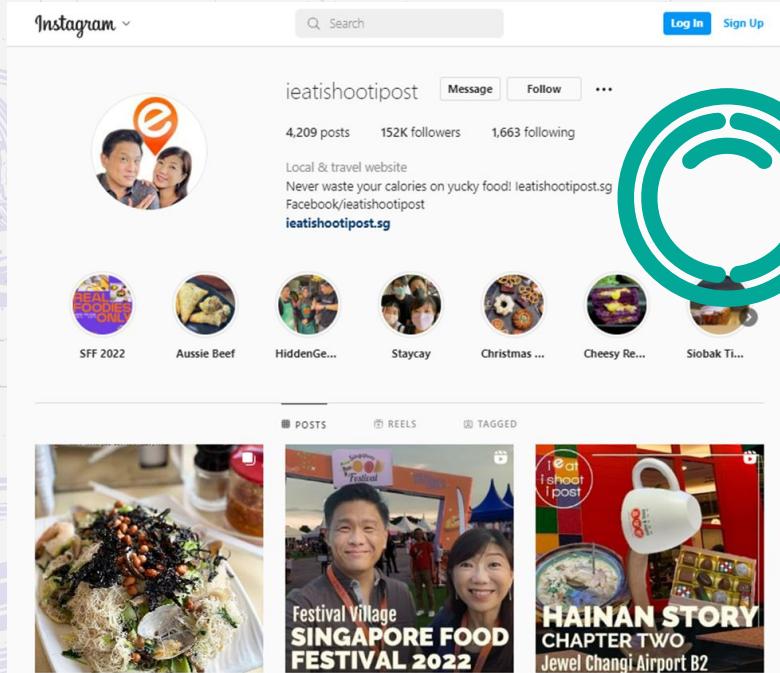


E-COMMERCE

35%

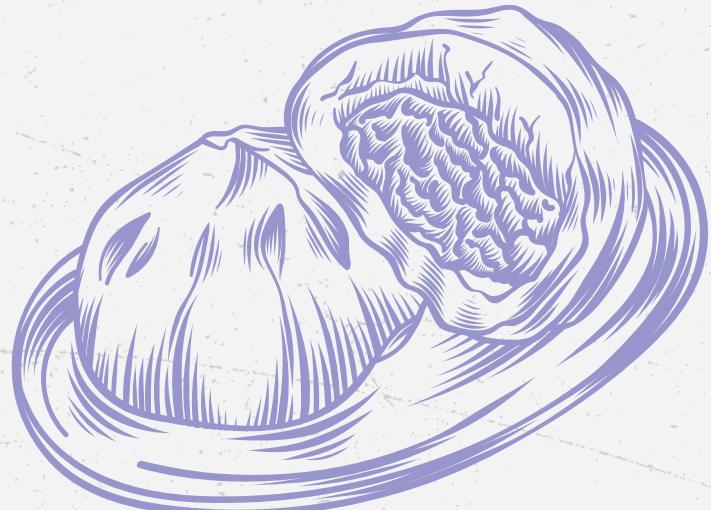
Increased in sales through
recommended items by
personalized
recommenders

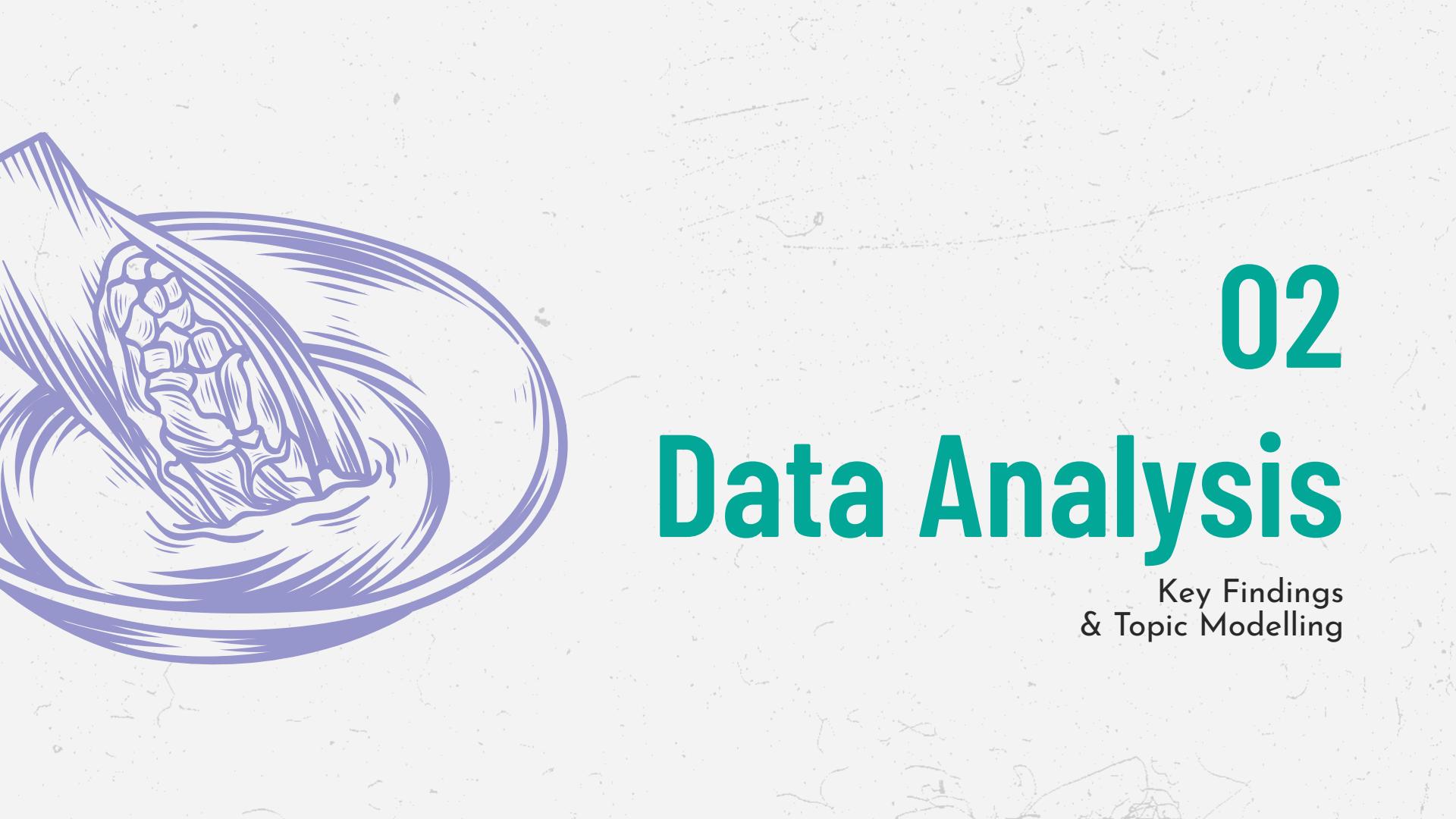
Background



Problem Statement

Build a SG Food Recommender
using user's reviews and ratings
from Yelp website





02

Data Analysis

Key Findings
& Topic Modelling

Webscraping



Restaurants Info

rest_name, href, address, postal_code,
latitude, longitude, opening, star_rating,
review_count, location



Users Info

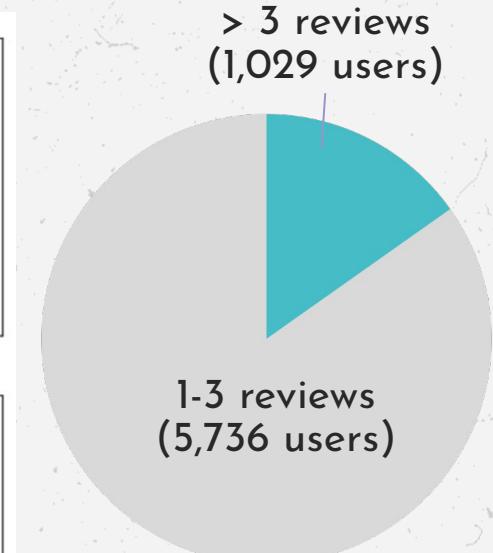
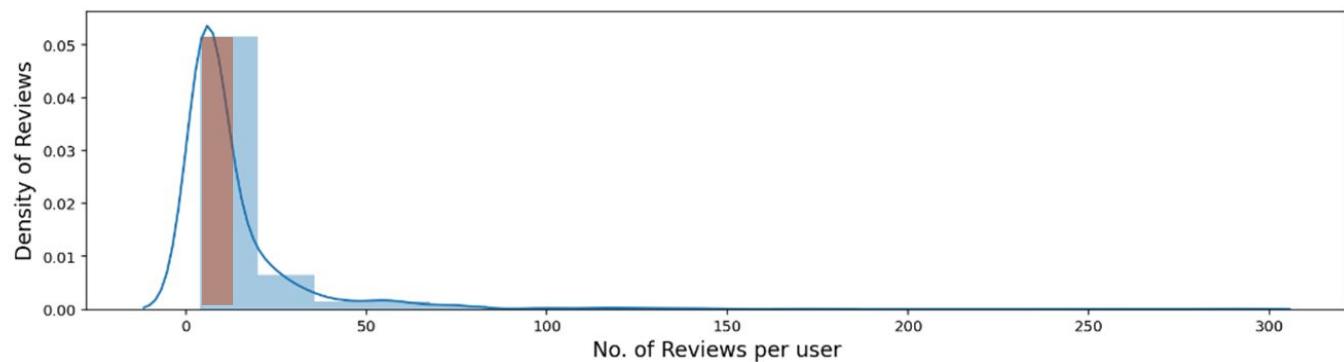
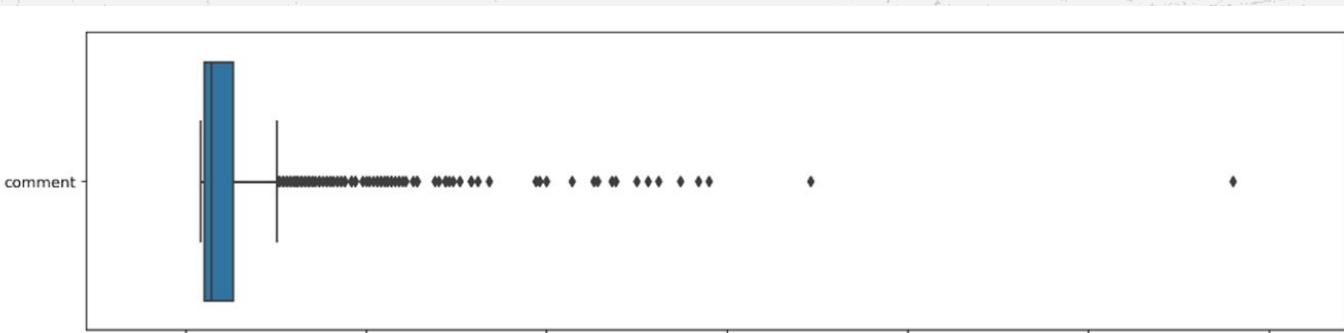
username, userid, user_location



Reviews Info

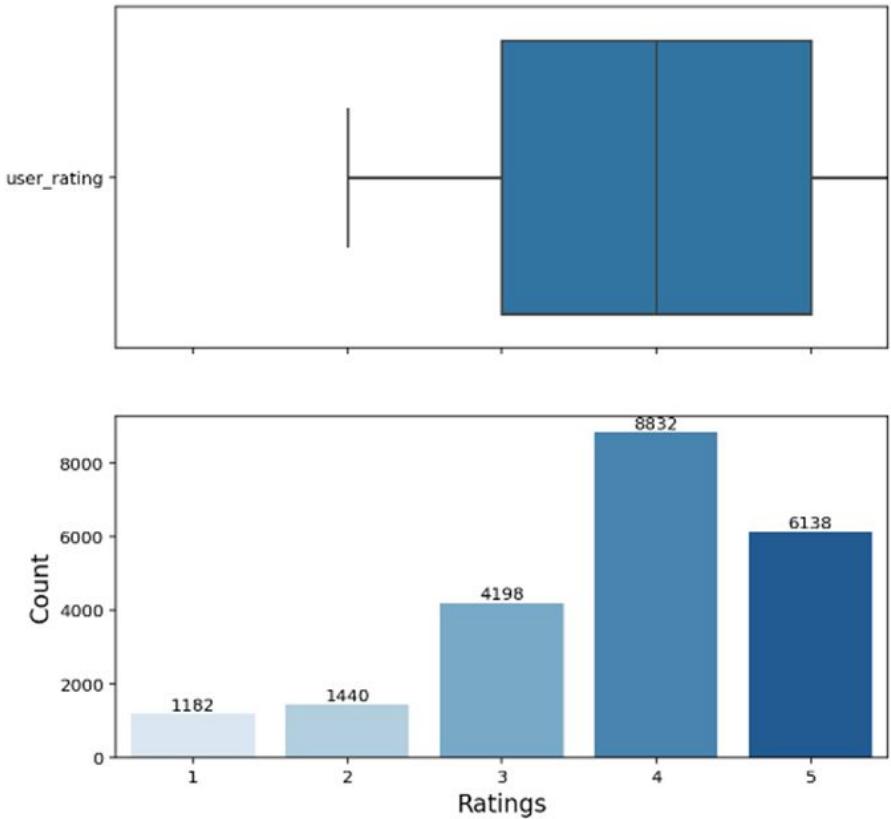
user_rating, review_date, comment

Majority of the Reviews Per User (with >3 Reviews) Ranges from 5 to 13



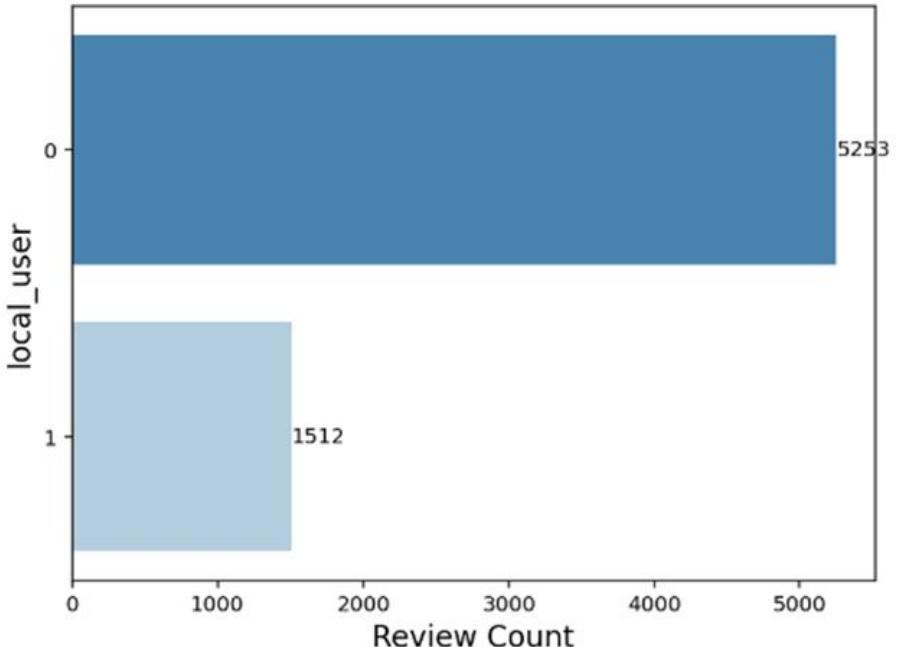
Rating Distribution

Average rating is 3.8

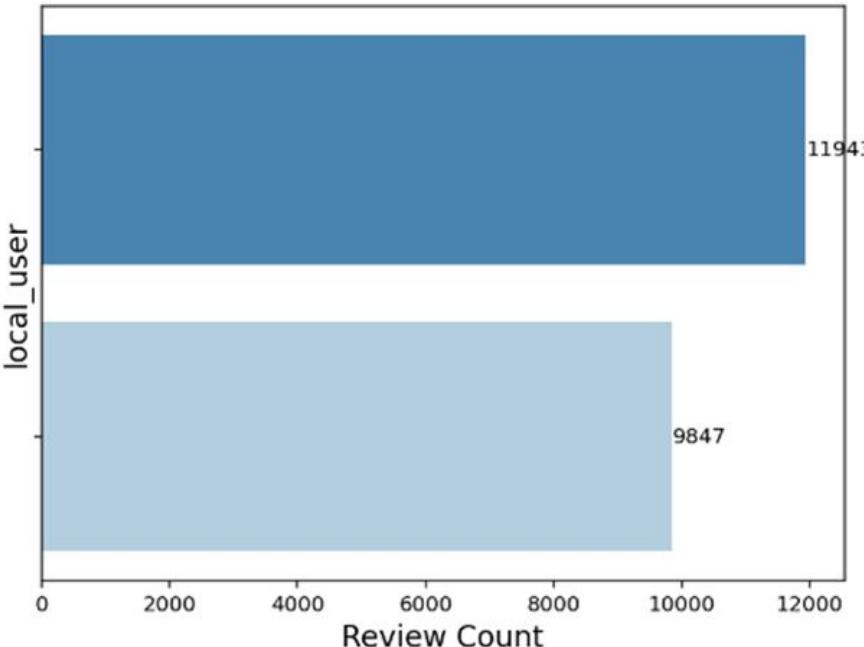


Locals submit more reviews than overseas users

Five Times More Overseas Users Reviews than Locals

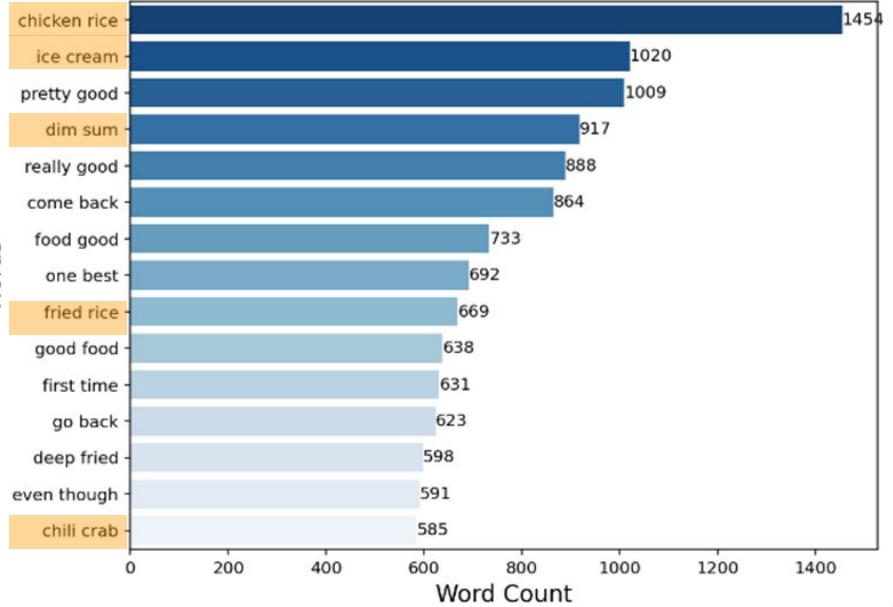


More Overseas Users Reviews than Locals'

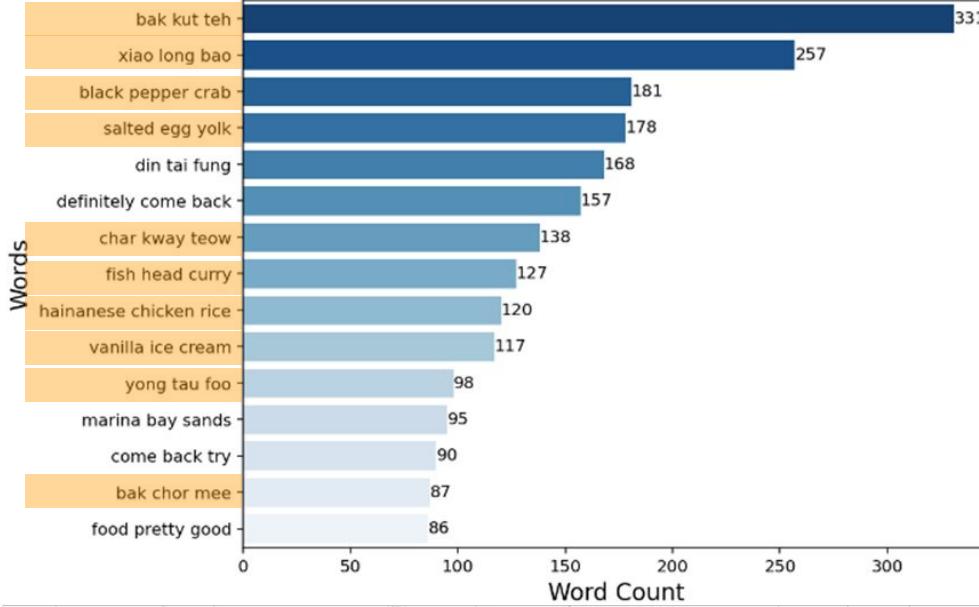


Top Bi-grams & Tri-grams

Top 15 Bigrams

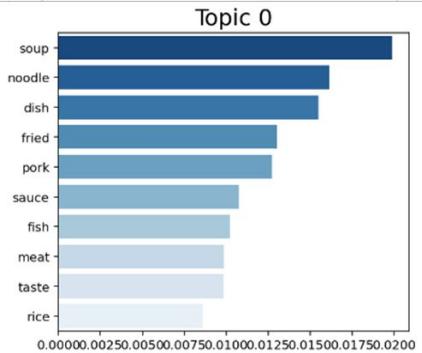


Top 15 Trigrams

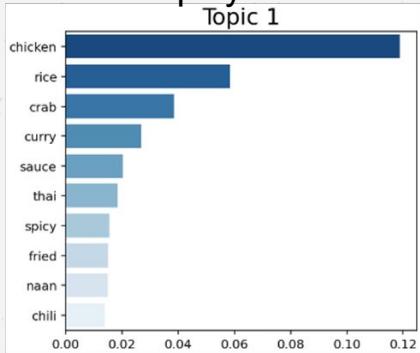


Topic Modelling

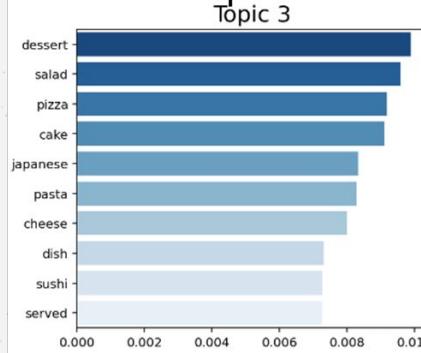
Chinese Food



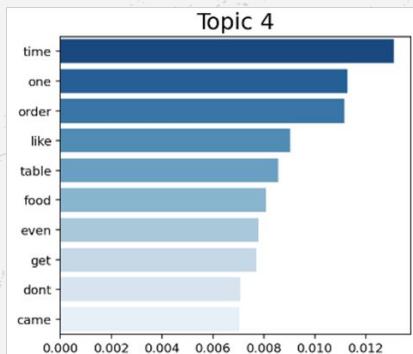
Spicy Food



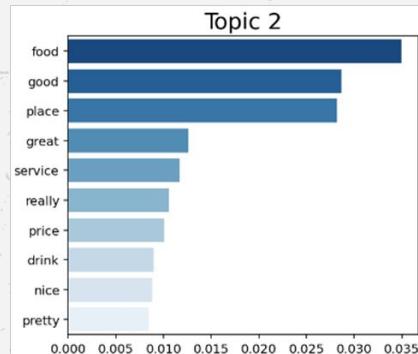
Western/Japanese Food



Ambience



Service



Intertopic Distance Map (via multidimensional scaling)



03

Recommendation System

Methodology
& Evaluation

SG Food Recommender

Stage 1: Location-based Filtering

- 1 Kmeansmodel Filtering
Restaurant clusters
- Filter top 10 restaurants around user



Filter set of restaurants

Stage 2: Collaborative Filtering

Memory Based Filtering

- 2 User-based Filtering
- Restaurants visited by the similar users
- 3 Item-based Filtering
- Restaurants similar to those visited by user.

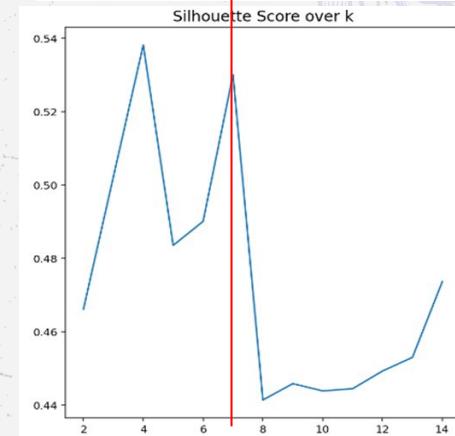
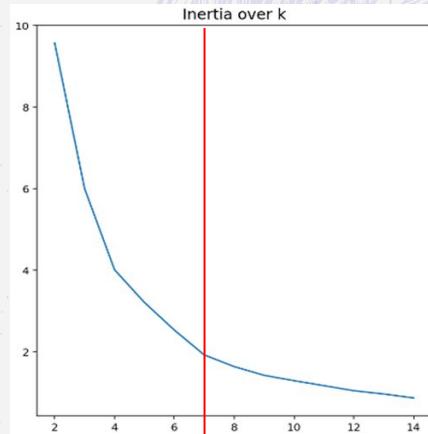
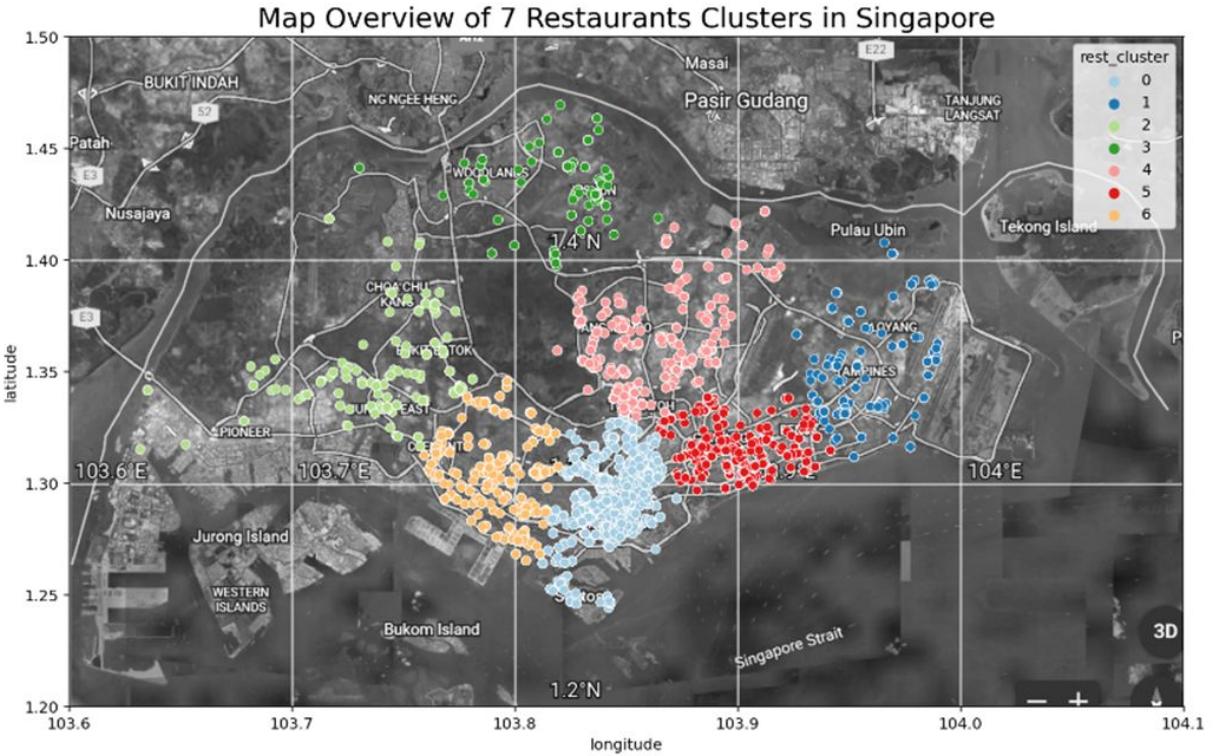


- 4 Model-based Prediction
- Predict user ratings on restaurants
- Recommend top rated restaurants



Optimal k is 7
(Low Inertia and High Silhouette Score)

1. Location-based Filtering



2. User-based Filtering



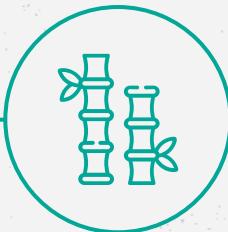
User-item matrix

Based on user ratings, fill NA with zero



User similarity matrix

Apply cosine similarity



Top 10 similar users

Based on user similarity matrix



Filter Restaurants

Get restaurants with more than 3.8 rating from similar users

3. Item-based Filtering



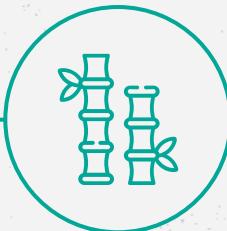
**Item-
features DF**

Based on price
range, review
count, restaurant
cluster, restaurant
categories



**Item similarity
matrix**

Apply cosine
similarity



**Restaurants
visited**

Based on reviews
submitted by user



**Filter
Restaurants**

Get top 10 similar
restaurants for each
visited restaurants

Model-based Prediction



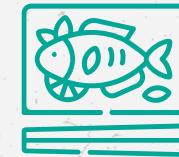
Basic Models

1. NormalPredictor
2. BaselineOnly



KNNS

1. Basic
2. With Means
3. With Z Score
4. Baseline



Matrix Factorisation

1. Singular Value Decomposition (SVD)

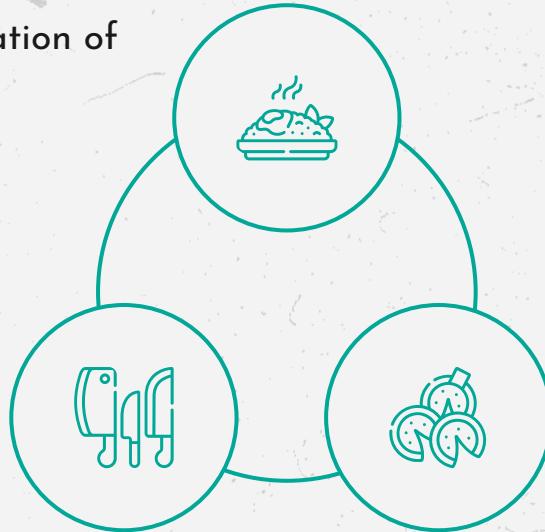
Evaluation Metric

RMSE

standard deviation of
the residuals

Precision@k

proportion of
recommended items that
are relevant to the number
of recommended items



Recall@k

proportion of
recommended items that
are relevant to the number
of relevant items.

Model Summary

	model_name	train_cross_val_rmse	train_rmse	test_rmse	average_model_precision	average_model_recall
6	svd	🏆	0.896	0.748	0.894	0.581
1	baselineonly		0.898	0.81	0.896	0.59
5	knn_baseline		0.908	0.49	0.916	0.594
2	knn_basic		0.95	0.552	0.953	0.168
3	knn_means		0.971	0.551	0.964	0.467
4	knn_zscore		0.971	0.557	0.966	0.466
0	normal_predictor		1.278	1.292	1.275	0.55

SVD

n_factors
(latent factors)

n_factors
(latent factors)

	Items		
Users			

Case Studies - 3 Users at GA



Kevin W.

Singapore, Singapore
Visited 7 restaurants

noodles | japanese | singaporean |
asian_fusion | cocktail_bars

1. Man Man Japanese Unagi
2. Two Men Bagel House
3. Rhubarb Le Restaurant
4. Ah Tai Hainanese Chicken Rice
5. The Coconut Club



Luciana Y.

Singapore, Singapore
Visited 8 restaurants

french |
breakfast_&_brunch |
pizza | japanese |
coffee_&_tea

1. Man Man Japanese Unagi
2. Rhubarb Le Restaurant
3. Two Men Bagel House
4. Trattoria Nonna Lina
5. Ramen Keisuke Tori King



Clarissa L.

Basel, Switzerland
Visited 35 restaurants

chinese | italian | thai |
coffee_&_tea | filipino

1. Man Man Japanese Unagi
2. Rhubarb Le Restaurant
3. The Coconut Club
4. Ah Tai Hainanese Chicken Rice
5. Terra Tokyo Italian

04

Recommendation & Conclusion

Recommendation
& Future Works

Conclusion & Recommendations



Singaporean Food

- EDA identified food unique to Singapore
- Overseas users are 5 times more than locals
- But locals post about 2 times more than overseas users.
- Market Recommender to overseas travellers



Recommendation System

- successfully developed a SG Food Recommender
- provided relevant recommendations



Future Works

- Build database
- Implicit Data
- Content of Users' Comment
- Neural NEtwork and Deep Learning

THANK YOU!

CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#) and infographics & images by [Freepik](#)

Please keep this slide for attribution