SINGAPORE RESTAURANT RECOMMENDER SYSTEM

Kun Won, DSI-21

Table of Contents

Why Restaurants?

~23,000

Food retail outlets in Singapore

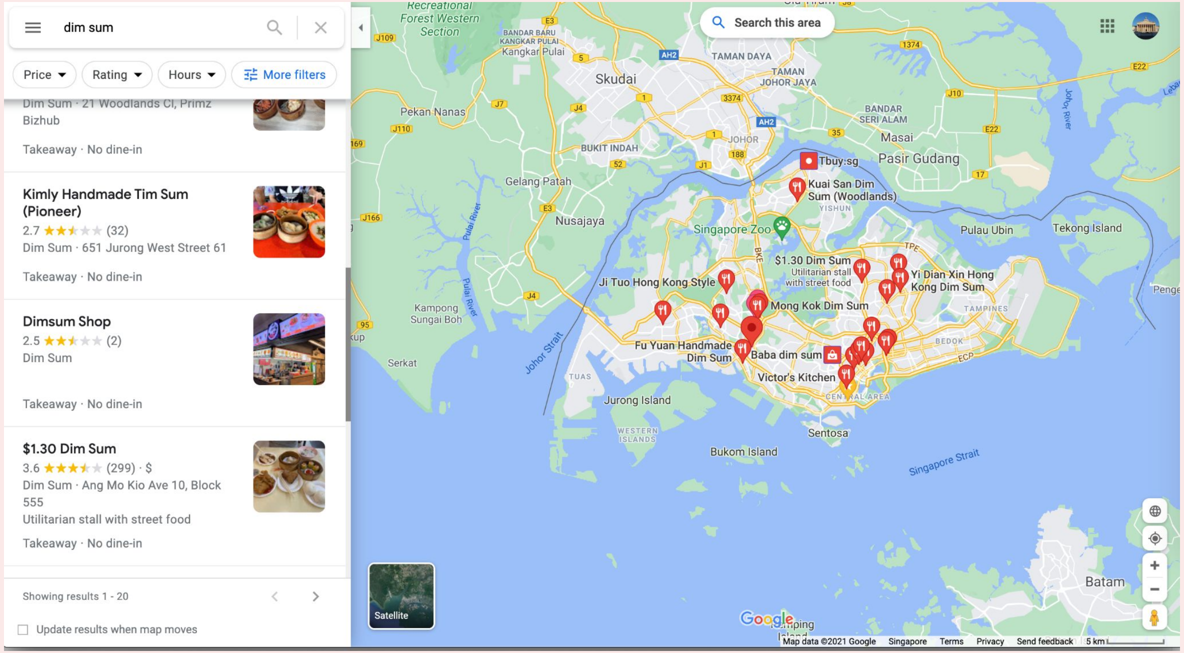
S$8.3bil

SG F&B Industry Net Worth

Consumers face choice overload

Reliance on word-of-mouth, brick and mortar advertising, food blogs

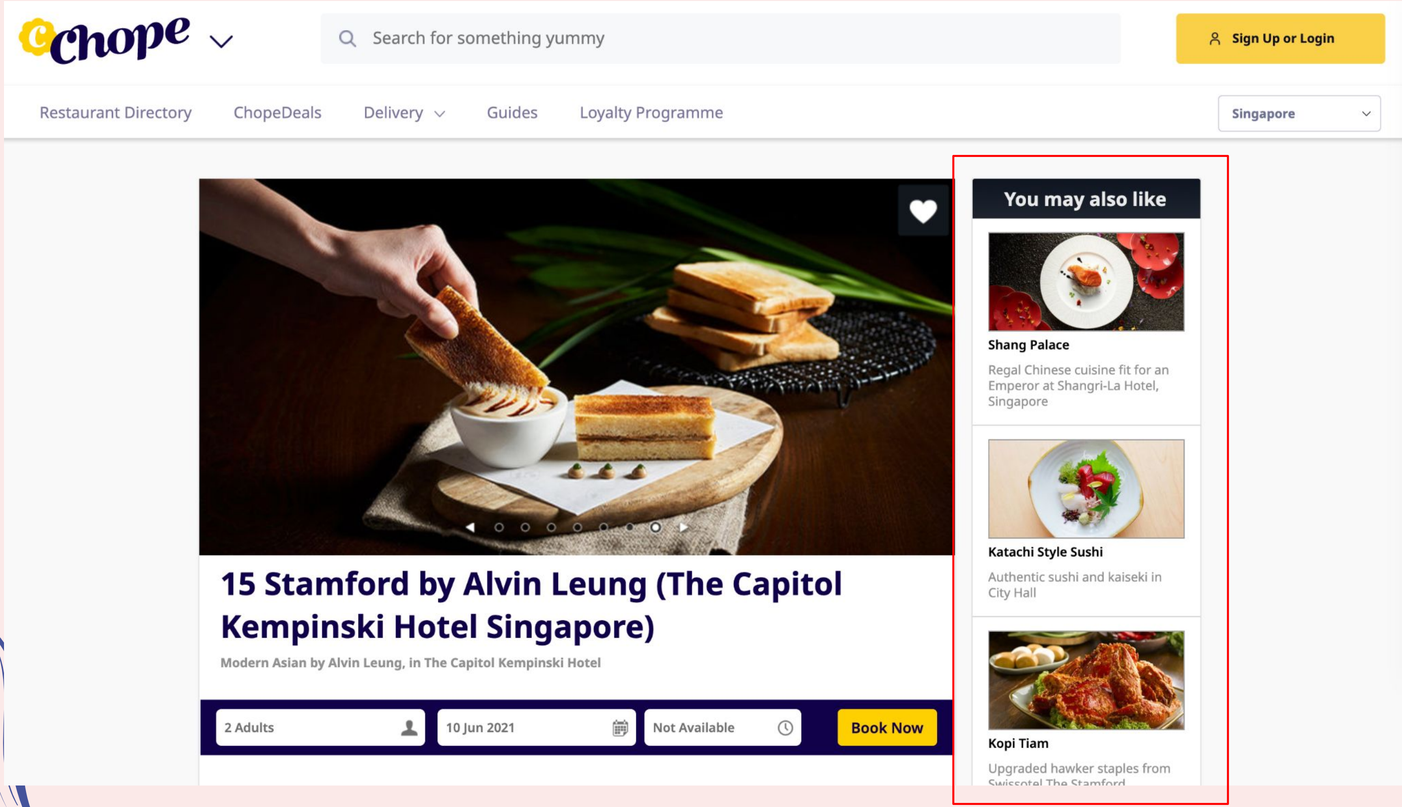
Existing commercial recommender systems have their weaknesses



Google Maps

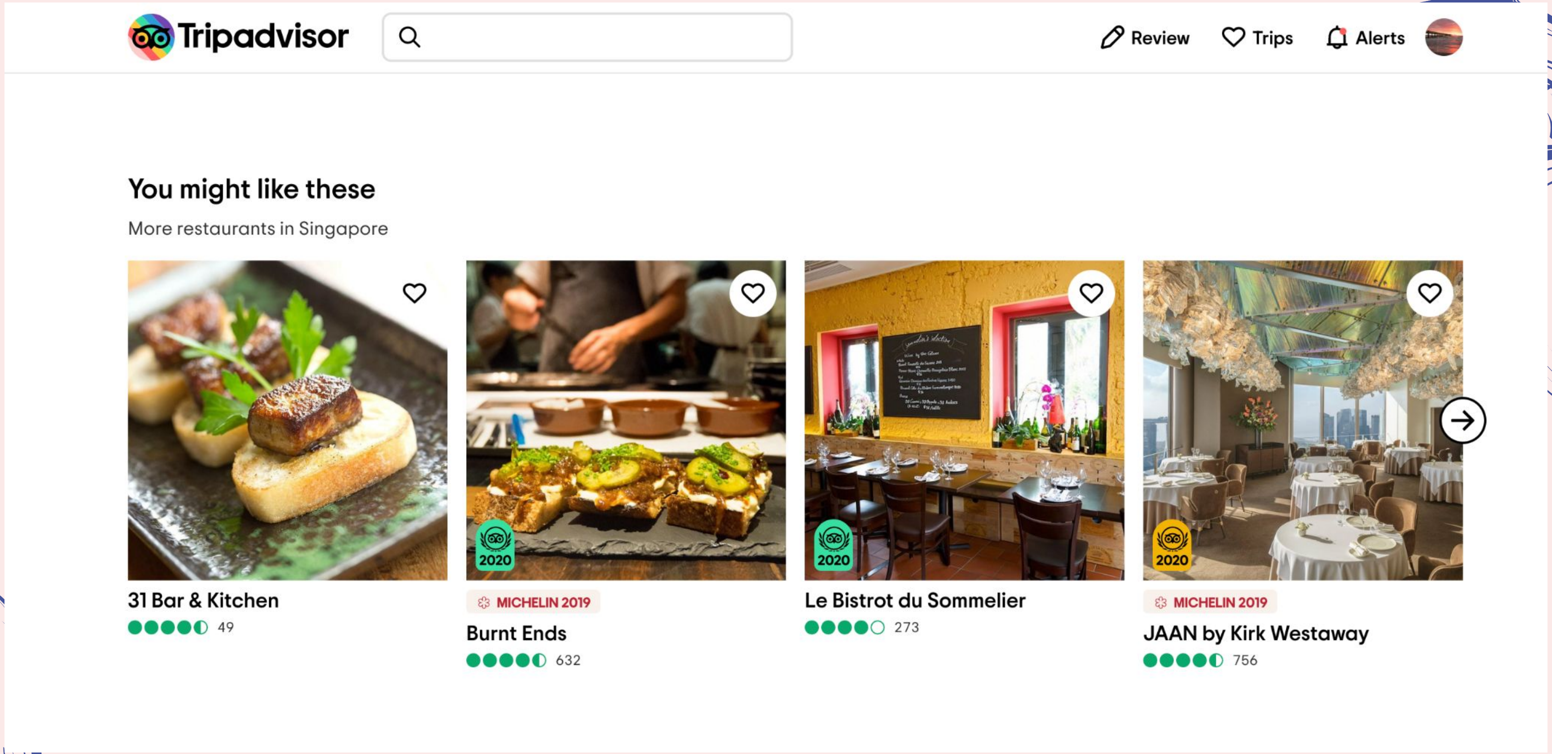
Difficult to use and resource intensive

Rankings prioritise distance and diversity over ratings



Chope

Focus on high end dining with smaller database of ~1,000 restaurants



Tripadvisor

Reviews mainly by tourists. May not be reflective of local tastes.

Project Goals

Build a recommender app which takes inputs of restaurant(s) a user likes and outputs similar restaurants

Simplicity

Easy to use and quickly find recommendation

Relevance

Recommendations are relevant (e.g. similar in cuisine to the input restaurant)

Quality

System should prioritise higher rated restaurants

TABLE OF CONTENTS

Process

01
Data Scraping

02
Data Cleaning and EDA

03
Topic Modelling

04
Build Recommender System

05
Deploy App and Gather Feedback

Datasets Used

6,000

Restaurants

220,000

Reviews

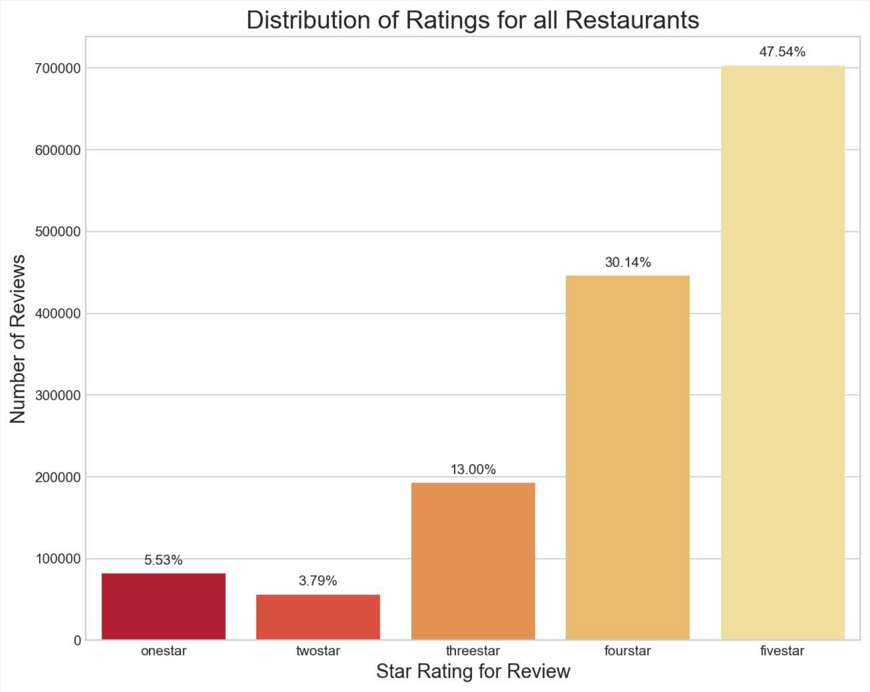
Scraped from Google Maps using Apify

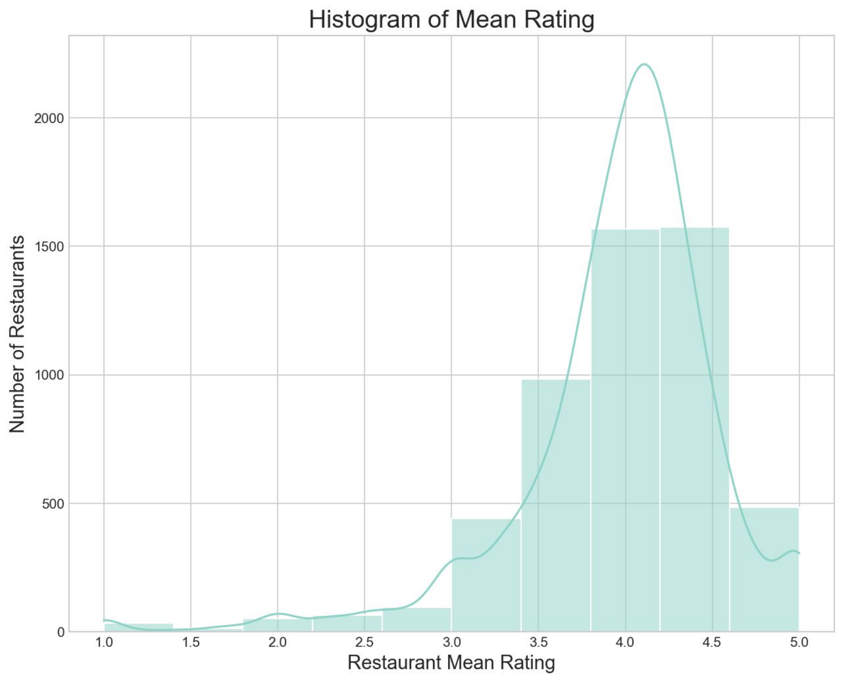
Including fields such as business category, restaurant ID, reviewer ID, restaurant closure status, mean rating

APIFY

Individual ratings tend to favour extremes

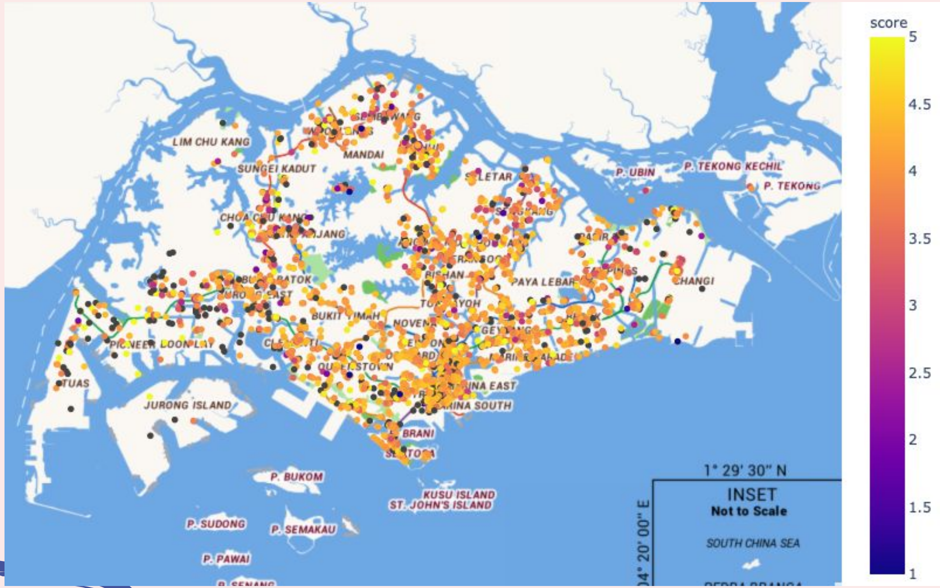
Better to use mean ratings in the model



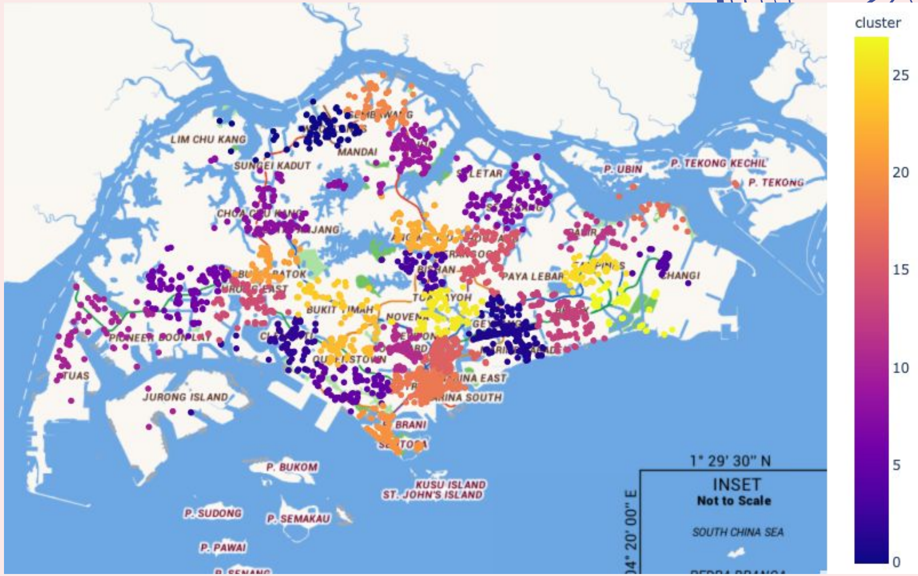


Good geographical spread of restaurants

Colour showing mean rating



Colour showing cluster

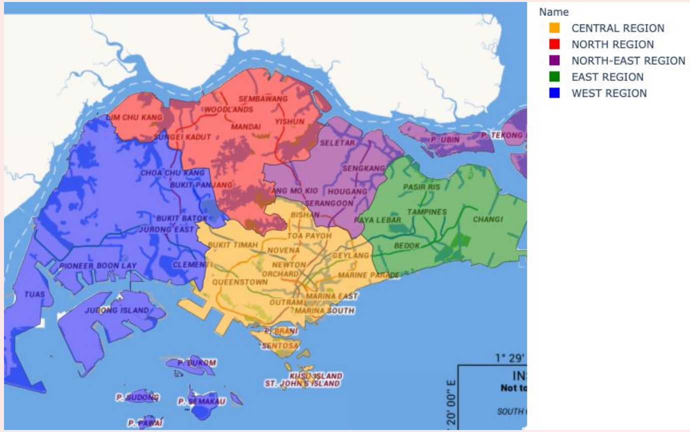


Plotting Geographical Data using Plotly
Used OneMap API to get coordinates and map data

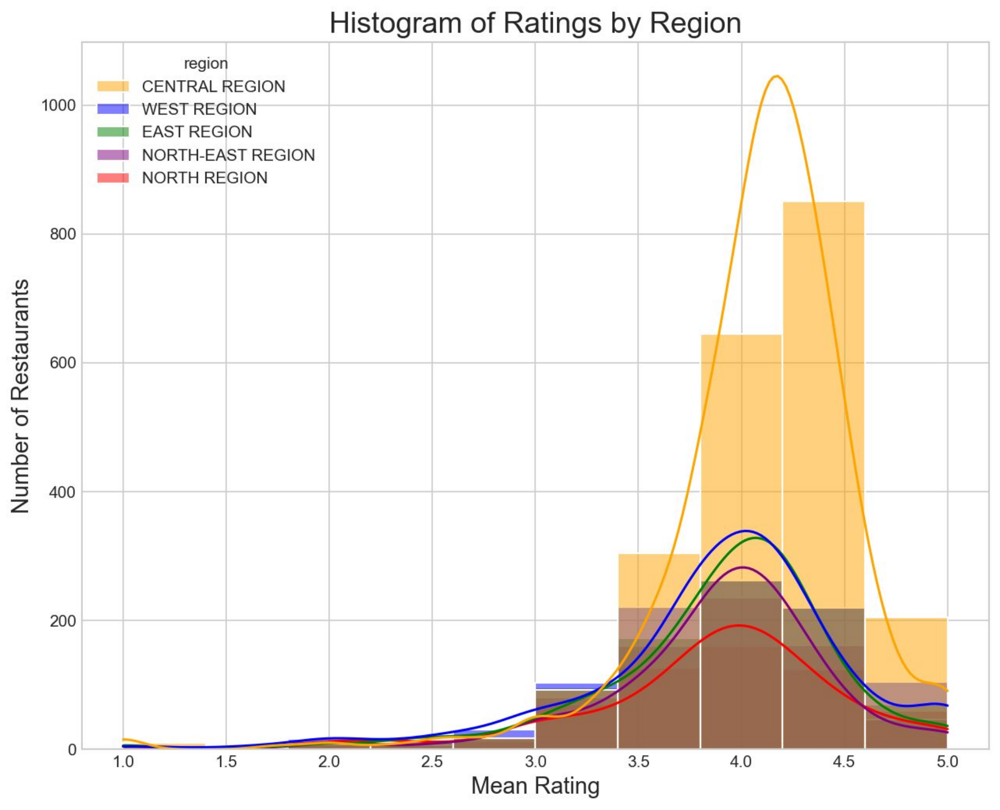
K-Means Clustering
Enable filtering by location in app

Which is best? East or West? Turns out, neither

URA Planning Regions

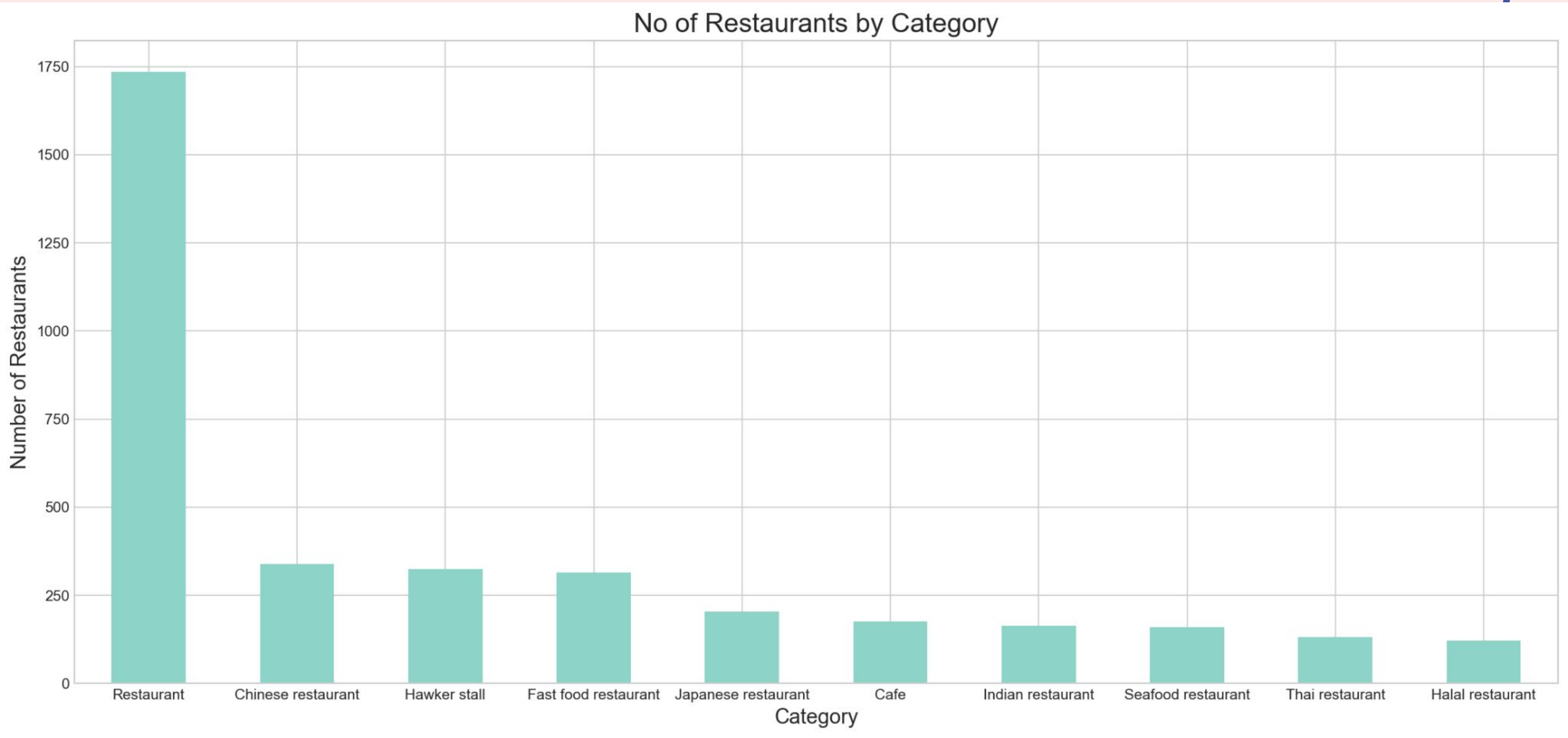


Ratings distribution for central region is more left skewed



Restaurant category data was unspecific

Having good categorical data would enable user to tell restaurant cuisine quickly



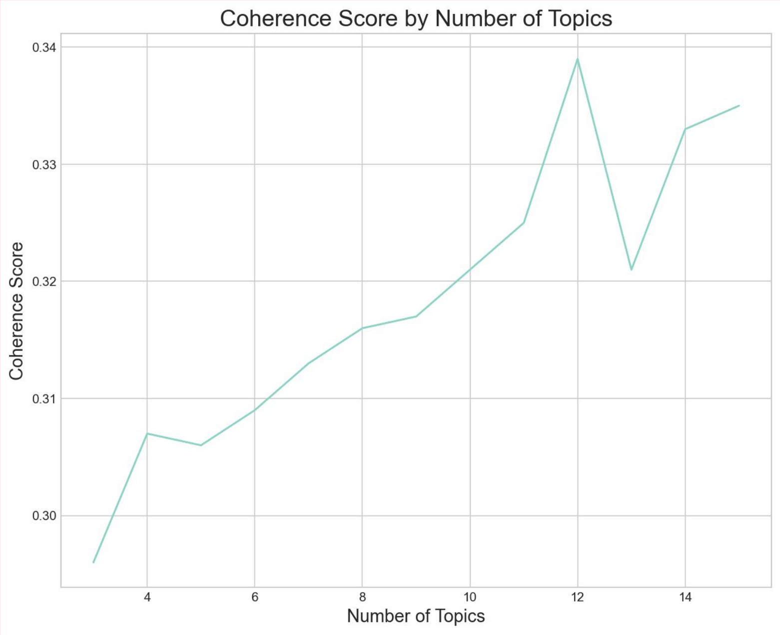
Impute category using topic modelling

Having good categorical data would enable user to tell restaurant cuisine quickly

Process reviews text data (Spacy, NLTK)
Remove stopwords, POS tagging just nouns and adjectives, generating bigrams and trigrams

Topic Modelling using Gensim
Sum reviews data into a document for each restaurant. Use topic modelling to find similarity between restaurants.

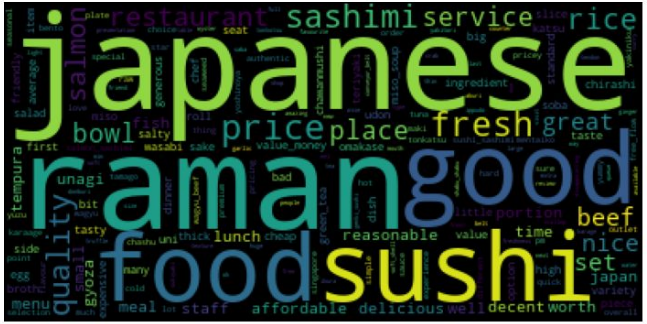
Chose k as 12 topics based on coherence score. High coherence means high similarity in keywords for documents in each topic



Majority of topics were useful

Topic #2

Topic #8



A few were ambiguous

Topic #4

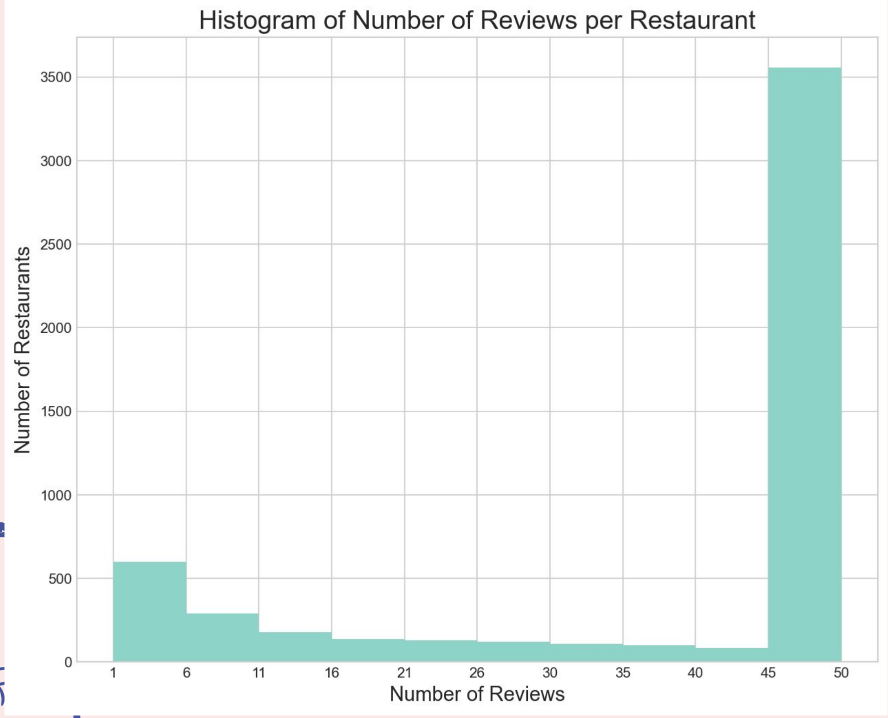
Topic #3

Imputed category using only clearly distinguishable topics

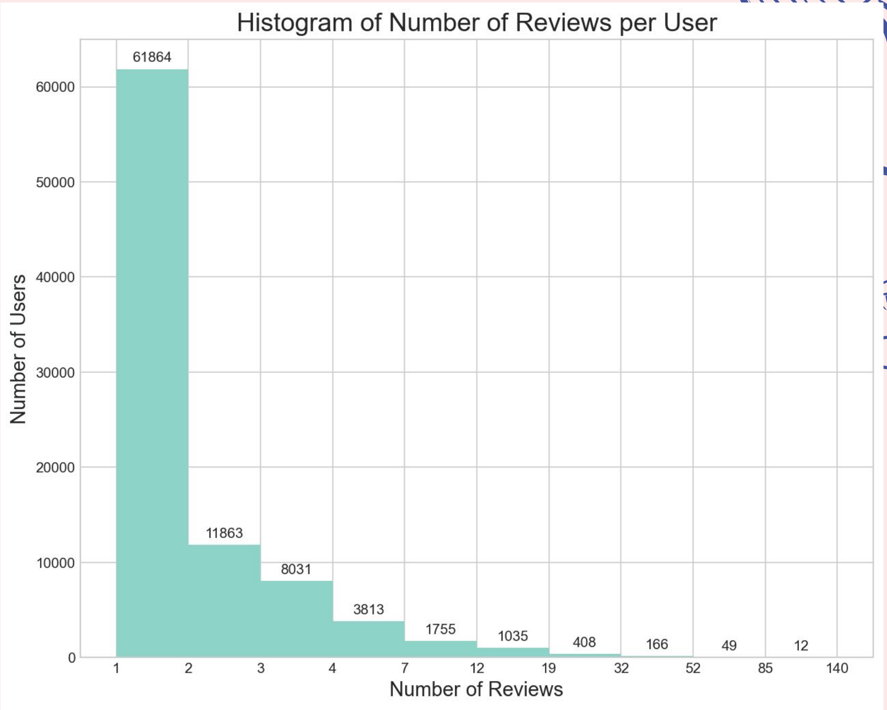
Restaurants with ambiguous category reduced from 1,700 to 300

Content-based filtering > Collaborative-based filtering

Many restaurants have the max scraped of 50 reviews



Number of reviews per user likely insufficient



Content Based Recommender System Workflow

01

TFIDFVectorizer on review data

02

Generate cosine similarity matrix for each restaurant

03

Average cosine similarity matrices, rank and take top 50 most similar

04

Within top 50, only show those within the selected cluster

05

Rank by mean rating and number of reviews. Only show top 10 by rank

TABLE OF CONTENTS

Evaluation Design and Metric

Metric

Precision@k where k is 10.

Precision

$$\text{precision} = \frac{|\{\text{relevant documents}\} \cap \{\text{retrieved documents}\}|}{|\{\text{retrieved documents}\}|}$$

Relevance

Does the output restaurant share similar main dishes with the input restaurant?

Design

Compute average precision@k for 5 queries
Build random recommender system
Compare both precision@k

Case example

Restaurant Name	**Restaurant Category**
L'éclair Pâtisserie (Jewel)	Dessert shop

Case Study for L'éclair Pâtisserie

My recommender does considerably better than random

My Recommender: 0.5 (Avg after 5 diff queries: 0.66)	
Good Bites	Cafe
Lemuel Chocolate (Westway)	Chocolate cafe
The Bread Shop	Cafe
Munchi Delights	Hawker stall
The Spot - Singapore	Fine Dining Restaurant
5 by Sans Façon	French restaurant
The English House by Marco Pierre White	Traditional restaurant
SweetSpot	Cafe
Tamago-EN (Northpoint City)	Cafe
15 Stamford by Alvin Leung	Fine Dining Restaurant

Random Recommender: 0.1 (Avg after 5 diff queries: 0.14)	
Uncle Penyet	Indonesian restaurant
Taj Indian Food	Hawker stall
Pasta Cucina	Hawker stall
Camden Hill Restaurant & Bar	Bistro
La Braceria	Italian restaurant
Hansik Restaurant	Korean restaurant
Tomi Sushi	Sushi restaurant
Yao Ba Cha Seafood Steamboat	Seafood Restaurant
The Tree Cafe (E!Hub)	Cafe
Red House Seafood at Prinsep	Seafood restaurant

Further areas for improvement

Simplicity

Use JavaScript to program a more intuitive UI

Relevance

Integrate other metrics such as cost similarity

Explore the performance of other algorithms such as Jaccard Similarity

Quality

Explore introducing some randomness in ranking to ensure novelty and reusability, while delivering quality

TABLE OF CONTENTS

FUN PART!

http://kw-restaurant-recommender-app.herokuapp.com/

Home

Kun Won's HangryGoWhere 2.0

Simple Restaurant Recommender

Recommend based on a single restaurant

Search for a restaurant you like (look at the map)!

Restaurant Name

Search

Advanced Restaurant Recommender

Recommend by inputting multiple restaurants IDs (copy from search page) and filtering by location

Restaurant ID 1

Fill Here

Restaurant ID 2

Optional

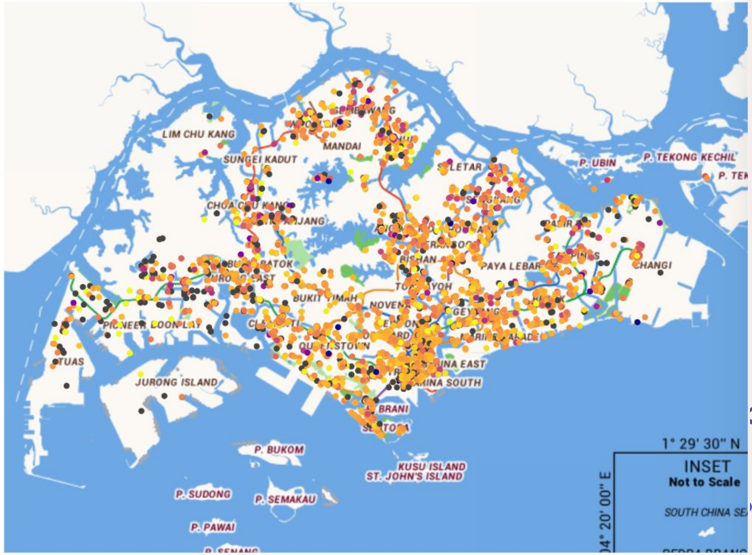
Restaurant ID 3

Optional

Location

All Areas

Recommend!



THANKS

Do you have any questions?

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

Please keep this slide for attribution