

# Find Similar Restaurants

---

Nirupam Purushothama

# Agenda

- Problem Statement
- Dataset Details
- Data Preparation
- Similarity Calculation
- Clustering - Final Results

# Problem Statement

- Given a restaurant, use Yelp Dataset and find restaurants that are similar to it
- Similarity assumptions
  - Restaurants that offer similar facilities
  - Restaurants that offer similar cuisines/services
  - Restaurants that have similar ratings and price ranges
- Complicated because
  - Yelp dataset has 23 different facilities and 471 different types of cuisines/services mentioned

# Dataset Details

- Location - Arizona
  - Has most number of restaurants: 11,072
  - Focusing on single location makes the computation manageable
- Attributes considered
  - Facilities (23) - Yes/No
    - Accepts-Insurance, BikeParking, BusinessAcceptsBitCoin etc.
  - Services / Cuisines (471) - Yes/No
    - Afghan, Chinese, Wraps, Yelp Events, Yoga, Wine Tasting Room etc.
  - Other attributes considered
    - Stars (Rating) - 1 to 5
    - Price-Range - 0 to 4

# Data Preparation

- Extracted restaurant data only for Arizona
- Facilities details and categories information had to be formatted
  - Used Regular expressions
  - And PreProcessing to convert attributes and categories to one-hot-vector format
- Normalized all data between 0 and 1

# Similarity Calculation

- Used notes to implement four functions
  - Jaccard Similarity
  - Cosine Similarity
  - Euclidean Dissimilarity
  - Pearson Correlation

# Experiment - Similarity

Selected Restaurant

## Wild Ginger Asian Bistro

Claimed



189 reviews

Details

\$\$ · Asian Fusion, Chinese, Bars

Edit



861 N Higley Rd  
Ste 111  
Gilbert, AZ 85234

Edit

Get Directions

(480) 907-7717

wildgingeraz.com

Message the business

Send to your Phone



Master Gee Chicken by Katrina L.

Write a Review



### More business info

Takes Reservations **Yes**

Delivery **No**

Take-out **Yes**

Accepts Credit Cards **Yes**

Accepts Apple Pay **No**

Accepts Cryptocurrency **No**

Good For **Lunch, Dinner, Dessert**

Bike Parking **Yes**

Wheelchair Accessible **Yes**

Good for Kids **Yes**

Good for Groups **Yes**

Attire **Casual**

Ambience **Casual, Classy**

Noise Level **Quiet**

Good For Dancing **No**

Alcohol **Full Bar**

Nirupam P

# Cosine Similarity - Results

## Bamboo Cafe

  142 reviews  Details

\$\$ · Chinese 


## Wong's Chinese Cuisine

  341 reviews  Details

\$\$ · Chinese, Asian Fusion 

## Singing Pandas Asian Bar

  483 reviews  Details

\$\$ · Bars, Chinese, Asian Fusion 

## Little Shanghai

  70 reviews  Details

\$\$ · Chinese, Asian Fusion 

## Pacific Gardens

  86 reviews  Details

\$\$ · Sushi Bars, Asian Fusion, Chinese

## Thai Basil Signature

  170 reviews  Details


\$\$ · Thai 



# Jaccard Similarity - Results

**Singing Pandas Asian Bar**  Claimed

 483 reviews  Details

\$\$ · [Bars](#), [Chinese](#), [Asian Fusion](#) 

**Little Shanghai**

 70 reviews  De

\$\$ · [Chinese](#), [Asian Fusion](#) 

**Wong's Chinese Cuisine**

 341 reviews  Details

\$\$ · [Chinese](#), [Asian Fusion](#) 

**Bamboo Cafe**

 142 reviews

\$\$ · [Chinese](#) 

**Dao Kitchen**

 85 reviews

\$ · [Bars](#), [Asian Fusion](#) 

**China Grill** 

 141 review

\$\$ · [Chinese](#) 

# Euclidean Dissimilarity - Results (Sorted in Reverse)

## Singing Pandas Asian Bar

Claimed

483 reviews

Bars, Chinese, Asian Fusion

## Little Shanghai

70 reviews

Chinese, Asian Fusion

## Wong's Chinese Cuisine

341 reviews

Chinese, Asian Fusion

## Dao Kitchen #2 Asian Bistro

116 reviews

Chinese, Asian Fusion

## Panda Garden

147 reviews

Chinese

## Flo's - New Asian Cuisine

390 reviews

Chinese, Asian Fusion

# Pearson Correlation - Results

## Wong's Chinese Cuisine



341 reviews

[Details](#)

\$\$ · [Chinese](#), [Asian Fusion](#)

[Edit](#)

## Singing Pandas Asian Bar

Claimed



483 reviews

[Details](#)

\$\$ · [Bars](#), [Chinese](#), [Asian Fusion](#)

[Edit](#)

## Little Shanghai



70 reviews

[Details](#)

\$\$ · [Chinese](#), [Asian Fusion](#)

[Edit](#)

## Pacific Gardens



86 reviews

[Details](#)

\$\$ · [Sushi Bars](#), [Asian Fusion](#), [Chinese](#)

## Thai Basil Signature



170 reviews

[Details](#)

\$\$ · [Thai](#)

[Edit](#)

## Bamboo Cafe



142 reviews

\$\$ · [Chinese](#)

[Edit](#)

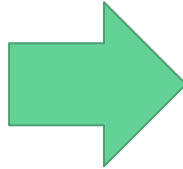


# Cosine Similarity Results - Attributes comparison

## Selected Restaurant

### More business info

Takes Reservations **Yes**  
Delivery **No**  
Take-out **Yes**  
Accepts Credit Cards **Yes**  
Accepts Apple Pay **No**  
Accepts Cryptocurrency **No**  
Good For **Lunch, Dinner, Dessert**  
Bike Parking **Yes**  
Wheelchair Accessible **Yes**  
Good for Kids **Yes**  
Good for Groups **Yes**  
Attire **Casual**  
Ambience **Casual, Classy**  
Noise Level **Quiet**  
Good For Dancing **No**  
Alcohol **Full Bar**



### More business info

Liked by Vegetarians **Yes**  
Takes Reservations **Yes**  
Delivery **No**  
Take-out **Yes**  
Accepts Credit Cards **Yes**  
Accepts Apple Pay **No**  
Accepts Google Pay **No**  
Good For **Lunch, Dinner**  
Parking **Private Lot**  
Bike Parking **Yes**  
Good for Kids **Yes**  
Good for Groups **Yes**  
Attire **Casual**  
Ambience **Casual**  
Noise Level **Quiet**  
Alcohol **Beer & Wine Only**

Liked by Vegans **Yes**

Liked by Vegetarians **Yes**

Takes Reservations **Yes**

Delivery **No**

Take-out **Yes**

Accepts Credit Cards **Yes**

Accepts Apple Pay **No**

Accepts Google Pay **No**

Accepts Cryptocurrency **No**

Good For **Lunch, Dinner**

Parking **Private Lot**

Bike Parking **No**

Wheelchair Accessible **Yes**

Good for Kids **Yes**

Good for Groups **Yes**

Attire **Casual**

Ambience **Casual, Classy**

Noise Level **Average**

### More business info

Takes Reservations **Yes**

Delivery **Yes**

Take-out **Yes**

Accepts Credit Cards **Yes**

Accepts Apple Pay **No**

Parking **Private Lot**

Bike Parking **Yes**

Good for Kids **Yes**

Good for Groups **Yes**

Attire **Casual**

Ambience **Casual**

Noise Level **Quiet**

Alcohol **Full Bar**

# Problem with relying on cosine similarity

## Tang Dynasty Restaurant



29 reviews

[Details](#)

\$\$\$

Karaoke, Chinese, Bars

[Edit](#)

Differences between selected and matched restaurants

- Price Mismatch
- Ratings Mismatch

## China Grill



141 reviews

[Details](#)

\$\$ · Chinese

[Edit](#)

## Taste of China



127 reviews

[Details](#)

\$ · Chinese

[Edit](#)

## King Wok



132 reviews

[Details](#)

\$ · Chinese

[Edit](#)

## Chef Chiang



98 reviews

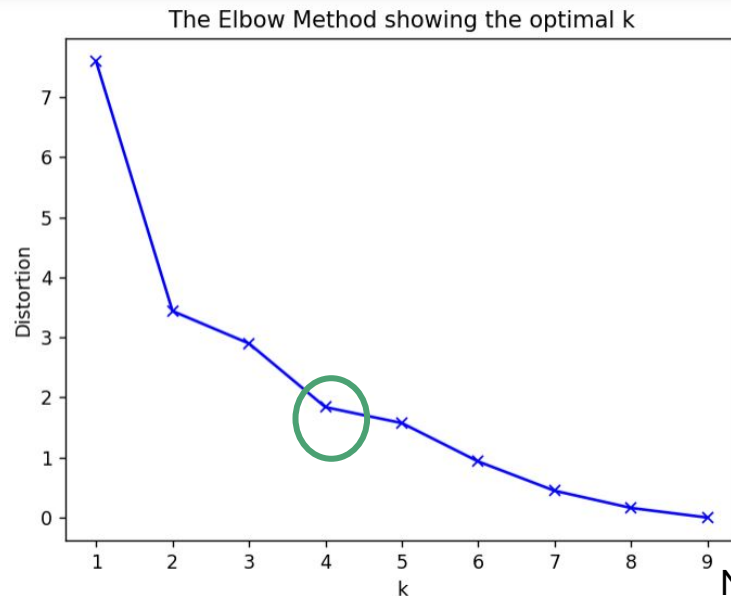
[Details](#)

\$\$ · Chinese

[Edit](#)

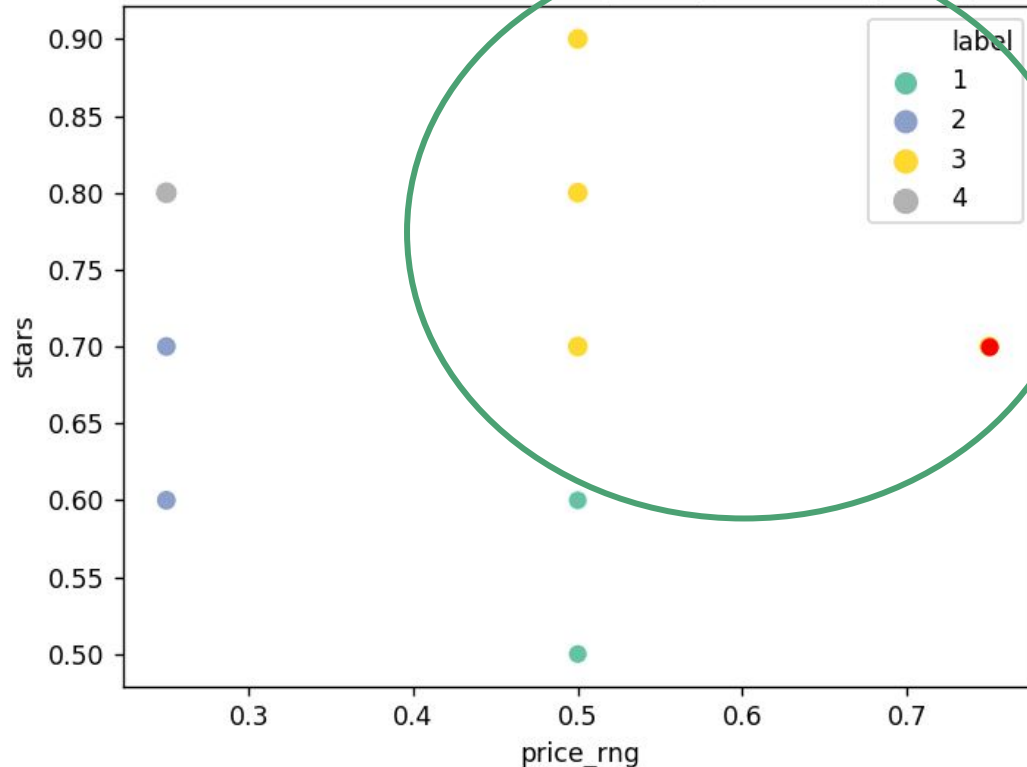
# Clustering

- Extract the top 50 similar restaurants (using Cosine - Similarity)
- Cluster on Ratings & Price-Range
  - Use elbow method to determine the optimal number of clusters
- Example:
  - Using the restaurant in the previous slide
  - Elbow method suggests  $K = 4$



# Visualizing the Restaurant clusters

Restaurant clusters based on similarity, weighted-rating and price-range



- Red is the selected Restaurant
- Selected restaurant is a high priced restaurant
- No other restaurants in the same price point
- Hence, identified with Yellow / Cluster - 3

# Final List - After Clustering

## Tang Dynasty Restaurant



Much better in terms of fit with Ratings and Price information

- No other restaurant exists with \$\$\$ and hence we see a match with \$\$

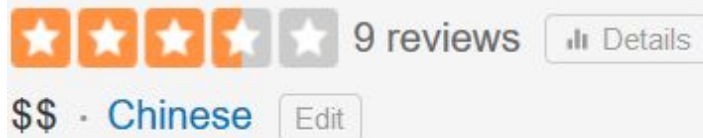
## Dragon Palace



## Chen's Chinese Bistro



## Golden Moon Palace



## Shanghai Club





# Q & A

---

# Thank You

---