
Recommender for Restaurants in Mexico

Stefanie Porcaro

The Goal

To create a system that will recommend a restaurant for a given user using two strategies:

1. **Collaborative Filtering** — based on the assumption that people who agreed in the past will agree in the future, and that they will like similar kinds of items as they liked in the past
2. **Content-Based Filtering** — based on a description of the item and a profile of the user's preferences. They treat recommendation as a user-specific classification problem and learn a classifier for the user's likes and dislikes based on product features

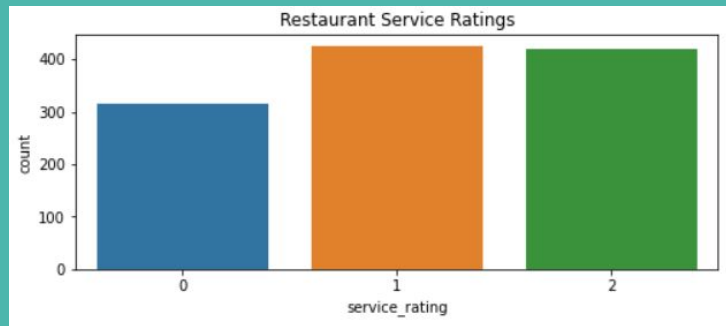
Data

- 1160 ratings from 133 users for 87 restaurants
- Restaurants were in 3 different cities in central Mexico: San Luis Potosi, Ciudad Victoria, Cuernavaca
- Created 3 dataframes:
 1. Restaurant Ratings: overall rating, food rating, service rating
 2. User Information: birth year, profession, budget, preferred ambience, etc.
 3. Restaurant Information: name, city, cuisine, price, ambience, etc.

<https://www.kaggle.com/uciml/restaurant-data-with-consumer-ratings?select=README>

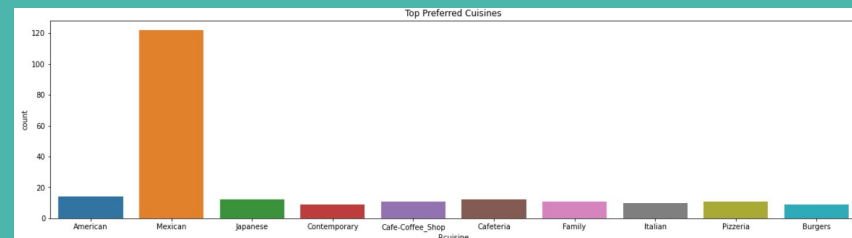
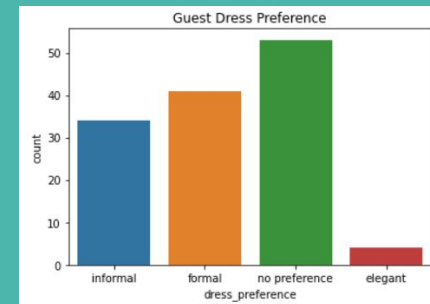
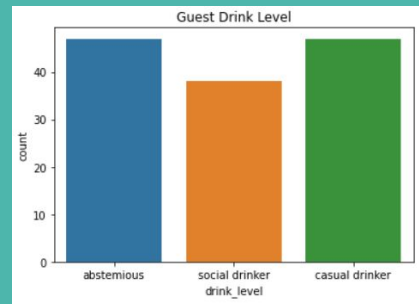
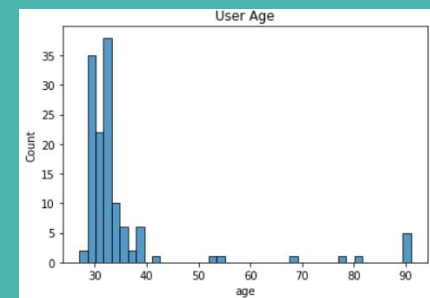
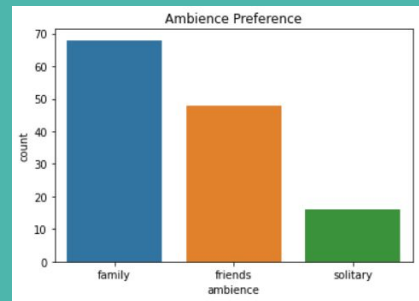
Restaurant Ratings

- Range: [0, 1, 2]
- Overall rating mean = 1.20
Food rating mean = 1.22
Service rating mean = 1.09
- Top 5 users averaged 16 ratings
Top 5 restaurants averaged 29 ratings



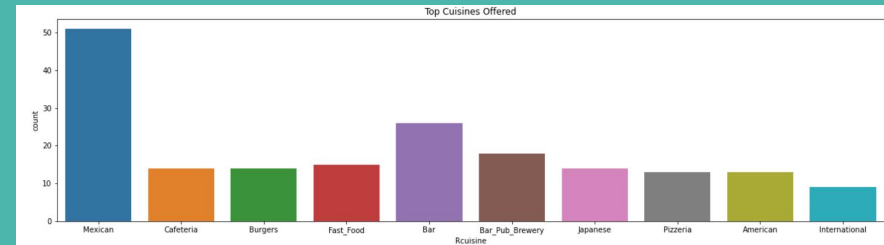
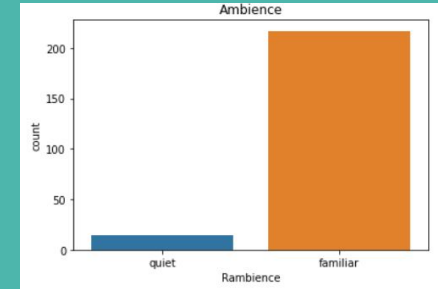
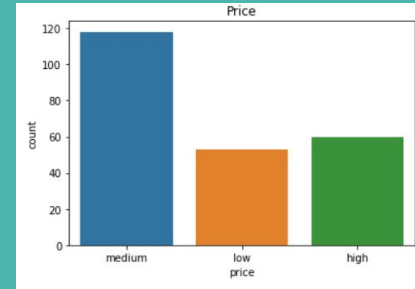
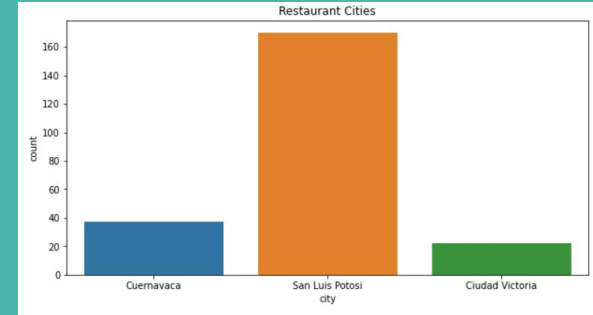
Users

- Changed birth year to age
- Average age = 37
- Converted units to metric system
- Filled in 17 rows with missing data:
 - missing 'budget' → 'medium'
 - missing 'transport' → 'on foot'
- Displayed in a clustermap, there were some slight correlations between attributes:
 - Married - Kids
 - Single = Social Drinker
 - Car owner - Non-smoker
 - Profession in technology - No dress-preference



Restaurants

- Fixed many variations in spelling of city
- Filled in 12 missing cities
- Displayed in a clustermap, there were some slight correlations between attributes:
 - pub/brewery - wine/beer service
 - Contemporary cuisine - alcohol served
 - International cuisine - informal dress code
 - Family ambience - informal dress code



Simple Recommender

Find recommendations based on weighted overall popularity scores

Examples:

Best Mexican restaurants in Cuernavaca:

rating	food_rating	service_rating	score	cuisine	rest_name	city	alcohol	smoking_area	dress_code	accessibility	price	ambience	franchise
1.75	1.75	1.00	1.619914	Mexican	Log Yin	Cuernavaca	No_Alcohol_Served	none	informal	no_accessibility	low	familiar	f
1.50	1.75	1.50	1.426931	Mexican	El Oceano Dorado	Cuernavaca	Full_Bar	permitted	informal	completely	medium	familiar	f
1.50	1.25	1.75	1.426931	Mexican	El cotorreo	Cuernavaca	Wine-Beer	permitted	informal	no_accessibility	low	familiar	f

Best restaurants that are quiet with full accessibility:

rating_count	rating	food_rating	service_rating	score	cuisine	rest_name	city	alcohol	smoking_area	dress_code	accessibility	price	ambience	franchise
6	1.666667	1.500000	1.500000	1.586520	American	tacos los volcanes	San Luis Potosi	No_Alcohol_Served	none	informal	completely	low	quiet	t
4	1.500000	1.000000	1.000000	1.426931	Bar	rockabilly	San Luis Potosi	Full_Bar	section	informal	completely	low	quiet	f
17	1.235294	1.117647	0.882353	1.231675	Bar	crudalia	San Luis Potosi	Wine-Beer	section	casual	completely	low	quiet	f

Collaborative Filtering - User-Based

Tested 10 predictive algorithms:

- KNN-Based: KNNBasic, KNNWithMeans, KNNwithZScore
- Matrix Factorization: SVD, SVDpp, NMF
- Other: SlopeOne, CoClustering

	SVD	KNN with Means	SVDpp	KNN with Z-Score	Baseline Only	Co-Clus tering	Slope One	KNN Baseline	NMF	KNN Basic
rmse	.754	.757	.769	.770	.781	.785	.796	.800	.801	.812
mae	.601	.583	.616	.596	.625	.612	.622	.647	.625	.659

Best choice \Rightarrow **SVD**

rmse = .746 and mae = .590 after hyperparameter tuning

Collaborative Filtering Recommendations

- Found recommendations for a specific user
- Used the 'surprise' library
- Found cosine distances between each restaurant
- Gathered recommendations by finding the user's top-rated restaurants and then found similar restaurants to them
- Can also input any restaurant and gather similar top restaurants

Example:

User #1006

userID	smoker	drink_level	dress_preference	ambience	transport	marital_status	
U1006	True	social drinker	no preference	friends	car owner	single	
hijos	age	interest	personality	religion	activity	weight	budget
independent	32	variety	hard-worker	none	student	165.3465	medium

Find recommendations:

rest_name	city	cuisine	price	ambience	alcohol
Restaurant Bar Hacienda los Martinez	San Luis Potosi	Bar	medium	familiar	Wine-Beer
Gorditas Dona Tota	Ciudad Victoria	Mexican	medium	familiar	No_Alcohol_Served
Hamburguesas La perica	Ciudad Victoria	Mexican	low	quiet	No_Alcohol_Served

Example:

Koye Sushi

rest_name	city	alcohol	smoking_area	dress_code	accessibility	price	ambience	franchise
Koye Sushi	San Luis Potosi	No_Alcohol_Served	none	informal	no_accessibility	high	familiar	f

Find similar restaurants:

rest_name	city	cuisine	price	ambience	alcohol
Restaurante y Pescaderia Tampico	San Luis Potosi	Seafood	high	familiar	No_Alcohol_Served
shi ro ie	San Luis Potosi	Japanese	high	familiar	Wine-Beer
La Estrella de Dimas	San Luis Potosi	Mexican	medium	familiar	No_Alcohol_Served
tacos de la estacion	Ciudad Victoria	Mexican	low	quiet	No_Alcohol_Served

Content-Based Recommendations

- Aggregated restaurant information and restaurant ratings
- Vectorized each row
- Found the Jaccard distances between each vector
- Produce recommendations

Example:

Restaurant given:

placeID	name	city	cuisine	price	ambience	alcohol	dress code	rating	food rating	service rating
132921	Crudalia	SLP	bar	low	quiet	wine-beer	casual	1.24	1.12	.88

Recommended restaurants:

placeID	name	city	cuisine	price	ambience	alcohol	dress code	rating	food rating	service rating
132937	Rockabilly	SLP	bar	low	quiet	Full bar	informal	1.5	1.0	1.0
132856	Unicols Pizza	SLP	Italian	low	familiar	wine-beer	informal	.857	1.14	.857
132951	VIPS	SLP	American	loq	familiar	No alcohol	casual	1.0	.9	.9

Buen Provecho!