Trip recommendation system

Edoardo Berardi Vittur
Ali Yassine
Ethan Greene

Restaurant data

Restaurant Id: Unique id of every restaurant across various cities of the world
• Restaurant Name: Name of the restaurant
• Country Code: Country in which restaurant is located
• City: City in which restaurant is located
• Address: Address of the restaurant
• Locality: Location in the city
• Locality Verbose: Detailed description of the locality
• Longitude: Longitude coordinate of the restaurant's location
• Latitude: Latitude coordinate of the restaurant's location
• Cuisines: Cuisines offered by the restaurant
• Average Cost for two: Cost for two people in different currencies
• Currency: Currency of the country
• Has Table booking: yes/no
• Has Online delivery: yes/ no
• Is delivering: yes/ no
• Switch to order menu: yes/no
• Price range: range of price of food
• Aggregate Rating: Average rating out of 5
• Rating color: depending upon the average rating color
• Rating text: text on the basis of rating of rating
• Votes: Number of ratings casted by people

Standardizing

- Average Cost for two: Cost for two people in different currencies
- Currency: Currency of the country
- Country Code: Country in which restaurant is located
- Price range: range of price of food
- Aggregate Rating: Average rating out of 5

Expense data frames

Restaurant Name	rating	votes	price_for_oneUSD	price_rating
Le Petit Souffle	4.8	314	41.40	3
Izakaya Kikufuji	4.5	591	45.16	3
Heat - Edsa Shangri-La	4.4	270	150.54	4
Ooma	4.9	365	56.45	4
Sambo Kojin	4.8	229	56.45	4
Din Tai Fung	4.4	336	37.64	3
Buffet 101	4.0	520	75.27	4
Vikings	4.2	677	75.27	4
Spiral - Sofitel Philippine Plaza Manila	4.9	621	225.82	4
Locavore	4.8	532	41.40	3
Silantro Fil-Mex	4.9	1070	30.11	3
Mad Mark's Creamery & Good Eats	4.2	488	33.87	3
Silantro Fil-Mex	4.8	294	30.11	3
Guevarra's	4.2	458	37.64	3
Sodam Korean Restaurant	4.3	223	26.35	3

Separating the cuisines

Cuisines
French, Japanese, Desserts
Japanese
Seafood, Asian, Filipino, Indian

cat1	cat2	cat3	cat4	cat5	cat6	cat7
Desserts	French	Japanese	N/A	N/A	N/A	N/A
Japanese	N/A	N/A	N/A	N/A	N/A	N/A
Asian	Filipino	Indian	Seafood	N/A	N/A	N/A
Japanese	Sushi	N/A	N/A	N/A	N/A	N/A
Japanese	Korean	N/A	N/A	N/A	N/A	N/A
Chinese	N/A	N/A	N/A	N/A	N/A	N/A

Separate into multiple categories?

```
def uniqueList(data):
    myCuis = df["Cuisines"].unique()
    new_list = []
    for i in myCuis:
        if i != i:
            continue
        listy = i.split(", ")
        for j in listy:
            new_list.append(j)

    best_list = []
    for food in new_list:
        found = False
        for cuisine in best_list:
            if food == cuisine:
                found = True
        if(not found):
            best_list.append(food)
    return best_list
```

```
def makeScore(data,restrauntCats):
    perRestraunt = []
    catScore = []
    index = 0
    found = True
    for cuisineList in data['Cuisines']:
        #print(cuisineList)
        if cuisineList != cuisineList:
            cuisineList = "None"
        cuisines = cuisineList.split(", ")
        catScore = []
        for restrauntCat in restrauntCats:
            found = False
            #print(restrauntCat)
            for cuisine in cuisines:
                if(found):
                    continue
                #print("Cuis and restCuis")
                #print(cuisine)
                #print(restrauntCat)
                if restrauntCat == cuisine:
                    catScore.append(1)
                    found = True
                    #print(catScore)
            if(not found):
                catScore.append(0)
                #print(catScore)
        perRestraunt.append(catScore)
    return perRestraunt
```

What I've been looking for

Restaurant Name	French	Japanese	Desserts	Seafood	Asian	Filipino	Indian	Sushi	Korean	Chinese	...	Drinks Only	Oriya	Bihari	Assamese	Andhra	Mangalorean
Le Petit Souffle	1	1	1	0	0	0	0	0	0	0	...	0	0	0	0	0	0
Izakaya Kikufuji	0	1	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
Heat - Edsa Shangri-La	0	0	0	1	1	1	1	0	0	0	...	0	0	0	0	0	0
Ooma	0	1	0	0	0	0	0	1	0	0	...	0	0	0	0	0	0
Sambo Kojin	0	1	0	0	0	0	0	0	1	0	...	0	0	0	0	0	0
Din Tai Fung	0	0	0	0	0	0	0	0	0	1	...	0	0	0	0	0	0
Buffet 101	0	0	0	0	1	0	0	0	0	0	...	0	0	0	0	0	0
Vikings	0	0	0	1	1	1	0	0	0	0	...	0	0	0	0	0	0
Spiral - Sofitel Philippine Plaza Manila	0	0	0	0	1	0	1	0	0	0	...	0	0	0	0	0	0
Locavore	0	0	0	0	0	1	0	0	0	0	...	0	0	0	0	0	0

Silhouette score before and after PCA

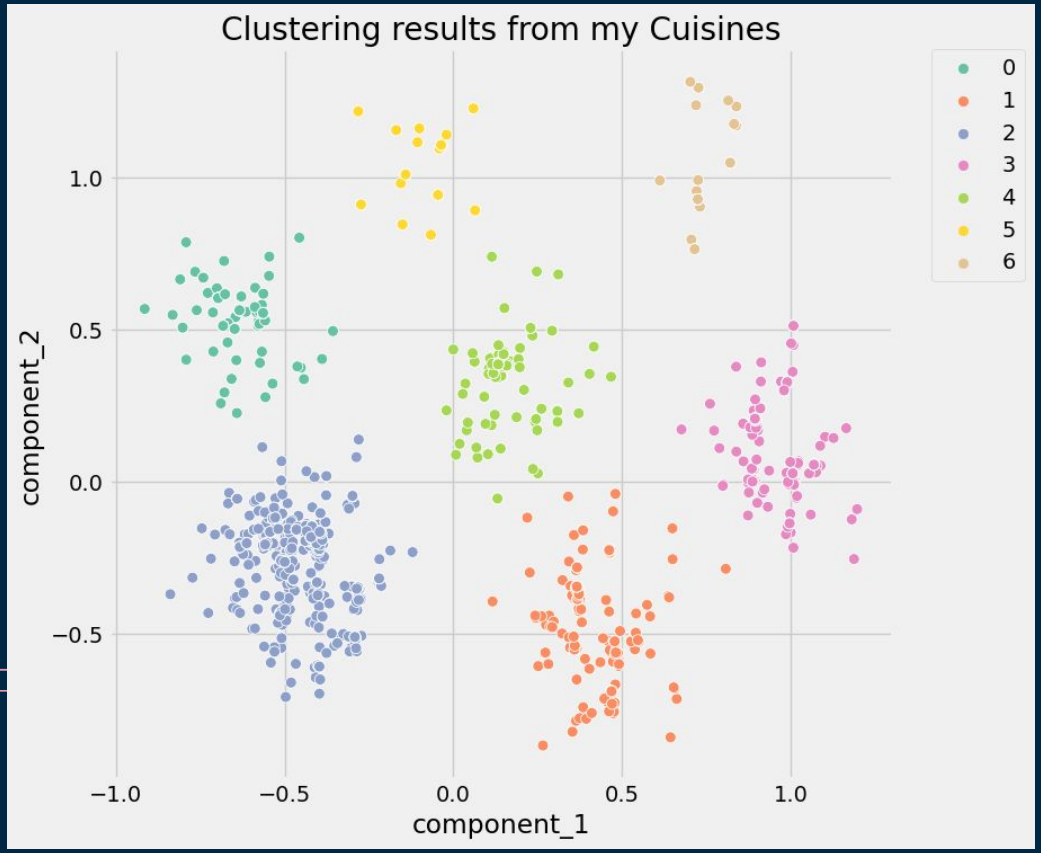
Before:

```
In [150]: kmeans = KMeans(n_clusters=7)

          kmeans.fit(scaled_features)


          kmeans_silhouette = silhouette_score(
              scaled_features, kmeans.labels_
          ).round(2)
          kmeans_silhouette

Out[150]: 0.08
```

After:

```
          pipe.fit(df_cuisines)

          preprocessed_data = pipe["preprocessor"].transform(df_cuisines)

          predicted_labels = pipe["clusterer"]["kmeans"].labels_

          silhouette_score(preprocessed_data, predicted_labels)

Out[92]: 0.77518093602708
```

Visualizing our Restaurant Clusters



Recommending from this:

```
def recc_restraunts(userInfo, data):
```

```
userInfo = makeScore(["Burger, Sushi"], uniqueCuisines)
recc_restraunts(userInfo, data)
```

```
Kinoshita
TT Burger
Green Truck Pub
Shorts Burger and Shine
Sumo Sushi
Jim's Burgers
Monkeypod Kitchen by Merriman
Tokyo Sushi
```