December 12, 2022 | Monday

American Dine-in Culture

Database Foundation for Business Analytics | BUAN 6320.002

THE UNIVERSITY
OF TEXAS AT DALLAS

Team Members:

Dalia Debbarma

Neha Patidar

Paritosh Tiwari

Shagun Gupta

Index

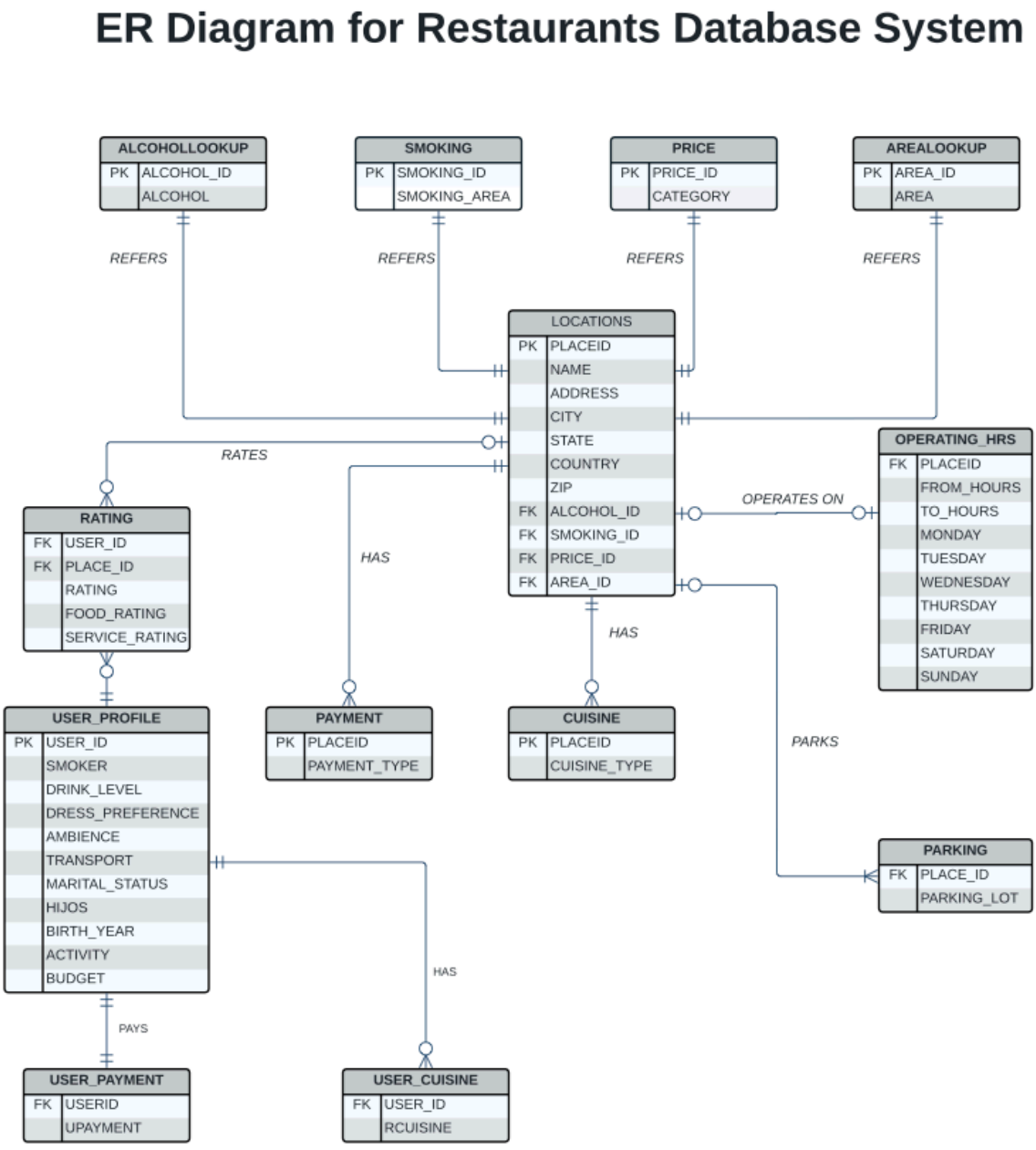
Description of project

The project focuses on:

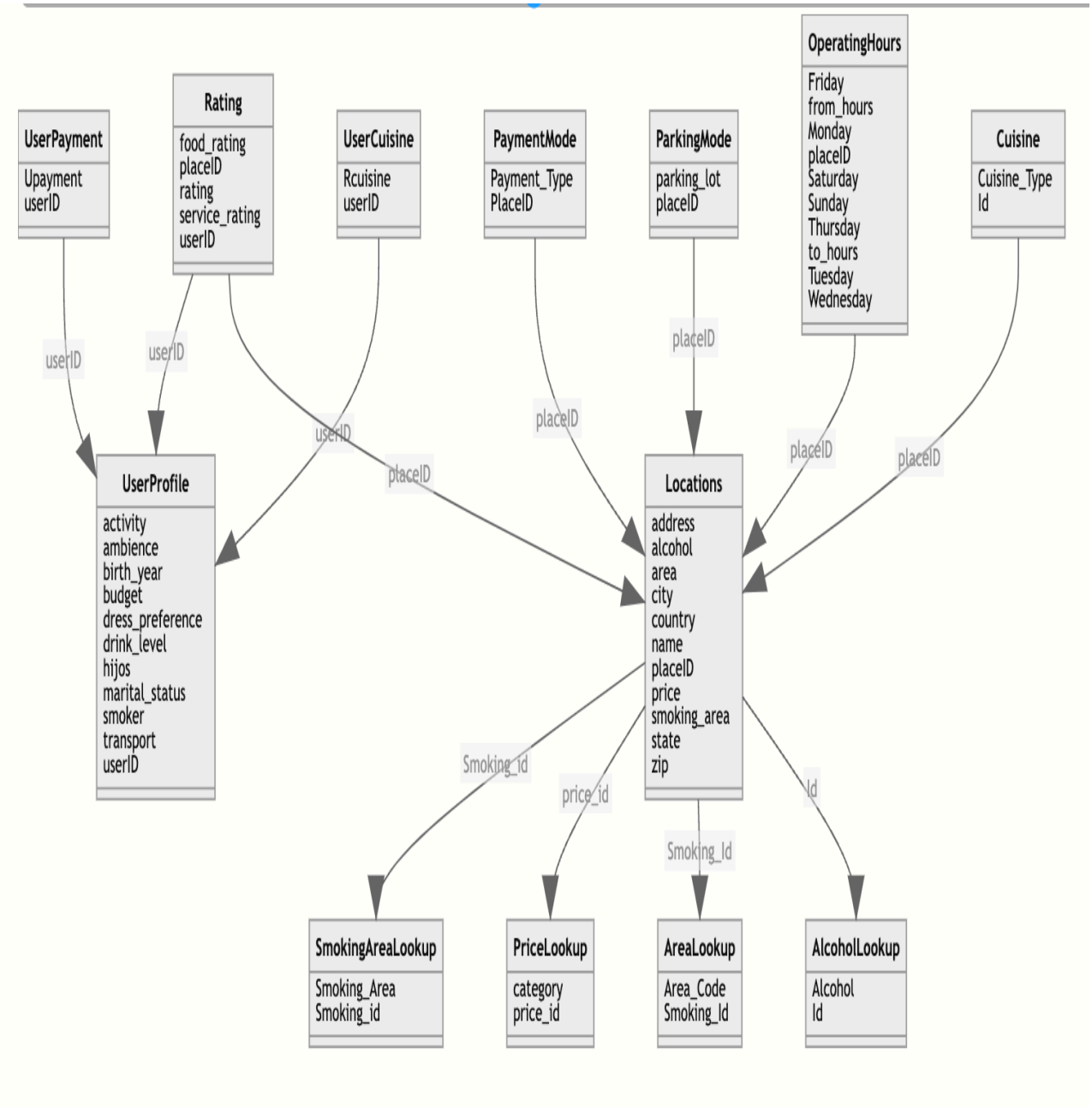
- Deriving key insights for US eateries **trend analysis**
- Easier management of data for eateries across the States for **online food delivery** and **dine-in** options
- Creating a recommender system capable database to determine popularity of restaurants as per given attributes.

The project uses secondary data collected from Kaggle. The dataset contains total of 1452 records and there are total 13 entities, which are Restaurant, Payment options, Alcohol smoking, location, Price, Cuisine, User Profile, Parking, Rating, Accessibility.

Conceptual Design/Logical Data Model



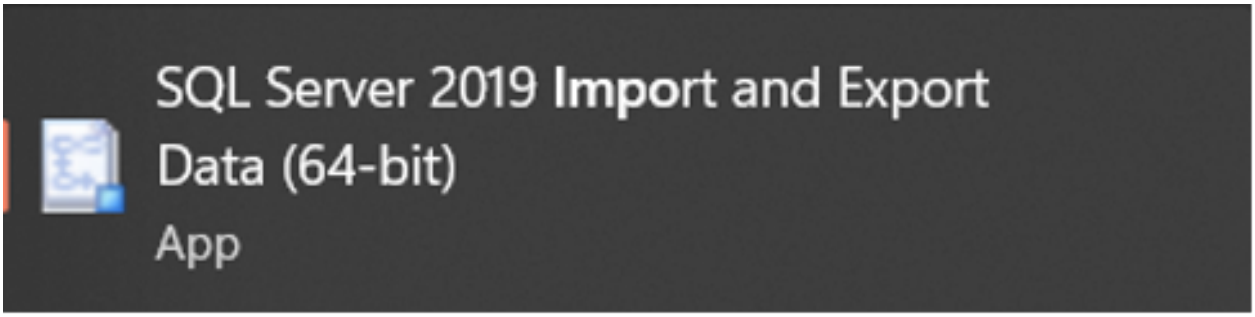
Physical Data Model



Data loading concept used

1. **For managing the data, we created SQL Server in Azure.**

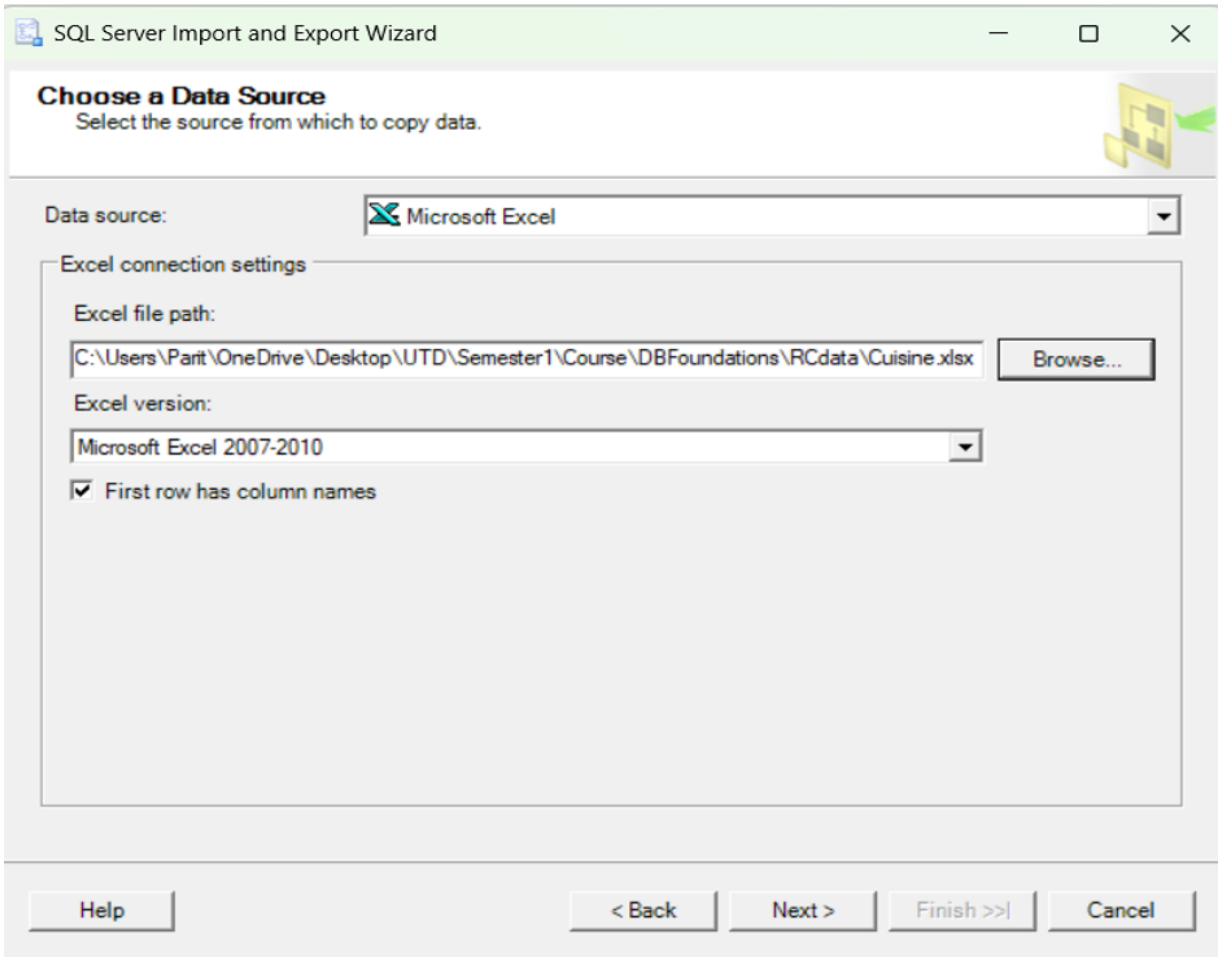
We initially used SQL Server Management studio, we faced challenges for the team to access the database and then we came up with the solution to create the server in Azure and host it, so that we can access the database across the team.
For loading the data in SQL Server Management studio, we used the inbuilt "SQL Server Import and Export Data" in Windows system.



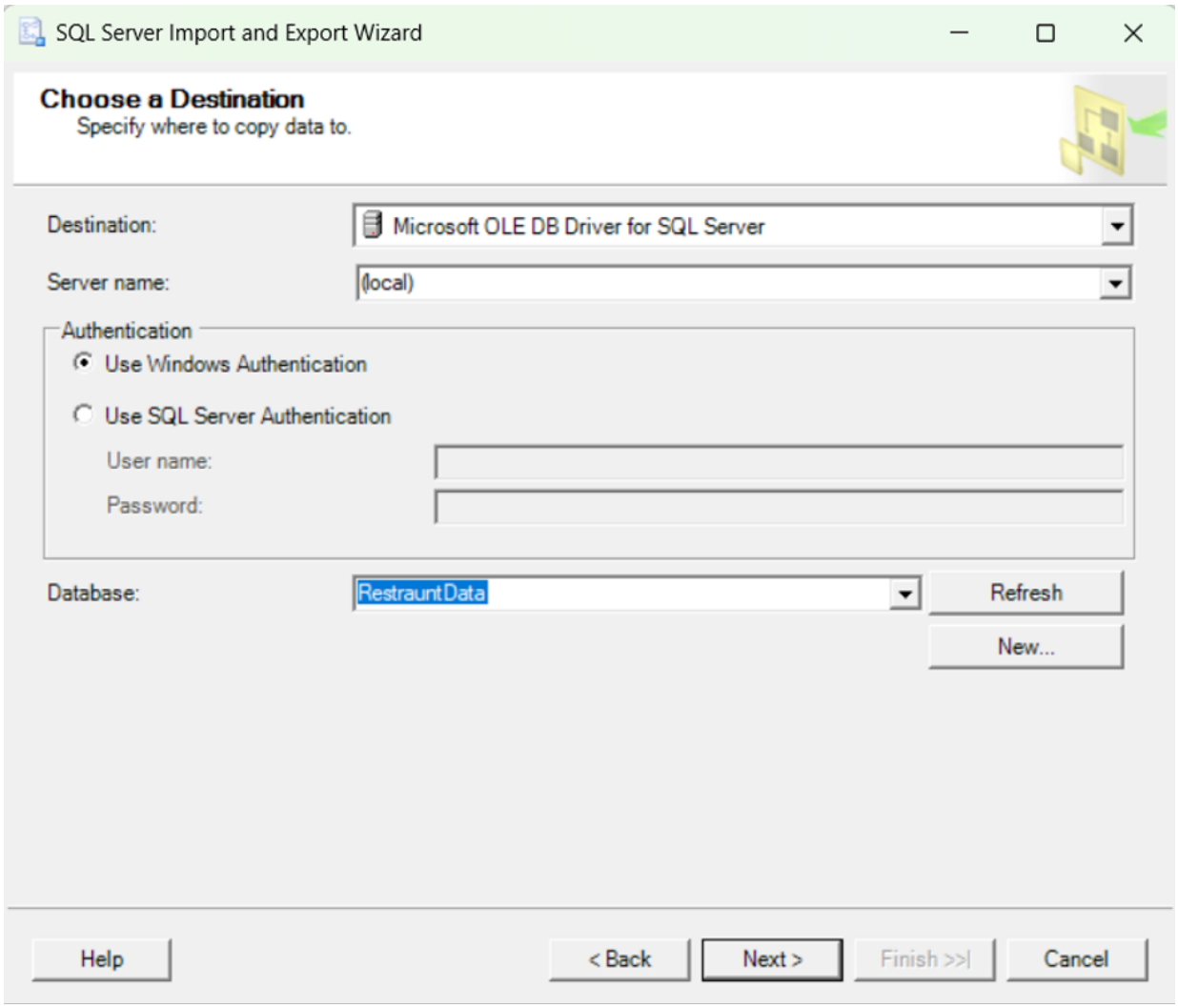
To load the data, we selected the Comma Separated Values files (csv) and exported it to Database.

Steps to load the data:

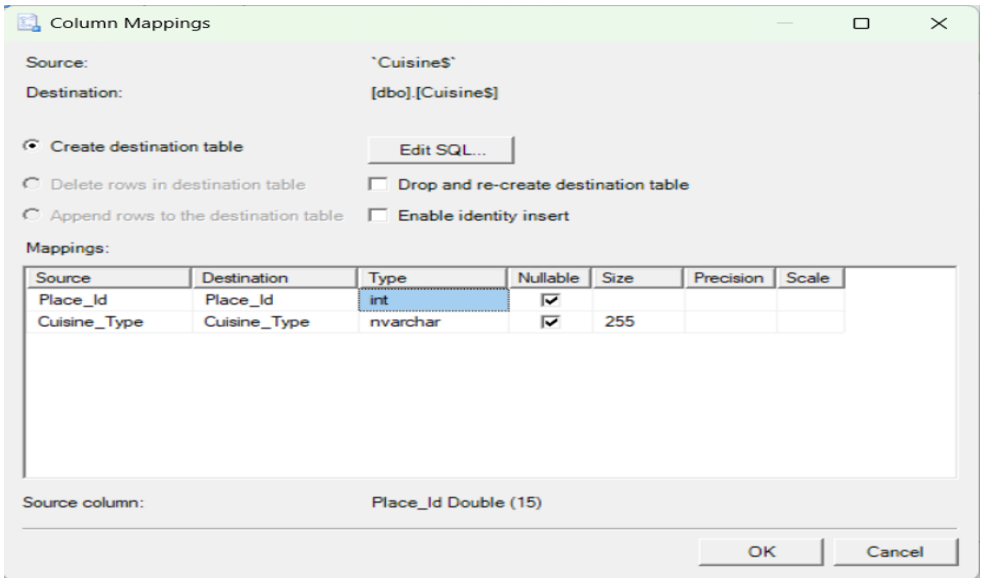
Step 1: Launch SQL Server Import and Export Wizard and select the data source as Microsoft Excel, our data has first row as Column Names



Step 2: On clicking next, we can select the destination as Microsoft OLE DB Driver for SQL server and select the server's name, which in our case was localhost.
We can select the authentication type and the database in which we wish the data to store.



Step 3: We map the columns to the columns in database and change the datatype as required.

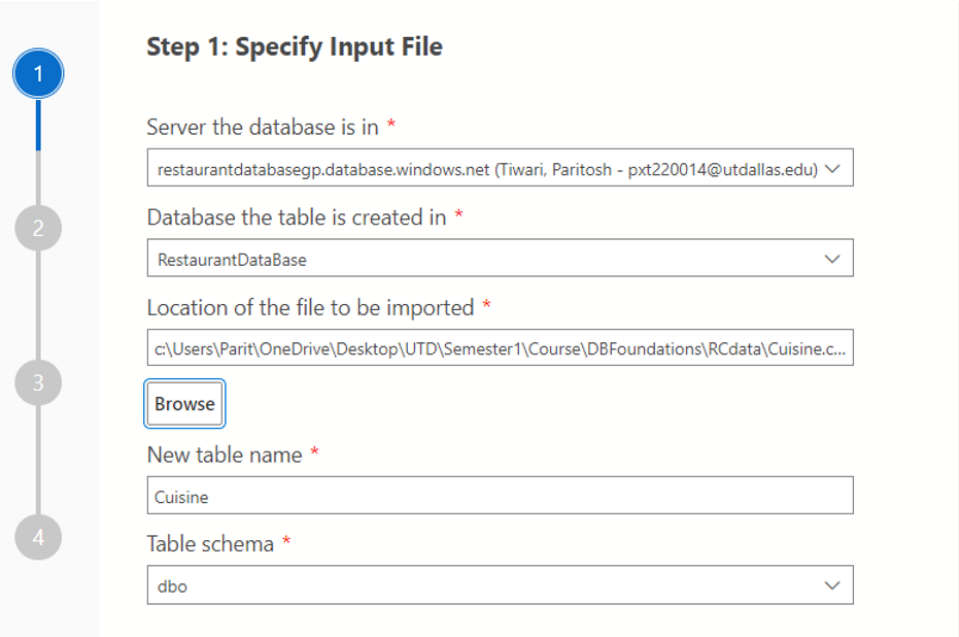


Step 4: We finish the import, and we will be able to see the table in the database.

Since we faced the issue of accessing the data for all team members, we hosted the data in Azure SQL Server and provided Azure Active directory and SQL login to the users.

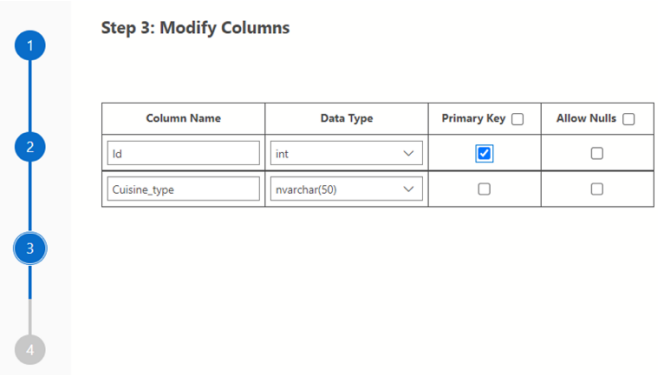
2. **For importing the data, we used the SQL Server import extension for Azure Database.**

Step 1: Launch the importing wizard and provide server name, database name, file path from which we will be importing data, table name and schema name.



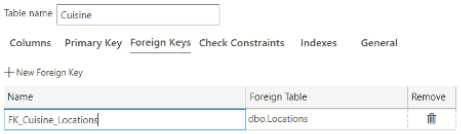
Step 2: Clicking on Next, we can preview the data format in which the data will be uploaded to database.

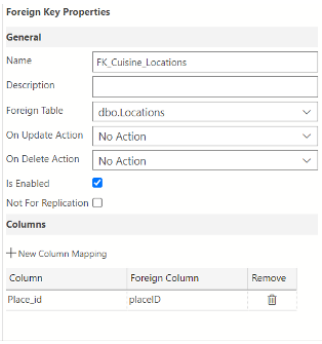
Step 3: We can specify primary key and column names.



Step 4: Import Data, and the data in the CSV or excel file will be uploaded to the database.

Step 5: After the import, we open the design of the table and can add the constraints.





Insights

We have driven few insights from the database, American Dine-in culture.

1. Which payment types are frequently used by the customers in decreasing order, group by payment type.

```
SELECT COUNT(Payment_Type) as "Number of transactions", Payment_Type
FROM [dbo].[PaymentMode]
GROUP by Payment_Type
ORDER BY COUNT(Payment_Type) desc;
```

Query results:

Results | Messages

	Number of transactions	Payment_Type
1	500	cash
2	255	VISA
3	194	MasterCard-Eurocard
4	153	American_Express
5	130	bank_debit_cards
6	42	Diners_Club
7	11	Discover
8	10	checks
9	7	gift_certificates
10	7	Carte_Blanche
11	5	Japan_Credit_Bureau

The top five payments methods are CASH, VISA, MasterCard-Eurocard, America_Express & bank_debit_cards.

2. At what time of the day are the prices high.

```
SELECT DISTINCT from_hours, to_hours, Price.ID, Price.Price
FROM [dbo].[OperatingHours] Operatinghrs
INNER JOIN [dbo].[Locations] Locations
ON Operatinghrs.placeID = Locations.placeID
INNER JOIN [dbo].[PriceLookup] Price
ON Price.ID = Locations.Price
WHERE Price.Price = 'high';
```

Query results:

Results Messages

	from_hours	to_hours	ID	Price
1	00:00:00	00:00:00	1	high
2	00:00:00	23:30:00	1	high
3	07:00:00	19:00:00	1	high
4	07:00:00	22:30:00	1	high
5	07:00:00	23:00:00	1	high
6	07:00:00	23:30:00	1	high
7	08:00:00	20:00:00	1	high
8	08:00:00	21:30:00	1	high
9	08:00:00	22:00:00	1	high
10	08:00:00	23:30:00	1	high
11	09:30:00	23:30:00	1	high
12	11:00:00	22:00:00	1	high
13	11:00:00	22:30:00	1	high
14	12:00:00	19:30:00	1	high
15	12:00:00	23:30:00	1	high
16	13:00:00	19:00:00	1	high
17	13:00:00	22:30:00	1	high
18	13:00:00	23:30:00	1	high
19	13:30:00	18:30:00	1	high
20	13:30:00	22:30:00	1	high
21	13:30:00	23:30:00	1	high
22	14:00:00	18:00:00	1	high
23	14:00:00	20:30:00	1	high
24	14:00:00	22:00:00	1	high
25	14:00:00	23:30:00	1	high

To decide the timmings of going out for dining according to the prices range, so in most cases prices are high after 10:30 pm.

3. Which places serves Alcohol and has smoking area. To get insights regarding the places which have both alcohol and smoking areas to help in deciding according to people's preferences.

```
SELECT Locations.name, alcohol.Alcohol, smoke.Smoking_Area
FROM [dbo].[AlcoholLookup] alcohol
INNER JOIN [dbo].[Locations] Locations
ON alcohol.Id = Locations.alcohol
INNER JOIN [dbo].[SmokingAreaLookup] smoke
ON Locations.smoking_area = smoke.Id
WHERE smoke.Smoking_Area ='permitted' AND  alcohol.Alcohol IN ('Full_Bar','Wine-
Beer');
```

Query results:

Results Messages

	name	Alcohol	Smoking_Area
1	Restaurant Familiar El Ch…	Wine-Beer	permitted
2	El cotorreo	Wine-Beer	permitted
3	El Oceano Dorado	Full_Bar	permitted

People have only three choices for dine-in which serve alcohol and have a smoking area.

4. Display names of the places which have valet parking.

```
SELECT locations.name, Locations.address, locations.placeID
FROM [dbo].[Locations] locations
INNER JOIN [dbo].[ParkingMode] parking
ON locations.placeID = parking.placeID
WHERE parking.parking_lot = 'valet parking';
```

Query results: -

Results Messages

	NAME	ADDRESS	PLACE_ID
1	Rincon del Bife	null	134975
2	La Posada del Virrey	Av. V. Carranza	132862

Above given two restaurants which have valet parking which will increase customers' convenience.

5. Total number of places where smoking is permitted for selecting option where to visit.

```
SELECT COUNT(smoke.Smoking_Area) As placesHaving_SmokeArea
FROM [dbo].[SmokingAreaLookup] smoke
INNER JOIN [dbo].[Locations] locations
ON smoke.Smoking_id= locations.smoking_area
WHERE smoke.Smoking_Area IN ('permitted','only at bar','section');
```

Query result:-

Results Messages

	placesHaving_SmokeArea
1	34

6.Total number of people who prefer to do payment using CASH.

```
SELECT COUNT(Upayment) AS "Total no of CASH payments"
FROM [dbo].[UserPayment]
WHERE Upayment = 'cash'
```

Query results: –

Results Messages

	Total no of CASH payments
1	131

There are 131 users who paid using CASH.

7. Update the address as "850 Cecil drive", city as "Richardson", state as "TX", country as "United States" and zip as "75080" where name is equal to "Cafe Chaires".

```
UPDATE [dbo].[locations]
SET address = '850 Cecil drive', city='Richardson',state='TX',country='United
States',zip='75080'
WHERE name ='cafe ambar';
```

Messages

9:00:10 PM Started executing query at Line 51
(1 row affected)
Total execution time: 00:00:00.081

This allows easier management of restaurants data.

```
--Check if columns are updates or not–
```

```
SELECT name,address, city, state, country, zip
from [dbo].[locations]
WHERE name ='cafe ambar';
```

Results Messages

	name	address	city	state	country	zip
1	cafe ambar	850 Cecil drive	Richardson	TX	United States	75080

8.Display count of the cuisine and name of the cuisine where rating of food is highest.

```
SELECT COUNT(cuisine.Rcuisine) "Total count of cuisine", cuisine.Rcuisine
FROM [dbo].[Rating] rating
INNER JOIN [dbo].[UserCuisine] cuisine
ON rating.userID = cuisine.userID
WHERE rating.food_rating= 2
GROUP BY cuisine.Rcuisine
ORDER BY COUNT(cuisine.Rcuisine);
```

Query results: –

Results Messages

	Total count of cuisine	Rcuisine
1	2	Turkish
2	2	Doughnuts
3	2	Hawaiian
4	2	Australian
5	2	Korean
6	3	Spanish
7	3	Soup
8	5	Juice
9	6	Seafood
10	6	Cuban
11	6	Asian
12	6	Diner
13	6	Game
14	6	Middle_Eastern
15	6	Mongolian
16	6	Dessert-Ice_C…
17	7	Bagels
18	7	Continental-E…
19	7	Bar
20	8	Sushi
21	9	Bakery
22	10	Deli-Sandwich…
23	12	Tex-Mex
24	13	Indian-Pakist…
25	13	Polish
26	13	Hot_Dogs
27	13	Fusion
28	13	Polynesian
29	13	Portuguese
30	13	Eastern_Europ…
31	13	Lebanese
32	14	Afghan
33	15	Latin_American
34	17	Breakfast-Bru…
35	17	Family
36	17	Fast_Food
37	18	Chinese
38	19	Burgers
39	19	Moroccan
40	20	Barbecue
41	22	Italian
42	22	Regional
43	22	Pizzeria
44	24	Cafeteria
45	27	Contemporary
46	34	American
47	37	Cafe-Coffee_S…
48	42	Japanese
49	365	Mexican

There are multiple restaurants with the highest food ratings.

9. Display the type of cuisine and no of cuisine types in decreasing order of total number of cuisines.

```
SELECT DISTINCT Cuisine_Type, count(Cuisine_Type) as  total_number_of_cuisine
FROM [dbo].[Cuisine]
Group by Cuisine_Type
ORDER BY total_number_of_cuisine DESC;
```

Results Messages

	Cuisine_Type	total_number_of_cuisine
1	Mexican	956
2	International	248
3	American	236
4	Dutch-Belgian	220
5	Italian	168
6	Greek	132
7	Bar	128
8	French	124
9	Cafe-Coffee_Sh…	108
10	Pizzeria	100
11	Bar_Pub_Brewery	96
12	Cafeteria	92
13	Chinese	84
14	Fast_Food	80
15	Seafood	72
16	Japanese	68
17	Family	56
18	German	56
19	Mediterranean	52
20	Burgers	52
21	Vegetarian	40
22	Contemporary	36
23	Deli-Sandwiches	36
24	Steaks	32
25	Latin_American	28
26	Asian	28
27	Hot_Dogs	28
28	Bakery	24
29	Sushi	24
30	Juice	24

31	Armenian	20
32	Polish	20
33	Continental-Eu…	16
34	Regional	12
35	Breakfast-Brun…	12
36	African	12
37	Southwestern	12
38	Dessert-Ice_Cr…	12
39	Spanish	12
40	Diner	12
41	Barbecue	12
42	Vietnamese	8
43	Eastern_Europe…	8
44	Game	8
45	Bagels	4
46	Thai	4
47	California	4
48	Organic-Healthy	4
49	Caribbean	4
50	Turkish	4
51	Persian	4
52	Ethiopian	4
53	Brazilian	4
54	Soup	4
55	Mongolian	4
56	Afghan	4
57	Fine_Dining	4
58	Southern	4
59	Korean	4

The most common cuisine served in USA are Mexica, Continental, American.

10. Display the names and ratings of the restaurant whose food and service rating are 2.

```
SELECT DISTINCT locations.name, ratings.rating, ratings.food_rating,
ratings.service_rating
FROM [dbo].[locations] locations
INNER JOIN [dbo].[Rating] ratings
ON locations.placeID =ratings.placeID
WHERE ratings.food_rating = 2 AND ratings.service_rating=2 AND ratings.rating =2;
```

Query Result:

	name	rating	food_rating	service_rating
1	Cabana Huasteca	2	2	2
2	Cafe Chaires	2	2	2
3	cafe punta del cielo	2	2	2
4	Cafeteria cenidet	2	2	2
5	Cafeteria y Restaurant El…	2	2	2
6	Carls Jr	2	2	2
7	carnitas mata calle Emili…	2	2	2
8	carnitas_mata	2	2	2
9	Chaires	2	2	2
10	Chilis Cuernavaca	2	2	2
11	churchs	2	2	2
12	crudalia	2	2	2
13	dairy queen	2	2	2
14	Dominos Pizza	2	2	2
15	don burguers	2	2	2
16	El angel Restaurante	2	2	2
17	El Club	2	2	2
18	El Herradero Restaurante …	2	2	2
19	el lechon potosino	2	2	2
20	El Mundo de la Pasta	2	2	2
21	El Oceano Dorado	2	2	2
22	el pueblito	2	2	2
23	El Rincon de San Francisco	2	2	2
24	El Rincón de San Francisco	2	2	2
25	emilianos	2	2	2
26	Giovannis	2	2	2
27	Gordas de morales	2	2	2
28	Gorditas Doa Gloria	2	2	2
29	Gorditas Dona Tota	2	2	2
30	Hamburguesas La perica	2	2	2

31	KFC	2	2	2
32	Kiku Cuernavaca	2	2	2
33	Koye Sushi	2	2	2
34	la Cantina	2	2	2
35	La Cantina Restaurante	2	2	2
36	la Cochinita Pibil Restau…	2	2	2
37	la Estrella de Dimas	2	2	2
38	La Fontana Pizza Restaura…	2	2	2
39	la parroquia	2	2	2
40	la perica hamburguesa	2	2	2
41	La Posada del Virrey	2	2	2
42	La Virreina	2	2	2
43	Little Cesarz	2	2	2
44	Log Yin	2	2	2
45	los Toneles	2	2	2
46	Luna Cafe	2	2	2
47	Mariscos El Pescador	2	2	2
48	Mariscos Tia Licha	2	2	2
49	McDonalds Centro	2	2	2
50	Michiko Restaurant Japones	2	2	2
51	palomo tec	2	2	2
52	pizza clasica	2	2	2
53	Pizzeria Julios	2	2	2
54	Potzocalli	2	2	2
55	Preambulo Wifi Zone Cafe	2	2	2
56	puesto de tacos	2	2	2
57	Restaurant and Bar and Cl…	2	2	2
58	Restaurant Bar Coty y Pab…	2	2	2
59	Restaurant Bar Hacienda l…	2	2	2
60	Restaurant de Mariscos de…	2	2	2

So, there are a total of 60 restaurants which are rated 2 in service, food and rating, which is the highest rating.

11. Display the no of locations and the zip code for that location where the address and zip code are not null and having no of cuisines more than or equal to 3.

```
SELECT count(locations.name) AS no_of_cuisines, Locations.zip
FROM [dbo].[Locations] locations
INNER JOIN [dbo].[PriceLookup] price
ON locations.price = price.ID
WHERE locations.address != 'null' AND locations.zip != 'null'
GROUP BY locations.zip
HAVING count(locations.name)>=3
ORDER BY locations.zip,no_of_cuisines desc;
```

Query results:

Results Messages

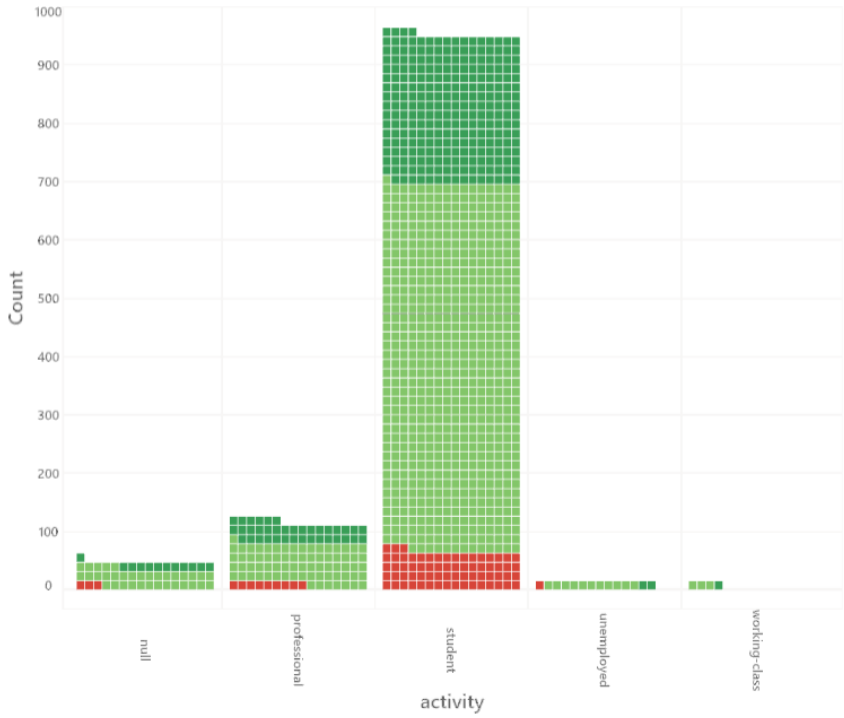
	no_of_cuisines	zip
1	14	78000
2	3	78250
3	3	78269

So, the zip code 78000 has the highest number of multi cuisine restaurants.

Visualizations:-

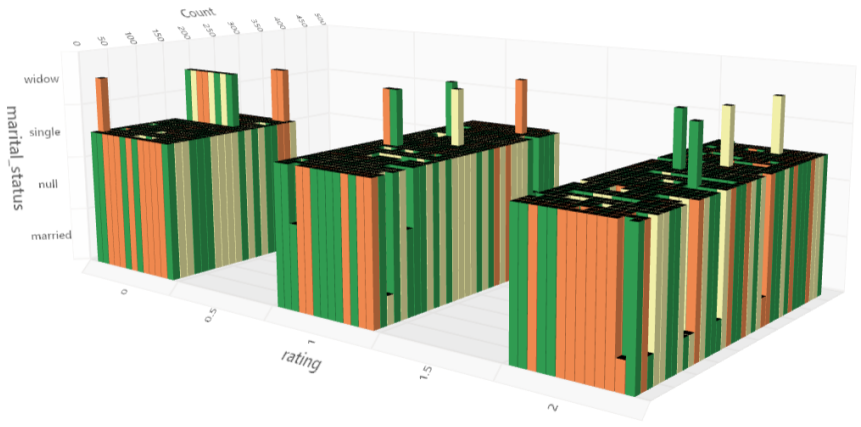
Carried out few visualizations based on restaurant database system.

1. From the below graph we can conclude that most of the customers are Students who prefer going out for dining.



Visualization Picture. 1

2. From the below visualization we can conclude that the restaurants with the highest rating have significantly more number of customers as compared to low rated restaurants.



Picture. 2

References: -

https://www.kaggle.com/datasets/uciml/restaurant-data-with-consumer-ratings

https://www.geeksforgeeks.org/sql-tutorial/?ref=gcse

https://azure.microsoft.com/en-us/products/data-studio/

Presentation Link:

https://cometmail-my.sharepoint.com/:v:/g/personal/sxg210168_utdallas_edu/EfLnklbpiStJvzkYzzv_L28BUeIugjKzP_u87kRIFbm89w