



**Insight.**

## Analysis of Reviews on College Park Restaurants

This proposal contains detailed and relevant  
information regarding the plan and database design  
to better use restaurants' reviews in College Park

**Prepared by:**

Rui Jin,  
Aditya Jadeja,  
Sai Vaishnav Vinjamuri,  
Yun-Jung Fan,  
Ying-Ting Lin

# About INSIGHT. Co & Project Introduction

---

Dining plays an essential role in our daily life, yet sometimes it is hard for people to make decisions about when, where, and what to eat. Therefore, by collecting and integrating the data from major review platforms, such as Google Maps, Yelp, etc., the goal of this project is to design a database, in which insights of restaurants can be provided to dining industry and people using these problems.

## Mission Statement

From the customers' points of view, our value is to provide reliable information about restaurants for them to make better and quicker decisions.

From the restaurants' points of view, we hope we can help them improve by analyzing and getting insight from customers' feedbacks.

## Mission Objectives

For restaurants:

- To find out what customers value the most by checking the reviews coming with the highest and the lowest star rate and review like.
- To find out customers who tend to leave good or bad comments and take the following steps regarding customer relationship.

For customers:

- To compare evaluations about a certain restaurant from different platforms to see if the evaluations are consistent.
- To find out the top ten famous restaurants in each category to help customers to make decision in a timely manner.

# Project Background

---

## Users

Our database user will be mainly divided into two groups: customers and restaurants. The definition for each group is presented as follows:

- Customers: people who go to restaurants and whatever leave or read comments on the review platforms.
- Restaurants: restaurants in College Park, which have been registered on the review platforms.

## Data Sources

The data is collected by using web-scraping with Python. The four main packages applied in this project are Google API, Yelp API, Selenium, and Request. Due to the time limit, only three reviews from each restaurant are collected. The details of the ways to collect, clean, organize, and utilize data are further shown in the Readme file as provided.

# Database Design

## Conceptual Database Design (ER Diagram)

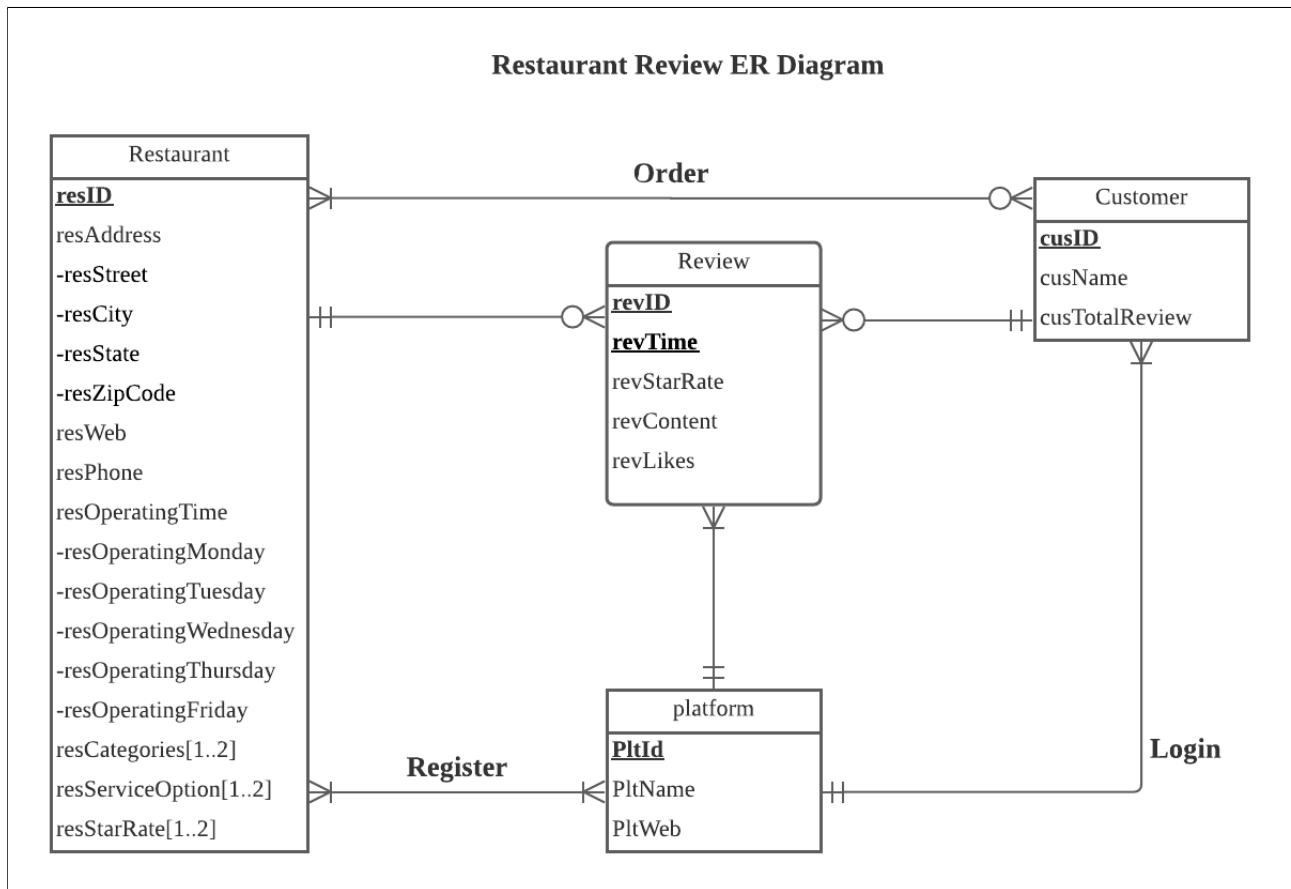


Figure 1: Restaurant Review ER Diagram

The ER diagram of this project consists of four entities: 1) Restaurant; 2) Customer; 3) Review; and 4) platform. The attributes and the relations between entities are shown in Figure 1 above. Restaurant, Customer, and Platform are 3 strong entities, whereas Review are associated entity. As stated in the mission objectives, the focal points of the ER diagram are Review and its cardinality, thus greater efforts are focused on this part from which and the business rule can be better explained.

The relationship between Customer and Review: one customer can have zero or multiple reviews, and one Review is written by one customer only. As for the relation between Restaurant and Review, one review is written for only one restaurant, and one restaurant could have zero or multiple reviews. As for the relation between Platform and Review, one platform could have at least one or more reviews and one review could only show on one platform. Furthermore, there are some details related to the dataset taken into consideration when designing the ER-diagram. For instance, in restaurant entity, reviewstartRate is multivalue attributes since different ratings from different platforms are considered. Another example is that the restaurant total review is not a derived attribute, since rather than collecting from the database, the exact number of reviews from platform dataset can be directly collected.

## Logical Database Design (Relational Schema)

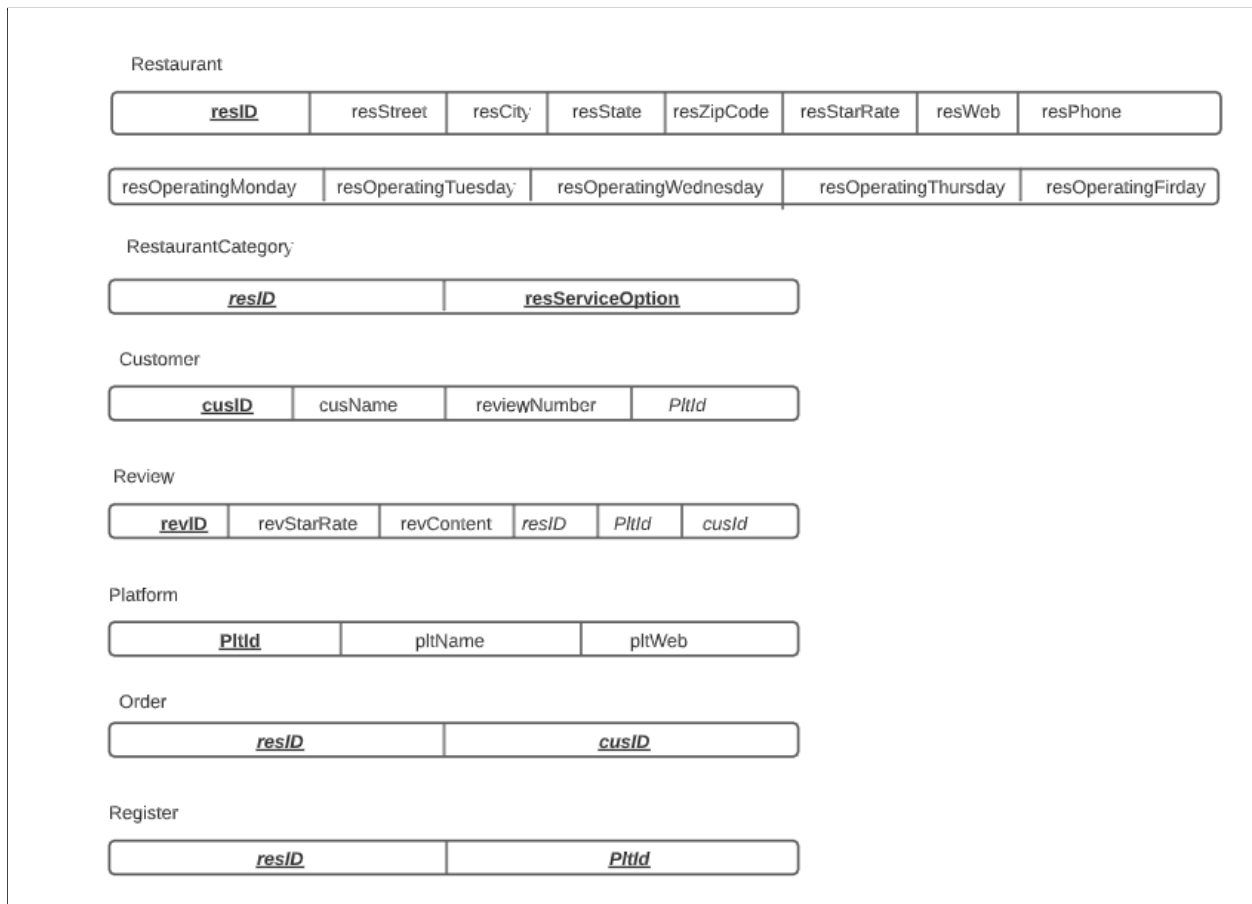


Figure 2: Logical Database Design (Relational Schema)

Figure 2 above illustrates the logical database design (Relational Schema). As has been stated from the ER-diagram that there are four major Entity tables (i.e., Restaurant, Customer, Review, and Platform). Exclusively, there are 3 relations that can be drawn from the ER-diagram, which are Order, Register, and Login. However, as there is “One to Many” Cardinality Relationship between Platform and Customer, the Login table is discarded in the relational schema as there is no need of including it. The relations embedded in the logical database design are presented as follows:

- Firstly, In Restaurant table, resID is the primary key, and resStreet, resCity, and resZipCode are the composite attributes of attribute ‘resAddress’. Similarly, resOperatingMonday, resOperatingTuesday, resOperatingWednesday, resOperatingThursday, resOperatingFriday are composite attributes of attribute ‘resOperatingTime’.
- Secondly, RestaurantCategory comprises composite Primary keys. Additionally, in the customer table, cusID is the primary key and PltId is the foreign key. In the Review table, revID is the primary key with a set of foreign keys consisting of resID, PltId, and cusId. In the Platform table, PltId is the primary key and it does not have foreign key.
- In the relations, Order and Register both consist of a set of composite primary keys in their respective tables, indicating that there are three sets of composite primary key pairs in the Relational Schema.

# Implementing Database

With the development of the database, four business transactions, as examples of using the developed database, are presented by showing the SQL code, SQL result, and how they can be presented in Tableau.

1. "If a restaurant appears in more than 1 platforms, what are the average rating of each restaurant on all platforms and the rating of each restaurant on each platform?"

## SQL Code

Figure 3 below demonstrates the "Select Statement" and "Create View" query for this question, this statement helps the customers in checking the consistency of ratings across platforms and thus help them making better decisions while choosing a restaurant based on ratings.

For "Create View", subquery is used to select from only those restaurants who have reviews on more than one platform, after which the ratings from the different platforms can be compared. Also, the restaurants are also sorted in the descending order according to their average rating across all platforms.

```
--Q1) If a restaurant appears on more than one platforms, what are the average rating of each restaurant on all platforms and the rating of each restaurant on each platform?
CREATE VIEW [Restaurant Rating On Each Platform]
AS
SELECT r.resID, r.resName, avr.[AVG. Rating], t.resRating, p.pltName
FROM [Insight.Restaurant] r, [Insight.Register] g, [Insight.ResRating] t, [Insight.Platform] p, (SELECT t.resID, AVG(t.resRating) AS 'AVG. Rating'
FROM [Insight.ResRating] t
GROUP BY t.resID
HAVING COUNT(t.resID) > 1) avr
WHERE r.resID = g.resID AND g.pltId = t.pltId AND r.resID = t.resID AND g.pltId = p.pltId AND avr.resID = t.resID
GO

SELECT *
FROM [Restaurant Rating On Each Platform] r
ORDER BY r.[AVG. Rating] DESC
```

Figure 3: Screenshot of the SQL code for Transaction 1

## Results from SQL and Tableau

The results for the query, as shown in Figure 4 and Figure 5, indicate the average ratings of restaurants across platforms accompanied by the ratings from each platform.

	resID	resName	AVG. Rating	resRating	pltName
1	R40	QU JAPAN	4.500000	4.5	GoogleMap
2	R40	QU JAPAN	4.500000	4.5	Yelp
3	R29	Marathon Deli	4.300000	4.6	GoogleMap
4	R29	Marathon Deli	4.300000	4.0	Yelp
5	R32	The Board and Brew	4.300000	4.6	GoogleMap
6	R32	The Board and Brew	4.300000	4.0	Yelp
7	R07	Blaze Pizza	4.250000	4.5	GoogleMap
8	R07	Blaze Pizza	4.250000	4.0	Yelp
9	R44	Taqueria Habanero	4.250000	4.5	GoogleMap
10	R44	Taqueria Habanero	4.250000	4.0	Yelp
11	R72	Pizza Roma	4.200000	4.4	GoogleMap
12	R72	Pizza Roma	4.200000	4.0	Yelp
13	R82	Krazi Kebob	4.200000	4.4	GoogleMap
14	R82	Krazi Kebob	4.200000	4.0	Yelp
15	R27	NuVegan Cafe, College Park	4.200000	4.4	GoogleMap

Figure 4: Results of Transaction 1 from SQL

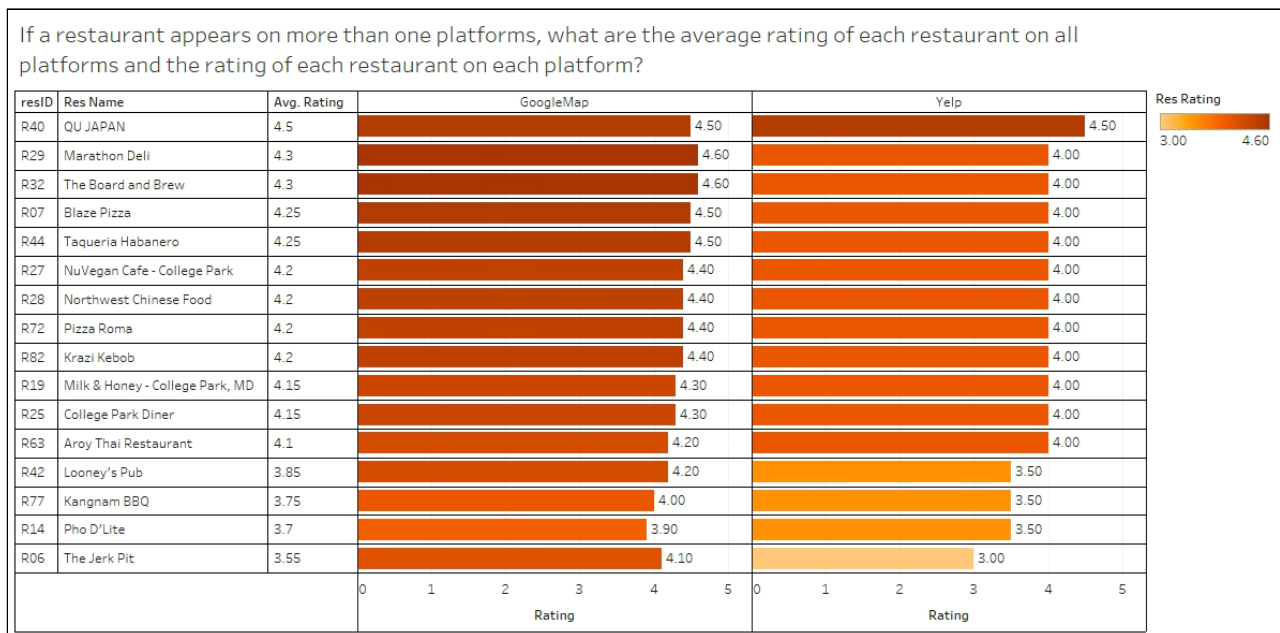


Figure 5: Results of Transaction 1 from Tableau

2. "If there are more than three restaurants in a category (Cuisine), what is the information of the top three restaurants in each of those categories?"

### SQL Code

Figure 6 below shows the "Select Statement" and "Create View" query for this question, this question gives information about the top 3 restaurants in each category (Cuisine) according to their average rating. This can help customers choose restaurants of their liking quickly and more efficiently.

For the "Create View" query for this question, subquery is used to display only those categories which contain more than three restaurants. Also, "Partition by" and "Order by" are used to segregate restaurants according to their categories and then sort them according to their ratings.

```
--Q2) If there are more than three restaurant in the category, what is the information of the top three restaurants in each category?
CREATE VIEW [Top 3 Restaurant In Each Platform] AS
SELECT *
FROM (SELECT T.resCategories,
             ROW_NUMBER() OVER (PARTITION BY T.resCategories ORDER BY T.resStarRate DESC) AS 'rank',
             T.resID, T.resName,
             T.resStarRate,
             T.resStreet,
             T.resCity,
             T.resState,
             T.resZip,
             T.resPhone,
             T.resWeb
      FROM (SELECT r.resID, avg(a.resRating) as resStarRate, c.resCategories, r.resName, r.resCity, r.resPhone, r.resState, r.resStreet, r.resWeb, r.resZip
            FROM [Insight.Restaurant] r, [Insight.RestaurantCategory] c, [Insight.ResRating] a
            WHERE r.resID=c.resID AND r.resID = a.resID AND c.resCategories IN (SELECT c.resCategories
                                     FROM [Insight.RestaurantCategory] c
                                     GROUP BY c.resCategories
                                     HAVING COUNT(c.resCategories)>3)
            GROUP BY r.resID, c.resCategories, r.resName, r.resCity, r.resPhone, r.resState, r.resStreet, r.resWeb, r.resZip
            ) AS T
      ) AS T
WHERE T.rank <= 3
WITH CHECK OPTION
GO
```

Figure 6: Screenshot of the SQL code for Transaction 2

## Results from SQL and Tableau

The result for the query shows the top 3 rated restaurants in each category (Cuisine), as illustrated in Figure 7 and Figure 8.

Results Messages											
	resCategories	rank	resID	resName	resStarRate	resStreet	resCity	resState	resZip	resPhone	resWeb
1	Chinese restaurant	1	R28	Northwest Chinese Food	4.200000	7313 Baltimore Ave suite E	College Park	MD	20740	+1 240-714-4473	http://www.northwestchinesefood.com/
2	Chinese restaurant	2	R51	China Buffet	4.000000	9098 Baltimore Ave	College Park	MD	20740	+1 301-982-6600	NULL
3	Chinese restaurant	3	R55	Panda Express	3.900000	3972 Campus Dr	College Park	MD	20742	+1 301-314-6111	http://www.pandaexpress.com/
4	Fast food restaurant	1	R64	Chick-fil-A	4.300000	Stamp Union Food Court 3972 Campus Dr	College Park	MD	20742	+1 301-314-6568	https://www.chick-fil-a.com/maryland-collegepark
5	Fast food restaurant	2	R41	Taco Bell	4.000000	8428 Baltimore Ave	College Park	MD	20740	+1 301-982-7198	https://locations.tacobell.com/md/college-park/8...
6	Fast food restaurant	3	R71	Taco Bell	3.600000	Glenn L. Martin Hall Campus Dr	College Park	MD	20740	NULL	https://locations.tacobell.com/md/college-park/8...
7	Mexican restaurant	1	R44	Taqueria Habanero	4.250000	8145 Baltimore Ave Suite A&B	College Park	MD	20740	+1 240-241-4486	http://taqueriahabanero.com/
8	Mexican restaurant	2	R21	Chipotle Mexican Grill	4.200000	10250 Baltimore Ave Ste G	College Park	MD	20740	+1 240-387-7140	https://locations.chipotle.com/md/college-park/1...
9	Mexican restaurant	3	R43	Chipotle Mexican Grill	4.000000	7332 Baltimore Ave	College Park	MD	20740	+1 240-582-0015	https://locations.chipotle.com/md/college-park/7...
10	Pizza restaurant	1	R69	Potomac Pizza - College Park	4.300000	7777 Baltimore Ave	College Park	MD	20740	+1 240-582-5242	http://www.potomacpizz.com/
11	Pizza restaurant	2	R07	Blaze Pizza	4.250000	7419 Baltimore Ave	College Park	MD	20740	+1 301-392-2324	http://www.blazepizza.com/locations/college-park/
12	Pizza restaurant	3	R17	Ledo Pizza	4.200000	4509 Knox Rd	College Park	MD	20740	+1 301-422-8122	https://order.ledopizza.com/menu/ledo-pizza-colle...
13	Restaurant	1	R01	Nene's Restaurant	4.700000	9905 Rhode Island Ave	College Park	MD	20740	+1 301-254-9537	https://www.nenesrestaurant.com/
14	Restaurant	2	R30	Caribbean Cove DC	4.700000	9348 Cherry Hill Rd	College Park	MD	20740	+1 240-542-4965	http://caribbeancovecatering.com/
15	Restaurant	3	R47	Ritchie's Colombian Restaurant	4.700000	7313 Baltimore Ave suite H	College Park	MD	20740	+1 240-764-7422	https://ritchiescolombian.wixsite.com/mysite
Query executed successfully. doitsqlx.rhsmith.umd.edu,97... BUDT703_Student_253 (166) BUDT703_Project_0507.06 00:00:00 18 rows											

Figure 7: Results of Transaction 2 from SQL

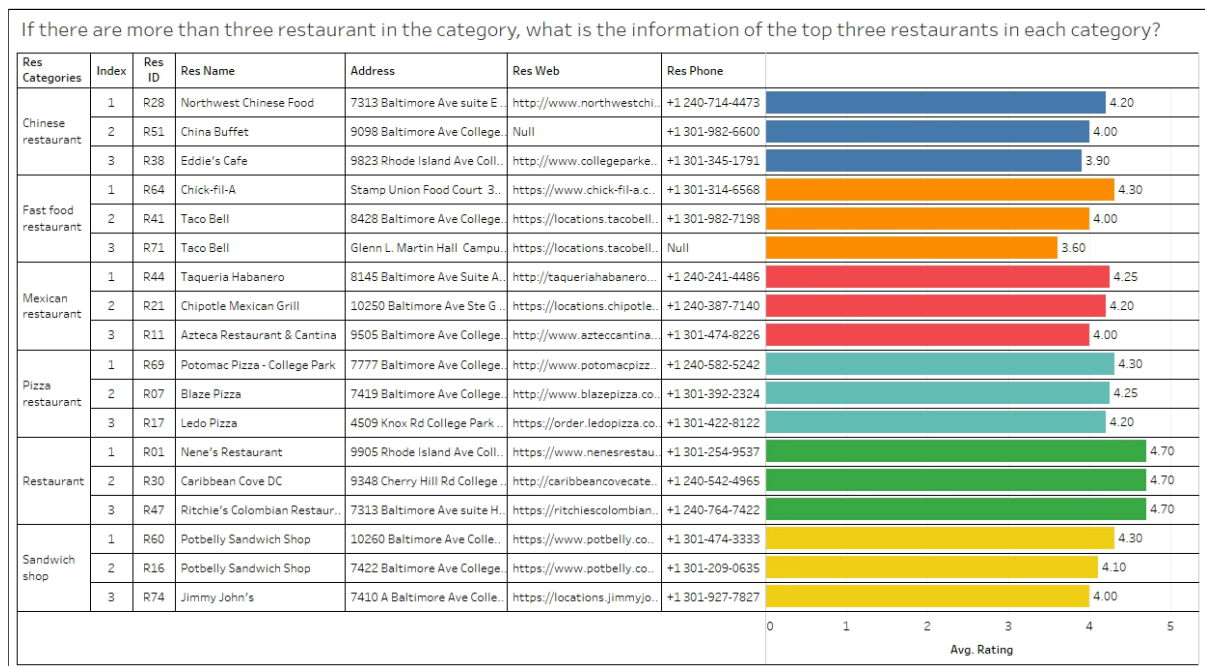


Figure 8: Results of Transaction 2 from Tableau

3. "Who are the customers that leave more than one review for restaurants in College Park and what are their average star rating and average number of Review likes?"

## SQL Code

Figure 9 below shows the "Select Statement" and "Create View" query for this question. This question gives information about the customers who leave the greatest number of reviews and the average number of review likes for those customers. This shows which customers are more active on the platforms.

In the "Create View" query for this question, subquery is also used to display only those customers whose id occurs more than once (i.e., customers with more than one review).



Additionally, those customers are sorted in the descending order according to their number of reviews.

```
CREATE VIEW [Customers with more than 1 review ] AS
SELECT c.cusID, c.cusName, avr.[Number of Reviews], avr.[AVG. Star Rate], ROUND(AVG (CAST (r.revLike AS float)),2) as 'AVG. Review Likes'
FROM [Insight.Customer] c, (select r.cusID, count(r.cusID) as 'Number of Reviews', ROUND(AVG(r.revStarRate),2) as 'AVG. Star Rate'
FROM [Insight.Review] r
GROUP BY r.cusID
HAVING count(r.cusID) > 1
) avr
WHERE c.cusID = r.cusID AND c.cusID = avr.cusID
GROUP BY c.cusID, c.cusName, avr.[Number of Reviews], avr.[AVG. Star Rate]
GO
```

Figure 9: Screenshot of the SQL code for Transaction 3

### Results from SQL and Tableau

The results for the query show the customers who have the highest number of reviews, as demonstrated in Figure 10 and Figure 11.

	cusID	cusName	Number of Reviews	AVG. Star Rate	AVG. Review Likes
1	C005	Krishna Bhavithavya Kidambi	7	4.140000	0.57
2	C108	Ronak Patel	4	4.750000	0.25
3	C174	Jasmin Hollywood	3	4.670000	0
4	C056	Brent L	2	4.000000	1
5	C081	Chris W.	2	5.000000	0.5
6	C104	Jack Bug	2	4.000000	0

doitsqlx.rhsmith.umd.edu,97... | BUDT703\_Student\_253 (63) | BUDT703\_Project\_0507\_06 | 00:00:00 | 6 rows

Figure 10: Results of Transaction 3 from SQL

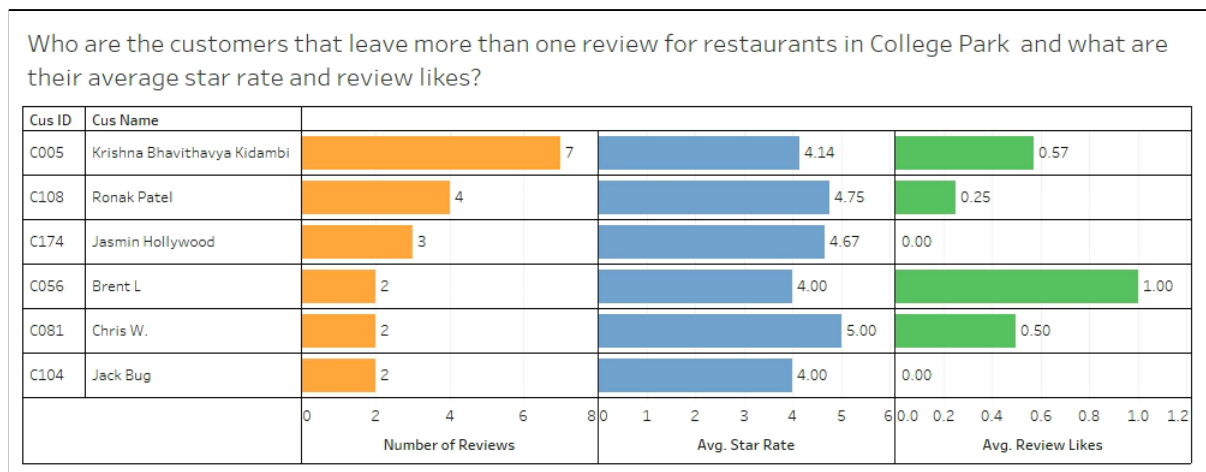


Figure 11: Results of Transaction 3 from Tableau

#### 4. "What are the reviews that come with the highest review like?"

##### SQL Code

Figure 12 below shows the "Select Statement" and "Create View" query for this question. This question gives information about reviews with the highest number of likes. This can help the customers in finding out the most helpful reviews when selecting restaurants. Similarly,



for the “Create View” query for this question, subquery, Partition by, and Order by are used to segregate according to restaurant id and sort by number of reviews.

```
CREATE VIEW [Reviews with the highest review like ] AS
SELECT res.resID, res.resName, T.revContent, T.revStarRate, T.revLike
FROM (
    SELECT *,
        row_number() over(partition by T.resID order by T.revLike desc) as rn
    FROM dbo.[Insight.Review] as T
    ) T, [Insight.Restaurant] res
WHERE T.rn = 1 and T.revLike != 0 and res.resID = T.resID
WITH CHECK OPTION
GO

SELECT *
FROM [Reviews with the highest review like]
ORDER BY revLike DESC
```

Figure 12: Screenshot of the SQL code for Transaction 4

### Results from SQL and Tableau

The results for the query show the restaurant reviews with highest number of likes, as shown in Figure 13 and Figure 14.

Results

Messages

	resID	resName	revContent	revStarRate	revLike	
1	R08	KOITÉ GRILL	Let me first say, that we did not visit this restaurant location. But, we stoppe...	5.0	8	<div> <div></div> <div></div> </div>
2	R17	Ledo Pizza	First and foremost, this is NO longer the original Ledo(s). This is now a chai...	2.0	5	
3	R11	Azteca Restaurant & Cantina	Delicious, inexpensive food! The food is great! The drinks are a little on the...	5.0	4	
4	R01	Nene's Restaurant	I drove by this place several times during the pandemic in 2020 but didn't st...	5.0	3	
5	R25	College Park Diner	Clean place, friendly, good food and service. You won't be disappointed.	5.0	3	
6	R52	LaTao Hotpot College park	Lovely place and delicious food. Very entertaining to cook the food yoursel...	5.0	3	
7	R54	Krazi Burrito	Decided to give them a try and we loved it. Staff was very friendly and wel...	4.0	3	
8	R58	Paisano's	Love there food. Would eat here often before COVID-19 and teleworking. I...	5.0	3	
9	R68	Noodles and Company	If you want to having dining hall food for \$10, this is your place. The bad re...	1.0	3	
10	R06	The Jerk Pit	There are a rack (a lot) of Jamaican culinary spots in the DC area. I've defi...	5.0	3	
11	R73	sweetgreen	The audacity for this location to charge \$2 to "add hot roasted sweet potat...	2.0	2	
12	R55	Panda Express	Made online purchase and went to the store. It was closed an hour earlier. ...	1.0	2	
13	R49	Mamma Lucia	Amazing Italian food if you call it what a joke I ordered veggie egg plant par...	1.0	2	
14	R32	The Board and Brew	Friendly service and fun concept with the board game playing option!	4.0	2	
15	R42	Leona's Pub	One of my favorite spots to hang out with college friends. It's a college hangout...	5.0	2	

Query executed successfully.

doitsqlx.rhsmith.umd.edu,97...

BUDT703\_Student\_253 (67)

BUDT703\_Project\_0507\_06

00:00:00

41 rows

Figure 13: Results of Transaction 4 from SQL

What are the reviews that come with the highest review like?

Res ID	Res Name	Rev Content	Rev Star Rate	Rev Like	Restaurant Name
R08	KOITÉ GRILL	Let me first say, that we did not visit this restaurant location. But, we stopped by Koite Grill Stand at RFK Stadium Open Air Farmers' Market in DC and we ordered Lamb, Chicken with JOLLOF Rice, Purple Tea and a upside down Pineapple ...	5.000	8.000	(All)
R17	Ledo Pizza	First and foremost, this is NO longer the original Ledo(s). This is now a chain Ledo(s) restaurant, plain and simple. Now if that's your thing, good on you. But those of us who patronized the original and have eaten at the chains, know the ...	2.000	5.000	
R11	Azteca Restaurant & Cantina	Delicious, inexpensive food! The food is great! The drinks are a little on the strong side (for me), but still good! The service here was a little slow even though there weren't many people in there, but they may have had a lot of carry out ...	5.000	4.000	
R01	Nene's Restaurant	I drove by this place several times during the pandemic in 2020 but didn't stop in until 2021. Now I'm hooked ...a regular, so to speak. Delicious Nigeria food cooked to perfection. I enjoyed the spicy goat meat with Jollof rice and ...	5.000	3.000	
R06	The Jerk Pit	There are a rack (a lot) of Jamaican culinary spots in the DC area. I've definitely reviewed a couple of the more popular eateries in the area...except this one @jerkpit!! Located at 9078 Baltimore Ave., College Park M.D., this place can be ...	5.000	3.000	
R25	College Park Diner	Clean place, friendly, good food and service. You won't be disappointed.	5.000	3.000	
R52	LaTao Hotpot College park	Lovely place and delicious food. Very entertaining to cook the food yourself, mix the sauces and experiment. The staff is very friendly and quick. They also have karaoke rooms, in case you love to sing and want to spend couple of hours ...	5.000	3.000	

Figure 14: Results of Transaction 4 from Tableau

5. "What are the reviews that come with the highest review like?"

### SQL Code

Figure 15 below shows the "Select Statement" and "Create View" query for this question. This question gives information about the restaurants which have the highest number of reviews. This can help customers in finding the most popular/trending restaurants. For the "Create View" query for this question, we use subquery to get the restaurants with multiple reviews and the sort them according to the number of reviews.

```
CREATE VIEW [Restaurants with highest number of reviews] AS
SELECT r.resID, r.resName, rev.[Number of Reviews]
FROM [Insight.Restaurant] r, (select r.resID, count(r.resID) as 'Number of Reviews'
FROM [Insight.Review] r
GROUP BY r.resID
HAVING count(r.resID) > 1
) rev
WHERE r.resID = rev.resID
GROUP BY r.resID, r.resName, rev.[Number of Reviews]
GO

SELECT *
FROM [Restaurants with highest number of reviews]
Order by [Number of Reviews] DESC
```

Figure 15: Screenshot of the SQL code for Transaction 5

### Results from SQL

The results for the query show the restaurants with the highest number of reviews, as demonstrated in Figure 16.

Results		Messages	
	resID	resName	Number of Reviews
1	R06	The Jerk Pit	6
2	R07	Blaze Pizza	6
3	R14	Pho D'Lite	6
4	R19	Milk & Honey - College Park, MD	6
5	R25	College Park Diner	6
6	R27	NuVegan Cafe - College Park	6
7	R28	Northwest Chinese Food	6
8	R29	Marathon Deli	6
9	R32	The Board and Brew	6
10	R40	QU JAPAN	6
11	R42	Looney's Pub	6
12	R44	Taqueria Habanero	6
13	R63	Aroy Thai Restaurant	6
14	R72	Pizza Roma	6
15	R77	Korean BBQ	6

✓ Query executed successfully.

Figure 16: Results of Transaction 5 from SQL