

# **Problem Statement**



In a competitive market like restaurant, understanding the factor that influence business success is crucial for stakeholders.

This project aims to investigate the relationship between user engagement and business success matrices for restaurant.

The primary goal of this project is to identify the best locations to establish a new restaurant. Additionally, we aim to determine peak hours and levels of customer engagement at these locations, enabling us to make informed arrangements and optimize our operations accordingly.

## ***Data source:- Yelp***

Yelp is a popular online platform that connects people with local businesses. Founded in 2004 by Jeremy Stoppelman and Russel Simmons, Yelp provides a space where users can read and write reviews about local businesses, including restaurants, bars, shops, and more.

**Link:** <https://www.yelp.com/dataset/download>

**Kaggle Download link:** <https://www.kaggle.com/datasets/yelp-dataset/yelp-dataset>

## **Data Overview**

1. This dataset is a subset of Yelp and contains information about businesses in eight metropolitan areas in the USA and Canada.
2. The original data is provided by Yelp in JSON format.
3. The five JSON files include business, review, user, tip, and check-in data.
4. These JSON files are stored in a database for easy data retrieval.

## **Hypothesis**

1. Higher level of user engagement (more reviews, more tips and check-in's) correlate with higher review counts and ratings of restaurant.
2. Consistent engagement over time is positively associated with sustained business success for restaurants.
3. A high number of reviews and higher review ratings are directly correlated with the success of a restaurant.
4. High number of tips and check-ins also indicate a restaurant's success.

## **Analysis and Findings**

- Out of approximately 150k businesses, around 52k are restaurants. However, only 35k of these restaurants are currently open.

```
count(*)  
0      150346
```

**Total Businesses**

```
52286 rows × 1 columns
```

**Restaurant Business**

```
35004 rows × 1 columns
```

**Currently working**

## • Top 20 Locations for Establishing a New Restaurant:

	city	state	longitude	latitude	avg_rating	review_count	restaurant_count	success_score
0	Philadelphia	PA	-75.155564	39.955505	3.585200	495404	3527	47.013195
1	New Orleans	LA	-90.074427	29.950647	3.764009	386992	1392	48.428344
2	Nashville	TN	-86.768170	36.208102	3.571089	260980	1681	44.539342
3	Tampa	FL	-82.502346	27.890814	3.626972	236140	1965	44.873566
4	Tucson	AZ	-111.010760	32.338572	3.460134	192649	1643	42.105090
5	Indianapolis	IN	-86.127217	39.637133	3.482143	186614	1904	42.262082
6	Reno	NV	-119.789339	39.476117	3.584586	151460	798	42.757246
7	Santa Barbara	CA	-119.690672	34.414445	3.913677	127068	446	45.995434
8	Saint Louis	MO	-90.407187	38.583223	3.502088	125866	958	41.124949
9	Boise	ID	-116.206275	43.611192	3.597015	55551	603	39.297657
10	Edmonton	AB	-113.604288	53.436403	3.511912	48476	1553	37.889478
11	Clearwater	FL	-82.798901	27.965747	3.608945	44768	436	38.649170
12	Saint Petersburg	FL	-82.749065	27.806517	3.663333	39370	300	38.760942
13	St. Louis	MO	-90.221942	38.612495	3.630807	38146	409	38.302116
14	Metairie	LA	-90.183650	29.994939	3.383621	33363	348	35.241197
15	Sparks	NV	-119.754857	39.535111	3.404167	32698	240	35.386652
16	Franklin	TN	-86.802720	35.956533	3.466667	30718	300	35.819807
17	St. Petersburg	FL	-82.727831	27.796953	3.756917	27800	253	38.443883
18	Brandon	FL	-82.307700	27.937370	3.453744	24669	227	34.928903
19	Wilmington	DE	-75.625673	39.732089	3.504566	23975	438	35.342879

we have identified the top 20 cities where restaurants are the most successful. Success is measured by average review ratings, the total number of reviews, and the number of restaurants in each city.

Among these, the top 5 cities stand out:

1. **New Orleans, LA:** With an average rating of 3.76 and 386,992 reviews, this city has a high success score of 48.43.
2. **Philadelphia, PA:** This city has the highest number of reviews (495,404) for its 3,527 restaurants, giving it a success score of 47.01, despite a slightly lower average rating of 3.59.
3. **Santa Barbara, CA:** Known for high customer satisfaction, Santa Barbara has the highest average rating of 3.91, with a success score of 45.99.
4. **Nashville, TN:** With 260,980 reviews and an average rating of 3.57, Nashville earns a success score of 44.54.
5. **Tampa, FL:** Tampa has an average rating of 3.63 and 236,140 reviews, resulting in a success score of 44.87.

These cities are great places to start a new restaurant due to their high customer engagement and positive reviews.

- The table below shows the business success matrix:

	0
avg_review_count	44.866561
min_review_count	5.000000
max_review_count	7568.000000
median_review_count	15.000000
avg_star	3.596724
min_star	1.000000
max_star	5.000000
median_star	3.500000

- The table below shows the restaurant with the highest rating:

	name	review_count	avg_rating
0	ā café	48	5.0
1	two birds cafe	77	5.0
2	the brewers cabinet production	13	5.0
3	taqueria la cañada	17	5.0
4	la bamba	44	5.0
5	la 5th av tacos	24	5.0
6	el sabor mexican and chinese food	21	5.0
7	eat.drink.Om...YOGA CAFE	7	5.0
8	d4 Tabletop Gaming Cafe	8	5.0
9	cabbage vegetarian cafe	12	5.0

- The table below shows the restaurant with the highest reviews:

	name	review_count	avg_rating
0	McDonald's	16490	1.868702
1	Chipotle Mexican Grill	9071	2.381757
2	First Watch	8688	3.896552
3	Acme Oyster House	8343	4.000000
4	Taco Bell	8017	2.141813
5	Chick-fil-A	7943	3.381410
6	Oceana Grill	7400	4.000000
7	Buffalo Wild Wings	6810	2.347458
8	Panera Bread	6613	2.661905
9	Hattie B's Hot Chicken - Nashville	6093	4.500000
10	Domino's Pizza	6091	2.290210
11	Wendy's	5930	2.030159
12	Ruby Slipper Cafe	5800	4.166667
13	Chili's	5744	2.514706
14	Reading Terminal Market	5721	4.500000
15	Applebee's Grill + Bar	5515	2.534314
16	IHOP	5513	2.533708
17	Outback Steakhouse	5317	2.967391
18	Ruby Slipper - New Orleans	5193	4.500000
19	Mother's Restaurant	5185	3.500000
20	Los Agaves	5160	4.125000
21	Texas Roadhouse	5070	3.428571
22	Royal House	5070	4.000000
23	Olive Garden Italian Restaurant	4949	3.066667
24	Burger King	4890	2.027869

Higher ratings do no guarantee a higher review count, or vice versa.

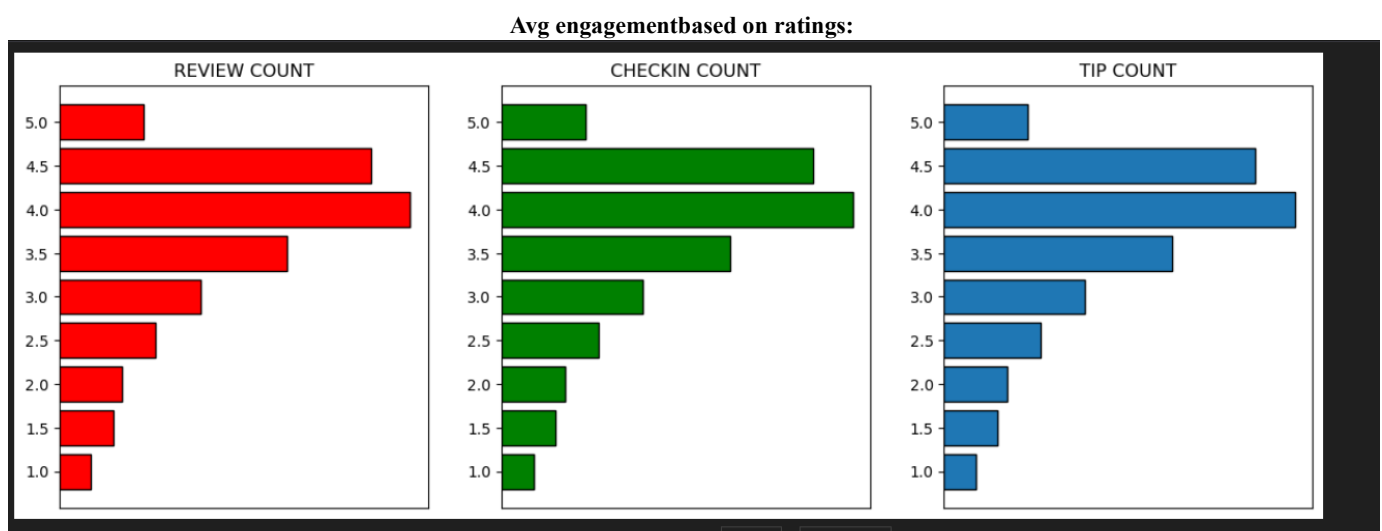
Success of Restaurants is not solely determined by ratings or review counts.

Review count reflects user engagement but not necessarily overall customer satisfaction or business performance

## # Checking whether restaurants with higher engagement tend to have higher ratings:

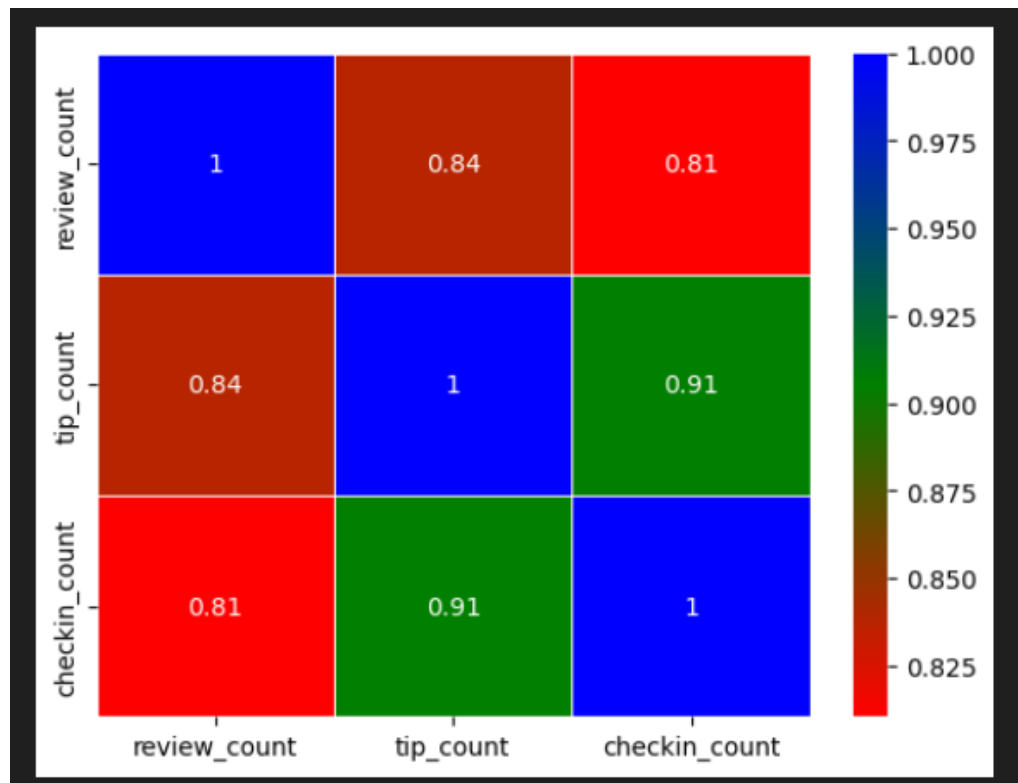
	rating	avg_review_count	avg_checkin_count	avg_tip_count
0	1.0	14.365079	17.518072	2.781513
1	1.5	24.358459	34.480969	3.884654
2	2.0	28.677613	53.553242	4.696448
3	2.5	43.694356	92.919049	7.170221
4	3.0	63.903960	140.232167	10.522143
5	3.5	103.351377	208.645312	15.710198
6	4.0	159.216587	310.735849	23.209835
7	4.5	141.149835	214.915277	19.190083
8	5.0	38.101562	37.269621	5.472222

- Data shows a general increase in average review, check-in, and tip counts as ratings improve from 1 to 4 stars.
- Restaurants rated 4 stars exhibit the highest engagement and shows a downward trend for rating above 4.
- The drop in engagement at 5.0 stars might suggest either a saturation point where fewer customers feel compelled to add their reviews, or a selectivity where only a small, satisfied audience frequents these establishments.



- **The relationship between the number of reviews, tip count, and check-ins is being analyzed to understand the correlation between these engagement metrics.**

	review_count	tip_count	checkin_count	avg_rating
review_count	1.000000	0.837270	0.810453	0.168707
tip_count	0.837270	1.000000	0.905120	0.151168
checkin_count	0.810453	0.905120	1.000000	0.106538
avg_rating	0.168707	0.151168	0.106538	1.000000



These correlations suggest that user engagement across different platforms (reviews, tips, and check-ins) is interlinked; higher activity in one area tends to be associated with higher activity in others.

Businesses should focus on strategies that boost all types of user engagement, as increases in one type of engagement are likely to drive increases in others.

- **Difference in user engagement between high-rated ,Avg rated and low-rated businesses:**

The data shows that businesses with higher ratings have more user engagement, such as more reviews, tips, and check-ins.

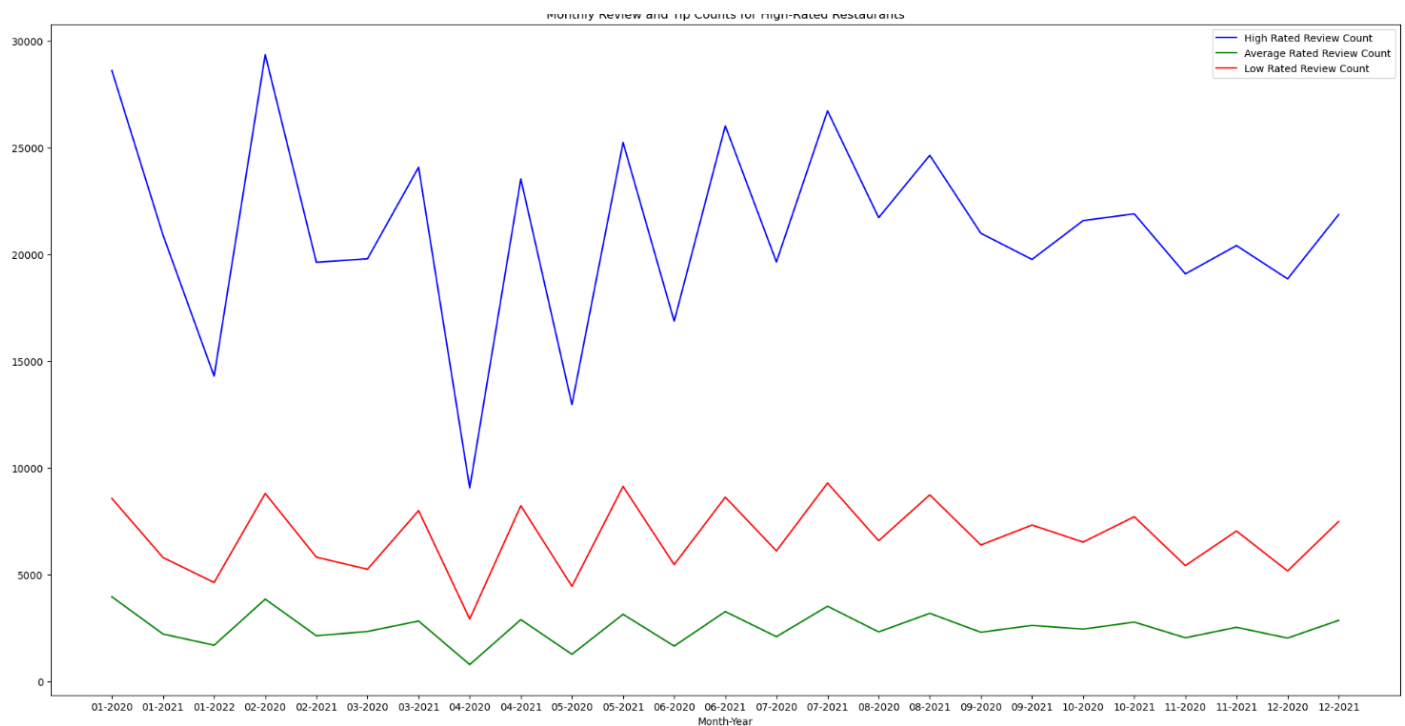
This trend shows that providing good service and quality is important because it leads to more reviews, check-ins, and tips, which are key indicators of customer engagement and satisfaction.

```
category
HIGH_RATED    16400
AVG_RATED     11905
LOW_RATED      6699
Name: count, dtype: int64
```

```
R_df.groupby('category')[['review_count', 'checkin_count', 'tip_count']]
```

	review_count	checkin_count	tip_count
category			
AVG_RATED	88.205292	182.447743	13.765272
HIGH_RATED	144.033110	256.972090	20.774834
LOW_RATED	34.574116	67.987630	5.705924

- Monthly review and tip counts for restaurants:**

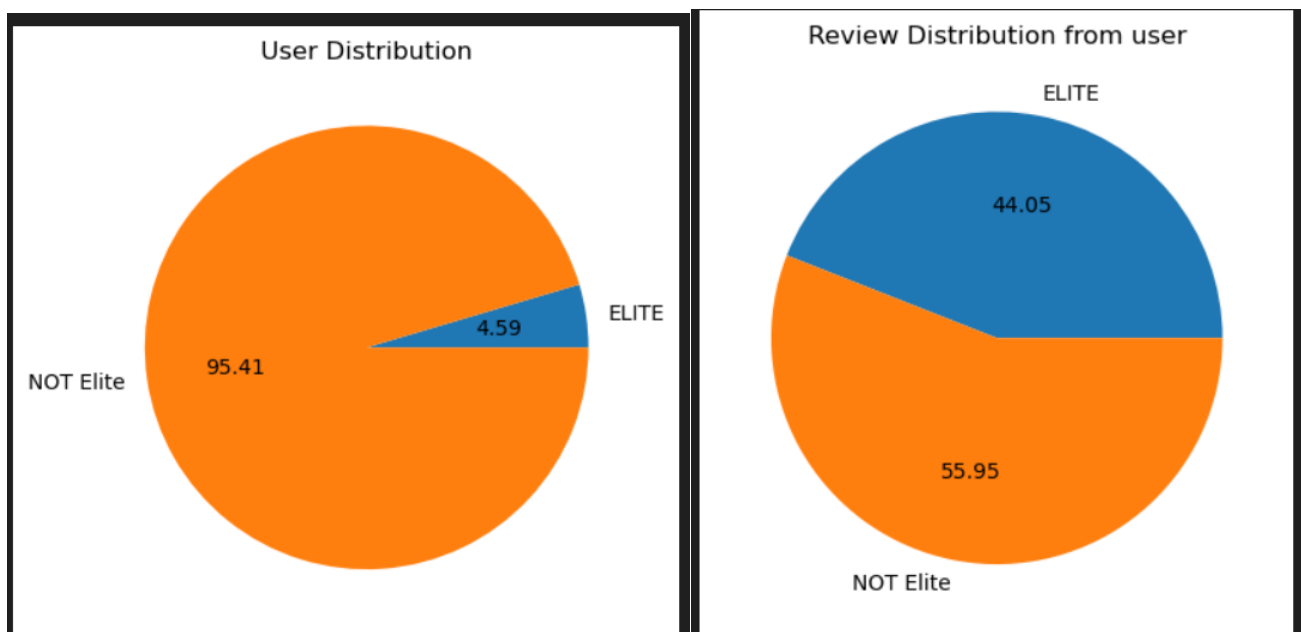




- **Elite users :**

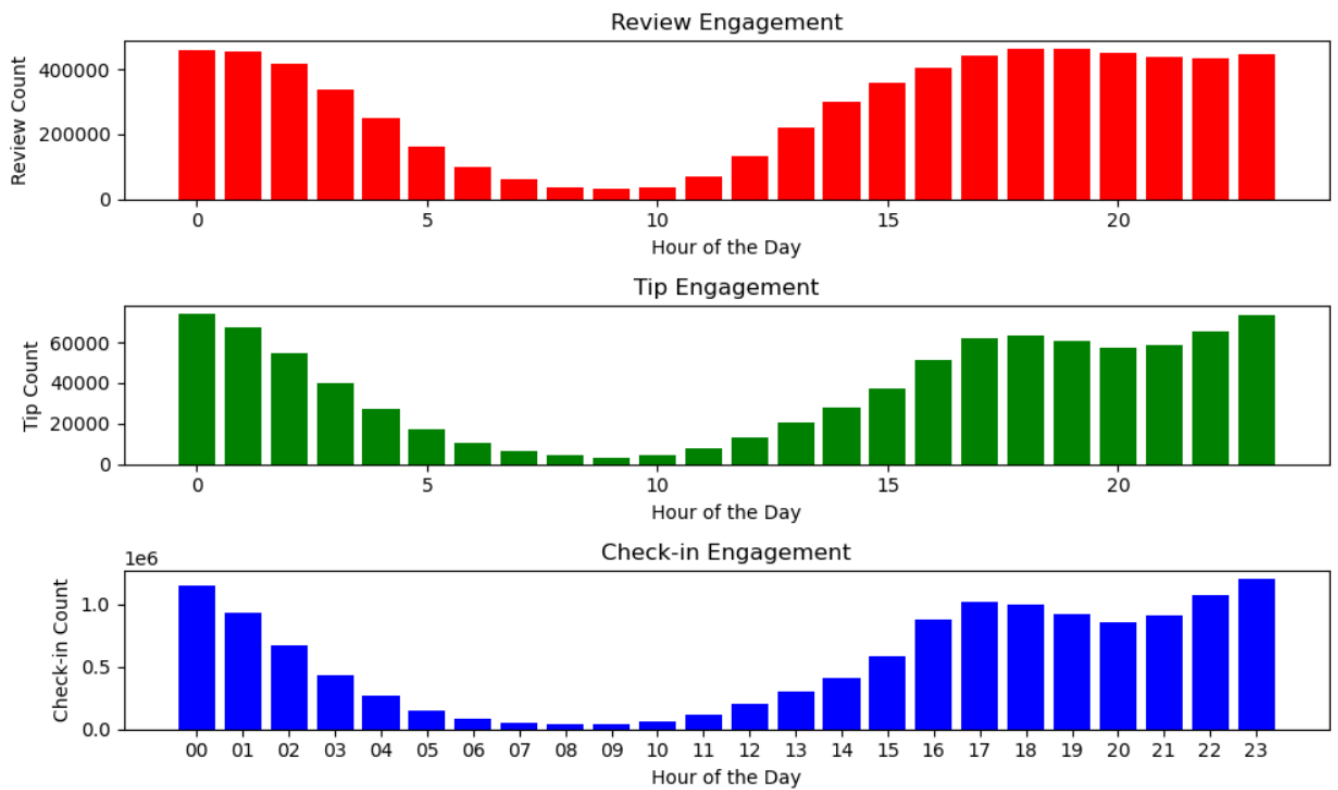
- Elite users are individuals who have been recognized and awarded the "Elite" status by Yelp for their active and high-quality contributions.
- Elite users, despite being significantly fewer in number, contribute a substantial proportion of the total review count compared to non-elite users.
- Establishing a positive relationship with elite users can lead to repeat visits and loyalty, as they are more likely to continue supporting businesses they have had good experiences with.

	elite	row_count	total_review_count
0	ELITE	91198	20484441
1	NOT Elite	1896699	26021235



- **Rush hours:**

- The busiest hours for restaurants, based on user engagement, span from 4 pm to 1 am.
- Knowing the peak hours allows businesses to optimize their staffing levels and resource allocation during these times to ensure efficient operations and quality service delivery.



## Findings

### 1. Top Cities for Restaurant Success:

- **New Orleans, LA:** High success score of 48.43 with an average rating of 3.76 and 386,992 reviews.
- **Philadelphia, PA:** Success score of 47.01 with the highest number of reviews (495,404) and an average rating of 3.59 across 3,527 restaurants.
- **Santa Barbara, CA:** Highest average rating of 3.91 and a success score of 45.99.
- **Nashville, TN:** Success score of 44.54 with 260,980 reviews and an average rating of 3.57.
- **Tampa, FL:** Success score of 44.87 with an average rating of 3.63 and 236,140 reviews.

### 2. Peak Hours:

- The busiest hours for restaurants are from 4 pm to 1 am.
- Knowing peak hours helps businesses optimize staffing and resources during these times.

### 3. Engagement and Ratings:

- Higher engagement metrics are observed in businesses with higher ratings.
- Providing good service and quality is important as it leads to more reviews, check-ins, and tips.