



Dataset Overview

The dataset contains detailed information about restaurants globally, including location, cuisine types, pricing, ratings, and service features such as online delivery and table booking.



Problem Statement

1. Problem:

- The global restaurant industry lacks clear insights into what drives customer satisfaction and performance across different markets, cuisines, price points, and service offerings.
- Operators and investors cannot determine which cities offer growth potential, whether service investments (online delivery, table booking) improve ratings, or how pricing and cuisine variety impact success.

2. Business Impact:

- This intelligence gap leads to capital misallocation, poor expansion decisions, suboptimal pricing strategies, and service investments without proven ROI.
- Stakeholders miss high-value opportunities, enter saturated markets, and cannot identify why certain restaurants outperform competitors in similar conditions.

3. Expected Solution:

- Data-driven analysis will reveal actionable patterns in cuisine preferences, pricing optimization, service adoption impact, and geographic opportunities across multiple countries.
 - The insights will enable evidence-based decisions for market entry, pricing strategy, service investments, and identification of high-performing vs underperforming market segments.
-



Tools

Python – Data cleaning, EDA, feature engineering

PostgreSQL – Business intelligence queries and analysis

Power BI – Interactive dashboard & visualization.

```

8
9 /* Q1: What are the top 5 most popular cuisines by restaurant count,
10      and what is their average rating?*/
11 select
12   cuisines,
13   count(restaurant_id) as total_restaurants,
14   round(avg(rating)::numeric,2) as avg_ratings
15   from restaurant
16   group by cuisines
17   order by total_restaurants desc limit 5;
18

```

Data Output Messages Notifications

	cuisines	total_restaurants	avg_ratings
	text	bigint	numeric
1	North Indian	936	1.67
2	North Indian, Chine...	511	2.42
3	Fast Food	354	2.12
4	Chinese	354	2.04
5	North Indian, Mugh...	334	2.89

```

19
20 /* Q2: Identify the top 10 cities by restaurant count and
21      calculate each city's average rating and total votes.*/
22 select
23   city,
24   count(restaurant_id) as total_restaurants,
25   round(avg(rating)::numeric,2) as avg_ratings,
26   sum(votes) as total_votes
27   from restaurant
28   group by city
29   order by total_restaurants desc limit 10;
30

```

Data Output Messages Notifications

	city	total_restaurants	avg_ratings	total_votes
	text	bigint	numeric	numeric
1	New Delhi	5473	2.44	628340
2	Gurgaon	1118	2.65	132160
3	Noida	1080	2.04	73488
4	Faridabad	251	1.87	6486
5	Ghaziabad	25	2.85	2366
6	Lucknow	21	4.20	9130
7	Guwahati	21	4.19	4964
8	Bhubaneshw...	21	3.98	4243
9	Amritsar	21	3.69	3665
10	Ahmedabad	21	4.16	12266

```

32 /* Q3: What is the distribution of restaurants across price ranges and
33     price categories, and what percentage does each represent? */
34 select
35     price_range,
36     price_category,
37     count(restaurant_id) as restaurants,
38     Round((count(restaurant_id)/sum(count(restaurant_id))over())*100,2) as restaurant_percentage
39     from restaurant
40     group by price_range,price_category
41     order by price_range;

```

Data Output Messages Notifications

The screenshot shows a SQL query interface with various icons at the top. The main area displays a table with four rows of data. The columns are labeled: price_range, price_category, restaurants, and restaurant_percentage. The data shows the following distribution:

	price_range	price_category	restaurants	restaurant_percentage
1	1	Cheap	4444	46.53
2	2	Moderate	3113	32.59
3	3	Expensive	1408	14.74
4	4	Luxury	586	6.14

```

43 /* Q4: Calculate the percentage of restaurants offering online delivery,
44     table booking, and both services. */
45 with totals as(
46     select
47         count(*) as total_restaurants
48         from restaurant
49 )
50 select
51     'Online Delivery' as category,
52     count(*) as restaurants,
53     round((count(*)*100.0/(select total_restaurants from totals)),2) as percentage
54     from restaurant
55     where online_delivery = 'Yes'
56 UNION ALL
57 select
58     'Table Booking' as category,
59     count(*) as restaurants,
60     round((count(*)*100.0/(select total_restaurants from totals)),2) as percentage
61     from restaurant
62     where table_booking = 'Yes'
63 UNION ALL
64 select
65     'Both Services' as category,
66     count(*) as restaurants,
67     round((count(*)*100.0/(select total_restaurants from totals)),2) as percentage
68     from restaurant
69     where online_delivery = 'Yes' and table_booking = 'Yes';

```

	category text	restaurants bigint	percentage numeric
1	Online Delivery	2451	25.66
2	Table Booking	1158	12.12
3	Both Services	435	4.55

```

72 /*Q5: What is the average number of votes
73 per restaurant in each rating category?*/
74 select
75 rating_text,
76 count(restaurant_id) as restaurant_count,
77 round(avg(votes)) as avg_votes
78 from restaurant
79 group by rating_text
80 order by avg_votes desc;

```

Data Output Messages Notifications

≡+ 📁 ⏮ 🗂️ ⏮ 🗑️ 📁 SQL

	rating_text text	restaurant_count bigint	avg_votes numeric
1	Excellent	301	852
2	Very Good	1079	520
3	Good	2100	229
4	Poor	186	91
5	Average	3737	48
6	Not rated	2148	1

```

87 /*Q6: Compare average ratings and median votes between restaurants with
88 and without online delivery.*/
89 select
90 'Online Delivery' as delivery_mode,
91 count(*) as total_restaurants,
92 round(avg(rating)::numeric,2) as avg_rating,
93 percentile_cont(0.5) within group(order by votes) as median_votes
94 from restaurant
95 where online_delivery = 'Yes'
96 UNION ALL
97 select
98 'No-Online Delivery' as delivery_mode,
99 count(*) as total_restaurants,
100 round(avg(rating)::numeric,2) as avg_rating,
101 percentile_cont(0.5) within group(order by votes) as median_votes
102 from restaurant
103 where online_delivery = 'No';

```

Data Output Messages Notifications

≡+ 📁 ⏮ 🗂️ ⏮ 🗑️ 📁 SQL

	delivery_mode text	total_restaurants bigint	avg_rating numeric	median_votes double precision
1	Online Delivery	2451	3.25	79
2	No-Online Deliv...	7100	2.47	18

```

106 -- Q7: Find the top 3 cities in each country ranked by average rating.
107 with top_city as (
108 select
109   country,
110   trim(city) as city,
111   avg(rating) as avg_rating,
112   row_number()over(partition by country order by avg(rating) desc) as rn
113   from restaurant
114   group by country, city
115 )
116 select
117   country, city,
118   round(avg_rating::numeric,1) as avg_rating
119   from top_city where rn <= 3;

```

Data Output Messages Notifications

	country text	city text	avg_rating numeric
1	Australia	Beechworth	4.6
2	Australia	Tanunda	4.4
3	Australia	Palm Cove	4.4
4	Brazil	Rio de Janeiro	4.3
5	Brazil	São Paulo	3.5
6	Brazil	Brasília	3.5
7	Canada	Vineland Stati...	4.3
8	Canada	Chatham-Kent	3.7
9	Canada	Waterloo	3.7

Total rows: 38 Query complete 00:00:00.069

```

122 /*Q8: What is the relationship between num_cuisines (variety) and average rating?
123   Do multi-cuisine restaurants perform better?*/
124 select
125   count(restaurant_name) as total_restaurants,
126   num_cuisines,
127   round(avg(rating)::numeric,2) as avg_rating,
128   case
129   when num_cuisines>=5 and avg(rating)>=3 then 'strong positive'
130   when num_cuisines>=5 and avg(rating)<3 then 'weak positive'
131   when num_cuisines between 3 and 4 and avg(rating)>=3 then 'moderate positive'
132   when num_cuisines<3 and avg(rating)>=3 then 'specialised strong'
133   else 'no relation'
134   end as relation
135   from restaurant
136   group by num_cuisines order by avg_rating desc;

```

Data Output Messages Notifications

	total_restaurants bigint	num_cuisines bigint	avg_rating numeric	relation text
1	28	7	3.67	strong positive
2	74	6	3.67	strong positive
3	164	5	3.51	strong positive
4	584	4	3.42	moderate posit...
5	14	8	3.42	strong positive
6	1840	3	3.04	moderate posit...
7	3444	2	2.69	no relation
8	3403	1	2.23	no relation

```

139 /* Q9: Calculate the average cost difference between price categories and
140     identify which category offers the best rating-to-cost ratio.*/
141 select
142     price_range,
143     price_category,
144     round(avg(avg_cost)::numeric,1)as avg_cost,
145     round(avg(rating)::numeric,2) as avg_rating,
146     round((avg(avg_cost)/avg(rating))::numeric,2) as rating_to_cost_ratio
147     from restaurant
148     group by price_range, price_category
149     order by price_range;

```

Data Output Messages Notifications

	price_range bigint	price_category text	avg_cost numeric	avg_rating numeric	rating_to_cost_ratio numeric
1	1	Cheap	137.6	2.00	68.80
2	2	Moderate	298.4	2.94	101.46
3	3	Expensive	2588.6	3.68	702.79
4	4	Luxury	924.3	3.82	242.09

```

152 /*Q10: Identify restaurants that are priced above their city's average cost
153     but have ratings above 4.0 (premium performers).*/
154 with cte as (
155     select
156         restaurant_name,
157         city,
158         avg_cost AS restaurant_cost,
159         round(avg(avg_cost) over(partition by city)) as city_avg_cost,
160         rating
161     from restaurant
162 )
163 select
164     *
165     from cte
166     where restaurant_cost > city_avg_cost and rating > 4.0;
167

```

Data Output Messages Notifications

	restaurant_name text	city text	restaurant_cost double precision	city_avg_cost double precision	rating double precision
1	Punjab Grill	Abu Dhabi	165	91	4.9
2	Tamba	Abu Dhabi	250	91	4.7
3	P.F. Chang's	Abu Dhabi	125	91	4.2
4	The Cheesecake Factory	Abu Dhabi	100	91	4.6
5	Cho Gao - Crowne Plaza Abu Dhabi	Abu Dhabi	175	91	4.4
6	Olive Garden	Abu Dhabi	115	91	4.1
7	Famous Dave's Barbecue	Abu Dhabi	130	91	4.6
8	Denny's	Abu Dhabi	95	91	4.6

Total rows: 554 Query complete 00:00:00.139

```

174 /*Q11: Rank restaurants within each city by votes and
175 identify the top 3 most popular restaurants per city with their key metrics.*/
176 with top_rest as (
177 select
178 restaurant_name, city, votes, rating, avg_cost,
179 num_cuisines, online_delivery, table_booking,
180 row_number() over(partition by city order by votes desc) as ranks
181 from restaurant
182 )
183 select
184 *
185 from top_rest
186 where ranks <=3;
187

```

Data Output Messages Notifications

	restaurant_name text	city text	votes bigint	rating double precision	avg_cost double precision	num_cuisines bigint	online_delivery text	table_booking text	ranks bigint
1	The Cheesecake Factory	Abu Dhabi	586	4.6	100	2	No	No	1
2	Via Delhi	Abu Dhabi	525	4	50	3	Yes	No	2
3	Pizza Di Rocco	Abu Dhabi	471	4.4	75	2	Yes	Yes	3
4	Pinch Of Spice	Agra	177	4.2	500	3	No	No	1
5	Pind Balluchi	Agra	175	3.7	450	2	No	No	2
6	Thaaliwala	Agra	168	4.1	350	2	No	No	3
7	650 - The Global Kitchen	Ahmedabad	1582	4.2	450	6	No	No	1
8	Patang - The Revolving Restaurant	Ahmedabad	1315	3.7	900	3	No	No	2

Total rows: 323 Query complete 00:00:00.190

CRLF L

```

199 -- Q12: Find cities that have either above-average restaurant density OR above-average rating.
200 with top_city as (
201 select
202 city,
203 count(restaurant_name) as restaurant_density,
204 avg(rating) as avg_rating
205 from restaurant
206 group by city
207 ),
208 avgs as(
209 select
210 avg(restaurant_density) as avg_density,
211 avg(avg_rating) as avg_city_rating
212 from top_city
213 )
214 select
215 city,
216 restaurant_density,
217 round(avg_rating) as rating
218 from top_city
219 where restaurant_density > (select avg_density from avgs)
220 or avg_rating >=(select avg_city_rating from avgs)
221 order by restaurant_density desc;

```

	city	restaurant_density	rating
		bigint	double precision
1	New Delhi	5473	2
2	Gurgaon	1118	3
3	Noida	1080	2
4	Faridabad	251	2
5	Ahmedabad	21	4
6	Bhubaneshwar	21	4
7	Guwahati	21	4
8	Lucknow	21	4
9	Goa	20	4
10	Des Moines	20	4

Total rows: 86 Query complete 00:00:00.128

```

224 -- Q13: Analyze service adoption: does price_range correlate with the likelihood of offering table_booking and online_delivery services?
225 select
226 price_range,
227 price_category,
228 count(*) as total_restaurants,
229 sum(case when table_booking = 'Yes' then 1 else 0 end) as table_booking_count,
230 sum(case when online_delivery = 'Yes' then 1 else 0 end) as online_delivery_count,
231 concat(round(sum(case when table_booking = 'Yes' then 1 else 0 end)*100.0/count(*),1),'%') as table_booking_pct,
232 concat(round(sum(case when online_delivery = 'Yes' then 1 else 0 end)*100.0/count(*),1),'%') as online_delivery_pct
233 from restaurant
234 group by price_range,price_category
235 order by price_range;
236

```

Data Output Messages Notifications

	price_range	price_category	total_restaurants	table_booking_count	online_delivery_count	table_booking_pct	online_delivery_pct
	bigint	text	bigint	bigint	bigint	text	text
1	1	Cheap	4444	1	701	0.0%	15.8%
2	2	Moderate	3113	239	1286	7.7%	41.3%
3	3	Expensive	1408	644	411	45.7%	29.2%
4	4	Luxury	586	274	53	46.8%	9.0%

```

238 /*Q14: Identify underperforming opportunities: cities with high restaurant counts(>20)
239 but average ratings below 3.5, grouped by price category */
240 with cte as (
241 select
242 city,
243 price_category,
244 count(restaurant_name) as restaurant_counts,
245 round(avg(rating)::numeric,2) as avg_rating
246 from restaurant
247 group by city,price_category
248 )
249 select
250 *
251 from cte
252 where restaurant_counts > 20 and avg_rating < 3.5;

```

Data Output Messages Notifications

	city	price_category	restaurant_counts	avg_rating
	text	text	bigint	numeric
1	New Del...	Moderate	1656	2.84
2	Gurgaon	Moderate	440	2.84
3	Gurgaon	Cheap	453	2.01
4	New Del...	Cheap	3052	1.96
5	Faridab...	Moderate	86	2.29
6	Noida	Moderate	401	2.38
7	Noida	Luxury	21	2.71
8	Noida	Expensive	93	3.18

Total rows: 11 Query complete 00:00:00.141