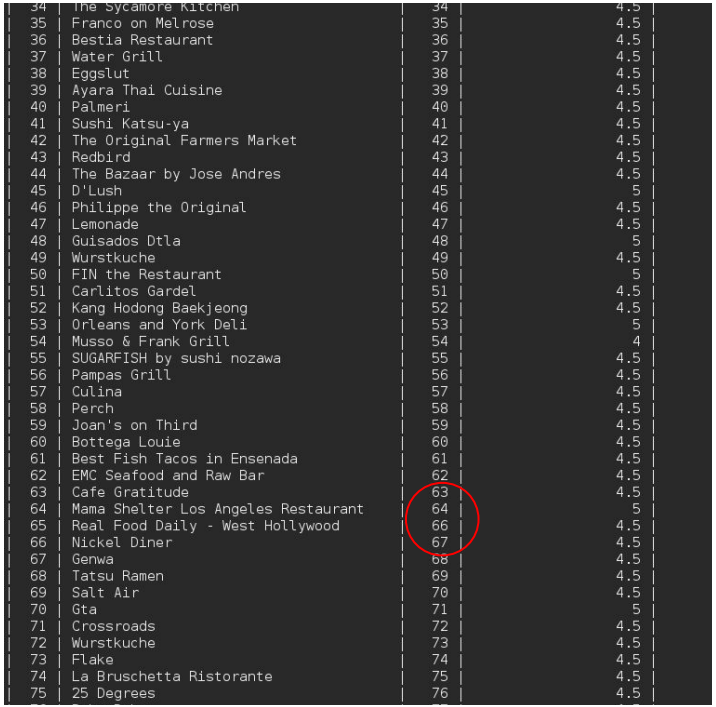
Data Discovery: Parser Project

Tripadvisor Restaurant Ratings

SELECT * FROM restaurant ORDER BY rank;



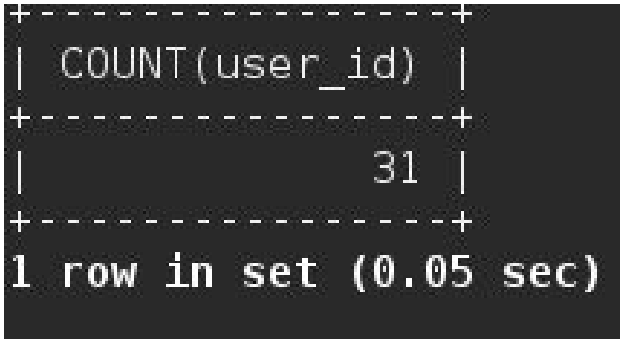
34	The Sycamore Kitchen	34	4.5	
35	Franco on Melrose	35	4.5	
36	Bestia Restaurant	36	4.5	
37	Water Grill	37	4.5	
38	Eggslut	38	4.5	
39	Ayara Thai Cuisine	39	4.5	
40	Palmeri	40	4.5	
41	Sushi Katsu-ya	41	4.5	
42	The Original Farmers Market	42	4.5	
43	Redbird	43	4.5	
44	The Bazaar by Jose Andres	44	4.5	
45	D'Lush	45	5	
46	Philippe the Original	46	4.5	
47	Lemonade	47	4.5	
48	Guisados Dtla	48	5	
49	Wurstkuche	49	4.5	
50	FIN the Restaurant	50	5	
51	Carlitos Gardel	51	4.5	
52	Kang Hodong Baekjeong	52	4.5	
53	Orleans and York Deli	53	5	
54	Musso & Frank Grill	54	4	
55	SUGARFISH by sushi nozawa	55	4.5	
56	Pampas Grill	56	4.5	
57	Culina	57	4.5	
58	Perch	58	4.5	
59	Joan's on Third	59	4.5	
60	Bottega Louie	60	4.5	
61	Best Fish Tacos in Ensenada	61	4.5	
62	EMC Seafood and Raw Bar	62	4.5	
63	Cafe Gratitude	63	4.5	
64	Mama Shelter Los Angeles Restaurant	64	5	
65	Real Food Daily - West Hollywood	66	4.5	
66	Nickel Diner	67	4.5	
67	Genwa	68	4.5	
68	Tatsu Ramen	69	4.5	
69	Salt Air	70	4.5	
70	Gta	71	5	
71	Crossroads	72	4.5	
72	Wurstkuche	73	4.5	
73	Flake	74	4.5	
74	La Bruschetta Ristorante	75	4.5	
75	25 Degrees	76	4.5	

- It seems that there is no relation between the average amount of stars a restaurant gets and the ranking it holds.

- There is no restaurant ranked 65th

Variations in review per user

```
SELECT COUNT(user_id)
FROM (
      SELECT user_id, COUNT(DISTINCT review_stars) ct
      FROM review
      GROUP BY user_id) c
WHERE ct > 1;
```



```
+------------------+
| COUNT(user_id)   |
+------------------+
|               31 |
+------------------+
1 row in set (0.05 sec)
```

In changing the number for the comparison, we find that *no user* has given out more than two amounts of stars for their reviews.

Perhaps our users are very lazy, or only ever visit restaurants that are pretty good.

Our users seem to be creatures of habit:

Out of the 928 reviewers present in the data, only 31 of them have given out at least two different amounts of stars for their reviews.

And this may be the case...

It seems that our users only eat at good restaurants.

93.68% of all reviews are at least 4 stars.

```
SELECT review_stars, COUNT(review_stars)
FROM review
GROUP BY review_stars;
```

```
+--------------+--------------------+
| review_stars | COUNT(review_stars) |
+--------------+--------------------+
|            1 |                142 |
|            2 |                350 |
|            3 |               1836 |
|            4 |               7971 |
|            5 |              26541 |
+--------------+--------------------+
5 rows in set (0.03 sec)

MariaDB [restaurants]> SELECT 26541 / 36840 + 7971 / 36840;
+-----------------------------+
| 26541 / 36840 + 7971 / 36840 |
+-----------------------------+
|                      0.9368 |
+-----------------------------+
1 row in set (0.00 sec)

MariaDB [restaurants]>
```