Data Scientist Role Play: Profiling and Analyzing the Yelp Dataset Coursera Worksheet

This is a 2-part assignment. In the first part, you are asked a series of questions that will help you profile and understand the data just like a data scientist would. For this first part of the assignment, you will be assessed both on the correctness of your findings, as well as the code you used to arrive at your answer. You will be graded on how easy your code is to read, so remember to use proper formatting and comments where necessary.

In the second part of the assignment, you are asked to come up with your own inferences and analysis of the data for a particular research question you want to answer. You will be required to prepare the dataset for the analysis you choose to do. As with the first part, you will be graded, in part, on how easy your code is to read, so use proper formatting and comments to illustrate and communicate your intent as required.

For both parts of this assignment, use this "worksheet." It provides all the questions you are being asked, and your job will be to transfer your answers and SQL coding where indicated into this worksheet so that your peers can review your work. You should be able to use any Text Editor (Windows Notepad, Apple TextEdit, Notepad ++, Sublime Text, etc.) to copy and paste your answers. If you are going to use Word or some other page layout application, just be careful to make sure your answers and code are lined appropriately. In this case, you may want to save as a PDF to ensure your formatting remains intact for you reviewer.

Part 1: Yelp Dataset Profiling and Understanding

- 1. Profile the data by finding the total number of records for each of the tables below:
- i. Attribute table :
- --to find the total number of data in the attribute table select * from attribute;

Output : 10000

Output : 10000

ii. Business table :

⁻⁻to find the total number of data in the business table select * from business;

```
iii. Category table :
--to find the total number of data in the category table
     select * from category;
Output : 10000
iv. Checkin table :
--to find the total number of data in the checkin table
     select * from checkin;
Output : 10000
v. elite years table :
--to find the total number of data in the elite years table
   select * from elite years;
Output : 10000
vi. friend table :
--to find the total number of data in the friend table
   select * from friend;
Output : 10000
vii. hours table :
--to find the total number of data in the hours table
     select * from hours;
Output : 10000
```

viii. photo table:

--to find the total number of data in the photo table select * from photo;

Output : 10000

ix. review table :

--to find the total number of data in the review table select * from review;

Output : 10000

x. tip table :

--to find the total number of data in the tip table select * from tip;

Output : 10000

xi. user table :

--to find the total number of data in the tip table
 select * from tip;

Output : 10000

2. Find the total distinct records by either the foreign key or primary key for each table. If two foreign keys are listed

```
in the table, please specify which foreign key.
i. Business:
--to find the total number of distinct data in the Business
table
   select distinct * from Business;
Output : 10000
ii. Hours:
--to find the total number of distinct data in the hours table
   select distinct * from hours;
Output : 10000
iii. Category :
--to find the total number of distinct data in the Category
table
   select distinct * from Category;
Output : 10000
iv. Attribute:
--to find the total number of distinct data in the Attribute
table
   select distinct * from Attribute;
Output : 10000
v. Review:
--to find the total number of distinct data in the Review
   select distinct * from Review;
Output : 10000
vi. Checkin:
--to find the total number of distinct data in the Checkin
   select distinct * from Checkin;
Output : 10000
vii. Photo:
--to find the total number of distinct data in the Photo table
   select distinct * from Photo;
Output : 10000
viii. Tip:
--to find the total number of distinct data in the Tip table
   select distinct * from Tip;
Output : 10000
ix. User :
```

--to find the total number of distinct data in the User table

select distinct * from User;

Output : 10000

x. Friend:

--to find the total number of distinct data in the Friend table

select distinct * from Friend;

Output : 10000

xi. Elite years:

--to find the total number of distinct data in the Elite_years table

select distinct * from Elite_years;

Output : 10000

Note: Primary Keys are denoted in the ER-Diagram with a yellow key icon.

3. Are there any columns with null values in the Users table? Indicate "yes," or "no."

Answer: Yes

SOL code used to arrive at answer:

 $\mbox{--to}$ find the description of the datatypes for the user table

pragma table info('user');

4. For each table and column listed below, display the smallest (minimum), largest (maximum), and average (mean) value for the following fields:

i. Table: Review, Column: Stars

min:1 max:5 avg:3.7082

ii. Table: Business, Column: Stars

min:1.0 max: 5.0 avg: 3.6549

iii. Table: Tip, Column: Likes

min: 0 max: 2 avg:0.0144

iv. Table: Checkin, Column: Count

min: 1 max: 53 avg: 1.9414

v. Table: User, Column: Review_count

min:0 max:2000 avg:24.2995

5. List the cities with the most reviews in descending order:

SQL code used to arrive at answer:
 select city, sum(review_count) as reviews
 from business
 group by city
 order by sum(review_count) desc;

Copy and Paste the Result Below:

city	reviews l
l Las Vegas	++ 82854
l Phoenix	l 34503 l
l Toronto	24113
Scottsdale	l 20614 l
Charlotte	l 12523 l
l Henderson	l 10871 l
I Tempe	l 10504 l
l Pittsburgh	l 9798 l
l Montréal	9448
l Chandler	8112
l Mesa	l 6875 l
Gilbert	l 6380 l
l Cleveland	l 5593 l
l Madison	l 5265 l
Glendale	l 4406 l
l Mississauga	l 3814 l
l Edinburgh	l 2792 l
l Peoria	l 2624 l
l North Las Vegas	l 2438 l
l Markham	l 2352 l

6. Find the distribution of star ratings to the business in the following cities:

i. Avon

SQL code used to arrive at answer:

Copy and Paste the Resulting Table Below (2 columns — star rating and count):

+	iting	Count_of_stars	+
	1.5		†
1	2.5	2	
1	3.5	3	
1	4.0	2	
1	4.5	1	
I	5.0	1	
+			+

ii. Beachwood

SQL code used to arrive at answer:

```
select stars as Rating, count(stars) as
Count_of_stars
from business b
where city = 'Beachwood'
group by stars
```

Copy and Paste the Resulting Table Below (2 columns — star rating and count):

+	+
l Rating	Count_of_stars
+	++
2.0	1 1
1 2.5	1
1 3.0	2
1 3.5	2 1
4.0	1 1
1 4.5	2 1
5.0	I 5 I
+	+

7. Find the top 3 users based on their total number of reviews:

```
SQL code used to arrive at answer:
    select name, review_count as reviews
    from user
    order by review_count desc
    limit 3;
```

Copy and Paste the Result Below:

+		+		+
Ì	name	ĺ	reviews	I
+		+		+
1	Gerald	Ī	2000	
	Sara	Ī	1629	
1	Yuri		1339	
Δ.		ь.		
-		-		т

8. Does posing more reviews correlate with more fans?

Please explain your findings and interpretation of the results:

More reviews and fans are not correlated. Below representation will explain my stand:-

-SQL Query to find the fans and review counts and then sorting the fans in descending order:-

Output table :-

++	· 	+-		+
I name I	review_count	1	fans l	
++ Amy	 609		 503	F
Mimi	968		497 l	
Harald	1153		311	
Gerald	2000		253 l	
Christine	930		173 l	
Lisa	813		159 l	
Cat	377		133 l	
William	1215		126 l	
l Fran l	862		124 l	
l Lissa l	834		120 l	
l Mark l	861		115 l	
l Tiffany l	408		111	
bernice	255		105 l	
l Roanna l	1039		104 l	
l Angela l	694		101	
l .Hon l	1246	1	101 l	
l Ben l	307		96 l	
l Linda l	584	1	89 l	
Christina	842		85 l	
l Jessica l	220		84	
l Greg l	408		81 l	
l Nieves l	178		80 l	
l Sui l	754		78 I	
l Yuri l	1339		76 I	
Nicole	161		73 l	
++		+-		H

(Output limit exceeded, 25 of 10000 total rows shown)

In this we see Amy with review count 603 has 503 fans while that of Harald has 1153 review count yet gas 311 fans.

9. Are there more reviews with the word "love" or with the word "hate" in them?

Answer: There are more review for the word "Love".

SQL code used to arrive at answer:

```
SELECT 'love' as Word, COUNT(text) as Count_of_text
FROM review
WHERE text LIKE '%love%'
UNION
SELECT 'hate'as Word, COUNT(text) as Count_of_text
FROM review
WHERE text LIKE '%hate%';
```

Output table :-

İ	Word	İ	Count_of_text	İ
 	hate love	 	232 1780	i I

10. Find the top 10 users with the most fans:

```
SQL code used to arrive at answer:
    select name, fans
    from user
    order by fans desc
    limit 10;
```

Copy and Paste the Result Below:

+	+-		+
l name		fans	Ī
+	+-		+
l Amy		503	
l Mimi		497	I
l Harald		311	
Gerald		253	
Christine		173	
l Lisa		159	
l Cat		133	
l William		126	
l Fran		124	
l Lissa		120	

+----+

11. Is there a strong relationship (or correlation) between having a high number of fans and being listed as "useful" or "funny?" Out of the top 10 users with the highest number of fans, what percent are also listed as "useful" or "funny"?

Key:

0% - 25% - Low relationship 26% - 75% - Medium relationship

76% - 100% - Strong relationship

SQL code used to arrive at answer:
select name, fans, useful, funny
 from user
 order by fans desc, useful desc, funny desc
 limit 10;

Copy and Paste the Result Below:

		+		4		+	
						funny	
+	+		+		+		+
l Amy		503		3226		2554	
l Mimi		497		257		138	
Harald		311		122921		122419	
Gerald	1	253		17524		2324	1
Christine	1	173		4834		6646	1
l Lisa	1	159		48		13	1
l Cat	-	133	Ī	1062		672	
William		126	Ī	9363		9361	
l Fran				9851		7606	1
l Lissa		120		455		150	1
+	-+		+-		+		-+

Please explain your findings and interpretation of the results:

It is not necessary that you have a high fan base to be funny or useful. In the result we see that Amy with 503 fans have been pointed by 2k+ as funny and 3k+ as useful.

Whilst Harald with a fan base of 311 has been pointed by 12k+ as funny and 12k+ as useful.

Part 2: Inferences and Analysis

- 1. Pick one city and category of your choice and group the businesses in that city or category by their overall star rating. Compare the businesses with 2-3 stars to the businesses with 4-5 stars and answer the following questions. Include your code.
- i. Do the two groups you chose to analyze have a different distribution of hours?

The 4-5 star business group seems to have shorter hours of business then the 2-3 star group. As the sample size was quite small so further analysing couldn't be predicted.

ii. Do the two groups you chose to analyze have a different number of reviews?

This is a yes no answer question. Let me tell you why so? one of the 4-5 star business had lots of reviews while the other in the same group had some reviews near about being same to that of 2-3 star business group.

iii. Are you able to infer anything from the location data provided between these two groups? Explain.

I am unable to infer anything from the business location since they are all from different areas.

SQL code used for analysis:

```
SELECT B.name,
                  B.review count,
                  H.hours,
                  postal code,
                  CASE
                      WHEN hours LIKE "%monday%" THEN 1
                      WHEN hours LIKE "%tuesday%" THEN 2
                      WHEN hours LIKE "%wednesday%" THEN 3
                      WHEN hours LIKE "%thursday%" THEN 4
                      WHEN hours LIKE "%friday%" THEN 5
                      WHEN hours LIKE "%saturday%" THEN 6
                      WHEN hours LIKE "%sunday%" THEN 7
                  END AS ord,
                  CASE
                      WHEN B.stars BETWEEN 2 AND 3 THEN '2-3
stars'
                      WHEN B.stars BETWEEN 4 AND 5 THEN '4-5
stars'
                  END AS star rating
```

FROM business B INNER JOIN hours H
ON B.id = H.business_id
INNER JOIN category C
ON C.business_id = B.id
WHERE (B.city == 'Las Vegas'
AND
C.category LIKE 'shopping')
AND
(B.stars BETWEEN 2 AND 3
OR
B.stars BETWEEN 4 AND 5)
GROUP BY stars,ord
ORDER BY ord,star_rating ASC;

2. Group business based on the ones that are open and the ones that are closed. What differences can you find between the ones that are still open and the ones that are closed? List at least two differences and the SQL code you used to arrive at your answer.

i. Difference 1:

Review is an important aspect. The business which had a bad or low review are now closed while the good reviews are now running well.

ii. Difference 2:

The user rating for the business should also be considered as they need to be good enough to keep the business up and running.

3. For this last part of your analysis, you are going to choose the type of analysis you want to conduct on the Yelp dataset and are going to prepare the data for analysis.

Ideas for analysis include: Parsing out keywords and business attributes for sentiment analysis, clustering businesses to find commonalities or anomalies between them, predicting the overall star rating for a business, predicting the number of

fans a user will have, and so on. These are just a few examples to get you started, so feel free to be creative and come up with your own problem you want to solve. Provide answers, in-line, to all of the following:

- i. Indicate the type of analysis you chose to do:

 predicting if the business will stay open or close.
- ii. Write 1-2 brief paragraphs on the type of data you will need for your analysis and why you chose that data:

To better help businesses understand the importance of different factors which will help their business stay open. Some data that may be important like number of reviews, star rating of business, hours open, and of course location. We will gather the city, state, postal_code, and address to make processing easier later on. Categories and attributes will be used to better distinguish between different types of businesses. `is_open` will determine which business is open and which business have closed (not hours) but permanently.

iii. Output of your finished dataset: id l name address l city | state | pin_code long | count_of_review | review_stars | monday_hours | tuesday_hours | wednesday_hours | thursday_hours | friday_hours | saturday_hours | sunday_hours I categories I attributes l is_open l +-----+-----+-----+----+-----+----+-----______

```
| -0DET7VdEQ0JVJ_v6klEug | Flaming Kitchen
3235 York Regional Road 7 | Markham
                                          I ON I L3R 3P9
| 43.8484 | -79.3487 |
                                   25 I
                                                 3.0 |
12:00-23:00 | 12:00-23:00 | 12:00-23:00
                                              12:00-23:00
| 12:00-23:00 | 12:00-23:00
                             | 12:00-23:00
                                              I Asian
Fusion, Restaurants
RestaurantsTableService,GoodForMeal,Alcohol,Caters,HasTV,Resta
urantsGoodForGroups,NoiseLevel,WiFi,RestaurantsAttire,Restaura
ntsReservations,OutdoorSeating,RestaurantsPriceRange2,BikePark
ing, RestaurantsDelivery, Ambience, RestaurantsTakeOut, GoodForKid
s.BusinessParkina
1 |
1 -2HjuT4yjLZ3b5f_abD87Q | Freeman's Car Stereo
4821 South Blvd
                         | Charlotte
                                          | NC | 28217
| 35.1727 | -80.8755 |
                                                 3.5 L
                                    8
9:00-19:00 | 9:00-19:00 | 9:00-19:00
                                              9:00-19:00
| 9:00-19:00 | 9:00-17:00 | None
Electronics, Shopping, Automotive, Car Stereo Installation
BusinessAcceptsCreditCards, RestaurantsPriceRange2, BusinessPark
ing, Wheelchair Accessible
       1 |
| -CdstAUdEvci8GeJG8owpQ | Motors & More
                                                         2315 Highland Dr
                           l Las Vegas
                                           | NV | 89102
| 36.1465 | -115.167 |
                                    7 I
                                                 5.0 |
7:00-17:00 | 7:00-17:00
                            1 7:00-17:00
                                              1 7:00-17:00
| 7:00-17:00 | 8:00-12:00
                               l None
                                              1 Home
Services, Solar Installation, Heating & Air Conditioning/HVAC
BusinessAcceptsCreditCards, BusinessAcceptsBitcoin, ByAppointmen
t0nly
       1 |
I -K4gAv8_vjx8-2BxkVeRkA | Baby Cakes
4145 Erie St
                           | Willoughby
                                           I OH | 44094
                                                3.5 | None
| 41.6399 | -81.4064 |
                                    5 I
| 11:00-17:00 | 11:00-17:00
                               1 11:00-20:00
11:00-17:00 | 10:00-17:00 | None
                                           | Bakeries,Food
BusinessAcceptsCreditCards, RestaurantsTakeOut, WheelchairAccess
```

```
ible, Restaurants Delivery
       1 I
| -PtTGvWsckUL8tTutHr6Ew | Snip-its Rocky River
21609 Center Ridge Rd | Rocky River | OH | 44116
| 41.4595 | -81.8587 |
                                  18 I
                                               2.5 |
10:00-19:00 | 10:00-19:00 | 10:00-19:00
                                            10:00-19:00
| 10:00-19:00 | 9:00-17:30 | 10:00-16:00 | Beauty &
Spas, Hair Salons
Business Accepts Credit Cards, Restaurants Price Range 2, Good For Kids,\\
BusinessParking, ByAppointmentOnly
       1 |
| -ayZoW_iNDsunYXX_0x1YQ | Standard Restaurant Supply
2922 E McDowell Rd
                          l Phoenix
                                         | AZ | 85008
| 33.4664 | -112.018 |
                                  15 I
                                               3.5 L
8:00-18:00 | 8:00-18:00
                           8:00-18:00
                                            8:00-18:00
| 8:00-18:00 | 9:00-17:00
                          l None
Shopping, Wholesalers, Restaurant Supplies, Professional
Services, Wholesale Stores
BusinessAcceptsCreditCards,RestaurantsPriceRange2,BusinessPark
ing, BikeParking, WheelchairAccessible
I -d9qyfNhLMQwVVg_raBKeg | What A Bagel
973 Eglinton Avenue W
                     l York
                                         I ON I M6C 2C4
| 43.6999 | -79.4295 |
                                   8 |
                                               3.0 |
6:00-15:30 | 6:00-15:30
                           | 6:00-15:30
                                            | 6:00-15:30
| 6:00-15:30 | 6:00-15:30
                             l None
Restaurants, Bagels, Breakfast & Brunch, Food
NoiseLevel, RestaurantsAttire, RestaurantsTableService, OutdoorSe
ating
       1 I
| -hjbcaxaU9yYXY2iI-49sw | Pinnacle Fencing Solutions
| Phoenix
           13 |
             4.0 | 8:00-16:00 | 8:00-16:00
                                             8:00-16:00
| 8:00-16:00 | 8:00-16:00 | None
                                              l None
| Home Services, Contractors, Fences & Gates
| BusinessAcceptsCreditCards, ByAppointmentOnly
I -iu4FxdfxN4rU4Fu9BjiFw | Alterations Express
```

```
17240 Royalton Rd | Strongsville | OH | 44136 | 41.3141 | -81.8207 | 3 | 4.0 |
8:00-19:00 | 8:00-19:00 | 8:00-19:00
                                           8:00-19:00
| 8:00-19:00 | 8:00-18:00 | None
Shopping, Bridal, Dry Cleaning & Laundry, Local Services, Sewing &
Alterations
BusinessParking, BusinessAcceptsCreditCards, RestaurantsPriceRan
ge2, BusinessAcceptsBitcoin, BikeParking, ByAppointmentOnly, Wheel
chairAccessible
       1 |
I -j4NsiRzSMrMk2N_bGH_SA | Extra Space Storage
2880 W Elliot Rd | Chandler | AZ | 85224
| 33.3496 | -111.892 |
                                 5 I
                                             4.0
8:00-17:30 | 8:00-17:30 | 8:00-17:30
                                          8:00-17:30
Services, Self Storage, Movers, Shopping, Local Services, Home
Decor, Home & Garden
| BusinessAcceptsCreditCards
| -uiBBVWI6tMDm2JFbZFrOw | Gussied Up
1090 Bathurst St
                        | Toronto | ON | M5R 1W5
| 43.6727 | -79.4142 |
                                 6 I
                                             4.5 | None
| 11:00-19:00 | 11:00-19:00 | 11:00-19:00
11:00-19:00 | 11:00-17:00 | 12:00-16:00 | Women's
Clothing, Shopping, Fashion
BusinessAcceptsCreditCards,RestaurantsPriceRange2,BusinessPark
ing, BikeParking
       1 |
| O-aPEeNc2zVb5Gp-i7Ckqq | Buddy's Muffler & Exhaust |
1509 Hickory Grove Rd | Gastonia | NC | 28056
| 35.2772 | -81.06 |
                                 4 |
                                             5.0 |
8:30-17:00 | 8:30-17:00 | 8:30-17:00
                                          8:30-17:00
| 8:30-17:00 | 9:00-15:00 | None
Automotive, Auto Repair
| BusinessAcceptsCreditCards
       1 |
| 01xXe2m_z048W5gcBFpoJA | Five Guys
2641 N 44th St, Ste 100 | Phoenix
                                    | AZ | 85008
                                63 I
| 33.478 | -111.986 |
                                             3.5 l
```

```
10:00-22:00 | 10:00-22:00
                           10:00-22:00
                                                10:00-22:00
| 10:00-22:00 | 10:00-22:00
                                10:00-22:00
                                                | American
(New), Burgers, Fast Food, Restaurants
RestaurantsTableService,GoodForMeal,Alcohol,Caters,HasTV,Resta
urantsGoodForGroups,NoiseLevel,WiFi,RestaurantsAttire,Restaura
ntsReservations,OutdoorSeating,BusinessAcceptsCreditCards,Rest
aurantsPriceRange2,BikeParking,RestaurantsDelivery,Ambience,Re
staurantsTakeOut,GoodForKids,DriveThru,BusinessParking
        1 I
| 06I2r8S3tHP_LwGnnkk6Uw | All Storage - Anthem
2620 W Horizon Ridge Pkwy
                            l Henderson
                                             l NV
| 36.0021 | -115.102 |
                                     3 |
                                                   3.5 |
9:00-16:30 | 9:00-16:30
                             9:00-16:30
                                                1 9:00-16:30
| 9:00-16:30 | 9:00-16:30
                                l None
Rental, Local Services, Self Storage, Parking, Automotive
| BusinessAcceptsCreditCards,BusinessAcceptsBitcoin
        1 |
| 07h3mGtTovPJE660nX6E-A | Mood
                                                           1 1
                          | Edinburgh
Greenside Place
                                           I EDH
                                                   I EH1 3AA
55.957 | -3.18502 |
                                 11 I
                                                2.0 | None
1 None
                                 1 22:30-3:00
                l None
22:00-3:00
             1 22:00-3:00
                              1 22:30-3:00
                                            l Dance
Clubs, Nightlife
Alcohol, OutdoorSeating, BusinessAcceptsCreditCards, RestaurantsP
riceRange2, AgesAllowed, Music, Smoking, RestaurantsGoodForGroups,
WheelchairAccessible
        0 1
| OAJF-USLN6K5T4caooDdjw | Starbucks
                                                           4605 E Chandler Blvd, Ste A | Phoenix
                                                     1 85048
                                             l AZ
| 33.3044 | -111.984 |
                                    52 I
                                                   3.0 |
                                                1 5:00-20:30
5:00-20:00
           | 5:00-20:00
                             1 5:00-20:00
1 5:00-20:00
             1 5:00-20:00
                               | 5:00-20:00
                                              | Coffee &
Tea, Food
BusinessParking, Caters, WiFi, OutdoorSeating, BusinessAcceptsCred
itCards, RestaurantsPriceRange2, BikeParking, RestaurantsTakeOut
| OB3W6KxkD3o4W4l6cq735w | Big Smoke Burger
                                                           260 Yonge Street
                                             I ON
                            | Toronto
                                                     I M4B 2L9
```

```
| 43.6546 | -79.3805 |
                                   47 I
                                                3.0 |
10:30-21:00 | 10:30-21:00 | 10:30-21:00
                                             10:30-21:00
| 10:30-21:00 | 10:30-21:00
                              | 11:00-19:00
Poutineries, Burgers, Restaurants
RestaurantsTableService,GoodForMeal,Alcohol,Caters,HasTV,Resta
urantsGoodForGroups,NoiseLevel,WiFi,RestaurantsAttire,Restaura
ntsReservations,OutdoorSeating,BusinessAcceptsCreditCards,Rest
aurantsPriceRange2,WheelchairAccessible,BikeParking,Restaurant
sDelivery, Ambience, RestaurantsTakeOut, GoodForKids, DriveThru, Bu
sinessParkina |
                     1 |
| 0IySwcfqwJjpHPsYwjpAkg | Subway
2904 Yorkmont Rd
                           | Charlotte
                                          I NC
                                                 1 28208
                                                3.5 |
| 35.1903 | -80.9288 |
                                    7 |
6:00-22:00 | 6:00-22:00
                            | 6:00-22:00
                                             1 6:00-22:00
| 6:00-22:00 | 10:00-21:00
                              l None
                                             | Fast
Food, Restaurants, Sandwiches
Ambience,RestaurantsPriceRange2,GoodForKids
1000 Scenic Loop Dr
                           l Las Vegas
                                          | NV | 89161
                                   32 L
| 36.1357 | -115.428 |
                                                4.5 L
8:00-16:30 | 8:00-16:30
                            1 8:00-16:30
                                             8:00-16:30
8:00-16:30
            1 8:00-16:30
                              8:00-16:30
Education, Visitor Centers, Professional Services, Special
Education, Local Services, Community Service/Non-Profit, Hotels &
Travel, Travel Services, Gift
Shops, Shopping, Parks, Hiking, Flowers & Gifts, Active Life |
BusinessAcceptsCreditCards,GoodForKids
       1 |
| ONi7Stqt4RFWDGjOYRi2Bw | Scent From Above Company
2501 W Behrend Dr, Ste 67 | Scottsdale
                                          l AZ
                                                1 85027
| 33.6656 | -112.111 |
                                   14 I
                                                4.5 L
6:00-16:00 | 6:00-16:00
                            6:00-16:00
                                             | 6:00-16:00
| 6:00-16:00 | None
                               l None
Cleaning, Local Services, Professional Services, Carpet
Cleaning, Home Services, Office Cleaning, Window Washing
BusinessAcceptsCreditCards,ByAppointmentOnly
| OWBMEfqXQnEOAIkV-uCW6w | The Charlotte Room
                                                        I 19
Charlotte Street
                        l Toronto
                                       I ON
                                               I M5V 2H5 I
```

```
43.6466 | -79.3938 |
                                              3.5 l
                                 10 |
15:00-1:00 | 15:00-1:00
                            | 15:00-1:00
                                             15:00-1:00
| 15:00-2:00 | 18:00-2:00
                              l None
                                             | Event
Planning & Services, Bars, Nightlife, Lounges, Pool Halls, Venues &
Event Spaces
BusinessParking, HasTV, CoatCheck, NoiseLevel, OutdoorSeating, Busi
nessAcceptsCreditCards,RestaurantsPriceRange2,Music,Wheelchair
Accessible, Smoking, Ambience, BestNights, RestaurantsGoodForGroup
s, HappyHour, GoodForDancing, Alcohol
       0 |
| 0Y3lHyqRHfWOBuQlS1bM0q | PC Savants
11966 W Candelaria Ct | Sun City
                                         | AZ | 85373
| 33.6901 | -112.319 |
                                   11 |
                                                5.0 l
10:00-19:00 | 10:00-19:00 | 10:00-19:00
                                            10:00-19:00
| 10:00-19:00 | 11:00-18:00
                             | 11:00-18:00 | IT Services &
Computer Repair, Electronics Repair, Local Services, Mobile Phone
Repair
| BusinessAcceptsCreditCards,BusinessAcceptsBitcoin
| OaKsGxx7XP2TMs_fn_9xVw | Sweet Ruby Jane Confections
8975 S Eastern Ave, Ste 3-B | Las Vegas
                                                  89123
  36.015 | -115.118 |
                                   30 L
                                                4.0 |
10:00-19:00 | 10:00-19:00
                            10:00-19:00
                                             10:00-19:00
| 10:00-19:00 | 10:00-19:00
                              l None
Food, Chocolatiers & Shops, Bakeries, Specialty Food, Desserts
BusinessAcceptsCreditCards, RestaurantsPriceRange2, BusinessPark
ing, Wheelchair Accessible
       0 |
| OcxO1Lx2Pi7u6ftWX3Wksq | Oinky's Pork Chop Heaven
22483 Emery Rd
                          | North Randall | OH
                                               | 44128
| 41.4352 | -81.5214 |
                                    3 L
                                                3.0 L
6:00-23:00 | 6:00-23:00 | 6:00-23:00
                                            | 6:00-23:00
Food, Restaurants
RestaurantsAttire, RestaurantsGoodForGroups, GoodForKids, Restaur
antsReservations,RestaurantsTakeOut
I 0e-j5VcEn54EZT-FKCUZdw I Sushi Osaka
```

```
5084 Dundas Street W | Toronto | ON | M9A 1C2 | 43.6452 | -79.5324 | 8 | 4.5 |
11:00-23:00 | 11:00-23:00 | 11:00-23:00 | 11:00-23:00
| 11:00-23:00 | 11:00-23:00 | 14:00-23:00 | Sushi
Bars, Restaurants, Japanese, Korean
RestaurantsTakeOut,WiFi,RestaurantsGoodForGroups,RestaurantsRe
servations
     1 |
+-----
+-----
+-----
+-----
+-----
+-----
+-----
+-----
______
_____
----+
(Output limit exceeded, 25 of 70 total rows shown)
iv. Provide the SQL code you used to create your final
dataset:
SELECT B.id,
           B.name,
           B.address,
           B.city,
           B.state,
           B.postal_code as pin_code,
           B.latitude as lat,
           B.longitude as long,
           B.review_count as count of review,
           B.stars as review stars,
           MAX (CASE
           WHEN H.hours LIKE "%monday%" THEN
TRIM(H.hours, '%MondayTuesWednesThursFriSatSun |%')
           END) AS monday_hours,
```

```
MAX (CASE
                  WHEN H.hours LIKE "%tuesday%" THEN
TRIM(H.hours, '%MondayTuesWednesThursFriSatSun |%')
                  END) AS tuesday hours,
                  MAX (CASE
                  WHEN H.hours LIKE "%wednesday%" THEN
TRIM(H.hours, '%MondayTuesWednesThursFriSatSun |%')
                  END) AS wednesday hours,
                  MAX(CASE
                  WHEN H.hours LIKE "%thursday%" THEN
TRIM(H.hours,'%MondayTuesWednesThursFriSatSun|%')
                  END) AS thursday hours,
                  MAX (CASE
                  WHEN H.hours LIKE "%friday%" THEN
TRIM(H.hours, '%MondayTuesWednesThursFriSatSun |%')
                  END) AS friday hours,
                  MAX (CASE
                  WHEN H.hours LIKE "%saturday%" THEN
TRIM(H.hours,'%MondayTuesWednesThursFriSatSun|%')
                  END) AS saturday hours,
                  MAX (CASE
                  WHEN H.hours LIKE "%sunday%" THEN
TRIM(H.hours, '%MondayTuesWednesThursFriSatSun |%')
                  END) AS sunday hours,
                  GROUP CONCAT(DISTINCT(C.category)) AS
categories,
                  GROUP CONCAT(DISTINCT(A.name)) AS
attributes,
                  B.is open
          FROM business B
          INNER JOIN hours H
          ON B.id = H.business id
          INNER JOIN category C
          ON B.id = C.business id
          INNER JOIN attribute A
          ON B.id = A.business id
          GROUP BY B.id;
```