Cognifyz Technologies

Data Analytics Internship – Dataset, Tasks Info & Output Screenshots

Dataset :

Restaurant ID	Restaurant Name	Country Code	City	Address	Locality
6317637	Le Petit Souffle	162	Makati City	Third Floor, Century City Mall, Kalayaan Avenue, Poblacion, Makati City	Century City Mall, Poblacion, Makati City
6304287	Izakaya Kikufuji	162	Makati City	Little Tokyo, 2277 Chino Roces Avenue, Legaspi Village, Makati City	Little Tokyo, Legaspi Village, Makati City
6300002	Heat - Edsa Shangri-La	162	Mandaluyong City	Edsa Shangri-La, 1 Garden Way, Ortigas, Mandaluyong City	Edsa Shangri-La, Ortigas, Mandaluyong City
6318506	Ooma	162	Mandaluyong City	Third Floor, Mega Fashion Hall, SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City
6314302	Sambo Kojin	162	Mandaluyong City	Third Floor, Mega Atrium, SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City
18189371	Din Tai Fung	162	Mandaluyong City	Ground Floor, Mega Fashion Hall, SM Megamall, Ortigas, Mandaluyong City	SM Megamall, Ortigas, Mandaluyong City
6300781	Buffet 101	162	Pasay City	Building K, SM By The Bay, Sunset Boulevard, Mall of Asia Complex (MOA), Pasay City	SM by the Bay, Mall of Asia Complex, Pasay Ci
6301290	Vikings	162	Pasay City	Building B, By The Bay, Seaside Boulevard, Mall of Asia Complex (MOA), Pasay City	SM by the Bay, Mall of Asia Complex, Pasay Ci
6300010	Spiral - Sofitel Philippine Plaza Manila	162	Pasay City	Plaza Level, Sofitel Philippine Plaza Manila, CCP Complex, Pasay City	Sofitel Philippine Plaza Manila, Pasay City
6314987	Locavore	162	Pasig City	Brixton Technology Center, 10 Brixton Street, Kapitolyo, Pasig City	Kapitolyo

more 9553 rows.

Tasks Information and Output Screenshots :-

#Level 1 :

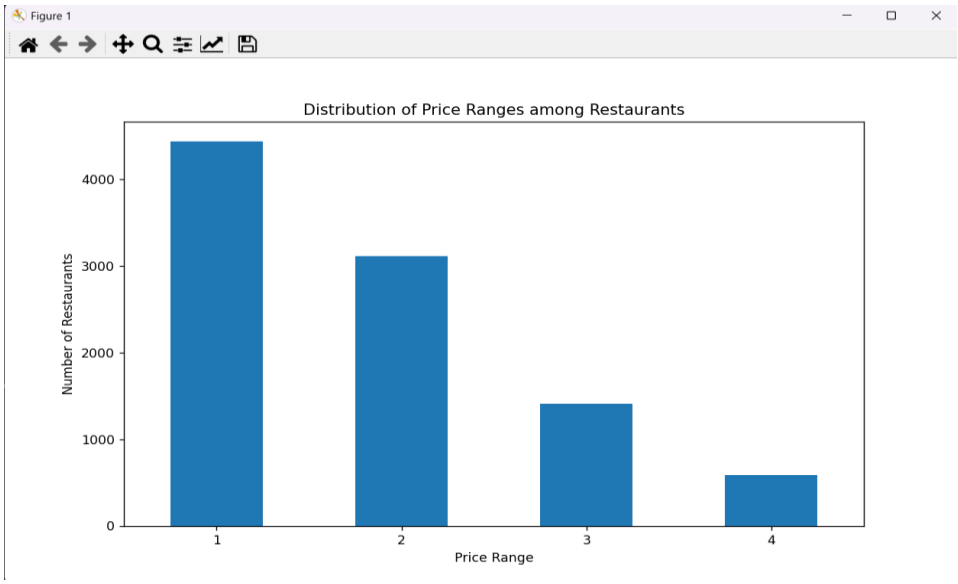
Task 1: Top Cuisines, Determine the top three most common cuisines in the dataset. Calculate the percentage of restaurants that serve each of the top cuisines.

```
Top Three Most Common Cuisines:
North Indian: 3960 restaurants (41.46%)
Chinese: 2735 restaurants (28.64%)
Fast Food: 1986 restaurants (20.79%)
```

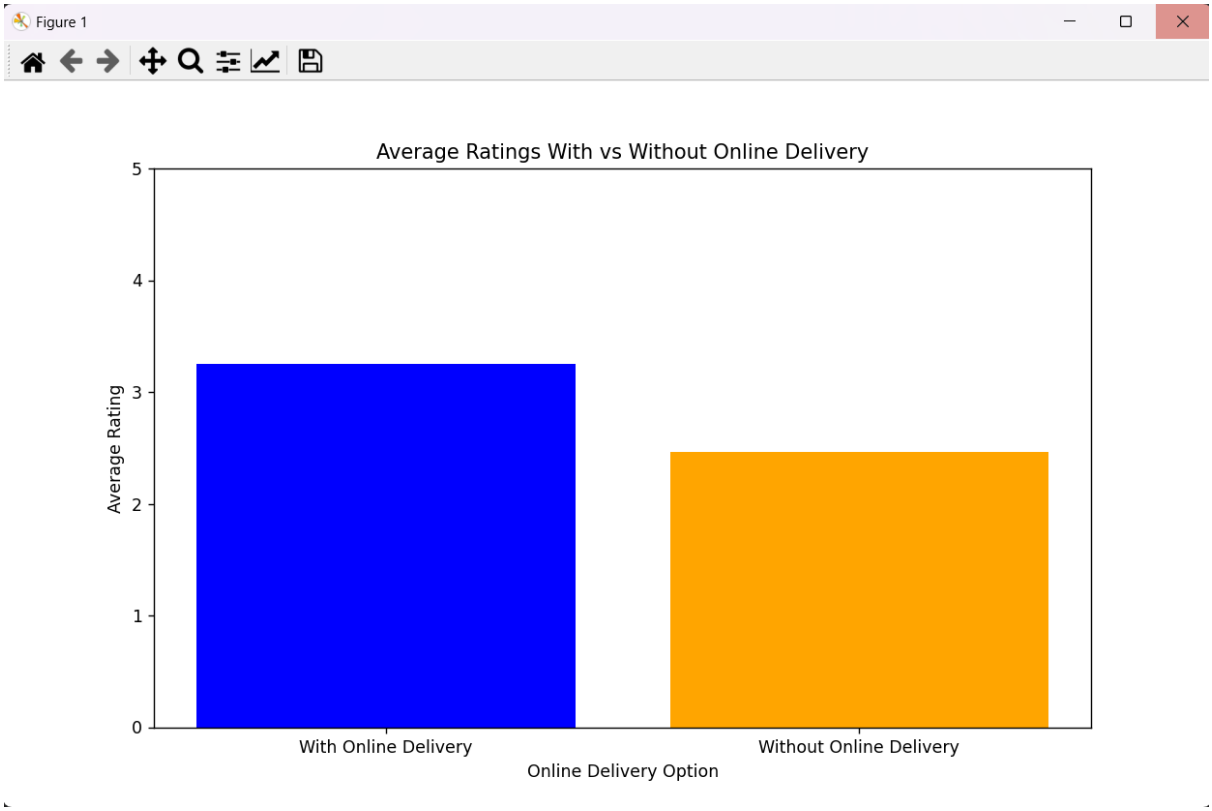
Task 2: City Analysis, Identify the city with the highest number of restaurants in the dataset. Calculate the average rating for restaurants in each city. Determine the city with the highest average rating.

```
City with the most restaurants: New Delhi
Number of restaurants: 5473
City with the highest average rating: Inner City
Highest average rating: 4.9
```

Task 3: Price Range Distribution, Create a histogram or bar chart to visualize the distribution of price ranges among the restaurants.

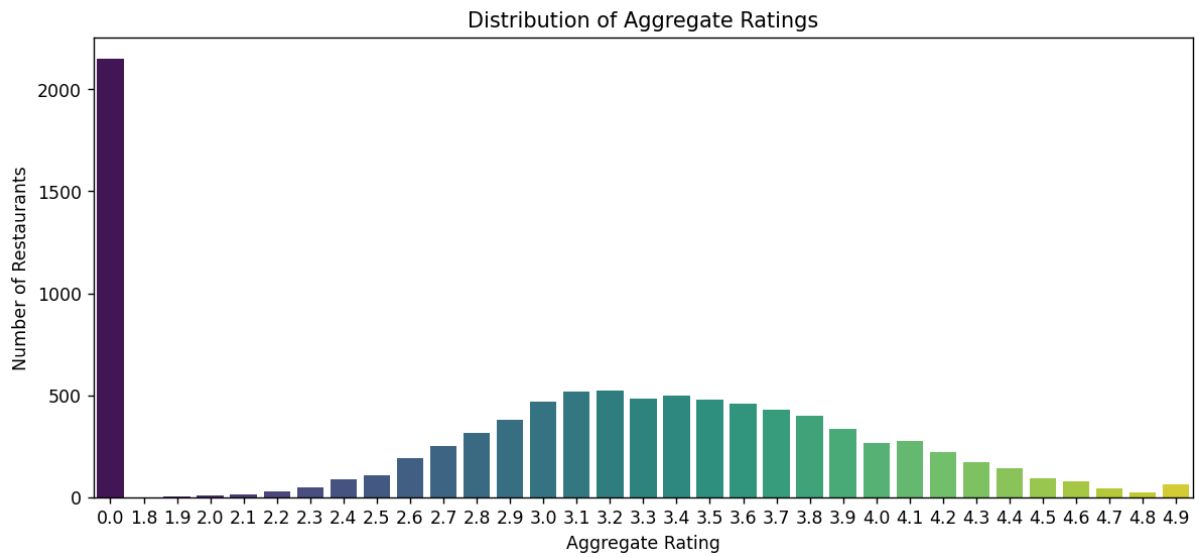


Task 4: Online Delivery, Determine the percentage of restaurants that offer online delivery.



#Level 2 :

Task 1: Restaurant Ratings, Analyze the distribution of aggregate ratings and determine the most common rating range Calculate the average number of votes received by restaurants.

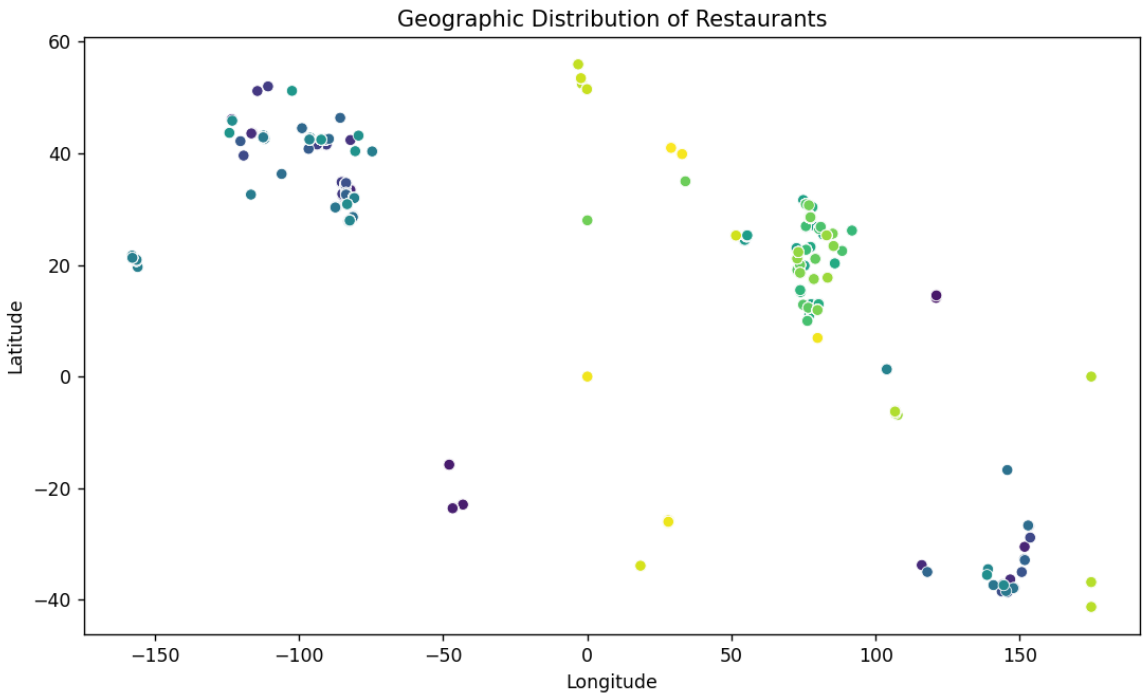


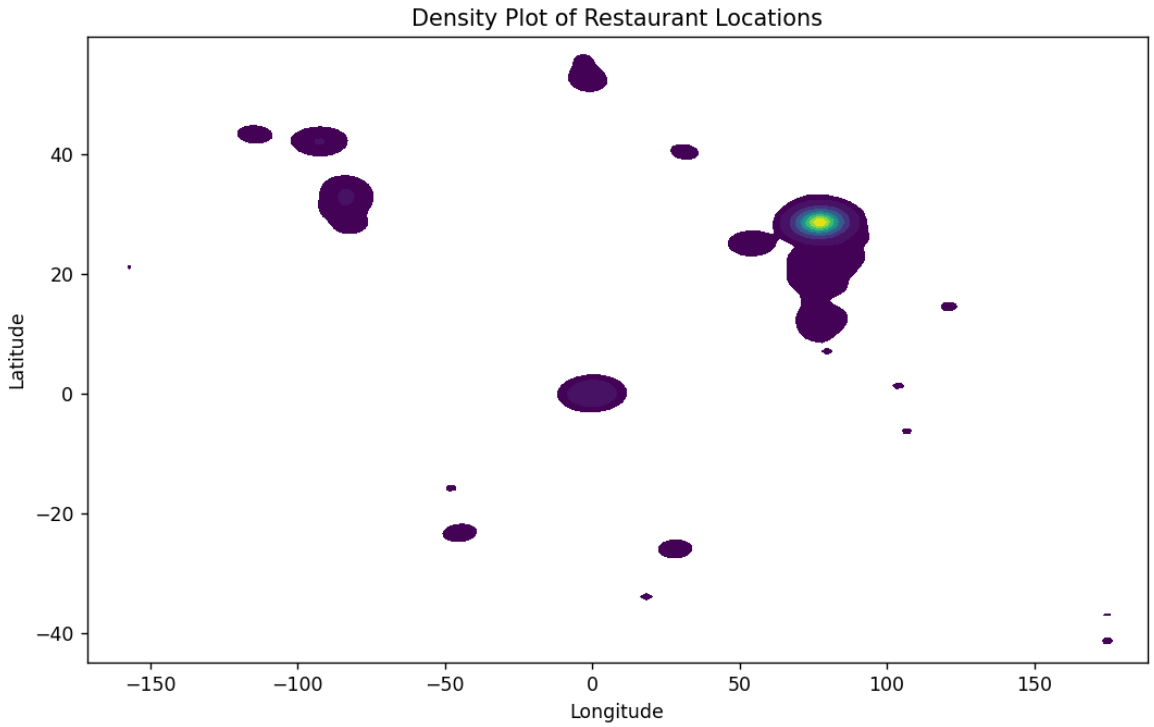
Task 2: Cuisine Combination, Identify the most common combinations of cuisines in the dataset. Determine if certain cuisine combinations tend to have higher ratings.

```
Most Common Combinations of Cuisines:
Cuisines
North Indian                        936
North Indian, Chinese               511
Chinese                             354
Fast Food                           354
North Indian, Mughlai               334
Cafe                                299
Bakery                              218
North Indian, Mughlai, Chinese      197
Bakery, Desserts                    170
Street Food                         149
Name: count, dtype: int64

Average Ratings of Top Cuisine Combinations:
Cuisines
North Indian                        1.672329
North Indian, Chinese               2.421722
Chinese                             2.042090
Fast Food                           2.118362
North Indian, Mughlai               2.888623
Cafe                                2.890970
Bakery                              1.924312
North Indian, Mughlai, Chinese      2.568528
Bakery, Desserts                    2.317647
Street Food                         2.161745
Name: Aggregate rating, dtype: float64
```

Task 3: Geographic Analysis, Plot the locations of restaurants on a map using longitude and latitude coordinates. Identify any patterns or clusters of restaurants in specific areas.





Task 4: Restaurant Chains, Identify if there are any restaurant chains present in the dataset. Analyze the ratings and popularity of different restaurant chains.

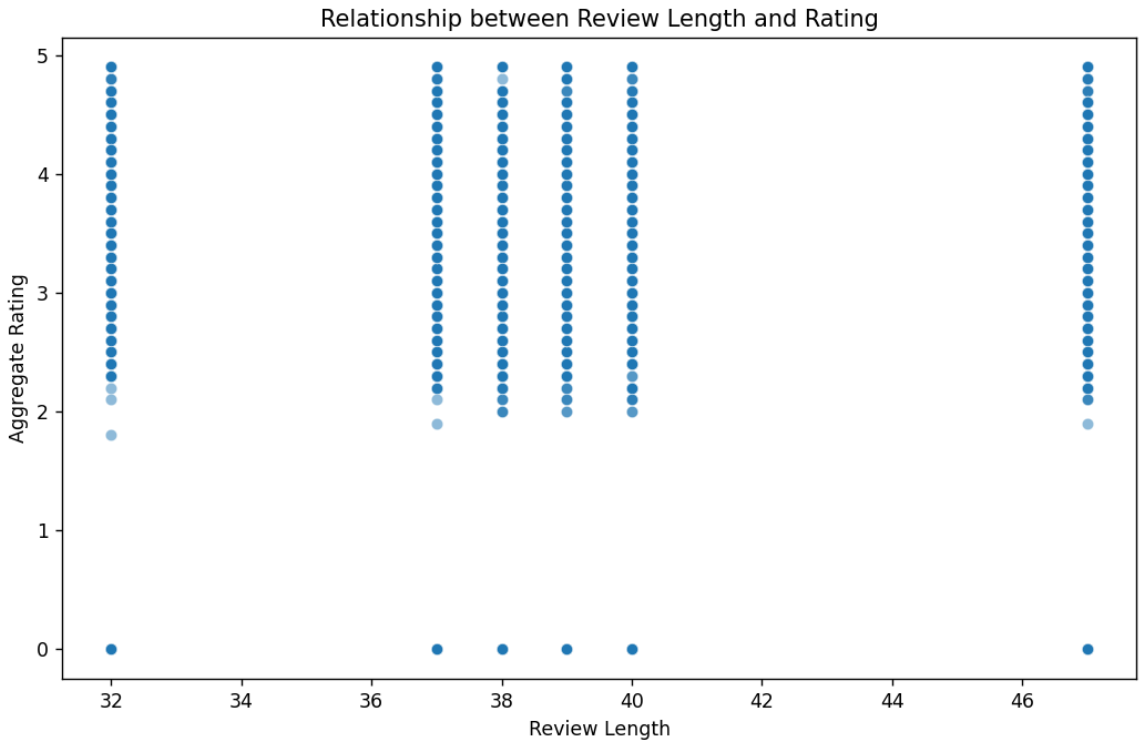
```
Restaurant Chains:
Restaurant Name
Cafe Coffee Day       83
Domino's Pizza        79
Subway                63
Green Chick Chop      51
McDonald's            48
                      ..
Town Hall              2
Halki Aanch            2
Snack Junction         2
Delhi Biryani Hut      2
Beliram Degchiwala     2
Name: count, Length: 734, dtype: int64

Chain Analysis:
                        Average Rating  Total Votes
Restaurant Name
Barbeque Nation               4.353846        28142
AB's - Absolute Barbecues     4.825000        13400
Big Chill                     4.475000        10853
Farzi Cafe                    4.366667        10098
Truffles                      3.950000         9682
...                                ...          ...
Bikaner Misthan Bhandar       0.000000            0
Aap Ki Khatir                 0.000000            0
Street Cafe                   0.000000            0
Jyoti Sweets                  0.000000            0
Firangi Bake                  0.000000            0

[734 rows x 2 columns]
```

#Level 3 :

Task 1: Restaurant Reviews, Analyze the text reviews to identify the most common positive and negative keywords. Calculate the average length of reviews and explore if there is a relationship between review length and rating.



Task 2: Votes Analysis, Identify the restaurants with the highest and lowest number of votes. Analyze if there is a correlation between the number of votes and the rating of a restaurant.

```
Restaurant with the highest votes:
Name: Toit
Votes: 10934
Rating: 4.8

Restaurant with the lowest votes:
Name: Cantinho da Gula
Votes: 0
Rating: 0.0

Correlation between votes and rating:
0.31369058419541135
```

Task 3: Price Range vs. Online Delivery and Table Booking, Analyze if there is a relationship between the price range and the availability of online delivery and table booking. Determine if higher-priced restaurants are more likely to offer these services.

There is a separate file containing python codes for each level and task.

For ex : Level 1 Task 1 (L1_t1.py)

Level 1 Task 2 (L1_t2.py)

Level 3 Task 1 (L3_t1.py)... like these files are named.