COGNIFYZ DATA SCIENCE INTERNSHIP LEVEL 2 REPORT

About the Level

Level 2 of the Cognifyz Data Science Internship focuses on the following tasks:

1. Table Booking and Online Delivery
2. Price Range Analysis, and
3. Feature Engineering.

Task 1: Table Booking and Online Delivery

- Determine the percentage of restaurants that offer table booking and online delivery.
- Compare the average ratings of restaurants with table booking and those without.
- Analyse the availability of online delivery among restaurants with different price ranges.

Task 2: Price Range Analysis

- Determine the most common price range among all the restaurants.
- Calculate the average rating for each price range.
- Identify the colour that represents the highest average rating among different price ranges.

Task 3: Feature Engineering

- Extract additional features from the existing columns, such as the length of the restaurant name or address.
- Create new features like "Has Table Booking" or "Has Online Delivery" by encoding categorical variables.

RESULTS

Task 1: Table Booking and Online Delivery

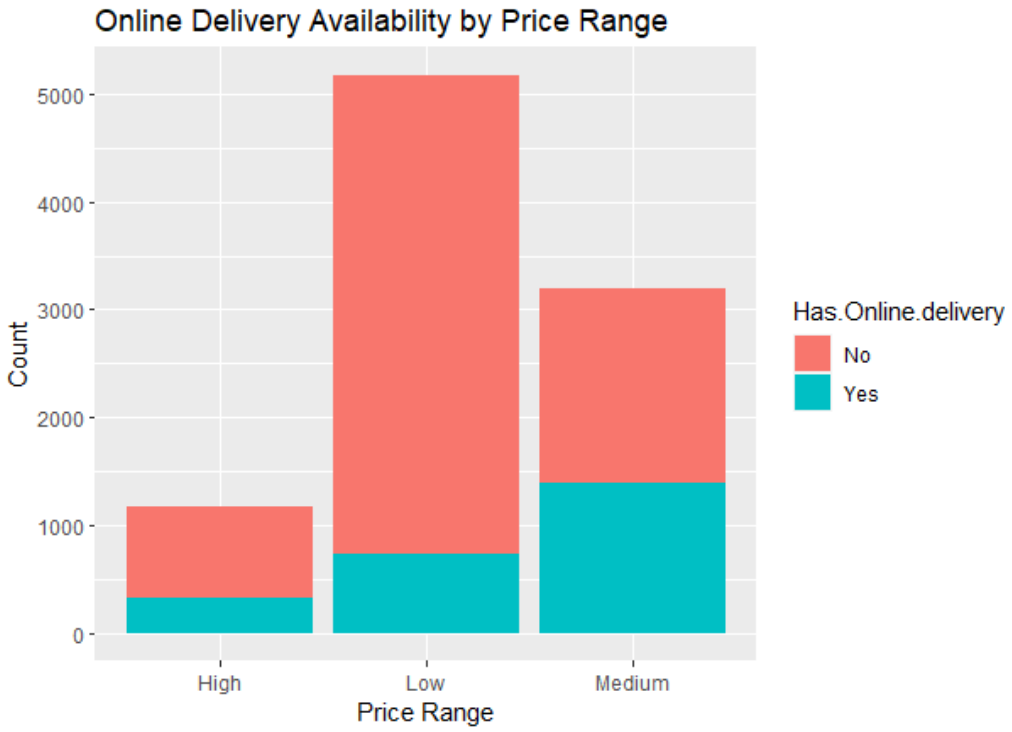
The percentage of restaurants that offer table booking is 12.12% while the percentage of restaurants that offer online delivery is 25.66%. Clearly, restaurants that offer online delivery have a higher percentage than those that offer table booking. On the other hand, restaurants with table bookings have a higher average rating (3.44) than those without (2.56).

Also, the availability of online delivery among restaurants with medium price ranges is higher than those with low and high prices.

```
Online Delivery Availability by Price Range:
> print(online_delivery_by_price_range)

price_ranges        No       Yes
      High   0.7191489 0.2808511
      Low    0.8585234 0.1414766
      Medium 0.5662086 0.4337914
```

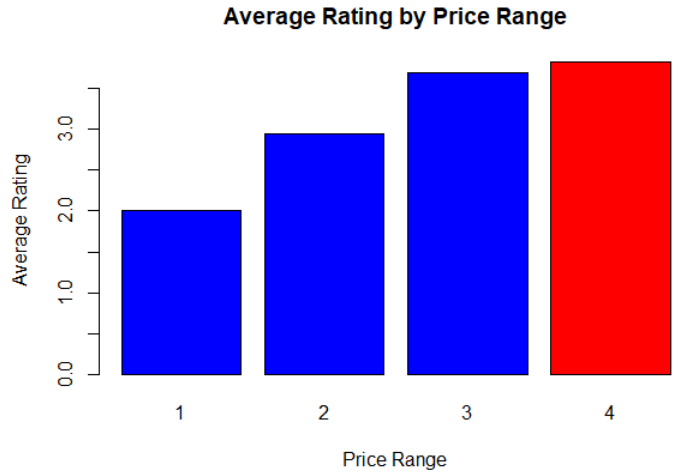
Below is a bar plot to represent it well.



Task 2: Price Range Analysis

The most common price range among all the restaurants is 1.

Price range 4 gets the highest average rating of 3.82 followed by price range 3 with an average rating of 3.68, price range 2 with an average rating of 2.94 and price range 1 with an average rating of 2.00. Below is a bar plot with the highest average rating in red.



Task 3: Feature Engineering

In this task, I created two new columns “Restaurant Name Length” and “Address Length” based on the length of the restaurant names and addresses respectively.

I also encoded the columns "Has Table Booking" and "Has Online Delivery" with binary numbers “1” for “Yes” and “0” for “No” as additional two new columns.

Conclusion

This level of the project emphasized the significance of leveraging advanced data science techniques to optimize analysis.

The price range analysis provided insights into the common price range for all restaurants and the price range with the highest average rating. This uncovered opportunities to maximize revenue while maintaining competitive pricing structures.

In addition, the implementation of feature engineering techniques also enriches the dataset with meaningful predictors. This will improve the performance and interpretability of predictive models that will be developed with the data.