**Project Summary**

**Level 1: Data Exploration and Analysis**

* Conducted comprehensive data exploration and preprocessing to ensure data integrity.
* Performed descriptive analysis, extracting key statistical measures, and identified popular cuisines and cities.
* Explored geospatial insights, visualizing restaurant locations and investigating correlations with ratings.

**Level 2: Advanced Analysis**

* Analyzed table booking and online delivery services, uncovering insights on customer preferences and availability.
* Determined the most common price range and associated it with the highest average rating.
* Employed advanced feature engineering to enhance dataset intelligence.

**Level 3: Predictive Modeling and Insights**

* Built regression models to predict restaurant aggregate ratings, with Random Forest emerging as the top performer.
* Explored the relationship between cuisine types and restaurant ratings.
* Analyzed data visualizations to uncover rating distributions and other insights.

**Conclusion**

**1. Data Overview and Exploration:**

* The dataset encompasses 9,551 restaurant records with 21 columns.
* Minimal null values were detected, predominantly within the 'Cuisines' column.
* No duplicates existed, and data type conversion was unnecessary.
* The 'Aggregate rating' distribution displayed a balanced pattern.

**2. Descriptive Insights:**

* Key statistical metrics for numerical columns were identified.
* The most prominent country codes were 1 and 216, while cities like New Delhi, Gurgaon, and Noida led in restaurant counts.
* Popular cuisines included North Indian and Chinese.

**3. Geospatial Analysis:**

* North America and Asia, especially India, were prominent for restaurant presence.
* New Delhi emerged as the city with the most restaurants, followed by Gurgaon, Noida, and Faridabad.
* Latitude exhibited no correlation with ratings, while longitude showed a negative correlation.

**4. Table Booking and Online Delivery Analysis:**

* Approximately 12.12% of restaurants offered table booking, and 25.66% provided online delivery services.
* Restaurants with table booking displayed a significantly higher average rating of 3.44 compared to 2.56 for those without this service.
* Online delivery was more prevalent in restaurants with medium-priced food products.

**5. Price Range Analysis:**

* Price range 1 was the most common among restaurants.
* Restaurants in price range 4 achieved the highest average rating, followed by price ranges 3, 2, and 1.

**6. Feature Engineering:**

* Two new columns, 'Restaurant Name Length' and 'Address Length,' were created based on the length of restaurant names and addresses.
* Two binary columns, 'Has Table Booking' and 'Has Online Delivery,' were introduced through categorical variable encoding.

**7. Predictive Modeling Insights:**

* Three regression models, Linear Regression, Decision Tree, and Random Forest were employed to predict restaurant aggregate ratings.
* Random Forest outperformed other models, demonstrating the lowest Mean Squared Error (MSE) and the highest R-squared value.

**8. Customer Preference Analysis:**

* Various cuisines such as cafe, mughlai, north Indian, and fast food had a significant impact on restaurant ratings, with varying performance.
* North Indian and Chinese cuisines displayed greater rating variability, while cafe and fast food cuisines maintained more consistent ratings.
* Based on the number of votes, North Indian, Mughlai, and Chinese cuisines emerged as the most popular.
* Italian, Hawaiian, Seafood, Tea, Sandwich, Continental, and Indian cuisines received the highest average ratings.

**9. Data Visualization Highlights:**

* Restaurant ratings followed a negatively skewed distribution.
* Italian, Hawaiian, Seafood, Tea, Sandwich, Continental, and Indian cuisines secured the top spots in terms of the highest average ratings.
* Cities like Inner City, Quezon City, and Makati City were identified as the most popular based on the highest average rating.
* A positive correlation between votes and restaurant ratings was observed.