**Food\_Hub\_Data\_Analysis**

**Data**

The food aggregator company has stored the data of the different orders made by the registered customers in their online portal. They want to analyze the data to get a fair idea about the demand of different restaurants which will help them in enhancing their customer experience. Suppose you are hired as a Data Scientist in this company and the Data Science team has shared some of the key questions that need to be answered. Perform the data analysis to find answers to these questions that will help the company to improve the business.

**Data Description**

The data contains the different data related to a food order. The detailed data dictionary is given below.

Data Dictionary

* order\_id: Unique ID of the order
* customer\_id: ID of the customer who ordered the food
* restaurant\_name: Name of the restaurant
* cuisine\_type: Cuisine ordered by the customer
* cost: Cost of the order
* day\_of\_the\_week: Indicates whether the order is placed on a weekday or weekend (The weekday is from Monday to Friday and the weekend is Saturday and Sunday)
* rating: Rating given by the customer out of 5
* food\_preparation\_time: Time (in minutes) taken by the restaurant to prepare the food. This is calculated by taking the difference between the timestamps of the restaurant's order confirmation and the delivery person's pick-up confirmation.
* delivery\_time: Time (in minutes) taken by the delivery person to deliver the food package. This is calculated by taking the difference between the timestamps of the delivery person's pick-up confirmation and drop-off information

**Basic Steps:**

* + 1. Display the top 5 rows.
    2. Display the last 5 rows
    3. Check the shape of dataset.
    4. Check the datatypes of each feature.
    5. Check the Statistical summary
    6. Check the null values
    7. Check the duplicate values
    8. Check the anomalies or wrong entries.
    9. Check the outliers and their authenticity.
    10. Do the necessary data cleaning steps like dropping duplicates, unnecessary columns, null value imputation, outliers treatment etc.

1. **Order Analysis**
   * What is the total number of orders in the dataset?
   * What is the average cost of an order?
   * How many unique customers have placed orders?
   * Which restaurant has received the highest number of orders?
2. **Customer Behavior**
   * What is the average rating given by customers?
   * How does the rating vary between weekdays and weekends?
   * Which cuisine type is ordered the most?
   * What is the distribution of orders across different days of the week?
3. **Restaurant Performance**
   * What is the average food preparation time for each restaurant?
   * Which restaurant has the shortest average food preparation time?
   * How does the average delivery time compare across different restaurants?
   * Is there a correlation between the cost of the order and the rating given?
4. **Demand Patterns**
   * How does the demand for different cuisine types vary on weekdays versus weekends?
   * Which day of the week has the highest average order cost?
   * What is the most common day for orders to be placed?
   * How does the average rating vary by cuisine type?
5. **Operational Efficiency**
   * What is the average delivery time for all orders?
   * Which restaurant has the longest average delivery time?
   * Is there a relationship between food preparation time and delivery time?
   * How does the delivery time impact customer ratings?
6. **Customer Insights**
   * What is the repeat order rate (number of customers who have placed more than one order)?
   * What percentage of orders receive a rating of 4 or higher?