**Project**

**Zomato Data Analytics Training Case Study**

**Background Story:**

You’ve just joined the analytics team at **Zomato**, one of India's leading food delivery platforms. Recently, the company has been facing multiple challenges, particularly related to declining customer satisfaction, fluctuating delivery times, and inconsistent sales performance across different regions. The leadership is worried about customer retention, as data indicates a drop in repeat customers and many order cancellations. They’ve tasked you, a junior data analyst, with analyzing the data to uncover actionable insights.

Imagine yourself sitting in a meeting with Zomato's management team. The pressure is on, and everyone is waiting for answers to help Zomato reverse these downward trends. Your mission is to dive deep into the sales data, find key metrics that will help improve the customer experience, and ultimately drive the company's growth.

The dataset you’ll be analyzing consists of customer orders, restaurant ratings, delivery times, and several other attributes. The company needs insights that will guide decisions related to customer engagement, marketing strategies, and operational efficiency.

**Problem Selection:**

Zomato has experienced a decline in customer engagement and an increase in order cancellations.

**The problem:** Why are customers leaving? What is causing delivery delays and order cancellations?

**Why is it important:** These issues can affect Zomato’s bottom line by reducing customer lifetime value, impacting its reputation, and losing market share to competitors like Swiggy.

Your task is to identify the pain points within the data. What are the primary reasons for customer dissatisfaction? Which operational inefficiencies can be improved to enhance the delivery process? By tackling this problem, you'll learn how to approach real-world data challenges and develop the skills necessary to extract valuable insights from the data.

**Stakeholder Involvement:**

* **Internal Stakeholders:**
  + **Marketing Team:** Responsible for customer retention strategies.
  + **Operations Team:** Manages the delivery fleet and works to reduce delivery times.
  + **Customer Support:** Handles customer complaints, including refunds and cancellations.
  + **Executive Leadership:** Needs a summary of findings to make high-level decisions.
* **External Stakeholders:**
  + **Customers:** Their satisfaction drives Zomato’s revenue and brand loyalty.
  + **Restaurant Partners:** Their performance affects customer satisfaction and the speed of order preparation.

**Problem Definition:**

The main problem that you, as a data analyst, need to address is: **"What factors are driving customer dissatisfaction, leading to increased order cancellations and reduced repeat customers?"**

* **Data Requirements:** You already have access to key data like customer order history, restaurant performance, delivery times, and payment methods.
* **Metrics Development:** Important metrics will include average delivery time, order cancellation rates, customer ratings, and order frequency.

**Resources:**

* Google Sheet Link:
  + [Zomato Looker Case Data](https://docs.google.com/spreadsheets/d/1uvDEVe3jpn_II6SUZcIimsK23zC8i-INr6uCkZleheA/edit?usp=sharing)
  + When you open this sheet. Please create a copy of this sheet otherwise you will not be able to use this data.
  + Go to File —-> Create a Copy —--> Name it as “Zomato Data”

**Data Dictionary:**

Here is a quick reference to understand the dataset you’ll be working with:

|  |  |  |
| --- | --- | --- |
| **Column Name** | **Data Type** | **Description** |
| Order\_ID | String | Unique identifier for each order. |
| Customer\_ID | String | Unique identifier for each customer. |
| Order\_Date | Date | The date and time when the order was placed. |
| Restaurant\_ID | String | Unique identifier for each restaurant. |
| Restaurant\_Name | String | Name of the restaurant from which the order was placed. |
| Cuisine\_Type | String | Type of cuisine (e.g., Indian, Chinese, Italian). |
| Order\_Amount | Float | The total amount spent by the customer on the order. |
| Delivery\_Fee | Float | Fee charged for delivery. |
| Total\_Amount | Float | Final amount paid by the customer (Order\_Amount + Delivery\_Fee). |
| Payment\_Method | String | Method of payment (e.g., Credit Card, Debit Card, UPI, Cash). |
| Discount\_Amount | Float | Discount applied to the order. |
| Order\_Status | String | Status of the order (e.g., Completed, Cancelled, Failed). |
| Delivery\_Time | Integer | Time taken for the delivery in minutes. |
| Customer\_Rating | Integer | Customer rating of the order (scale of 1 to 5). |
| City | String | The city in which the order was placed. |
| Restaurant\_Rating | Float | Average rating of the restaurant based on all customer reviews (scale of 1 to 5). |
| Number\_of\_Items | Integer | Total number of items in the order. |
| Order\_Type | String | Type of order (e.g., Delivery, Dine-in, Takeaway). |

**1. What is the total number of completed orders?**

* **Hint:** Filter by Order\_Status = "Completed".
* **Visualization Type:** **Scorecard**
* **Why Use This Visualization:** A **Scorecard** is ideal for displaying key metrics at a glance, such as total completed orders. It provides a quick overview without clutter.

**2. What is the average delivery time for completed orders across all cities?**

* **Hint:** Filter for Order\_Status = "Completed" and calculate the average from Delivery\_Time.
* **Visualization Type:** **Bar Chart**
* **Why Use This Visualization:** A **Bar Chart** clearly shows comparison across categories (in this case, cities), allowing for easy analysis of which city has higher delivery times.

**3. Which payment method is the most used?**

* **Hint:** Group by Payment\_Method and count the occurrences.
* **Visualization Type:** **Pie Chart**
* **Why Use This Visualization:** A **Pie Chart** is effective when you need to show proportions and compare how each payment method contributes to the total number of orders.

**4. What is the trend of total revenue generated over time?**

* **Hint:** Sum the Total\_Amount and group by Order\_Date to see the trend.
* **Visualization Type:** **Time Series Chart**
* **Why Use This Visualization:** A **Time Series Chart** is ideal for tracking trends over time. It visually represents fluctuations in revenue, helping stakeholders identify peak periods and trends.

**5. Which city generates the most revenue?**

* **Hint:** Group by City and sum the Total\_Amount.
* **Visualization Type:** **Geo Map**
* **Why Use This Visualization:** A **Geo Map** is effective when analyzing location-based data. It visually displays which cities contribute the most to Zomato’s revenue.

**6. What is the distribution of customer ratings across all orders?**

* **Hint:** Group by Customer\_Rating and count the number of orders.
* **Visualization Type:** **Histogram**
* **Why Use This Visualization:** A **Histogram** is perfect for showing the distribution of data, such as customer ratings. It helps analyze how frequently different ratings are given.

**7. Which cuisine type has the highest average order value?**

* **Hint:** Group by Cuisine\_Type and calculate the average of Total\_Amount.
* **Visualization Type:** **Column Chart**
* **Why Use This Visualization:** A **Column Chart** helps compare average order values across different cuisine types in a clear and easy-to-read format.

**8. What are the total delivery times across various order statuses?**

* **Hint:** Group by Order\_Status and calculate the sum of Delivery\_Time.
* **Visualization Type:** **Stacked Bar Chart**
* **Why Use This Visualization:** A **Stacked Bar Chart** allows for easy comparison of total delivery times across different order statuses, making it easier to visualize operational inefficiencies.

**9. How do delivery times differ by cuisine type?**

* **Hint:** Group by Cuisine\_Type and calculate the average delivery time.
* **Visualization Type:** **Heat Map**
* **Why Use This Visualization:** A **Heat Map** visually shows variations in data. It’s useful for analyzing how delivery times differ across cuisines, with the intensity of the color representing delivery speed.

**10. What is the average number of items per order for different payment methods?**

* **Hint:** Group by Payment\_Method and calculate the average of Number\_of\_Items.
* **Visualization Type:** **Treemap**
* **Why Use This Visualization:** A **Treemap** provides a hierarchical view of the data and is great for comparing multiple categories like payment methods and the average number of items in each order.