Final Report: Restaurant Analysis for Zomato

1. Introduction

The objective of this analysis is to evaluate the performance of restaurants listed on Zomato by exploring key metrics such as popularity (measured by rating counts) and revenue generation. The study aims to uncover patterns and correlations between customer reviews and financial performance. Through this analysis, I provide insights into which restaurants are performing well, identify top-performing segments, and offer recommendations for potential business improvements.

2. Data Overview

Four datasets were used in this analysis:

- **Restaurant Table:** Contains details such as restaurant_id, name, city, rating, rating_count, cost, cuisine, lic_no, link, address, and menu link.
- **Orders Table:** Records order_date, sales_qty, sales_amount, currency, user_id, and restaurant_id.
- **Menu Table:** Includes menu_id, restaurant_id, food_id, cuisine, and price.
- **Food Table:** Consists of food_id, item, and veg_or_non_veg.

The **Users** table was deemed irrelevant to this analysis and was therefore excluded.

3. Data Cleaning and Preparation

Data cleaning was conducted in Power BI Desktop using the following steps:

- **Validation of Data Types:** Ensured that numerical, text, and date fields were correctly identified.
- **Handling Missing Values:** Filtered out or replaced null values in critical fields such as `restaurant_id` and `order_date`.
- **Removing Duplicates:** Verified unique identifiers (e.g., `restaurant_id` in the restaurant table) and removed duplicate entries.
- **Standardizing Formats:** Applied consistent formatting to text fields like city names and cuisine types using functions such as `TRIM` and case conversion.

Relationships between tables were set up in the model view with:

- **One-to-Many:** Between the restaurant table and orders table, and between the restaurant table and menu table.
- **One-to-Many:** Between the food table and menu table.

These steps ensured a clean and reliable dataset for further analysis.

4. Exploratory Analysis & Visualizations

Several visualizations were developed to extract insights:

A. Top 5 Most Popular Restaurants

- **Visualization:** Bar Chart
- **Metric:** Rating Count
- **Insight:** This chart highlights the restaurants with the highest number of customer reviews, serving as a proxy for popularity.

B. Top 5 Restaurants by Revenue

- **Visualization:** Bar Chart
- **Metric:** Total Revenue (calculated as the sum of sales_amount from the orders table)
- **Insight:** This chart identifies restaurants generating the highest revenue, revealing key performers from a financial perspective.

C. Popularity vs. Revenue Comparison

- **Visualization:** Scatter Plot
- **Metrics:** Rating Count (x-axis) vs. Total Revenue (y-axis)
- **Insight:** The scatter plot was used to examine the correlation between customer reviews and revenue. It addresses the hypothesis that the most popular restaurants may also generate the highest revenue.

D. Price Distribution by Cuisine

- **Visualization:** Treemap
- **Metrics:** Total Price per Cuisine
- **Insight:** The treemap provides an overview of how price is distributed across various cuisines. This visualization helps identify which cuisine types contribute most significantly to overall sales.

E. Revenue Distribution by Cuisine

- **Visualization:** Treemap
- **Metrics:** Total Revenue per Cuisine
- **Insight:** The treemap provides an overview of how revenue is distributed across various cuisines. This visualization helps identify which cuisine types contribute most significantly to overall sales.

Note: A geospatial analysis using built-in geocoding was considered but ultimately excluded due to its complexity. Instead, the focus remained on the above visualizations for clarity and impact.

5. Key Findings

- **Alignment of Popularity and Revenue:** The analysis showed that some of the restaurants with high rating counts also rank among the top revenue generators. However, discrepancies exist, suggesting that while popularity often leads to increased revenue, other factors like pricing strategy and location also play a role.
- **Cuisine Impact:** The treemap revealed that certain cuisines are major revenue contributors, which may indicate market preferences and demand trends.
- **Opportunities for Improvement:** Some restaurants with high popularity did not generate commensurate revenue, suggesting opportunities to optimize menu pricing or operational efficiency to capture greater financial performance.

6. Conclusions and Recommendations

Based on the analysis, the following conclusions and recommendations are made:

- **Strategic Investment:** Focus on supporting restaurants that demonstrate high popularity and strong revenue performance, as they likely have the best operational practices.
- **Menu Optimization:** For restaurants that are popular but lag in revenue, a review of pricing strategies and menu offerings is recommended to better align customer interest with profitability.
- **Cuisine Trends:** Expand or emphasize cuisine types that are identified as revenue drivers to capture market demand.
- **Continuous Monitoring:** Establish regular tracking of key performance indicators (KPIs) to continuously adapt strategies based on evolving consumer preferences and sales trends.

Attachments:
Restaurant.csv
Menu.csv
Orders.csv
Food.csv
Visualizations
Final Report
README