Report to Project Lead: Customer Analysis Segmentation Overview

To: Project Lead
From: Maxwell Afram
Date: 09/15/2024
Subject: Decomposition Report on Customer Data for Zomato Business Performance Analysis

Introduction

As part of my onboarding with Zomato's BI-Analytics Team, I was tasked with analyzing customer data to segment users and understand their purchasing behavior. This report presents the decomposition of the customer dataset and outlines an initial strategy for customer segmentation.

Decomposition of Fields

1. User ID:
    - Purpose: Primary key to track customer records.
    - Next Steps: Verify uniqueness, ensure no duplicates, and correctly link to transaction data.
2. Name & Email:
    - Purpose: Identification and communication management.
    - Next Steps: Clean data to ensure valid email formats; anonymize names for privacy where necessary.
3. Age:
    - Purpose: Segment customers by age groups to analyze behavior.
    - Next Steps: Categorize into age ranges (e.g., 18-24, 25-34) for targeted analysis.
4. Gender:
    - Purpose: Demographic segmentation for potential gender-based trends.
    - Next Steps: Standardize values and manage missing data.
5. Marital Status:
    - Purpose: May provide insights into household composition affecting order behavior.
    - Next Steps: Segment as "Single," "Married," or "Other" and explore correlations with spending patterns.
6. Occupation:
    - Purpose: Occupational categories may indicate different purchasing powers.
    - Next Steps: Categorize broader occupational fields (e.g., "student," "professional") for analysis.
7. Monthly Income:
    - Purpose: Key indicator of purchasing power.

    - Next Steps: Standardize values, convert text to numeric ranges, and segment into income groups.
8. Education Qualification:
    - Purpose: Educational background may inform dining habits.
    - Next Steps: Group into education levels (e.g., "High School," "Graduate") for pattern analysis.
9. Family Size:
    - Purpose: Household size could impact order frequency and value.
    - Next Steps: Segment by family size for behavioral analysis.

Key Questions for the Dashboard

1. Who are Zomato's customers?
    - What are their primary demographic characteristics?
2. What occupational and income-related segments exist?
    - How do income and occupation affect purchasing behavior?
3. How can customers be segmented by age, family size, and income?
    - What are the behavioral differences among these segments?
    - How does income influence restaurant preferences (premium vs. budget)?

Hypotheses

- Income & Restaurant Choice: Higher-income customers may prefer premium dining, while budget-conscious customers might opt for cost-effective restaurants.
- Age & Ordering Behavior: Younger customers are inclined toward fast food, while older customers may favor healthier or fine dining options.
- Family Size & Order Frequency: Larger households are likely to place bulk or frequent orders.
- Occupation Influence: Professionals may order more frequently during workdays, whereas students may show flexible patterns.

Planned Visualizations

1. Demographic Overview:
    - Bar/Pie Charts: Display distribution by age, gender, marital status, education, and occupation.
2. Customer Segmentation:
    - Cluster Analysis (Tableau): Group customers by income, occupation, family size, and age.
    - Heatmaps: Show purchasing behavior by segment.
3. Income vs. Restaurant Type:

    - Scatter/Stacked Bar Charts: Show relationships between income levels and restaurant preferences.
4. Purchasing Behavior:
    - Line/Bubble Charts: Visualize order frequency and size across segments.
5. Family Size & Order Behavior:
    - Box Plots: Depict order size distribution across family sizes and marital status.

Data Preparation and Tools

- Alteryx will handle:
    - Data Cleaning: Normalize monthly income, occupation, and demographic data.
    - Transformation: Segment customers by age, income, and family size.
- SQL Server:
    - Data will be stored and queried using SQL Server, linking customer data with order information for analysis.
- Tableau:
    - For visualizing insights from the cleansed and segmented data.

Data Preparation Steps

1. Data Cleaning in Alteryx:
    - Standardize fields like income and occupation.
    - Handle inconsistencies and missing values in demographic data.
2. Data Transformation:
    - Group age, income, and family size into relevant categories.
    - Merge cleaned customer data with transaction data to analyze behavior.
3. Data Validation & Loading to SQL Server:
    - Ensure data accuracy and load into SQL Server for querying and visualization.

Dashboard Overview

- Customer Demographics: Bar and pie charts will present the overall breakdown.
- Customer Segmentation: Cluster and heatmap visualizations will group customers by key attributes.
- Purchasing Behavior: Line and bubble charts will track order frequency and size.
- Income vs. Restaurant Type: Scatter and stacked bar charts will show how income relates to restaurant choice.
- Family Size & Order Patterns: Box plots will explore how family size influences orders.

Next Steps

- Finalize data cleaning and transformations.
- Conduct customer segmentation analysis.
- Build visualizations in Tableau based on insights.
- Submit the draft dashboard for feedback and revision.

Data Model: Joining Tables To combine the tables for analysis:

1. Common Columns:
    - user_id: To link Users with Orders.
    - r_id: To link Orders with Restaurants.
2. Joining Strategy:
    - Inner Join between Users and Orders on user_id.
    - Inner Join between Orders and Restaurants on r_id.
3. Cleaning Considerations:
    - Ensure consistent data types and handle missing values.

Conclusion

This decomposition report highlights the approach for customer segmentation and analysis. The insights gained will drive informed decisions on customer behavior, spending patterns, and preferences. I look forward to your feedback.

Best regards,
Maxwell Afram
Junior Analyst, Zomato