zomato

Customer Segmentation Analysis

by

Kristy Murry

zomato

BUSINESS DESCRIPTION:

Zomato is a multinational restaurant aggregator company.

BUSINESS OBJECTIVE:

Identify and profile Zomato's customer segments based on demographics (gender, age, monthly income, occupation) and spending behaviors to enhance target marketing efficiency and loyalty programs.

BUSINESS QUESTIONS:

1. What is the demographic distribution (gender, age, monthly income, occupation) of Zomato's customer base?
2. How do spending behaviors differ among customers?
3. What are the different customer segments based on recency, frequency, and monetary value?
4. What marketing strategies could Zomato implement in targeting specific customer segments to retain customers and increase sales?

DATA OVERVIEW

SOURCE: Zomato's internal sales database comprised of customer sales records from 2017 to 2020. Specifically used the following tables:

1. Users: customer demographic data
2. Orders: order records for customers

OVERALL STRATEGY:

1. Explored the demographic and sales data in the User and Orders tables to create Tableau visualizations identifying customers' spending behaviors based on gender, age, monthly, income, and occupation.

2. Examined customer order data in the Orders table and created Power BI visualizations identifying customer segments based on recency, frequency, and monetary analysis.

* Specifics included in Appendices.

Demographic Segmentation Dashboard Description:

The dashboard offers insights into various customer segments, categorized by the following criteria:

Demographic Attributes: Includes gender, age, monthly income, and occupation.

Business Metrics: Covers total customers, total orders, total revenue, average order size, average order frequency, average order value, and customer lifetime value.

Trend Analysis: Examines monthly average order quantities and average sales amounts (USD) to uncover seasonal patterns in customer behavior.

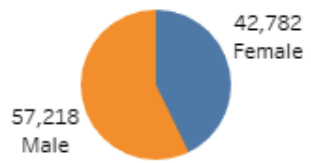
Sales Comparisons: Highlights the relationships between sales amounts (USD) and different age groups, as well as, the average order value and gender.

* Additionally, the data can be filtered by customer occupation, age group, and gender for a more focused analysis.

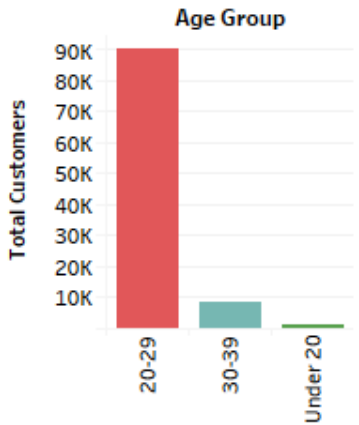
Demographics Overview

Occupation: (All) | Age Group: (All) | Gender: (All)

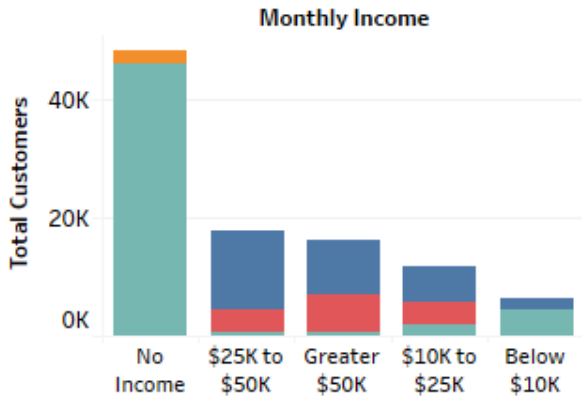
Gender Distribution



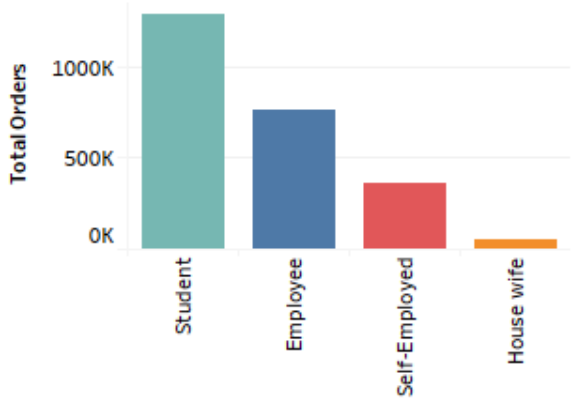
Age Distribution



Monthly Income Distribution

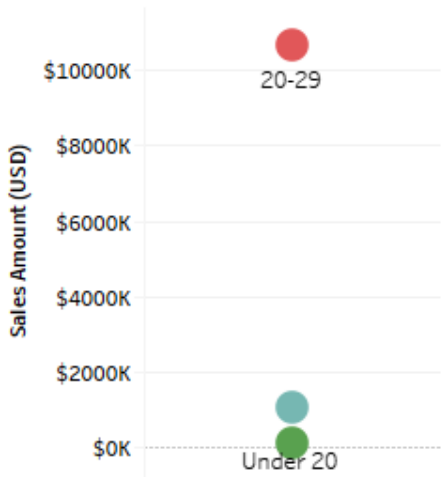


Occupation Distribution

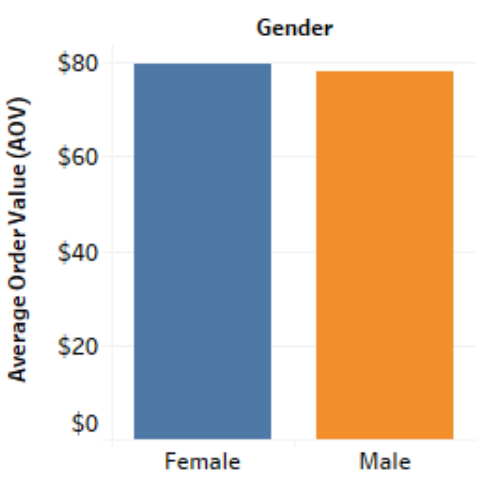


Purchase Behavior by Demographics

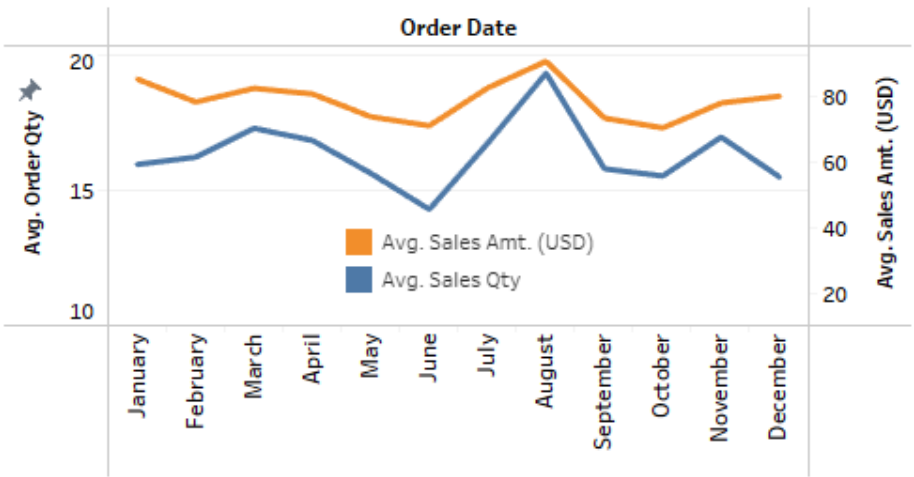
Sales ($) by Age Group



AOV by Gender



Orders Over Time



Sales & Revenue Analysis

Total Customers	Total Orders	Total Revenue	AVG Order Size	AVG (AOF)	AVG Order Value (AOV)	Customer Lifetime Value (CLV)
100,000	2,444,320	$11,839,521	16	1.503	$78.78	$355.20/3 Years

Demographic Insights:

- Majority of customers are **MALES** (57+K) compared to females (42+K) out of 100K. However, females have a higher average order value ($80) compared to males ($78).
- The age group of **20-29** has the greatest number of total customers (89+K) and has the highest total spend ($10+M).
- Customers with **"no monthly income"** make up the bulk at 48+K; with customers making between $25K and $50K coming in second at 17+K.
- **Students** comprise the largest occupation distribution of total orders at 1.2+M out of 2.4+M.
- **August** has the highest average order quantity (19) and the average sales amount ($90).
- Overall, the **average order size** is 16; **average order frequency** (AOF) is 1.5; **average order value** (AOV) is $78.78; and the **customer lifetime value** is $355.20/3 years.

Demographic Recommendations:

- **Male** customers are more likely to respond to promotions tailored to fast/easy meals, so target marketing efforts accordingly to grow this market share. Females have a higher average order value, so upsell higher-value menu items/restaurants to them.
- The age group of **20-29** often prioritizes convenience and speed, leading to more frequent orders from quick-service restaurants. They also engage more with app-based promotions and loyalty reward programs. So, target marketing to this group with app-based promotions and loyalty reward programs from quick-service restaurants.
- Most **students** with **"no monthly income"** have restricted budgets, so prefer restaurants marketing "meal deals" and discounts. They also spend more during breaks (school, holidays, social events) when they're free from school commitments (Ex. **August**). Therefore, target "meal deal" discounts and promotions to students during August to increase sales and order frequency.

RFM Customer Segmentation Dashboard Description: (pg. 1)

The dashboard provides an overview of customer segmentation using RFM analysis. Customers are categorized based on their recency, frequency, and monetary scores, which reflect their purchasing behavior patterns.

RFM Definitions:

- **Recency:** Measures how recently a customer has made a purchase.
- **Frequency:** Measures how often a customer purchases within a given period.
- **Monetary:** Measures how much money (USD) a customer spends during a specific period.

Customer Segments Attributes & Score Ranges:

- **Champions:** Recent buyers, frequent purchases, spend a lot. (**R** 4-5, **F** 4-5, **M** 4-5)
- **Loyal:** Not always recent, buy often, spend consistently. (**R** 3-5, **F** 3-5, **M** 4 3-5)
- **Potential Loyalists:** Recent brand engagement, avg. purchase frequency, avg. spend. (**R** 3-5, **F** 2-5, **M** 1-3)
- **New:** Recently made first purchase, low frequency, low spend. (**R** 3-5, **F** 1-2, **M** 1-2)
- **Promising:** Moderate recency, not yet frequent buyers or top spenders. (**R** 3-5, **F** 1-2, **M** 3-5)
- **At-Risk:** Haven't engaged recently, but bought frequently, and spent consistently in the past. (**R** 1-2, **F** 2-5, **M** 2-5)
- **Cannot Lose Them:** Low recency, but high purchase frequency and spend in the past. (**R** 1-2, **F** 1-5, **M** 3-5
- **Hibernating:** Low recency, low purchase frequency, low spenders.(**R** 1-3, **F** 1-5, **M** 1-3)

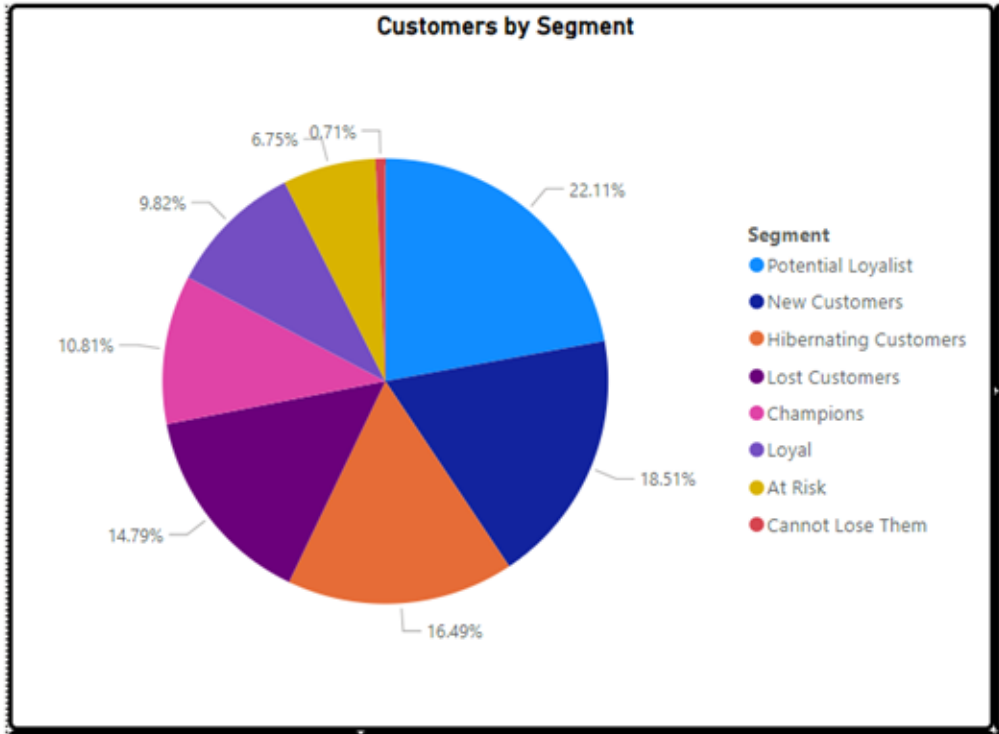
RFM Customer Segmentation Dashboard Description: (pg. 2)

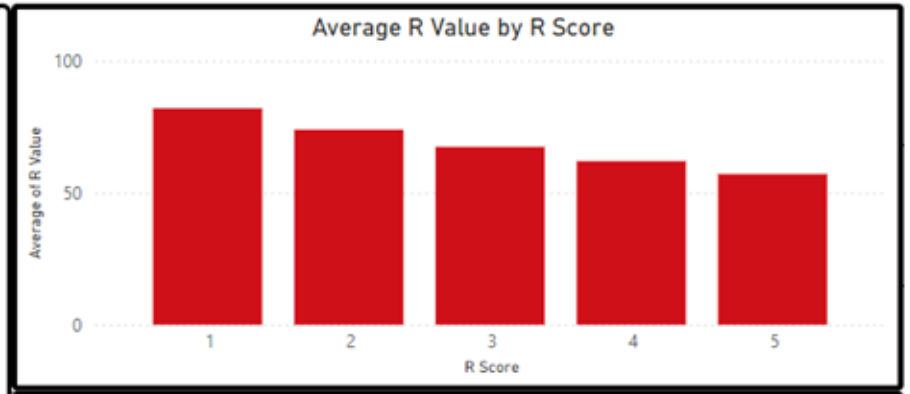
Business Metrics: Covers average recency, avg. frequency, and avg. spend of customers (USD).

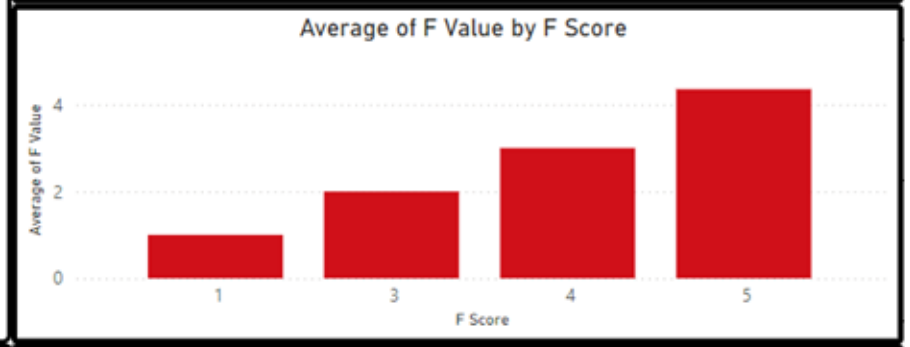
RFM Comparisons: Includes recency value vs. corresponding recency score, frequency value vs. corresponding frequency score, and monetary value vs. corresponding monetary score.

RFM Customer Segmentation Dashboard

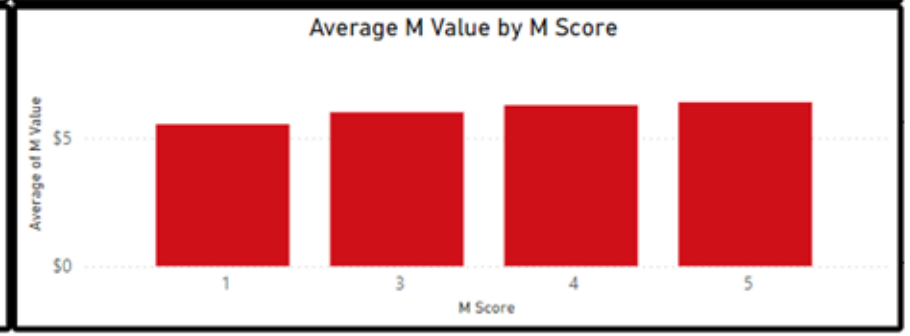
zomato







AVG Recency	AVG Frequency	AVG Spend (USD)
68.28	1.93	$5.87



RFM Insights:

- Top 2 customer segments are the **"Potential Loyalists"** with 22% and **"New Customers"** with 18% of total customers.
  - Zomato may have invested heavily in customer acquisition through marketing campaigns, partnerships with restaurants, and introductory discounts, leading to a significant influx of new users.
  - **"Potential loyalists"** likely result from customers responding positively to these campaigns but haven't yet fully transitioned into loyal or champion customers.
- Bottom 2 customer segments are the **"Cannot Lose Them"** with <1% and **"At-Risk"** with 7% of total customers.
  - Customers at the lower end of the value chain may churn early (before reaching the **"Cannot Lose Them"** stage) due to dissatisfaction, switching to competitors, or reduced interest in online food delivery.
  - This would reduce the size of the **"At-Risk"** segment, as many customers leave the ecosystem entirely before reaching this stage.
- Avg. recency of orders is 68 months (length may be due to older data), avg. frequency is 2 orders during that time, and avg. spend is $5.87.

RFM Recommendations:

- A large **"potential loyalist"** group is a prime opportunity to nurture and convert these customers into loyalists and champions through targeted engagement, personalized communication, and loyalty programs.
- **"New"** customers represent a vital pool for long-term growth if nurtured properly with strong onboarding, easy-to-use interfaces, and attractive offers.
- Be aware of threats, in other words, the small size of the **"Cannot Lose Them"** and **"At-Risk"** groups may create a false sense of security. If they churn, it could significantly impact revenue and brand perception.
- Implement personalized marketing campaigns for **"Potential Loyalists"** and **"New"** customers to accelerate their movement into higher-value segments like **"Champions"**.
- Continuously monitor and address the concerns of **"At-Risk"** customers to prevent churn, even if the segment appears small.

APPENDIX I: Data Specifics

1. Demographics & Sales Analysis Dashboard:

a. All monetary amounts (Rupees and US Dollars) were converted to US Dollars in Excel worksheet before Tableau and Power BI imports because the report is being presented to a US audience.

b. Tables & Relationships: Created relationship between the customer (users) table (user_id column) and the orders table (user_id column) in Tableau.

c. KPI's & Calculations:

- **Total Customers:** CNT(Zomato Users 2)
- **Total Orders/Sales:** SUM(Sales Qty)
- **Total Revenue:** SUM(Sales_Amount_USD)
- **Average Order Size:** AVG(Sales Qty)
- **Purchase Frequency:** COUNT([Order Date]) / COUNTD([User Id])
- **Average Order Frequency (AOF):** AVG(COUNT([Order Date]) / COUNTD([User Id])
- **Average Order Value (AOV):** SUM([Sales_Amount_USD]) / COUNT([Order Date])
- **Customer Lifespan:** DATEDIFF('year', MIN([Order Date]), MAX([Order Date]))
- **Customer Lifetime Value (CLV):** [Average Order Value (AOV)]* [Purchase Frequency] * [Customer Lifespan (CL)]
- **Age Group:** IF [Age] < 20 THEN "Under 20"
  ELSEIF [Age] >= 20 AND [Age] < 30 THEN "20-29"
  ELSEIF [Age] >= 30 AND [Age] < 40 THEN "30-39"
  ELSE "40+"
  END
- **Corrected Occupation:** IF [Occupation] = "self employeed" THEN "Self-Employed"
  ELSE [Occupation]
  END

APPENDIX II: Data Specifics

2. **Recency, Frequency, Monetary (RFM) Customer Segmentation Analysis Dashboard:**

a. Tables & Relationships: Created 3 Tables (Zomato Orders 1, RFM, Segment) and created a Many-to-Many relationship between the RFM (RFM column) and the Segment Table (Scores column) with cross-filter direction of Both.

b. KPI's & Calculations:

- **Last Transaction Date** (Cal. Column):

```
last transaction date =
  MAXX(FILTER('Zomato Orders 1','Zomato Orders 1'[user_id]='Zomato
  Orders 1'[user_id]),'Zomato Orders 1'[order_date])
```

- **RFM Table** (Measure)**:**

```
RFM table = SUMMARIZE(
'Zomato Orders 1','Zomato Orders 1'[user_id],
"R Value",[R Value],
"F Value",[F Value],
"M Value",[M Value])
```

- **R Value** (Measure)**:**

```
R Value = DATEDIFF('Zomato Orders 1'[last transaction date],TODAY(),MONTH)
```

- **F Value** (Measure)**:**

```
F Value =
  CALCULATE(
  DISTINCTCOUNT('Zomato Orders 1'[order_date]),
  ALLEXCEPT('Zomato Orders 1', 'Zomato Orders 1'[user_id])
```

- **M Value** (Measure)**:**

```
M Value =
  var TotalSales = SUM('Zomato Orders 1'[sales_amount_usd])
  var TotalQuantity = sum('Zomato Orders 1'[sales_qty])
  Return
  DIVIDE (TotalSales,TotalQuantity,0)
```

APPENDIX III: Data Specifics

2. Recency, Frequency, Monetary (RFM) Customer Segmentation Analysis Dashboard:

a. KPI's & Calculations (con't):

- **Recency Score** (Cal. Column)**:**

```
R Score =
 SWITCH (
  TRUE (),
  [R value] <= PERCENTILE.INC ( 'RFM table'[R Value], 0.20 ), "5",
  [R value] <= PERCENTILE.INC ( 'RFM table'[R Value], 0.40 ), "4",
  [R Value] <= PERCENTILE.INC ( 'RFM table'[R Value], 0.60 ), "3",
  [R value] <= PERCENTILE.INC ( 'RFM table'[R Value], 0.80 ), "2",
  "1" )
```

- **Frequency Score** (Cal. Column):

```
F Score =
 SWITCH (
 TRUE (),
 [F value] <= PERCENTILE.INC ( 'RFM table'[F Value], 0.20 ), "1",
 [F value] <= PERCENTILE.INC ( 'RFM table'[F Value], 0.40 ), "2",
 [F Value] <= PERCENTILE.INC ( 'RFM table'[F Value], 0.60 ), "3",
 [F value] <= PERCENTILE.INC ( 'RFM table'[F Value], 0.80 ), "4",
 "5" )
```

- **Monetary Score** (Cal. Column):

```
M Score =
SWITCH (
TRUE (),
[F value] <= PERCENTILE.INC ( 'RFM table'[F Value], 0.20 ), "1",
[F value] <= PERCENTILE.INC ( 'RFM table'[F Value], 0.40 ), "2",
[F Value] <= PERCENTILE.INC ( 'RFM table'[F Value], 0.60 ), "3",
[F value] <= PERCENTILE.INC ( 'RFM table'[F Value], 0.80 ), "4",
"5")
```

- **RFM Scores** (Cal. Column):

```
RFM = 'RFM table'[R Score]& 'RFM table'[F Score]&'RFM table'[M Score]
```