

Project - Online Food Delivery Analysis

This project analyzes customer behavior and order data using Python for data analysis and visualization. It explores order frequency, platform preferences, discount usage, customer demographics, and other order-related metrics. The project applies various statistical and visual methods to uncover trends in customer satisfaction, payment methods, delivery speed preferences, and order sizes. By analyzing these factors, the project provides valuable insights into customer preferences, helping to inform marketing strategies, improve delivery services, and enhance overall customer experience.

Here are some potential analyses we are going to perform:

1. Discount Usage Trends by Age:

- Younger age groups (e.g., 18-24) tend to use discounts more frequently, while older groups (e.g., 55+) show a lower preference for discounts.
- There is a noticeable variation in discount usage based on age, which can be crucial for targeting specific age demographics with promotions.

2. Order Frequency and Preferences:

- Customers with frequent orders tend to prefer faster delivery times and are generally more satisfied with their service, highlighting the importance of maintaining timely deliveries for high-frequency customers.
- Order frequency varies across different platforms, with some platforms seeing higher activity than others.

3. Payment Method Analysis:

- The majority of customers prefer specific payment methods (like credit cards or digital wallets), which could inform how payment options are presented in marketing campaigns.
- Payment method preferences might also correlate with customer satisfaction and order sizes, suggesting that payment convenience is a factor in customer satisfaction.

4.Age Group and Satisfaction:

- Satisfaction ratings vary across age groups, with younger customers reporting slightly higher satisfaction levels. This insight could guide customer service improvements or adjustments in how different age groups are engaged.

5.Order Size Insights:

- Larger order sizes are often associated with higher satisfaction ratings and preferences for specific delivery speeds, suggesting that larger orders may require more careful attention to delivery logistics and customer expectations.

Dataset Overview:

```
CREATE TABLE dataset(  
Customer_ID VARCHAR(20) NOT NULL PRIMARY KEY,  
Age INT NOT NULL,  
Preferred_Platform VARCHAR(25) NOT NULL,  
Order_Frequency VARCHAR(25) NOT NULL,  
Order_Type VARCHAR(2) NOT NULL,  
Delivery_Time_Preference VARCHAR(25) NOT NULL,  
Delivery_Speed_Preference VARCHAR(25) NOT NULL,  
Satisfaction_Rating INT NOT NULL,  
Payment_Method VARCHAR(25) NOT NULL,  
Discount_Usage VARCHAR(25) NOT NULL,  
Restaurant_Rating INT NOT NULL,  
Order_Size VARCHAR(25) NOT NULL  
);
```

Database Connectivity:

To perform the analysis, we'll establish a connection to the MySQL database containing the ecommerce data. We'll use the appropriate Python libraries, such as pandas and mysql, to fetch and manipulate the data directly from the database.

```
import mysql.connector

import pandas as pd

# Connect to MySQL

connection = mysql.connector.connect(

    host="localhost",    # Replace with your host (e.g., "127.0.0.1")

    user="root", # Your MySQL username

    password="1234", # Your MySQL password

    database="project" # Optional: specify database to connect to

)

# Check connection

if connection.is_connected():

    print("Connected to MySQL!")

#create cursor object to execute sql queries

cursor = connection.cursor()

# Query the data from walmart data sale

cursor.execute("SELECT * FROM dataset")

#After fetching data from the database we are storing it into Pandas DataFrame

dataset = pd.DataFrame(cursor.fetchall(), columns=[desc[o] for desc in cursor.description])

print(dataset.head(5))
```

Output

Description:

```
customer_data = pd.DataFrame(cursor.fetchall(), columns=[desc[0] for desc in cursor.description])
```

cursor.fetchall():

This part of the code fetches all the rows from the result set obtained from a database query using the cursor object. The `fetchall()` method retrieves all the rows as a list of tuples.

columns=[desc[0] for desc in cursor.description]: This part creates a list of column names for the DataFrame. It uses a list comprehension to iterate over the `cursor.description`, which is a list of 7-item tuples describing the columns in the result set. The `[desc[0] for desc in cursor.description]` extracts the first element (column name) from each tuple in the `cursor.description` and creates a list of column names.

pd.DataFrame(...): This part creates a Pandas DataFrame using the `pd.DataFrame()` constructor. It takes the fetched data (result of the query) and the list of column names as arguments. The DataFrame is assigned to the variable `customer_data`.

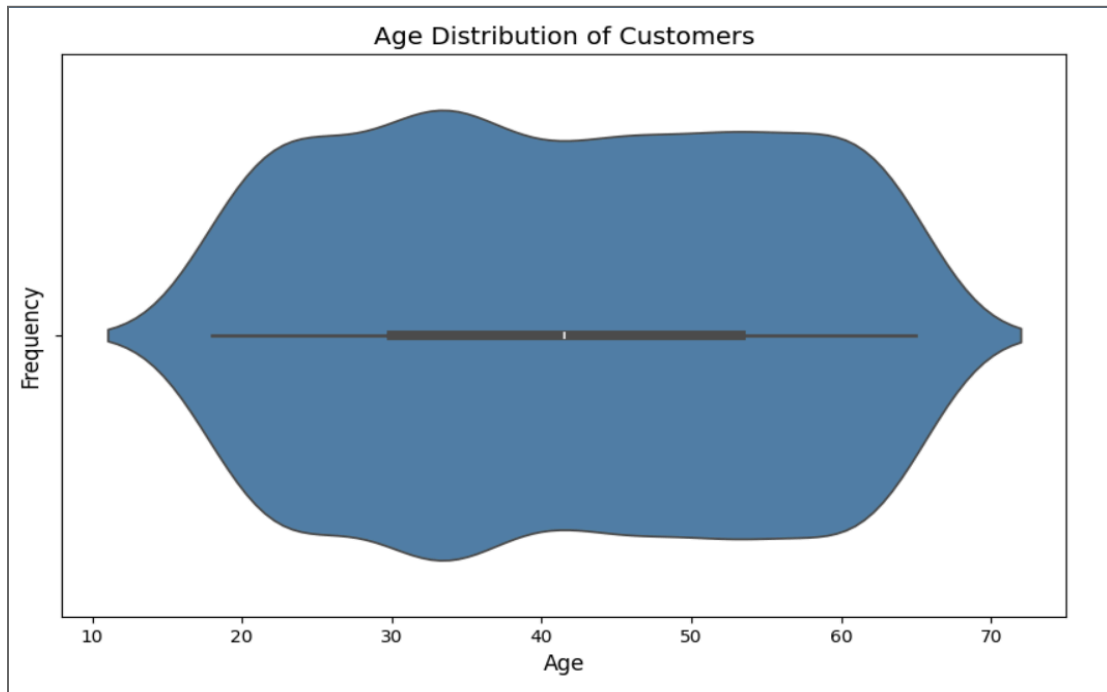
Data Cleaning:

Before proceeding with the analysis, let's perform some basic data cleaning:

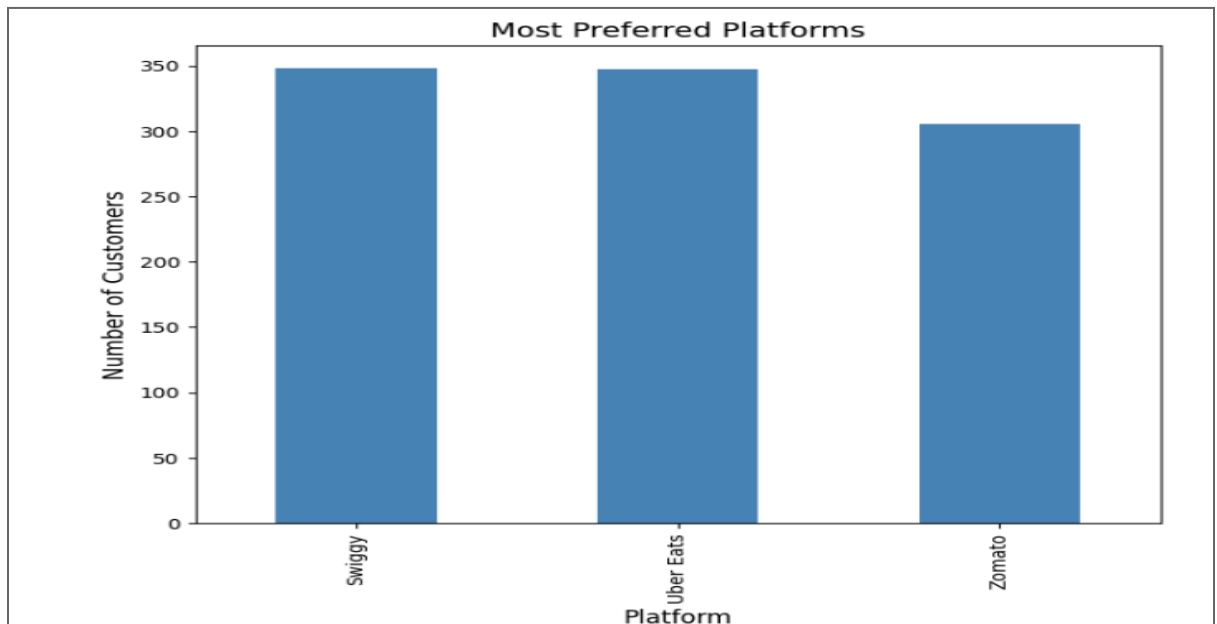
- Check for missing values in each table.
- Ensure data types are appropriate for each column.
- Handle any outliers or inconsistencies.

Exploratory Data Analysis (EDA) and Visualization:

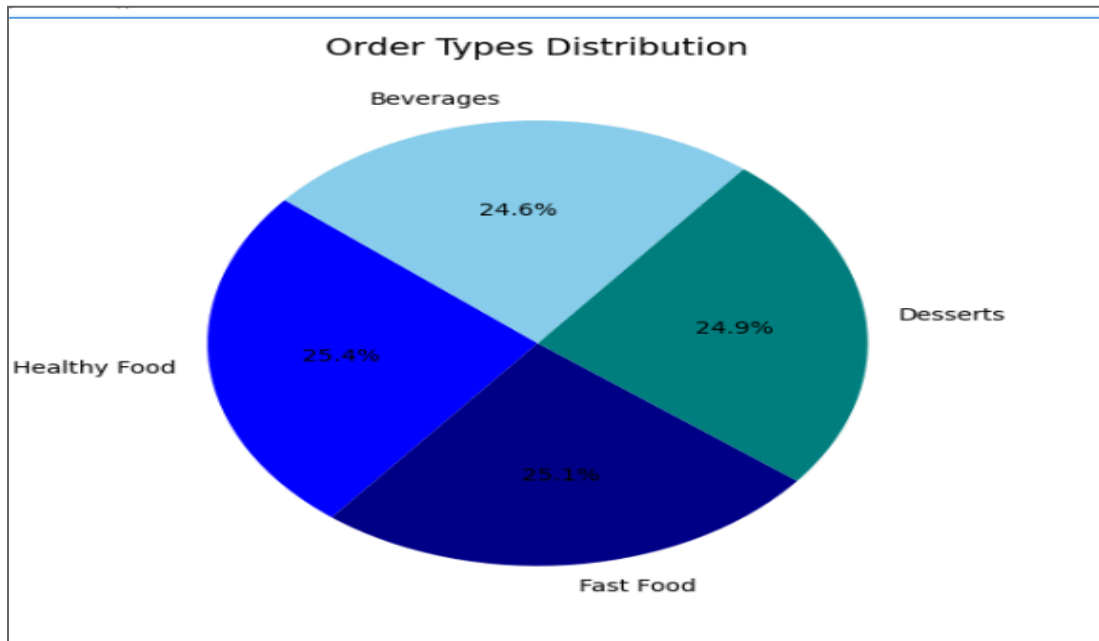
1.Find the age distribution of customers



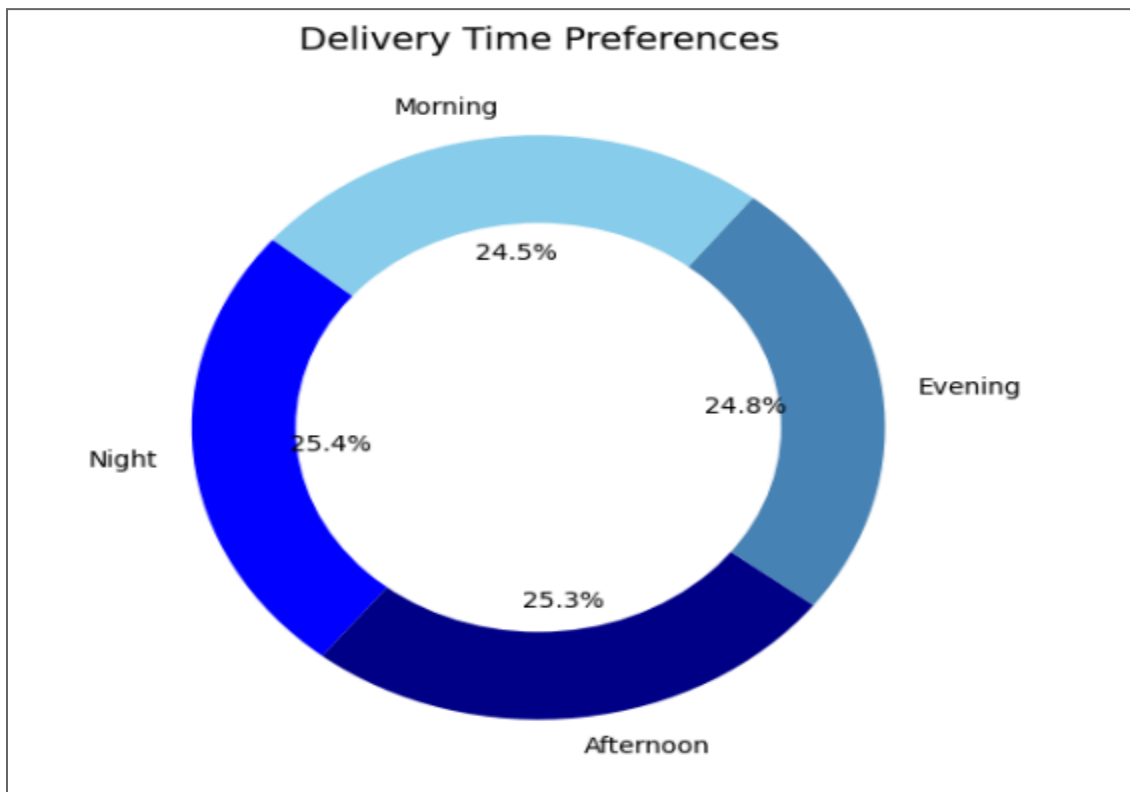
2.Find the most preferred platform for ordering food.



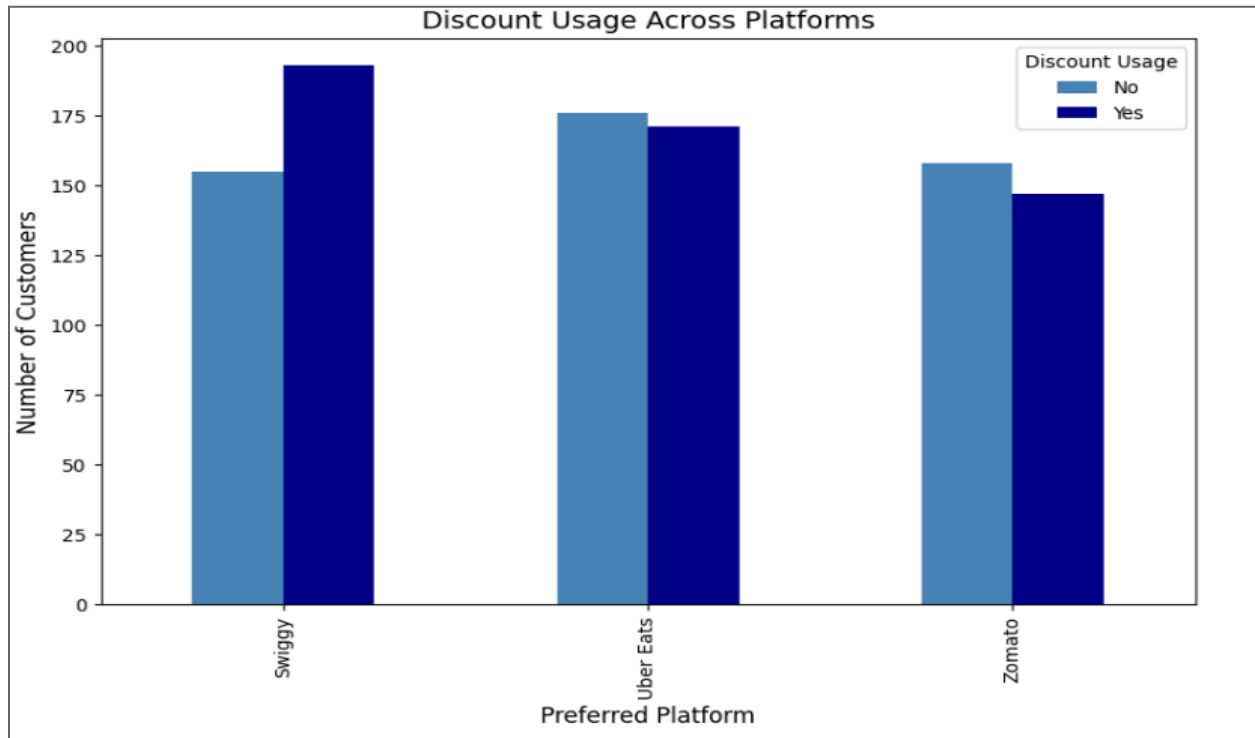
3. Find the frequency of order types.



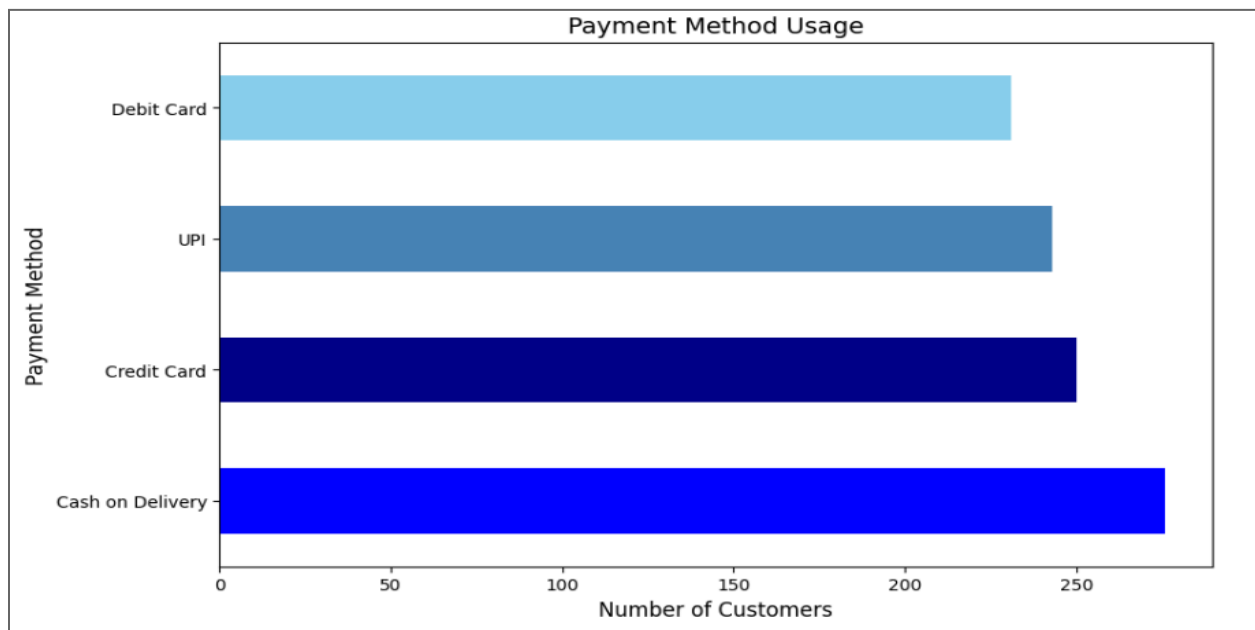
4. The most popular delivery time preference.



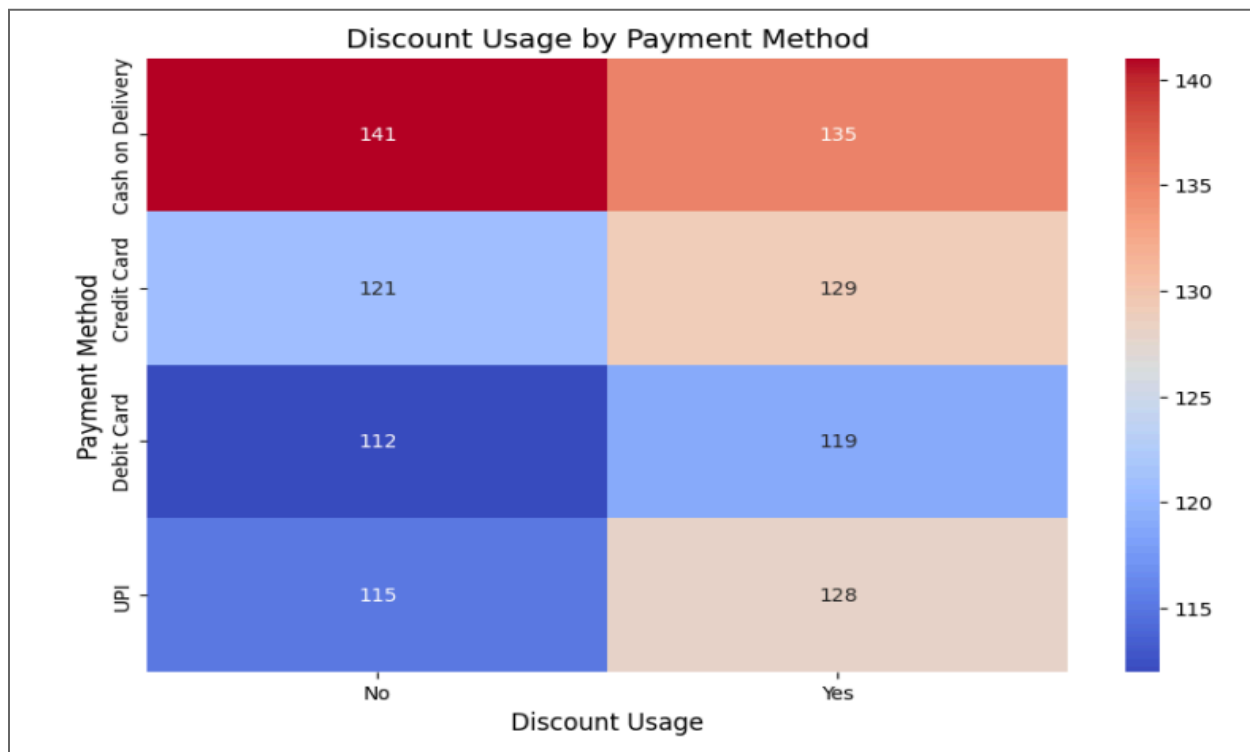
5. Discount usage across platforms.



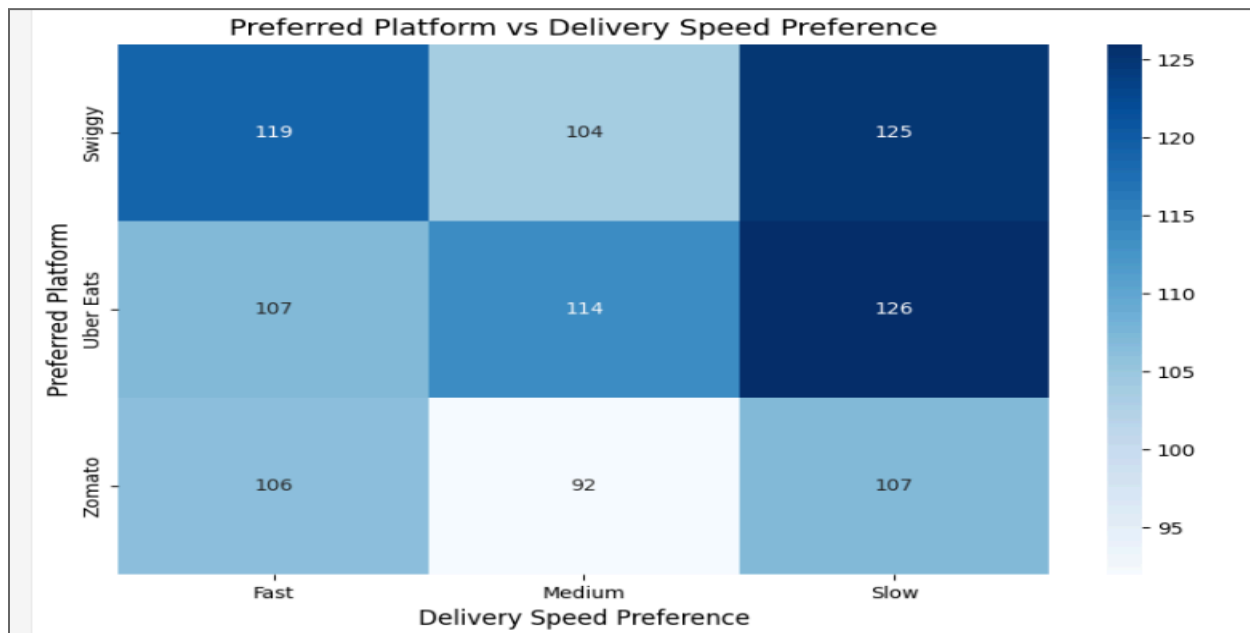
6. Analyzing payment method usage.



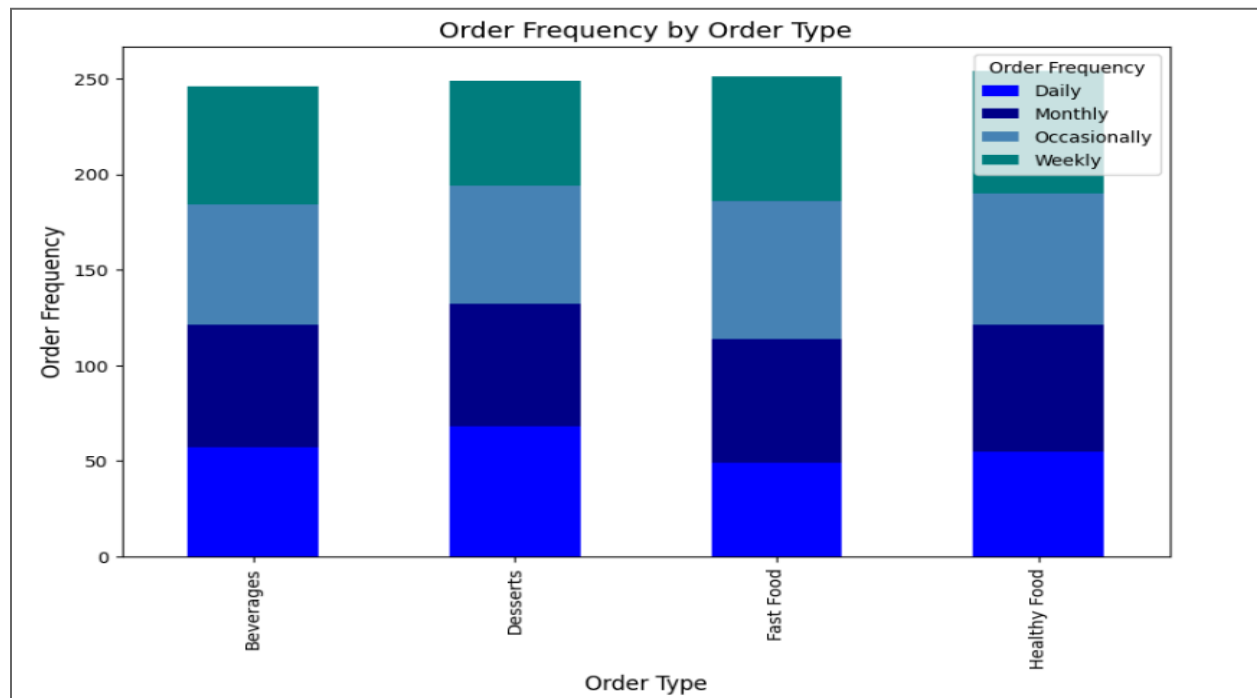
7. Discount usage varies by payment method.



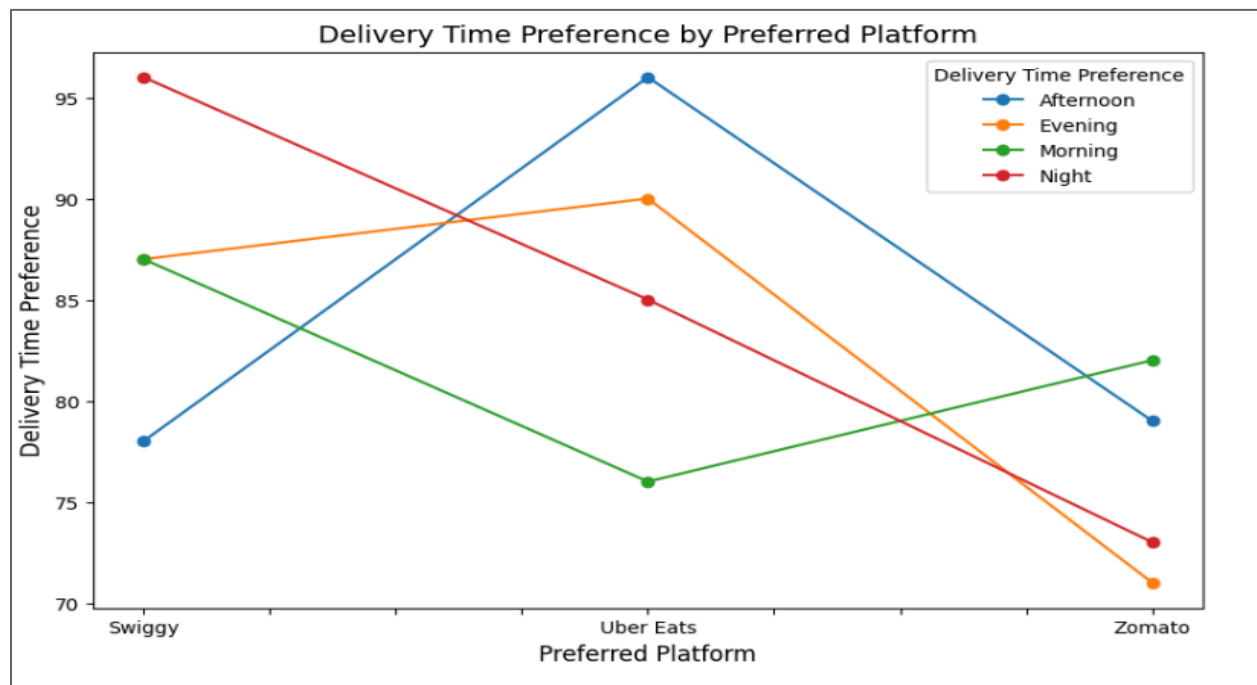
8. Preferred platform vs Delivery speed preference.



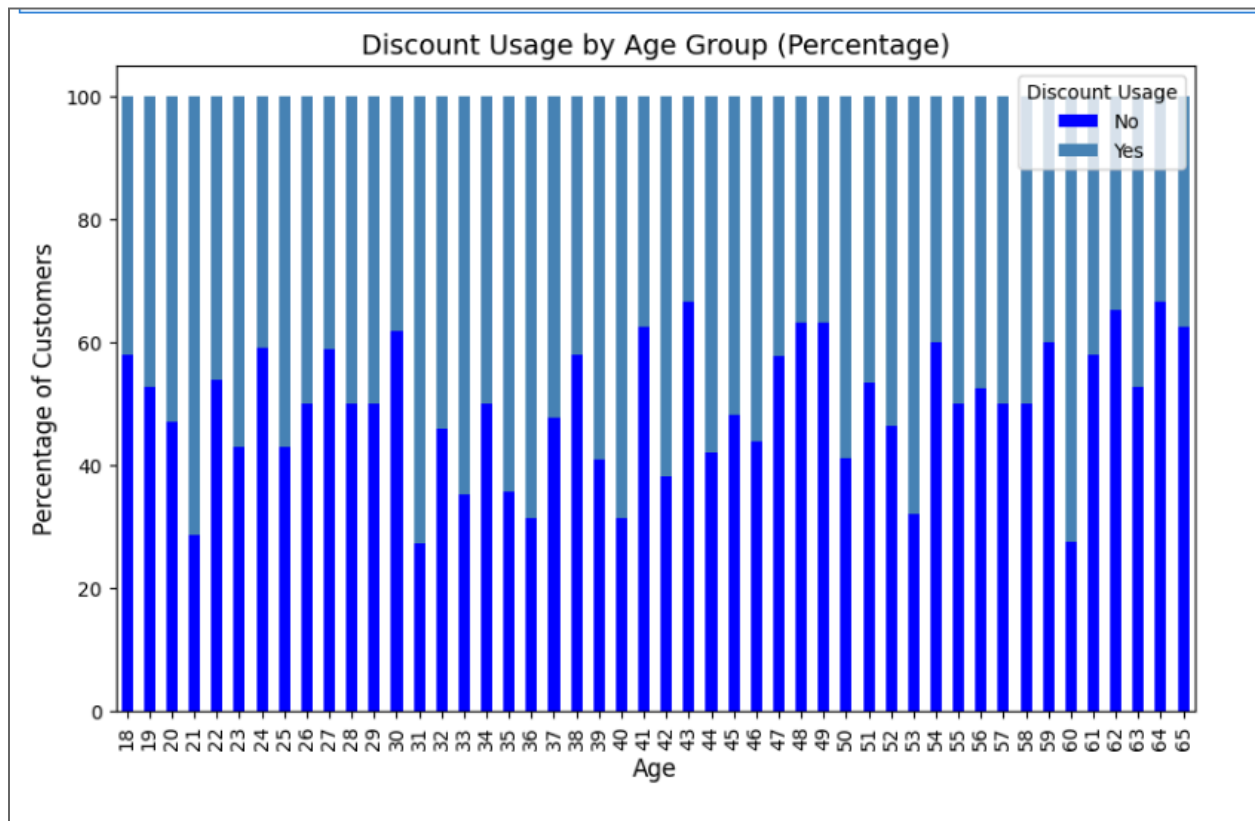
9. Order Frequency varies by Order Type.



10. Delivery Time Preference varies by Preferred Platform.



11. percentage of customers using discounts for different age groups.



Conclusion:

The analysis highlights key insights into customer preferences across different platforms. It reveals that order frequency and size vary, with some platforms showing higher engagement and larger orders. Customers also exhibit differing satisfaction levels, payment methods, and discount usage based on the platform they choose. By examining these trends, businesses can tailor their strategies for each platform, enhancing customer satisfaction and potentially driving greater sales. Additionally, understanding these patterns can guide decisions on promotional offers and service improvements to meet customer expectations more effectively.

