**Diwali Sales Analysis**

**Debojyoti Balial**

**Abstract:**

Our feature-based analysis of the database can help identify the key demographic groups contributing to maximum sales.

**Keywords*:*** Exploratory Data Analysis (EDA), Distribution.

# 

# 1. Problem Statement

We have a dataset which contains the below columns –

* User\_ID: An unique Id to identify the customer.
* Cust\_Name: Customer’s name against the User ID.
* Product\_ID: Unique Id against each product.
* Gender: Gender of the customer.
* Age Group: Age group the customer belongs to.
* Age: Exact age of the customer.
* Marital Status: Marital status of the customer. 1 represents married, 0 represents unmarried.
* State: The state of the customer.
* Zone: There are 5 zones. The zones are- Northern, Southern, Western, eastern, Central
* Occupation: Occupation of the customer
* Product category: Category of the product.
* Orders: How many orders one customer has placed.
* Amount: Spent amount on orders.
* Status: Blank column. Not relevant to the dataset.
* Unnamed1: Blank column. Not relevant to the dataset.

The main objective is to build an analysis, which could help the organization to understand which demography is actually contributing to sale during Diwali time. This would help the organization to target that key demographics with better deals and better understanding of the buying pattern.   
  
We are trying to find out here -

* Gender-wise demographics: The gender which contributing to maximum order and sale amount.
* Age Group: The age group which contributing to maximum order and sale amount.
* Occupation: How occupation affecting the sales pattern.
* State-wise sales: State-wise contribution to sales.
* Product Category: Product category-wise contribution to sales.
* Customer: Customers who has spent most during Diwali time.
* Age and Gender: How age and gender affecting the spending pattern.

**2. Introduction**

Diwali is one of the biggest festival in India. Diwali is celebrated in almost all the states in India. In this study we are trying to find out the spending pattern of the key demographies.

# 3. Exploring the database

**We have provided with one databases**

* Shape of this database is (11251, 13).
* Out of this five columns we have are numeric

# 4. Data Cleaning and Removing Null Values:

* We found out that there are 2 columns (Status, Unnamed1) are irrelevant to the dataset. So we have removed it.
* We have removed the 12 null values from the dataset.

# 5. Description of The Dataset:

* There are total 11239 entries in the dataset.
* Spending Amount on

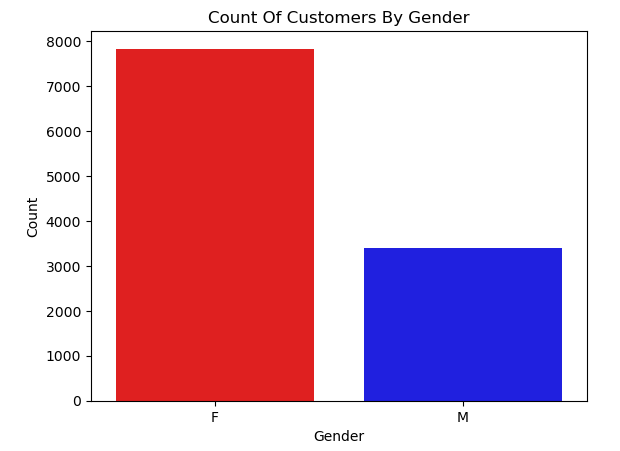
1. Average - 9453.610553
2. Maximum - 23952.000000

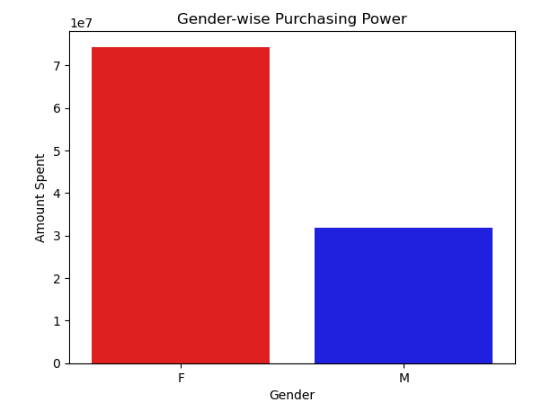
* Orders –

1. On average orders placed by a single customer - 2.489634
2. Maximum orders placed by a single customer – 4

# 6. Analysis on Gender -

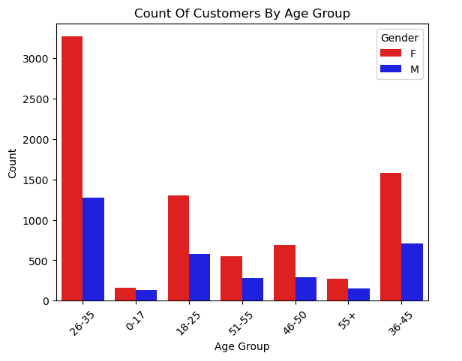
* From the below 2 carts we can see that Female customers are placing maximum number of orders and female customers are spending more amount.

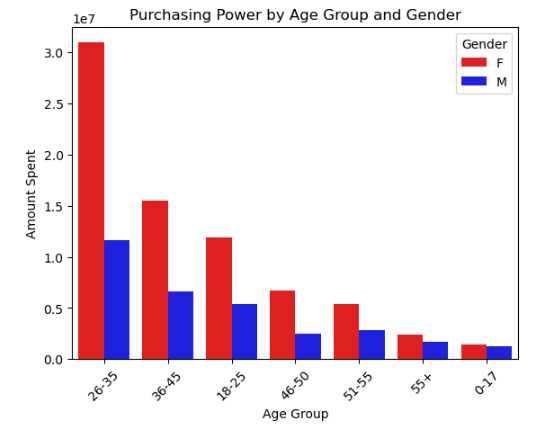




# 7. Analysis on Age Group:

* Age Group 26-35 and 36-45 are the two bins which are ordering and spending most.

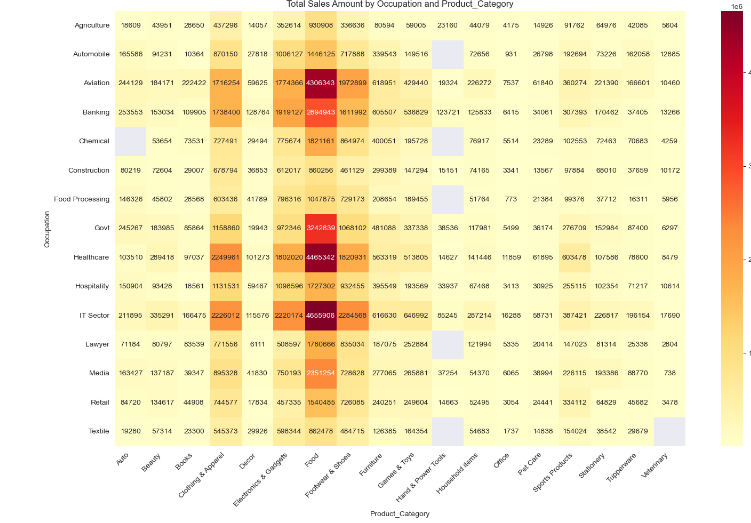




# 8. Analysis on Occupation and Product Category:

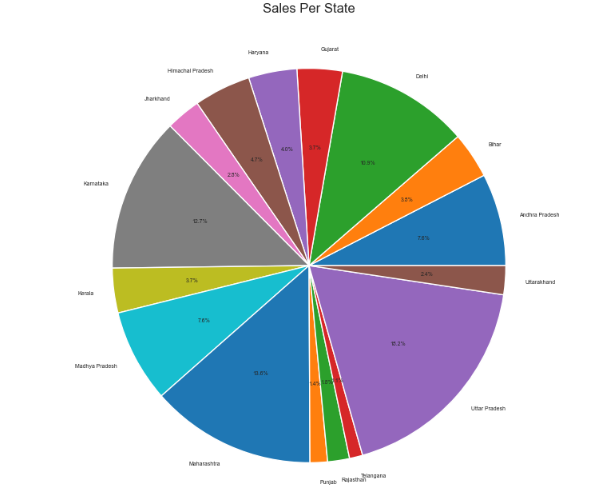
From the below Heat map we can Understand that -

* Members from different occupation are actually spending maximum on Food Items.
* It Sector, Healthcare, Banking and Aviation employees are the ones who are maximum spending and they are also contributing more in – Clothing & Apparel, Electronics & Gadgets, Furniture.
* Healthcare sector spending quite more in Stationeries.



**9. Analysis on States -**

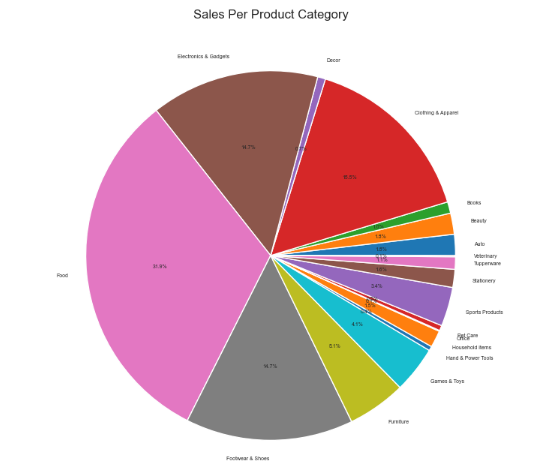
* + Since Uttar Pradesh is the most populous state so they are relatively spending more.
  + Karnataka is contributing quite a big portion



.

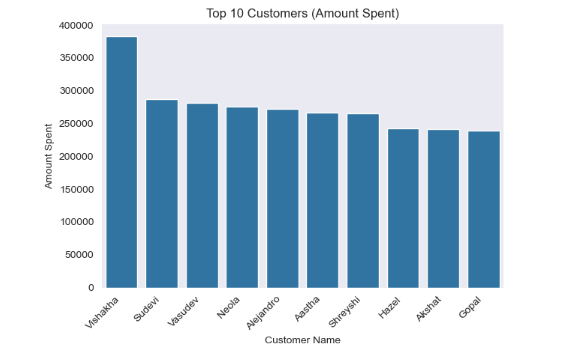
**10. Analysis on Product Category -**

* + While food is the largest category (since in Diwali time lot of the people actually gift food items to their neighbor, relatives), Clothing & Apparel comes in 2nd position.
  + Footwear & Shoes, Electronics & Gadgets are contributing 14.7% each in the total sales.



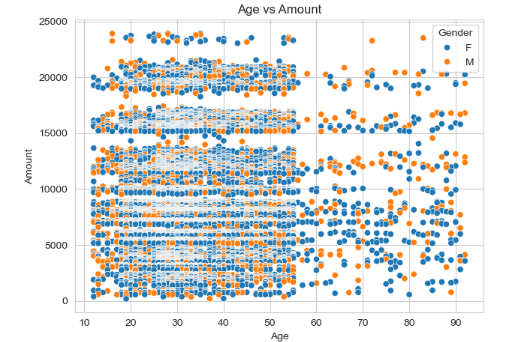
1. **Customers with Maximum Spending -**

* Top 10 customers who have spent most in Diwali time –



**12.Age-wise Spending Habit -**

* Find the age-wise spending habit –



**Conclusion:**

The key insights from the Diwali sales analysis are:

* Female customers are contributing to the maximum orders and highest spending during the Diwali season.
* Age groups 26-45 are the biggest spenders, accounting for the majority of orders and sales revenue.
* Certain occupations like IT, healthcare, banking, and aviation employees are the top spenders, particularly on categories like clothing, electronics, and furniture.
* Uttar Pradesh is the top contributing states in terms of Diwali sales.
* Food items, clothing, footwear, and electronics are the top-selling product categories during the Diwali festival.
* The top 10 highest-spending customers account for a significant portion of the total Diwali sales.

These insights can help the organization better target their Diwali offers and marketing campaigns to the key demographic groups and product categories that drive the maximum sales during this festival season.