

E-COMMERCE DATA ANALYSIS

To Integrate and analyse the datasets(Sales, Product, Customer and Store), Lets Consider sales table as a fact table and Product, Customer, store tables as a dimensional table. Customer table contains customer ID, Name, CITY, Country, Gender and Loyalty level. Product table contains Product ID, Product name, Brand, Category, Cost and Stocks. Store table contains Store ID, Store name, Region, City, and Store type. Sales table contains sales ID, Store ID, Product ID, Customer ID, Quantity, Unit price, Discount, Payment and Total amount.

In Customer table, we have to clean and transform the following columns. In customer ID column, replace CUST to C using find and replace, then in country column change United States of America to USA, then the customer name should be clean trim and proper, and handling missing values in loyalty level by using formula =if(isblank(h2),"Unknown",h2).

In product table, we have to change P as PROD in product ID column, and handling missing values in both cost and stock column using formulas [cost column = if(isblank(f2),avg(\$f\$2: \$f\$101),f2)] and stock column = if(isblank(g2),avg(\$g\$2: \$g\$101),g2)

In Sales table, we handling missing values in unit price, discount, quantity and total amount. For quantity = if(isblank(f2), 1,f2) consider customer a least buy one product, unit price=if(isblank(g2),avg(\$g\$2:\$g\$2001),g2) and for total price using formula Total amount = [quantity*unit price]*[1-discount].

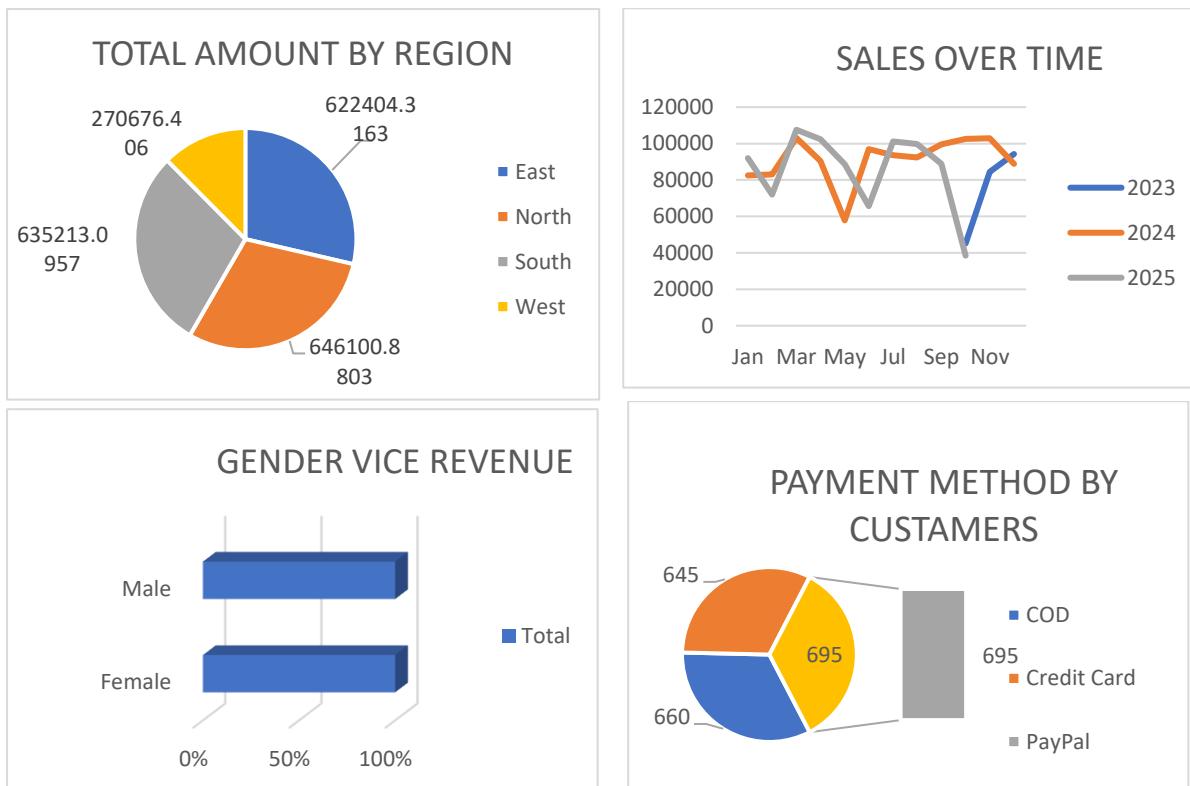
For further transformation, lookup the cost column in product table to sales table to find the profit, profit=vlookup(d2,product_dim[#all],6,false).and using data modelling build relation to all the table(sales as fact and others as dimension).

Creating new column called Profit,using formula profit=total amount-total cost. Using conditional formatting to the profit column.

E-COMMERCE DATA ANALYSIS

INSIGTHS FROM DASHBOARD

- 1) **Average Items per transaction** – 2.4 units most customer shop between 2 to 3 times.
- Average unit price** is \$513. This shows that stores sell mid to high level products.
- 2) **Region vice profit**, North generate more revenue west generate low.
- 3) **Sales Over the time** – March 2024 to 2025 are peak months, October 2023, October 2025 are low months.
- 4) **In Payment Method** – Paypal is highly used by the customers compared to credit card and COD.
- 5) **Sports category** generate more revenue, Beauty generate low.
- 6) This business has **more female custamers** compared to male



E-COMMERCE DATA ANALYSIS

1)This business generated **\$2170907 total revenue** and **\$879230 in total profit**.The more revenue generated in **NORTH REGION** and **PAYPAL** is most preferable payment method.

2)"**MARCH SPIKE**",Sales consistently peak in march and **SPORTS** sold more.

3)**SALES FORECAT**:Based on recent trends,forecast predicts sales around **75000 TO 80000** per month.

4)**THE CONCLUSION**:Focus sports promotion in **NORTH** and **SOUTH** region,wherethey are already in their peak.

Encourage **PAYPAL** over other payment methods and give some offers to promote the payment methods.