

Vrinda Store Data Analysis of year 2022

We are given a dataset of a store having different attributes including Order ID, Channel of sale, gender, age, amount and geographic region of the business.

OBJECTIVES:

Vrinda store wants to create an annual sales report for 2022. So that Vrinda store can understand their customers and grow more sales in 2023.

QUESTIONS TARGETED TO BE ANSWERED:

- Compare the sales and orders using a single chart.
- Which month got the highest sales and orders?
- Who purchased more – Men or Women in 2022?
- What is the different order status in 2022?
- List top 10 states contributing to the sales?
- Relation between age and gender based on number?
- Which channel is contributing to maximum sales?
- Highest selling category? Etc.

To answer these questions, we will go through all following process:

Data Cleaning, Data processing, Data analysis, Data visualization followed by report making and suggesting insights and next steps.

DATA CLEANING:

- First, we saw data, its data points if it can solve our business problems.
- Now, check if all data is correct by checking for missing data, duplicates and null values.
- Now, check for Customer ID column if all values are numeric and also check for null values.
- Gender ID are not consistent as they contain Men, M, Women and W. Apply filter and make all data consistent. i.e. M & W be converted to Men and Women using Filter and Find & Replace tools.
- Check for Age column if all the values are numeric value.
- In Date column check for outliers
- Check for consistency and null values of the columns status, channel, SKU, Category, Size, Quantity, Currency etc. We changed in quantity column from One to 1 and two to 2.

DATA PROCESSING:

- It would be more logical to process age data after diving in different groups rather than analyzing for each customer age.
- Create new column beside Age naming Age Group. Use formula:
=IF(E2>=50,"Senior", IF(E2>=30,"Adult", "Teenager"))
- After applying this formula copy all column and "paste as value" so that it will not let the system slow.
- We also extract month from Date column to answer a question from list. So we created a Month column using this formula: =TEXT(G2,"mmm"). It will change date to month eg., DEC, FEB etc.

DATA ANALYSIS:

- To analyze data, we will create different pivot tables. To do this, INSERT>Pivot Table> select attributes. Following pivot tables are created:
 - ❖ Orders Vs. Sales
 - ❖ Men Vs. Women
 - ❖ Order Status
 - ❖ Sales top 5 states
 - ❖ Age & Gender
 - ❖ Channels
- After selecting pivot chart data "Pivot Table Analyze" tab appears, select Pivot chart to create chart from table.
- Make required changes to create dashboard from all the charts.
- Now, add slicer on the dashboard data. We can only add slicer if it is a pivot table.

INSIGHTS:

- Women are more likely to buy compared to men (~65%)
- Maharashtra, Karnataka and Uttar Pradesh are in the top 3.
- Adult age group (39-49 yrs.) is max contributing (~50%)
- Amazon, Flipkart and Myntra channels are max contributors.

Final Conclusion to improve Vrinda store sales:

- Target women customers of age group (30-49 yrs.) living in Maharashtra, Karnataka and Uttar Pradesh by showing ads/offers/coupons available on Amazon, Flipkart and Myntra.