VRINDA STORE SALES ANALYSIS

Overview

Vrinda Store, established in 2005 in Punjab, is now a leading dress house in India, selling their products nationwide through online platforms such as Amazon, Flipkart, and Myntra etc.

About This Dataset

This dataset is a comprehensive collection of all sales data for the year 2023. It serves as a valuable resource for analysing sales trends, predicting future sales, and assessing the importance of various online platforms for the store.

This dataset contains comprehensive information on sales transactions for the year 2023, offering a diverse and extensive representation of the store's operations. It includes a variety of details such as Order ID, Customer ID, Gender, Age, Date of Order, Order Status, SKU, Online Platform, Category, Size, Quantity, Currency, Amount, and Shipping City. This rich dataset provides valuable insights into the store's sales performance, customer demographics, and the effectiveness of different sales channels.

Business Problem

Our client, Vrinda Store, aims to create an annual report on the sales of 2023. This report will help Vrinda gain a deeper understanding of their customers and develop strategies to grow their sales. By leveraging these insights, Vrinda Store aspires to become the biggest supplier of online dresses by the end of the 2024.

Objectives

The objective of this analysis is to create a comprehensive annual report on the sales of 2023 for Vrinda Store. This report aims to:

1. **Understand Customer Demographics and Behaviours**: Analyse the customer data to identify key demographics, purchasing patterns, and preferences.
2. **Evaluate Sales Performance**: Assess the overall sales performance across different categories, sizes, and sales channels.
3. **Identify High-Performing Products and Categories**: Determine which products and categories are driving sales and contributing the most to revenue.
4. **Analyse the Impact of Online Platforms**: Evaluate the effectiveness of various online sales channels (Amazon, Flipkart, Myntra, etc.) in reaching customers and generating sales.
5. **Track Geographic Distribution of Sales**: Examine the geographic distribution of sales to understand which regions are most lucrative and where there is potential for growth.
6. **Identify Areas for Improvement**: Highlight any gaps or areas where the store can improve its operations, marketing strategies, and customer engagement.
7. **Develop Growth Strategies**: Provide actionable insights and recommendations to help Vrinda Store enhance its market presence and achieve its goal of becoming the leading supplier of dresses in the state by the end of 2024.

Research Questions

- Compare the sales and orders?
- Which month got the highest sales and orders?
- Who purchased more men or women in year 2023?
- What are different order status in 2023?
- List Out the top 10 states the are massively contributing to our sales?
- Relation Between age and gender based on the number of sales?
- Which channel is contributing to most of the sales?
- Highest selling category?

Hypothesis

1. **Demographic Influence**: Certain demographic factors, such as age and gender, significantly impact purchasing behaviors and sales volume.
2. **Platform Impact**: Sales performance varies across different online platforms, with some contributing more to overall sales.
3. **Geographic Trends**: Sales are concentrated in specific geographic regions, indicating higher demand in those areas.

Tools

- **Excel**

Purpose: Initial Data cleaning and preparation and visualization.

Usage: Removing Duplicates, Handling missing values, data formatting and Dashboard Creation.

Data Source

I received this dataset from my client through Google Drive. I will upload it to my drive and share the link with you in case you need to check the dataset. You can follow the link below to access it.

https://drive.google.com/drive/home

Information regarding the columns used for the analysis

- **Index**: A unique identifier for each record.
- **Order ID**: The unique identification number for each order.
- **Customer ID (Cust ID)**: The unique identifier for each customer.
- **Gender**: The gender of the customer.
- **Age**: The age of the customer.
- **Date**: The date of the transaction.
- **Status**: The status of the order (e.g., completed, pending, canceled).
- **Channel**: The sales channel through which the order was placed (e.g., Amazon, Flipkart, Myntra).
- **SKU**: The Stock Keeping Unit, identifying the product.
- **Category**: The category of the product.
- **Size**: The size of the product.
- **Quantity (Qty)**: The quantity of the product ordered.
- **Currency**: The currency in which the transaction was made.
- **Amount**: The total amount of the transaction.
- **Shipping City (ship-city)**: The city to which the order was shipped.
- **Shipping State (ship-state)**: The state to which the order was shipped.
- **Shipping Postal Code (ship-postal-code)**: The postal code of the shipping address.
- **Shipping Country (ship-country)**: The country to which the order was shipped.
- **B2B**: Indicator if the order was a Business-to-Business transaction.

Data Preprocessing

Data processing involves collecting, cleaning, transforming, and organizing data to make it suitable for analysis. It is needed for data analysis to ensure accuracy, consistency, and completeness of data, allowing meaningful insights to be drawn. Without proper data processing, analysis results may be unreliable and lead to incorrect conclusions.

Data Cleaning:

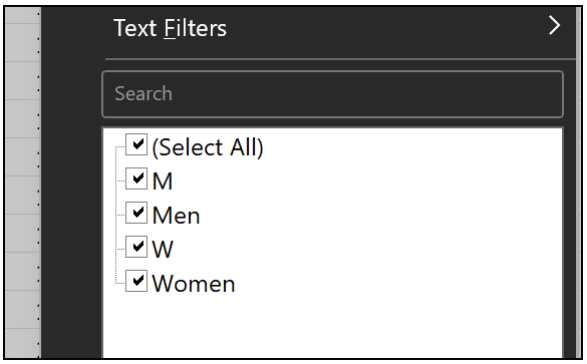
I have meticulously checked each column for missing values, ensuring a thorough review. I'm pleased to report that after examining all columns, not a single null or blank value remains in our dataset. Please refer to the attached screenshot for verification.

Counting Number Of values for Each Column		
Column_Name	**FormulaText**	**Count**
Index	=COUNT(A2:A31048)	31047
OrderID	=COUNTA(B2:B31048)	31047
CustID	=COUNTA(C2:C31048)	31047
Gender	=COUNTA(D2:D31048)	31047
Age	=COUNTA(E2:E31048)	31047
Date	=COUNTA(F2:F31048)	31047
Status	=COUNTA(G2:G31048)	31047
Channel	=COUNTA(H2:H31048)	31047
SKU	=COUNTA(I2:I31048)	31047
Category	=COUNTA(J2:J31048)	31047
Size	=COUNTA(K2:K31048)	31047
Qty	=COUNTA(L2:L31048)	31047
Currency	=COUNTA(M2:M31048)	31047
Amount	=COUNTA(N2:N31048)	31047
Ship-City	=COUNTA(O2:O31048)	31047
Ship-State	=COUNTA(P2:P31048)	31047
Ship-Postal-Code	=COUNTA(Q2:Q31048)	31047
Ship-Country	=COUNTA(R2:R31048)	31047
B2B	=COUNTA(S2:S31048)	31047

Currently, I am scrutinizing the data for any errors or misplaced values, aiming to either remove them or replace them with suitable alternatives where feasible.

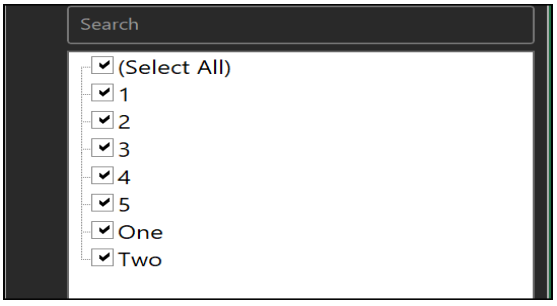
Gender:

This column contains additional types of values ('M' and 'W') that should only be 'Men' and 'Women', as shown below. Therefore, I will replace 'M' with 'Men' and 'W' with 'Women' to ensure consistency and accuracy in the data.



Quantity:

The quantity column also contains some misplaced values, such as '1', '2', '3', '4', '5', 'one', and 'two'. To ensure better consistency, I will replace 'one' with '1' and 'two' with '2'. Please refer to the attached picture for a clearer understanding.



Data Transformation:

Data transformation is the process of converting data from its original format or structure into a new format or structure that is suitable for analysis. This includes tasks like normalization, aggregation, filtering, and encoding data. The goal is to make the data more accessible, interpretable, and useful for generating insights.

Age: This column contains multiple ages, and analysing them individually would be time-consuming and result in a cluttered, hard-to-interpret analysis. To simplify and enhance the graphical representation, I will group ages into categories: below 20 as 'Teenager', below 35 as 'Young', below 50 as 'Adult', and above 50 as 'Senior'.

The formula to categorize the data based on the values in Age Column is as follows:

=IF(E2>=50,"Senior", IF(E2>=35,"Adult", IF(E2>=20,"Young","Teenager")))

Date: I will extract the month and day, storing them in two separate columns using the text function. This will provide categorical values, enabling more effective visualizations.

Ship-City: I noticed that this column contains a mix of uppercase and lowercase values, so I will convert all values to uppercase for consistency.

Now that we have completed the data preprocessing, we will proceed to the data analysis phase.

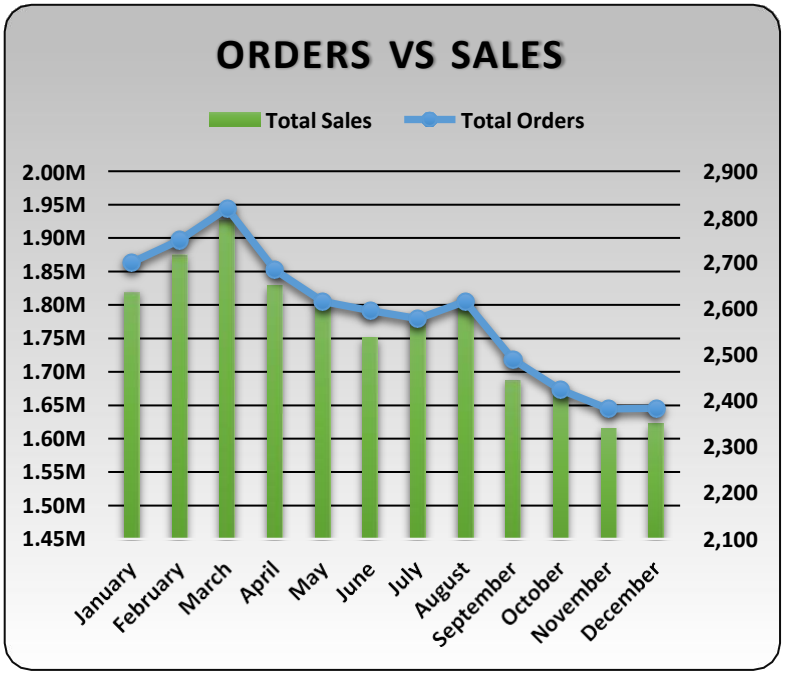
Data Analysis

In data analysis, the purpose of a pivot table is to summarize, sort, reorganize, group, count, total, or average data stored in a database. It allows us to transform data into a meaningful, easy-to-read table that helps in understanding patterns and relationships within the dataset.

Getting Insights and Analysing

Comparing the Orders and Sales:

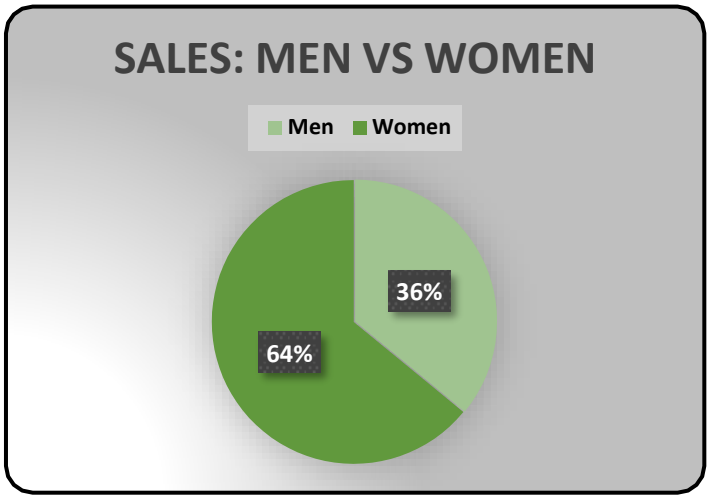
ORDERS VS SALES		
Row Labels	**Total Sales**	**Total Orders**
January	₹ 18,20,601.00	2702
February	₹ 18,75,932.00	2750
March	₹ 19,28,066.00	2819
April	₹ 18,29,263.00	2685
May	₹ 17,97,822.00	2617
June	₹ 17,50,966.00	2597
July	₹ 17,72,300.00	2579
August	₹ 18,08,505.00	2617
September	₹ 16,88,871.00	2490
October	₹ 16,66,662.00	2424
November	₹ 16,15,356.00	2383
December	₹ 16,22,033.00	2384
Grand Total	**₹ 2,11,76,377.00**	**31047**



After analysing the graphs and charts, I concluded that our sales were strong in the first quarter but then declined. From June to August, sales increased, but there was a steep decline afterward. The average monthly sales were ₹17,64,698.08. Sales for June, September, October, November, and December were notably poor, falling below the average.

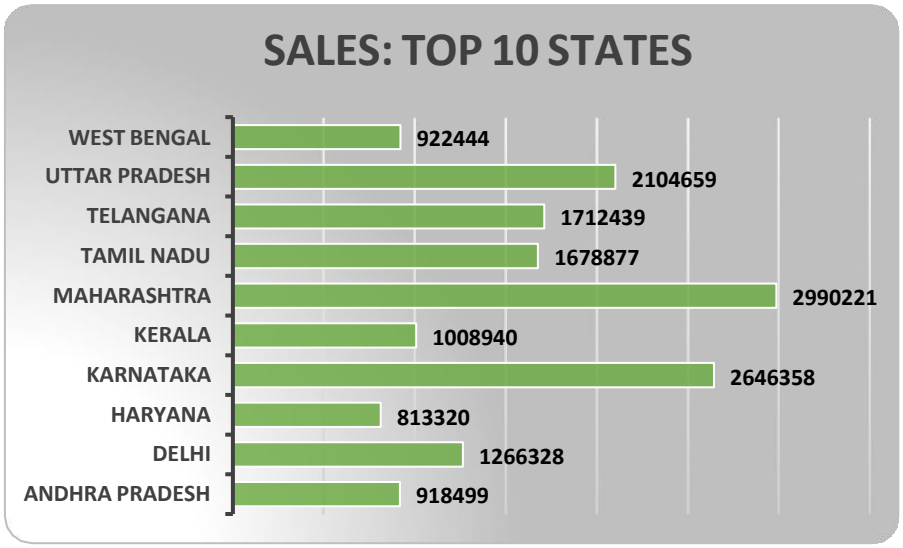
March was our best month. Comparatively, after June, we made significant progress, consistently surpassing our average monthly sales and achieving a revenue of nearly 1.8 million INR.

Revenue Through Men and Women:



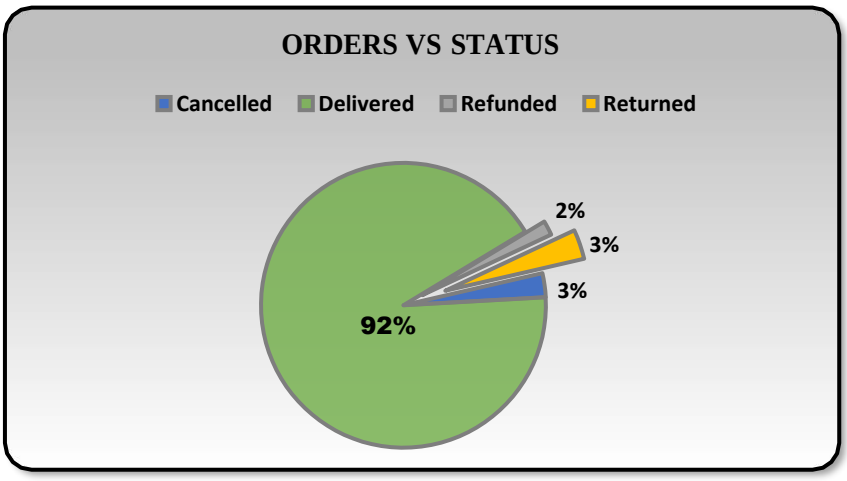
After analysing the images above, I can confidently state that women are the primary drivers of our sales, contributing to 64% of our revenue.

Top 10 states and their sales:



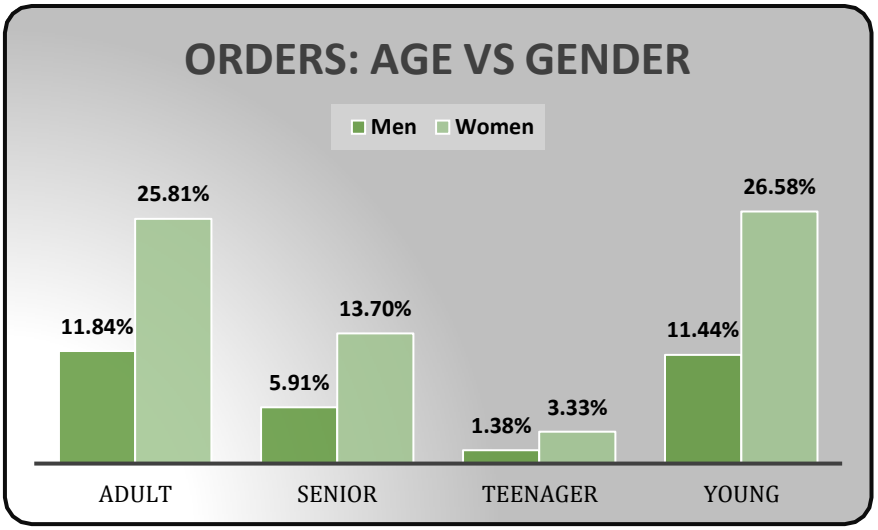
After examining the upper graph, I can confirm that the majority of our sales are coming from Uttar Pradesh, Maharashtra, and Karnataka.

Status wise Orders:



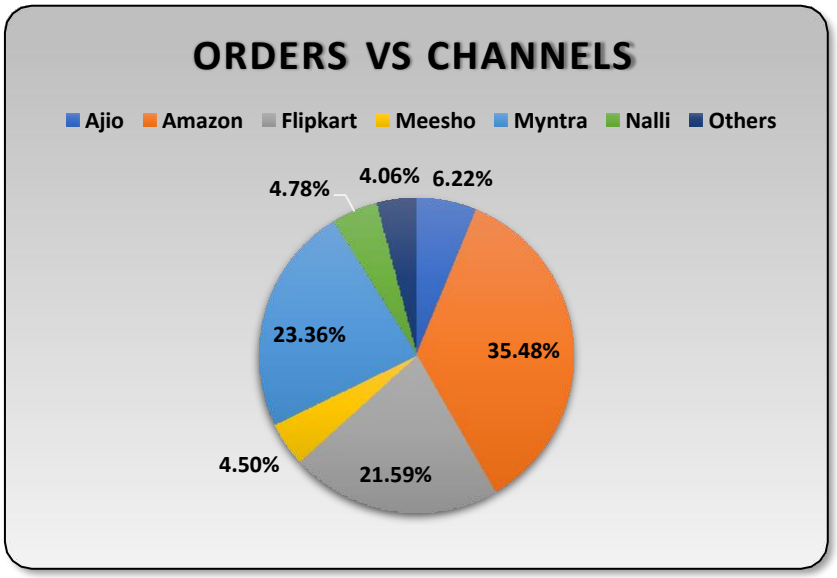
From the above graph, I can clearly see that most of our products are being successfully delivered. Only 5% of the orders are being returned and refunded, while just 3% are getting cancelled midway.

Gender wise Orders:



We can deduce that women aged between 20-50 are the primary buyers, contributing to 50% of our sales. In comparison, young and adult males each account for only about 11% of sales, which is still lower than senior citizen women, who contribute 13.7% of sales.

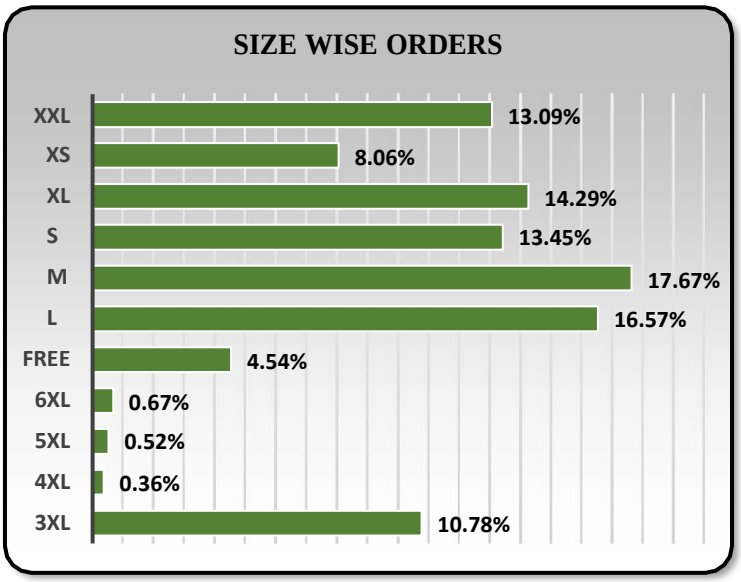
Online Platforms wise Orders:



From the graph above, it is evident that our largest supplier is Amazon, accounting for over 35% of our sales. Following Amazon, Flipkart and Myntra contribute significantly to our sales, with 21.59% and 23.36%, respectively.

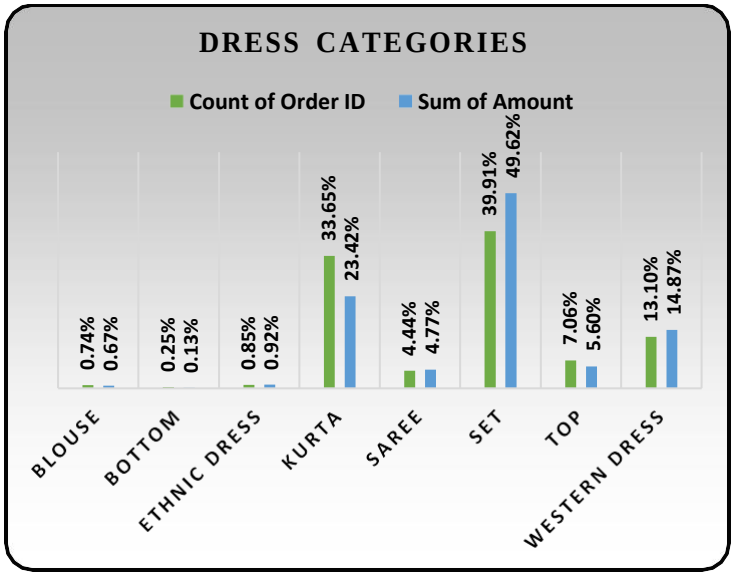
Size wise Orders:

By comparing all the dress sizes, we can identify which sizes are the most popular and should always be restocked. Conversely, we can determine which sizes are less popular and only need to be available for emergency cases. This approach ensures that customers can find their desired dress size without any issues.



From the graph, we can deduce that sizes 6XL, 5XL, and 4XL have a smaller customer base, resulting in lower sales. In contrast, sizes M, XL, and L are very popular, accounting for the majority of sales. Additionally, free-size dresses are not particularly popular, making up only 4.5% of orders. Sizes M and L each contribute to approximately 17% of orders.

Comparing Dress Categories:



From the graph above, I can confidently state that our highest sales come from SET, Kurtas, and Western dresses. In comparison, our least selling items are Bottoms, Ethnic Dresses, and Blouses, each contributing to less than 1% of our sales.

Insights Of the Data

- Women are likely to buy compared to men. (~65%)
- Maharashtra, Karnataka and Uttar Pradesh are the highest contributors in our sales. (~35%)
- Age group of 20-49 women are contributing almost 50% of the sales.
- Amazon, Flipkart and Myntra are the top 3 major suppliers of our product. (~80%)
- Kurtas and Set dresses are very popular as they are contributing majorly for our sales. (~73%)
- Most Popular dress sizes are L, M, XL, XXL and S. (~75.07%)
- 92% of our product are getting delivered while only 8% is either getting cancelled or returned.
- Up until middle of 3rd quarter of the year we had sales above average after that sales fall below margin.

Final conclusion to increase sales of store

- Target women customers of age group (20-49 yrs) Live in Maharashtra, Karnataka and Uttar Pradesh by shopping ads/coupons/offers available on Flipkart, Amazon and Myntra.
- Last year's total sales amounted to nearly ₹2.11 CR. This year, we should aim to increase sales by at least 15%, targeting a total sales figure of ₹2.42 CR.
- Target the festivals during the last four months by offering more discounts and promotions. This strategy will help boost sales as orders begin to accumulate.
- Increase the stock of Kurtas, Sets, and Western Dresses in sizes M, L, XL, XXL, and S, and offer a variety of designs to attract more female customers.

➢ https://github.com/RuchithTalanki/Dress-Shop-Sales-Analysis Link for the Report Dashboard.

The Author:

Ruchith Talanki