Sara Ray

Data Visualization Using Excel

For this assignment, I selected a dataset that shows retail sales trends over time. I used a Retail Sales dataset from Kaggle, which is a synthetic dataset created to simulate transactions in a retail environment. The dataset includes information such as transaction dates, customer demographics, product categories, quantities sold, and total sales amounts. Using this dataset, I created several pivot tables and visualizations to explore how sales vary by product type, customer characteristics, and seasonality across one year. These analyses allowed me to identify overall sales patterns, compare performance across different groups, and examine how trends change over time.

Dataset: https://www.kaggle.com/datasets/mohammadtalib786/retail-sales-dataset

Business Question

What are the key sales trends across products and customer demographics, and how do sales vary over time?

To answer this, I have broken this into different sub-questions to analyze.

1. Which product category generates the highest total sales?

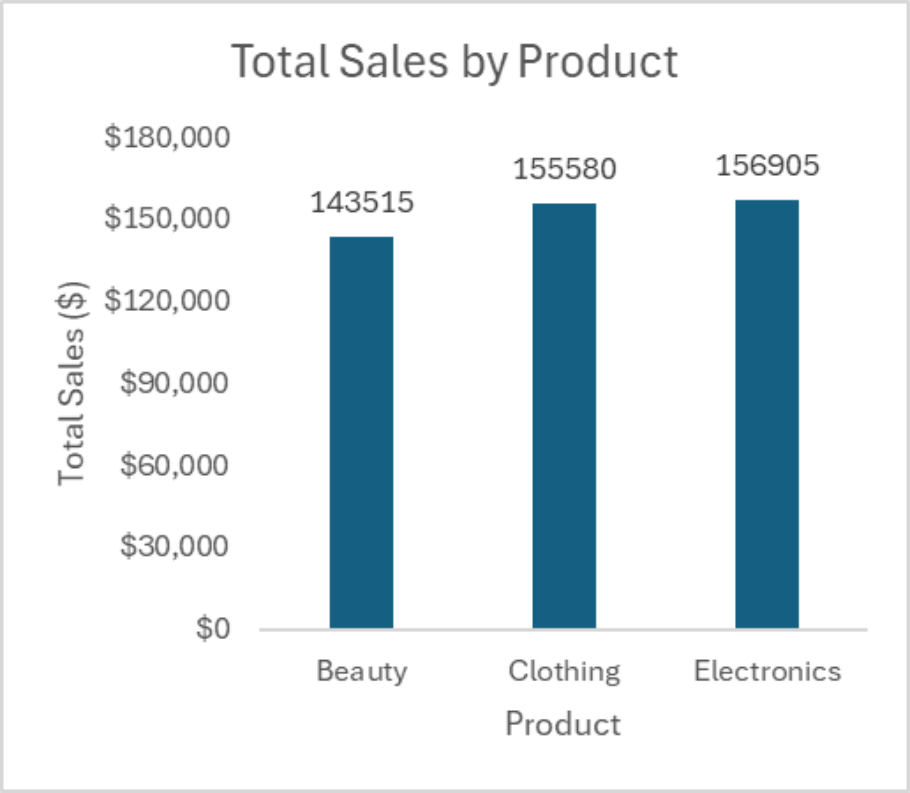

Figure 1

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Beauty	92	38230	415.5435	279523.9
Clothing	110	45610	414.6364	254091.6
Electronics	106	46735	440.8962	282695.6

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	46483.92	2	23241.96	0.085597	0.917986	3.02535
Within Groups	82815698	305	271526.9			
Total	82862182	307				

Figure 2

For this question, I created a pivot table and a column chart to compare total sales by product. The chart shows that Electronics had the highest total sales, while Clothing and Beauty had similar sales totals. To further analyze the data, I conducted a one-way ANOVA to determine whether there was a significant difference in mean sales across product categories. The ANOVA results produced a p-value of 0.92, which is greater than the significance level of 0.05, indicating that there is no statistically significant difference in sales among the three products. This suggests that although total sales differ across categories, the average transaction size is similar for all three products.

2. How do purchasing patterns differ by gender across product categories?

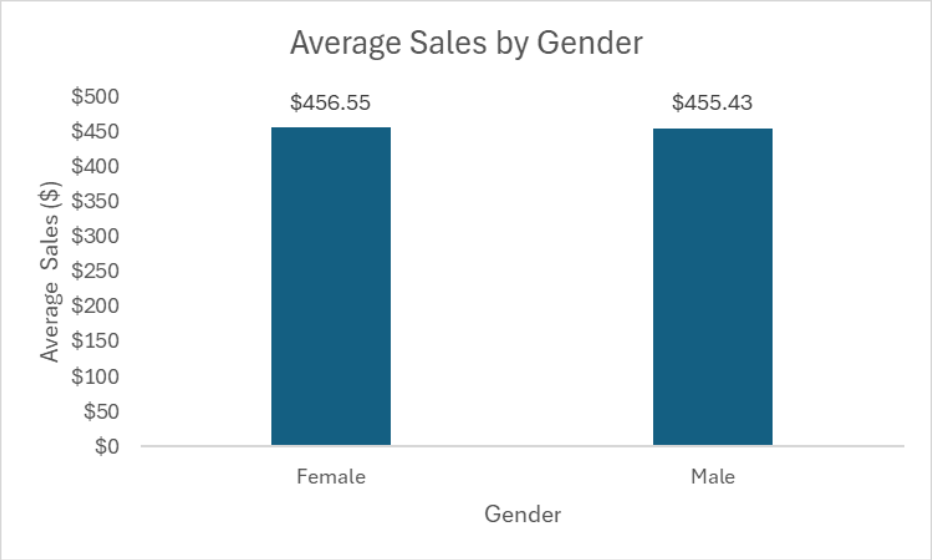

Figure 3

t-Test: Two-Sample Assuming Unequal Variances		
	Male	*Female*
Mean	455.4286	456.5490196
Variance	320847.3	307247.7923
Observations	490	510
Hypothesized Mean Difference	0	
df	994	
t Stat	-0.0316	
P(T<=t) one-tail	0.487399	
t Critical one-tail	1.646388	
P(T<=t) two-tail	0.974798	
t Critical two-tail	1.962353	

Figure 4

The chart shows the average sales per transaction by gender. The average sales for female customers ($456.55) and male customers ($455.43) are nearly identical. This is consistent with the two-sample t-test results, which indicate that there is no statistically significant difference in mean sales between male and female customers (p = 0.975).

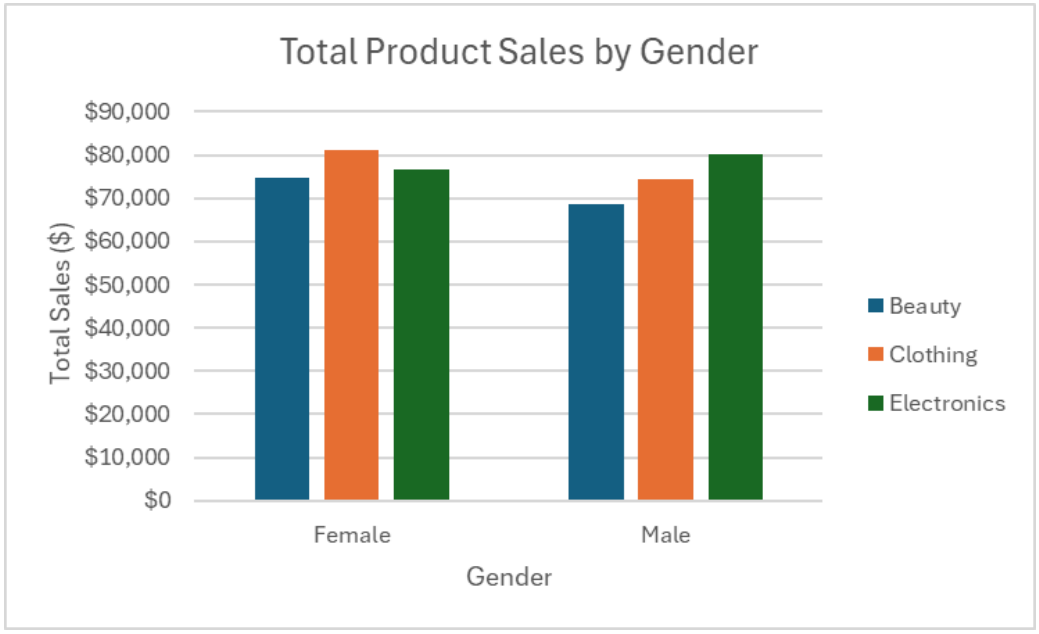

Figure 5

This chart shows total product sales by gender across the three product categories. Female customers generated higher total sales in Beauty and Clothing, while male customers generated higher total sales in Electronics. However, consistent with the t-test results, these differences reflect differences in the number of transactions rather than differences in average spending per transaction.

3. How do purchasing patterns differ by age across product sales?

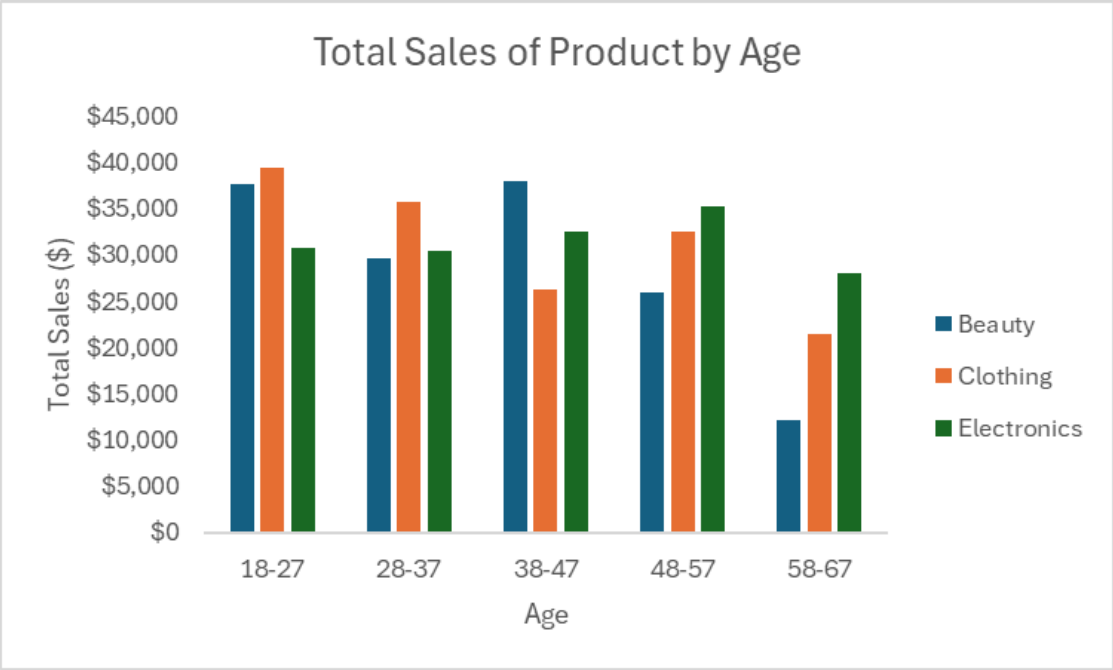

Figure 6

The chart shows that total sales vary by age group and product category. Clothing has the highest total sales among the 18–27 and 28–37 age groups, while electronics are higher in the 48–57 and 58–67 age groups. Beauty sales peak in the 38–47 group but decline in the oldest age group. This suggests that product preferences shift with age, with younger and middle-aged customers favoring clothing and beauty products, while older customers purchase more electronics.

4. How do monthly and quarterly sales change over time?

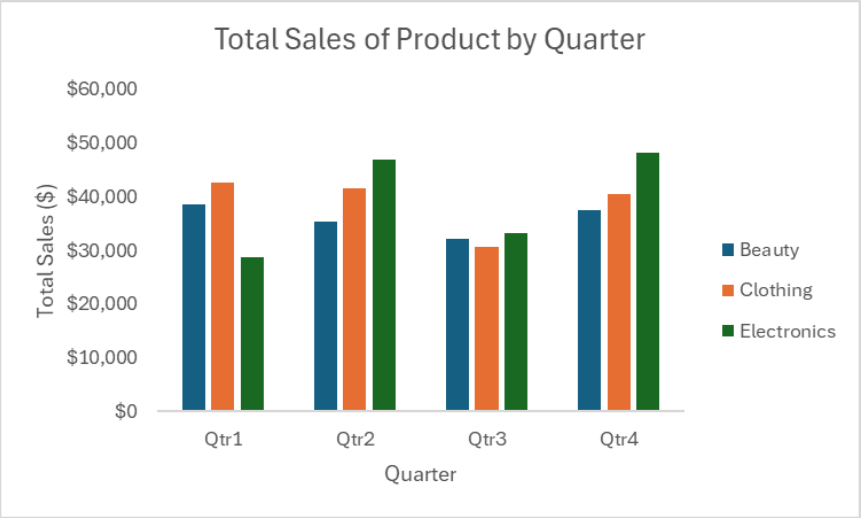

Figure 7

Total sales vary across quarters and product categories. Clothing generates the highest total sales in the first quarter, while electronics have the largest sales in both the second and fourth quarters. All three product categories declined in the third quarter, indicating a seasonal dip in overall sales before increasing in the fourth quarter.

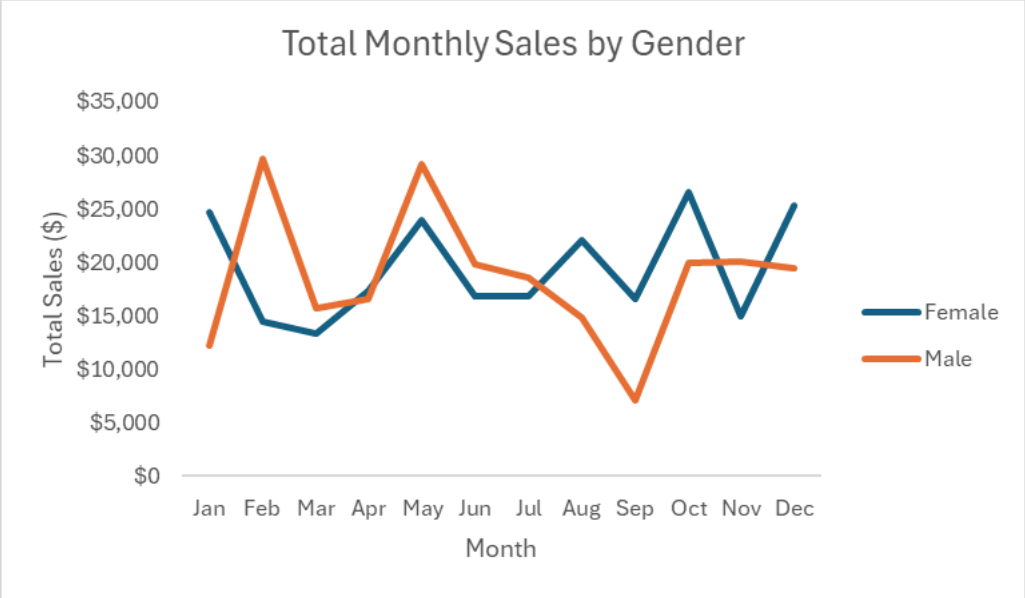

Figure 8

The chart shows that total monthly sales for males and females fluctuate throughout the year. Male sales peak in February and May, which could indicate spending for holidays such as Valentine's Day and Mother’s Day. Female sales peak in October and December. Overall, both genders show similar patterns with alternating months of higher total sales throughout the year.

Descriptive Statistics of Total Sales

	Total Sales
Mean	456
Standard Error	17.70868
Median	135
Mode	50
Standard Deviation	559.9976316
Sample Variance	313597.3473
Kurtosis	0.815031355
Skewness	1.376128244
Range	1975
Minimum	25
Maximum	2000
Sum	456000
Count	1000

Figure 9

Here is a descriptive statistics table of Total Sales. This table helps get an overall picture of the dataset and understand how sales values are spread across all transactions. The mean total sale is 456, while the median is much lower at 135, which suggests that a small number of large transactions are increasing the average. The minimum and maximum values range from 25 to 2000, showing a wide range in sales amounts. The large standard deviation also indicates that there is a lot of variability in transaction values. Overall, these statistics help identify extreme values in the data and provide a useful baseline for comparing different sales patterns later in the analysis.

Summary and Recommendations

Sales patterns differ more by customer demographics and seasonality than by differences in average spending. Younger customers contribute more to clothing sales, while older customers contribute more to electronics sales. Seasonality is strong, with the third quarter representing a consistent low point and the fourth quarter showing a clear recovery. Differences in total sales by gender are mainly driven by the number of transactions rather than by how much customers spend on average.

Different marketing strategies can be used by gearing promotions to specific age groups. For example, clothing promotions should focus on younger customers ages 18–37, electronics marketing should target older customers ages 48–67, and beauty campaigns should be aimed at middle-aged customers ages 38–47. Seasonal patterns should also guide planning, with increased inventory and marketing for electronics in the second and fourth quarters, careful inventory control during the third-quarter slowdown, and promotions designed to help increase sales during slower periods.

Overall, the results indicate that total sales are mainly influenced by customer demographics and seasonal trends rather than by significant differences in how much customers spend on average. This suggests that future revenue growth should focus on targeted marketing and increasing the number of transactions across key customer groups.