RETAIL SALES DATA ANALYSIS REPORT

Prepared by: Bola Abimbola

Oasis Infobyt: Level 1 Task 1

Project Title: Exploratory Data Analysis on Retail Sales Data

Tool Used: Power BI

Date: April 2025

Table of Contents

1. Introduction

This project involves conducting an Exploratory Data Analysis (EDA) on retail sales data using Power BI. The objective is to uncover insights into customer purchasing behavior, product performance, demographic patterns, and seasonal trends, with the aim of driving informed business decisions.

2. Objectives

- Understand the sales performance across different product categories.
- Analyze purchasing behavior by gender and age groups.
- Identify monthly and quarterly sales trends.
- Determine key descriptive statistics such as average, median, mode, and standard deviation.
- Provide actionable insights for retail strategy improvements.

3. Dataset Overview

The dataset contains the following columns:
Transaction ID, Date, Customer ID, Gender, Age, Product Category (Clothing, Beauty, Electronics), Quantity Sold, Price per Unit, Total Amoun.

Records are spread across the entire year 2023 and include both male and female customers from various age groups.

Transaction ID	Date	Customer ID	Gender	Age	Product Category	Quantity	Price per Unit	Total Amount	Month Number	Mor
7	Monday, 13 March 2023	CUST007	Male	46	Clothing	2	25	50	3	M
17	Saturday, 22 April 2023	CUST017	Female	27	Clothing	4	25	100	4	A
19	Saturday, 16 September 2023	CUST019	Female	62	Clothing	2	25	50	9	S
41	Wednesday, 22 February 2023	CUST041	Male	34	Clothing	2	25	50	2	F
44	Sunday, 19 February 2023	CUST044	Female	22	Clothing	1	25	25	2	F
64	Tuesday, 24 January 2023	CUST064	Male	49	Clothing	4	25	100	1	Ja
103	Tuesday, 17 January 2023	CUST103	Female	59	Clothing	1	25	25	1	Ja
127	Monday, 24 July 2023	CUST127	Female	33	Clothing	2	25	50	7	Ju
135	Sunday, 26 February 2023	CUST135	Male	20	Clothing	2	25	50	2	F
145	Thursday, 2 November 2023	CUST145	Female	39	Clothing	3	25	75	11	N
149	Wednesday, 11 October 2023	CUST149	Male	22	Clothing	3	25	75	10	O
156	Saturday, 25 November 2023	CUST156	Female	43	Clothing	4	25	100	11	N
170	Friday, 2 June 2023	CUST170	Female	25	Clothing	2	25	50	6	Ju
185	Monday, 27 February 2023	CUST185	Male	24	Clothing	1	25	25	2	F
188	Wednesday, 3 May 2023	CUST188	Male	40	Clothing	3	25	75	5	M
205	Tuesday, 7 November 2023	CUST205	Female	43	Clothing	1	25	25	11	N
206	Saturday, 5 August 2023	CUST206	Male	61	Clothing	1	25	25	8	A
223	Thursday, 2 February 2023	CUST223	Female	64	Clothing	1	25	25	2	F
236	Friday, 28 April 2023	CUST236	Female	54	Clothing	1	25	25	4	A
242	Tuesday, 2 May 2023	CUST242	Male	21	Clothing	1	25	25	5	M
261	Saturday, 5 August 2023	CUST261	Male	21	Clothing	2	25	50	8	A
277	Friday, 18 August 2023	CUST277	Male	36	Clothing	4	25	100	8	A

Brief Overview of Table

4. Methodology

- **Data Cleaning:** Performed in Power BI to ensure consistency and correctness.
- **Data Modeling:** Calculated columns and measures were created using DAX

  Age Group classification

```
Age Group = SWITCH(TRUE(),
[Age] >= 50, "Old (50-64)",
[Age] >= 30, "Middle (30-49)",
"Young (18-29)")
```

Month

```
Month = FORMAT([Date], "MMM")
```

- **Descriptive Analysis:** Measures such as mean, median, mode, and standard deviation were computed.

Mean

```
Average Age = ROUND(AVERAGE('retail_sales_dataset'[Age]),0)
```

```
Average Total Amount = AVERAGE('retail_sales_dataset'[Total Amount])
```

Median

```
Median Age = ROUND(MEDIAN('retail_sales_dataset'[Age]),0)
```

Mode

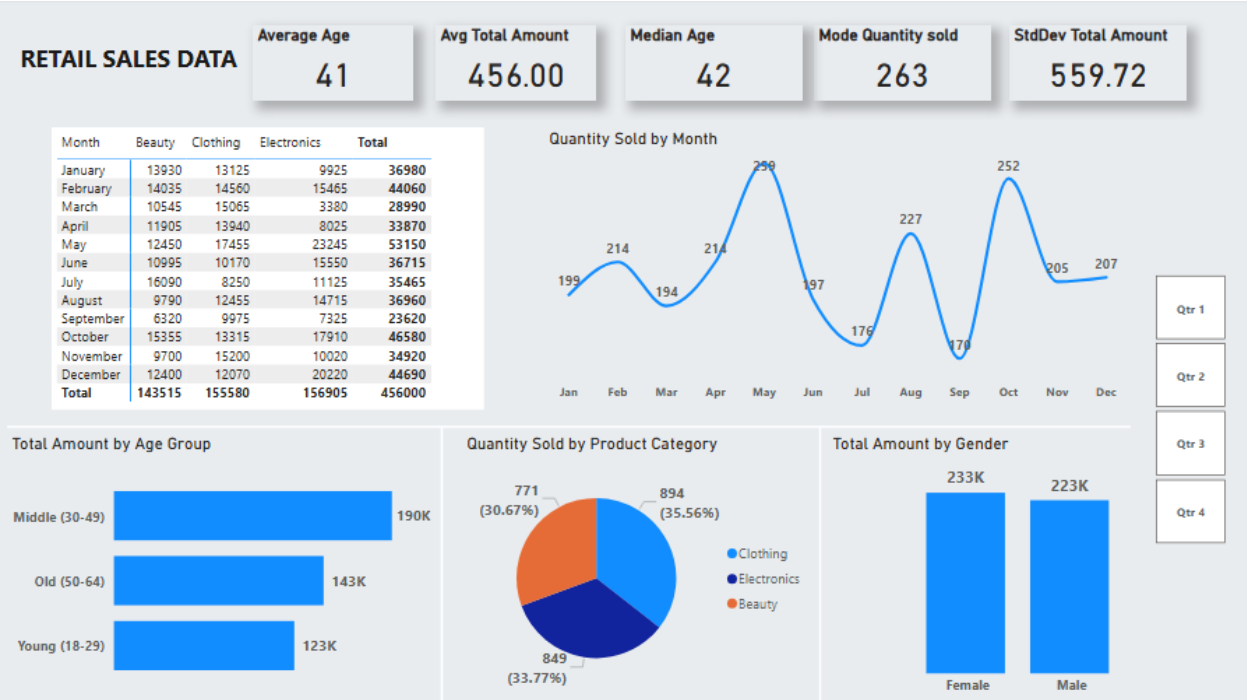
```
Mode Quantity =
CALCULATE(
    MAXX(
        VALUES('retail_sales_dataset'[Quantity]),
        CALCULATE(COUNTROWS('retail_sales_dataset'))
    )
)

```

Standard deviation

```
StdDev_TotalAmount = STDEV.P('retail_sales_dataset'[Total Amount])
```

- **Visualization:** Charts, cards, and tables were created to highlight insights in an interactive dashboard.



Dashboard Overview

Link to Dashboard [Microsoft Power BI]

5. Key Metrics and Insights

- **Average Age:** 41 years
- **Median Age:** 42 years
- **Average Total Amount:** ₦456.00
- **Mode Quantity Sold:** 263
- **Standard Deviation of Total Amount:** 559.72

6. Visualizations & Interpretation

- **Quantity Sold by Month:** Peak sales occurred in June (259 units), followed by October and August.

- **Total Amount by Age Group:**
    - Middle-aged customers (30-49) generated the highest revenue (₦190K).
    - Older adults (50-64): ₦143K
    - Young adults (18-29): ₦123K
- **Quantity Sold by Product Category:**
    - Clothing (35.56%) had the highest quantity sold.
    - Electronics and Beauty followed closely.
- **Total Amount by Gender:**
    - Female customers slightly outspent male customers (₦233K vs ₦223K).
- **Quarterly Filter:** Interactive slicers enable breakdown of data by Q1 to Q4.

7. Recommendations

- Focus marketing efforts on middle-aged customers as they contribute the most revenue.
- Increase promotions during Q2 and Q4, which show high sales performance.
- Expand clothing inventory or promotions, given its popularity.
- Leverage gender-based preferences to create targeted campaigns.

8. Conclusion

The retail sales EDA revealed valuable insights into customer demographics, product performance, and sales trends. With data-driven strategies, the

business can enhance decision-making, optimize inventory, and tailor marketing to maximize profitability.