

EXPLORATORY DATA ANALYSIS OF RETAIL SALES DATA

Excel-Based Data Analysis Project

Nurul Fatihah Binti Rizwan

PROJECT OVERVIEW

Purpose of the project

- To explore customer purchasing behavior using sales transaction data
- To identify patterns in spending across demographics and product categories

Tools Used:

- Microsoft Excel (Data Cleaning, Pivot Tables, Charts)

DATASET DESCRIPTION

- Customer demographics (Age, Gender)
- Product categories
- Purchase amount (RM)
- Transaction dates

Original Records: ~200+ transactions

DATA CLEANING PROCESS

Steps performed:

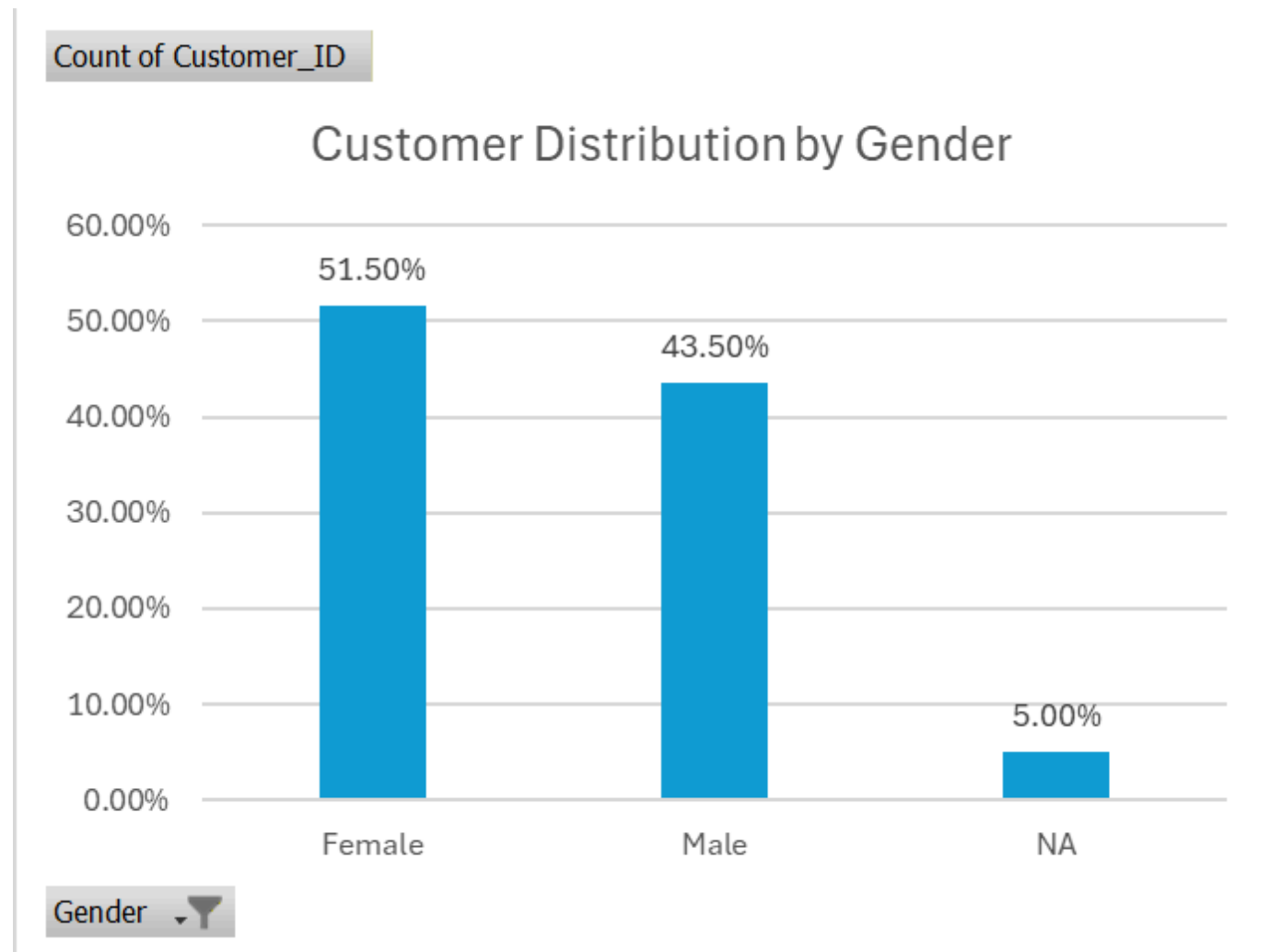
1. Identified and handled missing values (median for numerical, NA for categorical)
2. Removed duplicate records
3. Created derived features (Age Groups, Purchase Bands)

DESCRIPTIVE STATISTICS

- The dataset contains 200 complete transaction records, which is sufficient for identifying customer and sales patterns.
- The average customer age is approximately 29 years, with a median age of 28, indicating a relatively young customer base. Ages range from 18 to 45.
- The average purchase amount is RM137, while the median is RM100, suggesting that although most purchases are lower to mid-range, there are some higher-value transactions up to RM500.

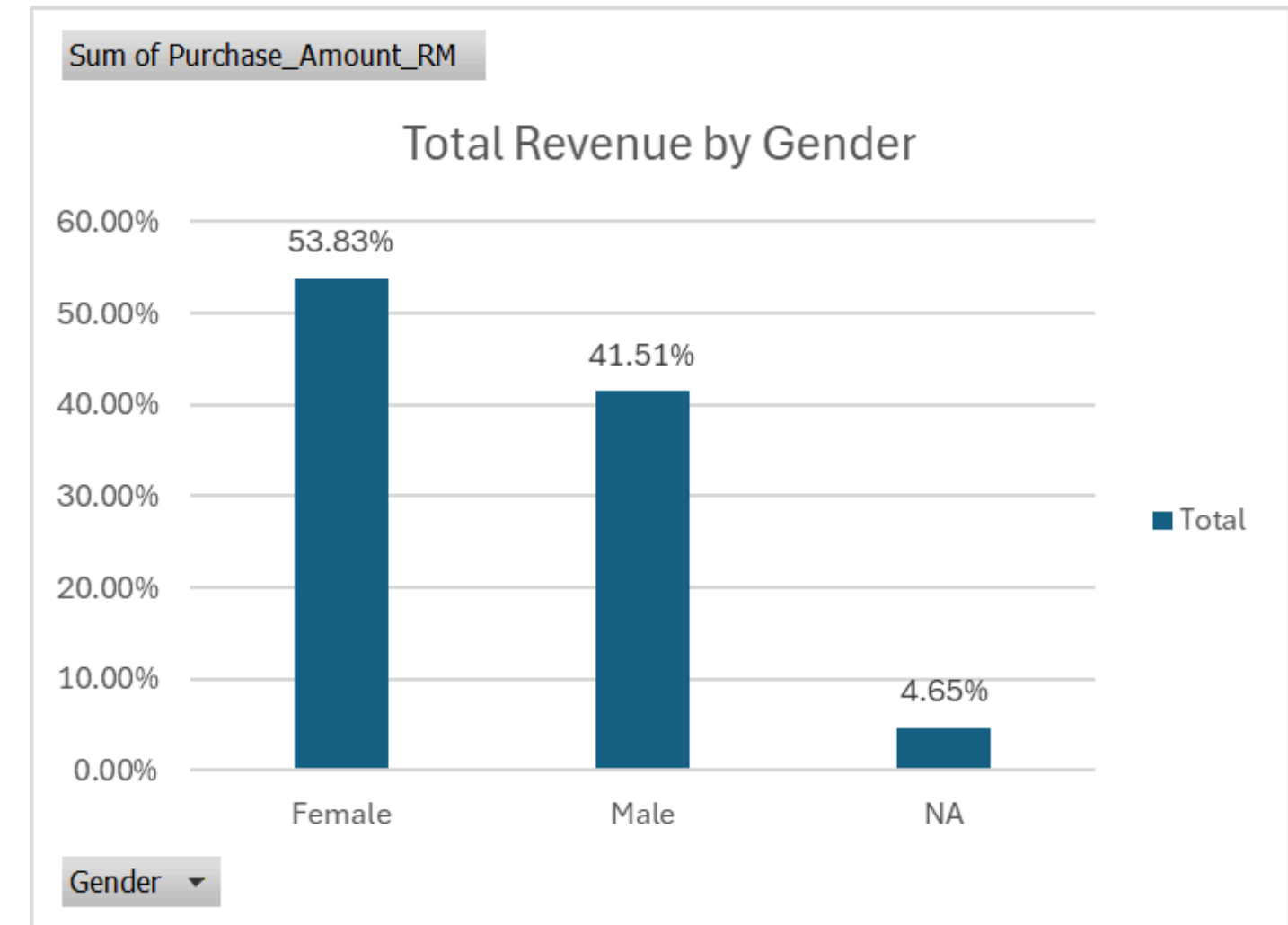
Statistic	Age	Purchase_Amount_RM
Count	200	200
Mean	29.195	137
Median	28	100
Min	18	50
Max	45	500

GENDER ANALYSIS



Insight:

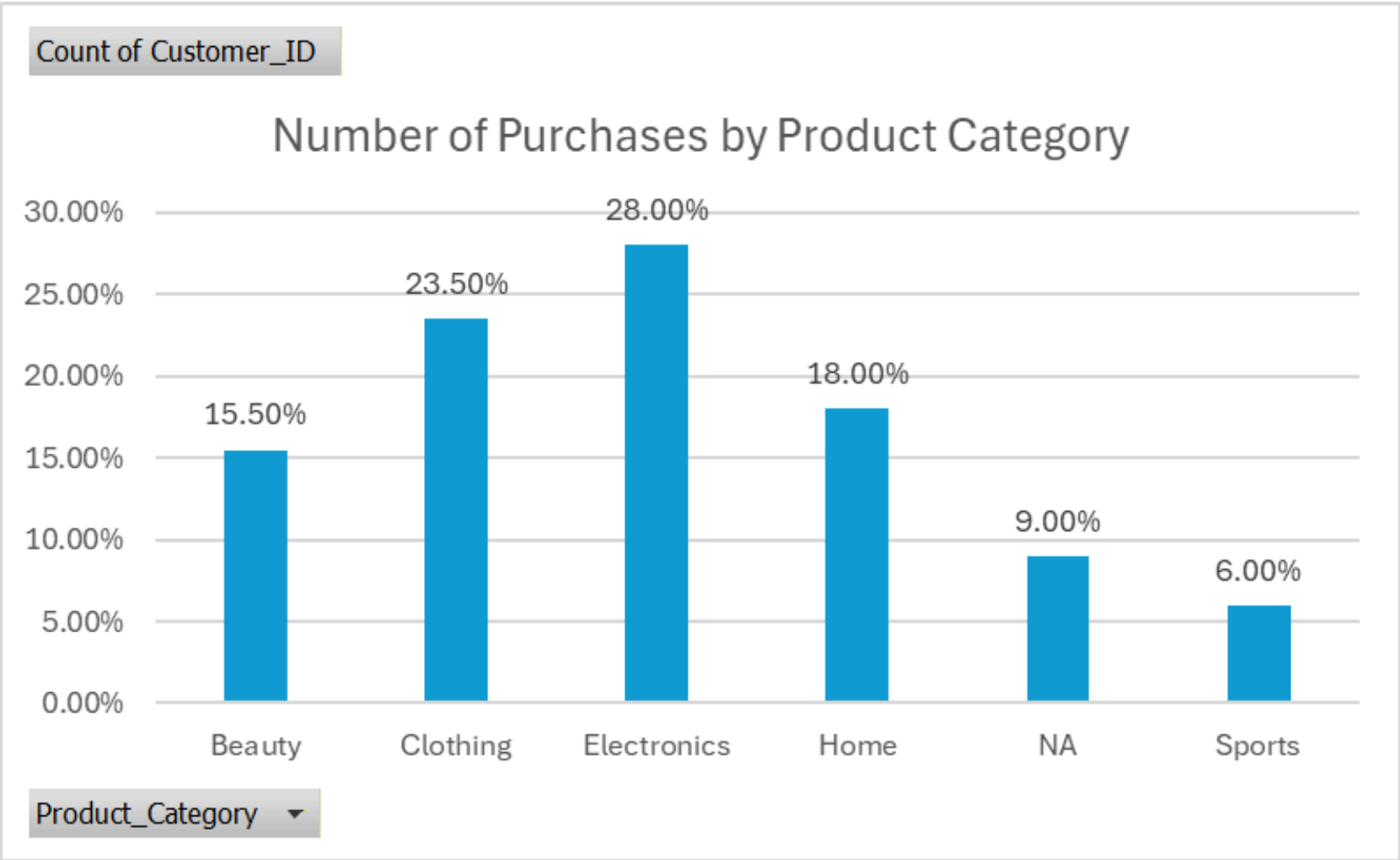
Female customers represent the largest proportion of the dataset, while a small percentage of transactions have unspecified gender information.



Insight:

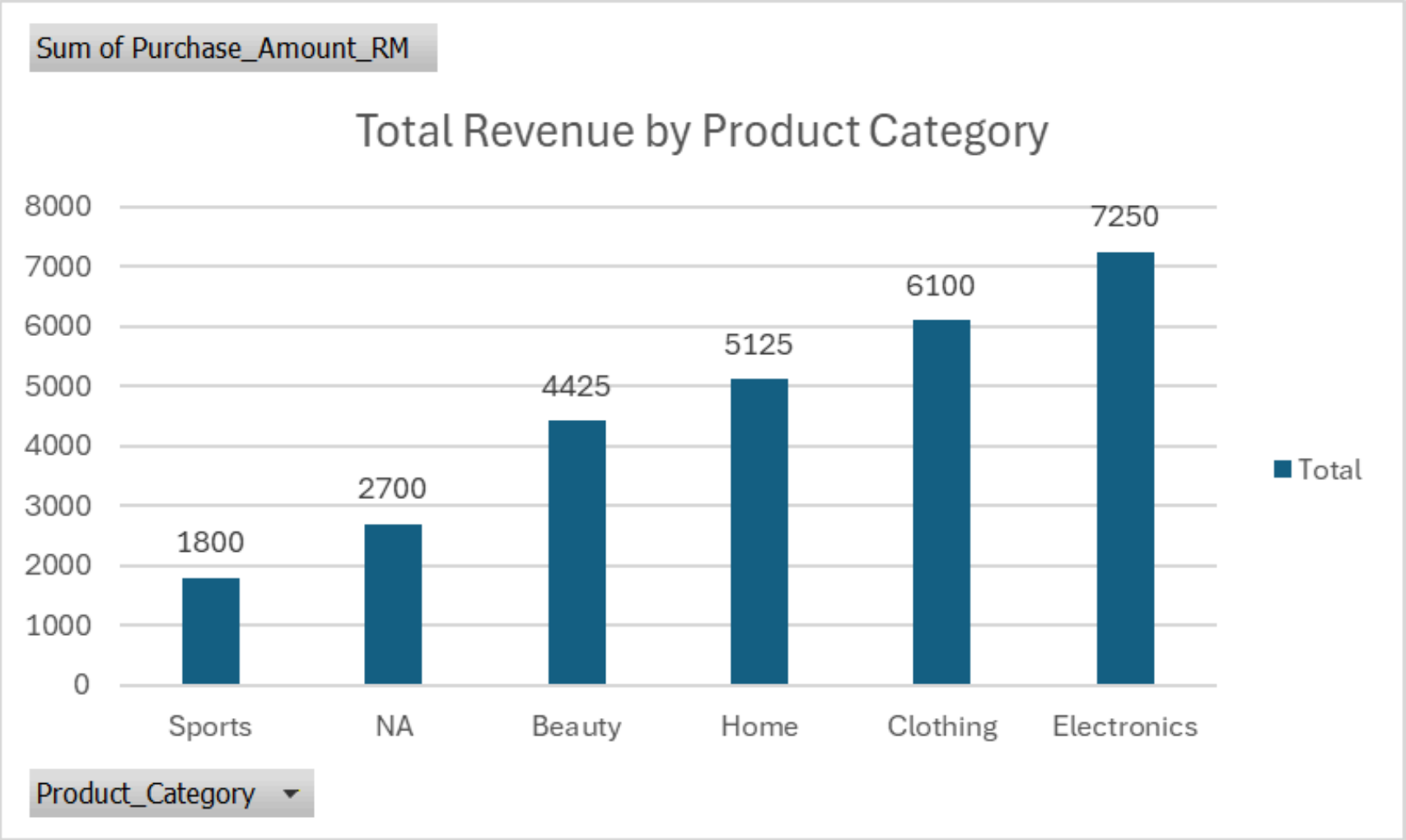
Female customers represent the largest proportion of the dataset, while a small percentage of transactions have unspecified gender information.

PRODUCT CATEGORY ANALYSIS



Insight:

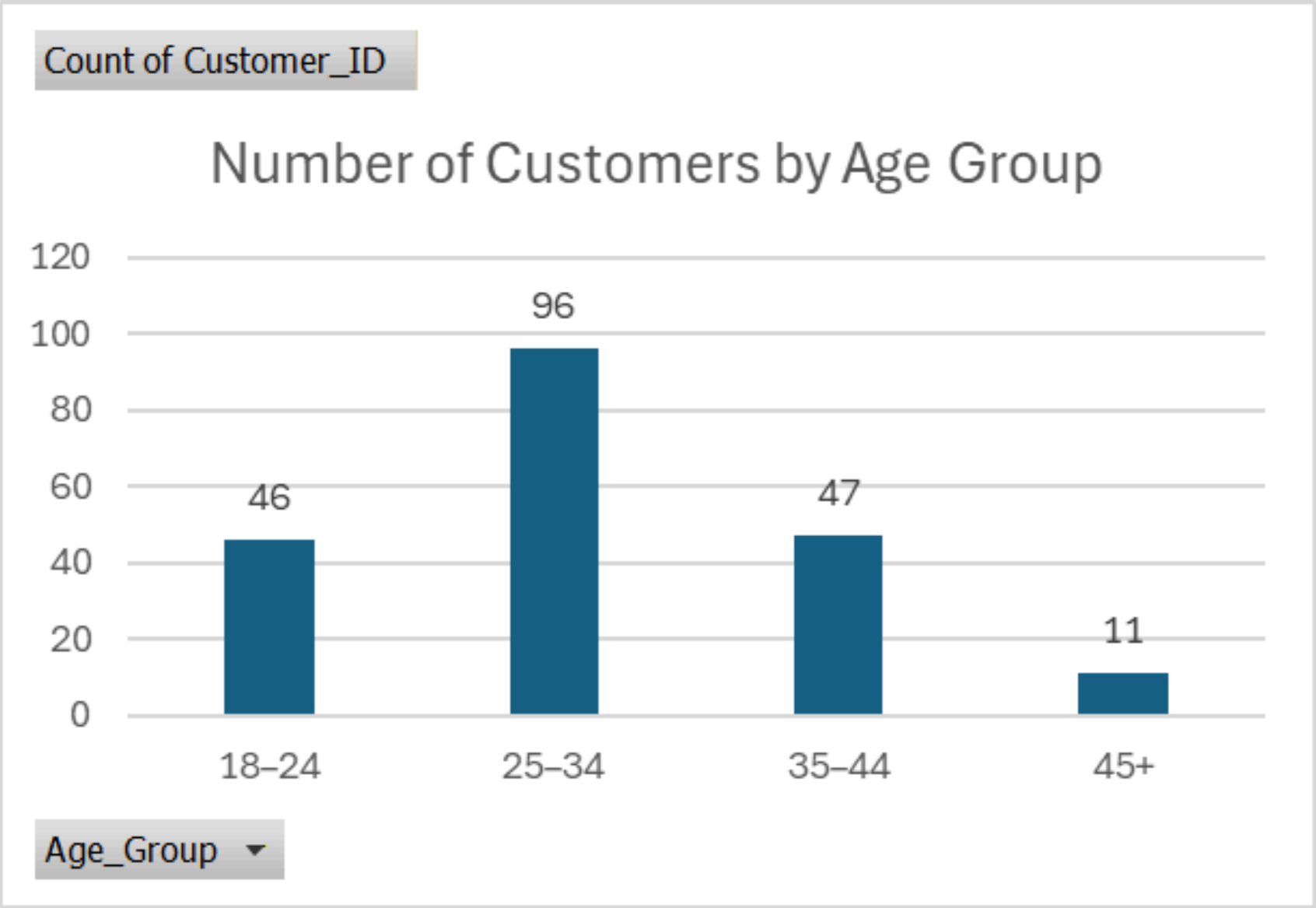
Electronics is the most frequently purchased product category, indicating strong customer demand in this segment.



Insight:

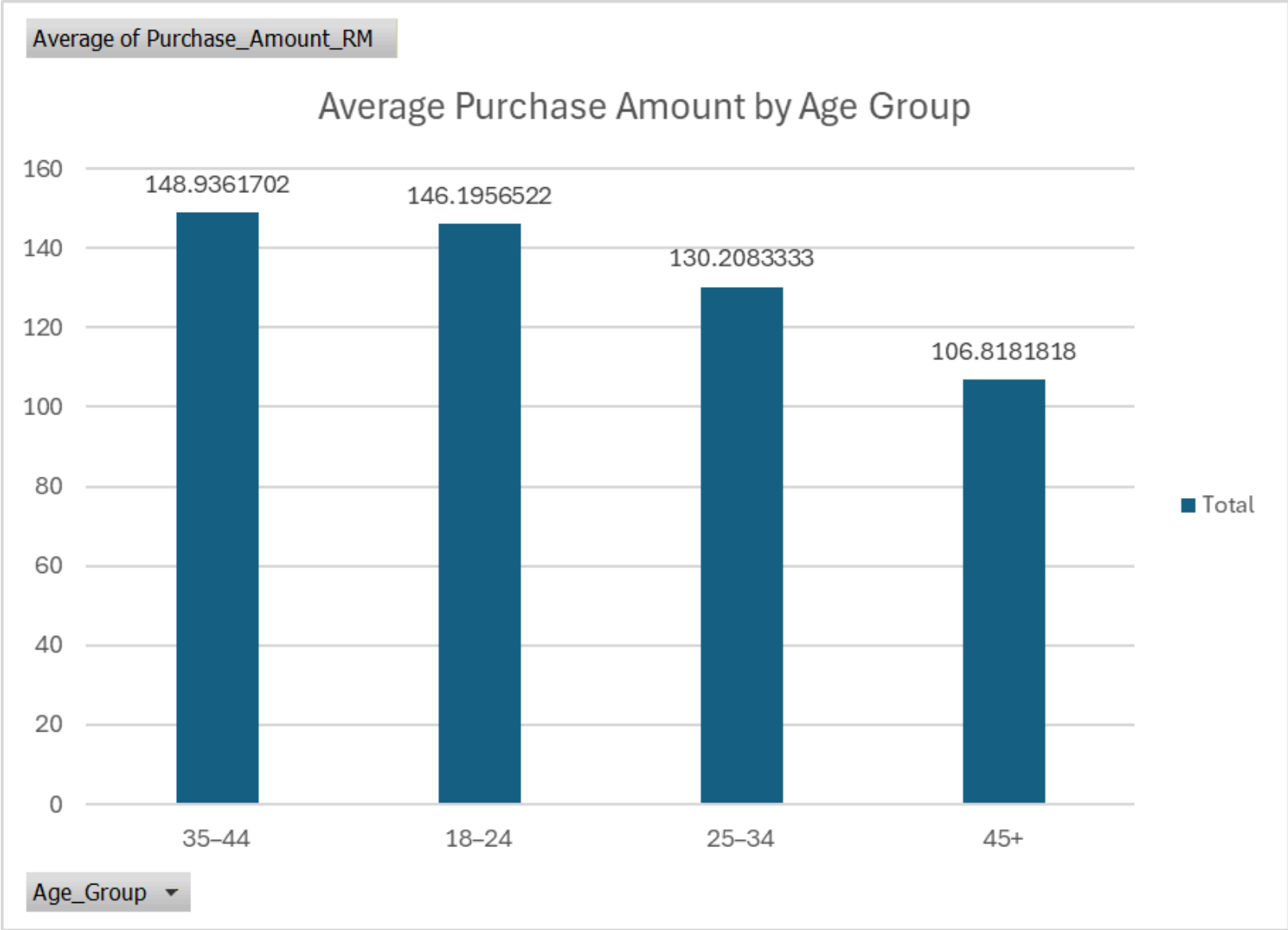
Electronics contributes the highest share of total revenue, making it the primary revenue driver compared to other product categories.

AGE GROUP ANALYSIS



Insight:

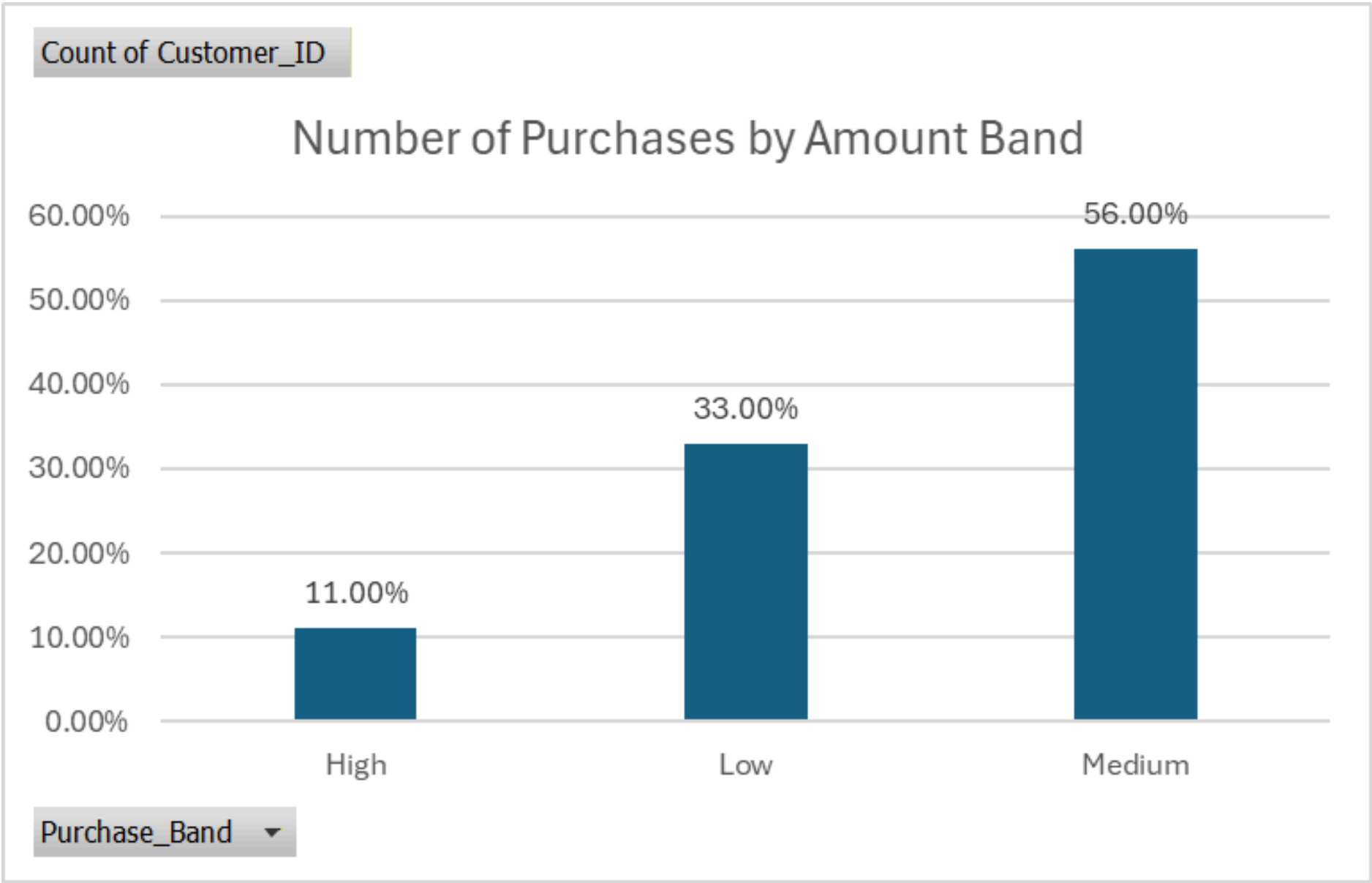
Customers aged 25–34 make up the largest customer segment, this group represents a key target for customer acquisition and engagement strategies.



Insight:

Customers aged 35–44 have the highest average purchase amount, indicating stronger spending power within this age group.

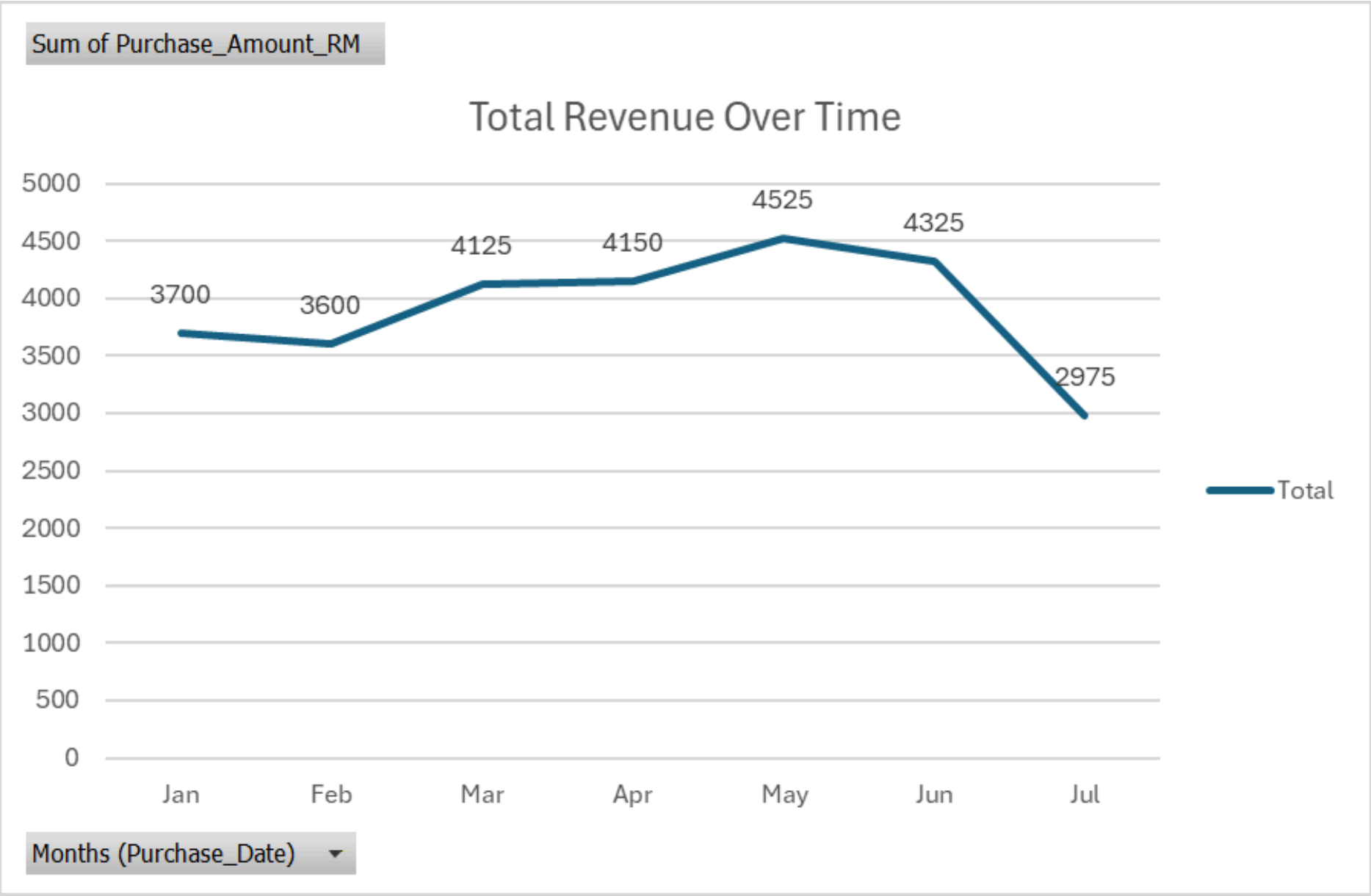
PURCHASE BAND ANALYSIS



Insight:

Medium-value purchases account for 56% of total transactions, indicating a strong customer preference for mid-range priced products.

MONTHLY SALES TREND



Insight:

Monthly sales show noticeable fluctuations, suggesting the presence of seasonal purchasing patterns rather than consistent sales throughout the year.

KEY BUSINESS INSIGHTS

Which product category generates the highest revenue?

Electronics generates the highest total revenue, making it the primary revenue-driving product category.

Do older customers spend more than younger ones?

Yes. While customers aged 25–34 represent the largest customer group, customers aged 35–44 have the highest average purchase amount, indicating stronger spending power.

How does purchase amount differ by gender?

Purchase behavior varies by gender, with female customers contributing the largest share of total revenue.

Are there seasonal trends in sales?

Monthly sales fluctuate over time, suggesting the presence of seasonal purchasing patterns rather than consistent sales across all months.

CONCLUSION:

Overall, the analysis highlights clear customer segments, key revenue drivers, and spending trends that can support targeted marketing and pricing strategies.