

Customer Transaction Analysis - Walmart

1. Executive Summary :

Product-Level Analysis :

City C emerged as the **top-performing market** at a product level

- 1.It recorded the **highest purchase frequency** and strong **average purchase amounts**, indicating high customer demand and purchasing power.
- 2.Products in **City C** also display characteristics of both **premium (high average spend + low frequency)** and **mass-retail (high average spend + high frequency)**, making it strategically important for Walmart's product placement and promotions.

Customer Segmentation by Demographics :

However, **FM segmentation revealed a key weakness in City C:**

- 1.**Risk segment overrepresentation at 130%**, along with **Need Attention at 119%**, signals **engagement and retention issues**. Despite high spend, customers in City C are less loyal and more at risk of churn compared to other cities.
- 2.**City B**, in contrast, showed **strong Best customers (170%)** with lower risk → a more stable premium base.
- 3.**Age 26–35** and **Male customers** were most aligned with the Best and Loyal segments.
- 4.**Older (55+) and Female customers** were significantly overrepresented in Risk and Need Attention, requiring targeted engagement strategies.

Key Takeaway: The City C

At a product level, City C looks like the strongest market.

At a customer segment level, City C shows high Risk and low Best representation.

This contradiction suggests that while customers in City C are buying frequently and spending heavily, they may not be emotionally loyal to Walmart → possibly price-sensitive, convenience-driven, or lacking targeted retention efforts.

Final Suggetions

Protect & Grow Premium Segments

Double down on **City B, Age 26–35, Males** → strongest Best customers, most stable loyalty base.

Invest in **premium and high-frequency products** here to drive long-term value.

Fix the City C Engagement Gap

Launch **retention and loyalty programs** in City C to convert high spenders into long-term Best customers.

Deep-dive into customer behavior (e.g., promotions, competition, service issues) to understand **why high spend is not translating into loyalty**.

Address High-Risk Groups

Target **55+ and Female customers** with personalized offers, better service touchpoints, and communication to reduce churn risk.

Consider education/onboarding for **0–17** to nurture them into future Best customers.

Product Strategy Alignment

Leverage **premium products (high avg. spend, low frequency)** for niche, high-margin segments.

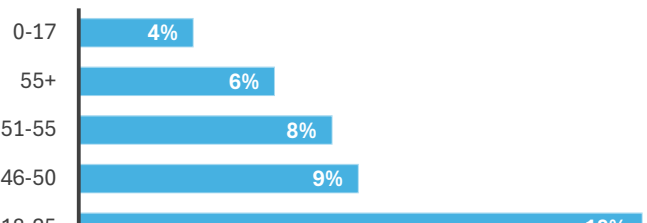
Push **popular high-frequency + high-avg products** in markets like City B to strengthen stable revenue streams.

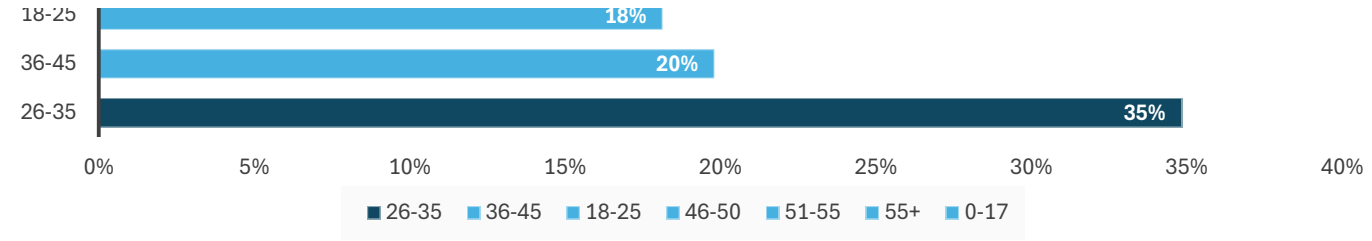
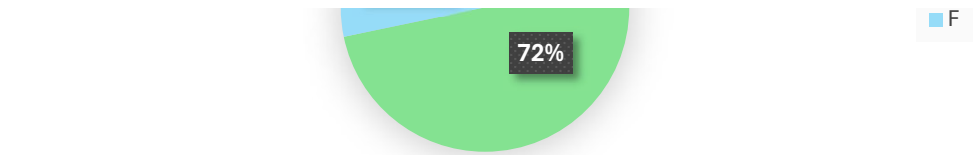
2. Demographic Profile of Customers

Gender Distribution (%Total Customers)

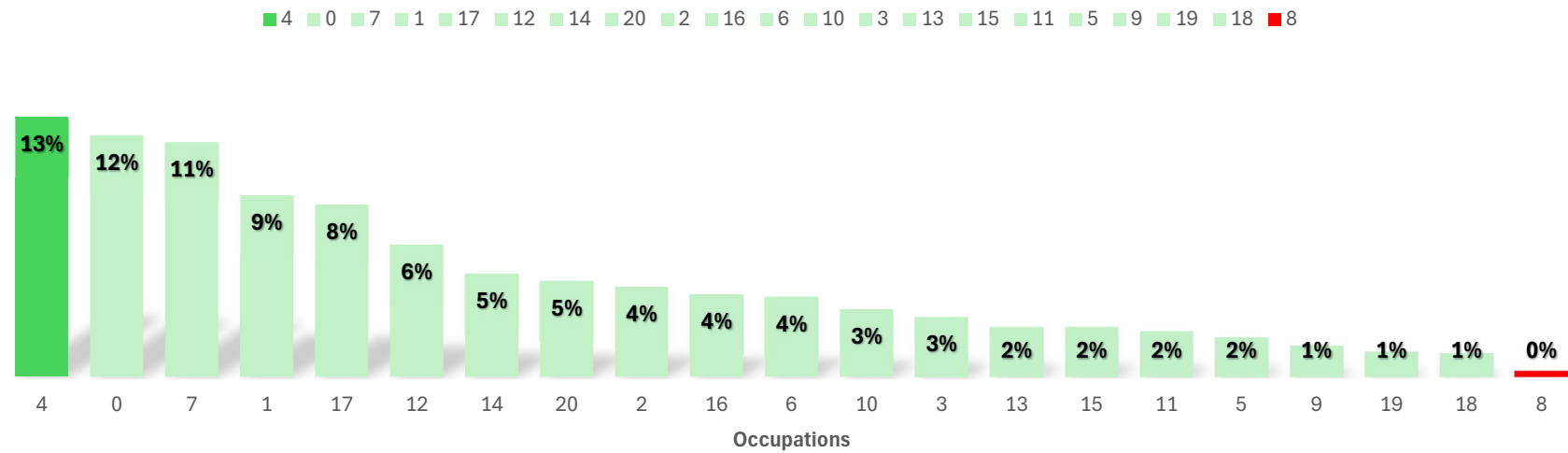


Age Group Distribution (% Total Customers)

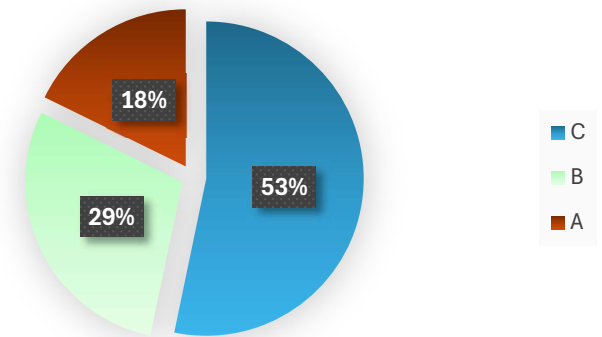




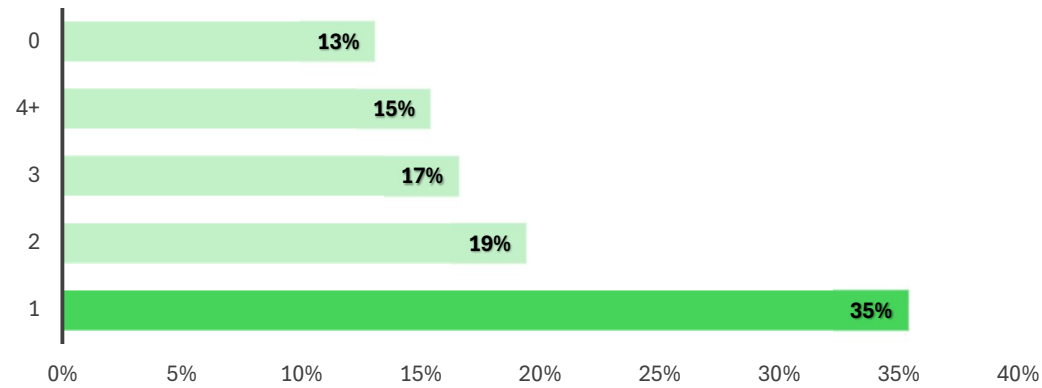
Distribution of Occupations (% Total Customers)



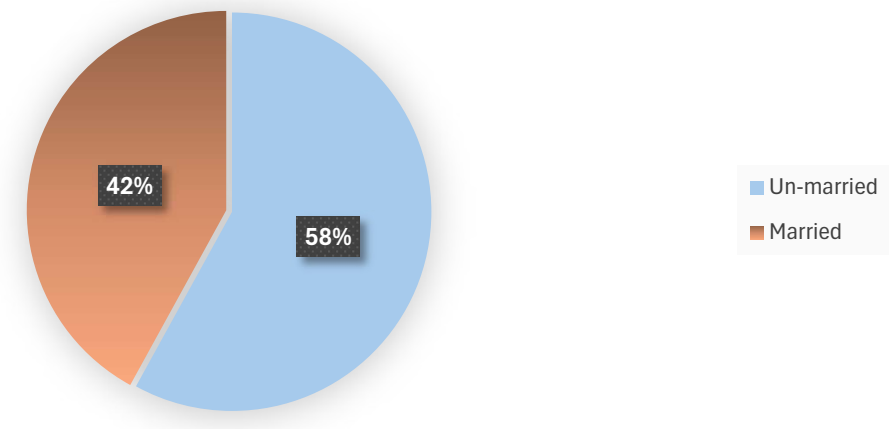
City Category Distribution (% Total Customers)



Duration of Customers staying in years



Distribution of Martial Status (% Total Customers)



■ 1 ■ 2 ■ 3 ■ 4+ ■ 0

- Years Stay** – Most customers (35%) have been in their current city for just 1 year
- Marital Status** – Unmarried customers form the majority at 58%
- City Category** – Over half of the customers (53%) come from Category C cities
- Occupations** – Occupation type 4 is the most common, making up 13% of customers
- Age Group** – The largest age group is 26–35 years, covering 35% of customers
- Gender** – The customer base is male-dominated, with 72% males and 28% females

Insights :

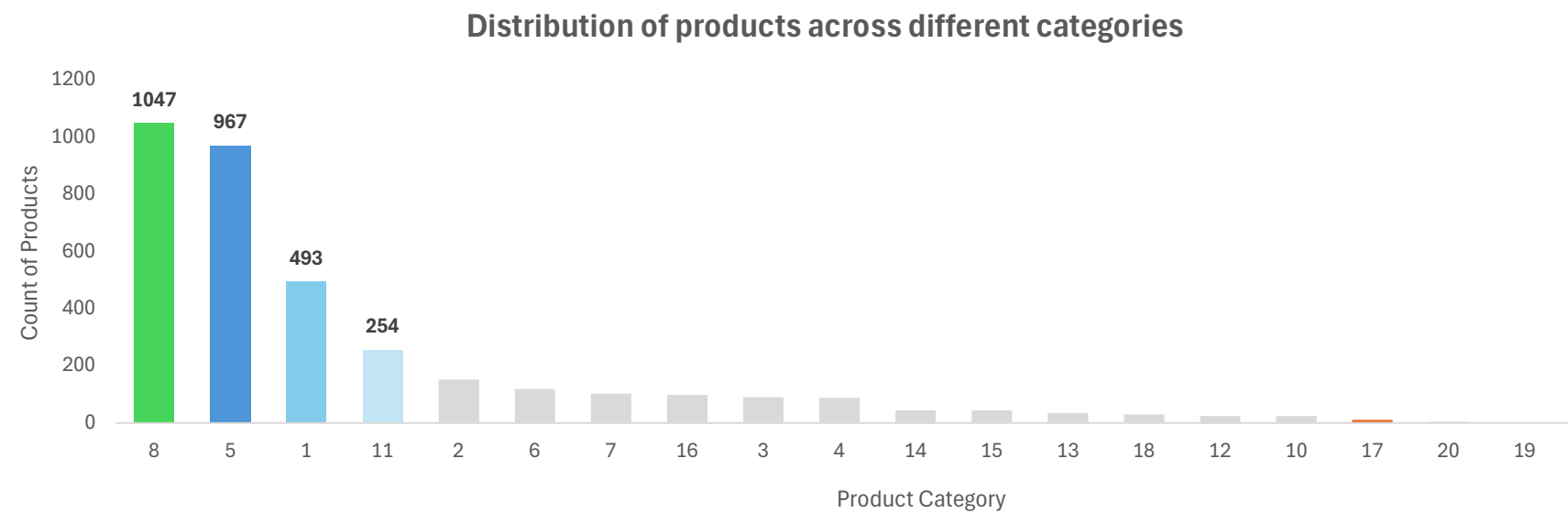
- Years Stay** – High share of recent movers (1 year stay) suggests customers may be in a transition phase or exploring new opportunities.
- Marital Status** – Majority unmarried customers might indicate a younger and more independent target audience.
- City Category** – Strong presence in Category C cities shows potential focus on smaller or developing urban markets.
- Occupations** – Occupation type 4 being the top group could help in creating targeted marketing for this profession.
- Age Group** – Peak in the 26–35 range shows core audience is young working adults.
- Gender** – Predominantly male customer base may indicate product appeal is stronger among men.

3. Product Category & Popularity Analysis

1) Distribution of Products by Category

Finding – Product categories **8** and **5** dominate the whole categories together making up over half of all products

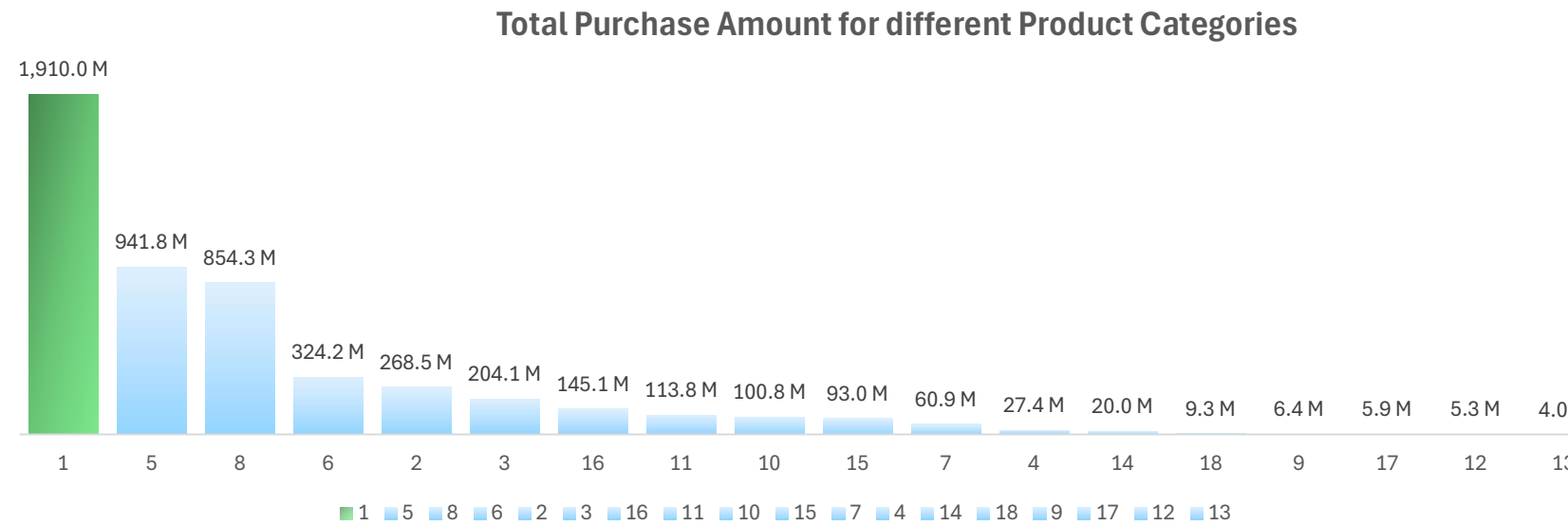
Insight– Focused marketing or stocking in these top categories could drive higher sales due to their wide availability.



2) Most Purchased Product Categories

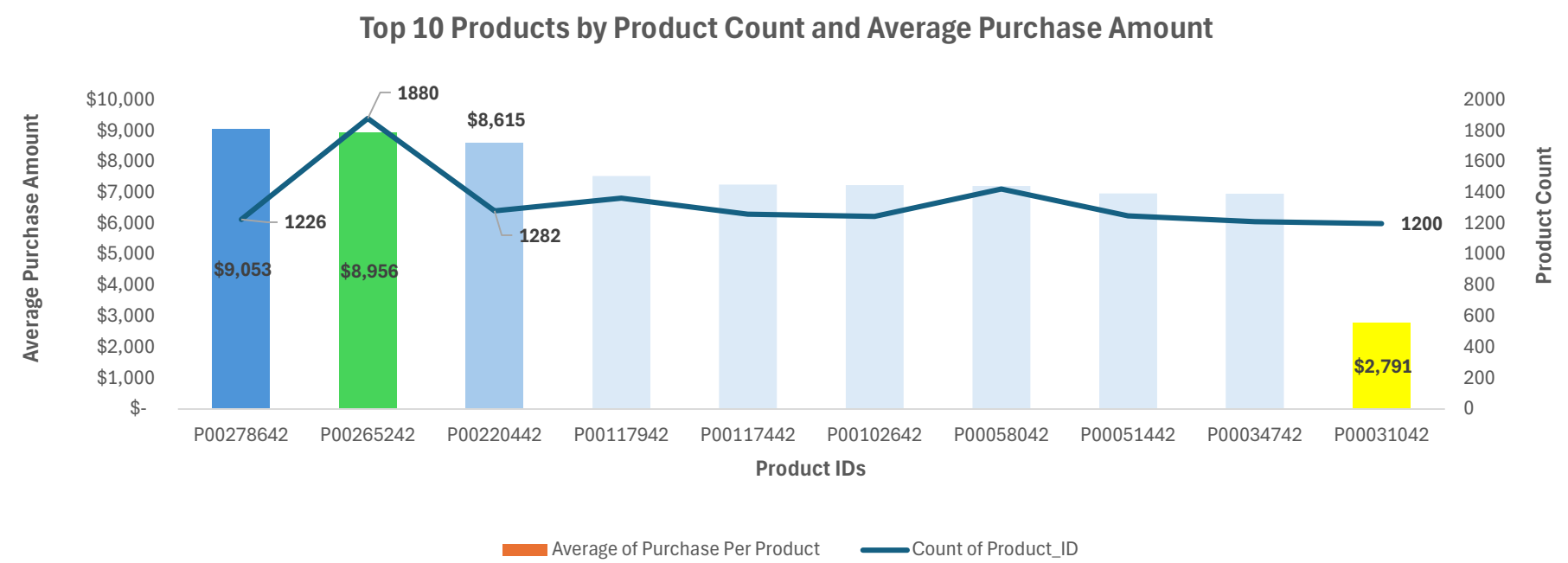
Finding : Category **1** leads with a purchase value of **\$1.91B**, followed by Category **5** (**\$941.8M**) and Category **8** (**\$854.3M**). The bottom categories (e.g., 13, 12, 17) each contribute less than **\$6M**.

Insight : Focusing on the top 3 categories could cover over **70%** of total revenue, allowing Walmart to concentrate marketing and stock efforts where they have the most



impact.

3) Overall Top 10 Products

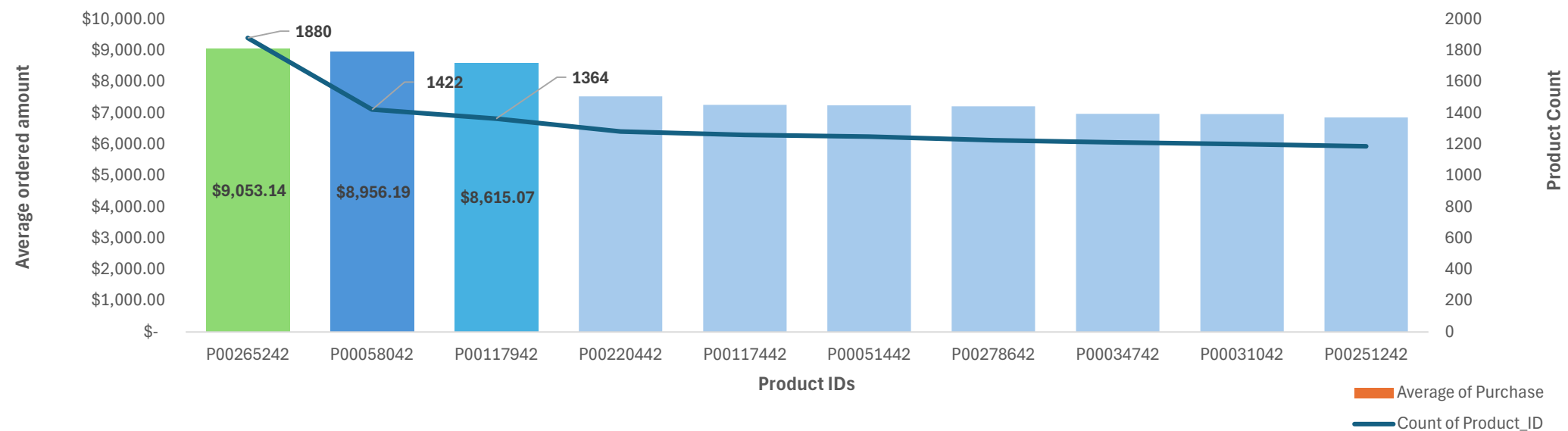


Finding :
Across all categories, **P00265242** is the top-selling product (**1,880** purchases) with an average purchase amount of **\$8,956**, while **P0031042** has a solid sales count (**1,226**) but the lowest average purchase amount (**\$2,791**).

Insight :
High-count, low-value products like **P0031042** could benefit from upselling or bundling to increase basket value, while star products like **P00265242** should remain heavily promoted and well-stocked.

4) Top Products within Categories

Top 10 Product Performance in Selected Categories by Product count and Average amount per product



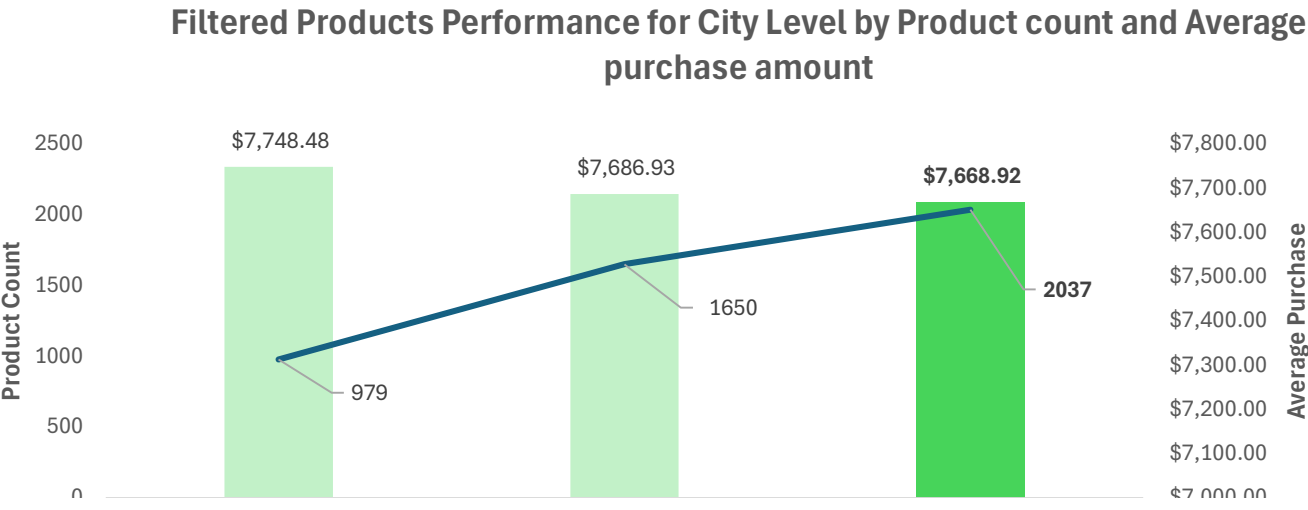
Finding: In Category 8, **P00265242** had the highest sales count (**1,890**) and average purchase amount (**\$9,053.14**), followed by **P00058042** (1,422 sales, **\$8,956.19**) and **P00117942** (1,364 sales, **\$8,615.07**).

Insight: These top products not only sell in high volumes but also have strong per-purchase values, making them prime candidates for hero-product marketing and bundle offers.

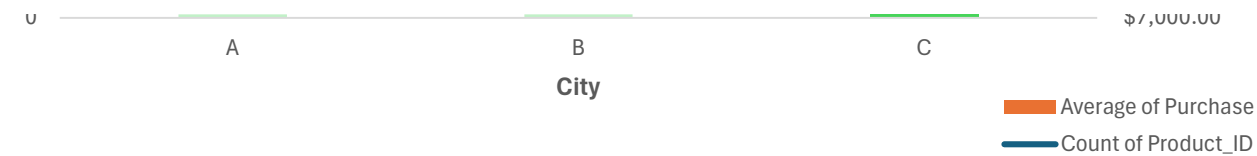
5) City Level - Performance

Finding: For selected products, **City C** recorded **2,037** purchases, followed by **City B** (**1,650**) and **City A** (**979**). Average purchase amounts across cities are close, around **\$7,660–\$7,750**.

Insight: Demand differences are driven by volume rather than pricing. Inventory planning should prioritize high-demand cities like City C to prevent stockouts.



cities like City C to prevent stockouts.



6) Age Group Preferences

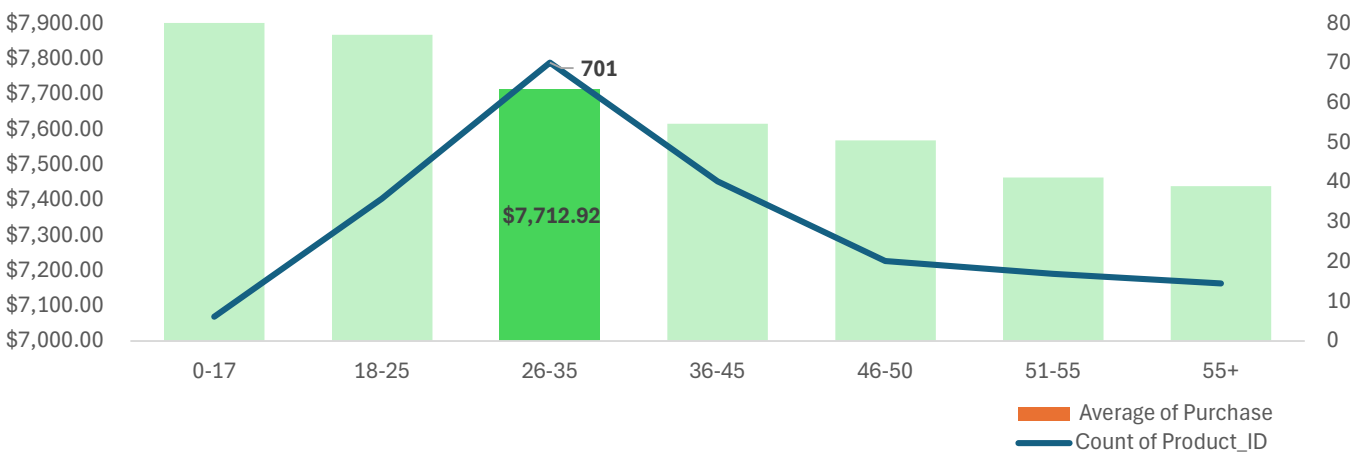
Finding:

The **26–35** age group had the highest purchase count (**701**) for selected products, with an average spend of **\$7,712**. while older groups (46+) had fewer purchases but similar average spend (~\$7,700).

Insight:

Younger customers dominate volume, making them the main target for promotional campaigns, while older customers’ steady spending offers an opportunity for targeted loyalty programs to boost volume.

Selected Products and City by Age groups Product Count count and Average ordered Amount



Summary --

By drilling down from **categories → top products → cities → age groups**, Walmart’s highest revenue comes from Category 1 and Category 8, led by products like P00265242 (1,890 sales, \$9,053 avg).

City C shows the strongest demand, especially for top products like P00265242 and P00058042, while the 26–35 age group leads in purchases across categories.

To boost growth, Walmart should focus promotions on high-performing SKUs in City C, target younger shoppers with tailored campaigns, and test bundling strategies for lower-value products like P0031042.

This layered analysis allows Walmart to create **category-focused, city-specific, and age-targeted strategies** that maximize revenue and optimize inventory allocation.

3. Frequency & Monetary Analysis

Frequency = how often products/customers appear in transactions.
Monetary = total purchase value.
Then ranked them and built segments: **Best, Loyal, Need Attention, Risk.**

Best = 27%
Loyal = 26%
Need Attention = 25%
Risk = 22%
(So ~50% of customers are stable, ~50% need recovery or attention.)

1) City Category Segementation

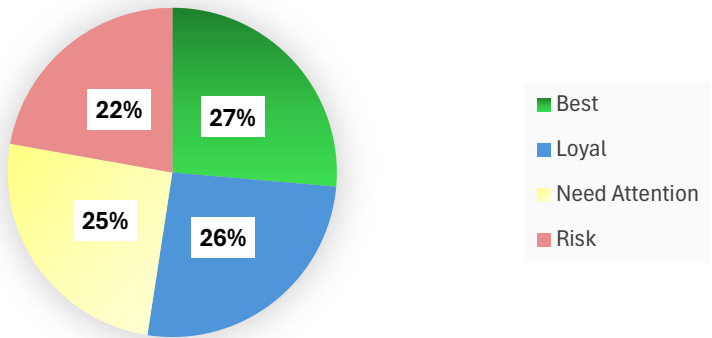
City B is the strongest contributor to the **Best segment (170%)** → overrepresented, a highly valuable customer base.

City A also performs well in Best (144%), but **slightly weaker in Loyal (92%)**.

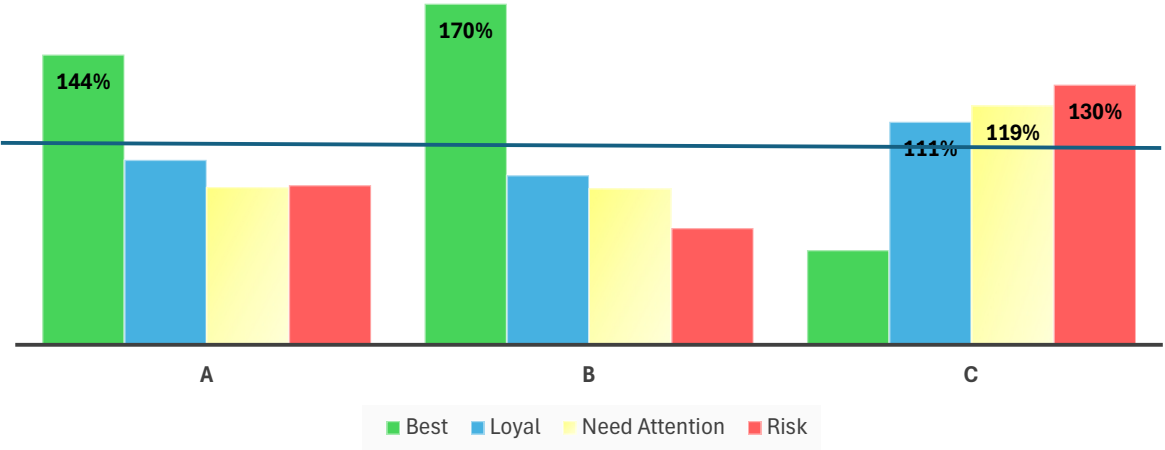
City C is problematic:

Risk is overrepresented (130%) → high churn potential.
Also **Need Attention (119%)** → they are struggling in engagement
Best is underrepresented (47%) → weaker quality customers here.

Segment Distribution - % of Total Customers



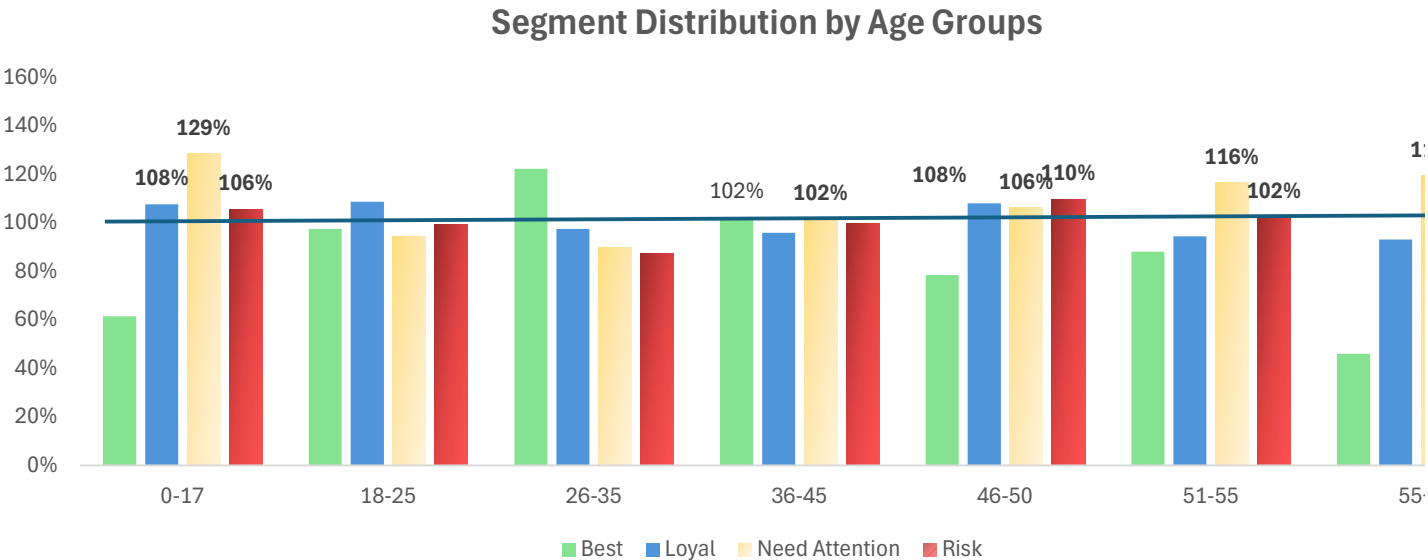
Segment Distribution by City Category



Takeaway: Prioritize **City B** (expand offers, retention). City C needs **recovery strategies** (reduce risk, engagement campaigns).

2) Age Group Segmentation

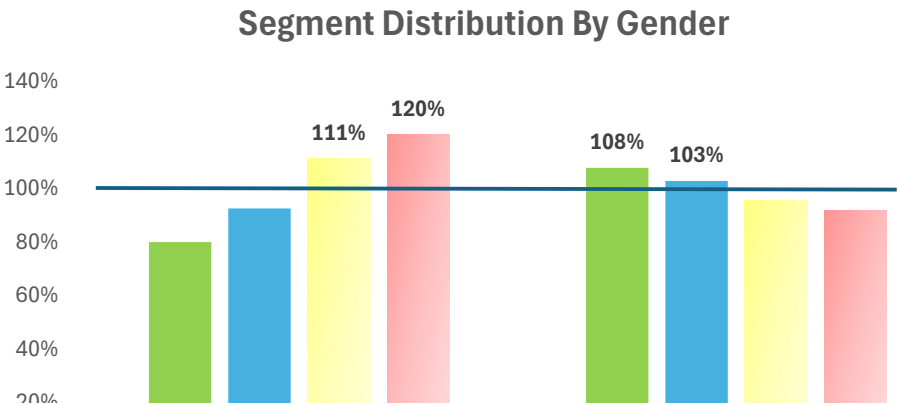
26–35: Strongest age group for **Best (122%)** → your most premium group.
18–25 & 36 -40: Balanced, slightly above 100% in Loyal,Best,Attention stable group.
0–17: Weak Best (61%) but **Need Attention (129%)** → young customers not fully converting to Best.
55+: Problematic:
Risk extremely high (152%).
Need Attention also overrepresented (119%).
Best only 46%.



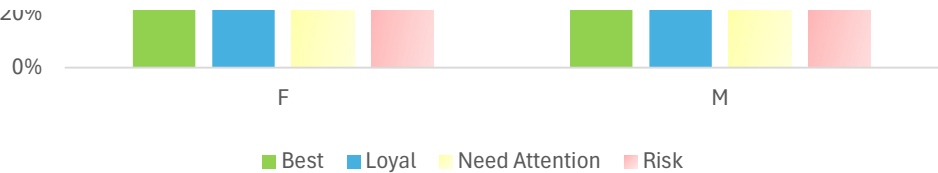
Takeaway: Focus heavily on **26–45** as your prime Best customers.
The **55+ group** is at risk → needs **special retention & engagement programs**.
The **0–17 group** shows opportunity if nurtured (future potential but currently weak Best).

3) Gender Segmentation

Males:
Overrepresented in **Best (108%)** and **Loyal (103%)** → stronger customer base.
Lower Risk (92%) → safer group.
Females:
Overrepresented in **Need Attention (111%)** and **Risk (120%)**.
Underrepresented in Best (80%).



Underrepresented in Best (33.3%).



Takeaway: Males are more stable & valuable. Females show **higher disengagement/risk**, so targeted **female-centric campaigns** are needed.

