Pet E-commerce Customer & Product Analysis

Datasets Used

- customers.csv
- products.csv
- orders.csv
- order_items.csv
- events.csv

All files were cleaned, merged where needed, and datetime columns were standardized.

Q1. Customer Summary Table

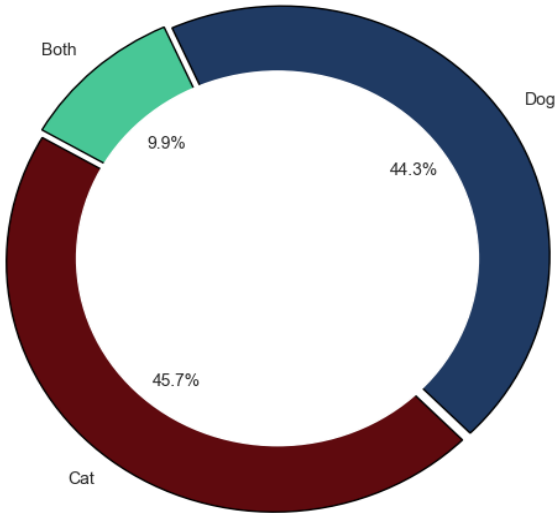
Objective: Generate customer metrics as:

- Total number of orders
- Total spending
- Premium membership status
- Pet type owned

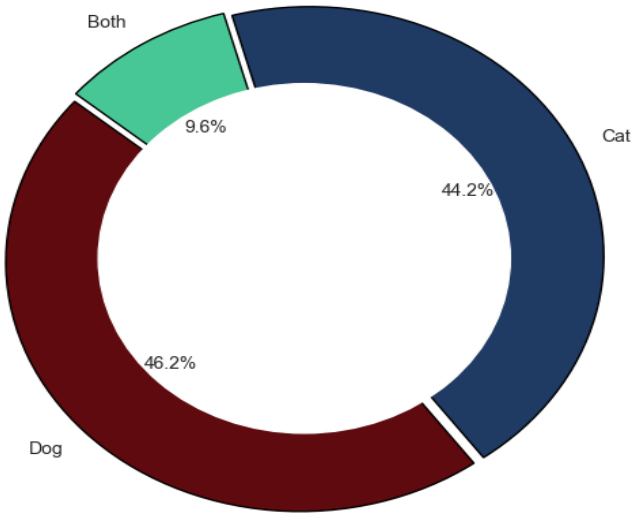
Approach:

- Merged customers.csv and orders.csv to count orders and sum spend.
- Extracted pet ownership from customer records.
- Added a “premium” column from original customer metadata.

Pet Type(Count) Distribution by Customers



Total Spends of Customers by Pet Type



Insights:

- **Customers spent more on dogs overall.**
- **Cat owners outnumber dog owners, but their spending is slightly lower.**
- **~10% of customers own both dogs and cats and contribute noticeably to revenue.**

Prepared: customer_summary.csv

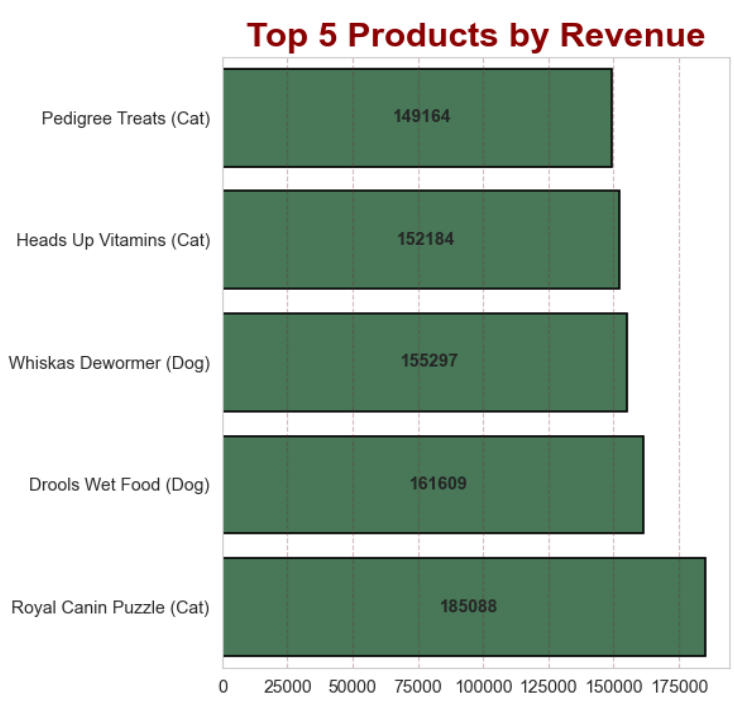
Q2. Product Performance Analysis

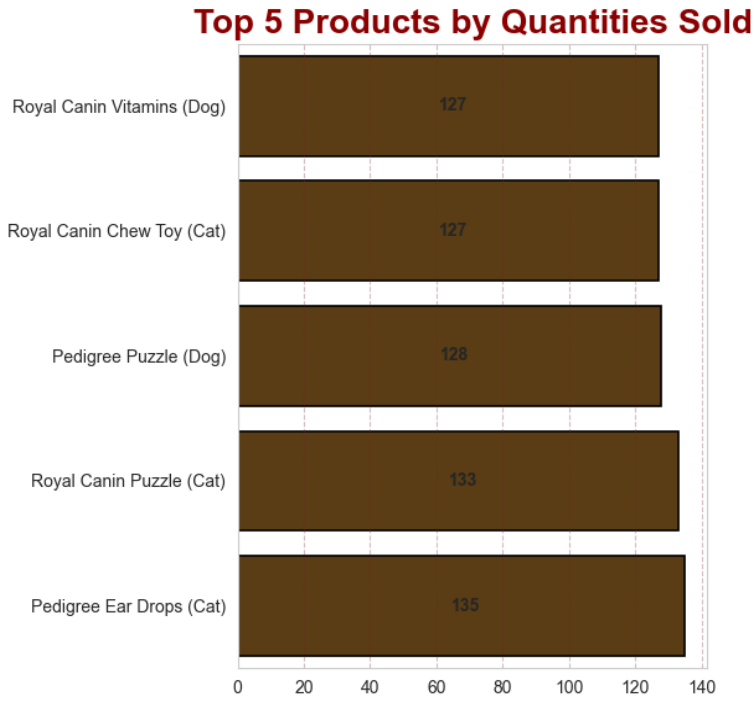
Objective: Identify top products by:

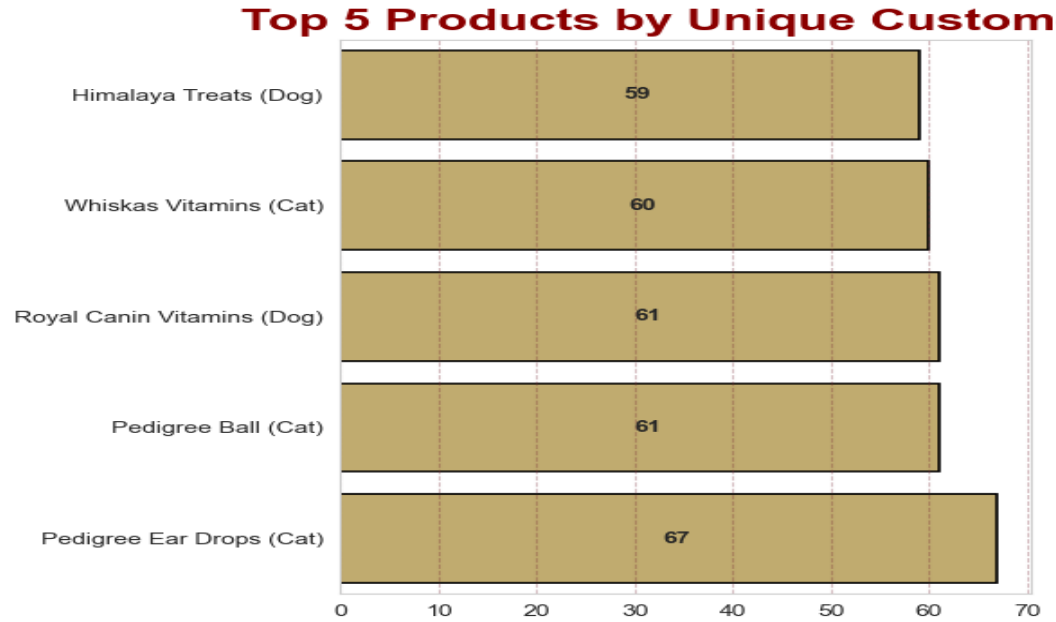
- Quantity sold
- Revenue generated
- Unique customer count
  Also, label each product as for Dog or Cat.

Approach:

- Merged order_items and products to calculate revenue and quantity sold.
- Renamed product names as "Product - (Pet Type)" for clarity.
- Aggregated sales and sorted by revenue, quantity, and unique buyers.







Insights:

- **The top products are mostly Cat-related.**
- **Cat Products have high purchase count.**

Prepared: top_products.csv

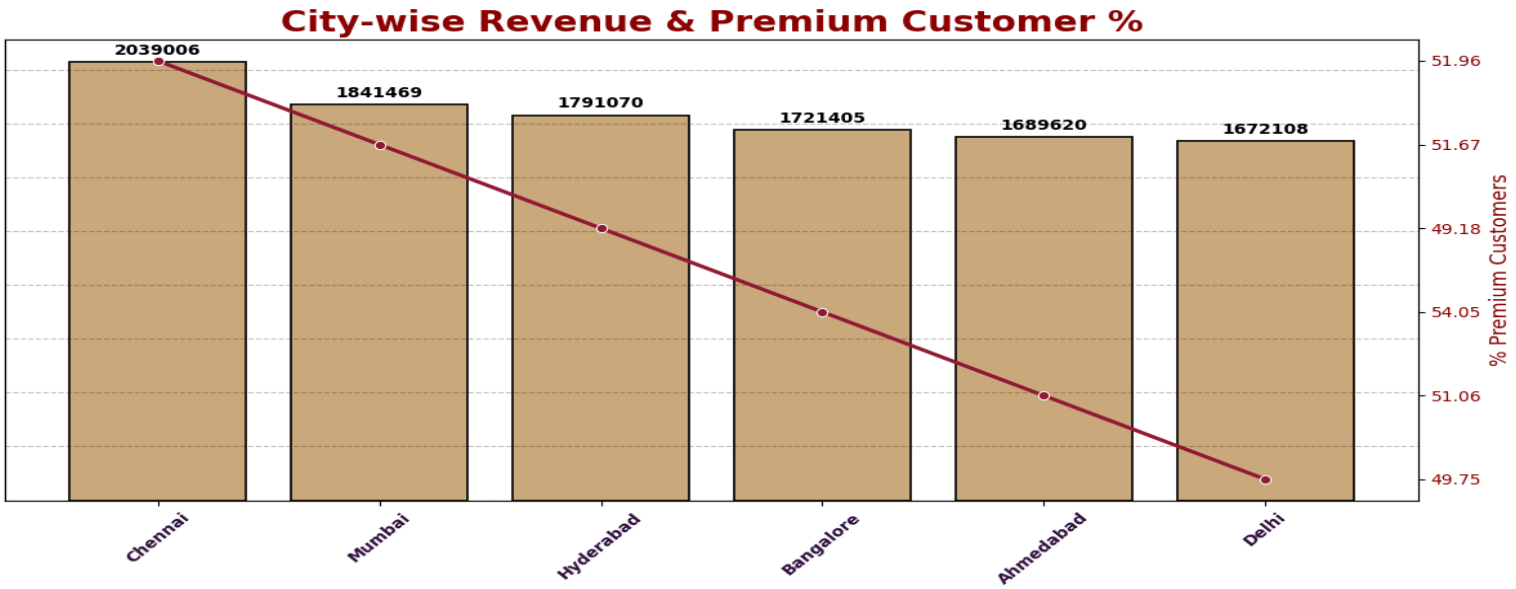
Q3. City-Level Revenue Metrics

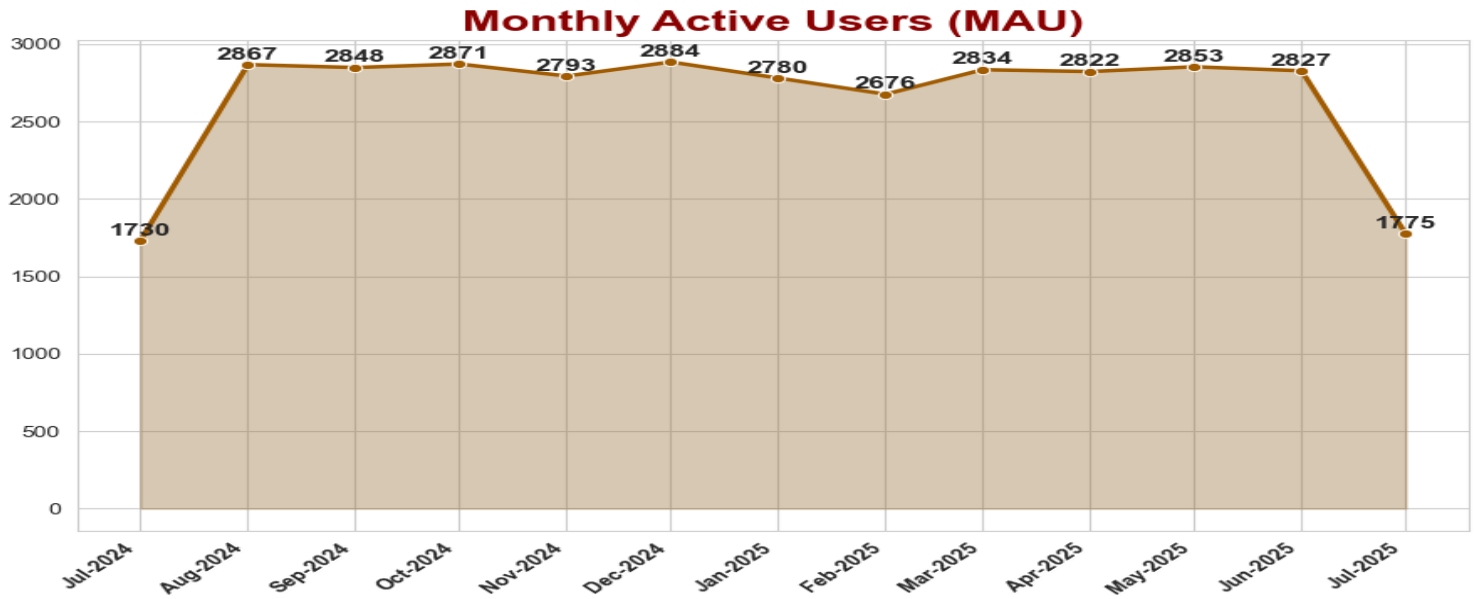
Objective: For each city:

- Number of customers
- Number of orders
- Total revenue
- % of premium customers

Approach:

- Grouped by city after merging customer and order data.
- Calculated revenue and order counts, flagged premium users.





Insights:

- **Chennai leads in Revenue and Premium Customer count (%).**
- **Hyderabad leads in total orders placed.**

- **Steady Monthly Active User (MAU) activity observed from Aug 2024 to June 2025.**

Prepared : city_metrics.csv

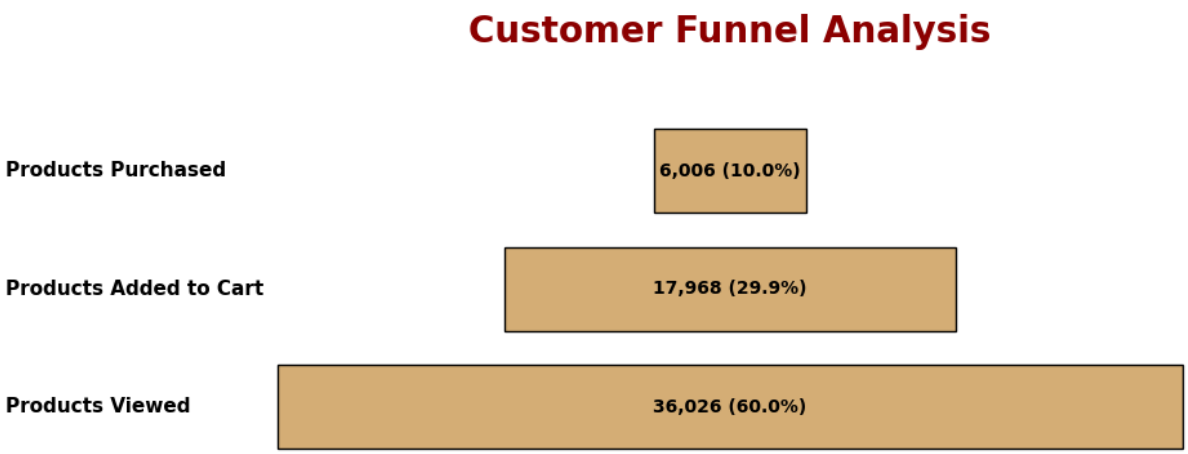
Q4. Funnel Conversion Metrics

Objective:

- For each customer, compute:
  - Views, add-to-cart, purchases
  - Conversion rates:
    - View → Cart
    - Cart → Purchase

Approach:

- Parsed events.csv by event type.
- Grouped and counted events per customer.
- Calculated conversion rates with denominator protection.



Insights:

- **By Count:**
  - **~30% of customers add products to cart**
  - **~10% completed purchasing a product**
  - **~60% of customers viewed Products**

Prepared : customer_funnel_metrics.csv

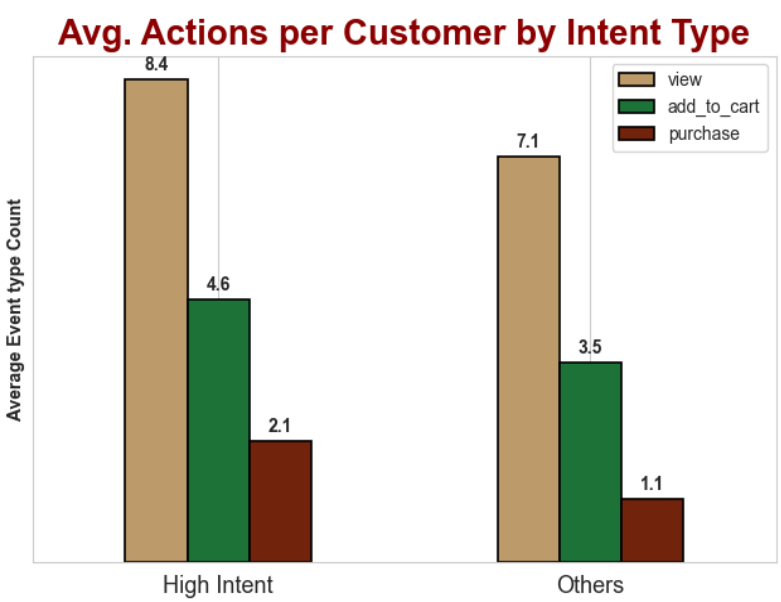
Q5. High-Intent Customers

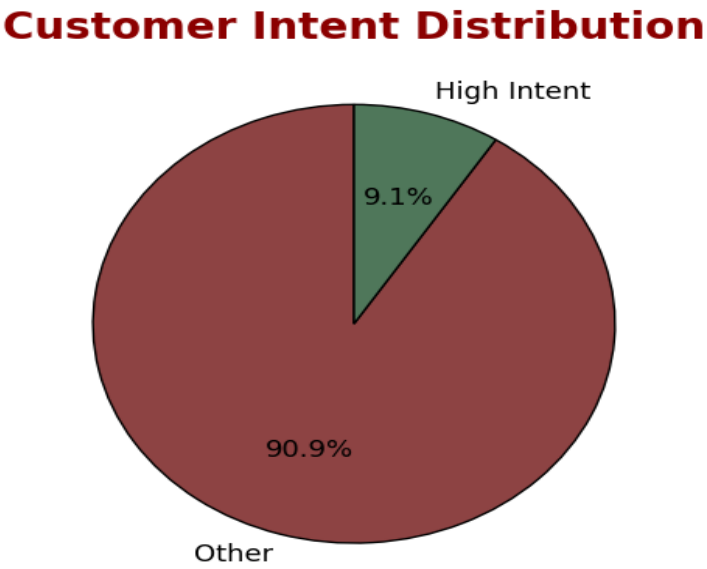
Objective: Flag customers as high intent if:

- 5 product views
- 2 add-to-cart events
- ≥1 purchase in the last 30 days

Approach:

- Defined a function to apply these conditions.
- Used latest event date as reference.
- Returned DataFrame with customer_id and is_high_intent flag.



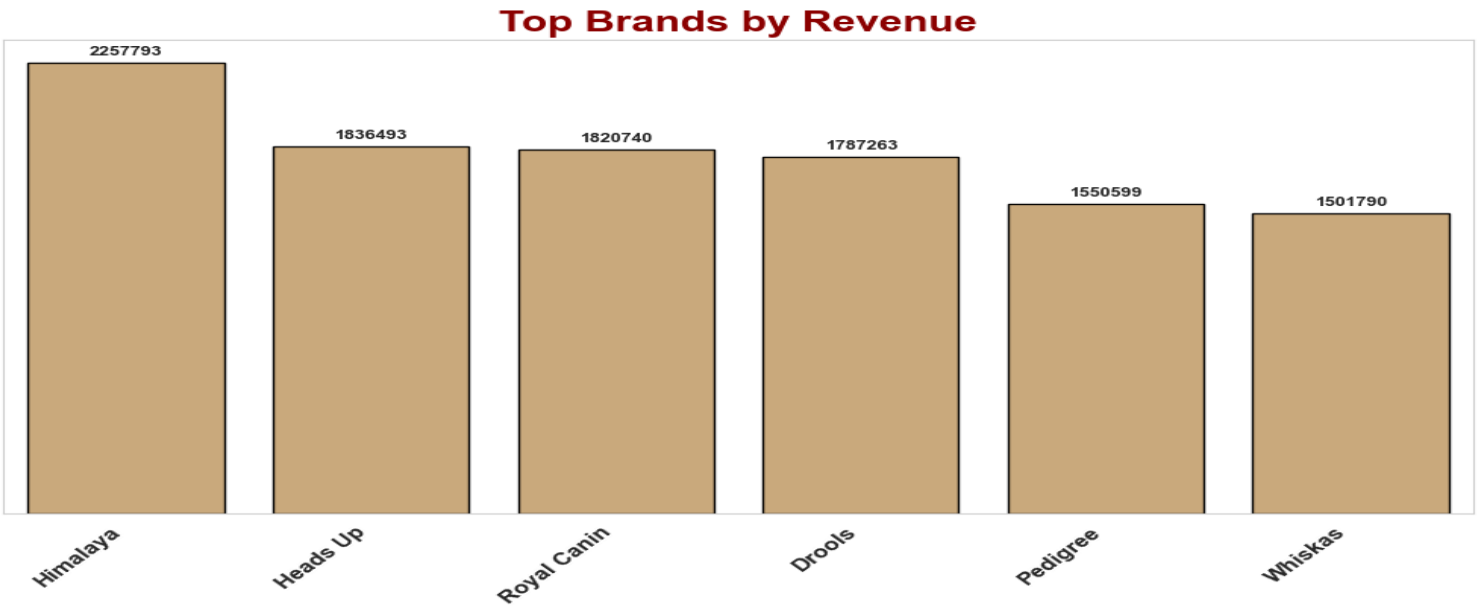


Insights:

- **Only 9.1% of users met this criteria out of a sample of 5000.**
- **High Intent users had higher average event frequency, suggesting stronger engagement and purchase potential.**

Printed: high_intent_df DataFrame

Brand-Level Analysis



- Analysed top brands by revenue**:**
  - **Himalaya was the top-performing brand in terms of sales.**
  - **Heads Up and Royal Canin followed closely afterwards Himalaya.**
  - **Whiskas showed the lowest performance, indicating lower customer demand.**
- **These insights help prioritize Inventory planning and Product stocking based on brand performance.**