

Data Visualization and Interpretation

2025

Autumn 2025

Assignment 01

Name:Sanya Agarwal

Roll no.:R2142230839

Sap_ID:500120145

Program: B.Tech CSE

Semester: 05

Batch: 01 (Data Science)

Date of assignment: 26-09-25



School of Computer Science,
University of Petroleum and Energy Studies,
Dehradun

Assignment-1

Dataset Used

DataSet details , the columns are as mentioned below :

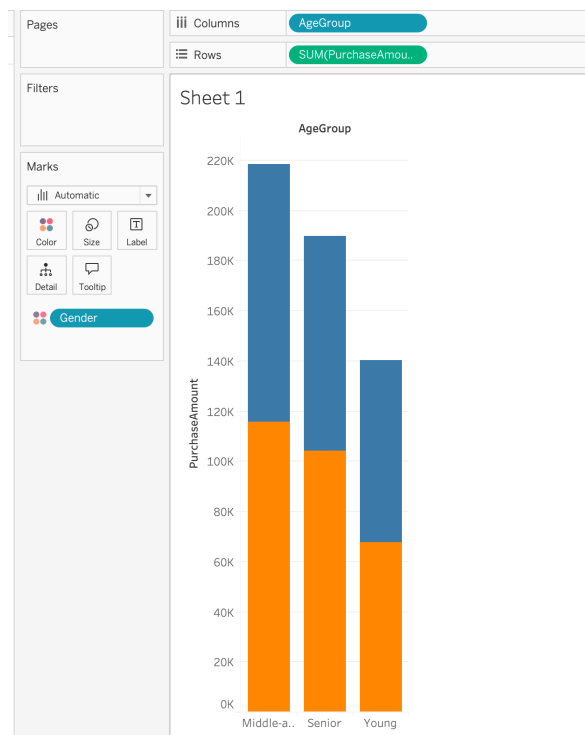
1. CustomerID – Unique ID for each customer
2. Gender – Male/Female
3. AgeGroup – Young, Middle-aged, Senior
4. ProductCategory – Groceries, Furniture,Clothing, Electronics
5. PurchaseAmount – Amount spent (in ₹)
6. PurchaseDate – Date of purchase
7. Region - East, South, North, Central, West
8. Store Type - Flagship, Outlet, Online

Data Set							
CustomerID	Gender	AgeGroup	ProductCategory	PurchaseAmount	PurchaseDate	Region	StoreType
B101	Female	Middle-aged	Groceries	18420	2025-01-13	East	Flagship
B102	Male	Senior	Furniture	9700	2025-02-11	South	Outlet
B103	Female	Young	Clothing	15720	2025-03-16	North	Online
B104	Male	Middle-aged	Electronics	21540	2025-03-21	Central	Flagship
B105	Female	Senior	Groceries	6500	2025-01-27	West	Outlet
B106	Male	Young	Furniture	10800	2025-02-06	East	Outlet
B107	Female	Middle-aged	Electronics	13400	2025-03-04	South	Online
B108	Male	Senior	Groceries	21800	2025-01-19	North	Flagship
B109	Female	Young	Clothing	2950	2025-03-11	Central	Outlet
B110	Male	Middle-aged	Furniture	18920	2025-01-28	West	Flagship
B111	Female	Senior	Groceries	12280	2025-02-25	East	Online
B112	Male	Young	Clothing	15900	2025-03-03	South	Outlet
B113	Female	Middle-aged	Electronics	23780	2025-02-10	North	Flagship
B114	Male	Senior	Furniture	5820	2025-01-31	Central	Outlet
B115	Female	Young	Groceries	8700	2025-02-18	West	Online
B116	Male	Middle-aged	Clothing	15290	2025-03-09	East	Flagship
B117	Female	Senior	Electronics	21000	2025-01-21	South	Outlet
B118	Male	Young	Furniture	6100	2025-03-15	North	Online
B119	Female	Middle-aged	Groceries	15220	2025-02-24	Central	Outlet
B120	Male	Senior	Clothing	7820	2025-01-30	West	Flagship
B121	Female	Young	Electronics	9000	2025-03-18	East	Online
B122	Male	Middle-aged	Furniture	11740	2025-02-12	South	Flagship

Bar Charts

1. Sum of Purchase Amount vs Age Group

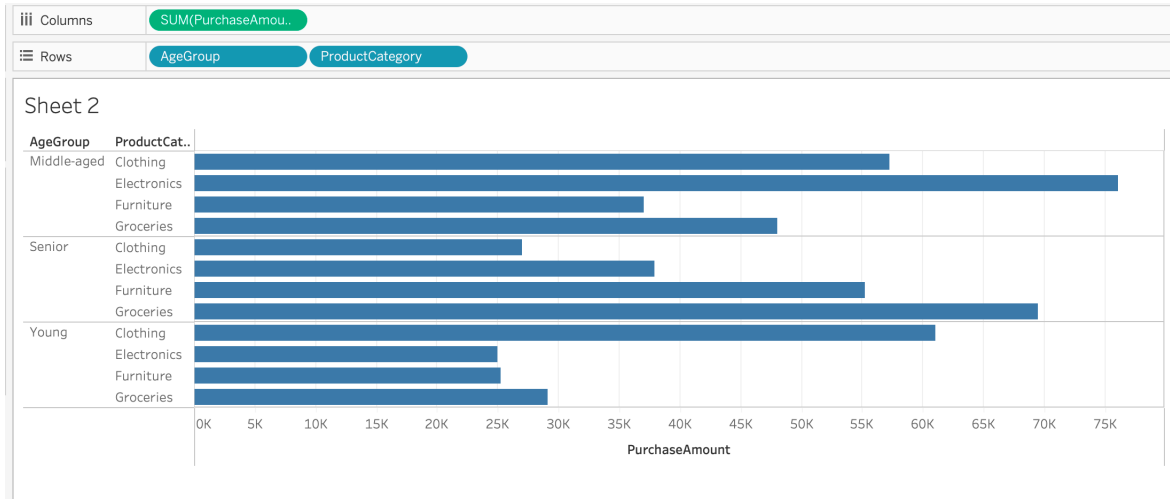
Reason: To compare the spending behavior of male and female customers, classified into different age groups



Interpretation: shows total purchase amounts by age group, broken down by gender. Middle-aged individuals contribute the most overall, with noticeable gender differences in spending across all age groups. This helps identify key demographics for targeted marketing.

2. Average Purchase Amount vs AgeGroup

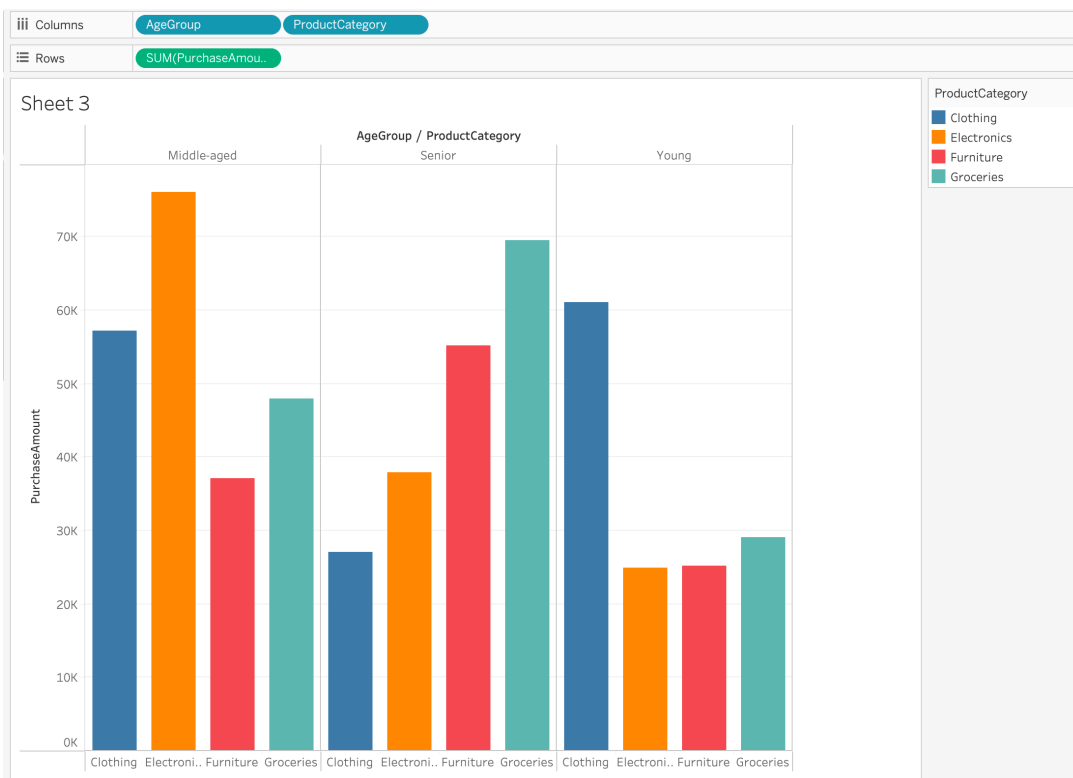
Reason: To identify which age group spends the most on average.



Interpretation: If middle-aged customers spend the most, they might be in stable income groups. If young customers spend less, they may prefer cheaper options like groceries or clothing.

3. Average Purchase Amount vs AgeGroup

Reason: To find which product category generates the highest revenue overall.

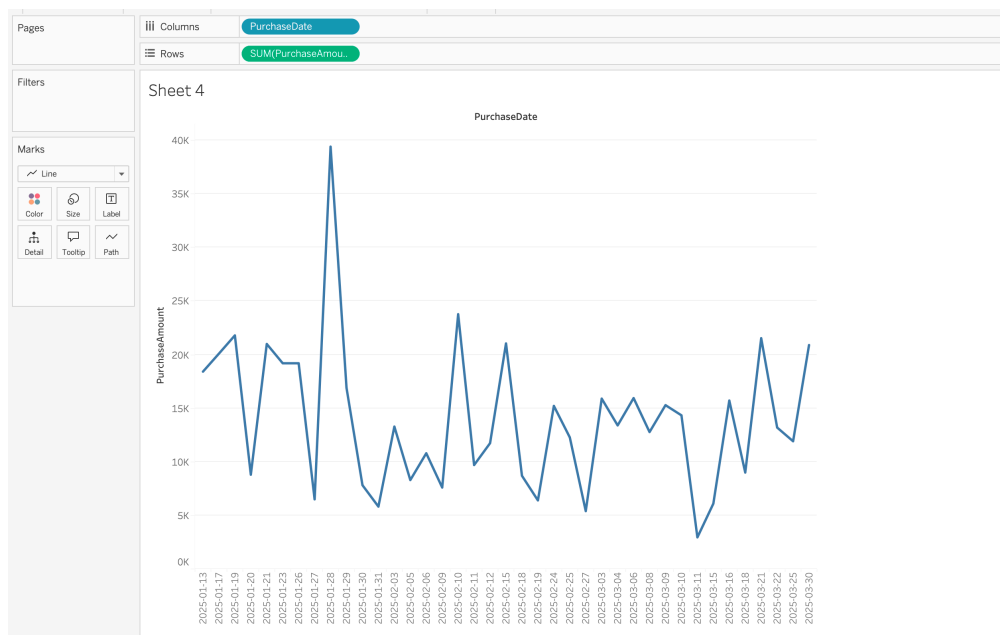


Interpretation: If electronics is the tallest bar, it is the strongest revenue driver. A smaller bar for clothing means they sell more frequently but with lower transaction values.

-Line Chart

1. Total Purchase Amount Over Time (PurchaseDate vs Sum)

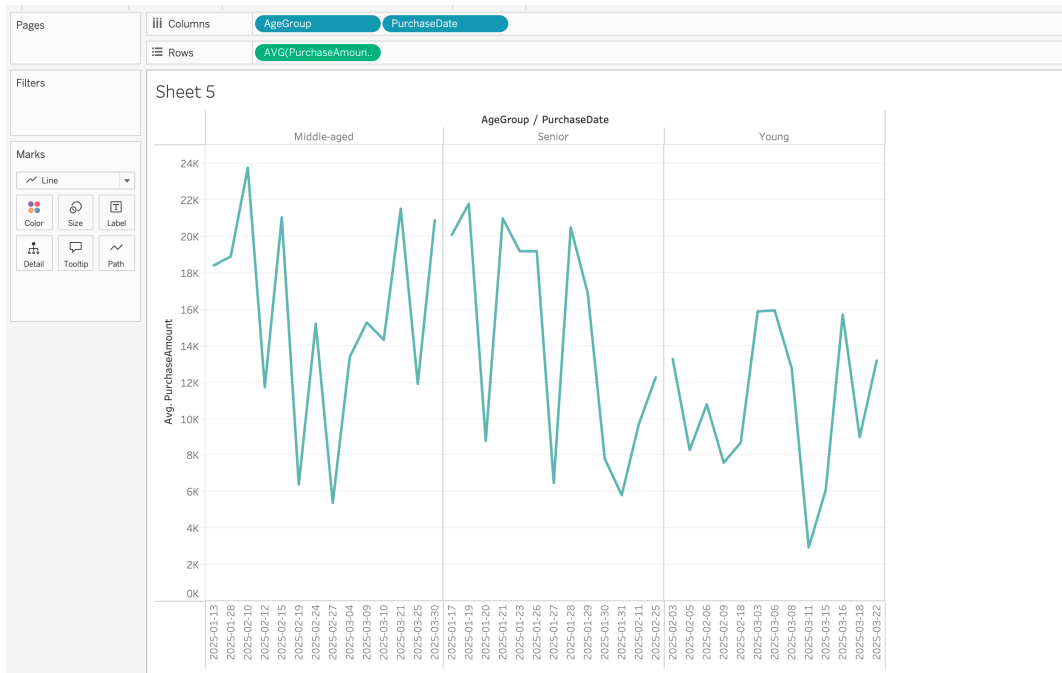
Reason: To check sales trends across days and see demand peaks.



Interpretation: Peaks might appear on weekends or festival dates, showing higher customer activity. Dips mean low shopping activity on certain weekdays.

2. Average Purchase Amount by AgeGroup Over Time

Reason: To see how spending by each age group changes with time.

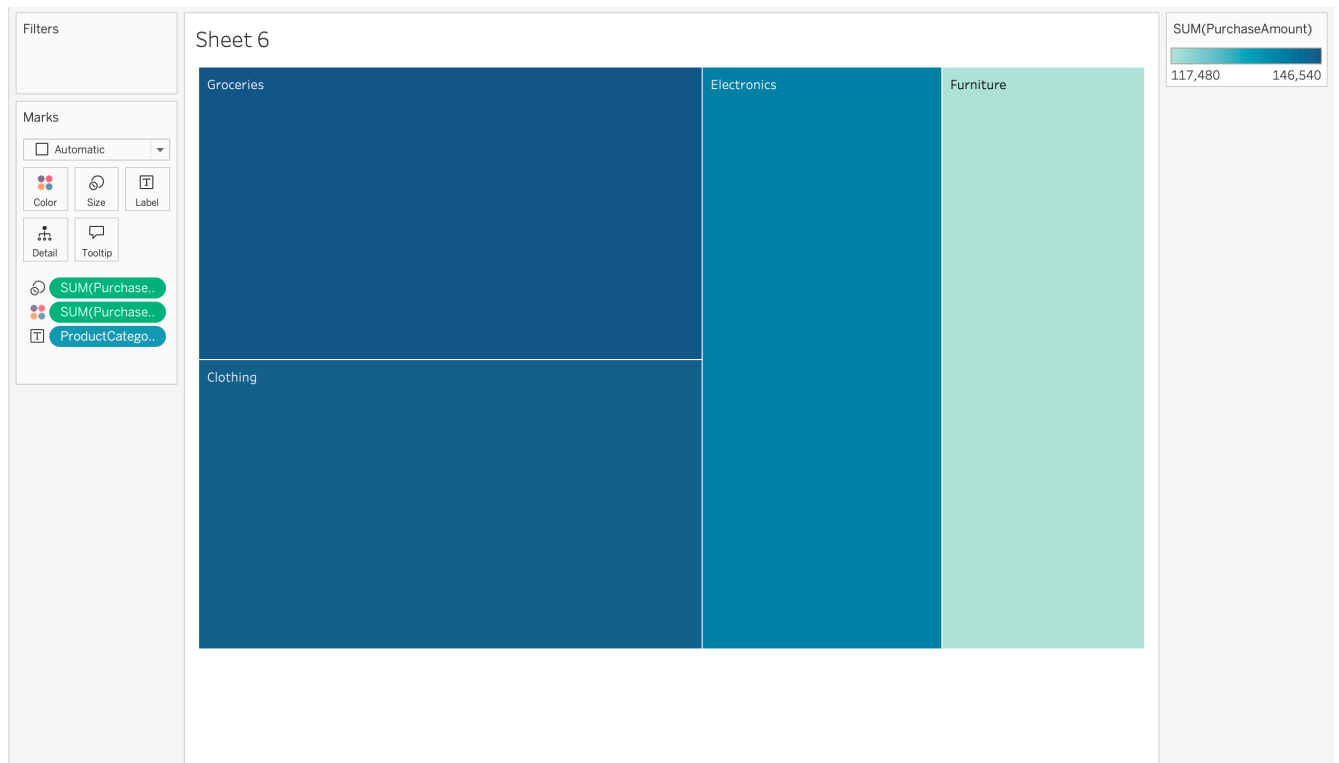


Interpretation: The chart shows that middle-aged customers spent more in January but declined in February, seniors maintained steady spending with a slight rise, while young customers showed a dip in February followed by a sharp increase in March. This indicates seniors are consistent buyers, while middle-aged and young customers show fluctuating patterns.

-Tree Maps

1. Purchase Distribution by ProductCategory

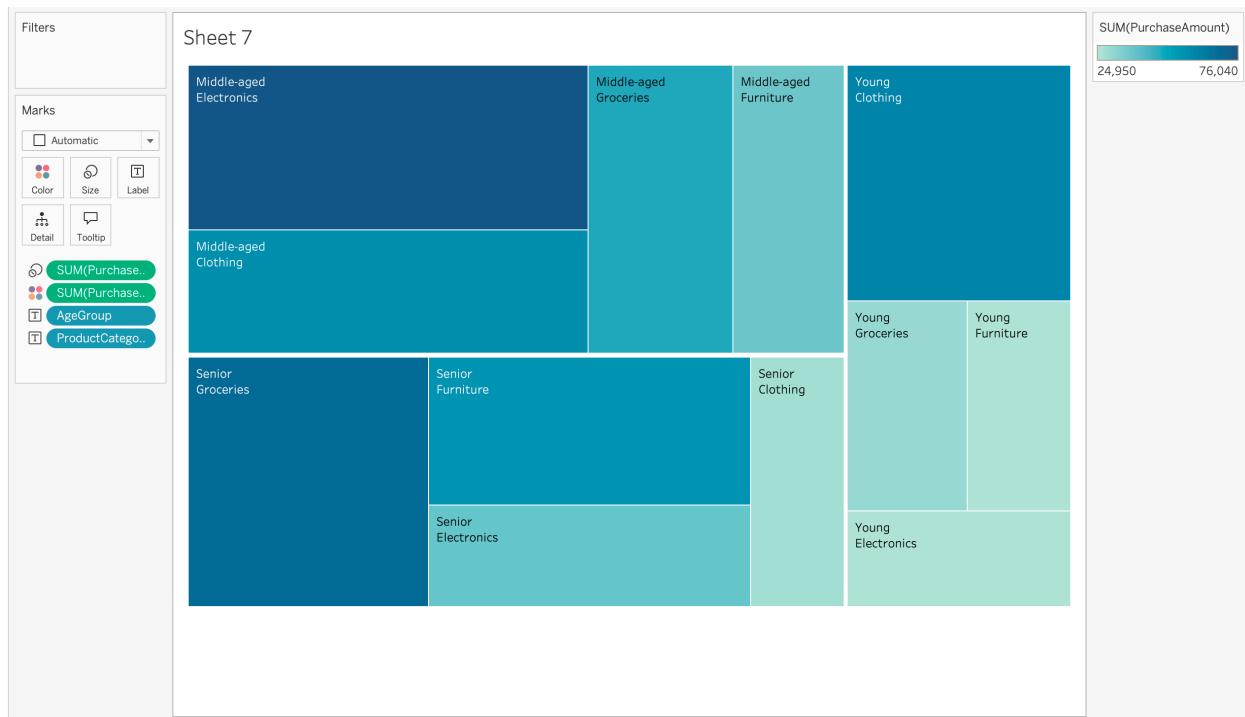
Reason: To visualize the share of each product category in total sales.



Interpretation: If groceries occupy the largest space, they dominate revenue. Smaller areas like electronics and furniture indicate they contribute less.

2. Purchase Distribution by AgeGroup

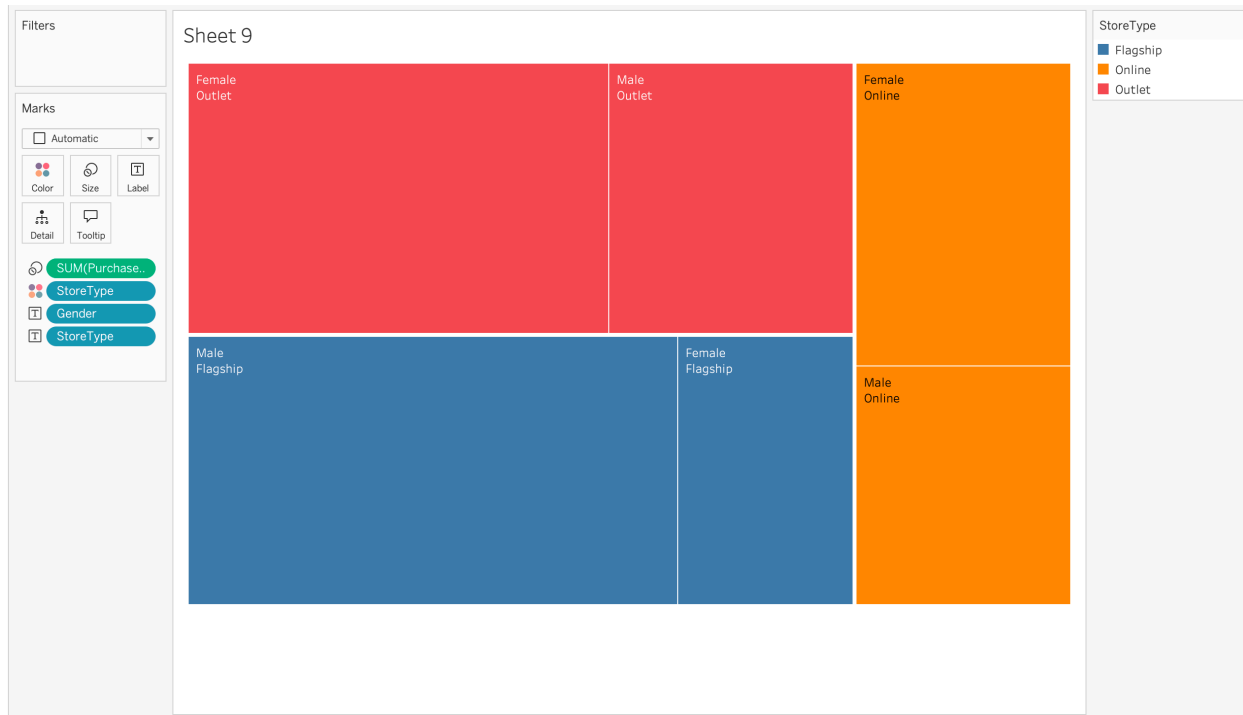
Reason: To see which age group contributes more to which category



Interpretation: Middle-aged customers are the biggest spenders, particularly on electronics and clothing. Seniors focus more on groceries and furniture, while young customers spend mainly on clothing but contribute the least across other categories.

3. Purchase Distribution by Gender and StoreType

Reason: To understand which gender uses which store for shopping



Interpretation: Females spend the most at outlet stores, followed by males at outlets, making outlets the top-performing channel. Males dominate flagship store purchases, while females lead in online shopping. Overall, online sales are the weakest, especially among men.