

# Business Analytics Project - GroupBy & Aggregation

Author: Tushar Kalra

Target Role: Business Analyst / Product Analyst / APM

Tech Stack: Python (Pandas), Jupyter Notebook, Data Analysis

## 1. Concept: GroupBy (Aggregation)

GroupBy = Splitting dataset into groups (by a column like customer, day, category) and applying aggregation functions (sum, avg, count, etc.) to generate summary insights.

Formula:

GroupBy(Key) = SUM / AVG / COUNT (Values)

Example:

- Total Spend per Customer = SUM(order\_amount) grouped by customer\_id
- Daily Revenue = SUM(order\_amount) grouped by order\_date

## 2. Sample Dataset (Raw Orders)

order_id	customer	day	category	amount
----------	----------	-----	----------	--------

1	A	Mon	Food	200
---	---	-----	------	-----

2	B	Mon	Food	300
---	---	-----	------	-----

3	A	Tue	Electronics	100
---	---	-----	-------------	-----

4	A	Tue	Food	400
---	---	-----	------	-----

5	B	Wed	Electronics	150
---	---	-----	-------------	-----

6	C	Wed	Food	500
---	---	-----	------	-----

### 3. Application (Aggregations)

#### (a) Total Spend per Customer

Formula: SUM(amount) grouped by customer

Code: `df.groupby('customer')['amount'].sum()`

Result:

customer	total_spend
A	700
B	450
C	500

Insight: Customer A spends the most (Rs700).

#### (b) Daily Revenue

Formula: SUM(amount) grouped by day

Code: `df.groupby('day')['amount'].sum()`

Result:

day	total_revenue
Mon	500
Tue	500
Wed	650

Insight: Wednesday has the highest sales (Rs650).

#### (c) Category Share of Revenue

Formula:  $(\text{SUM}(\text{amount\_category})/\text{SUM}(\text{amount\_total})) \times 100$

Code: `df.groupby('category')['amount'].sum()`

Result:

category	revenue	pct_share
Food	1400	84.85%
Electronics	250	15.15%

Insight: Food category contributes ~85% of revenue -> clear business focus area.

#### (d) Top Category per Day

Formula: MAX(amount) per day across categories

Code: `groupby(['day','category'])['amount'].sum()`

Result:

day	category	revenue
Mon	Food	500
Tue	Food	400
Wed	Food	500

Insight: Food dominates every day -> promotions should focus here.

## 4. Final Business Takeaways

- Customer A is high-value (Rs700) -> treat as premium.
- Food category is driving 85% of revenue -> main growth lever.

- Wednesday peak revenue -> run mid-week campaigns.
- Electronics weak (~15%) -> improve or reduce discounts.

## **5. Recruiter Note**

This project shows ability to:

- Use GroupBy for business summaries
- Write clear formulas + Pandas code
- Translate raw data -> insights & actions
- Present professional business-report style (GitHub/Portfolio ready)