Retail Sales Dataset (Sample)

You can copy this into a CSV named **retail_sales.csv**

```
OrderID,Date,Store,City,Product,Category,Quantity,UnitPrice,TotalPrice,Pay
mentMethod,CustomerType
1001,2024-01-05,Store A,Mumbai,Laptop,Electronics,1,55000,55000,Credit
Card,New
1002,2024-01-05,Store B,Delhi,Shampoo,Personal
Care,3,120,360,Cash,Returning
1003,2024-01-06,Store C,Bangalore,Jeans,Apparel,2,1500,3000,Credit
Card,New
1004,2024-01-06,Store
A,Mumbai,Smartphone,Electronics,1,25000,25000,UPI,Returning
1005,2024-01-07,Store B,Delhi,Bread,Grocery,5,40,200,Cash,New
1006,2024-01-07,Store C,Bangalore,T-Shirt,Apparel,4,800,3200,Credit
Card,Returning
1007,2024-01-08,Store A,Mumbai,Milk,Grocery,10,50,500,UPI,New
1008,2024-01-08,Store B,Delhi,Perfume,Personal Care,1,2500,2500,Credit
Card,Returning
1009,2024-01-09,Store
C,Bangalore,Headphones,Electronics,2,1500,3000,Cash,New
1010,2024-01-09,Store A,Mumbai,Rice,Grocery,3,90,270,Credit Card,Returning
1011,2024-01-10,Store B,Delhi,Shoes,Apparel,1,3000,3000,UPI,New
1012,2024-01-10,Store C,Bangalore,Milk,Grocery,12,48,576,Cash,Returning
1013,2024-01-11,Store A,Mumbai,Charger,Electronics,2,600,1200,Credit
Card,New
1014,2024-01-11,Store B,Delhi,Notebook,Stationery,10,35,350,Cash,Returning
1015,2024-01-12,Store
C,Bangalore,Smartwatch,Electronics,1,8000,8000,UPI,New
1016,2024-01-12,Store A,Mumbai,Biscuits,Grocery,6,25,150,Credit
Card,Returning
1017,2024-01-12,Store B,Delhi,Jacket,Apparel,1,4500,4500,UPI,New
1018,2024-01-13,Store C,Bangalore,Soap,Personal
Care,4,45,180,Cash,Returning
1019,2024-01-13,Store A,Mumbai,Keyboard,Electronics,1,1200,1200,UPI,New
1020,2024-01-13,Store B,Delhi,Shirt,Apparel,2,1100,2200,Credit
Card,Returning
```

Pandas Exercises

1. Load the dataset and display:

- first 5 rows
- last 5 rows
- column names
- shape

---

2. Convert the Date column to datetime and extract:

- Year
- Month
- Day

Add them as new columns.

---

3. Calculate total sales (sum of TotalPrice) for each:

- Store
- City
- Category

---

4. Find the top 5 highest-value orders by TotalPrice.

---

5. Filter the dataset to show only Electronics products with Quantity > 1.

---

6. Add a new column Discount:

- 10 percent discount for Returning customers
- 5 percent discount for New customers

Compute final price after discount.

---

7. Find how many orders were paid using:

- Cash
- Credit Card
- UPI

---

8. Group by Category and compute:

- Total quantity sold
- Total revenue

---

9. Identify the store with the highest total sales.

---

10. Filter rows where Product name contains the letter “a” or “A”.

---

11. Sort the dataset by Date and then by TotalPrice.

---

12. Find the average revenue per order for each CustomerType.

---

13. Create a pivot table:

Rows: Category
Columns: PaymentMethod
Values: TotalPrice (sum)

---

14. Write the filtered Electronics-only dataset into a new CSV file.

---

15. Use method chaining to:

- remove rows with Quantity < 2
- filter Category = Apparel
- compute TotalValue = Quantity * UnitPrice

- sort TotalValue descending
- reset index

Return the final DataFrame.
