Pandas Task 1

1. Display Basic Information.
2. How many transactions were made for each product category?
3. What is the count of each payment method used in the dataset?
4. Sort the dataset by Total Revenue in descending order and list the top 5 transactions.
5. Sort the dataset by Units Sold in ascending order and list the first 10 transactions.
6. Convert the dataset from wide format to long format with Transaction ID, Date, Product Category, Product Name, Region, and Payment Method as identifier variables, and Units Sold, Unit Price, and Total Revenue as measured variables.
7. Create a pivot table to summarize total revenue for each product category by region.
8. Create a pivot table to summarize the number of units sold for each product by date.
9. Check for any missing values in the dataset. If there are any, list the columns with missing values and their respective counts.
10. Verify that there are no missing values in the Date column.
11. Group the dataset by Product Category and calculate the total units sold for each category.
12. Group the dataset by Region and find the average unit price for each region.
13. Use the apply function to create a new column called Discounted Revenue which is 90% of the Total Revenue.
14. Apply a lambda function to the Units Sold column to increase each value by 10%.
15. Filter the dataset to include only transactions where the Total Revenue is greater than 5000.
16. Filter the dataset to include transactions from the North region only.
17. Drop the Payment Method column from the dataset and display the updated dataset.
18. Drop rows where the Units Sold is less than 10 and display the updated dataset.
19. Check for any duplicate rows in the dataset and list them.
20. Remove duplicate rows based on the Transaction ID column and display the cleaned dataset.

#Sample Data

data = { 'Transaction ID': range(1, 51),

'Date': pd.date\_range(start='2024-01-01', periods=50, freq='D'),

'Product Category': ['Electronics', 'Furniture', 'Clothing', 'Groceries', 'Books'] \* 10,

'Product Name': ['Laptop', 'Chair', 'T-Shirt', 'Milk', 'Novel'] \* 10,

'Units Sold': [10, 5, 20, 15, 7] \* 10,

'Unit Price': [1000, 200, 20, 2, 15] \* 10,

'Total Revenue': [10000, 1000, 400, 30, 105] \* 10,

'Region': ['North', 'South', 'East', 'West', 'Central'] \* 10,

'Payment Method': ['Credit Card', 'Debit Card', 'Cash', 'Credit Card', 'Cash'] \* 10 }