Task -1

Background:

You are provided with a dataset **(Raw_Data)** sheet containing records of field contacts made by users (agents). Each record logs who made the contact, where it occurred, and at what time.

The objective is to summarize and analyze the behavior of each user across dates based on this raw data.

The dataset consists of the following columns:

Raw_Data:

- **Region**: User's working region
- **User_ID:** Unique identifier for each user
- **User_Name:** Full name of the user
- **Contact_ID:** Unique identifier for each contact
- **Contact_Start:** Time when contact started
- **Contact_End:** Time when contact ended
- **Contact_Date:** Date of contact

Objectives:

All the solutions must be given in the summary sheet (except the first task) and the formulas used for the solution must be intact (Only values won't be accepted).

- **Covert to a table using the Raw_Data sheet**
- Add users name
- Calculate the total number of Contact made by each user.
- Calculate Average call duration of total contact made by user.
- Find users First contact date .
- Find users Last contact date

Task 2

Problem 1 : Sales Data Cleanup & Categorize Revenue

Tables

sales_transactions

transaction_id	product_code	customer_email	sale_date	amount	payment_status
1	p001	ali@gmail.com	2024-07-01	1,000	Paid
2	P001	ali@gmail.com	2024-07-01	1,000	Paid
3	p002	nabila@outlook.com	2024/07/02	500	paid
4	P003	NULL	01-07-2024	150	Failed
5	P003	sara@domain.com	2024-07-03	150	NULL
6	P003	sara@domain.com	2024-07-03	150.00	failed
7	p002	NULL	2024-13-01	0	Paid
8	p002	abir@example.com	2024-07-15	0	Completed
9	p001	ovi@example.com	2024-07-20	1,200.00	paid
10	P001	ovi@example.com	2024-07-20	1,200	paid

products_catalog

product_code	product_name	category
P003	Office Chair	Furniture
P001	Laptop Model A	Electronics
P002	Smartphone X	Electronics

Tasks:

- **Provide the DDL for sales_transactions and products_catalog.**
- **Clean and standardize the sales_transactions data.**
- **Remove duplicate transactions.**
- **Calculate total revenue per product category where payment_status is 'Paid'.**
- **Report how many duplicate transactions were removed.**

Problem 2: Dhaka Customer Purchase Behavior

Tables

customers

customer_id	name	region	signup_date
1	Alice	Dhaka	2023-01-10
2	Bob	Chittagong	2023-02-15
3	Carol	Dhaka	2023-03-05
4	David	Dhaka	2023-04-01
5	Eve	Sylhet	2023-05-20
6	Frank	Dhaka	2023-06-01
7	Grace	Dhaka	2023-06-15
8	Hannah	Dhaka	2023-07-01

orders

order_id	customer_id	order_date	total_amount
100	1	2024-06-01	600
101	1	2024-07-15	500
102	3	2024-07-10	1200
103	4	2024-07-12	0
104	2	2024-07-05	300
105	6	2024-07-01	600
106	6	2024-07-10	400
107	7	2024-07-05	300
108	7	2024-07-15	800
109	8	2024-07-20	0
110	8	2024-07-22	0

order_items

item_id	order_id	product_name	quantity	unit_price
1000	100	Laptop Model A	1	600
1001	101	Smartphone X	1	500
1002	102	Office Chair	2	600
1003	102	Desk Lamp	1	0
1004	104	Notebook	3	100
1005	105	Desk Lamp	1	600
1006	106	Notebook	1	400
1007	107	Laptop Model A	1	300
1008	108	Laptop Model A	1	800

Tasks:

- **Provide the DDL for customers, orders, and order_items.**
- **Find customers from Dhaka who have at least 2 orders, spent over 1000 in total, and bought at least 2 different products.**
- **List their name, number of orders, total spent, and first order date.**
- **Identify customers who signed up but never made a purchase.**

Deliverables:

Task 1: **The solution should use provided raw_data file link.** Download and Rename the file to **YourName-Excel.xlsx . Solution should use Excel functions to derive all the required insights. Ensure the dataset is processed in a table to generate actionable insights for further analysis and make sure to send the solution with the function intact used to solve this problem. and Submit the entire solution in Excel file .**

Task 2: **The solution should use in SQL queries to derive all the required insights.** Rename the file to **YourName-SQL.txt and Submit the entire solution (BY Problem) in a single text file .**