You are given raw customer data from a CRM system. Your task is to **clean the data and perform an initial exploratory analysis**. This is the kind of data you'd see in real life: messy, inconsistent, and incomplete.

Tasks:

Part A: Data Preprocessing (30 mins)

Perform the following **cleaning steps** and justify your decisions:

1. **Identify and handle missing values**
    - Count how many missing values exist in each column.
    - Describe how you imputed or removed them and why.
2. **Fix inconsistent data formats**
    - Some dates are in different formats, and numerical fields are strings with currency symbols — standardize them.
    - Convert `Customer_Since` to datetime. Convert `Income` to numeric.
3. **Handle categorical data**
    - Standardize inconsistent category labels (e.g., `Male`, `MALE`, `male` → `Male`)
    - Remove invalid or outlier entries (e.g., negative age, unknown cities)
4. **Remove duplicates and obvious data errors**
    - List how many duplicates or errors you found and how you fixed them.

Part B: Exploratory Data Analysis (30 mins)

Once the data is clean, answer the following:

1. **Summarize the dataset**
    - Mean, median, and mode of age and income.
    - Distribution of customers across cities and genders.
2. **Insights generation**
    - Find two interesting patterns in the data.
    - Example: Do certain age groups spend more? Is there a skew toward a certain gender or region?
3. **Data issues you noticed**
    - Are there any red flags or things that don't make sense even after cleaning?

Deliverable: All are mandatory.

- A Jupyter notebook or Excel sheet with your steps.
- Your **cleaned dataset** and **summary answers**.
- A brief explanation of **why you made each preprocessing decision** (in comments or cells).