AE–DM Reconciliation & Safety Reporting Using Base SAS & PROC SQL

PROJECT OVERVIEW

You are working as a Clinical SAS Programmer on **Study ONC101**, a Phase II oncology clinical trial.
Your task is to perform **data validation**, **subject reconciliation**, and **safety analysis** using two datasets:

- **DM (Demographics)**
- **AE (Adverse Events)**

You must use **Base SAS procedures** and **PROC SQL** to clean the data, identify discrepancies, and generate summary reports.

DATASETS PROVIDED

1. **DM Dataset (10 subjects)**
   STUDYID, USUBJID, AGE, SEX, RACE, ARM, COUNTRY
2. **AE Dataset (Adverse Events)**
   STUDYID, USUBJID, AEDECOD, AESTDTC, AEENDTC, AESER, SEVERITY

PROJECT OBJECTIVES

By the end of this project, you should be able to:

- Perform **data cleaning**
- Apply **range, consistency, and integrity checks**
- Reconcile subjects between DM and AE datasets
- Produce **common safety summary outputs**
- Use **Base SAS programming & PROC SQL** effectively
- Understand CR programming workflow

PROJECT TASKS

TASK 1: Import & Inspect Data

1. Create DM and AE datasets using **provided code**.
2. Print observations, check dataset structure, and ensure date formats are correct.

Deliverables:

- PROC CONTENTS output
- First 10 observations (PROC PRINT)

TASK 2: Data Cleaning & Validation

Perform the following checks in both **Base SAS** and **PROC SQL**:

1. AEs with missing AEDECOD
2. AEs where **AESTDTC > AEENDTC**
3. AEs for subjects **<18 years old**
4. Identify **duplicate AEs** using (USUBJID + AEDECOD + AESTDTC)
5. Identify missing values in:
    - AESTDTC
    - AEENDTC
    - AESER
    - SEVERITY

Deliverables:

- A cleaned AE dataset
- A report listing all data discrepancies

TASK 3: Subject Reconciliation

Use PROC SQL JOINs to identify:

1. Subjects in AE **but not** in DM
2. Subjects in DM **with no AEs**
3. Count of subjects in each category

Deliverables:

- Reconciliation report table
- List of unmatched subjects

TASK 4: Summary of AEs

Generate summary outputs:

A. Subject-Level AE Summary

- Total number of subjects
- Number of subjects with ≥1 AE
- % of subjects with AEs

B. Event-Level Summary

- Total number of AEs
- Total **serious AEs**
- Total **severe AEs**
- AE counts **by Preferred Term (AEDECOD)**
- AE counts **by Severity**

Deliverables:

- Summary tables
- Frequencies using PROC FREQ / PROC SQL

TASK 5: AE Incidence by Treatment Arm

Join DM and AE datasets.

For each **ARM** (Drug A / Drug B / Placebo), generate:

| ARM | Total Subjects | Subjects with ≥1 AE | % Subjects with AE |

Deliverables:

- Table generated using PROC SQL

TASK 6: Serious AE Listing

Create a listing of all subjects with **AESER = Yes**, sorted by:

1. USUBJID
2. AESTDTC

Include columns:

- USUBJID
- AEDECOD
- Severity
- Start/End dates

Deliverables:

- Serious AE Listing in PROC PRINT

TASK 7: Country-wise AE Distribution

Generate a summary of events by **Country** using DM+AE merge.

Deliverables:

- Table showing AE counts per country

TASK 8: Create Final Safety Report Dataset

Create a combined dataset including for each subject:

- USUBJID
- AGE
- SEX
- ARM
- Total AEs
- Serious AEs
- Severe AEs

Deliverables:

- Final ANALYSIS dataset named **AE_SUMMARY**

BONUS TASK

Create a **Top 5 Most Common AEs** table using SQL:

| Rank | AEDECOD | Count |

FINAL SUBMISSION PACKAGE SHOULD INCLUDE

Your submission should contain:

1. **SAS Program (.sas)** with all steps
2. **Output tables** for:
    - Data cleaning
    - Reconciliation
    - AE summaries
    - Arm-wise AE report
    - Serious AE listing
    - Country-wise summary
3. **AE_SUMMARY dataset**
4. PDF or PPT summarizing your findings