**Capstone Exercise: Applying Lessons in a Mini-Project**

This exercise applies everything from the last three lessons. Review your learning and follow the same workflow.

**A. Data and Files**

This project uses real clinical trial data from OUCRU. The data includes:

1. **Main Dataset**:
   * File: *"2-10-2020-\_03TS\_V1\_Data.xls"*
   * Contains most information. Variable explanations are in the statistical plan file (*"03TS analysis plan V1.5 November 2020 Accept changes.docx"*).
2. **Treatment Allocation**:
   * File: *"03TS\_Randlist.csv"*
   * Includes treatment arm allocations and a "dictionary" sheet explaining the four arms.
3. **Violations, Exclusions, Withdrawals**:
   * File: *"Protocol violations, exclusions, withdrawals.xlsx"*
4. **Statistical Plan**:
   * File: *"03TS analysis plan V1.5 November 2020 Accept changes.docx"*
   * Provides guidance on specific variables.

**B. Objectives**

**I. Data Preparation**

Prepare data for the following populations:

* **IT-ITT**, **IM\_ITT**, **IT\_PP**, **IM\_PP**, **IM\_ALL**
* Use *"03TS\_Randlist.csv"* for allocation of the first 272 participants.
* Use *"Protocol violations, exclusions, withdrawals.xlsx"* for filtering data per population.

**II. Table Creation**

1. **Baseline Characteristics**:
   * Summarize as follows:
     + **Numeric Data**: Report median, 1st/3rd quartiles, lowest, and highest values.
     + **Categorical Data**: Report count (n) and percentage (%).
   * No formal statistical comparisons between study arms.
   * Characteristics to summarize:  
     a. **Patient Details**:
     + Sex, age, BMI (*ENR.AGE, ENR.SEX, BMI = ADM.WEIGHT / (ADM.HEIGHT)^2*).  
       b. **Past Medical History**:
     + Variables like ADM.HYPERTENSION, ADM.MYOCARDIALINFART, ADM.SEVERELIVER, etc.  
       c. **Patient History**:
     + Duration of illness, incubation period, respiratory rate, platelet count, etc.  
       d. **Specific Severity Scores**:
     + Tetanus Severity, SOFA, APACHE II (calculated from variables).
2. **Adverse Events (AE)**:
   * Summarize proportions of individuals with AEs, categorized by grade (I-IV) and relation to treatment.
   * Exclude specific events (e.g., "Nasogastric tube", "Tracheostomy").
   * Use chi-square or Fisher’s exact test for comparisons.
   * Severe adverse events are summarized separately for each population.

**III. Plots**

1. Create histograms for:
   * **Pipecuronium**:
     1. Total dose during hospital stay (ventilated patients).
     2. Duration of use (ventilated patients).
   * **Diazepam**:
     1. Total dose during hospital stay.
   * **Midazolam**:
     1. Total dose during hospital stay.
   * **Benzodiazepines**:
     1. Total dose as diazepam equivalent.
     2. Total duration of use.
2. Create line plot for:

Create a line plot of the daily maximum temperature for each patient during the first 7 days after admission. Use one colour for each treatment arm and make the mean temperature for each day bold for each patient.

**Important:** refer to the statistical plan for further information. You might need to view the sample output in sample folder.