

## On the given dataset perform the following task:

1. Read Dataset and Explore the dataset by checking shape, columns, see the first/last 'n' rows using head/tail. (n= 5,15,30) (6)
2. Extract and document key insights by exploring column data types and applying info and describe and write down your insights inside the markdown.
3. Loc and iloc
  - Select all columns for the first clinical trial in the dataset.
  - Retrieve the Title and Status of the clinical trial with the NCT Number 'NCT04595136'.
  - Get the Sponsor/Collaborators and Start Date for clinical trials that are Recruiting.
  - Select the first 5 rows and columns Title, Conditions, and Outcome Measures.
  - Find the Completion Date and URL for the last 3 clinical trials in the dataset.
4. Determine the missing values in the whole dataset and analyze missing values in each column. (10)
5. Calculate the sum of duplicate rows (5)
6. Solve following question by using conditional statements (10)
  - How many studies have an enrollment greater than a certain threshold?
  - How many clinical trials have 'No Results Available'?
  - How many clinical trials are in an "Completed" and "Recruiting" status?
  - How many clinical trials are related to 'COVID-19'?
  - How many clinical trials started after January 1, 2020?