# Power BI Project Documentation: Clinical Trials Analysis

## **Report**

### 1. Project Overview

This Power BI project aims to analyze clinical trials conducted in India using a dataset of 5,000 records. The dashboard provides an interactive platform to explore trends in clinical trial phases, study types, sponsors, and geographical distribution. It helps users understand patterns in research activity, highlight top-performing sponsors, and identify potential gaps in trial coverage across therapeutic areas and regions.

### 2. Project Objectives

* To visualize the distribution of clinical trials by phase (Phase I–IV).
* To analyze sponsor performance by trial count and status.
* To understand regional and therapeutic area trends in Indian clinical research.
* To provide an interactive dashboard for decision-making and exploration.
* To derive insights into growth and challenges in India’s clinical research ecosystem.

### 3. Dataset Description

The dataset (clinical\_trial\_india\_5000.xlsx) contains detailed information about clinical trials conducted in India. It includes attributes such as trial ID, title, sponsor, study type, phase, status, therapeutic area, start and completion dates, and locations.

### 4. Data Cleaning (Using Power Query)

Data cleaning was performed in Power Query to ensure consistency and accuracy. The following transformations were applied:

* Removed duplicates based on Trial ID.
* Replaced null or blank values in categorical fields with “Not Specified.”
* Converted date fields (Start Date, Completion Date) to proper Date format.
* Trimmed text and standardized case for sponsor and phase names.
* Created calculated columns such as Trial Duration (Completion Date – Start Date).
* Filtered out trials outside the desired date range or missing key identifiers.

### 5. Data Modeling

The Power BI data model was designed to support flexible analysis of trial attributes and performance metrics. A star schema was implemented where the Trials table acts as the fact table, connected to dimension tables such as Sponsors, Locations, Phases, and Therapeutic Areas.

### 6. Dashboard and Visualization Summary

* Bar Chart: Number of trials by phase.
* Map Visualization: Geographic distribution by state.
* Donut Chart: Study type distribution.
* Table: Top 10 sponsors by total trials.
* Line Chart: Trend of trials initiated per year.