**Data Dictionary**

|  |  |  |  |
| --- | --- | --- | --- |
| **Column Name** | **Type** | **Calculated** | **Description** |
| NCT ID | object | n | The unique identification code given to each clinical trial upon registration at ClinicalTrials.gov. |
| Trial Status | object | n | The status of the clinical trial: Completed, not completed, or in-progress. |
| Phase | object | n | The stage of a clinical trial studying a drug or biological product, based on definitions developed by the U.S. Food and Drug Administration (FDA). |
| Start Date | datetime | n | The actual date on which the first participant was enrolled in a clinical study. |
| Completion Date | datetime | n | The date on which the last participant in a clinical trial was examined or received a treatment. |
| Trial Duration (Days) | integer | y | The duration between start and completion date of a clinical trial, expressed in days. |
| Trial Location | object | n | The country where the clinical trial(s) are being held. |
| Sponsor Name | object | n | The name of the company or organization sponsoring the trial. |
| Sponsor Type | object | n | The type of sponsor sponsoring the trial: Academic, industry, network, healthcare, cancer centers, etc. |
| Treated Condition | object | n | The name(s) of the disease(s) or condition(s) studied in the clinical trial, or the focus of the clinical study. |
| No of Participants | float | n | The number of participants in a clinical trial. |
| Healthy Participants | object | n | Indication that healthy participants are permitted to participate in the clinical trial: True or False. |
| Gender | object | n | The gender of the participants eligible to participate in the clinical study. |
| Clinical Trial Drug | object | n | The drug used on participants of a clinical trial to test the safety and efficacy of the drug. |
| First Approval Date | float | n | The date of first approval of the drug being tested. |
| Approval Status | object | n | The approval status of the drug being tested: Approved, not approved, or investigational. |
| ATC Name | object | n | The names given to the Anatomical Therapeutic Chemical (ATC) system, used to categorize drugs based on their target organ or system. |
| ATC Class | object | n | The codes given to the Anatomical Therapeutic Chemical (ATC) system, used to categorize drugs based on their target organ or system. |
| Therapeutic Area | object | y | The category formed from grouping the list of treated conditions into broader categories. |
| Drug-Therapeutic Area | object | y | Category formed from joining drug name to therapeutic area. |
| No. of Trial Locations | integer | y | The total number of unique countries where the clinical trials are conducted. |
| Sponsor\_Prior Completed Trials | integer | y | The sponsor previous track record for sponsoring completed clinical trials. |
| Sponsor\_Prior Not Completed Trials | integer | y | The sponsor previous track record for sponsoring not completed clinical trials. |
| Total COMPLETED PHASE1 Trials per Drug | float | y | The total number of phase 1 clinical trials with the ‘completed’ trial status that tested the same drug. |
| Total IN-PROGRESS PHASE1 Trials per Drug | float | y | The total number of phase 1 clinical trials with the ‘in-progress’ trial status that tested the same drug. |
| Total NOT COMPLETED PHASE1 Trials per Drug | float | y | The total number of phase 1 clinical trials with the ‘not completed’ trial status that tested the same drug. |
| Total COMPLETED PHASE1, PHASE2 Trials per Drug | float | y | The total number of phase 1 and 2 combined clinical trials with the ‘completed’ trial status that tested the same drug. |
| Total IN-PROGRESS PHASE1, PHASE2 Trials per Drug | float | y | The total number of phase 1 and 2 combined clinical trials with the ‘in-progress’ trial status that tested the same drug. |
| Total NOT COMPLETED PHASE1, PHASE2 Trials per Drug | float | y | The total number of phase 1 and 2 combined clinical trials with the ‘not completed’ trial status that tested the same drug. |
| Total COMPLETED PHASE2 Trials per Drug | float | y | The total number of phase 2 clinical trials with the ‘completed’ trial status that tested the same drug. |
| Total IN-PROGRESS PHASE2 Trials per Drug | float | y | The total number of phase 2 clinical trials with the ‘in-progress’ trial status that tested the same drug. |
| Total NOT COMPLETED PHASE2 Trials per Drug | float | y | The total number of phase 2 clinical trials with the ‘not completed’ trial status that tested the same drug. |
| Total COMPLETED PHASE2, PHASE3 Trials per Drug | float | y | The total number of phase 2 and 3 combined clinical trials with the ‘completed’ trial status that tested the same drug. |
| Total IN-PROGRESS PHASE2, PHASE3 Trials per Drug | float | y | The total number of phase 2 and 3 combined clinical trials with the ‘in-progress’ trial status that tested the same drug. |
| Total NOT COMPLETED PHASE2, PHASE3 Trials per Drug | float | y | The total number of phase 2 and 3 combined clinical trials with the ‘not completed’ trial status that tested the same drug. |
| Total COMPLETED PHASE3 Trials per Drug | float | y | The total number of phase 3 clinical trials with the ‘completed’ trial status that tested the same drug. |
| Total IN-PROGRESS PHASE3 Trials per Drug | float | y | The total number of phase 3 clinical trials with the ‘in-progress’ trial status that tested the same drug. |
| Total NOT COMPLETED PHASE3 Trials per Drug | float | y | The total number of phase 3 clinical trials with the ‘not completed’ trial status that tested the same drug. |
| Total ACADEMIC Sponsors per Drug | float | y | The total number of academic sponsors for clinical trials that tested the same drug. |
| Total CANCER CENTERS Sponsors per Drug | float | y | The total number of cancer center sponsors for clinical trials that tested the same drug. |
| Total FED Sponsors per Drug | float | y | The total number of FED sponsors for clinical trials that tested the same drug. |
| Total HEALTHCARE INSTITUTIONS Sponsors per Drug | float | y | The total number of healthcare institution sponsors for clinical trials that tested the same drug. |
| Total INDUSTRY Sponsors per Drug | float | y | The total number of industry sponsors for clinical trials that tested the same drug. |
| Total NETWORK Sponsors per Drug | float | y | The total number of network sponsors for clinical trials that tested the same drug. |
| Total NIH Sponsors per Drug | float | y | The total number of U.S. National Institutes of Health (NIH) sponsors for clinical trials that tested the same drug. |
| Total OTHER Sponsors per Drug | float | y | The total number of miscellaneous sponsors for clinical trials that tested the same drug. |
| Total TOP 10 PHARMA Sponsors per Drug | float | y | The total number of Top 10 pharma companies which sponsored clinical trials that tested the same drug. |
| No. of Drug-Therapeutic Pairs per Drug | integer | y | The total number of unique drug-therapeutic area combinations that are being tested for the same drug. |