Data Intake Report

Name: ABC Pharmaceutical - DataMinders

Report date: 6/19/24

Internship Batch: LISUM33

Version: 1.0

Data intake by: Data Minders

Data intake reviewer: Data Minders

Data storage location: https://github.com/S0n0f1saac/LISUM33-GroupProject

**Tabular data details:**

|  |  |
| --- | --- |
| **Total number of observations** | 3424 |
| **Total number of files** | 1 |
| **Total number of features** | 69 |
| **Base format of the file** | .xlsx |
| **Size of the data** | 899 KB |

**Proposed Approach:**

* Identification of duplicates:
  + We will conduct exact matching for features where we expect duplicates and clustering to identify potential duplicates for all other features.
* Assumptions: (if you assume any other thing for data quality analysis)

**Note: Convert this doc in pdf and provide the link of pdf file in your dashboard.**

**Please do not forget to remove this section while converting the file into pdf.**