

Data Intake Report

Name: Healthcare - Persistency of a Drug

Report date: 11/17/2024

Internship Batch: LISUM38

Version: 1.0

Data intake by: Kirtoria Ward

Data intake reviewer: <>

Data storage location: https://github.com/KTW04/Data_Glacier_Internship_LISUM38

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	899KB

Proposed Approach:

1. Data Cleaning:
 - Identify duplicate rows using unique patient IDs (Ptid) and remove them.
2. Assumptions:
 - All demographic and clinical data are accurate and valid.
 - Missing values will be handled using appropriate imputation or exclusion methods.
 - Outliers will be addressed if they significantly impact the results.