

Data Intake Report

Name: Data Science:: Healthcare - Persistency of a drug:: Group Project

Report date: 1/1/2025

Internship Batch: LISUM39

Version: 1.0

Data intake by: Devin Chau, Rohan Khatri, Ethan Dy

Data intake reviewer: Data Glacier

Data storage location:

<https://github.com/mynameisdevinchau/Data-Glacier-Internship/tree/main/Week%209>

Tabular data details:

Total number of observations	3424
Total number of files	1
Total number of features	69
Base format of the file	.csv
Size of the data	891 KB

Proposed Approach:

- For our approach, we did not need to deduplicate our values as the features in the dataset did not have any unnecessary features and all the values were fortunately not null.
- Assumptions we made for this dataset are that the doctor who gave the prescription was accurate and correct for the medication prescribed.