

# Data Intake Report

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

Report date: 12/17/2024

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%207>

## Tabular data details:

<b>Total number of observations</b>	3424
<b>Total number of files</b>	1
<b>Total number of features</b>	69
<b>Base format of the file</b>	.csv
<b>Size of the data</b>	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.