

Data Intake Report

Name: Drug persistence and Medical Adherence

Report date: 16/07/2022

Internship Batch: LISUM10:30

Version:1.0

Data intake by: Wangu Ndungu

Data intake reviewer: Nikola Andrejić

Data storage location: <https://www.kaggle.com/datasets/harbhajansingh21/persistent-vs-nonpersistent>

Tabular data details:

Persistent vs NonPersistent

Total number of observations	3425
Total number of files	1
Total number of features	69
Base format of the file	.csv
Size of the data	889KB

Proposed Approach:

- Exploratory data analysis to discover outliers, skewness, trends and correlation among various variables in the dataset.
- Building a classification model that'll predict a patient's chances of being persistent or non-persistent with their prescribed drugs.