Data Intake Report

Name: Data Science: Health Care: Persistence of Drug
Report date: 12/18/22
Internship Batch: LISUM15
Version:1.0
Data intake by: Emily Yao
Data intake reviewer: N/A
Data storage location: <location URL eg: github, cloud>

Tabular data details:

Total number of observations	3,425
Total number of files	1
Total number of features	69
Base format of the file	.xlsx
Size of the data	899 KB

Note: Replicate same table with file name if you have more than one file.

Proposed Approach:
- Mention approach of dedup validation (identification)

I will first check for NAs, duplicates, and outliers. I also would then do exploratory data analysis on several factors but most importantly being the persistency flag.
- Mention your assumptions (if you assume any other thing for data quality analysis)

No current assumptions 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.