Data Intake Report

Name: <Health care- Persistancy of a drug>

Report date: <02/09/2021>

Internship Batch:<Enter your batch code from Canvas course>

Version:<1.0>

Data intake by:<Kelvin Mpofu>

Data intake reviewer:<intern who reviewed the report>

Data storage location: < https://drive.google.com/file/d/1P\_oMc6gOBlhw6dY5PxaqxV2swdHMUooK/view>

**Tabular data details:**

|  |  |
| --- | --- |
| **Total number of observations** | <69> |
| **Total number of files** | <2> |
| **Total number of features** | <69> |
| **Base format of the file** | <.xlsx> |
| **Size of the data** | < 898 kb > |

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

**Proposed Approach:**

* Mention approach of dedup validation (identification)
* Mention your assumptions (if you assume any other thing for data quality analysis)

I assume the data is messy.

**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.**