Data Intake Report

Name: G2M insight for Cab Investment Firm

Report date: <10/13/2022>

Internship Batch: LISUM14

Version:<1.0>

Data intake by: In Kyung Cho

Data intake reviewer: N/A

Data storage location: https://github.com/joinn125/Intern\_Week2

**Tabular data details:**

|  |  |
| --- | --- |
| **Total number of observations** | <number of rows> |
| **Total number of files** | <Number of files received> |
| **Total number of features** | <Number of columns> |
| **Base format of the file** | <.csv,.txt etc> |
| **Size of the data** | <size in GB,TB,PB,MB etc> |

**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)