Data Intake Report

Name: G2M Cab Investment

Report date: 6/14/2023

Internship Batch: LISUM22

Version:<1.0>

Data intake by: Christopher Perng

Data intake reviewer:<intern who reviewed the report>

Data storage location: https://github.com/Chrispy1204/data\_glacier\_cab\_investment

**Tabular data details: data\_1**

|  |  |
| --- | --- |
| **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> |

**Tabular data details: City**

|  |  |
| --- | --- |
| **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)

**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.**