Data Intake Report

Name: Week 06 Data Ingestion and Schema Validation

Report date: 2022/12/11

Internship Batch: LISUM15

Version: 1.0

Data intake by: Flores, Richard

Data intake reviewer: N/A

Data storage location: Kaggle URL: https://www.kaggle.com/datasets/najzeko/steam-reviews-2021

**Tabular data details:**

|  |  |
| --- | --- |
| **Total number of observations** | 21,700,000 |
| **Total number of files** | 01 |
| **Total number of features** | 23 |
| **Base format of the file** | CSV |
| **Size of the data** | 8.17 GB |

**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.**