

Data Intake Report

Name: File ingestion and schema validation

Report date: 12th March 2023

Internship Batch: LISUM18

Data intake by: Syed Ahmad Sohail

Data storage location: <https://www.kaggle.com/datasets/mkechinov/ecommerce-behavior-data-from-multi-category-store?resource=download>

Tabular data details:

Total number of observations	67501979
Total number of files	1
Total number of features	9
Base format of the file	.csv
Size of the data	5.67 GB

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

Proposed Approach:

- Validated Columns with any special character , white spaces from the col name