

# Data Intake Report

Name: week 2

Report date: 3/19/2022

Internship Batch: Enter your batch code from Canvas course

Version:1.0

Data intake by: stephen williams

Data intake reviewer:<intern who reviewed the report>

Data storage location: <https://github.com/Svalentinow/G2M>

## Tabular data details: final\_data.csv

<b>Total number of observations</b>	4,672,096
<b>Total number of files</b>	4
<b>Total number of features</b>	13
<b>Base format of the file</b>	.csv
<b>Size of the data</b>	37.2 mb

**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)
- Checked for null and duplicated rows in final\_data.csv. found zero
- Did analysis on profit, seasonality and distribution