

Data Intake Report

Name: G2M insight for Cab Investment firm

Report date: 02/18/2024

Version: 1.0

Data intake by: Yan-Ping Yu

Data intake reviewer:

Data storage location: <https://github.com/YanPing0227/portfolio>

Tabular data details:

Total number of observations	49171
Total number of files	4
Total number of features	5
Base format of the file	.csv
Size of the data	1.23 MB

Total number of observations	20
Total number of files	4
Total number of features	3
Base format of the file	.csv
Size of the data	1 KB

Total number of observations	440098
Total number of files	4
Total number of features	3
Base format of the file	.csv
Size of the data	8.58 MB

Total number of observations	359392
Total number of files	4
Total number of features	8
Base format of the file	.csv
Size of the data	23.3 MB

1. In cases where duplicates contain partial or complementary information, the records of 3 tabular datasets are merged into 1 single, comprehensive dataset.
2. Also Check whether duplicated rows exist by using .duplicated() function