Data Intake Report

Name: Week 6: File ingestion and schema validation
Report date: November 8th, 2024
Internship Batch: LISUM38
Version:1.0
Data intake by: Bilgan Kiris
Data intake reviewer: Data Glacier
Data storage location: https://github.com/bilgandc/DataGlacierInternship/tree/main/Week6

Tabular data details:

Total number of observations	2000
Total number of files	1
Total number of features	11
Base format of the file	.csv
Size of the data	0.42 MB

Proposed Approach:

Validate data by identifying and removing duplicates, ensuring schema alignment, and assuming optional fields can be blank without impacting data quality.