

# Data Intake Report

Name: Data Ingestion Pipeline

Report date: July 11, 2022

Internship Batch: LISUM 10

Version:<1.0>

Data intake by: Jeeyeon Shon

Data intake reviewer:

Data storage location: <https://www.kaggle.com/datasets/hm-land-registry/uk-housing-prices-paid>

## Tabular data details:

price\_paid\_records.csv

Total number of observations	22489348
Total number of files	1
Total number of features	11
Base format of the file	.csv
Size of the data	2.41 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.**