

N.B. File structure: for each dataset, the list of columns is reported as English translation, original name and suggestions for improving data quality

Itineraries Caserta

- Name - *Nome* - keep it as String. Extra spaces might be removed in the data cleansing phase
- Municipality - *Comune* - align it with official municipality names (e.g., <https://www.istat.it/storage/codici-unita-amministrative/Elenco-comuni-italiani.xls>)
- Location/Building type - *Località* - check value consistency
- Address - *Via* - split it on multiple columns reporting (zip code, street, municipality, country, etc.)
- Building category/type - *Tipologia* - check value consistency
- Building century - *Secolo di costruzione* - replace empty values with NULL
- First year certification - *Anno prima attestazione* - replace NON DISPONIBILE/empty values with NULL
- Itinerary period - *Itinerario (periodo)* - check value consistency
- Style - *Stile* - check value consistency + replace empty values with NULL
- Details - *Informazioni* - replace NON DISPONIBILE/empty values with NULL + clean values (e.g., numbers)
- Localization - *Geolocalizzazione* - consider the possibility to split it in longitude and latitude + replace NON DISPONIBILE/empty values with NULL + check suspicious values
- Image - *URL Immagine* - replace NON DISPONIBILE/empty values with NULL
- Bibliography - *Bibliography* - replace NON DISPONIBILE/empty values with NULL
- Site - *Sitografia* - replace NON DISPONIBILE/empty values with NULL
- Novel usage - *Rifunzionalizzato* - check value consistency + replace empty values with NULL
- Usage type - *Tipo di rifunzionalizzazione* - replace NESSUNO/empty values with NULL

Itineraries - Nocera Inferiore

- Municipality - Città - align it with official municipality names (e.g., <https://www.istat.it/storage/codici-unita-amministrative/Elenco-comuni-italiani.xls>)
- Location - Località - remove extra spaces, remove abbreviations, deal with empty values as NULL
- Address - Via - split it on multiple columns reporting (zip code, street, municipality, country, etc.)
- Site type - Sito/definizione manufatto - check value consistency
- Epoch of first certification - Epoca prima attestazione - deal with empty values
- Itinerary period - Itinerario (periodo) - check consistency
- Details - Informazioni - replace NON DISPONIBILE/empty values with NULL + clean values (e.g., numbers)
- Localization - *Geolocalizzazione* - consider the possibility to split it in longitude and latitude + replace NON DISPONIBILE/empty values with NULL + check suspicious values
- Image - *URL Immagine* - replace NON DISPONIBILE/empty values with NULL
- Bibliography - *Bibliography* - replace NON DISPONIBILE/empty values with NULL
- Site - *Sitografia* - replace NON DISPONIBILE/empty values with NULL

- Building type - Tipologia edificio - check value consistency
- Novel usage - *Rifunzionalizzato* - check value consistency + replace empty values with NULL
- Usage type - *Tipo di rifunzionalizzazione* - replace NESSUNO/empty values with NULL

Itineraries - Nocera

- Municipality - Città - align it with official municipality names (e.g., <https://www.istat.it/storage/codici-unita-amministrative/Elenco-comuni-italiani.xls>)
- Location - Località - remove extra spaces, remove abbreviations, deal with empty values as NULL
- Address - Via - split it on multiple columns reporting (zip code, street, municipality, country, etc.)
- Register number - N. Registro generale
- Company name - Nome azienda
- filing date - Data deposito - check date format consistency
- trademark registration date - Data registrazione marchio - check date format consistency
- filing location - Luogo di Deposito - align it with official municipality names
- trademark name - Nome marchio - deal with empty, NON INSERITO values
- production field - Settore produzione - check consistency
- import - Importazione Si/No - check value consistency
- Period itinerary - Itinerario (periodo) - check value consistency
- trademark description - Descrizione del marchio - deal with empty values or MARCHIO NON PRESENTE / DISPONIBILE as NULL
- company website - Sito azienda - deal with empty values
- note - Note - deal with empty values
- Localization - *Geolocalizzazione* - consider the possibility to split it in longitude and latitude + replace NON DISPONIBILE/empty values with NULL + check suspicious values
- Bibliography - *Bibliography* - replace NON DISPONIBILE/empty values with NULL
- Site - *Sitografia* - replace NON DISPONIBILE/empty values with NULL
- Image - *URL Immagine (2)*- deal with multiple values + replace NON DISPONIBILE/empty values with NULL
- Building type - Tipologia edificio - check value consistency + empty values
- Novel usage - *Rifunzionalizzato* - check value consistency + replace empty values with NULL
- Usage type - *Tipo di rifunzionalizzazione* - replace NESSUNO/empty values with NULL

Itineraries - Costiera Amalfitana

- Name - NOME - keep it as String. Extra spaces might be removed in the data cleansing phase
- Location - LOCALITA' - align it with official municipality names (e.g., <https://www.istat.it/storage/codici-unita-amministrative/Elenco-comuni-italiani.xls>) + deal with multiple values
- Type - TIPOLOGIA - check value consistency
- Localization - GEOLOCALIZATION - consider the possibility to split it in longitude and

latitude + replace NON DISPONIBILE/empty values with NULL + check suspicious values

- Description - DESCRIZIONE - replace NON DISPONIBILE/empty values with NULL + clean values (e.g., numbers)
- Note - NOTE - replace NON DISPONIBILE/empty values with NULL + clean values (e.g., numbers)
- Image - URL IMMAGINE - replace NON DISPONIBILE/empty values with NULL