

| | | | | | | | | | | |
|---|------------------------|---|------------------------|--|------------------|----------------------------|------------------------------|------------------------|--|----------------------|
| Overview: | Merging | | | | | | | | | |
| Connection Info: | Index | Connection String | Credentials | | | | | | | |
| Source: | 1 | ../Resources/IMDB_data.csv | | | Target | local postgres/customer_db | Login:postgres, PW: postgres | | | |
| | 2 | ../Resources/Netflix_data.csv | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| Source index | Source Table | Source Column | Source Datatype | | Target DB | Target Table | Target Column | Target Datatype | | Business Rule |
| 1 | imdb_data_clean.csv | id | integer | | project_2_db | title | | integer | | 1 |
| 1 | imdb_data_clean.csv | series_title | string | | project_2_db | title | series_title | string | | 1 |
| 1 | imdb_data_clean.csv | runtime_of_series | float | | project_2_db | title | | string | | 5 |
| 1 | imdb_data_clean.csv | genre | string | | project_2_db | title | genre | timestamp | | 5 |
| 1 | imdb_data_clean.csv | imdb_rating | float | | project_2_db | title | imdb_rating | date | | 3 |
| | imdb_data_clean.csv | no_of_votes | integer | | project_2_db | title | | time | | 3 |
| | | | | | | | | | | |
| 2 | netflix_data_clean.csv | id | integer | | project_2_db | title | | | | |
| 2 | netflix_data_clean.csv | type | string | | project_2_db | title | | | | |
| 2 | netflix_data_clean.csv | title | string | | project_2_db | title | title | | | |
| 2 | netflix_data_clean.csv | release_year | integer | | project_2_db | title | | | | |
| 2 | netflix_data_clean.csv | director | string | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | cast | string | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | country | string | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | date_added | timestamp | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | rating | integer | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | year_added | integer | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | month_added | integer | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | show_id | string | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | duration | integer | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | listed_in | string | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | description | string | | project_2_db | | | | | 5 |
| 2 | netflix_data_clean.csv | season_count | integer | | project_2_db | title | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| Table Rules | | | | | | | | | | |
| Drop all null values | | | | | | | | | | |
| Drop unnamed columns | | | | | | | | | | |
| | | | | | | | | | | |
| Drop duplicate values from source indexes | | | | | | | | | | |
| Convert all timestamps to yyyy-mm-ddThh:mm:ss | | | | | | | | | | |
| Join source 1 to source 2 on Title | | | | | | | | | | |
| | | | | | | | | | | |
| Business Rules/ transformation | | | | | | | | | | |
| 1 | | passthrough | | | | | | | | |
| 2 | | Title case | | | | | | | | |
| 3 | | Add 2 additional columns splitting timestamp into a date column | | | | | | | | |
| 4 | | Split street address from source column | | | | | | | | |
| 5 | | Drop Column | | | | | | | | |