

IMDB_FINAL

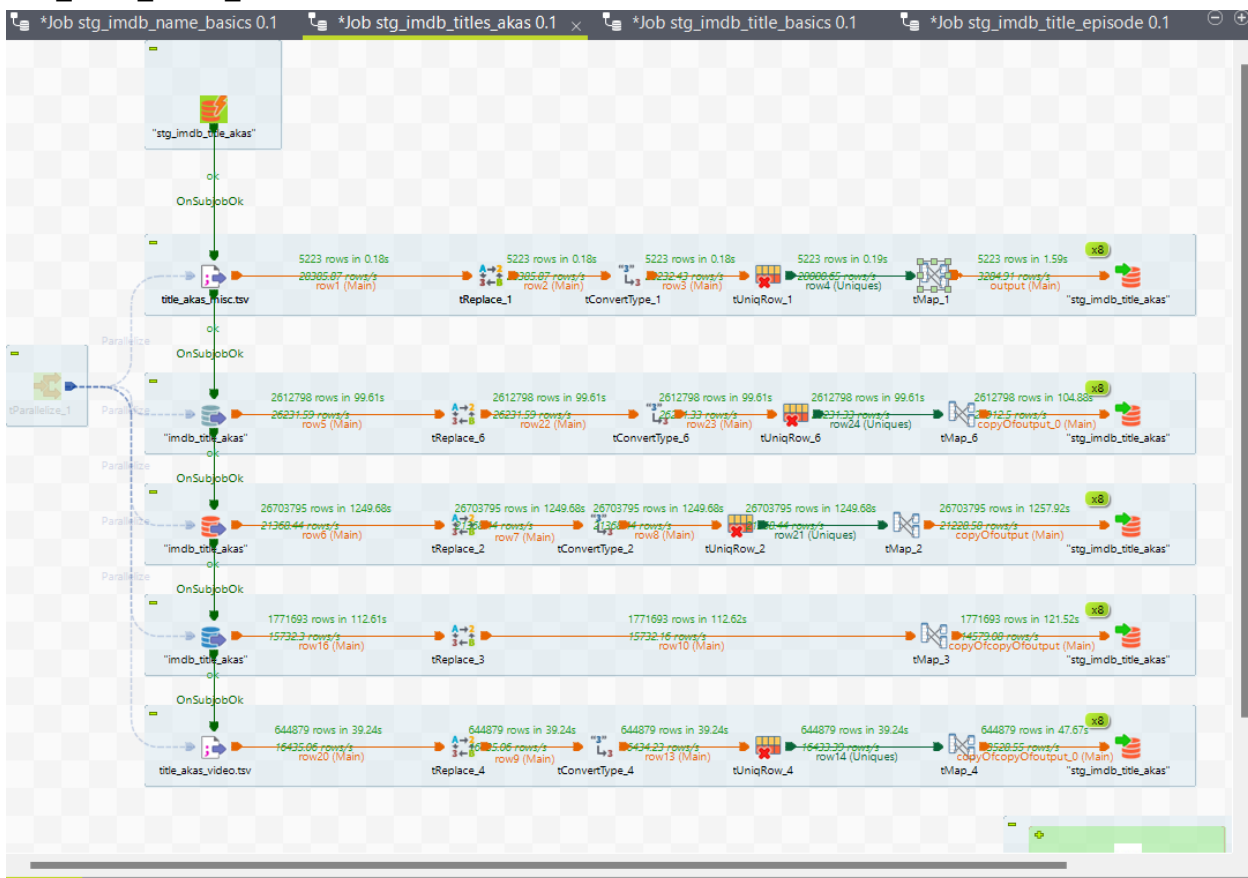
Group 19

1. STAGING

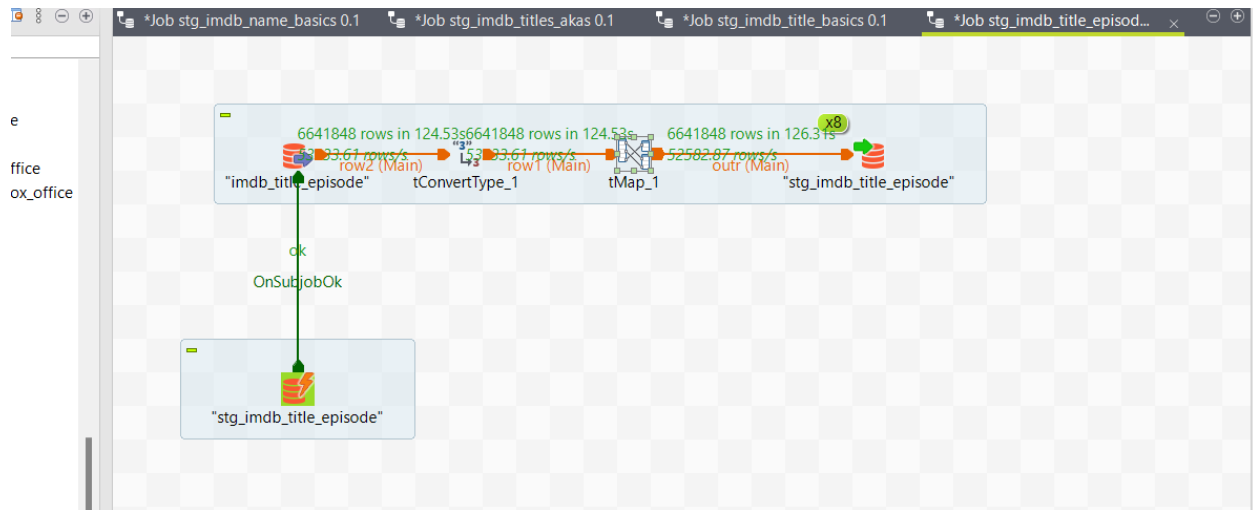
STG_IMDB_NAME_BASICS



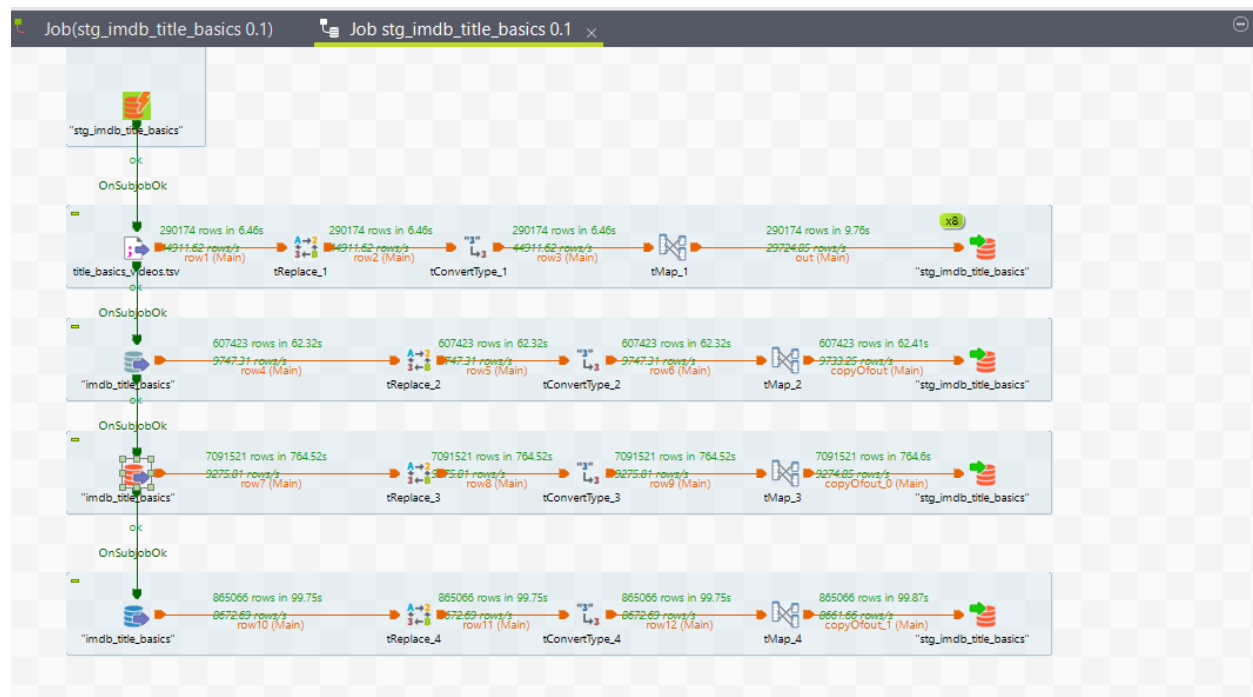
STG_IMDB_TITLE_AKAS



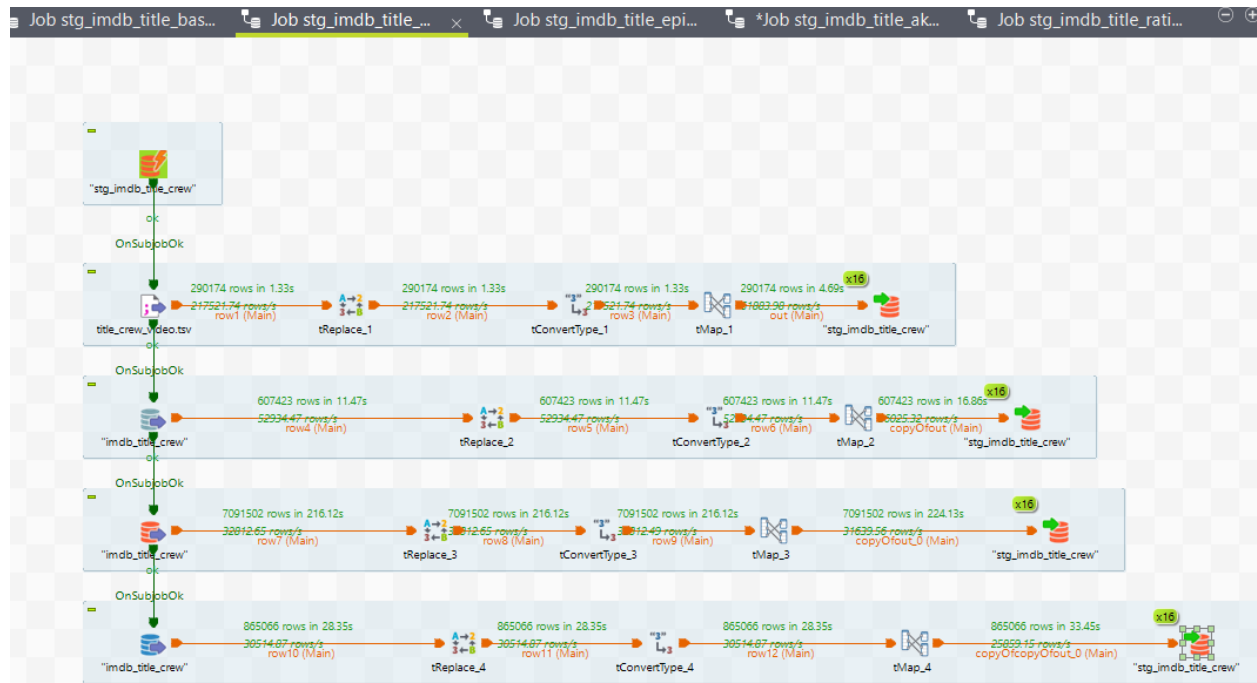
STG_IMDB_TITLE_EPISODE



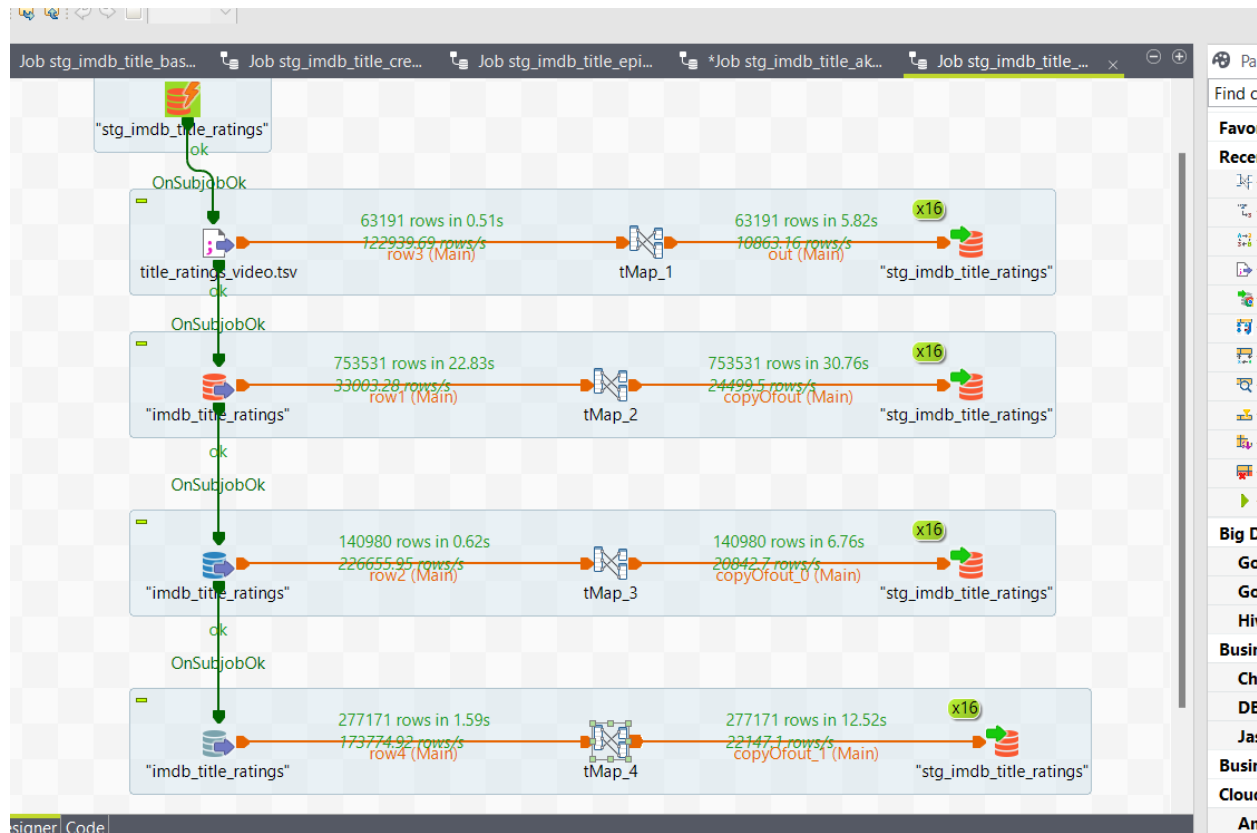
STG_IMDB_TITLE BASICS



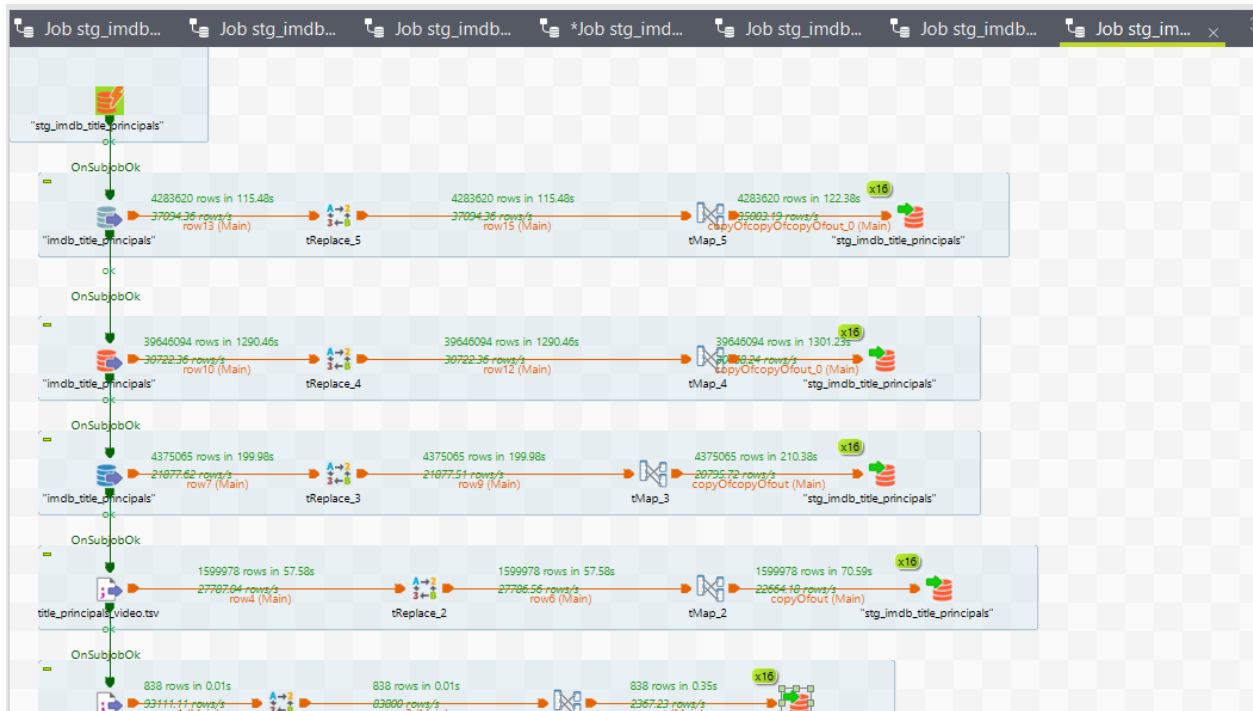
STG_IMDB_TITLE_CREW



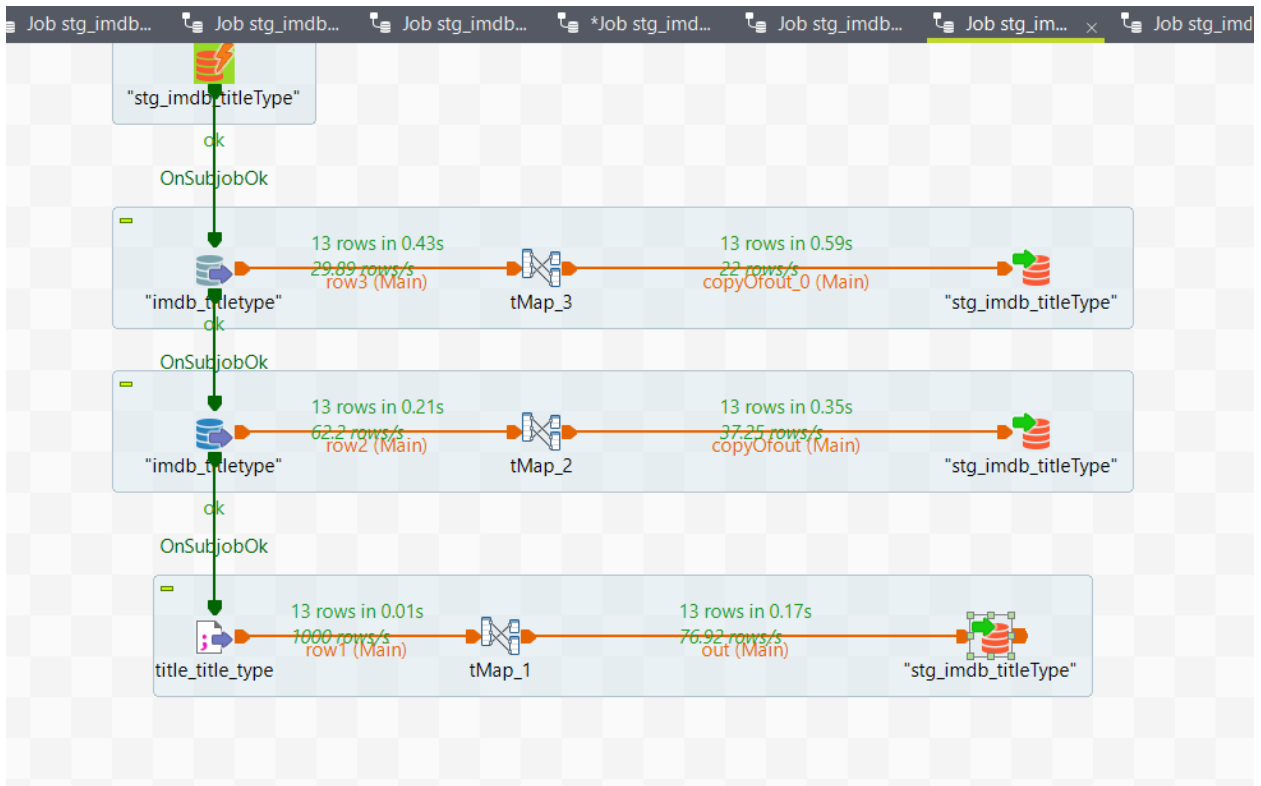
STG_IMDB_TITLE_RATINGS



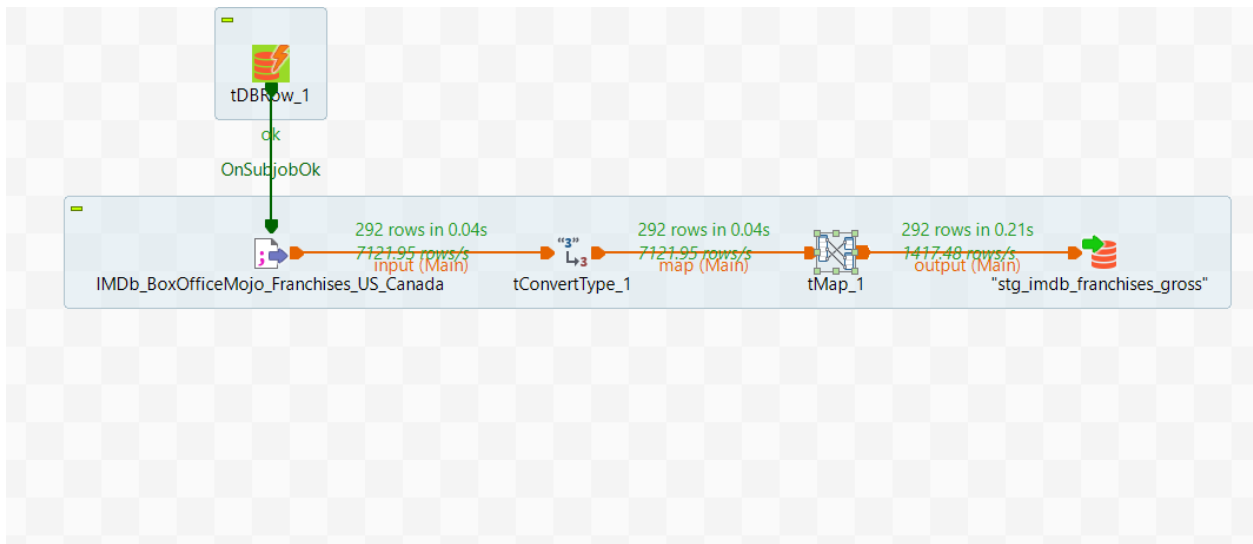
STG_IMDB_TITLE_PRINCIPLE



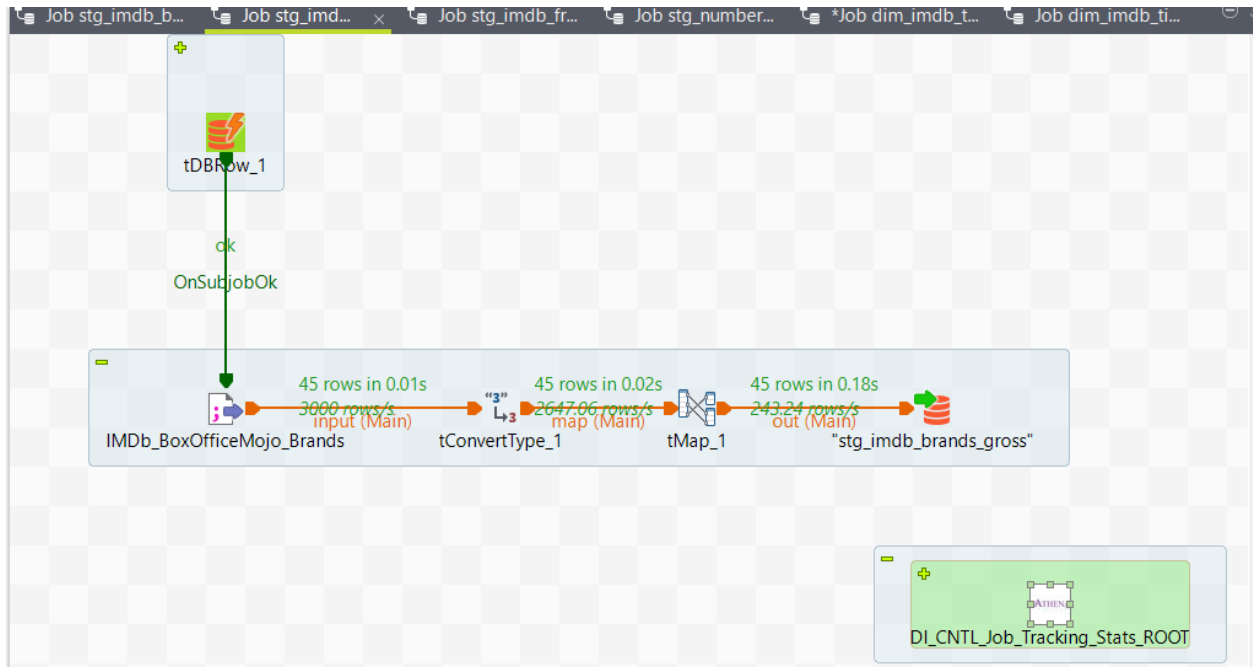
STG_IMDB_TITLE_TYPE



STG_FRANCHISE_GROSS

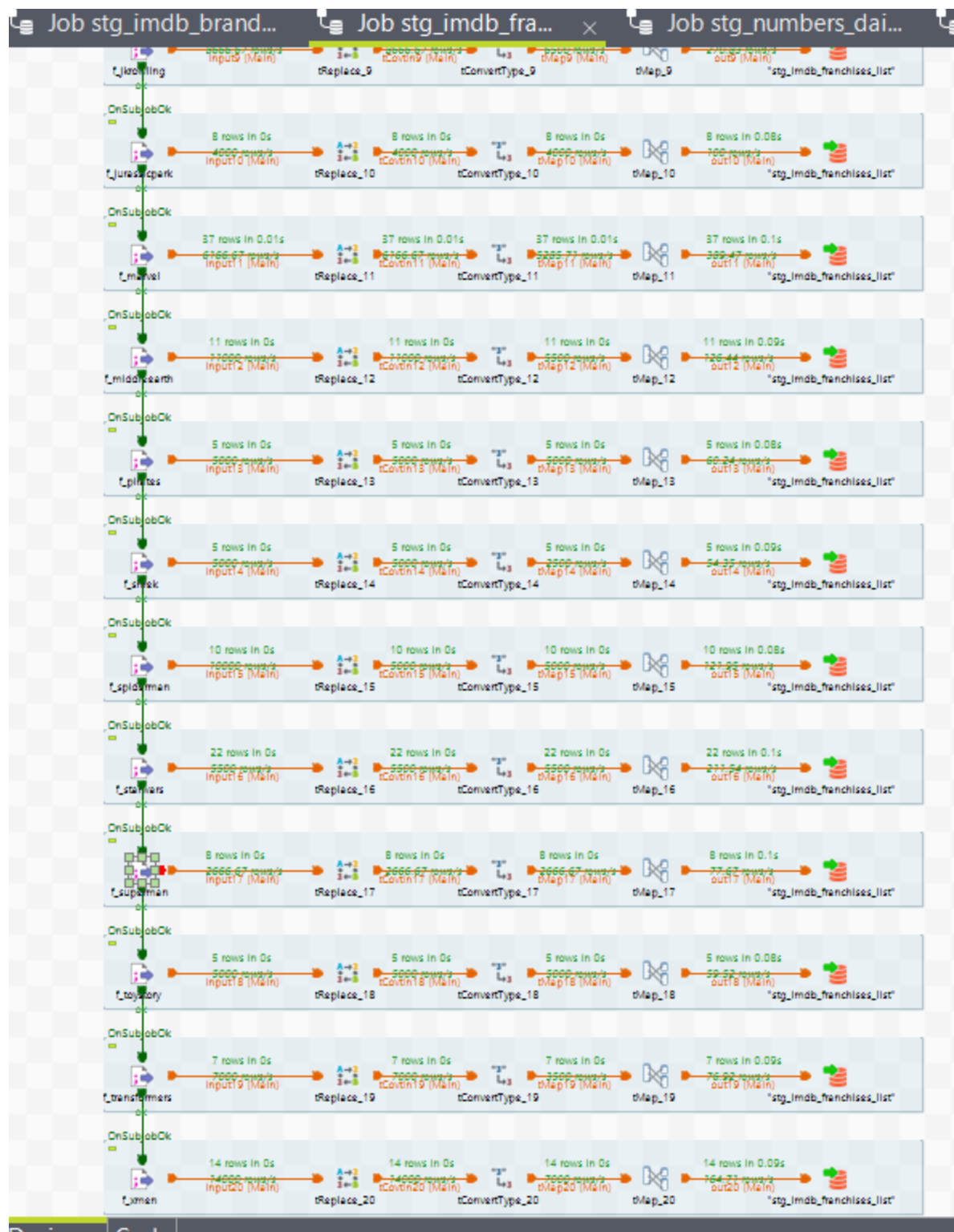


STG_BRAND_GROSS



STG_FRANCHISE_LISTS

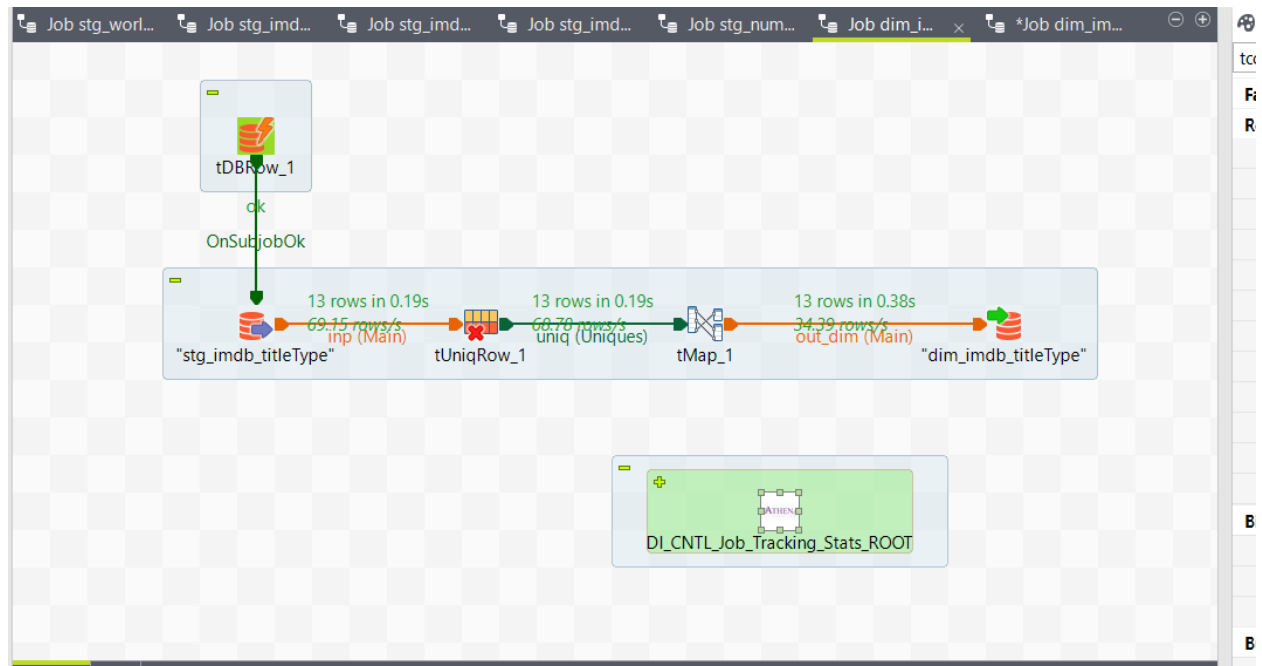




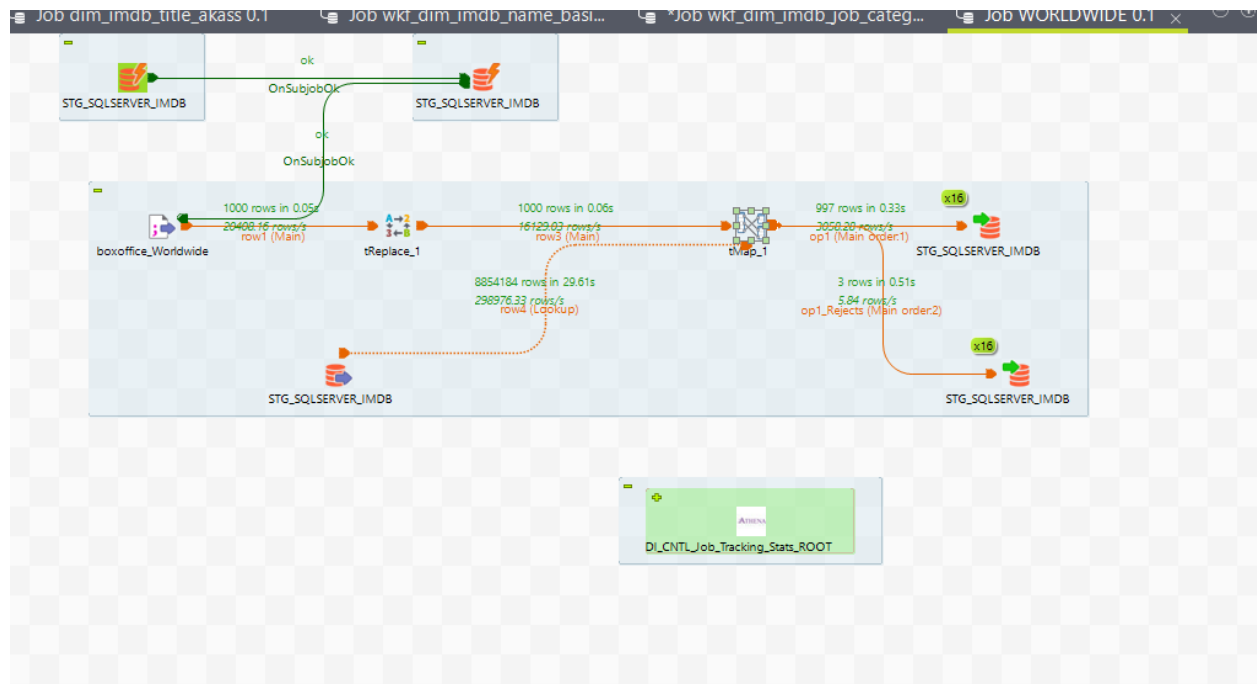
STG_DAILY_NUMBERS



STG_IMDB_TITLE_TYPE



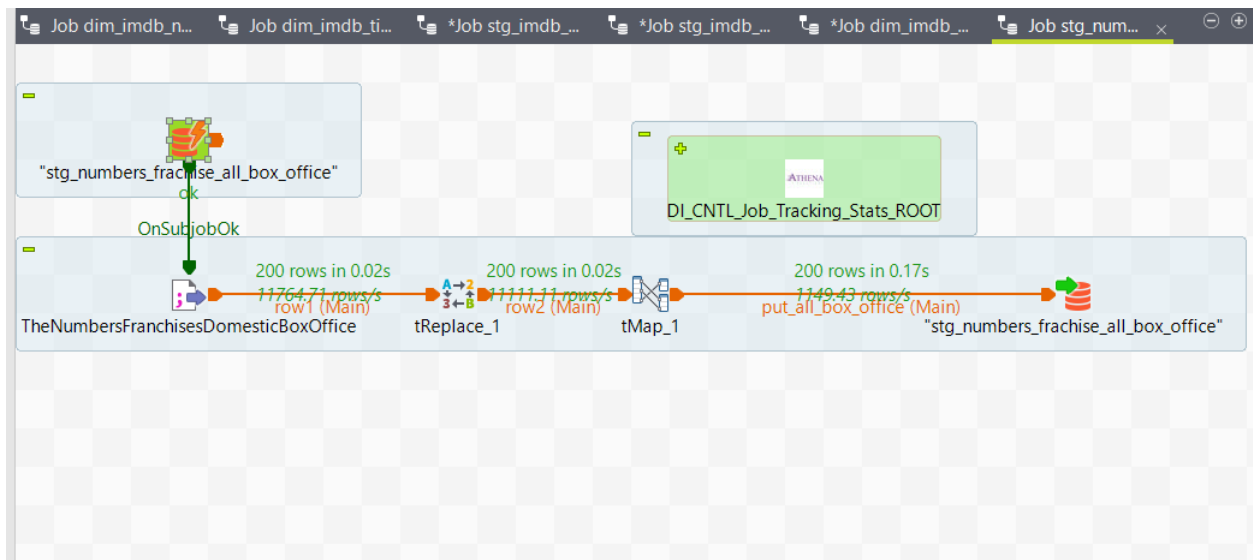
STG_WORLDWIDE_BOX_OFFICE



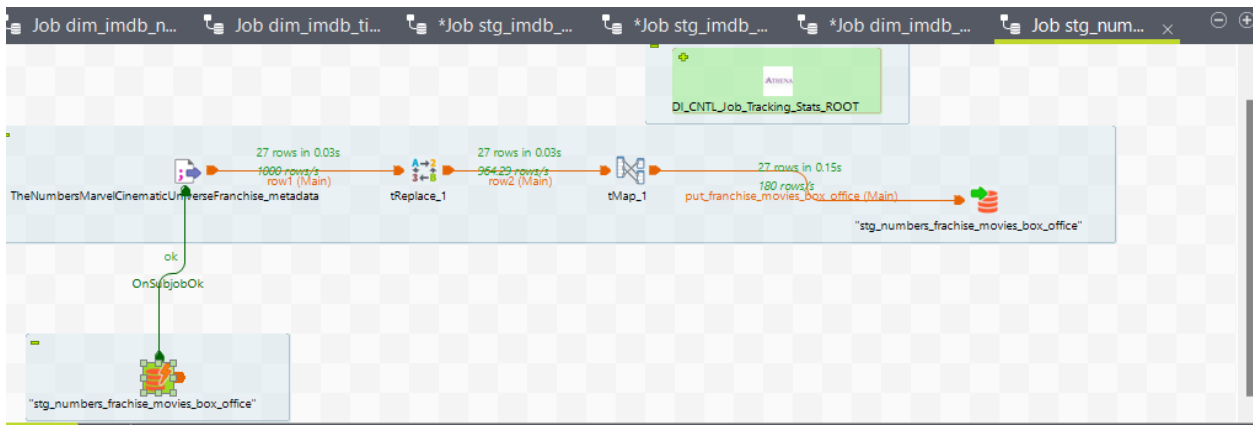
STG_NUMBERS_BOX_OFFICE



STG_FRANCHISE_ALL_BOX_OFFICE

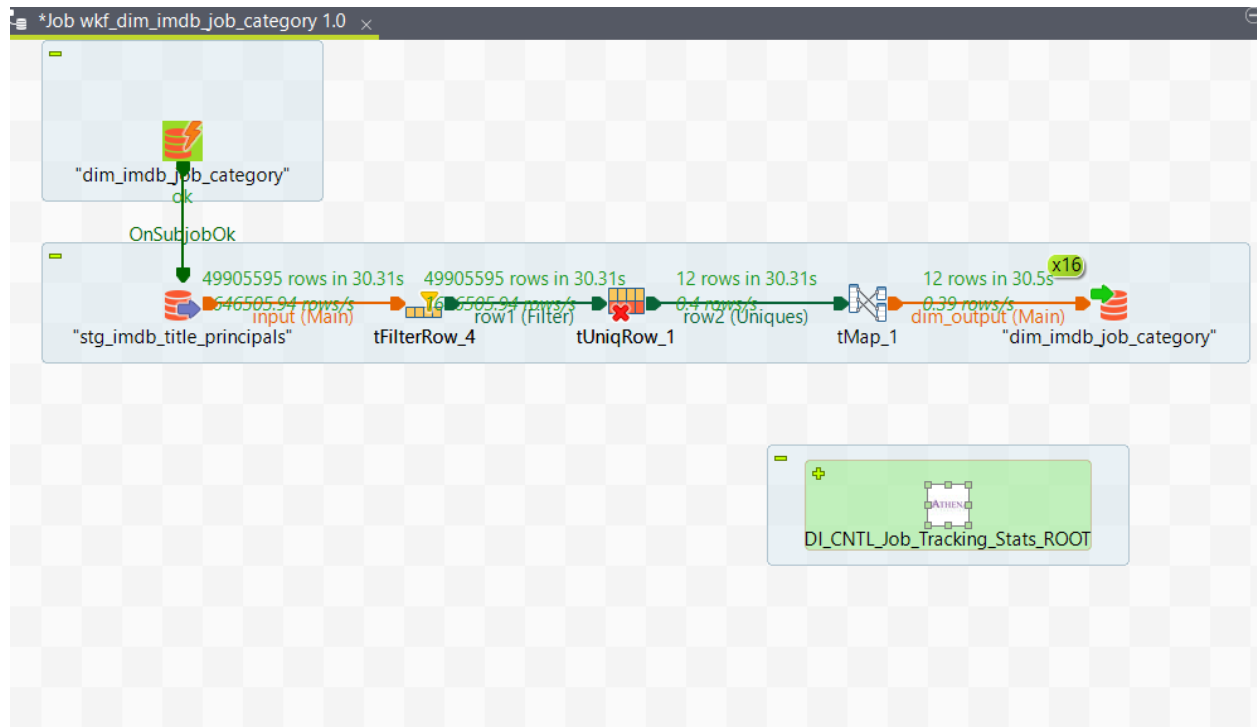


STG_NUMBERS_FRANCHISE_MOVIES_BOX_OFFICE

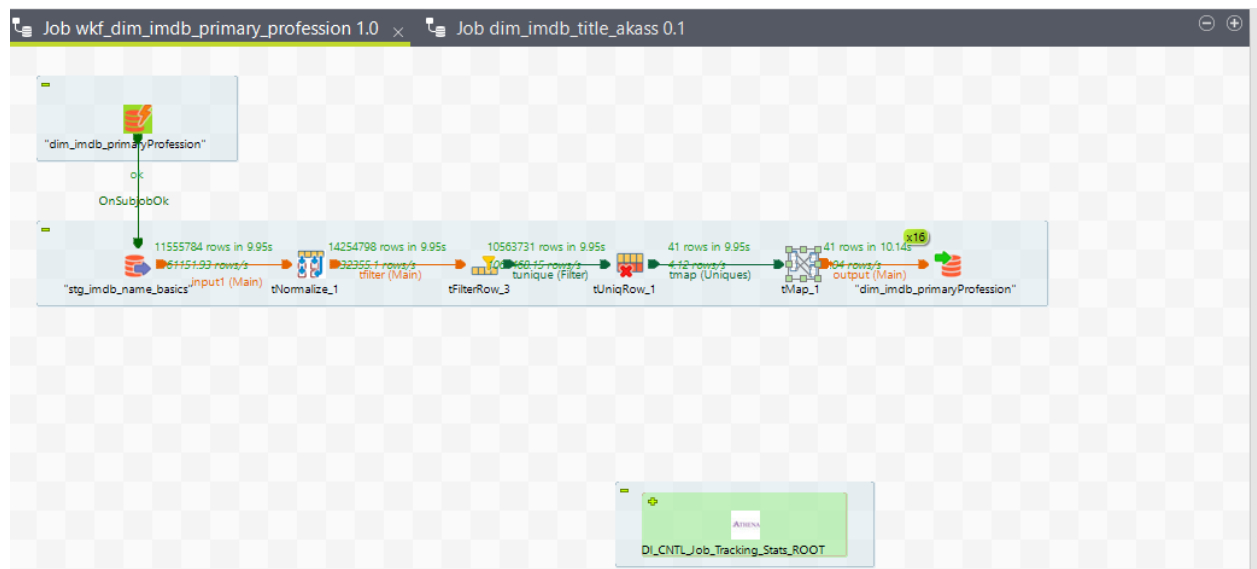


2. DIMENSION

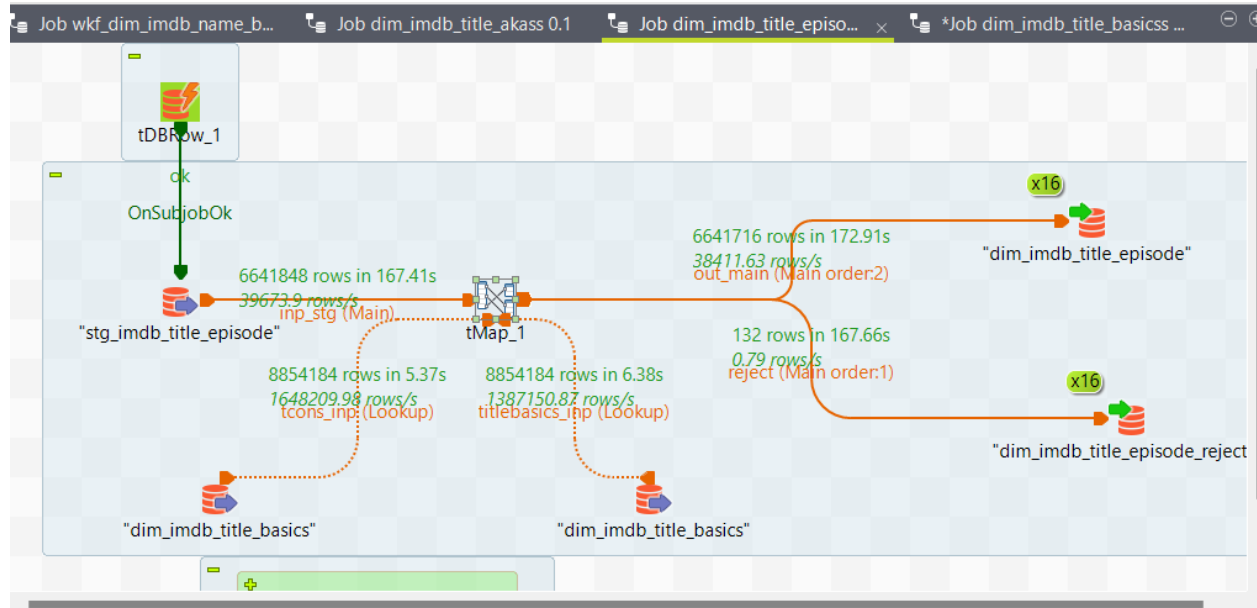
DIM_IMBD_JOB_CATEGORY



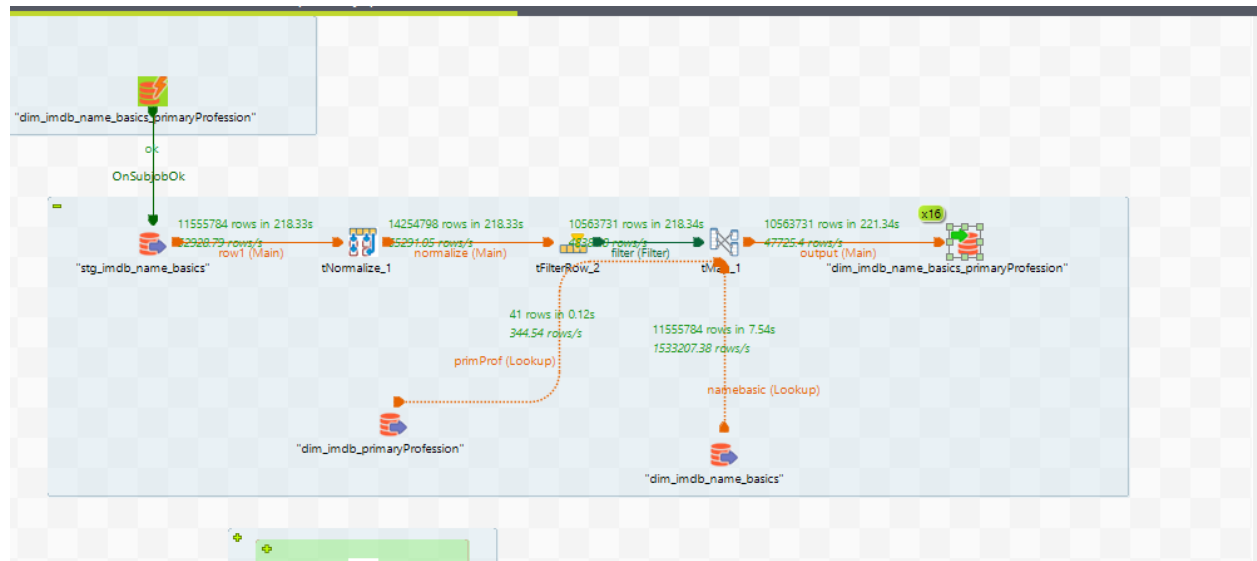
DIM_IMBD_PRIMARY_PROFESSION



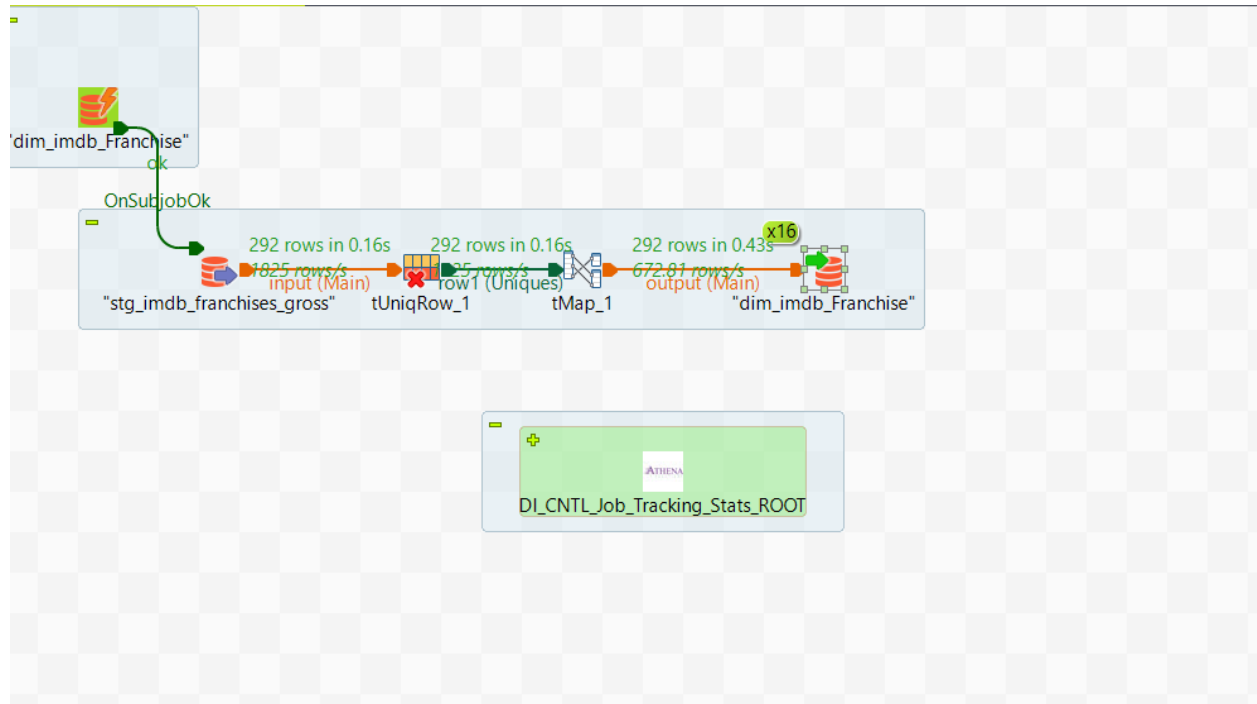
DIM_IMDB_TITLE_EPISODE



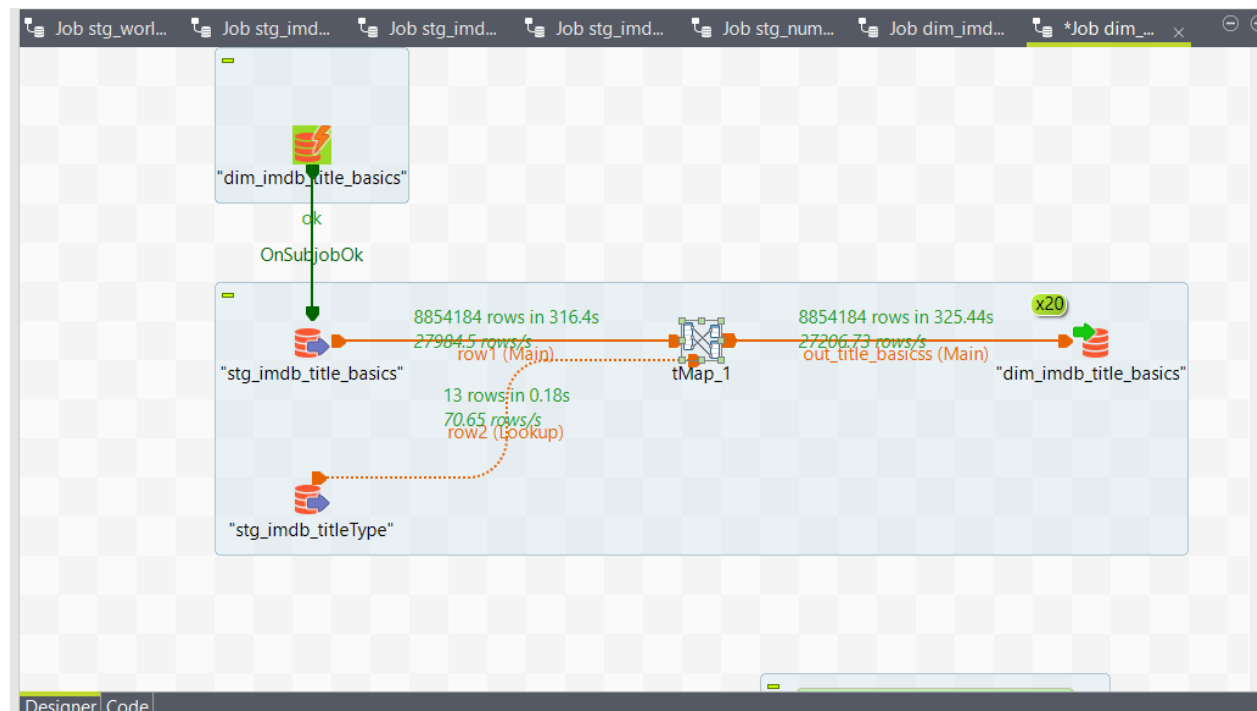
DIM_IMDB_NAME_BASICS_PRIMARY_PROFESSION



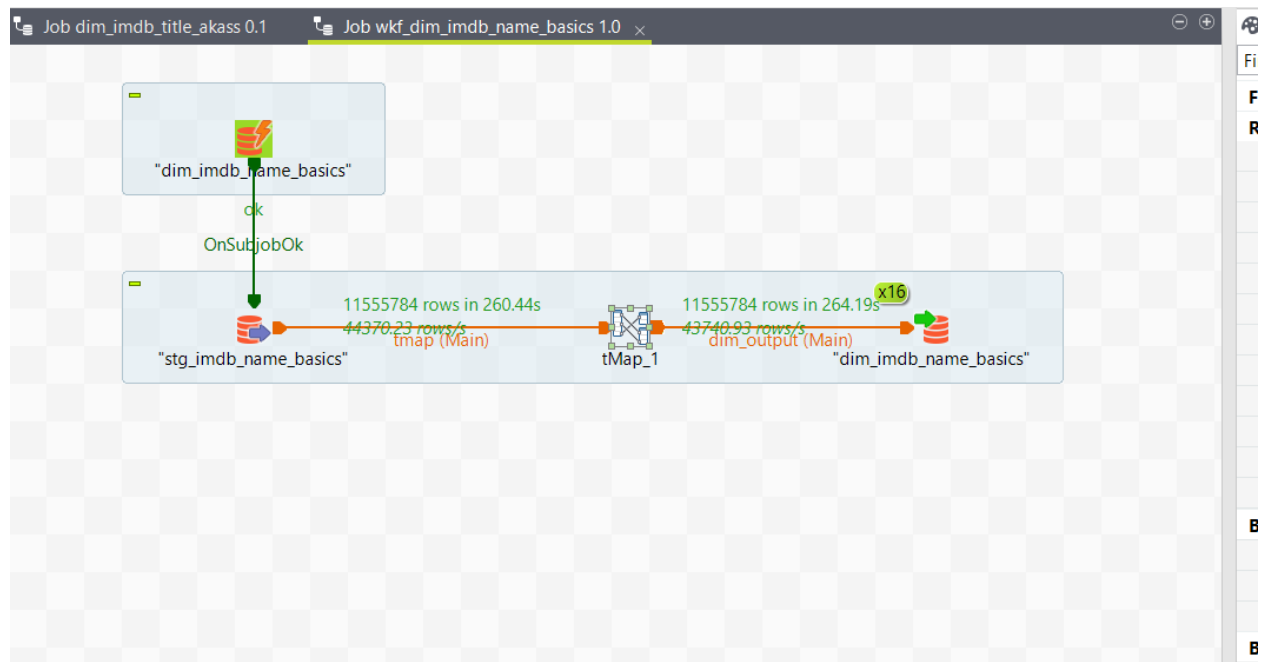
DIM_IMDB_FRANCHISE



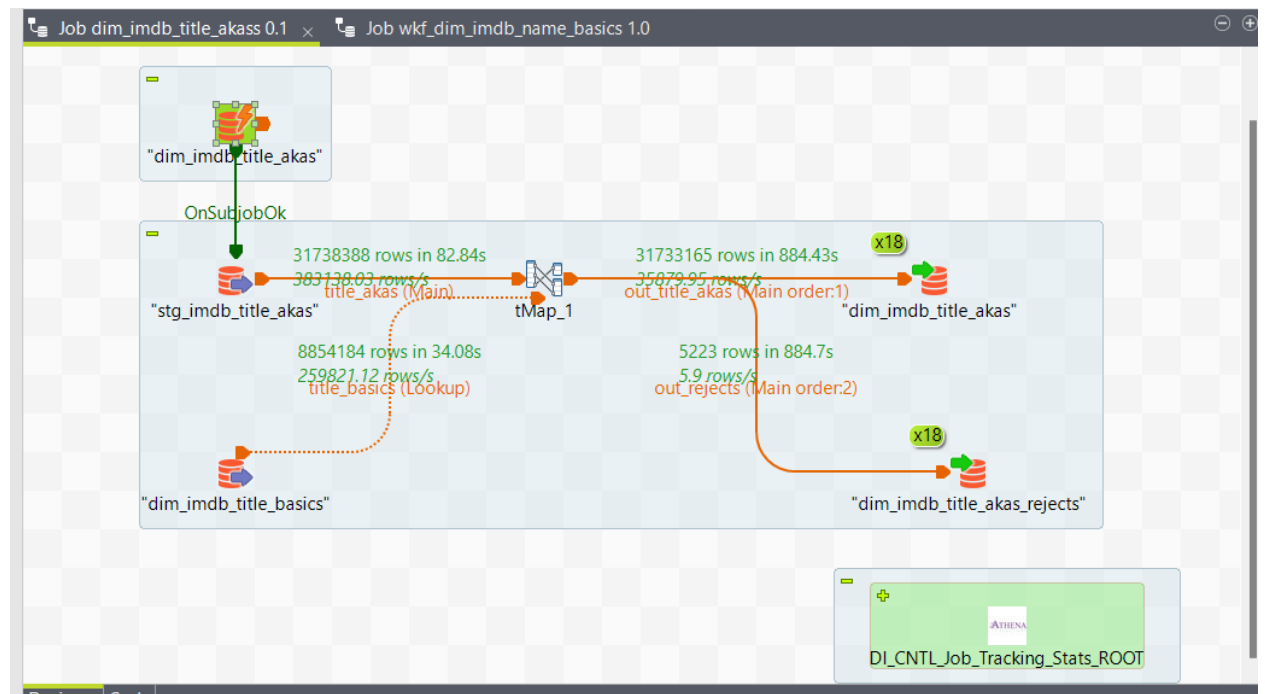
DIM_IMDB_TITLE_BASICS



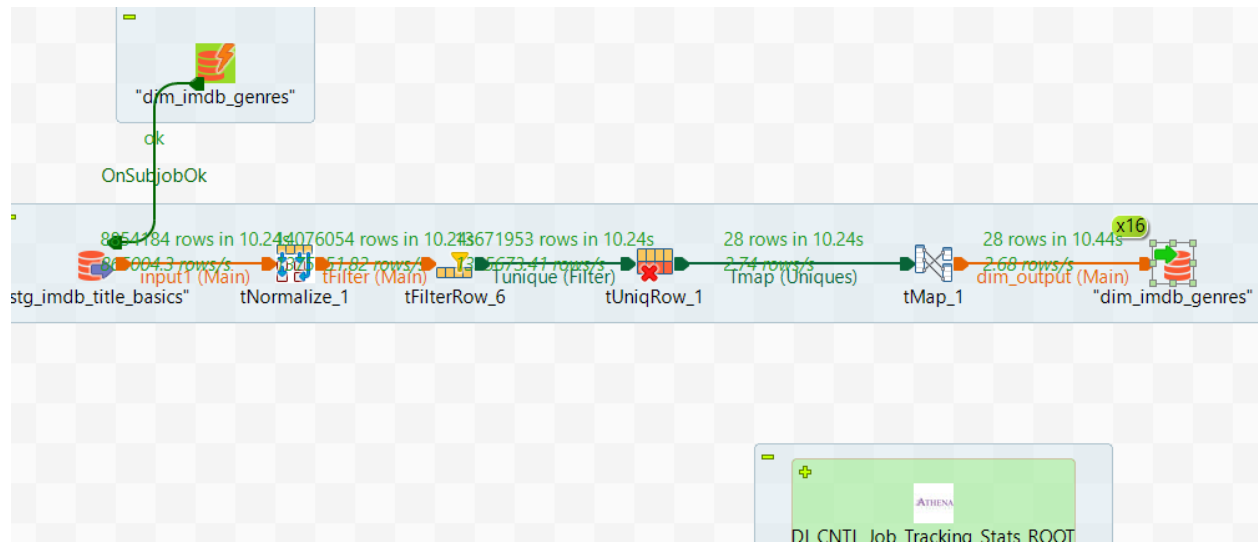
DIM_IMBD_NAME_BASICS



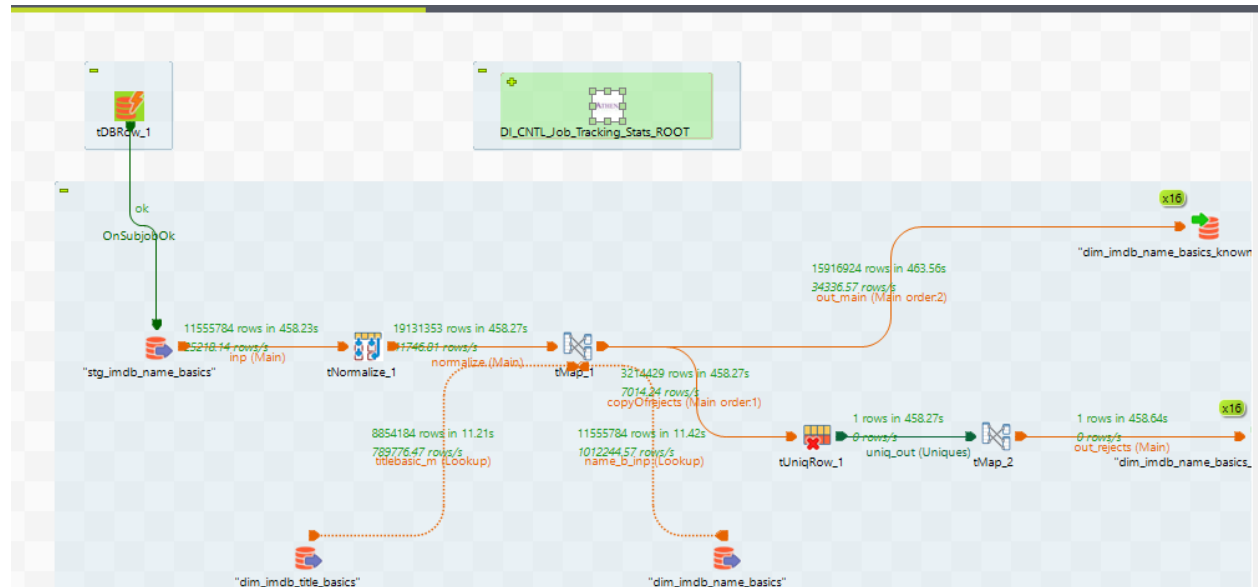
DIM_IMDB_TITLE_AKAS



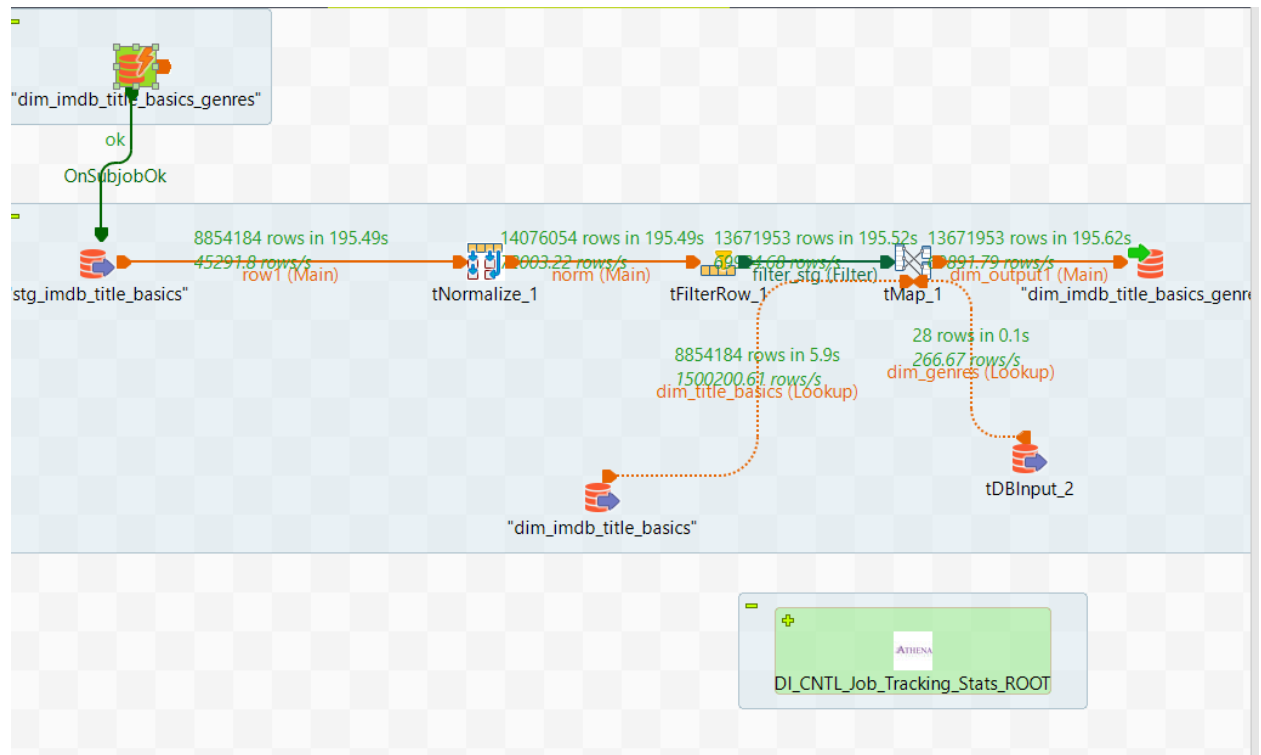
DIM_IMDB_GENRES



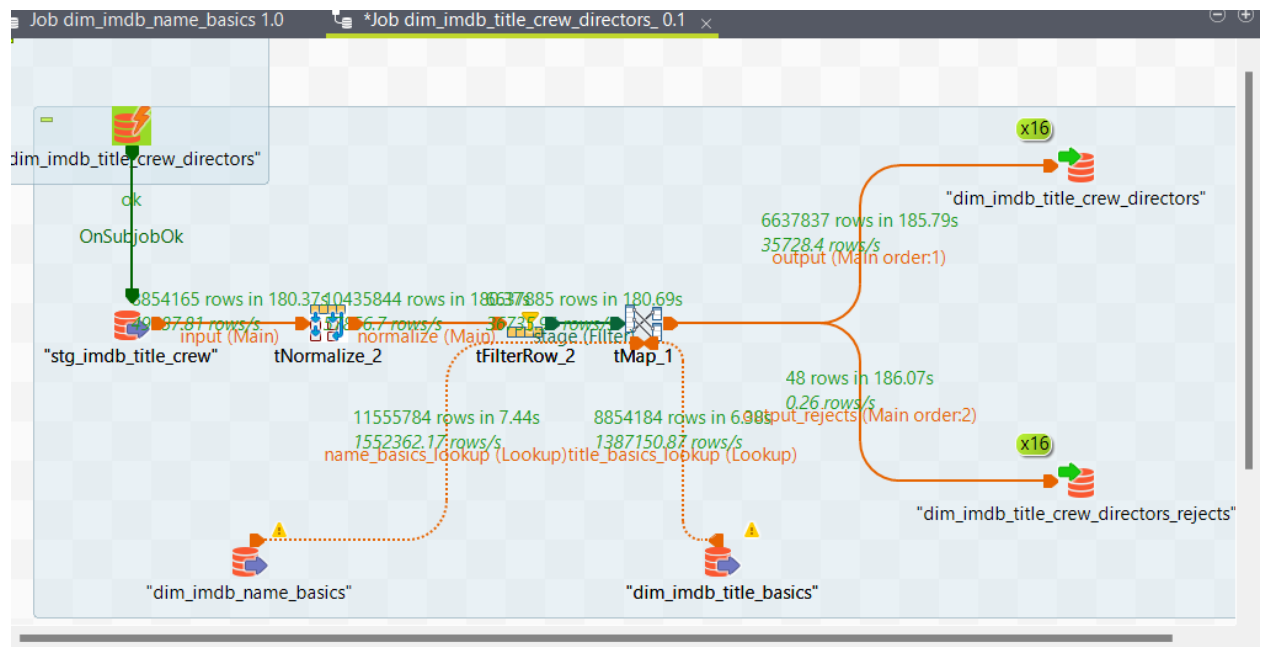
DIM_IMDB_NAME_BASICS_KNOWN_FOR_TITLES



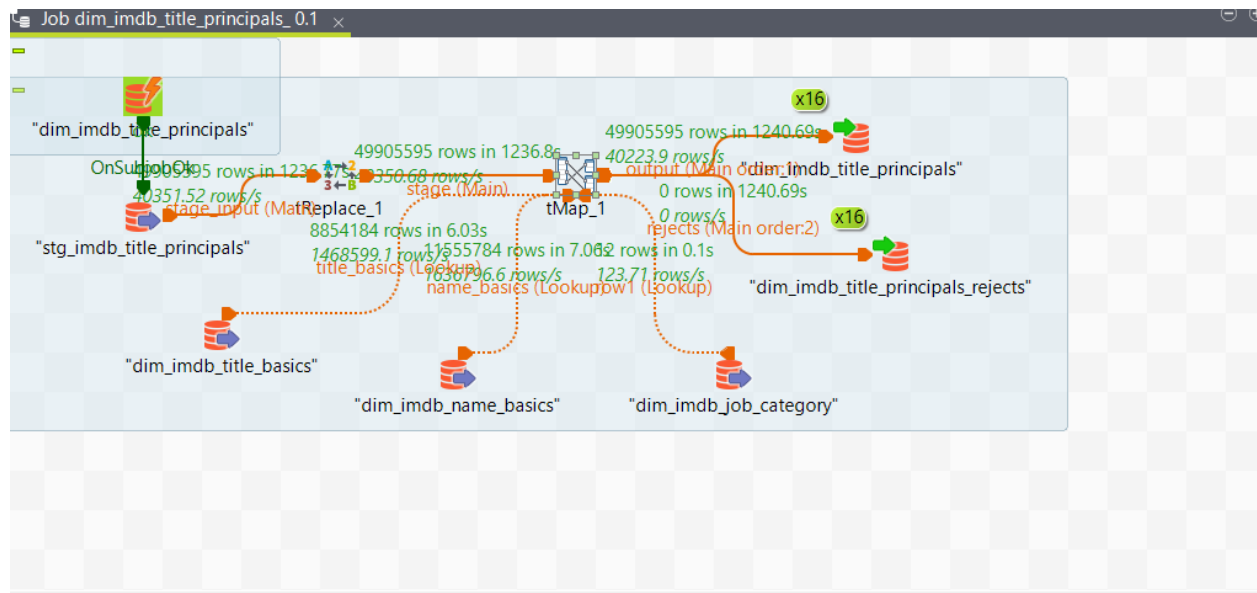
DIM_IMDB_TITLE_BASICS_GENRE



DIM_IMDB_TITLE_CREW_DIRECTORS

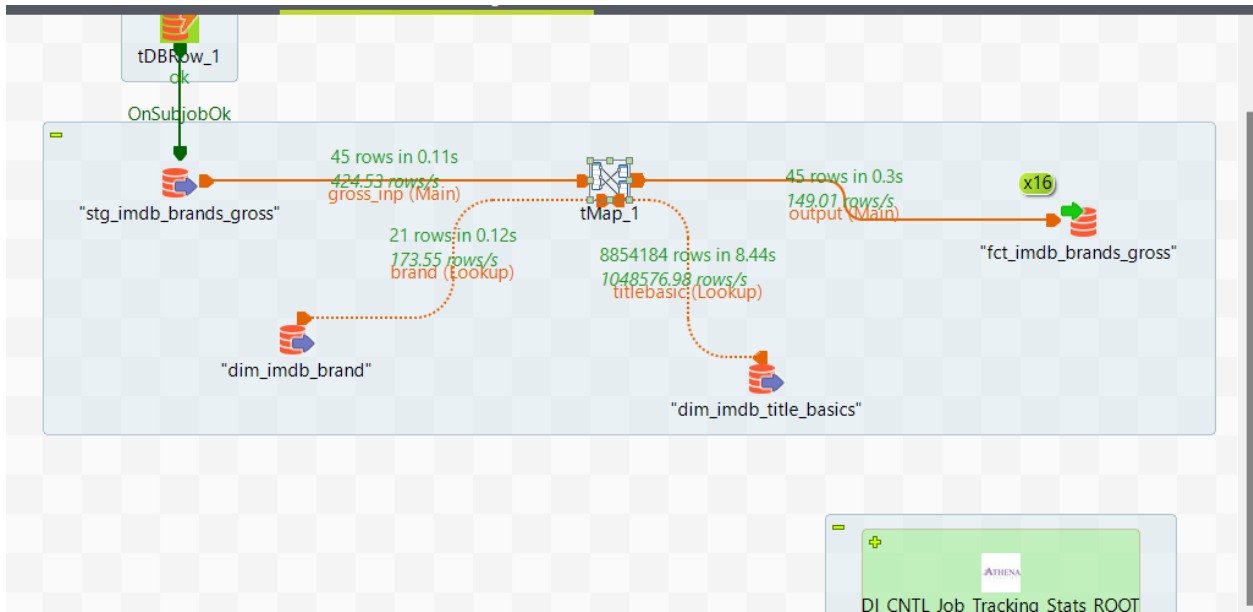


DIM_IMDB_TITLE_PRINCIPALS

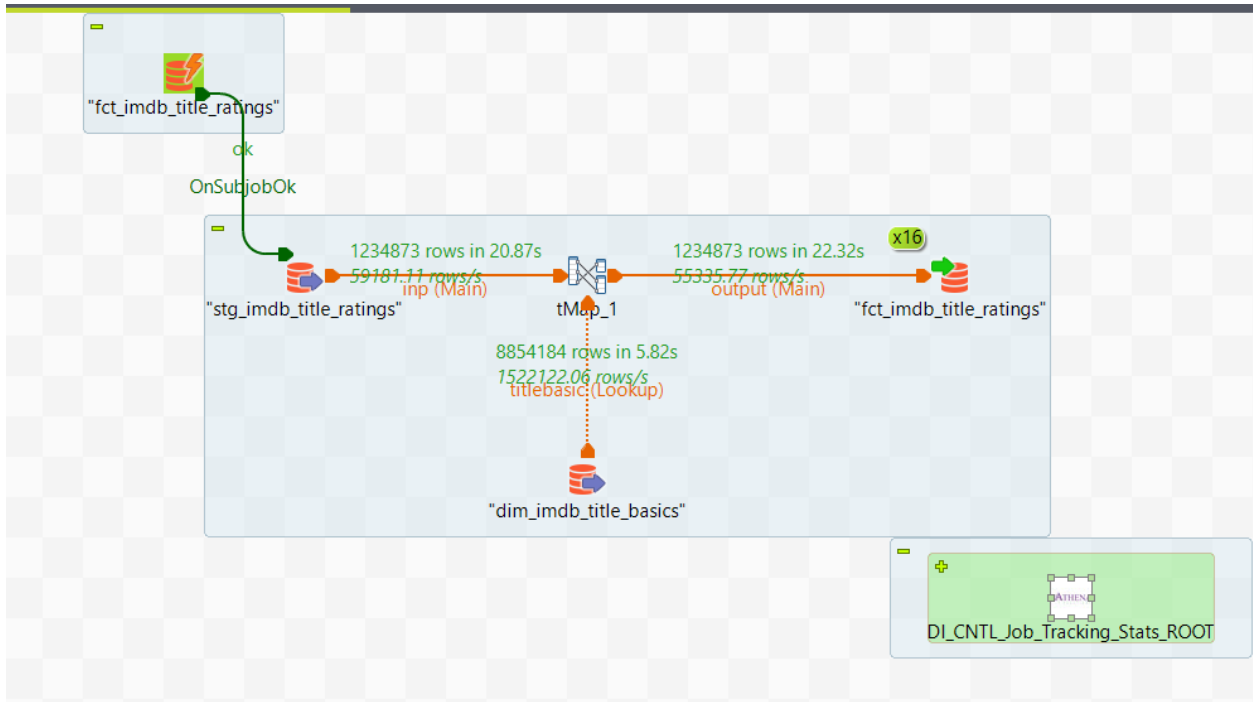


3. FACTS

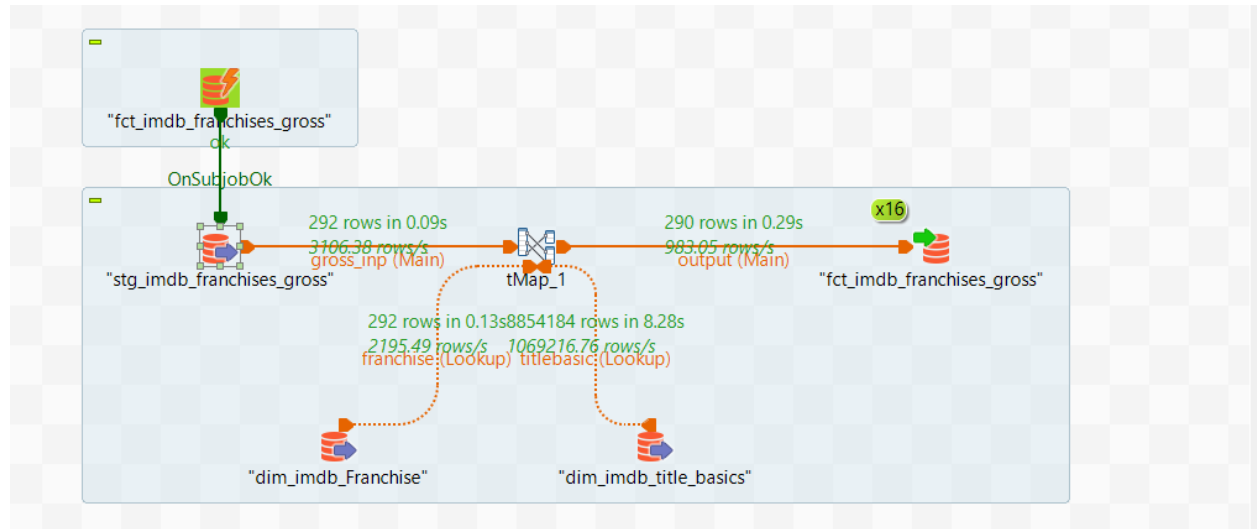
FCT_IMDB_BRAND_GROSS



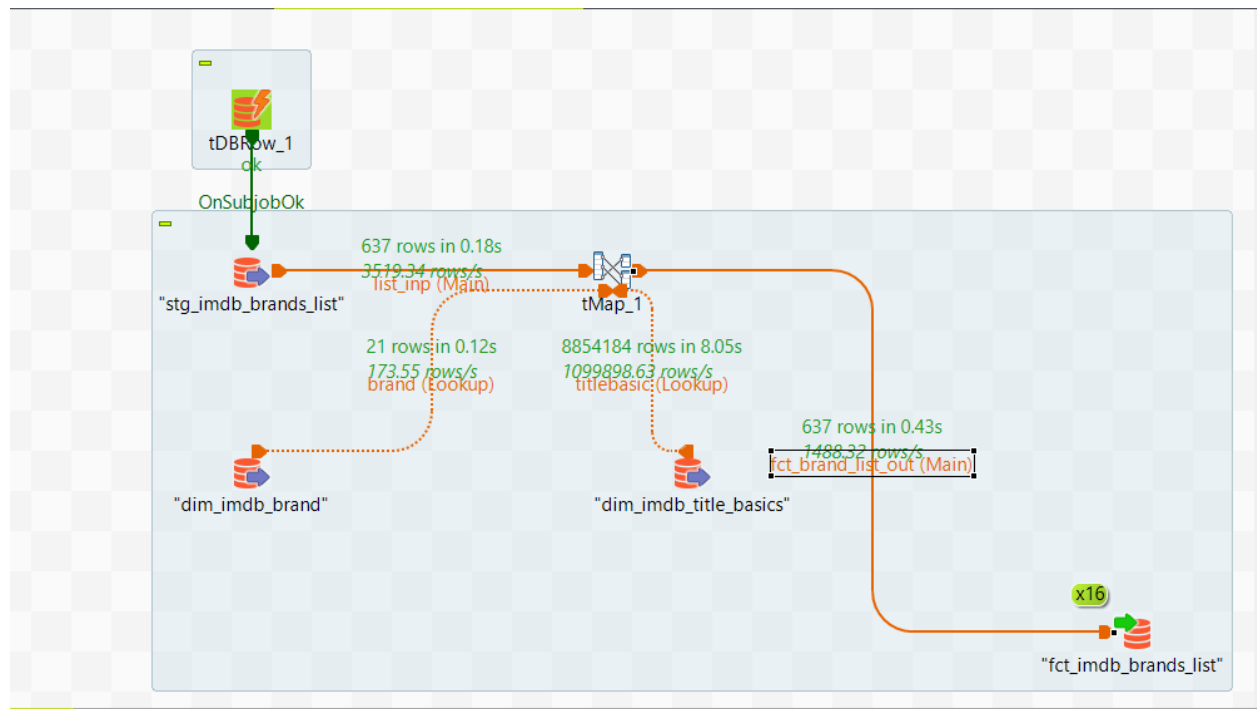
FCT_IMDB_TITLE_RATING



FACT_IMDB_FRANCHISE_GROSS



FCT_IMDB_BRANDS_LIST



4. STAGING ROW COUNTS

Staging	Row Count
STG_IMDB_NAME_BASICS	11,555,784
STG_IMDB_TITLE_AKAS	31,758,388
STG_IMDB_TITLE_EPISODE	6,641,848
STG_IMDB_TITLE_BASICS	8,854,184
STG_IMDB_TITLE_CREW	8,854,184
STG_IMDB_TITLE_RATINGS	1,234,873
STG_IMDB_TITLE_PRINCIPLE	49,905,595
STG_IMDB_TITLE_TYPE	13
STG_FRANCHISE_GROSS	292
STG_BRAND_GROSS	45
STG_FRANCHISE_LISTS	276
STG_DAILY_NUMBERS	786
STG_WORLDWIDE_BOX_OFFICE	3
STG_NUMBERS_BOX_OFFICE	948
STG_FRANCHISE_ALL_BOX_OFFICE	200
STG_NUMBERS_FRANCHISE_MOVIES_BOX_OFFICE	27

5. DIMENSION ROW COUNTS

Dimension	Row Count
DIM_IMDB_JOB_CATEGORY	12
DIM_IMDB_PRIMARY_PROFESSION	41
DIM_IMDB_TITLE_EPISODE	6641716
DIM_IMDB_TITLE_EPISODE_REJECTS	132
DIM_IMDB_NAME_BASICS_PRIMARY_PROFESSION	10563731
DIM_IMDB_FRANCHISE	292
DIM_IMDB_TITLE_BASICS	8854184
DIM_IMDB_TITLE_TYPE	13
DIM_IMDB_NAME_BASICS	11555784
DIM_IMDB_TITLE_AKAS	31733165
DIM_IMDB_TITLE_AKAS_REJECTS	5223
DIM_IMDB_GENRE	28
DIM_IMDB_NAME_BASICS_KNOWN_FOR_TITLES	15916924
DIM_IMDB_NAME_BASICS_KNOWN_FOR_TITLES_REJECTS	1
DIM_IMDB_TITLE_BASICS_GENRE	13671953
DIM_IMDB_TITLE_CREW_DIRECTORS	6637837
DIM_IMDB_TITLE_CREW_DIRECTORS_REJECTS	48
DIM_IMDB_TITLE_PRINCIPALS	49905595
DIM_IMDB_TITLE_PRINCIPALS_REJECTS	0

6. FACTS ROW COUNT

Fact	Row Count
FCT_IMDB_BRAND_GROSS	45
FCT_IMDB_TITLE_RATING	1234873
FACT_IMDB_FRANCHISE_GROSS	290
FCT_IMDB_BRANDS_LIST	637

7. RUNNING OF PARALLEL JOBS

The screenshot displays the Apache Airflow web interface. On the left, the 'Repository' pane shows a tree structure of job designs under 'LOCAL: imdb_final_main'. The main area shows a DAG named 'Job master_job' with a 'Designer' view. The DAG consists of several tasks connected by arrows, with 'Parallelize' labels indicating parallel execution. The tasks include 'tParallelize_1', 'stg_box_office_WORLDWIDE', 'stg_imdb_brands_gross', 'stg_imdb_brands_list', and 'stg_imdb_franchise_list_main'. Below the DAG, the 'Execution' tab is active, showing a list of job names and their corresponding job IDs. The 'Job master_job' is listed with a job ID of 'stg_imdb_title_rat'. The 'Execution' tab also includes buttons for 'Run', 'Kill', and 'Clear'. On the right, the 'Palette' pane shows various task types like 'Input', 'Output', 'Orchestration', and 'Misc'.

Job master_job

Execution

Name	Value
stg_imdb_title_rat	stg_imdb_title_rat
stg_numbers_franch	stg_numbers_franch
stg_imdb_title_typ	stg_imdb_title_typ
stg_box_office_WOR	stg_box_office_WOR
stg_imdb_brands_gr	stg_imdb_brands_gr