

NORTHEASTERN UNIVERSITY

INFORMATION SYSTEMS SPRING '20



IMDb Database Project

**DESIGNING ADVANCED DATA ARCHITECTURES FOR
BUSINESS INTELLIGENCE**

INFO 7370 – 35649

Guided By:

Prof. Richard Sherman

Group 10:

Priyanka Malpekar, Tanvi Gurav, Siddhi Prabhu

List of Tables

Sr. No.	Table Name	Table Type	Count
MOVIE LENS TABLES			
1	stg_ml_genome_scores	Staging Table	15,584,448
2	stg_ml_genome_tags	Staging Table	1,128
3	stg_ml_links	Staging Table	62,423
4	stg_ml_movies	Staging Table	62,423
5	stg_ml_ratings	Staging Table	25,000,095
6	stg_ml_tags	Staging Table	1,093,360 (1092869+Rejects: 491)
IMDB TABLES			
7	stg_iso_country	Staging Table	249
8	stg_iso_language	Staging Table	486
9	stg_imdb_name_basics	Staging Table	10,008,408
10	stg_imdb_title_episode	Staging Table	4,764,382
11	stg_imdb_title_ratings	Staging Table	1,023,897
12	stg_imdb_title_crew	Staging Table	6,704,790
13	stg_imdb_title_principals	Staging Table	38,699,152
14	stg_imdb_title_akas	Staging Table	21,614,617
15	stg_imdb_title_basics	Staging Table	6,704,790
16	stg_imdb_title_basics_genres	Staging Table	10,158,625

17	STG_box_office_worldwide	Staging Table	1,000
18	stg_imdb_title_crew_directors	Staging Table	5,170,628
19	stg_imdb_title_crew_writers	Staging Table	7,916,702
20	stg_imdb_name_basics_knownF orTitles	Staging Table	21,575,298
21	stg_imdb_brands_list	Staging Table	629
22	stg_imdb_franchises_list	Staging Table	282
23	stg_imdb_franchises_gross	Staging Table	289
24	stg_imdb_brands_gross	Staging Table	44
25	stg_imdb_name_basics_primary Profession	Staging Table	10,986,452
NUMBERS TABLES			
26	stg_numbers_daily_box_office	Staging Table	97
27	stg_numbers_franchise_all_box_o ffice	Staging Table	867
28	stg_numbers_franchise_movies_b ox_office	Staging Table	28
DIMENSIONS			
29	dim_imdb_titleType	Dimension	10
30	dim_imdb_title_basics	Dimension	6,704,790
31	dim_imdb_Franchise	Dimension	289
32	dim_imdb_primaryProfession	Dimension	40
33	dim_iso_language	Dimension	486

34	dim_imdb_name_basics	Dimension	10,008,408
35	dim_iso_country	Dimension	249
36	dim_imdb_genres	Dimension	28
37	dim_imdb_title_episode	Dimension	4,764,372
38	dim_imdb_brand	Dimension	44
39	dim_imdb_job_category	Dimension	12
40	dim_imdb_name_basics_primaryProfession	Dimension	10,986,452
41	dim_imdb_name_basics_knownForTitles	Dimension	21,565,259
42	dim_imdb_name_basics_knownForTitles_rejects	Reject_Dimension	10,039
43	dim_imdb_title_principals	Dimension	38,676,545
44	dim_imdb_title_principals_rejects	Reject_Dimension	22,607
45	dim_imdb_title_basics_genres	Dimension	10,158,625
46	imdb_projectdim_imdb_title_crew_writers	Dimension	7,916,672
47	imdb_projectdim_imdb_title_crew_writers_rejects	Reject_Dimension	30
48	imdb_projectdim_imdb_title_crew_directors	Dimension	5,170,597
49	imdb_projectdim_imdb_title_crew_directors_rejects	Reject_Dimension	31
50	dim_imdb_title_episode	Dimension	4,764,372
51	dim_imdb_title_episode_rejects	Reject_Dimension	10
Facts			

52	fct_ml_genome_tags	Fact	1128
53	fct_ml_genome_scores	Fact	15,560,760
54	fct_ml_genome_scores_rejects	Fact	23,688
55	fct_movies_box_office_worldwide	Fact	1000
56	fct_imdb_title_ratings	Fact	1,023,897
57	fct_ml_tags	Fact	1092260
58	fct_ml_tags_rejects	Fact	609
59	fct_imdb_brands_gross	Fact	44
60	fct_ml_ratings	Fact	24983941
61	fct_ml_ratings_rejects	Fact	16,154
62	fact_numbers_daily_box_office	Fact	97
63	fact_imdb_brands_gross	Fact	44
64	fct_numbers_franchise_all_box_office	Fact	867
65	fct_numbers_franchise_movies_box_office	Fact	28
66	fct_imdb_brands_list	Fact	44
67	fct_imdb_franchises_gross	Fact	289

Screenshots of Job Execution

Table: stg_ml_genome_tags

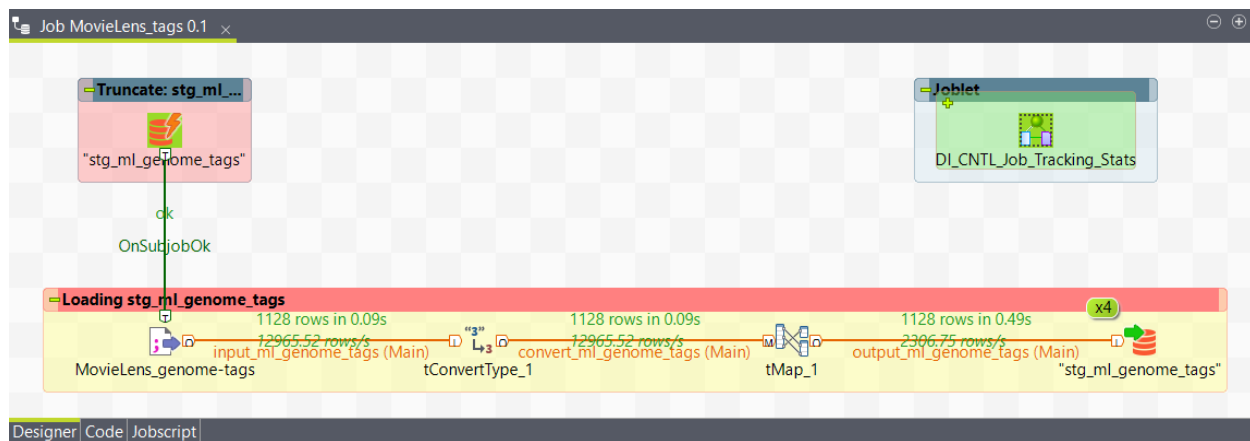


Table: stg_ml_movies

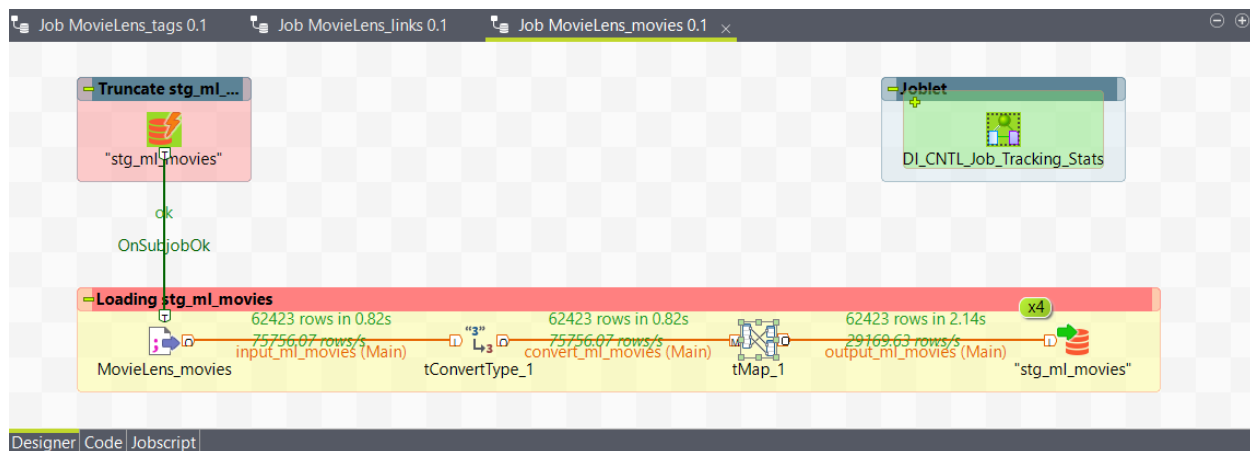


Table: stg_ml_genome_scores

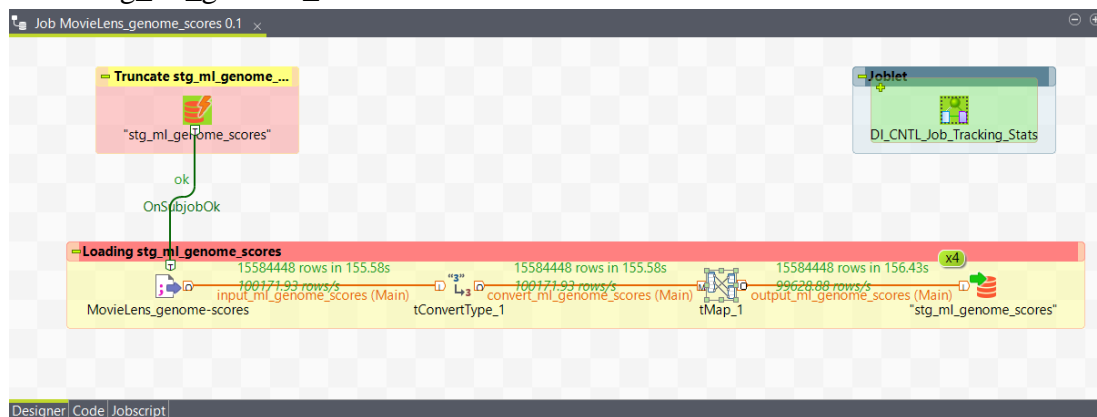


Table: stg_ml_links

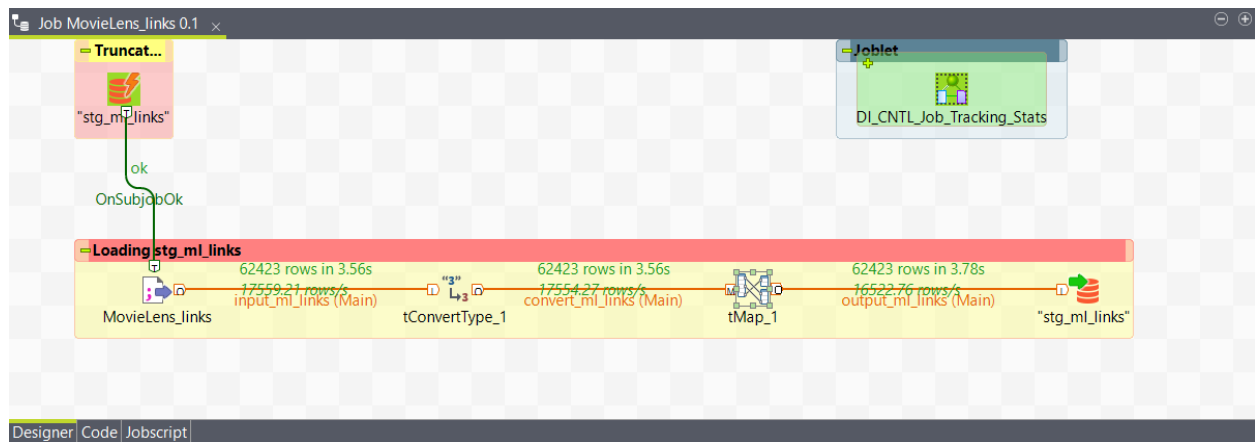


Table: stg_ml_ratings

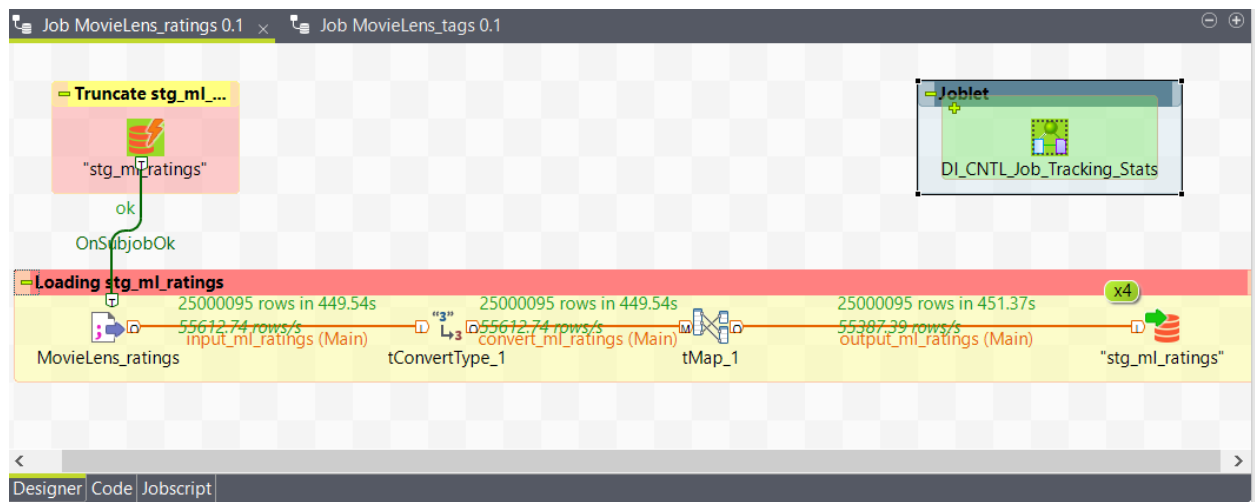


Table: stg_ml_tags

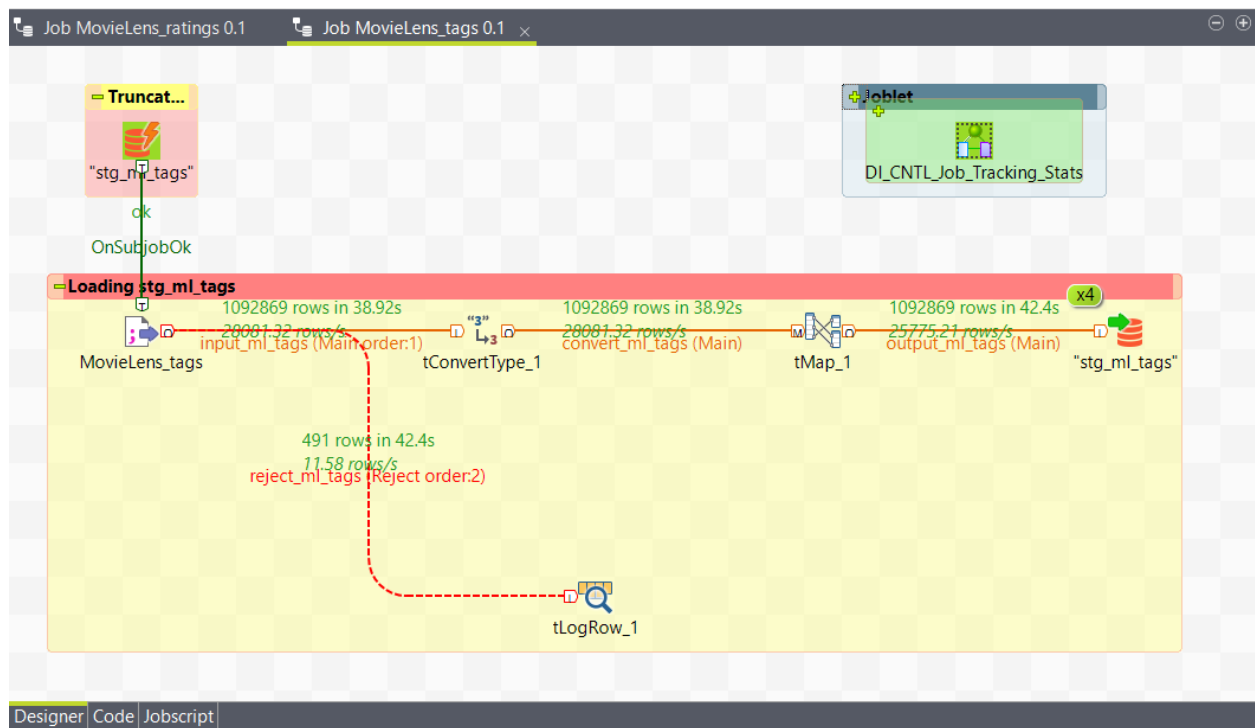


Table: stg_imdb_brands_gross

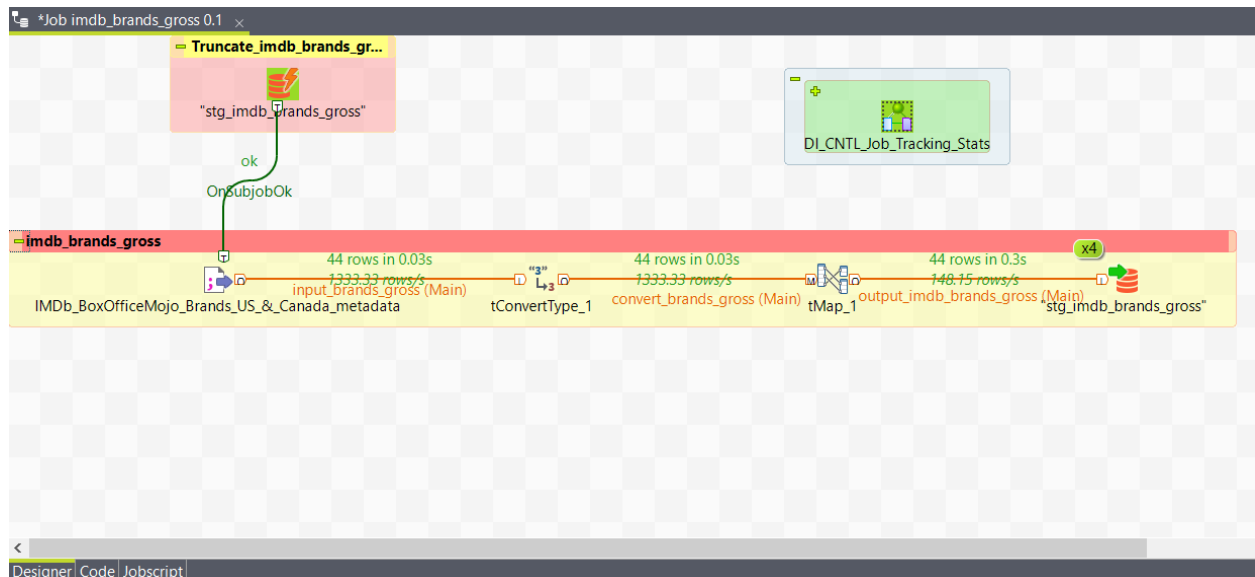


Table: stg_imdb_brands_list

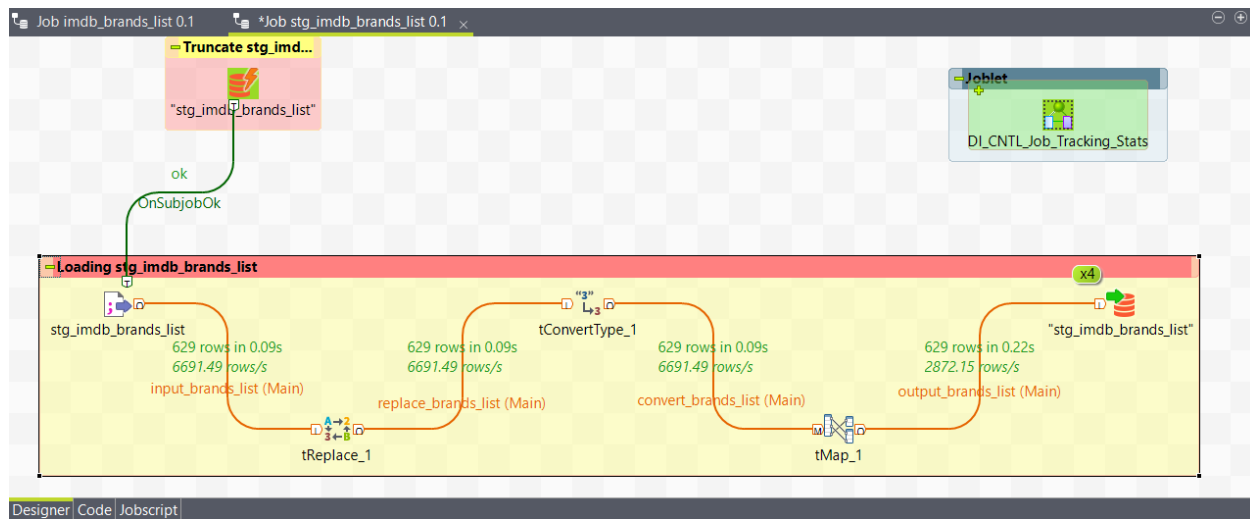


Table: imdb_franchises_gross

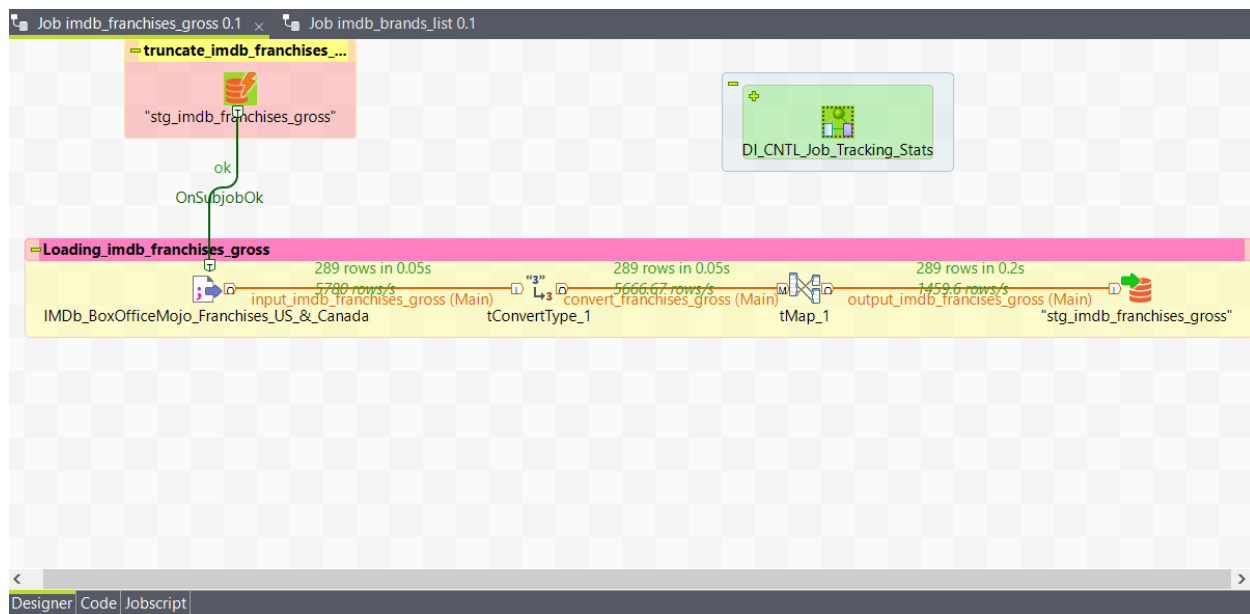


Table: stg_imdb_franchises_list

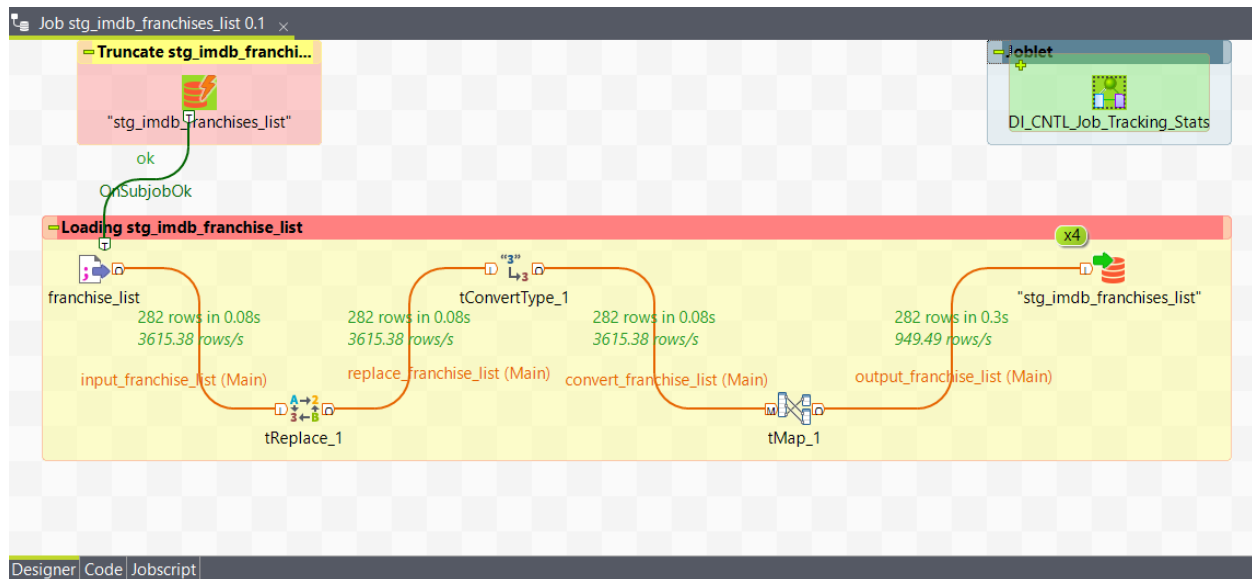


Table: stg_iso_country

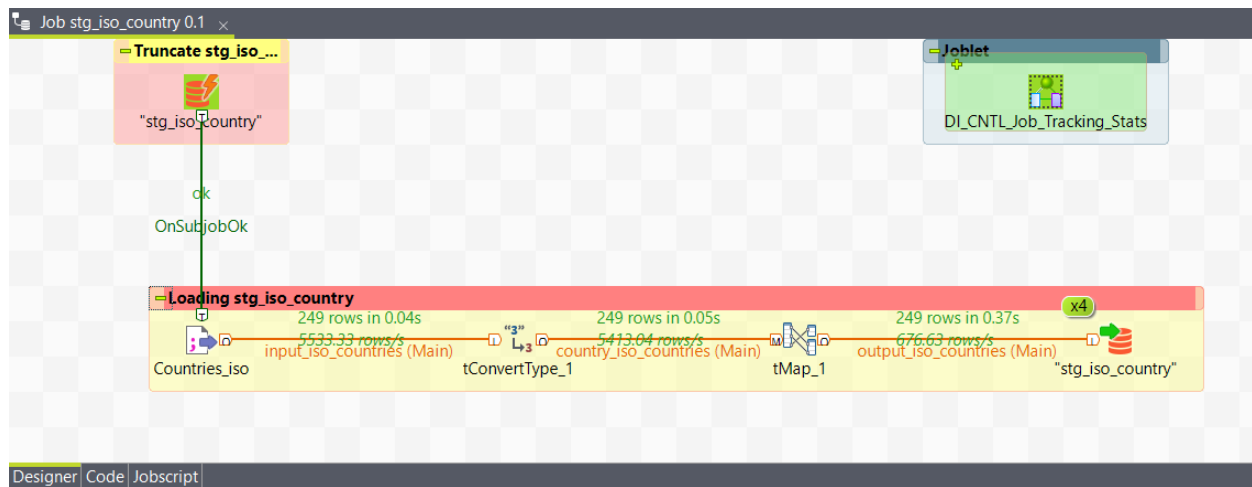


Table: stg_iso_language

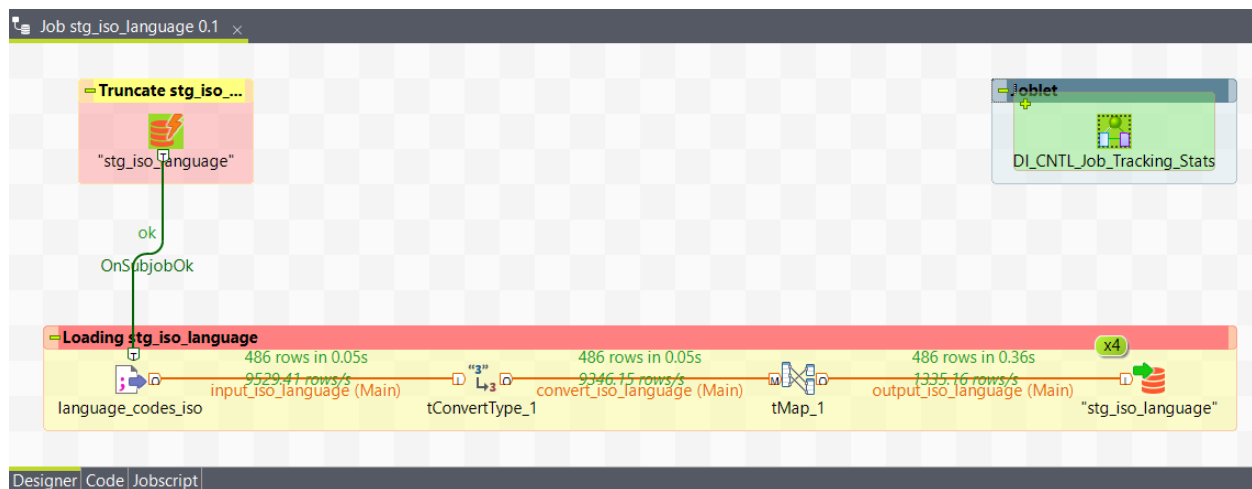


Table: stg_imdb_name_basics

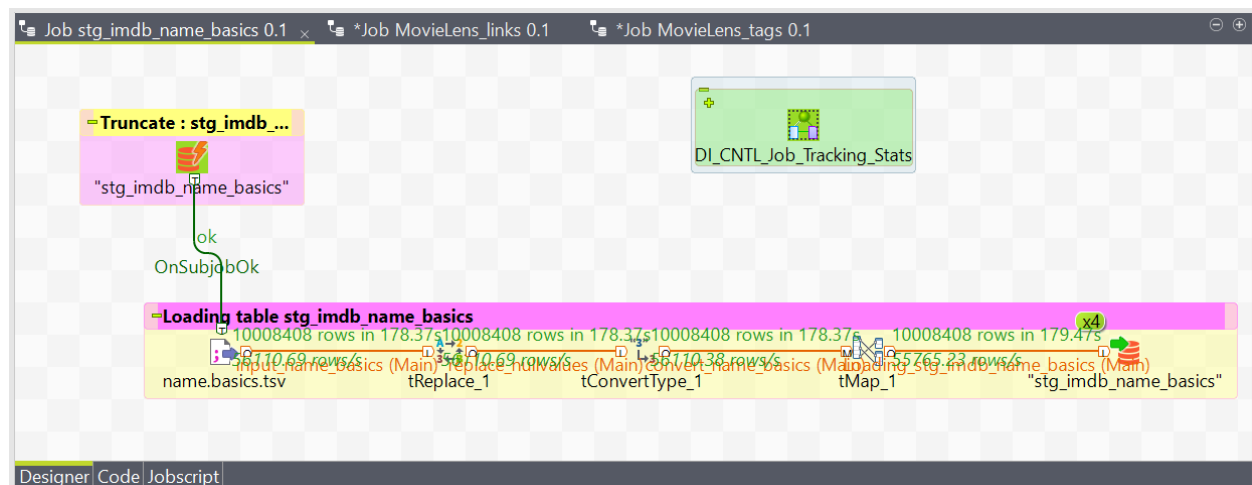


Table: stg_imdb_title_ratings

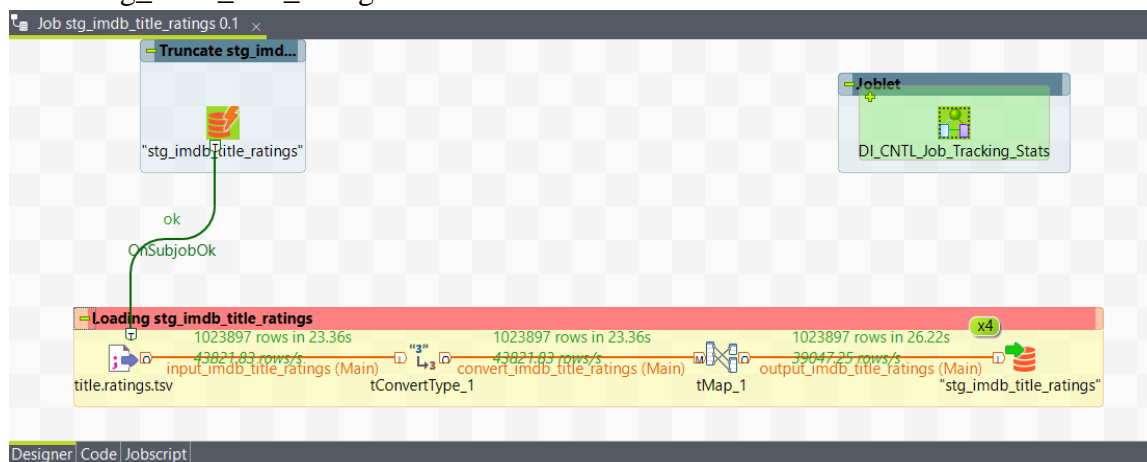


Table: stg_imdb_title_episodes

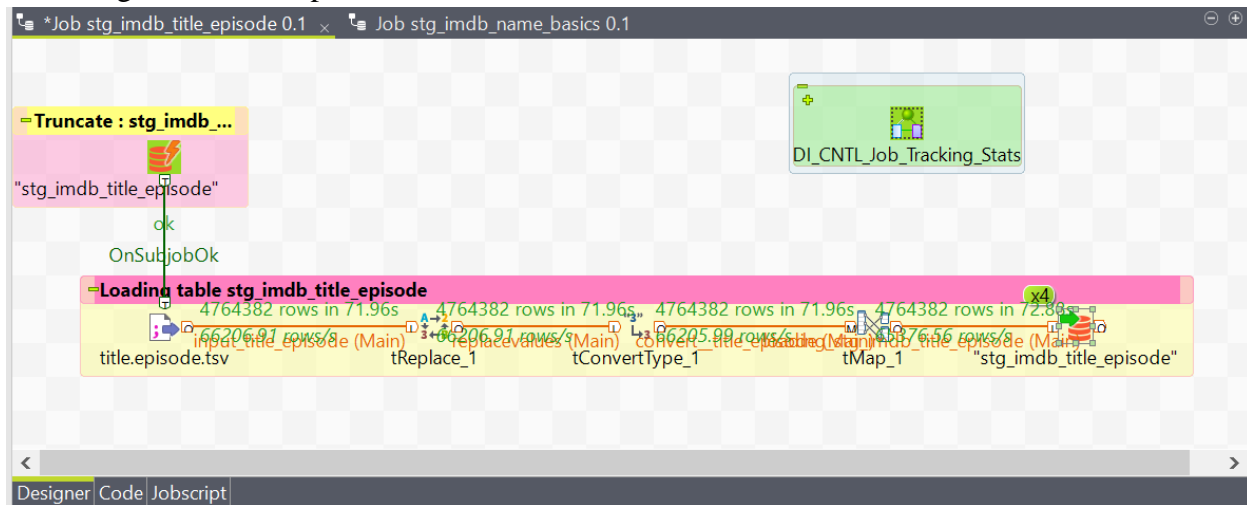


Table: stg_imdb_title_crew

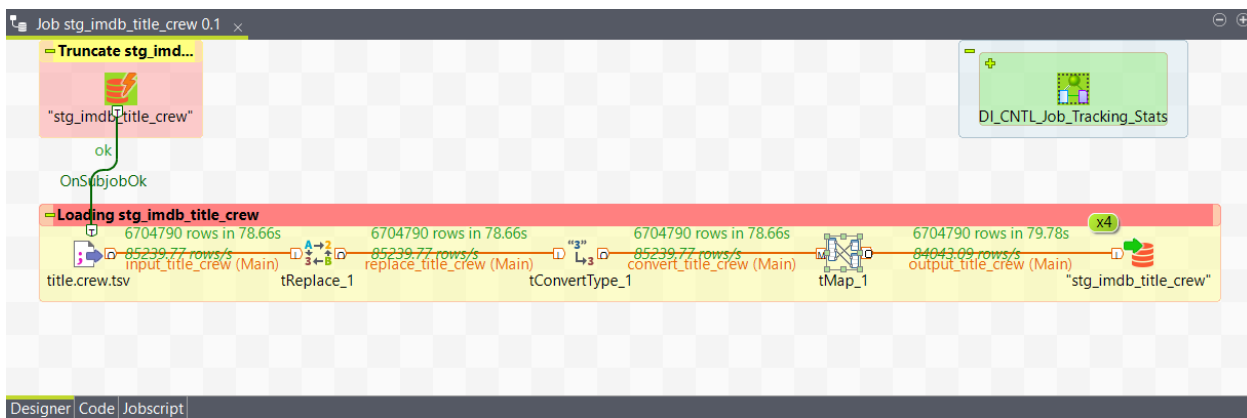


Table: stg_imdb_title_principals

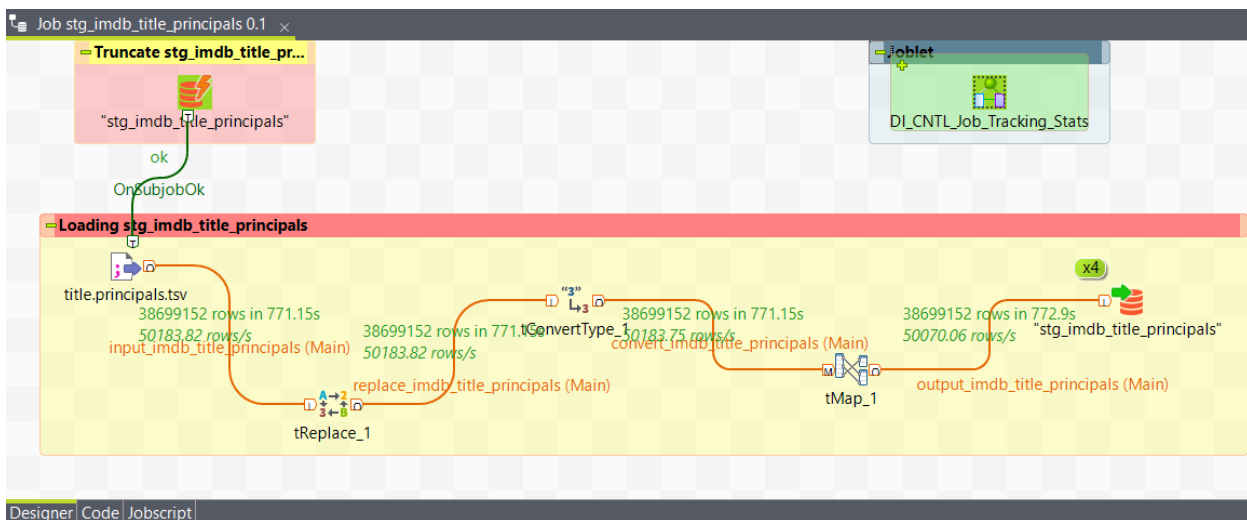


Table: stg_box_office_worldwide

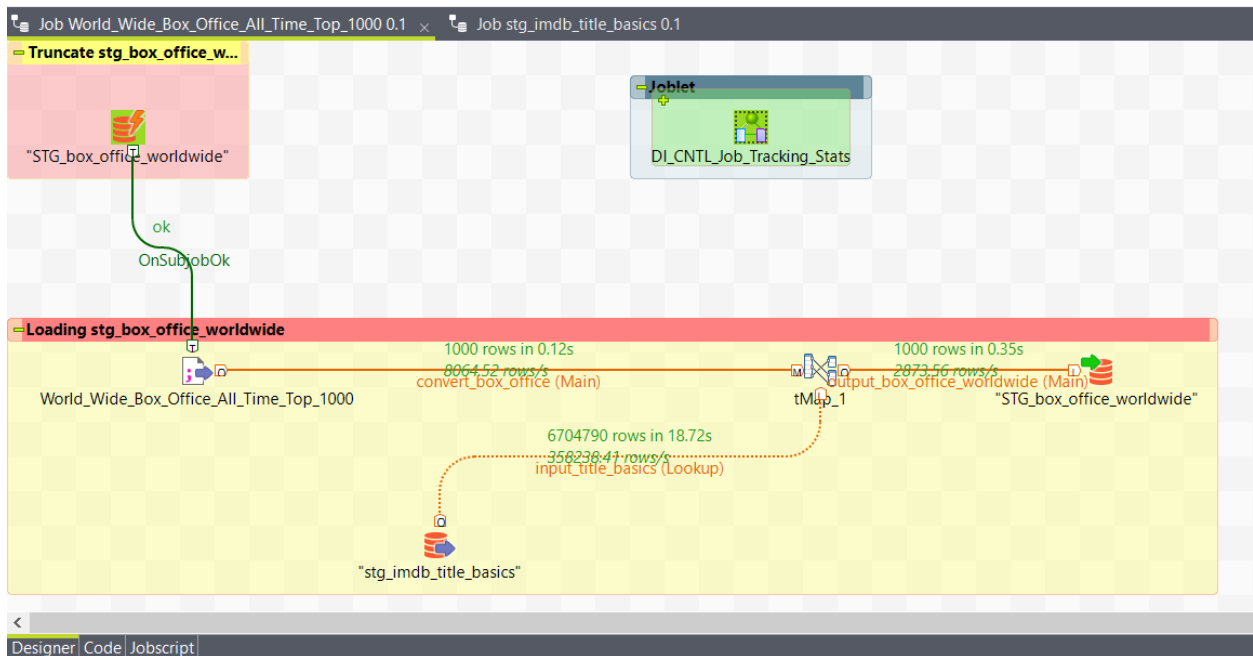


Table: stg_imdb_title_basics

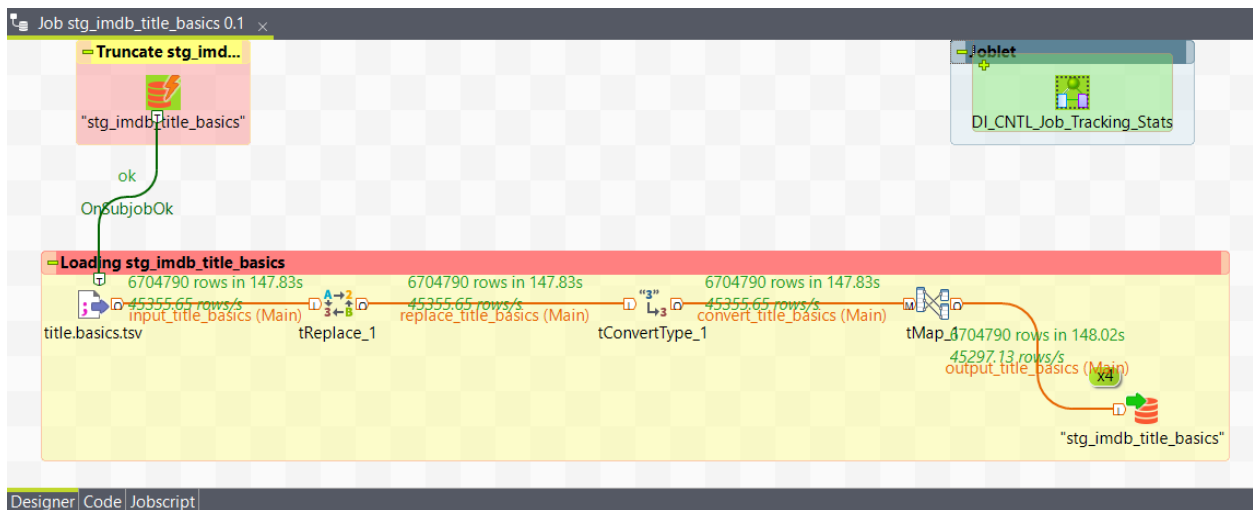


Table: stg_imdb_title_basics_genres

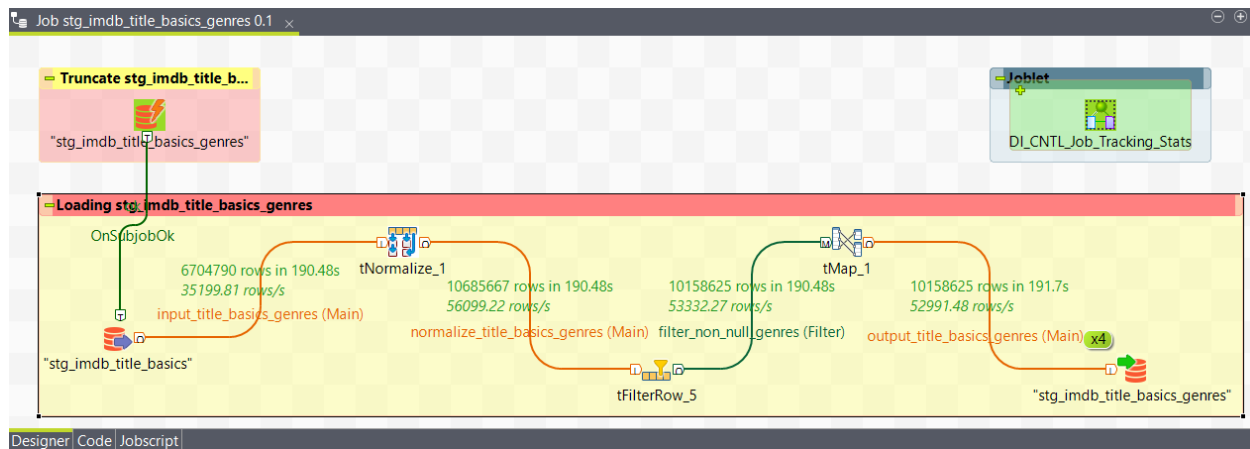


Table: stg_imdb_title_crew_directors

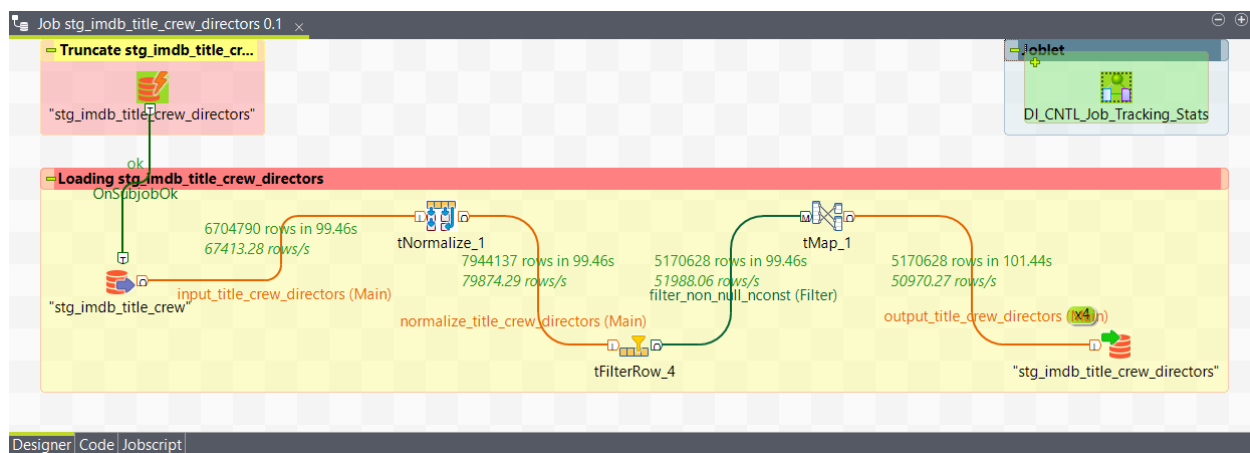


Table: stg_imdb_title_crew_writers

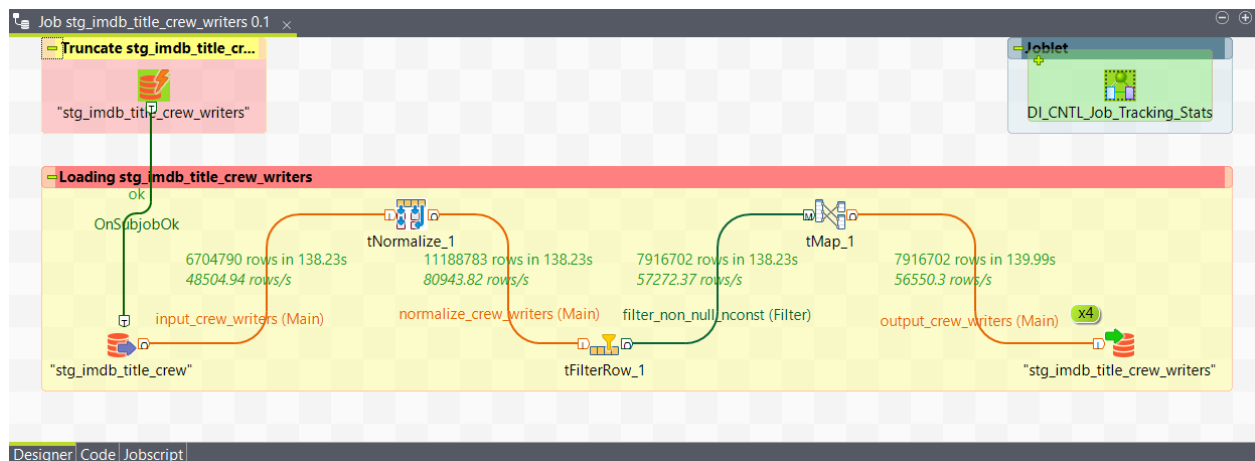


Table: stg_imdb_title_crew

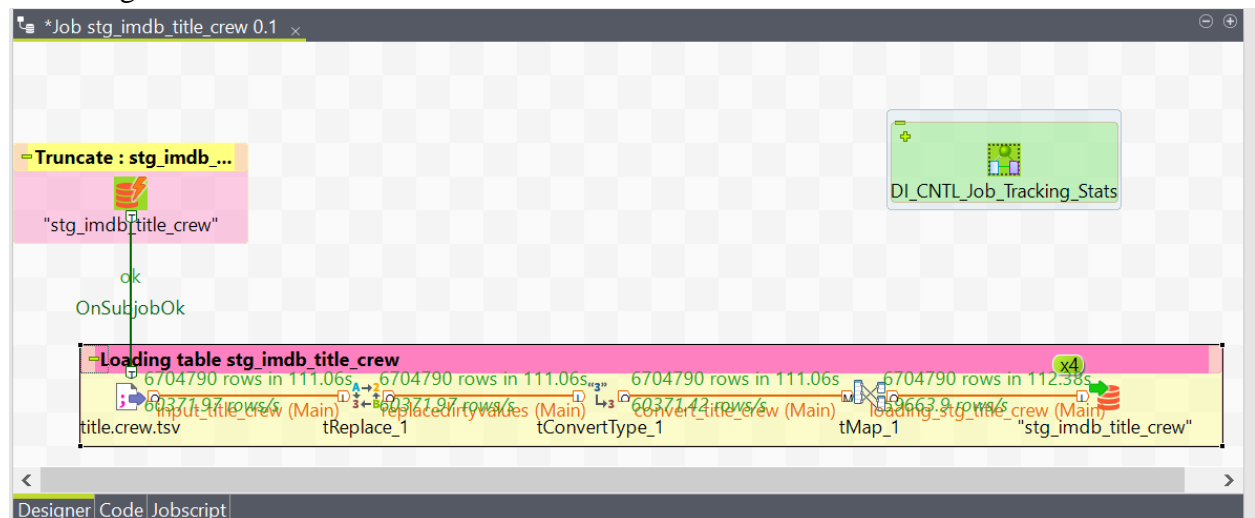


Table: stg_imdb_title_akas

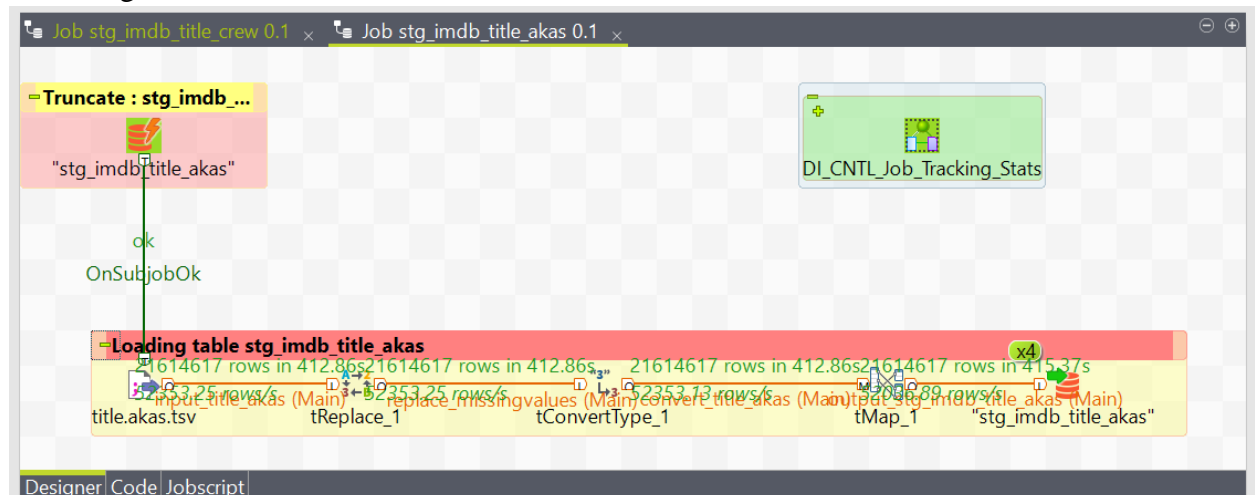


Table: stg_imdb_name_basics_primaryProfession

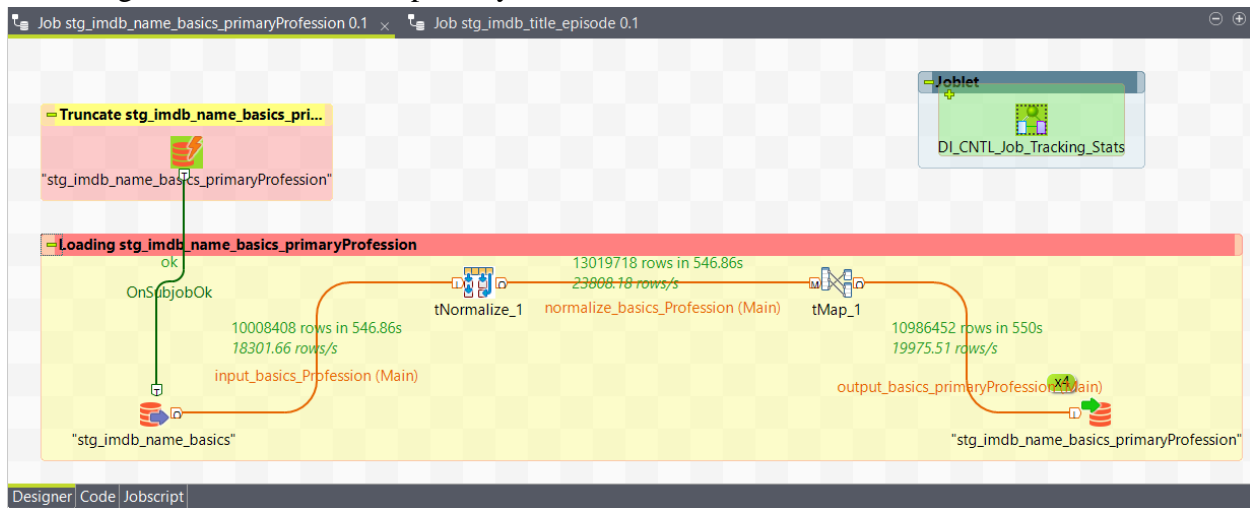


Table: stg_imdb_name_basics_knownForTitles

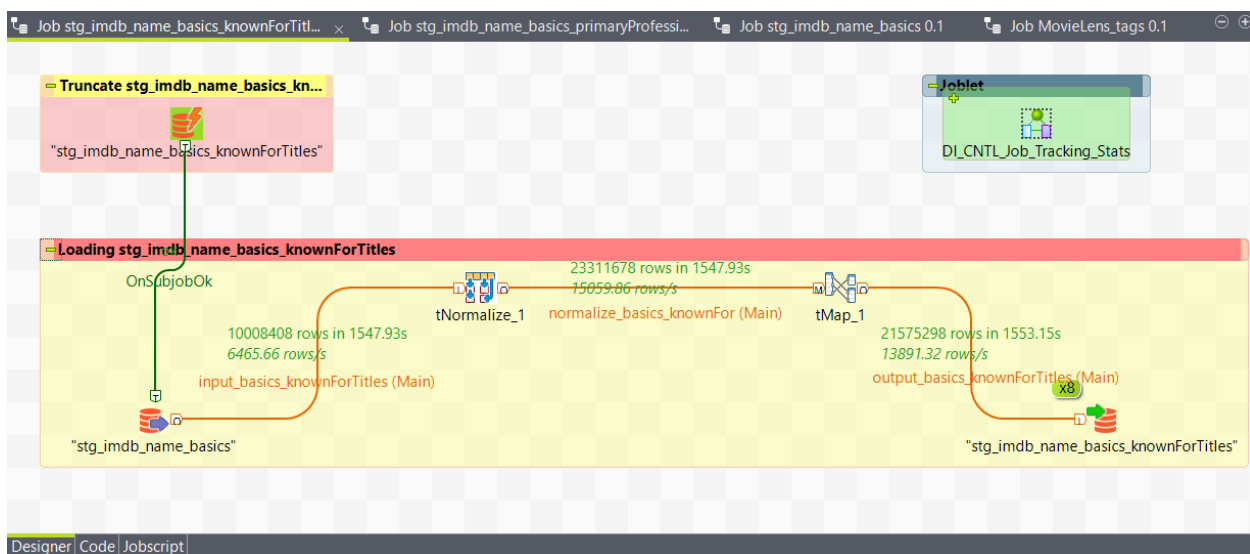


Table: stg_numbers_daily_box_office

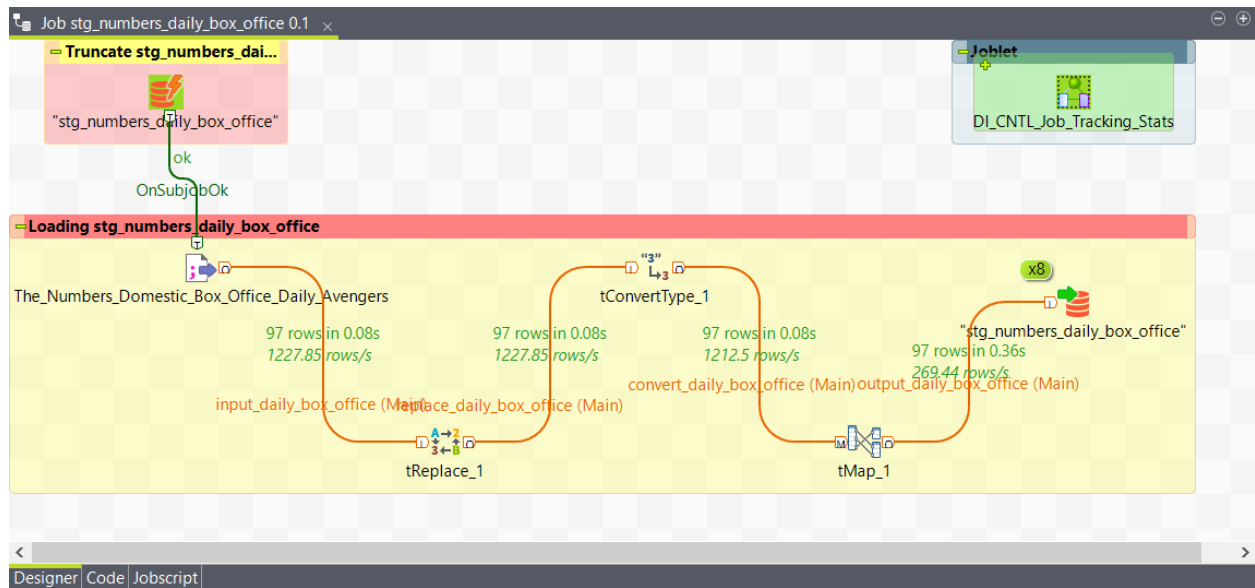


Table: stg_numbers_franchise_all_box_office

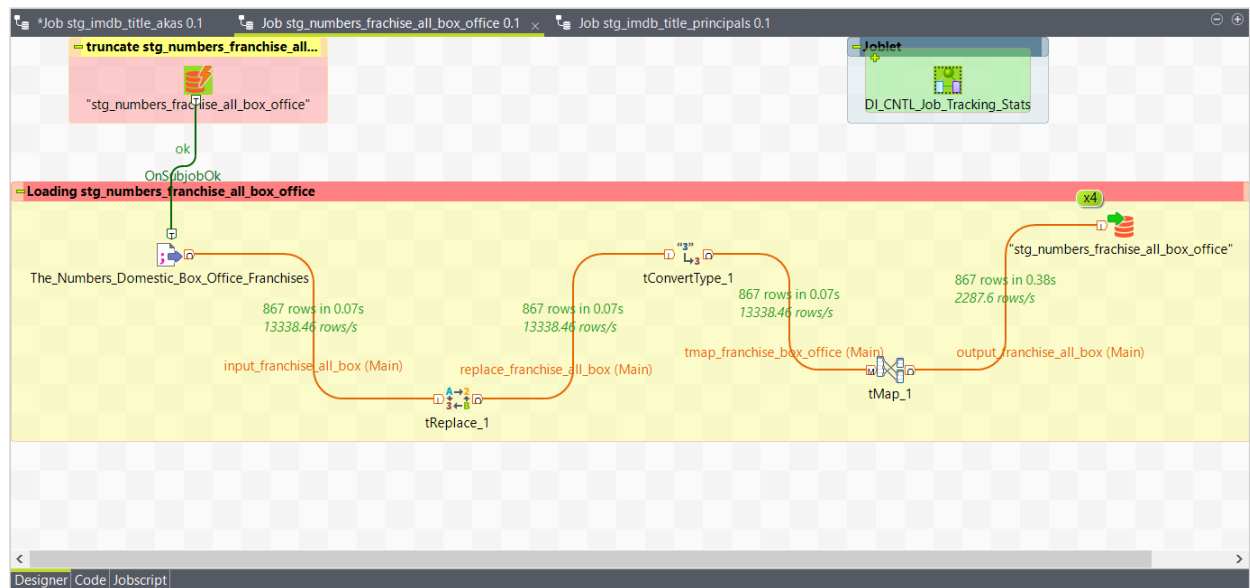
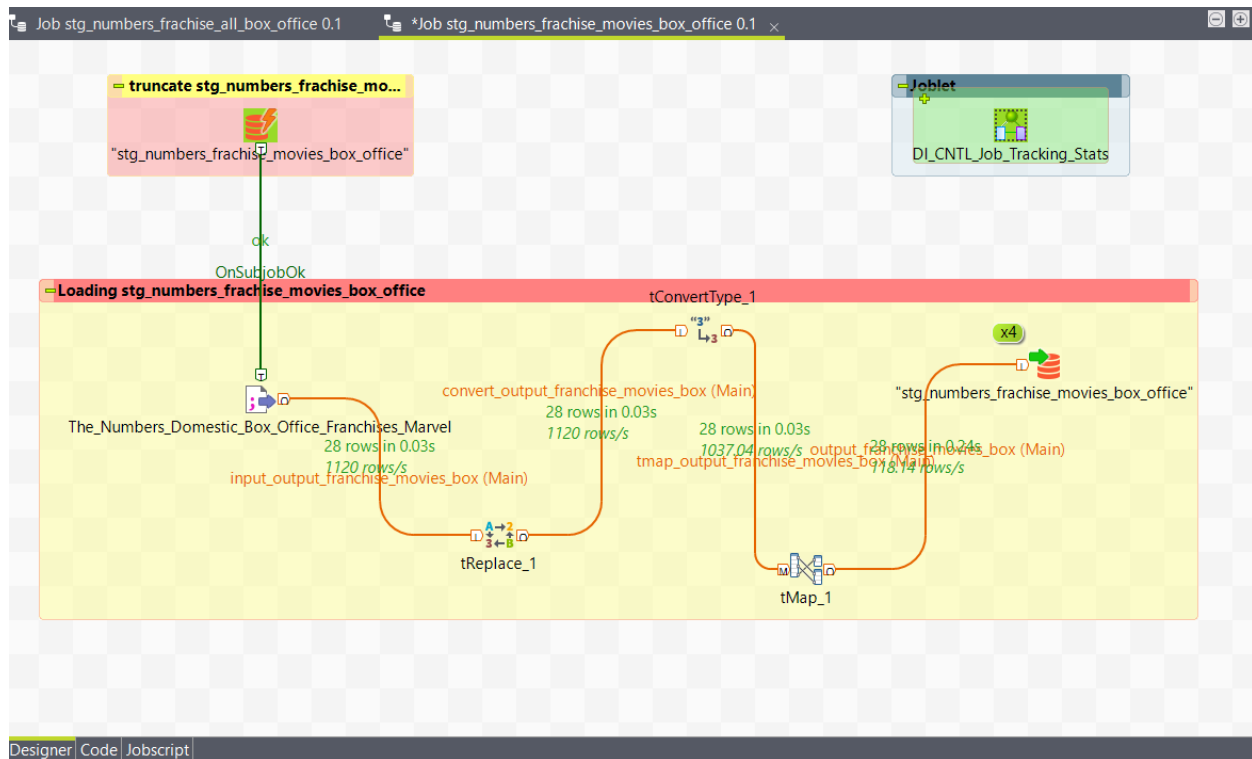


Table: stg_numbers_franchise_movies_box_office



DIMENSIONS:

Table : dim_imdb_titleType

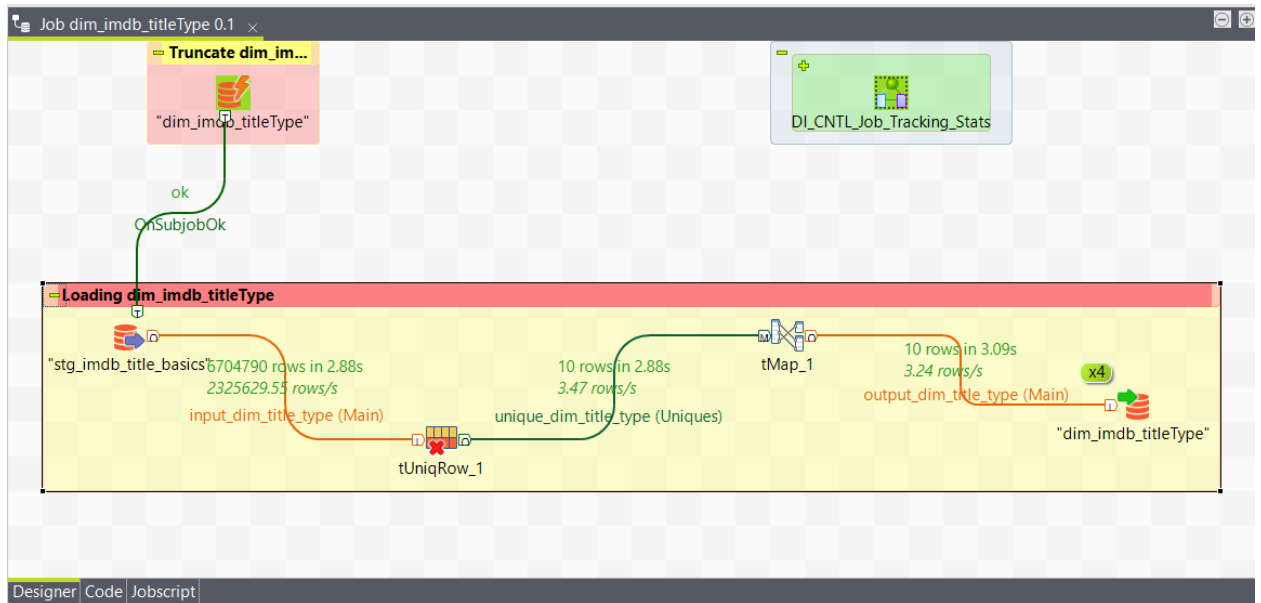


Table: dim_imdb_title_basics

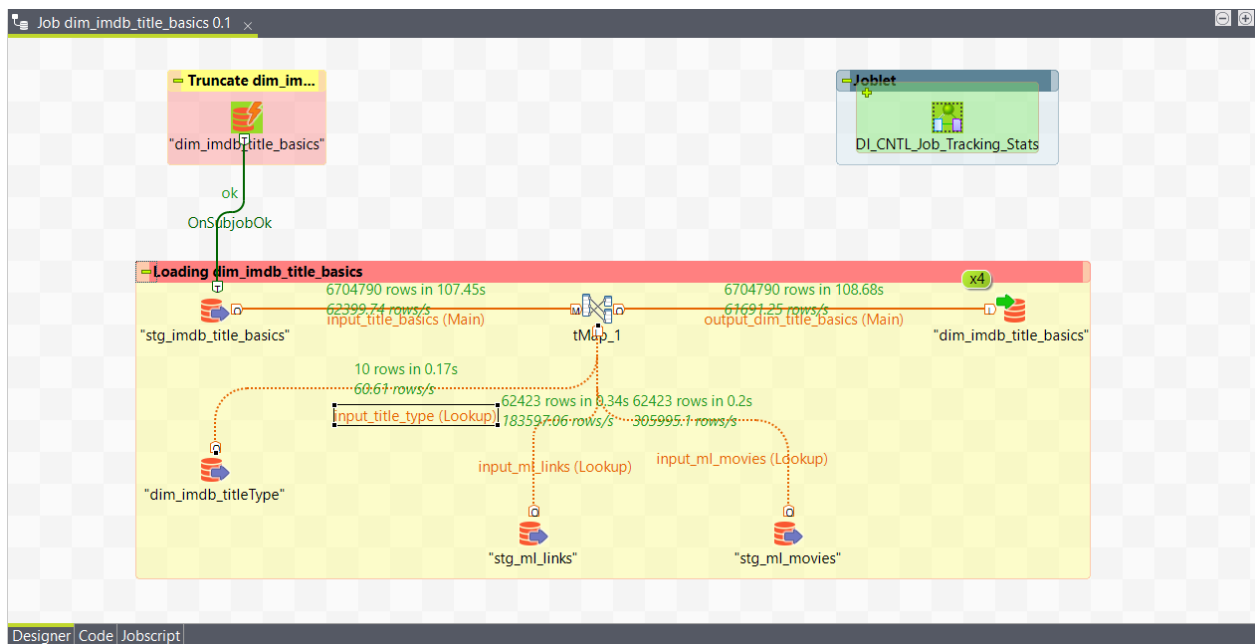


Table: dim_imdb_Franchise

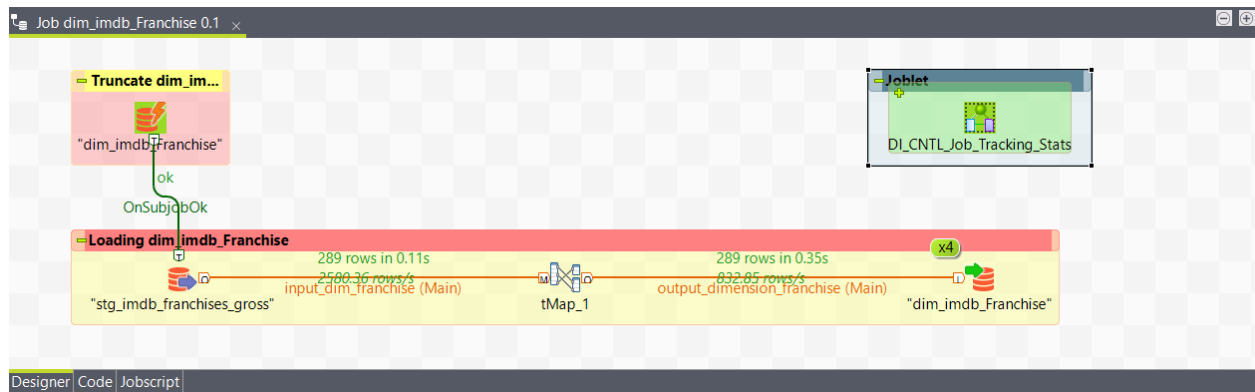


Table: dim_imdb_primary_profession

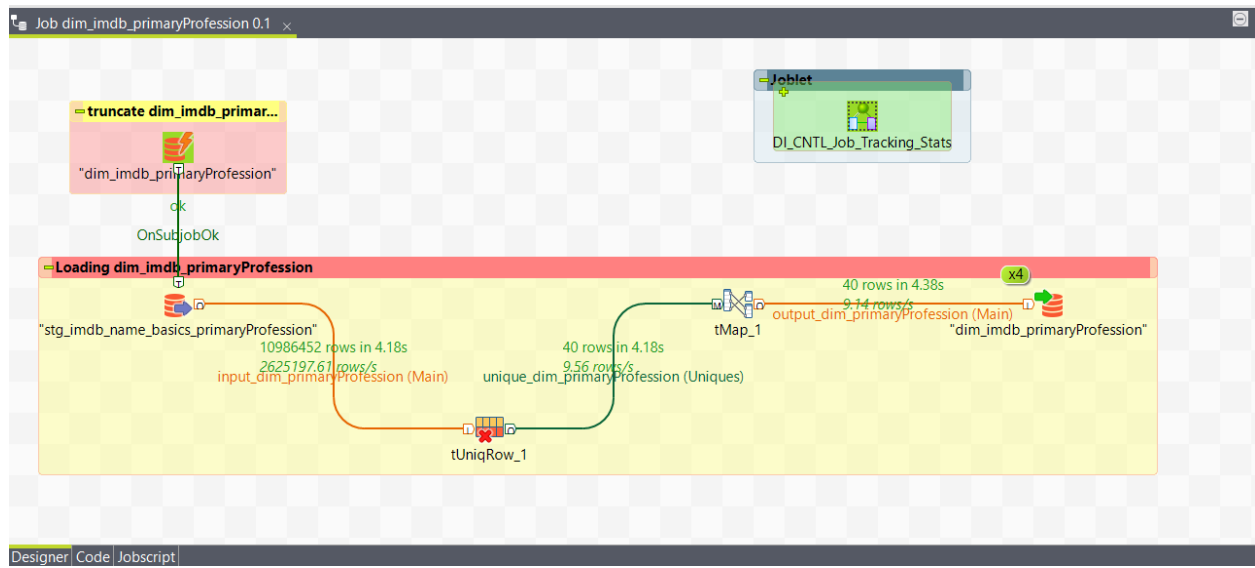


Table: dim_imdb_name_basics

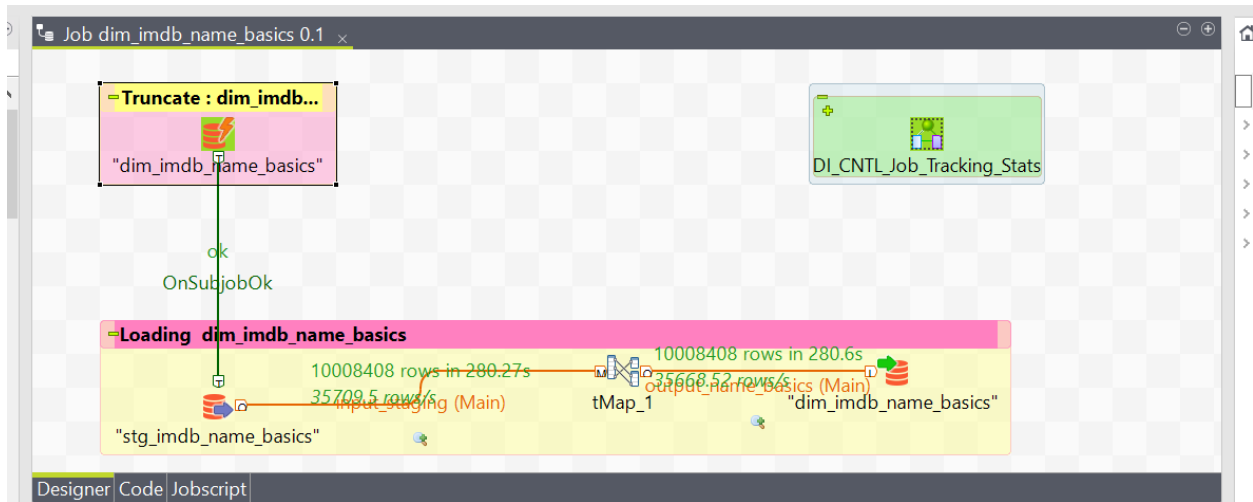


Table: dim_iso_language

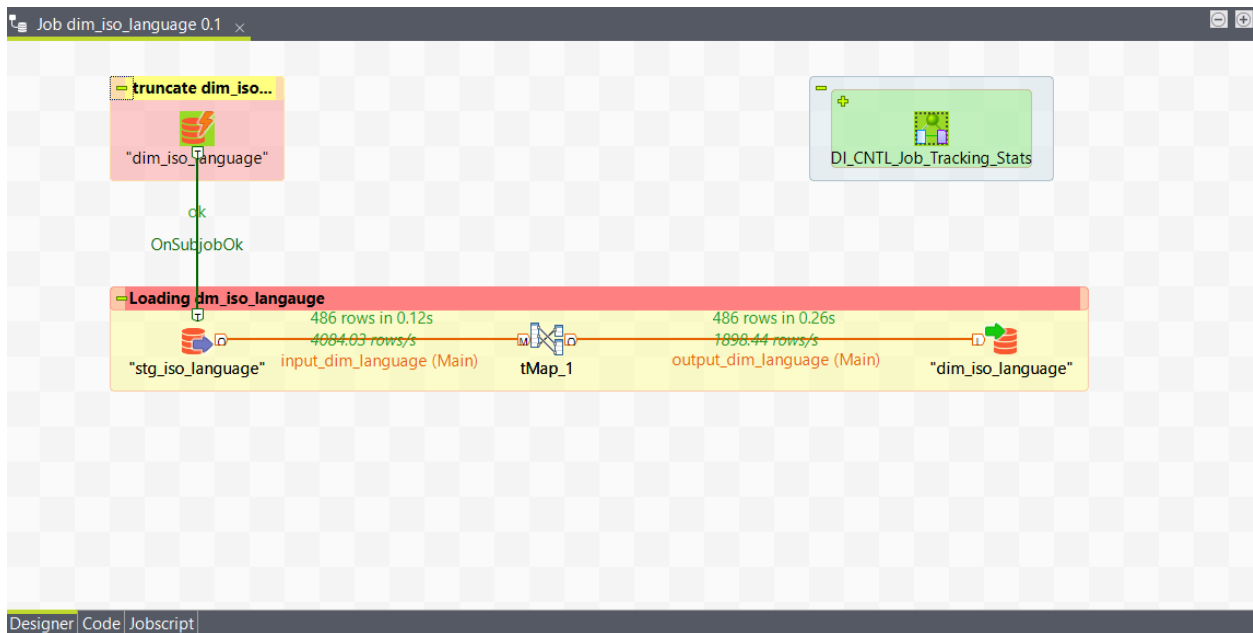


Table: dim_imdb_genres

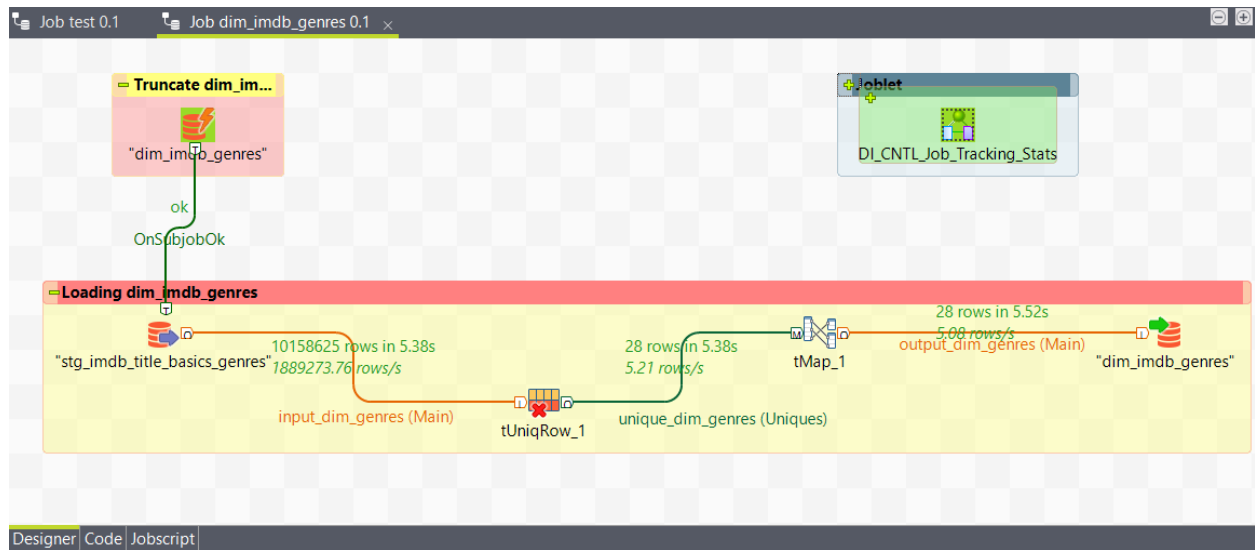


Table: dim_iso_country

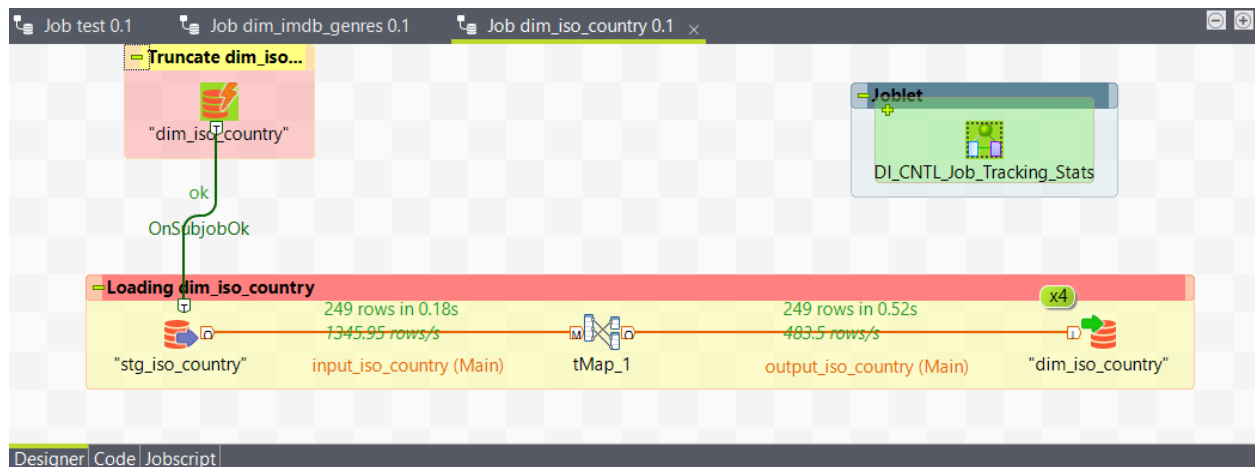


Table: dim_imdb_brand

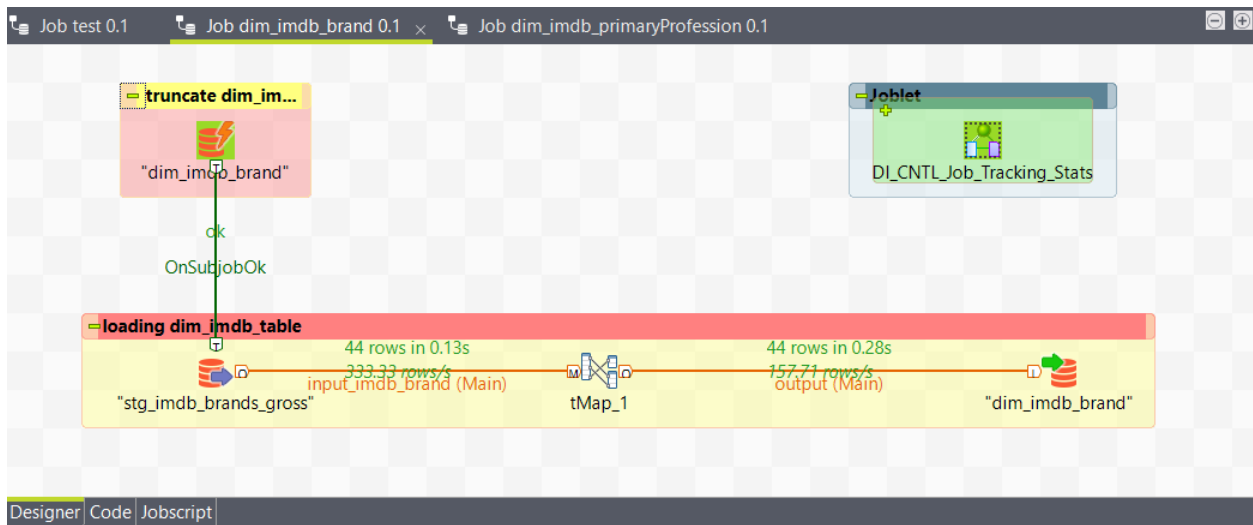


Table: dim_imdb_job_category

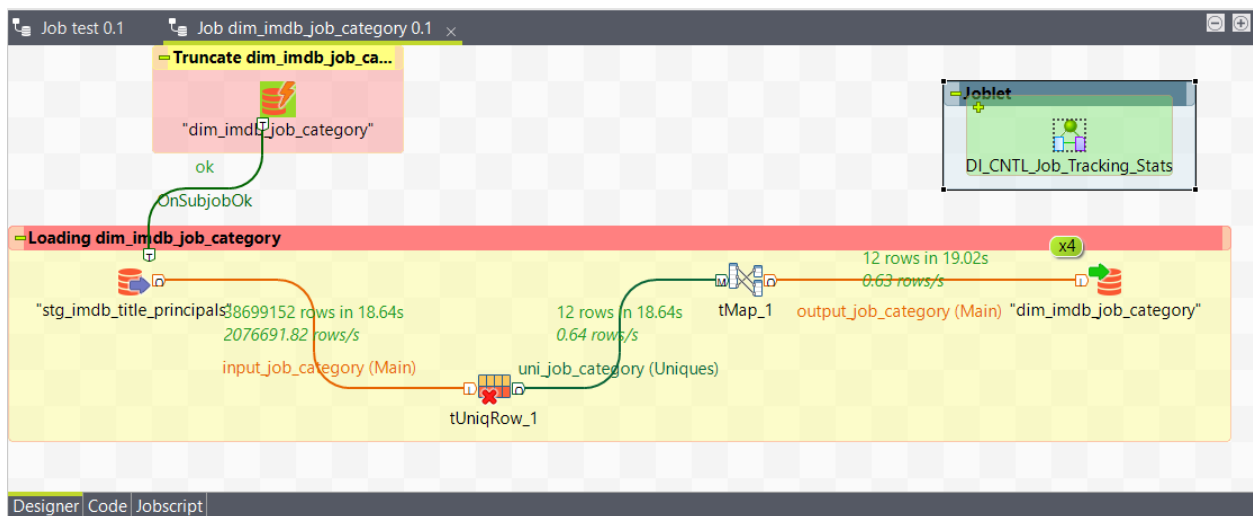


Table: dim_imdb_name_basics_PrimaryProfession

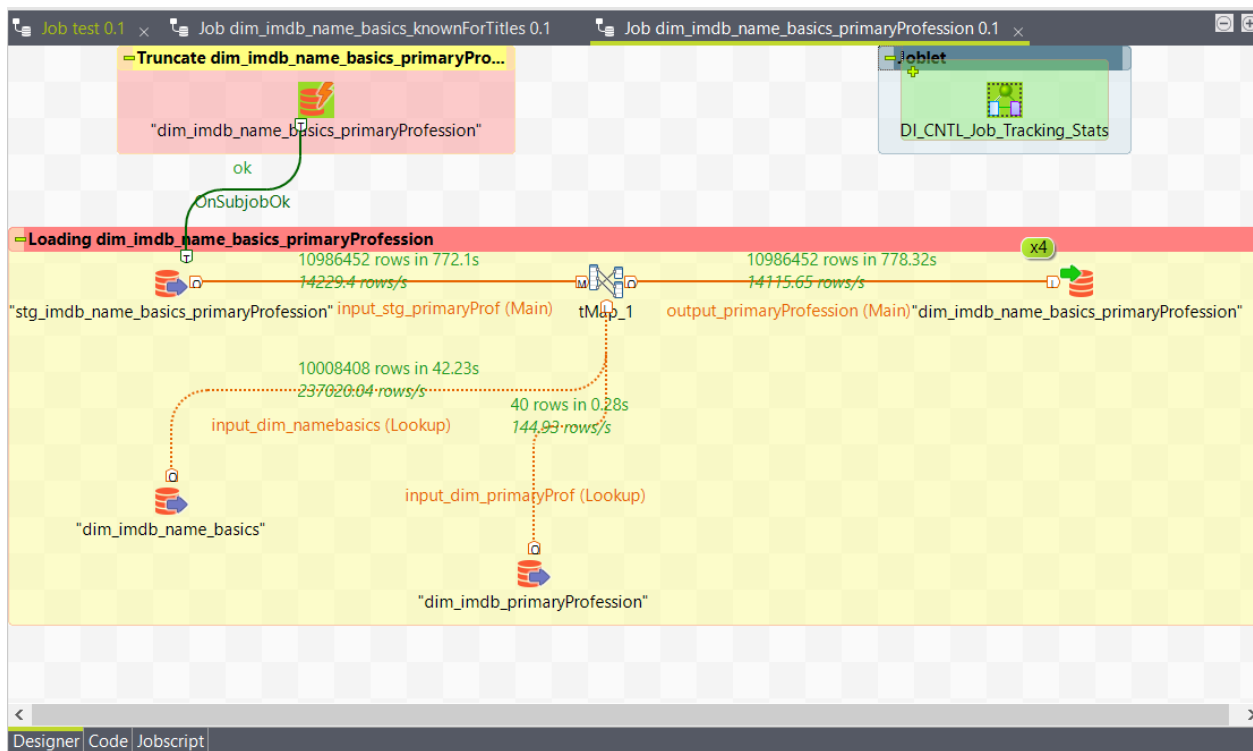


Table: dim_imdb_name_basics_knownForTitles

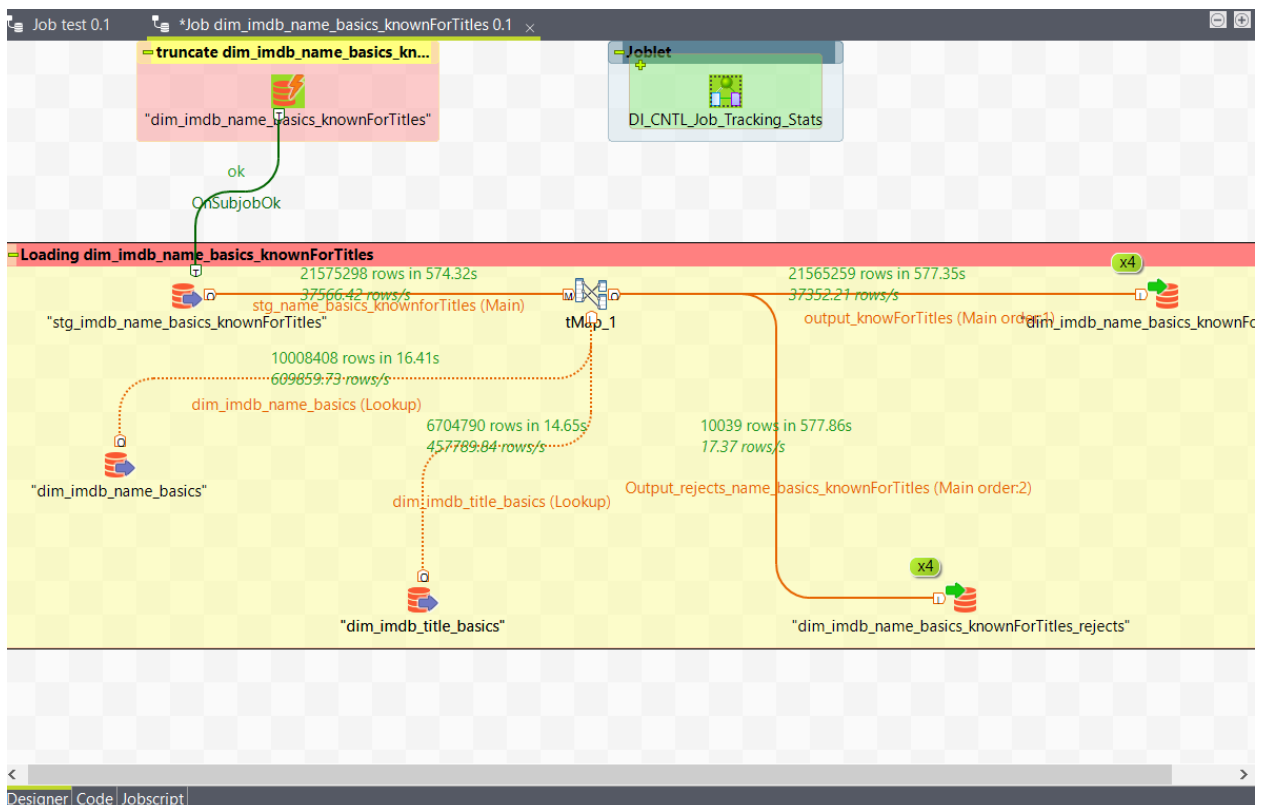


Table: dim_imdb_title_principals

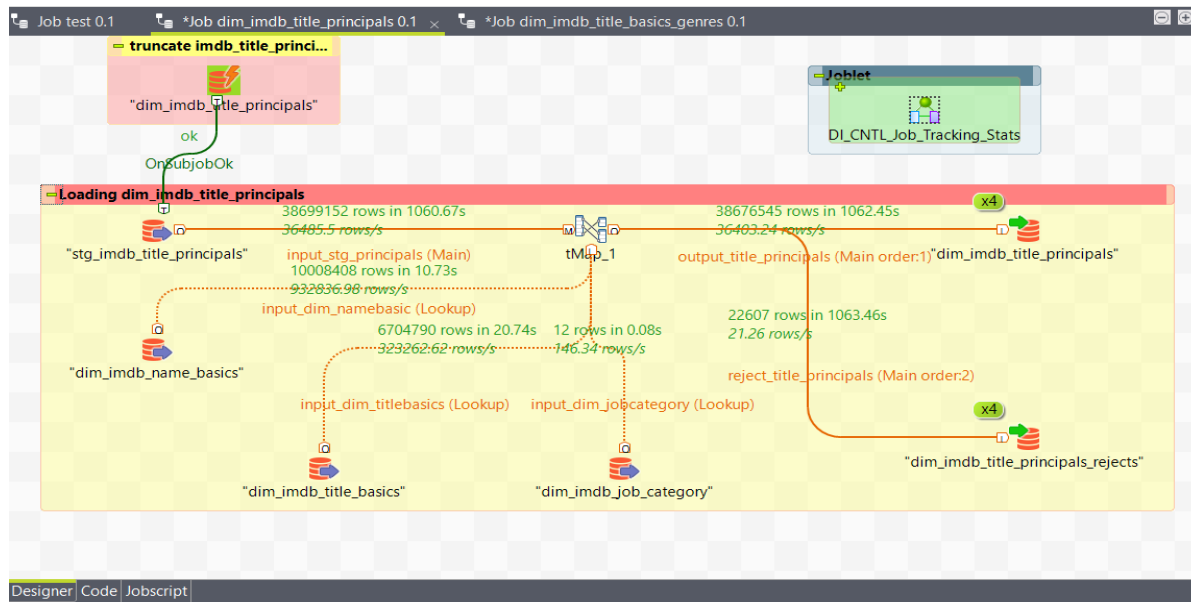


Table: dim_imdb_title_basics_genres

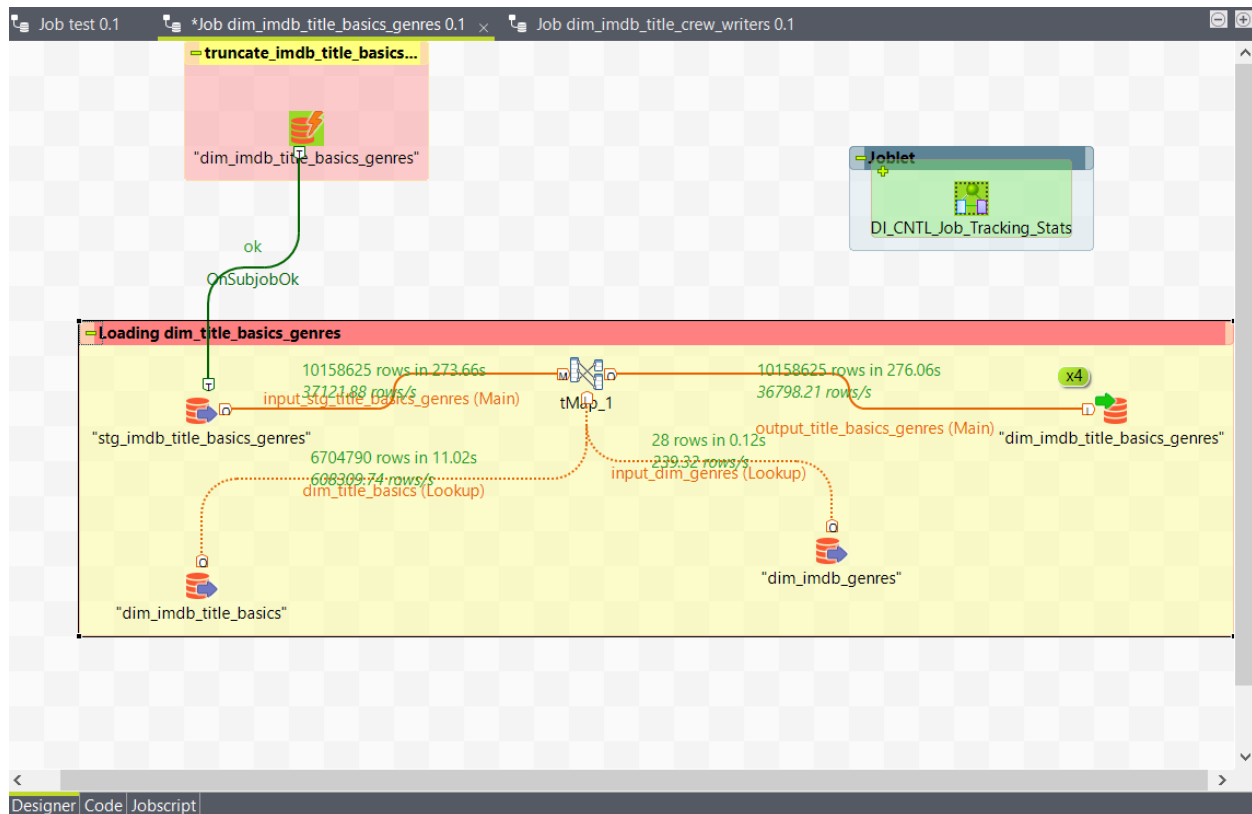


Table: dim_imdb_title_crew_writers

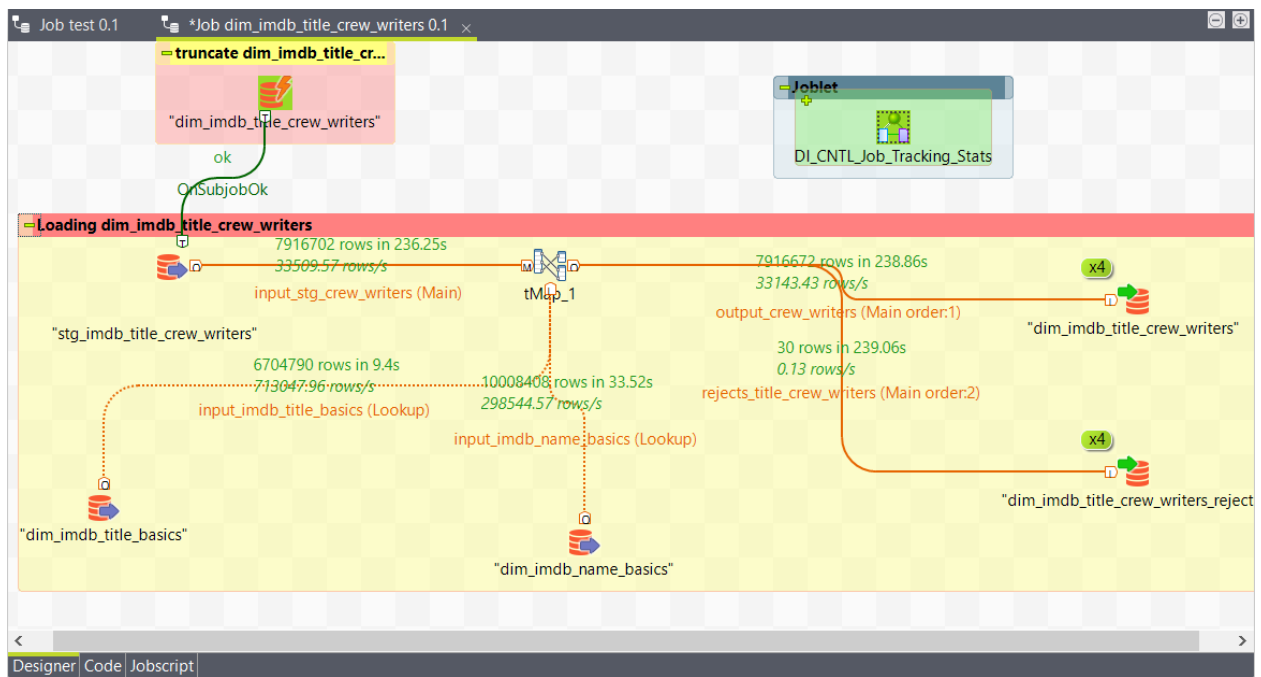


Table: dim_imdb_title_crew_writers

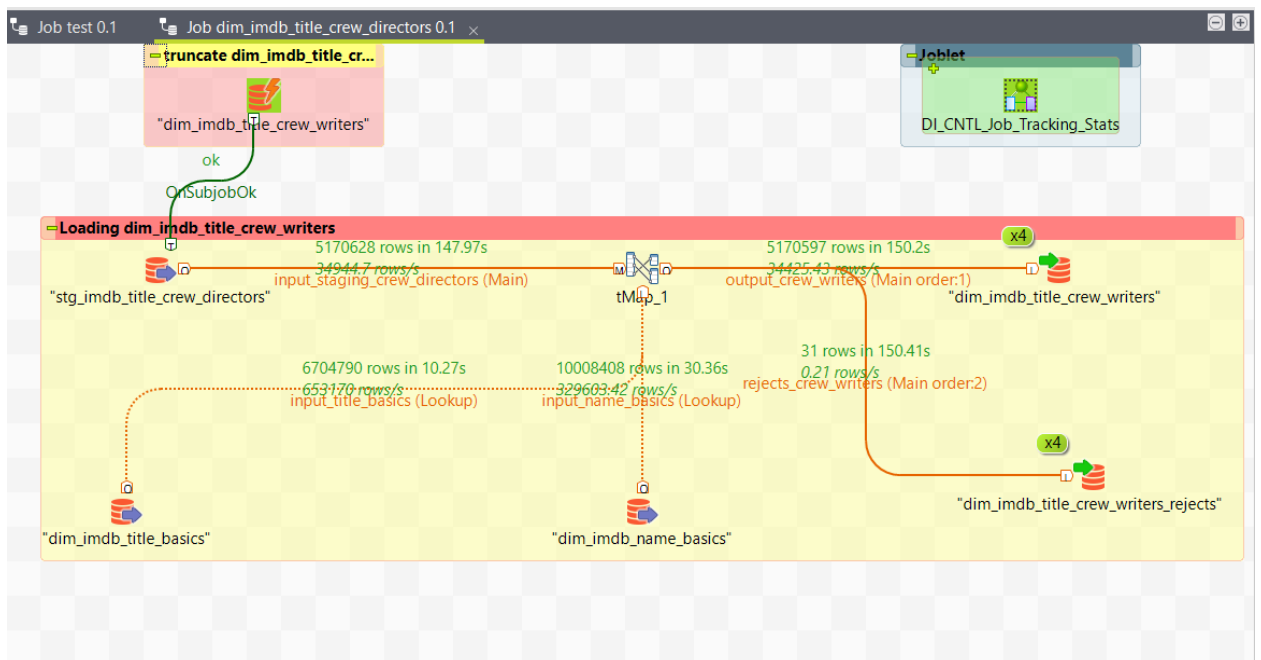
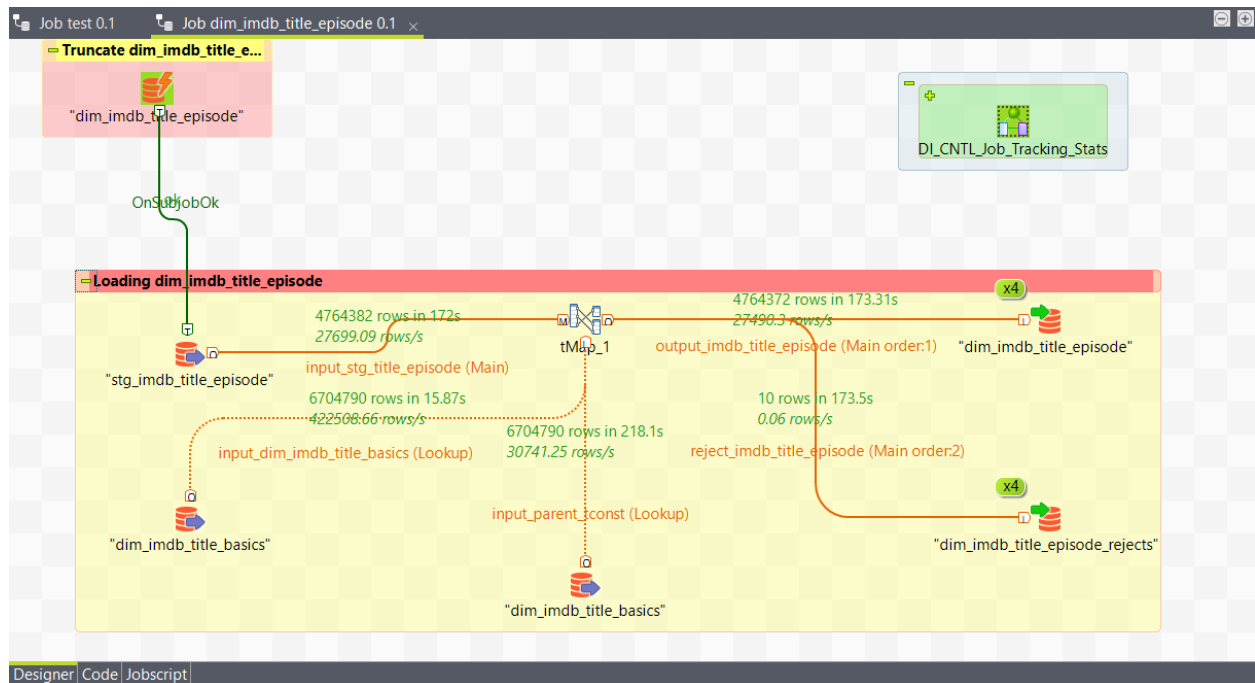


Table: dim_imdb_title_episode



Fct_ml_genome_tags

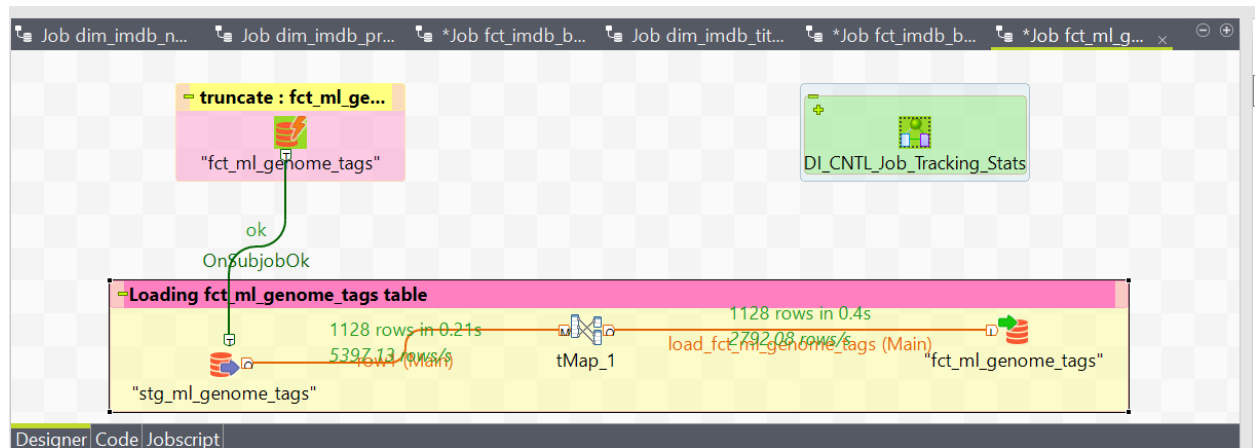


Table: fct_ml_genome_scores & fct_ml_genome_scores_rejects

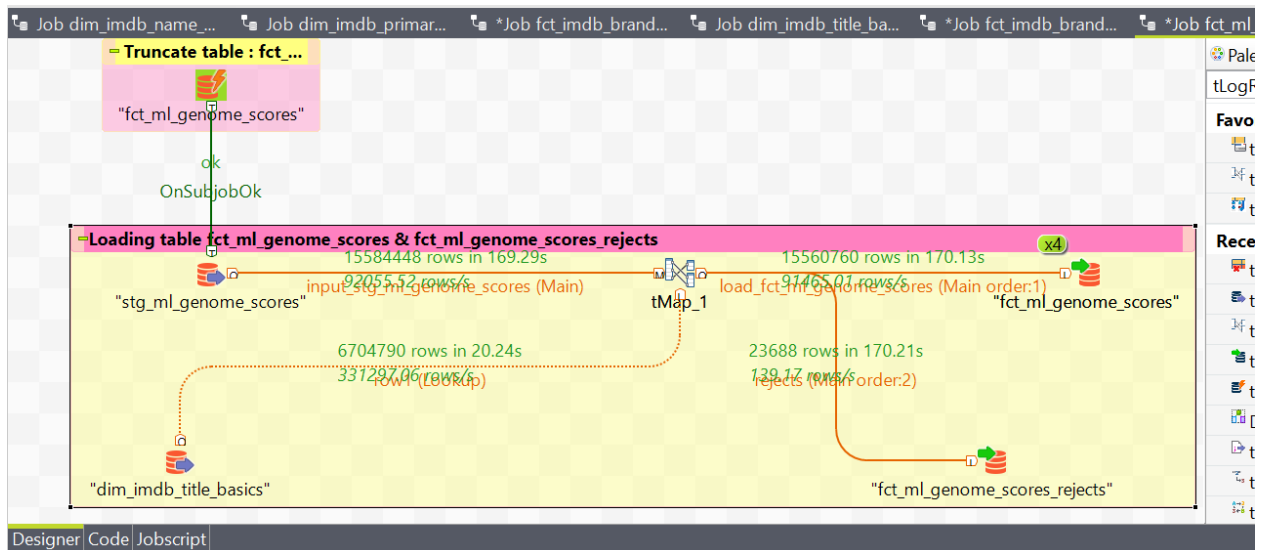
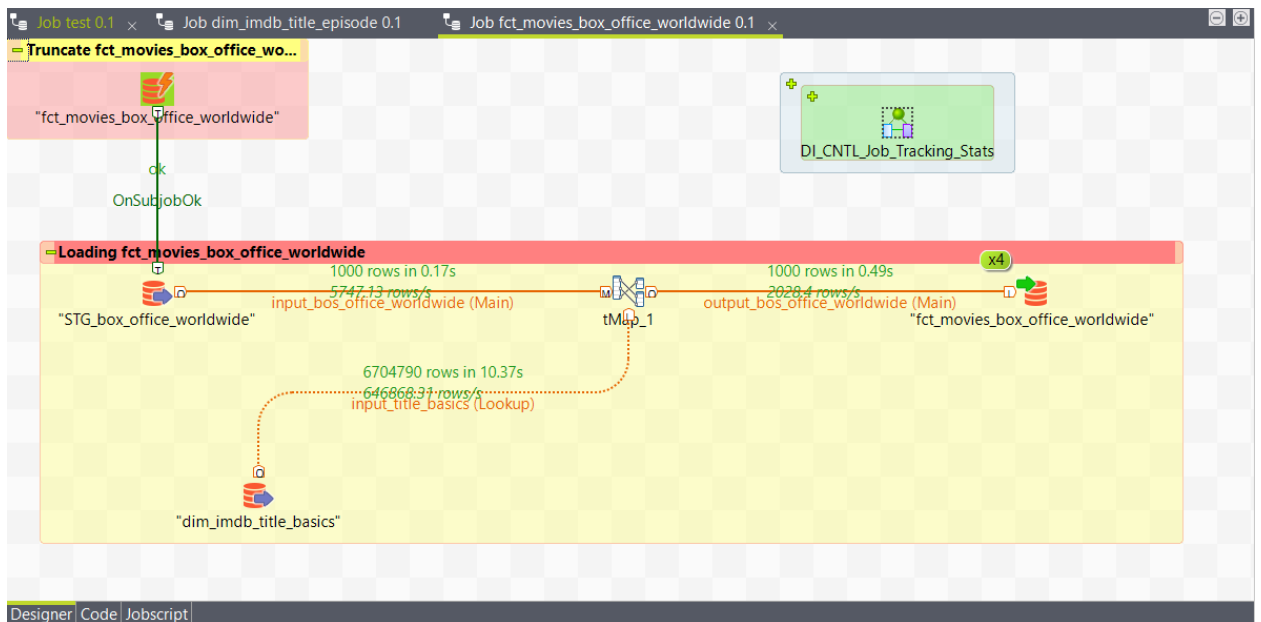
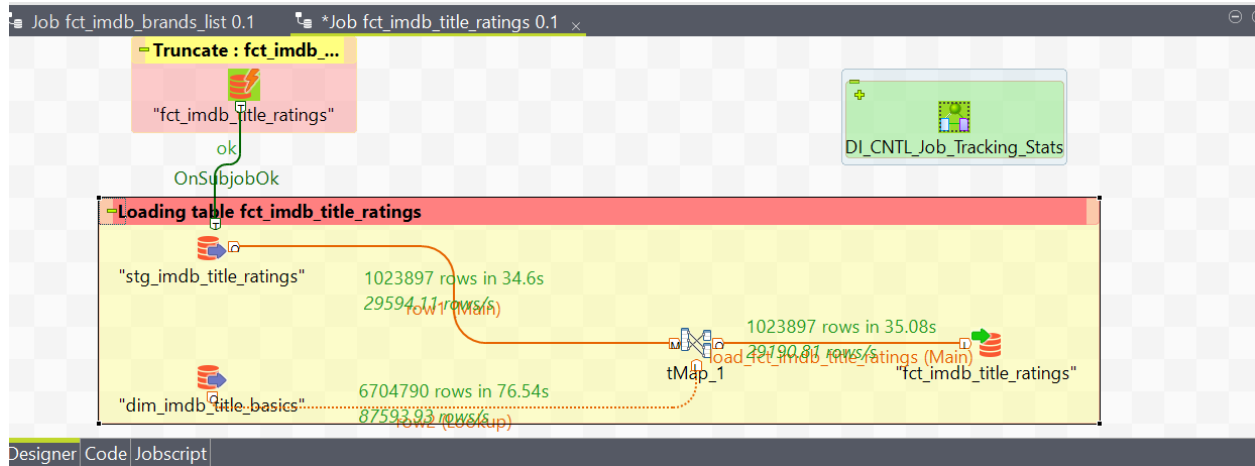


Table: fct_movies_box_office_worldwide



Fct_imdb_title_ratings



Fct_ml_ratings & rejects

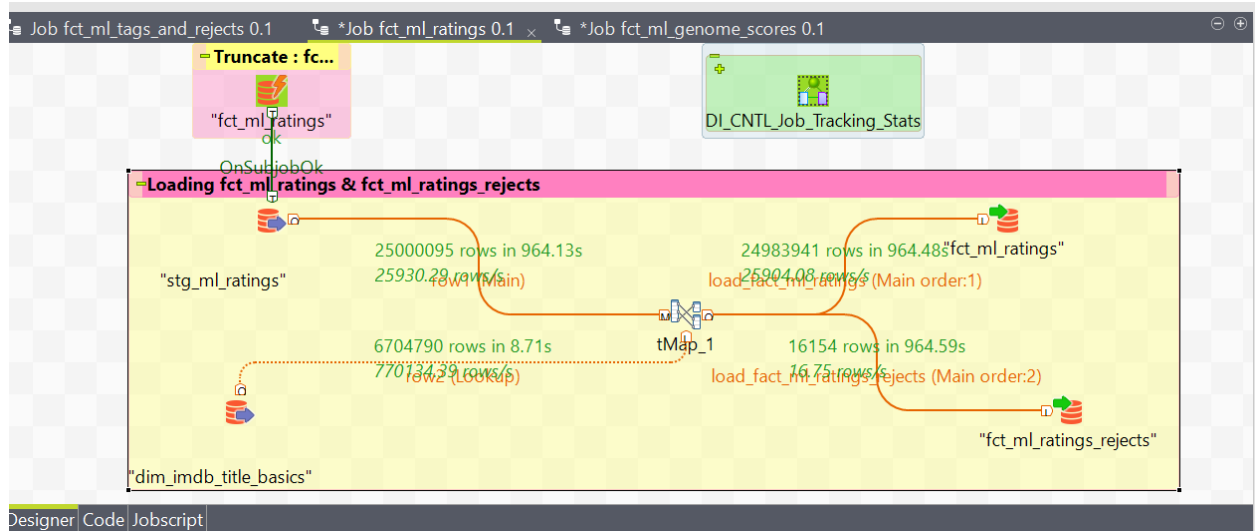


Table: fact_numbers_daily_box_office

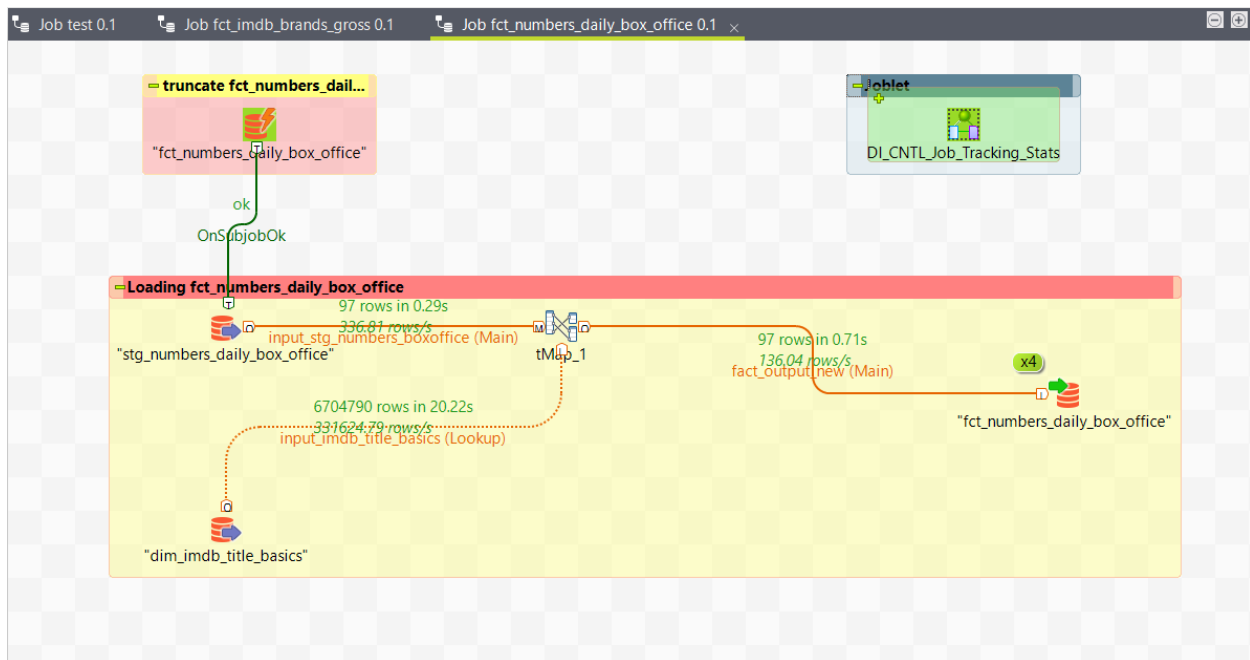


Table: fact_imdb_brands_gross

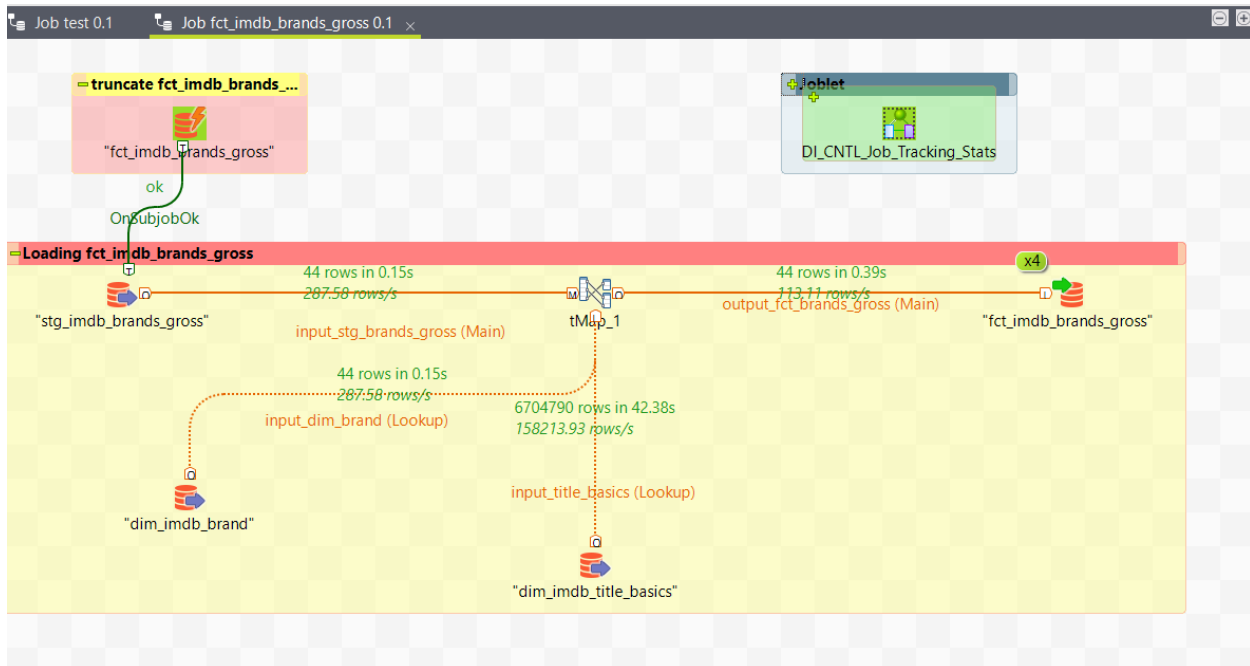
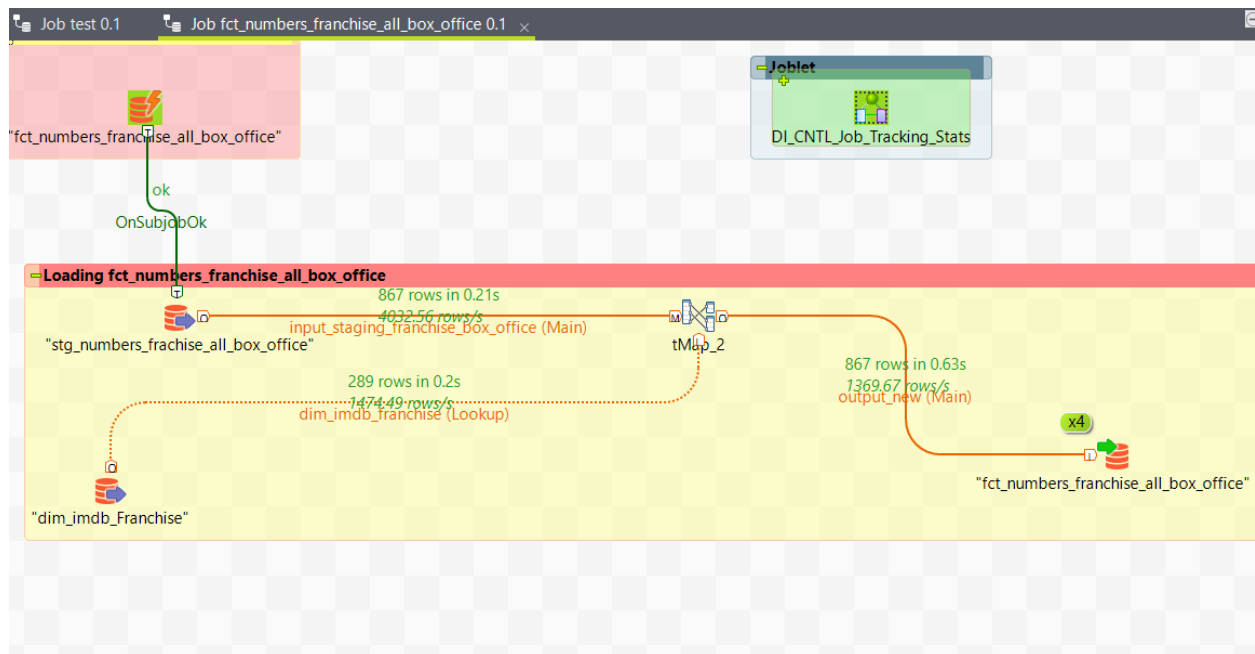


Table: fct_numbers_franchise_all_box_office



Fct_ml_tags & rejects

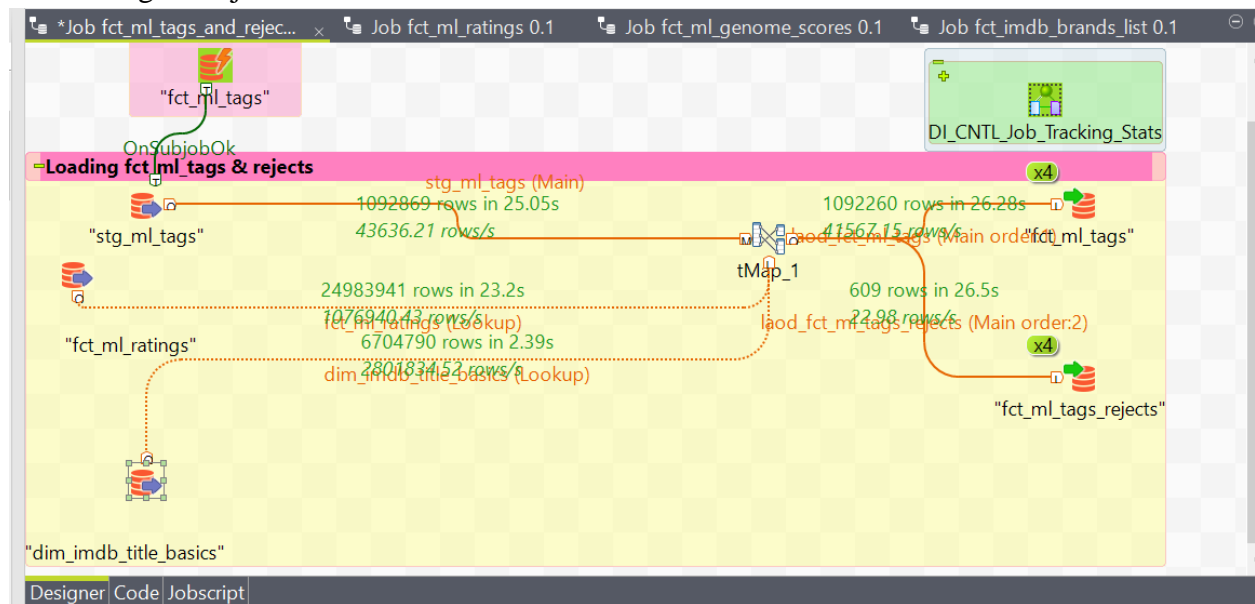


Table: fct_numbers_franchise_movies_box_office

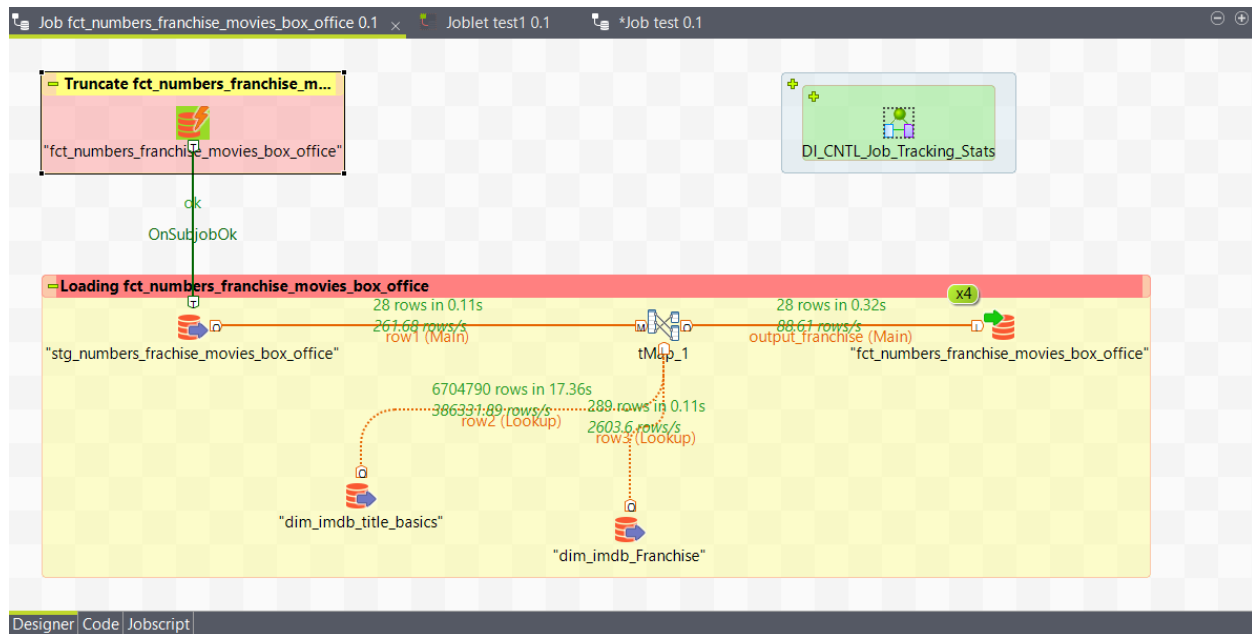


Table: fct_imdb_franchises_gross

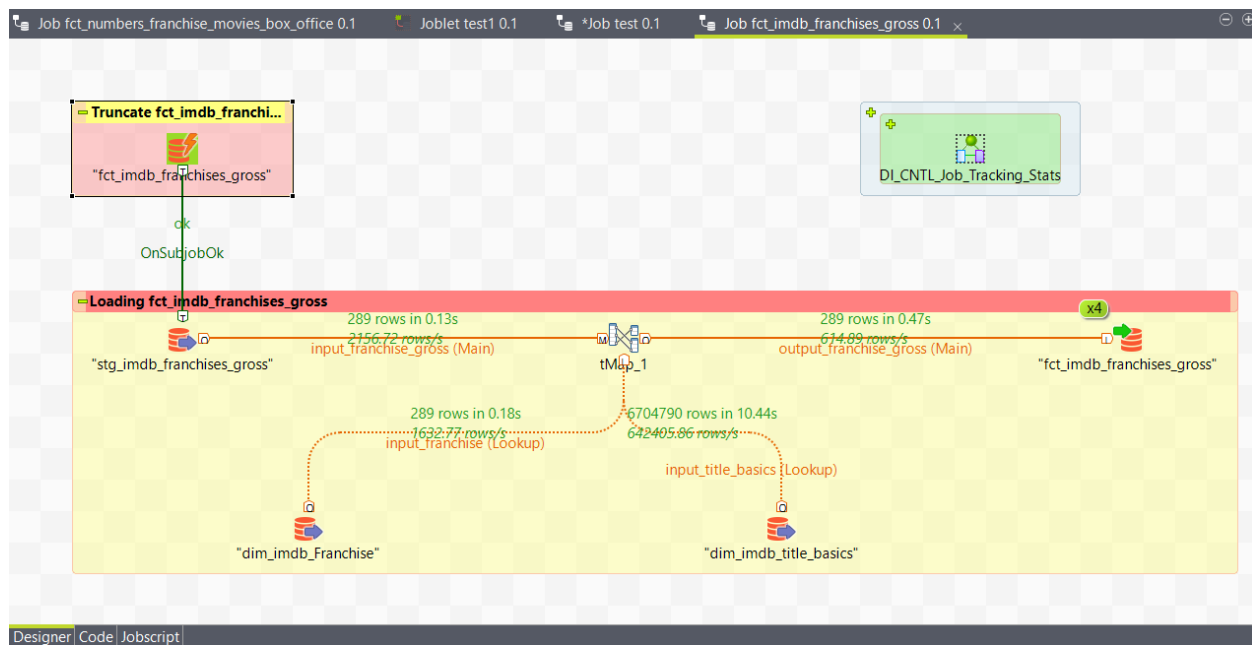


Table fct_imdb_franchises_list

