

# **IMDB Movies Analysis Project:**

## --Amey Arun Parange

### **Introduction**

Data Profiling and Analysis Report (Alteryx)

#### 2.1 Data Profiling

- Overview of Data Profiling
  - Purpose and Significance
  - Tools and Techniques Used (Alteryx)
- #### 2.2 Analysis Report (Inference)
- Exploratory Data Analysis (EDA)
  - Key Findings and Insights
  - Data Quality Assessment

### **Mapping Document**

#### 3.1 Source and Target Mapping

- Overview of Mapping
- Template Explanation
- Source Systems
- Target Systems

#### 3.2 Data Flow Diagrams

- Visual Representation of Data Flows
- Transformation Processes

- Connection Points

### **Data Modeling**

#### 4.1 Movie Titles Dimension Table

- Type 2 Implementation
  - Attributes and Hierarchies
  - Historical Data Handling
- #### 4.2 Entity-Relationship Diagram (ERD)
- Overview of ERD
  - Relationships and Cardinalities
  - Entity Descriptions

### **ETL Mappings Using Talend**

#### 5.1 Talend Introduction

- Overview of Talend
- Purpose in ETL Process

#### 5.2 Metadata Based Connections

- Explanation of Metadata Connections
- Implementation in Talend

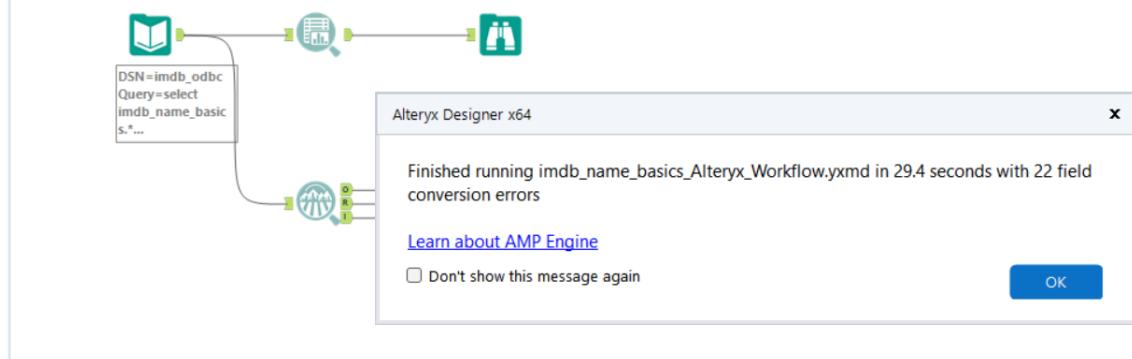
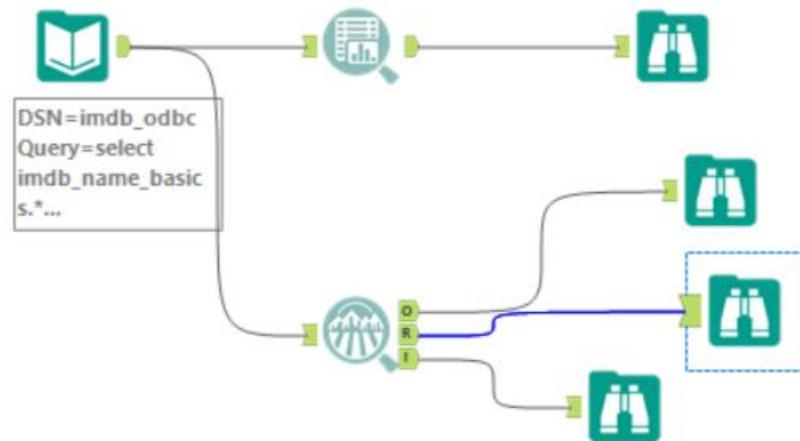
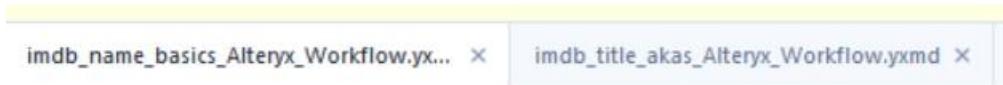
#### 5.3 Contexts in Talend

- Introduction to Contexts
  - Utilization in ETL Jobs
- 5.4 Environment Setup
- Configuration and Setup Steps
  - Dependencies and Requirements

## Data Profiling using Alteryx:

Workflow Screenshots : Up and running :

For Table **imdb\_name\_basics**

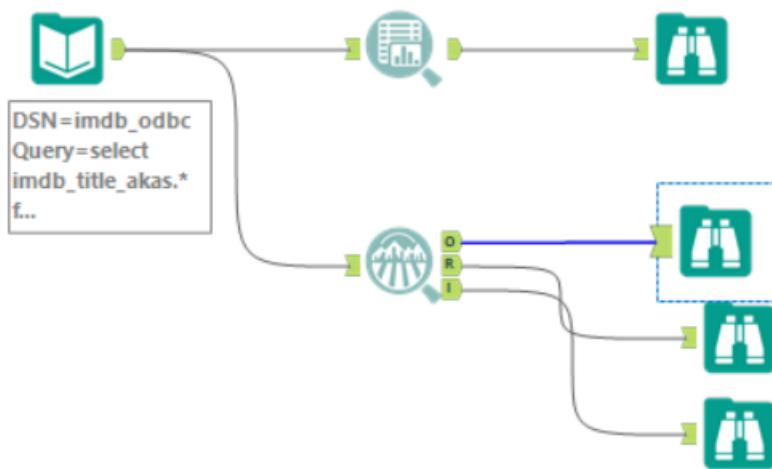


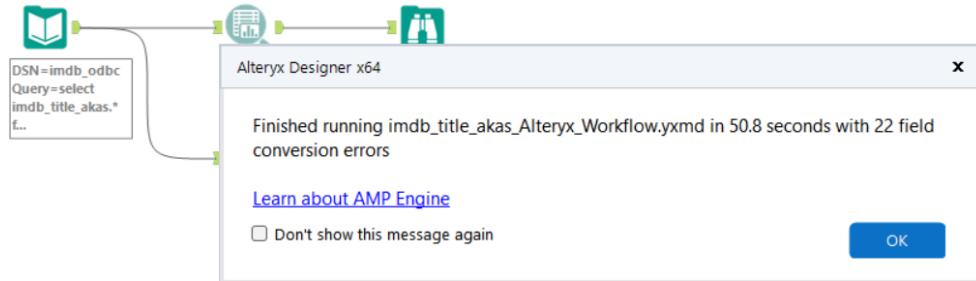
Name	Datatype	% missing	Unique values	Observation	Min Value	Max Value
birthYear	String	0%	376	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1,208,042
deathYear	String	0%	322	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1,358,373
knownForTitles	String	0%	1,172,487	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	76,778
nconst	String	0%	1,454,602	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1
primaryName	String	0%	1,349,758	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	36
primaryProfession	String	0%	12162	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together	1	248,050

DI_Create_DT	String	0%	76			
--------------	--------	----	----	--	--	--

For Table **imdb\_name\_akas**

S\_Alteryx\_Workflow.yx... X      imdb\_title\_akas\_Alteryx\_Workflow.yxmd X      imdb\_title\_basics\_Alteryx\_Workflow.yx

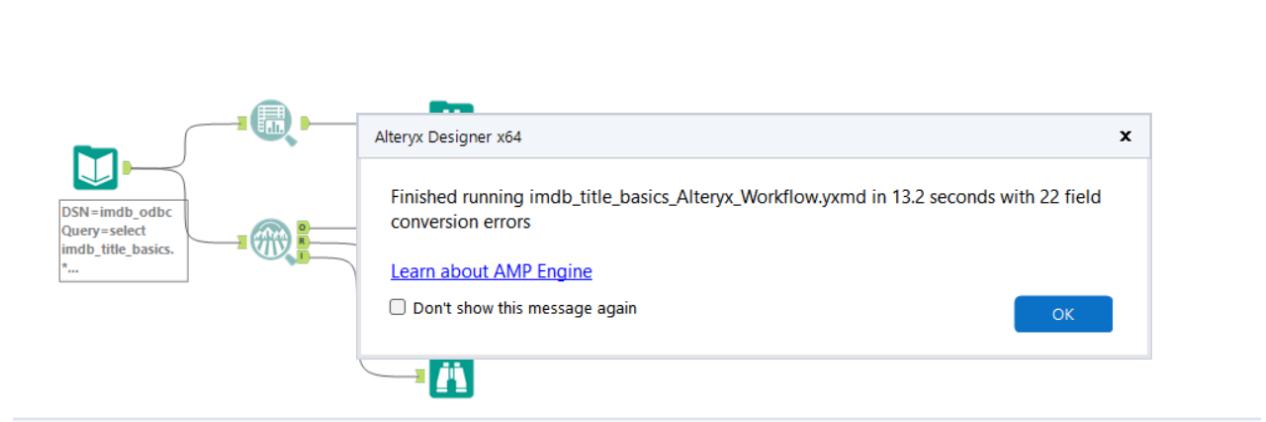
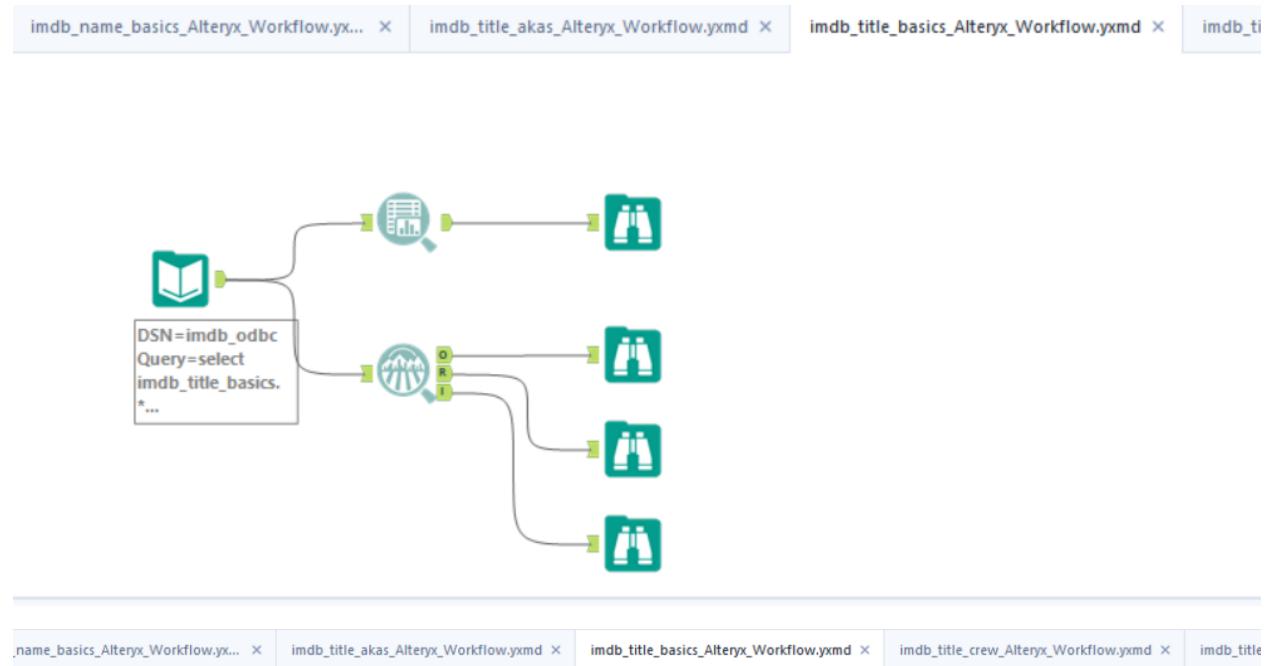




Name	Datatype	% missing	Unique values	Observation	Min Value	Max Value
Ordering	Numeric	0%	149		1	
Language	String	0%	101	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	2177430
Attributes	String	0%	130	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	2486126
title	String	0%	1497026	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels	1	189

				together.		
types	String	0%	24	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1468665
region	String	0%	235	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	531054
titleId	String	0%	592887	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.		149
isOriginalTitle	String	0%	3	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	37	2121643
DI_Create_DT	Date Time		126	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.		

## imdb\_title\_basics

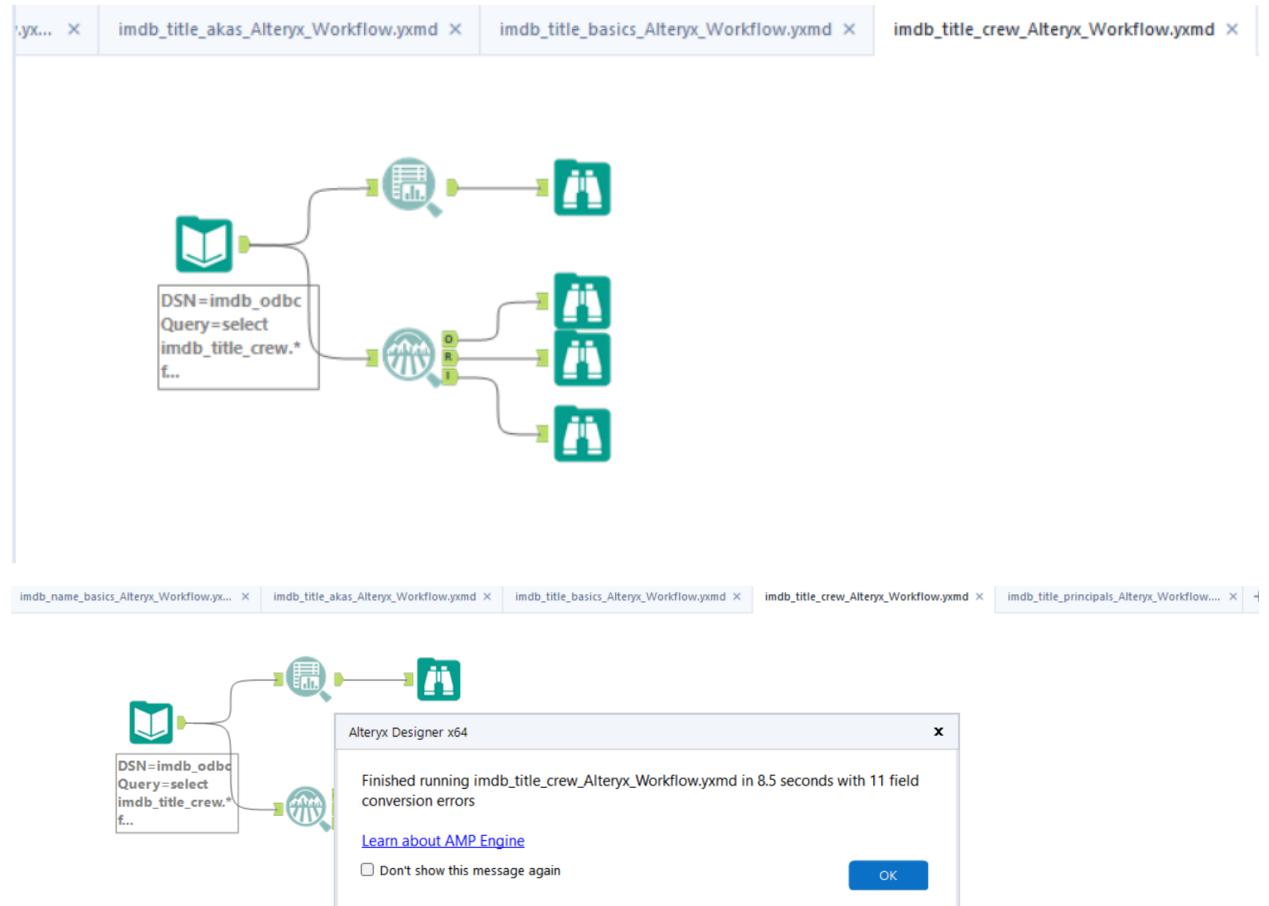


	Datatype	% missing	Unique values	Observation	Min Value	Max Value
isAdult	Numeric	0%	2	This field has a small number of unique values, and appears to be a categorical field. Consider changing the field data type to "string".	1	

DI_Create_DT	Date	0%	40	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.		
genres	String	0%	1,459	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	112,401
runtimeMinutes	String	0%	489	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	225,251
primaryTitle	String	0%	526,041	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	52
tconst	String	0%	607,423	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1
originalTitle	String	0%	537,134	Some values of this field have a small	1	48

				number of value counts. If Appropriate, consider combining some value levels together.		
startYear	String	0%	135	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	79,022
endYear	String	0%	1	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	607,423	607,423
titleType	String	0%	1		607,423	607,423

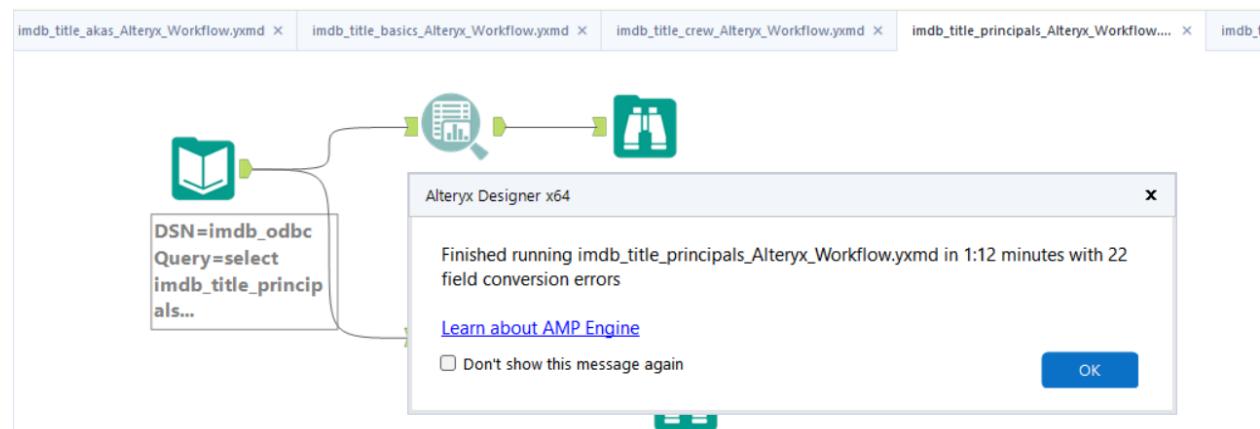
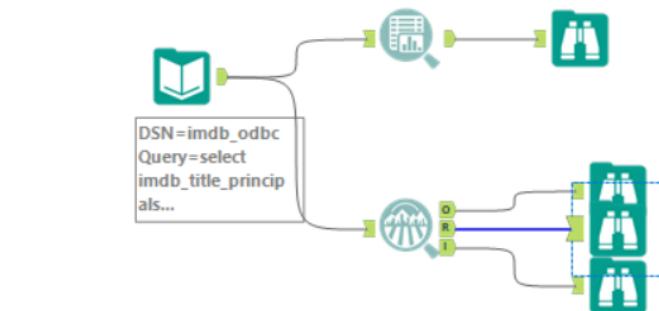
## imdb\_title\_crew



Name	Datatype	% missing	Unique values	Null values	Observation	Min Value	Max Value
DI_Create_DT	Date	0%	26		Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.		
directors	String	0%	320,758		Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	63,673

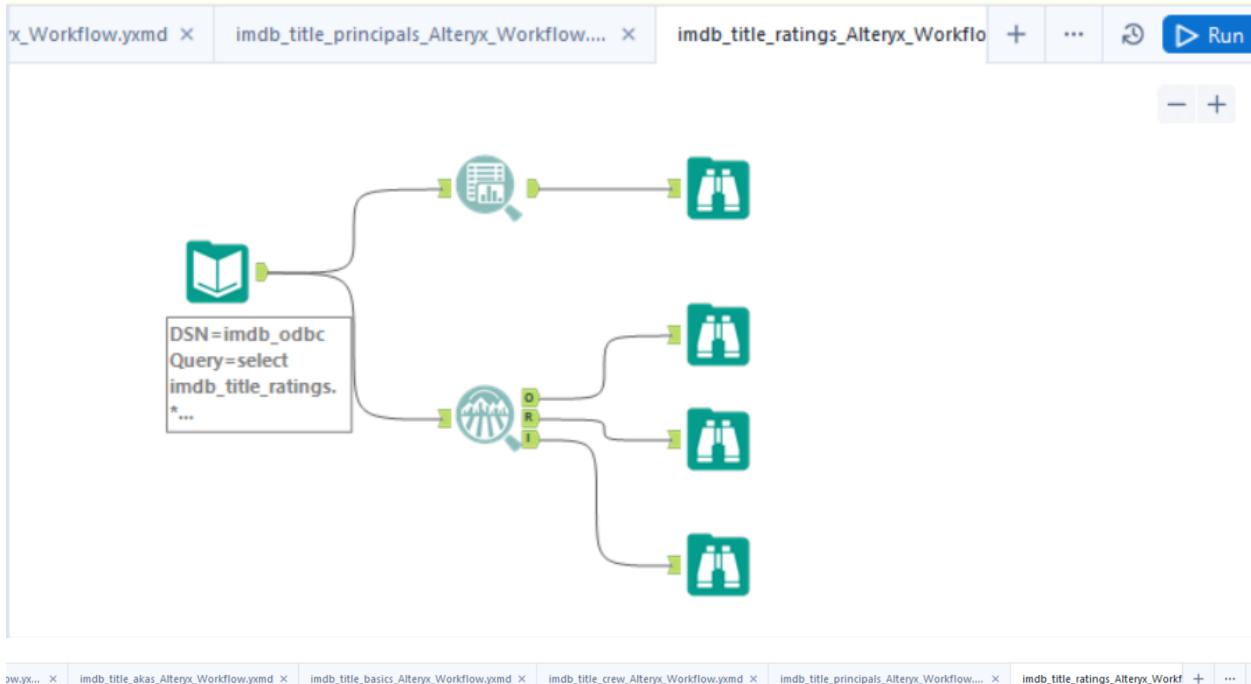
writers	String	0%	607,423		Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	48,775
tconst	String	0%			Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1

### imdb\_title\_principals

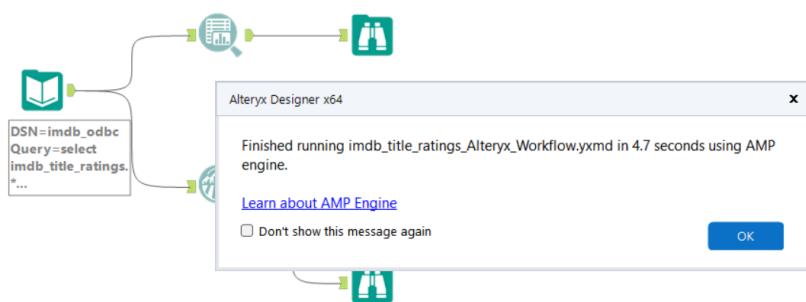


Name	Datatype	% missing	Unique values	Null values	Observation	Min Value	Max Value
ordering	Numeric	0%	10		This field has a small number of unique values, and appears to be a categorical field. Consider changing the field data type to "string".	1.000	10.000
DI_Create_DT	Date	0%	152		Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.		
category	String	0%	12		Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	60	1,176,368
job	String	0%	19,815		Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	3,550,474
nconst	String	0%	1,450,993		Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1,009
characters	String	0%	703,924			1	2,813,735
tconst	String	0%	591,300			1	10

## imdb\_title\_ratings



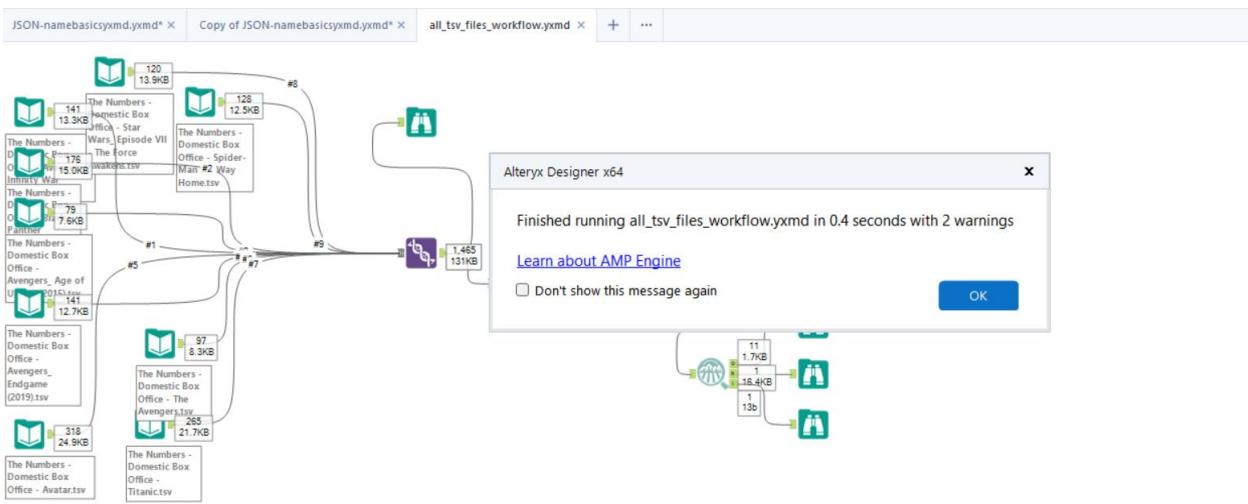
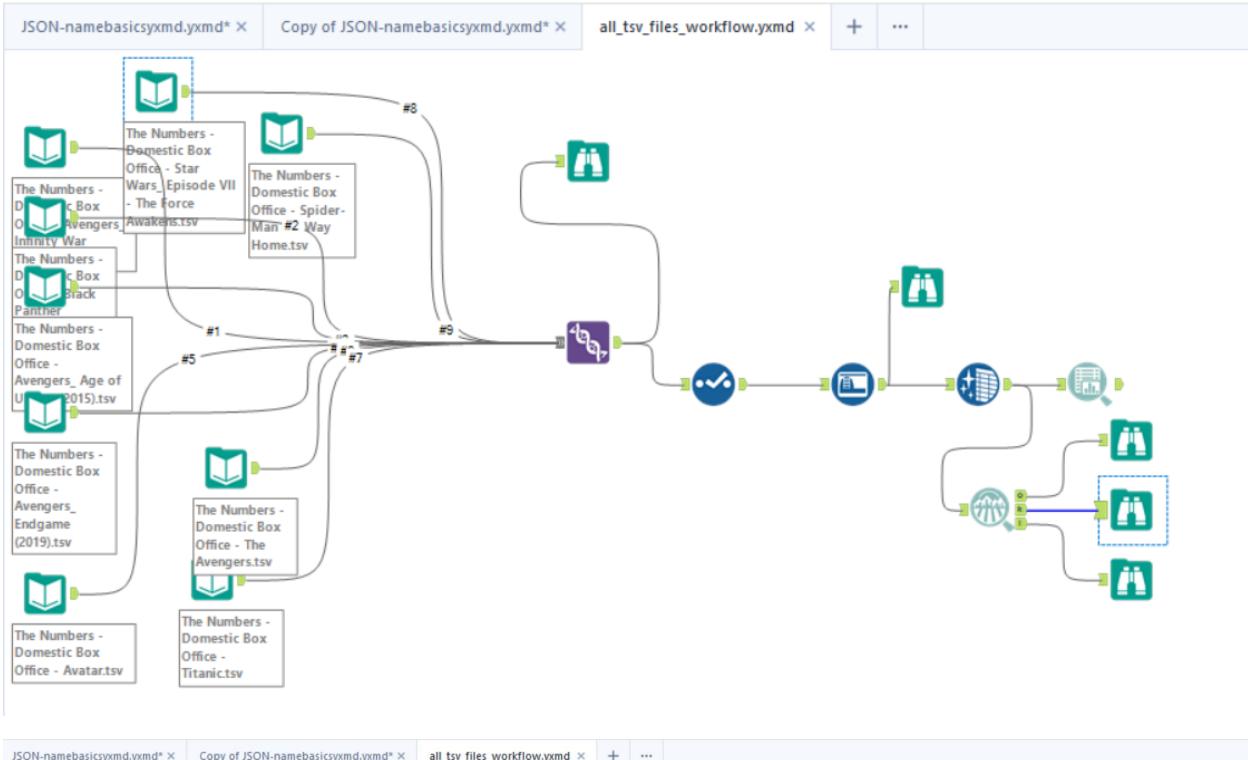
dw.yx... x imdb\_title\_akas\_Alteryx\_Workflow.yxmd x imdb\_title\_basics\_Alteryx\_Workflow.yxmd x imdb\_title\_crew\_Alteryx\_Workflow.yxmd x imdb\_title\_principals\_Alteryx\_Workflow.... x imdb\_title\_ratings\_Alteryx\_Workf + ...



Name	Datatype	% missing	Unique values	Observation	Min Value	Max Value
DI_Create_DT	Date	0%	10	This field has a small number of unique values, and appears to be a categorical field. Consider changing the field data type to "string".		
averageRati	Numeric	0%	91	Some values of this field have a small	1	10

ng				number of value counts. If Appropriate, consider combining some value levels together.		
numVotes	Numeric	0%	16,972	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	5	2,572,812
tconst	String	0%	277,171	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1

## tsv file Data Profiling:

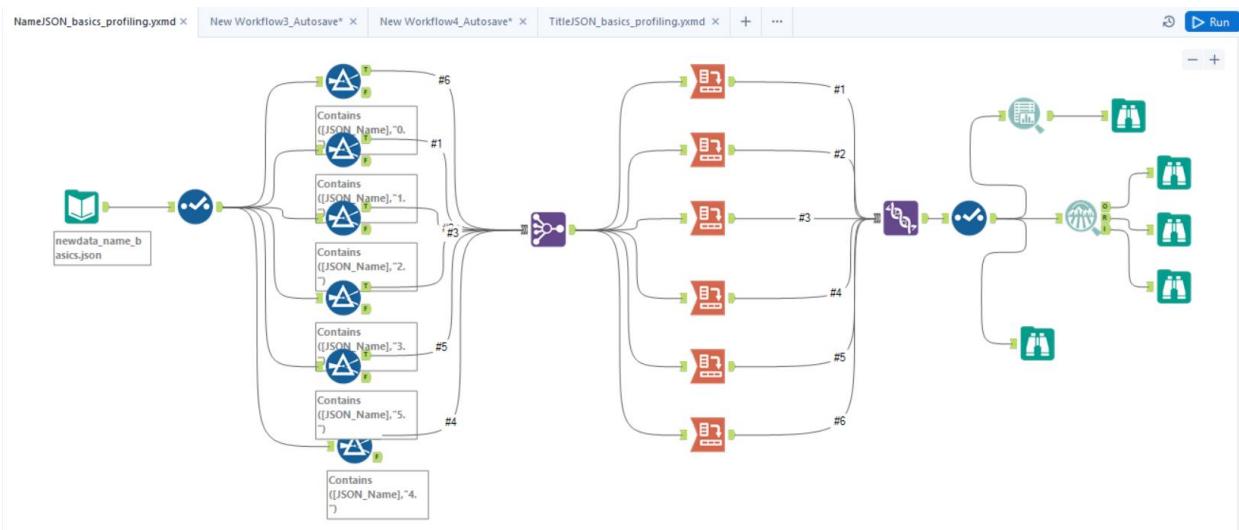
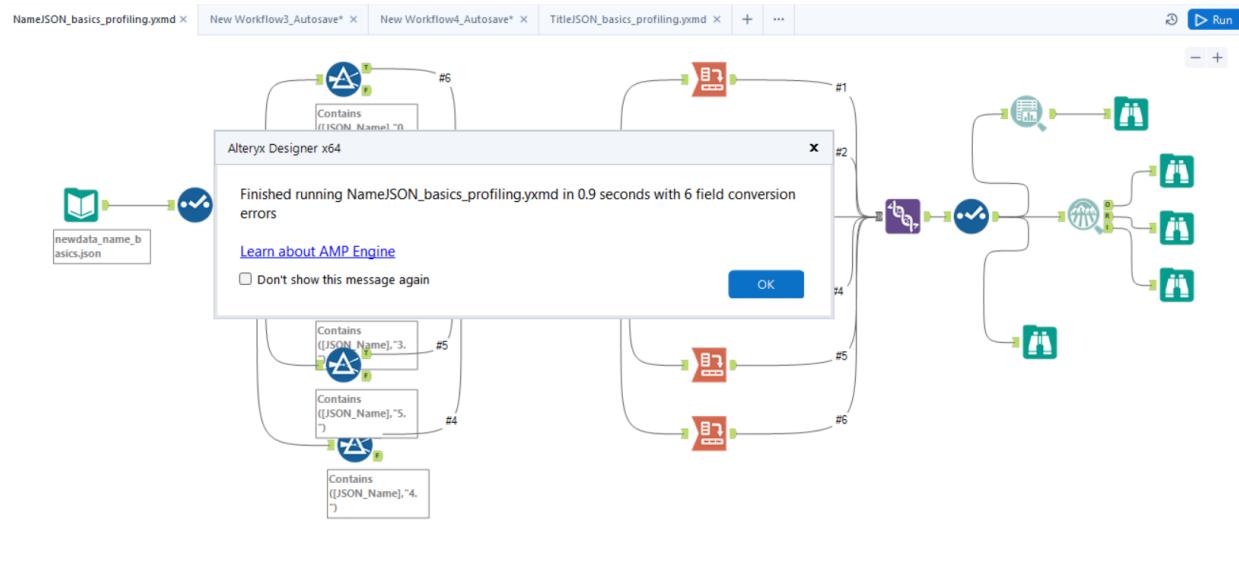


Name	Datatype	% missing	Unique values	Observation	Min Value	Max Value
Total Gross	String	0%	1,465	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	1
Rank	String	0%	20	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	473
% YD	String	1.5%	255	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	24
Date	Date	0%	1,323	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	2
Theaters	String	0%	191	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	28
Per Theater	String	0.3%	968		1	7

				Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.		
Days	String	0.2%	399	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	9
Gross	String	0.0%	1,462	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	2
%LW	String	7.0%	179	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	1	102
title	String	0.0%	9		79	318
tconst	String	8.7%	9		79	318

## JSON Files Data Profiling:

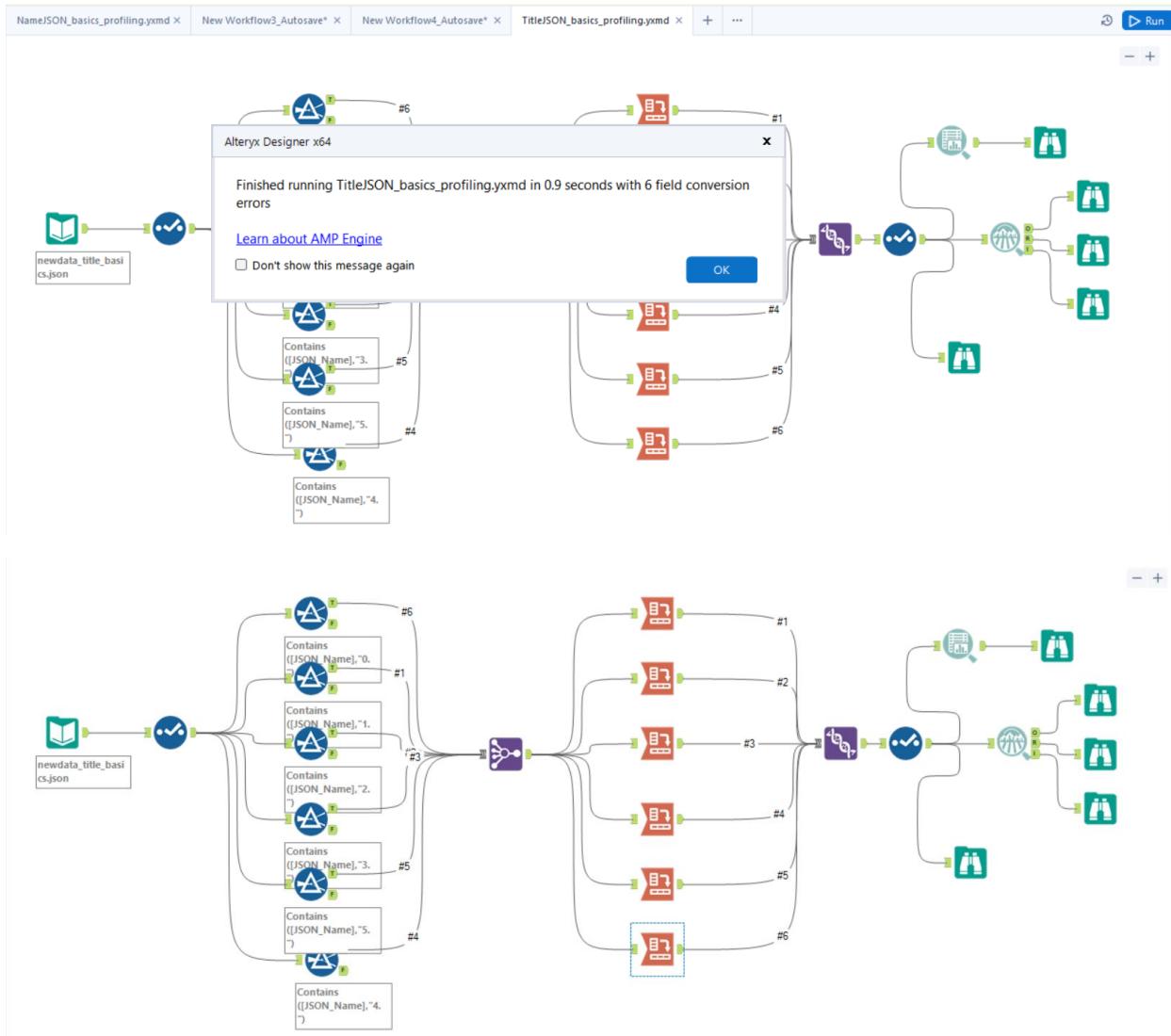
### name\_basics



Name	Datatype	% missing	Unique values	Observation	Min Value	Max Value
BIRTH Year	Numeric	0%	6	This field has a small number of unique values, and appears to be a categorical field.	1,965	1,984

				Consider changing the field data type to "string".		
Death Year	Numeric	100%	1	This field has over 10% missing values. Consider imputing these values. This field has a small number of unique values, and appears to be a categorical field. Consider changing the field data type to "string".		
Known For Titles	String	0%	6	This field has a small number of unique values, and appears to be a categorical field. Consider changing the field data type to "string".	1	1
Primary Name	String	0%	6		1	1
Primary Profession	String	0%	2		1	5
Nconst	String	0.0%	6		1	1

## newdata\_title\_basics

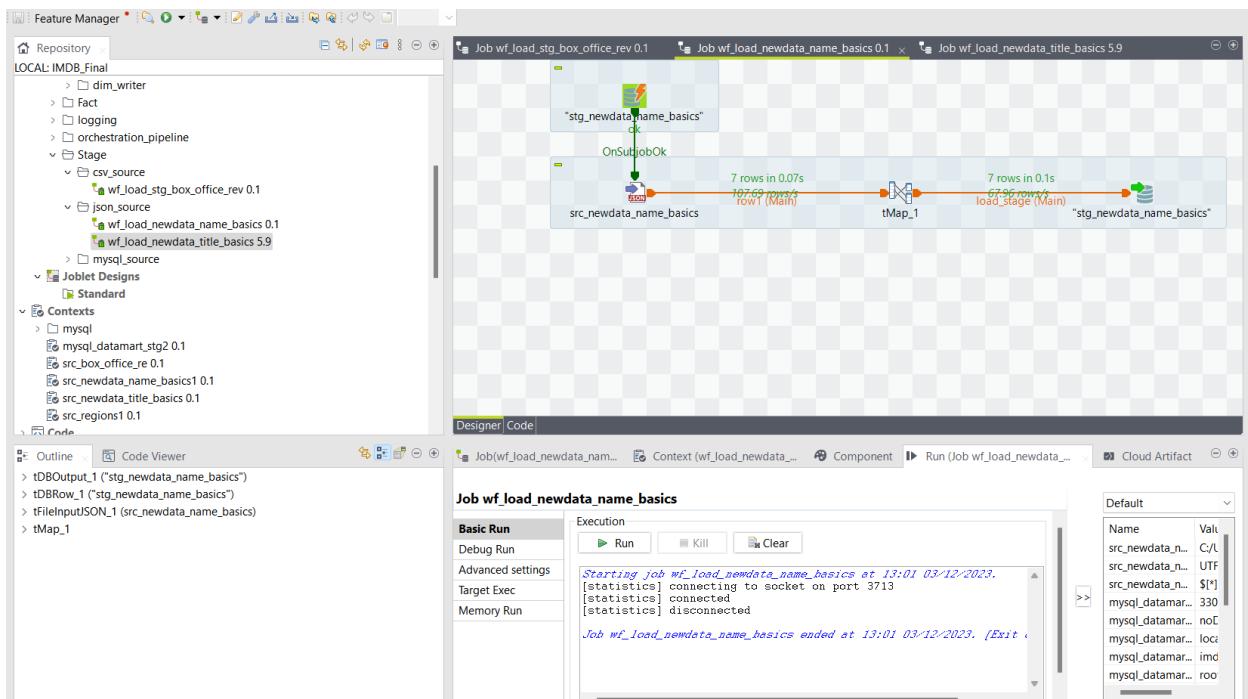
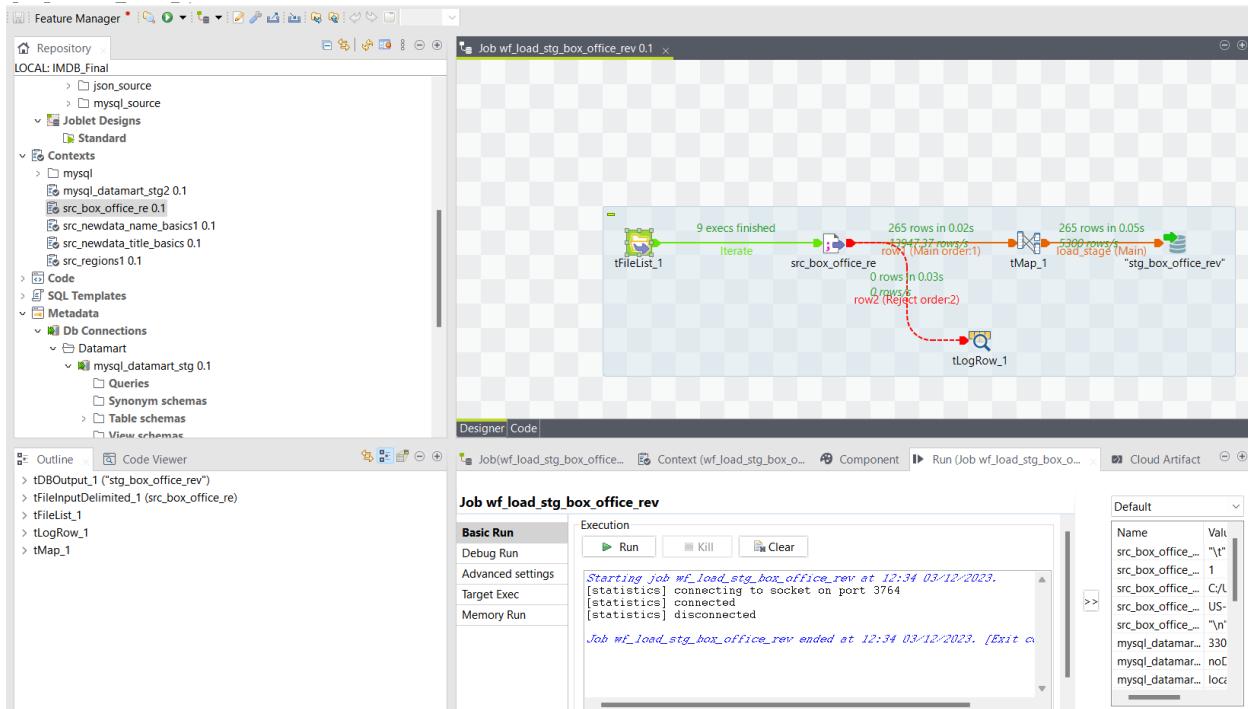


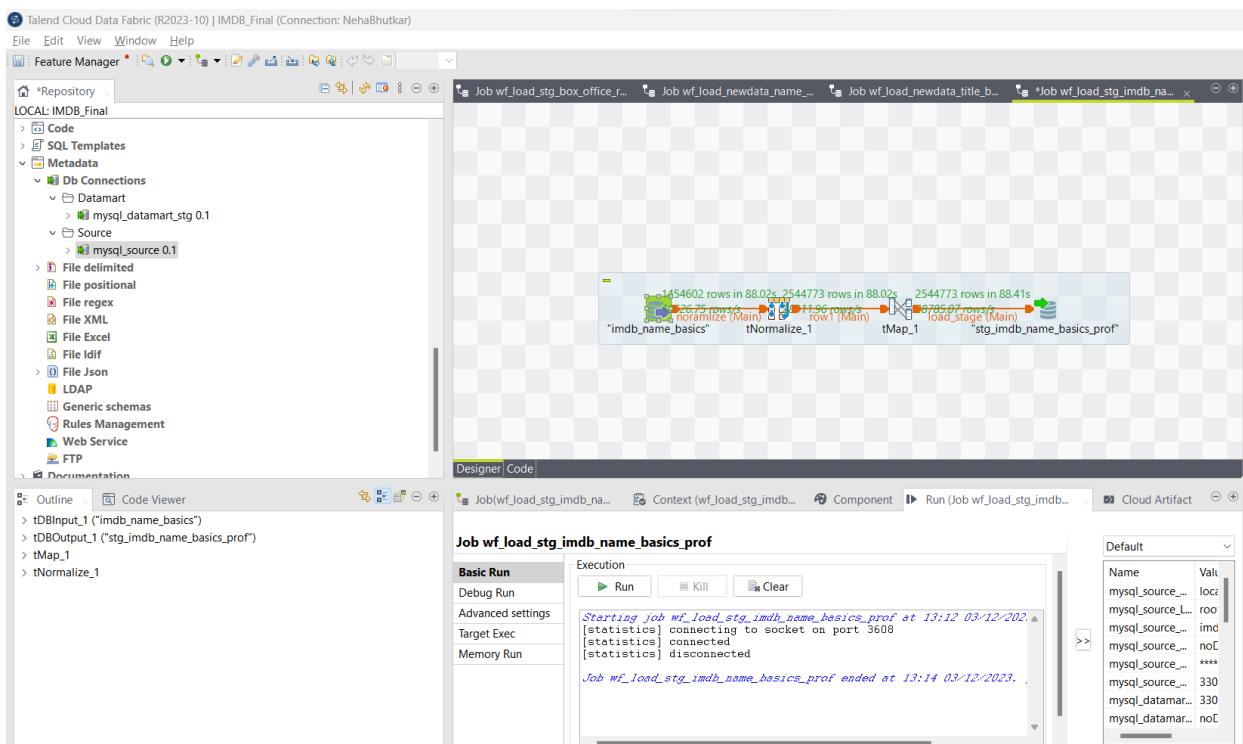
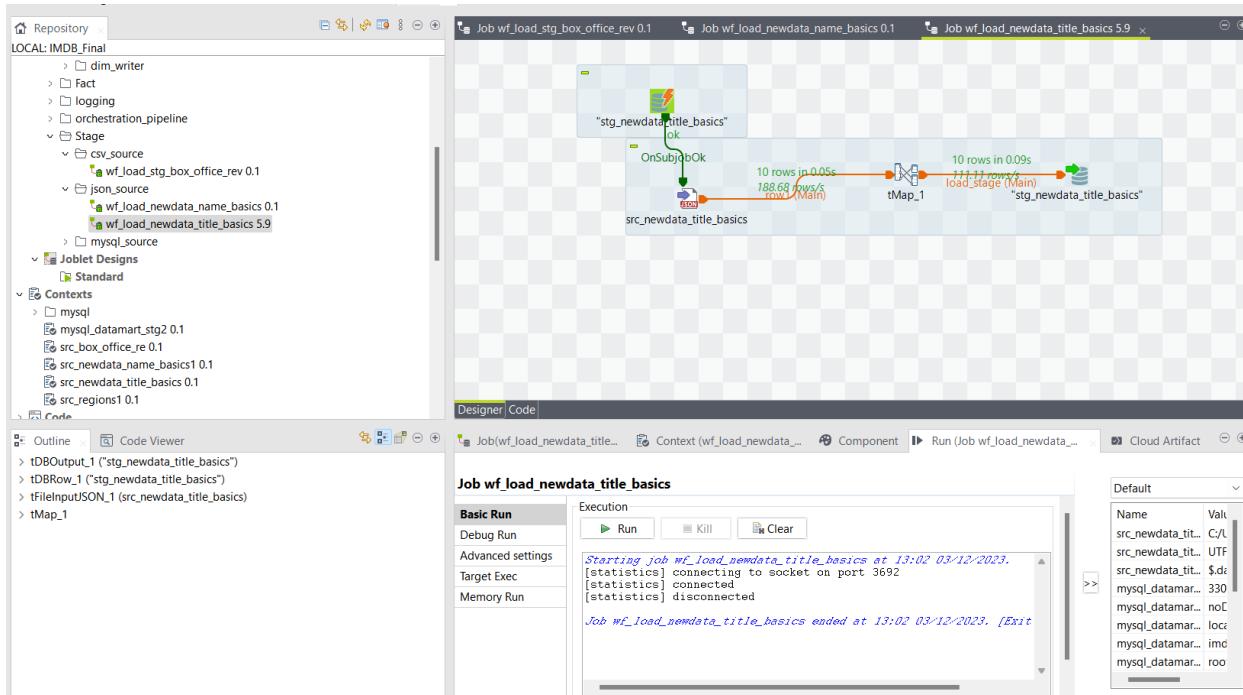
Name	Datatype	% missing	Unique values	Observation	Min Value	Max Value
endYear	Numeric	100.0%	1	This field has over 10% missing values. Consider imputing these values. This field has a small number of unique values, and appears to be a categorical field. Consider changing the field data type to		

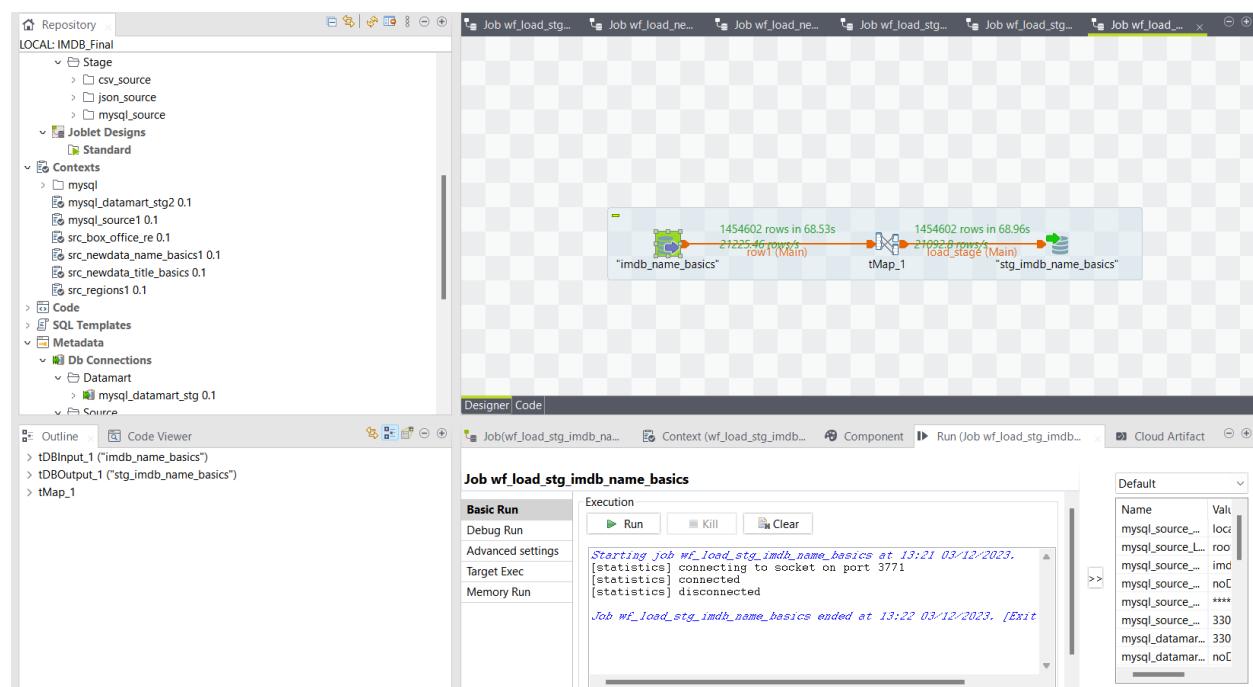
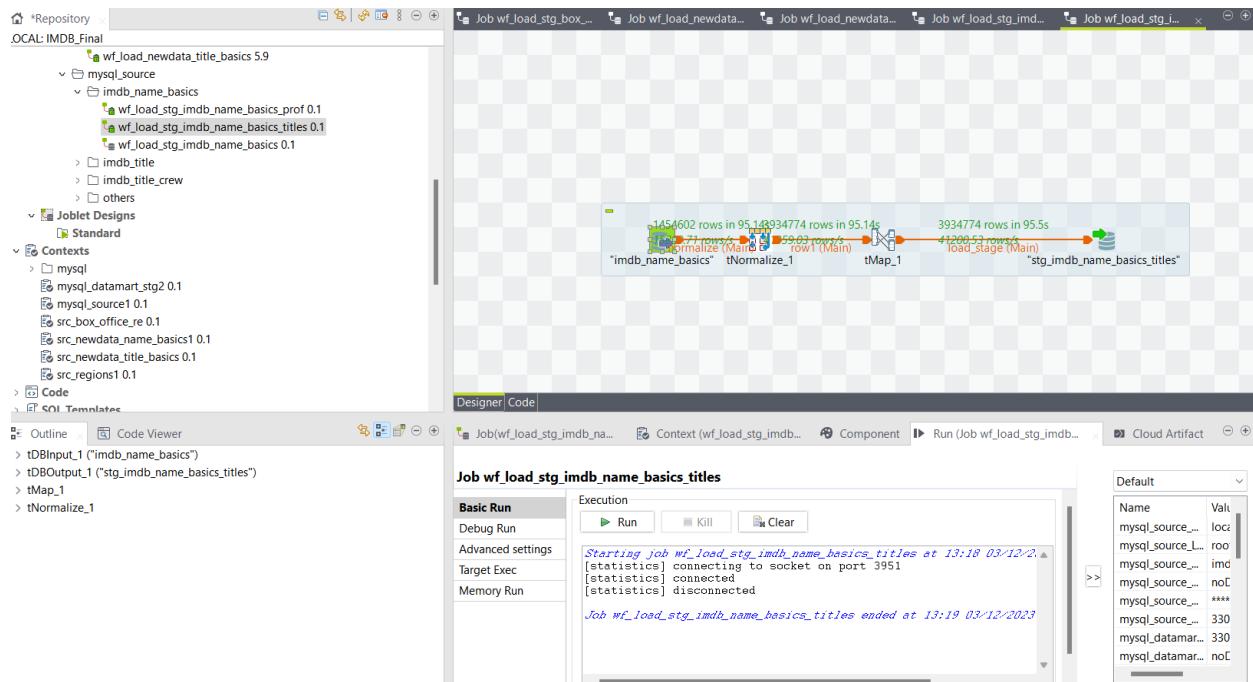
				"string".		
runtimeMinutes	Numeric	0%	6	This field has a small number of unique values, and appears to be a categorical field. Consider changing the field data type to "string".	134	194
startYear	Numeric	0%	6	This field has a small number of unique values, and appears to be a categorical field. Consider changing the field data type to "string".	1,997	2,021
titleType	String	0%	1		6	6
primaryTitle	String	0%	6		1	1
isAdult	String	0.0%	1		6	6
genres	String	0.0%	3		1	3
tconst	String	0.0%	6		1	1
originalTitle	String	0.0%	6		1	1

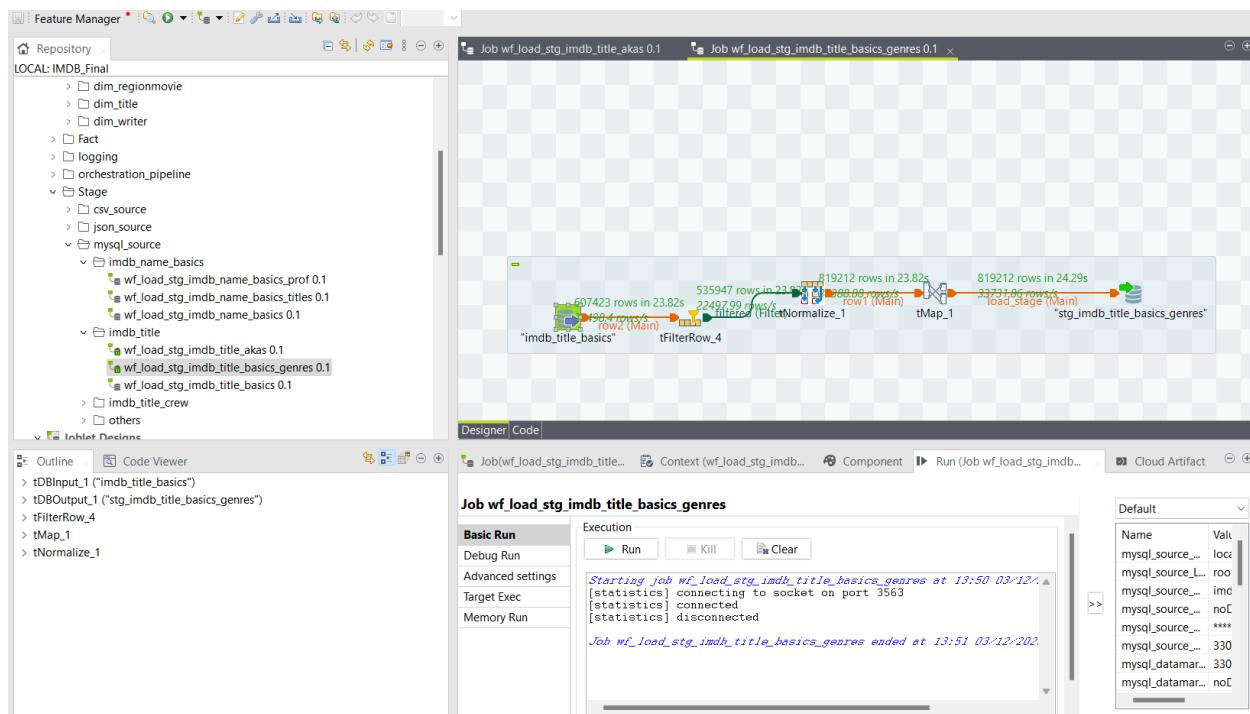
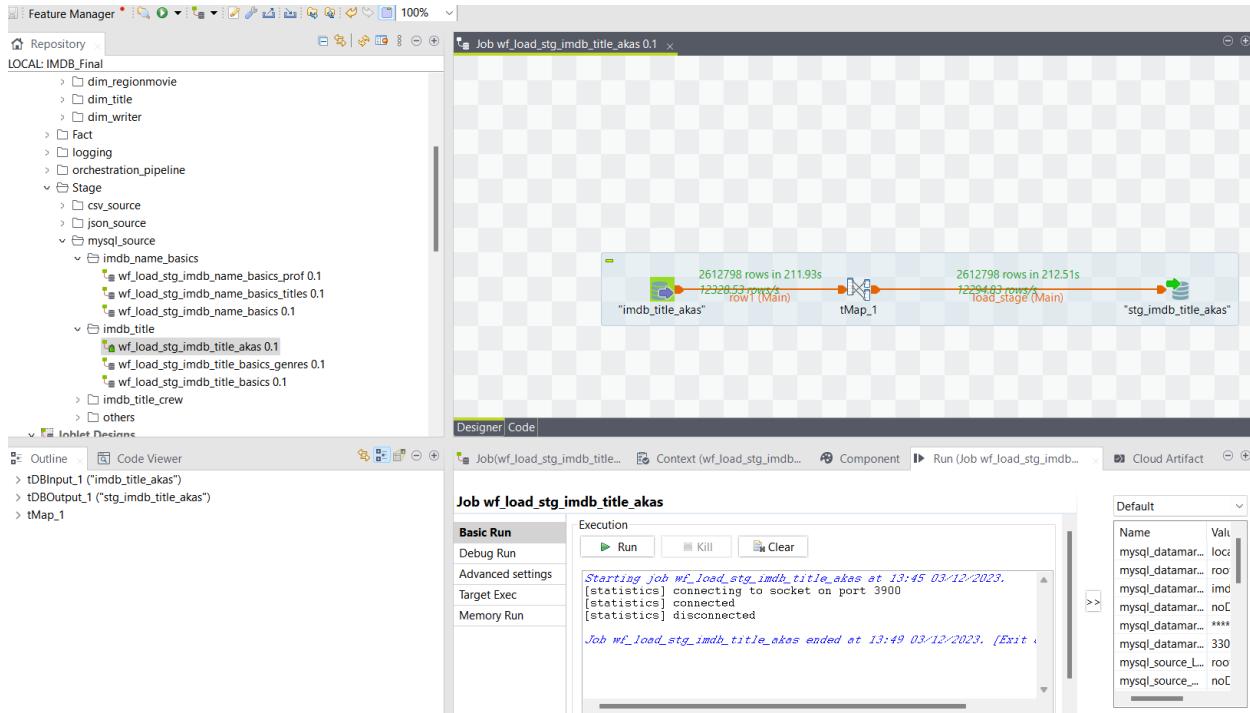
# Talend Workflows : Staging

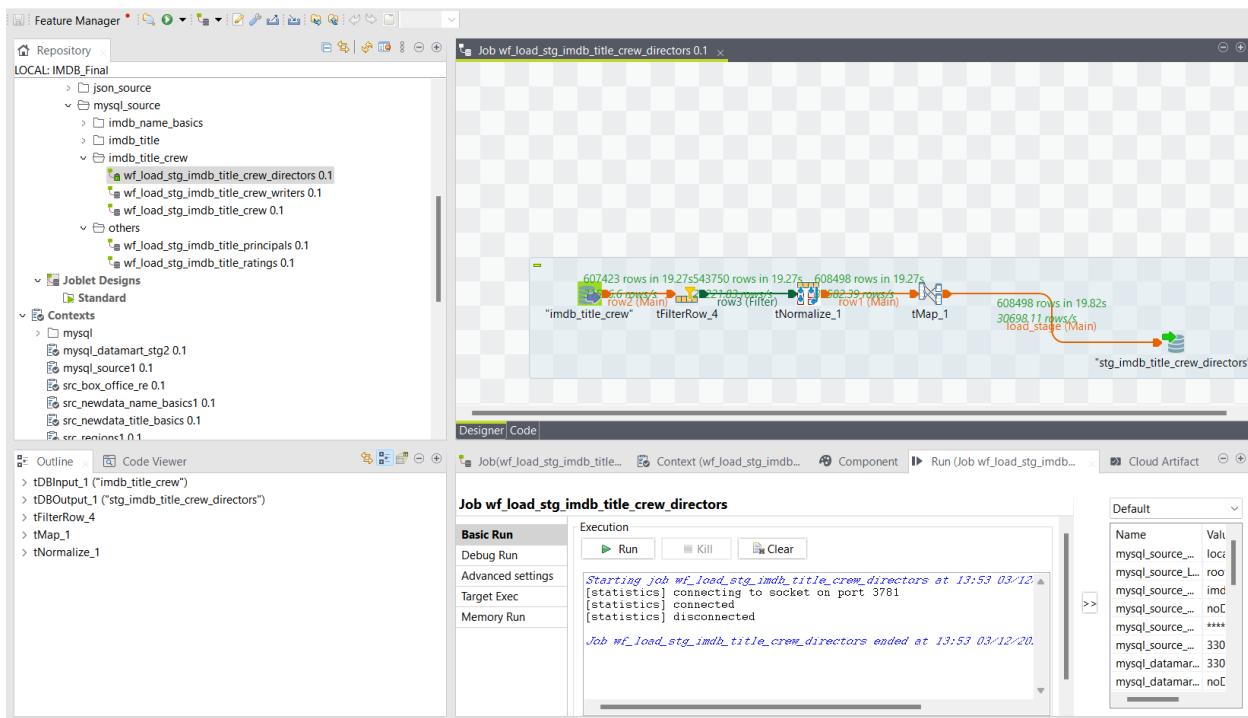
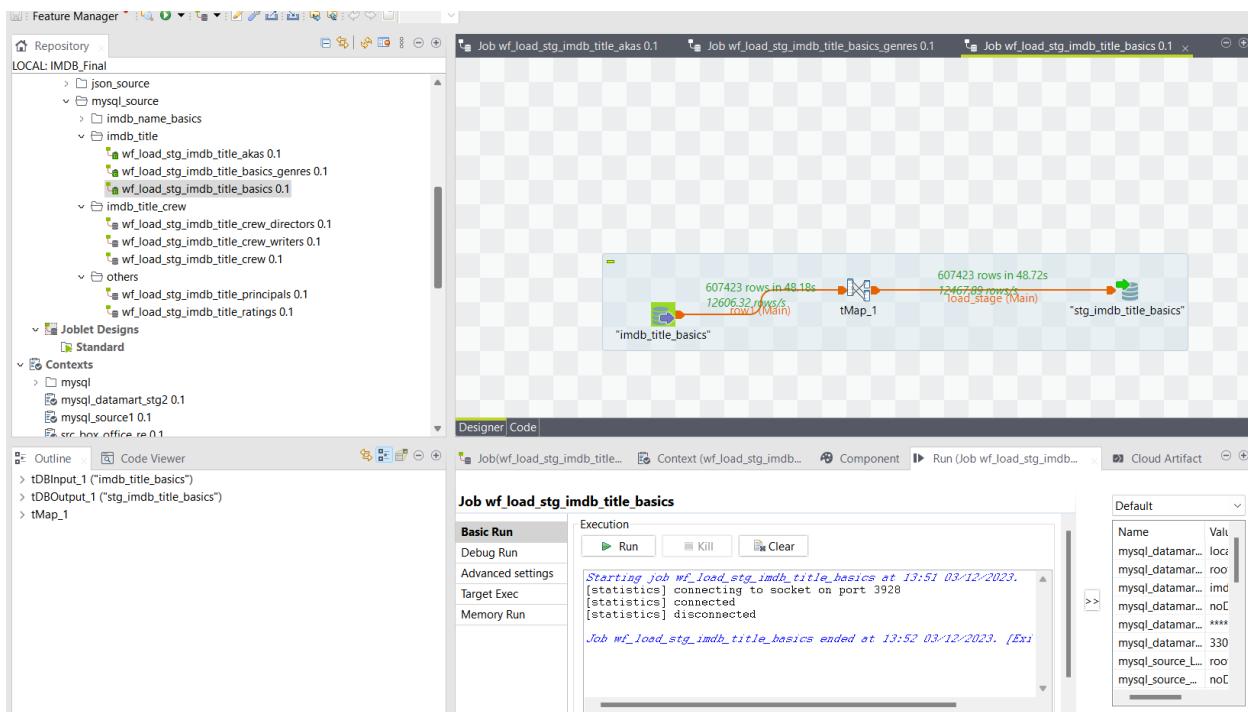
## Staging:

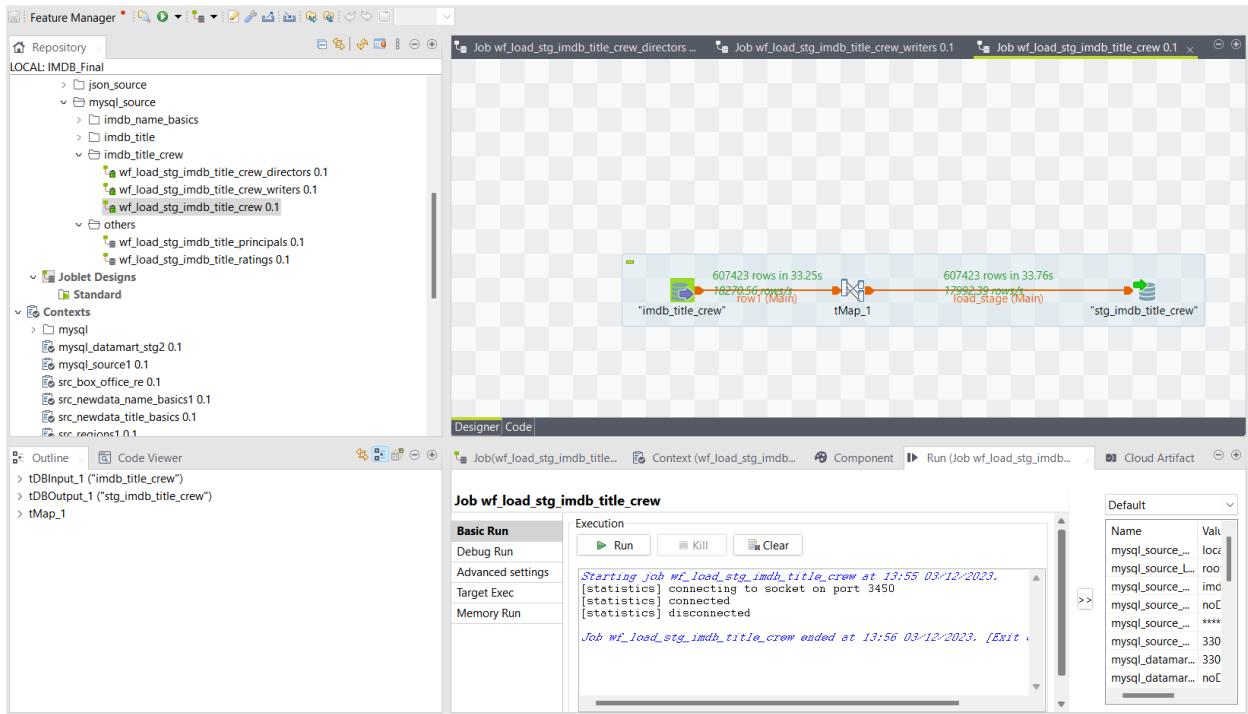
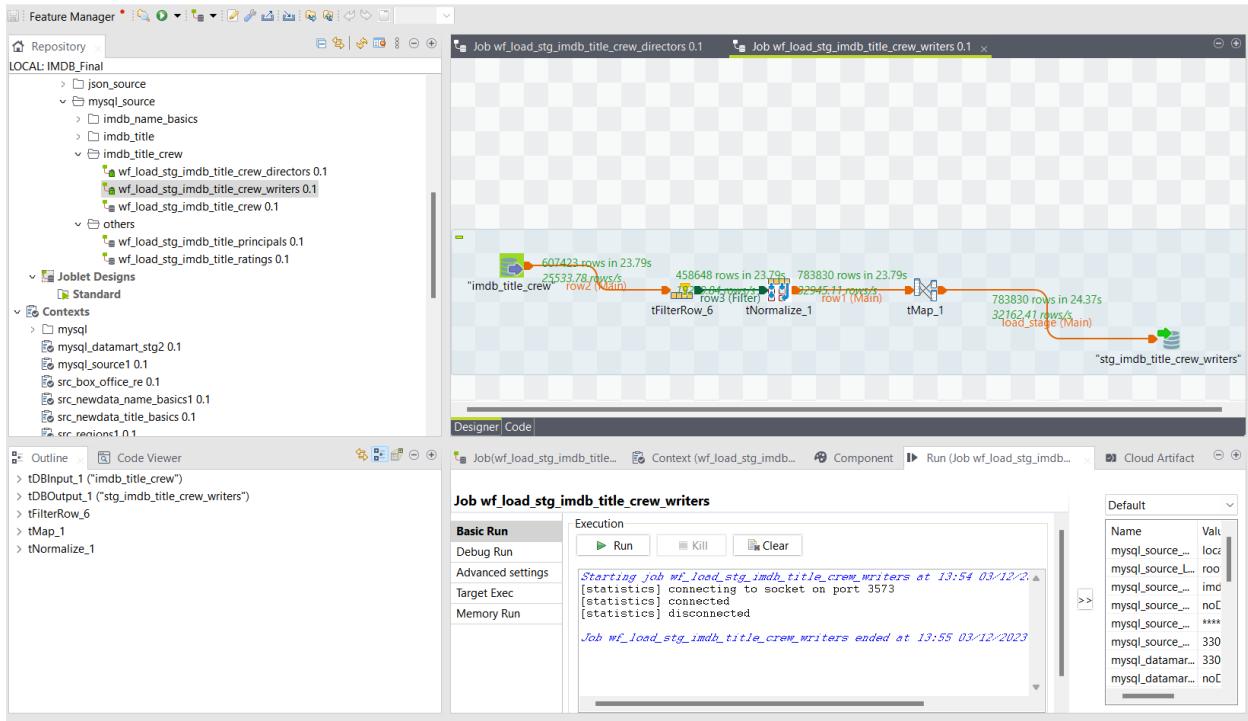


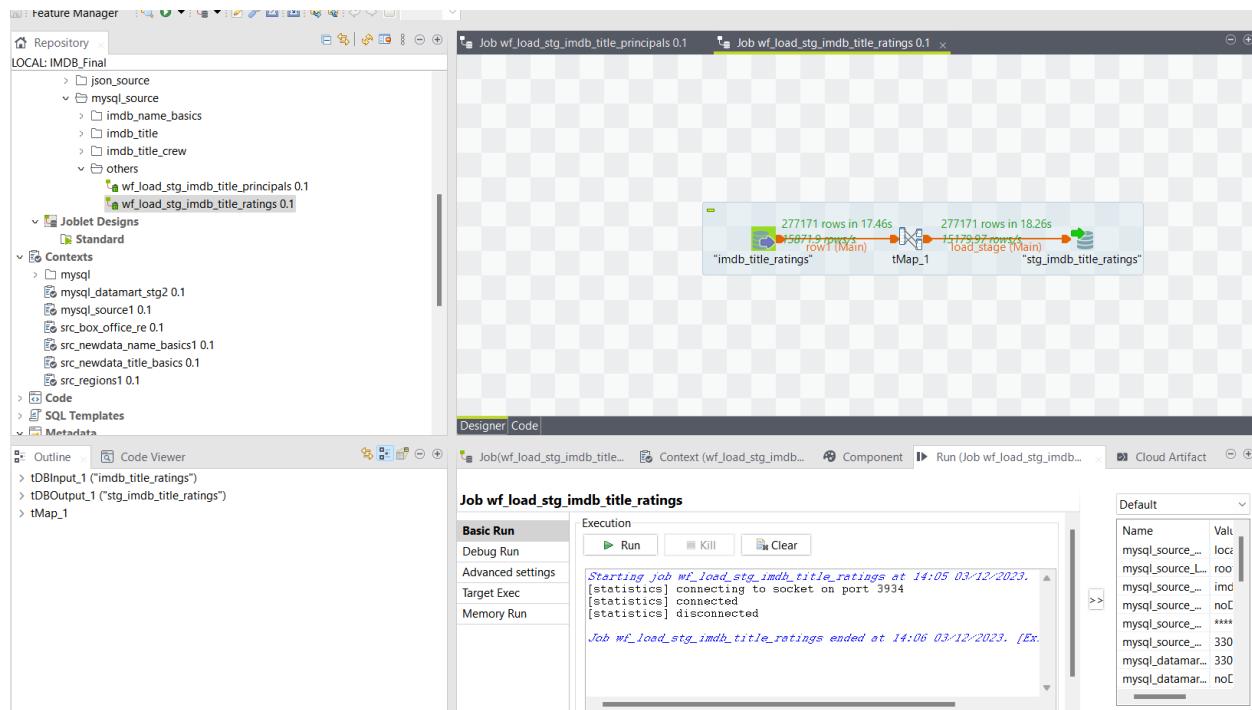
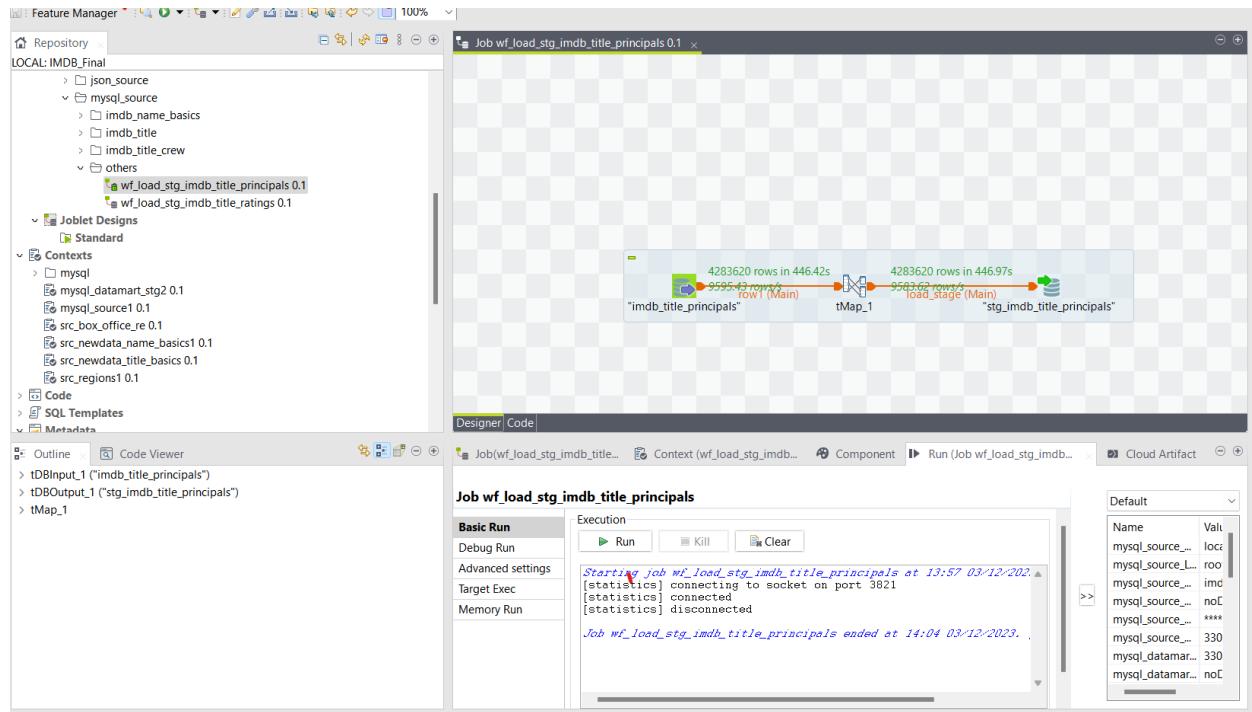












## SQL Query Data and Count for Staging Table –

movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas imdb\_tables\_backup SQL File 18\* SQL File 19\*

```

1 use imdb_movies_datamart;
2
3 select * from stg_box_office_rev;

```

**Result Grid**

sbor_id	tconst	title	Date	Rank1	Gross	YD	LW	Theatres	Per_Theatre	Total_Gross	Days	source_file_name
1	tt0499549	Avatar	Dec 18, 2009	1	\$26,752,099			3,452	\$7,750	\$26,752,099	1	The Numbers - Domestic Box Of
2	tt0499549	Avatar	Dec 19, 2009	1	\$25,529,036	-5%		3,452	\$7,395	\$52,281,135	2	The Numbers - Domestic Box Of
3	tt0499549	Avatar	Dec 20, 2009	1	\$24,744,346	-3%		3,452	\$7,168	\$77,025,481	3	The Numbers - Domestic Box Of
4	tt0499549	Avatar	Dec 21, 2009	1	\$16,385,820	-34%		3,452	\$4,747	\$93,411,301	4	The Numbers - Domestic Box Of
5	tt0499549	Avatar	Dec 22, 2009	1	\$16,086,461	-2%		3,452	\$4,660	\$109,497,762	5	The Numbers - Domestic Box Of
6	tt0499549	Avatar	Dec 23, 2009	2	\$16,445,291	0.02		3,452	\$4,764	\$125,943,053	6	The Numbers - Domestic Box Of
7	tt0499549	Avatar	Dec 24, 2009	1	\$11,150,998	-32%		3,452	\$3,230	\$137,094,051	7	The Numbers - Domestic Box Of
8	tt0499549	Avatar	Dec 25, 2009	2	\$23,095,046	1.07	-14%	3,456	\$6,683	\$160,189,097	8	The Numbers - Domestic Box Of
9	tt0499549	Avatar	Dec 26, 2009	1	\$28,274,406	0.22	0.11	3,456	\$8,181	\$188,463,503	9	The Numbers - Domestic Box Of
10	tt0499549	Avatar	Dec 27, 2009	1	\$24,247,681	-14%	-2%	3,456	\$7,016	\$212,711,184	10	The Numbers - Domestic Box Of
11	tt0499549	Avatar	Dec 28, 2009	1	\$19,418,139	-20%	0.19	3,456	\$5,619	\$232,129,323	11	The Numbers - Domestic Box Of

stg\_box\_office\_rev3 x

movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas imdb\_tables\_backup SQL File 18\* SQL File 19\*

```

1 use imdb_movies_datamart;
2
3 select * from stg_box_office_rev;
4 select count(*) from stg_box_office_rev;

```

Result 4 x

Read Only

count(*)
1465

## Stg\_imdb\_name\_basics –

movies SQL File 8 SQL File 10 dim\_imdb\_date dim\_imdb\_title tct\_movie imdb\_title\_akas IMDB\_tables\_backup SQL File 18 SQL File 19

```

1 use imdb_movies_datamart;
2
3 select * from stg_box_office_rev;
4 select count(*) from stg_box_office_rev;
5
6 select * from stg_imdb_name_basics;
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types |

nconst	primaryName	birthYear	deathYear	primaryProfession	knownForTitles	Source_Create_DT	Proc
nm0000001	Fred Astaire	1899	1987	soundtrack,actor,miscellaneous	tt0072308,tt0053137,tt0050419,tt0031983	2022-04-20 07:21:04	y5kjZ:
nm0000002	Lauren Bacall	1924	2014	actress,soundtrack	tt0038355,tt0117057,tt0071877,tt0037382	2022-04-20 07:21:04	y5kjZ:
nm0000003	Brigitte Bardot	1934	N	actress,soundtrack,music_department	tt0054452,tt0056404,tt0049189,tt0057345	2022-04-20 07:21:04	y5kjZ:
nm0000004	John Belushi	1949	1982	actor,soundtrack,writer	tt0080455,tt0077975,tt0078723,tt0072562	2022-04-20 07:21:04	y5kjZ:
nm0000005	Ingmar Bergman	1918	2007	writer,director,actor	tt0060827,tt0050986,tt0083922,tt0069467	2022-04-20 07:21:04	y5kjZ:
nm0000006	Ingrid Bergman	1915	1982	actress,soundtrack,producer	tt0077711,tt0036855,tt0034583,tt0038109	2022-04-20 07:21:04	y5kjZ:
nm0000007	Humphrey Bogart	1899	1957	actor,soundtrack,producer	tt0043265,tt0037382,tt0042593,tt0034583	2022-04-20 07:21:04	y5kjZ:
nm0000008	Marlon Brando	1924	2004	actor,soundtrack,director	tt0070849,tt0078788,tt0068646,tt0047296	2022-04-20 07:21:04	y5kjZ:
nm0000009	Richard Burton	1925	1984	actor,soundtrack,producer	tt0059749,tt0087803,tt0061184,tt0057877	2022-04-20 07:21:04	y5kjZ:
nm0000010	James Cagney	1899	1986	actor,soundtrack,director	tt0042041,tt0035575,tt0031867,tt0029870	2022-04-20 07:21:04	y5kjZ:
nm0000011	Gary Cooper	1901	1961	actor,soundtrack,producer	tt0027996,tt0044706,tt0034167,tt0035896	2022-04-20 07:21:04	y5kjZ:

stg\_imdb\_name\_basics5 x | Apply | Revert |

movies SQL File 8 SQL File 10 dim\_imdb\_date dim\_imdb\_title tct\_movie imdb\_title\_akas IMDB\_tables\_backup SQL File 18 SQL File 19

```

1 use imdb_movies_datamart;
2
3 select * from stg_box_office_rev;
4 select count(*) from stg_box_office_rev;
5
6 select * from stg_imdb_name_basics;
7 select count(*) from stg_imdb_name_basics;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types |

count(*)
1454602

Result 6 x | Read Only |

## Stg\_imdb\_name\_basics\_prof

movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas imdb\_tables\_backup SQL File 18\* SQL File 19\*

```

3   select * from stg_box_office_rev;
4   select count(*) from stg_box_office_rev;
5
6   select * from stg_imdb_name_basics;
7   select count(*) from stg_imdb_name_basics;
8
9   select * from stg_imdb_name_basics_prof;
```

```

**Result Grid** | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | **Result Grid**

| sinbp_id | nconst    | primaryProfession | Process_ID | DI_Create_DT        |
|----------|-----------|-------------------|------------|---------------------|
| 1        | nm0000001 | soundtrack        | 9ppihu     | 2023-12-03 13:12:50 |
| 2        | nm0000001 | actor             | 9ppihu     | 2023-12-03 13:12:50 |
| 3        | nm0000001 | miscellaneous     | 9ppihu     | 2023-12-03 13:12:50 |
| 4        | nm0000002 | actress           | 9ppihu     | 2023-12-03 13:12:50 |
| 5        | nm0000002 | soundtrack        | 9ppihu     | 2023-12-03 13:12:50 |
| 6        | nm0000003 | actress           | 9ppihu     | 2023-12-03 13:12:50 |
| 7        | nm0000003 | soundtrack        | 9ppihu     | 2023-12-03 13:12:50 |
| 8        | nm0000003 | music_department  | 9ppihu     | 2023-12-03 13:12:50 |
| 9        | nm0000004 | actor             | 9ppihu     | 2023-12-03 13:12:50 |
| 10       | nm0000004 | soundtrack        | 9ppihu     | 2023-12-03 13:12:50 |
| 11       | nm0000004 | writer            | 9ppihu     | 2023-12-03 13:12:50 |
| 12       | nm0000005 | writer            | 9ppihu     | 2023-12-03 13:12:50 |

stg\_imdb\_name\_basics\_prof 8 x | Apply | Revert

movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas imdb\_tables\_backup SQL File 18\* SQL File 19\*

```

4   select count(*) from stg_box_office_rev;
5
6   select * from stg_imdb_name_basics;
7   select count(*) from stg_imdb_name_basics;
8
9   select * from stg_imdb_name_basics_prof;
10  select count(*) from stg_imdb_name_basics_prof;
```

```

**Result Grid** | Filter Rows: | Export: | Wrap Cell Content: | **Result Grid**

count(*)
5089546

Result 9 x | Read Only

## stg\_imdb\_name\_basics\_titles

movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas imdb\_tables\_backup SQL File 18\* SQL File 19\*

```

7  select count(*) from stg_imdb_name_basics;
8
9  select * from stg_imdb_name_basics_prof;
10 select count(*) from stg_imdb_name_basics_prof;
11
12
13 select * from stg_imdb_name_basics_titles;

```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types |

sinbt_id	nconst	knownForTitles	Process_ID	DI_Create_DT
1	nm0000001	tt0072308	XocjwF	2023-12-03 13:18:09
2	nm0000001	tt0053137	XocjwF	2023-12-03 13:18:09
3	nm0000001	tt0050419	XocjwF	2023-12-03 13:18:09
4	nm0000001	tt0031983	XocjwF	2023-12-03 13:18:09
5	nm0000002	tt0038355	XocjwF	2023-12-03 13:18:09
6	nm0000002	tt0117057	XocjwF	2023-12-03 13:18:09
7	nm0000002	tt0071877	XocjwF	2023-12-03 13:18:09
8	nm0000002	tt0037382	XocjwF	2023-12-03 13:18:09
9	nm0000003	tt0054452	XocjwF	2023-12-03 13:18:09
10	nm0000003	tt0056404	XocjwF	2023-12-03 13:18:09
11	nm0000003	tt0049189	XocjwF	2023-12-03 13:18:09
12	nm0000003	tt0057345	XocjwF	2023-12-03 13:18:09

stg\_imdb\_name\_basics\_titles 10 x | Apply | Revert |

etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas

```

1 •  SELECT * FROM imdb_src_movie.imdb_title_akas;
2
3 •  use imdb_movies_datamart;
4
5 •  select count(*) from stg_imdb_name_basics_titles;

```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types |

count(*)
3934774

Result 1 x | Read Only |

**Stg\_imdb\_title\_akas –**

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas x

1 • SELECT \* FROM imdb\_src\_movie.imdb\_title\_akas;

2

3 • use imdb\_movies\_datamart;

4

5 • select count(\*) from stg\_imdb\_name\_basics\_titles;

6

7 • select \* from stg\_imdb\_title\_akas;

**Result Grid** | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types |

titleId	ordering	title	region	language	types	attributes	isOriginalTitle	Source_Create_DT	Process_ID	DI_Create_DT
tt0000502	1	Bohemios	W	W	original	W	1	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000502	2	Bohemios	ES	W	imdbDisplay	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	1	Kelly bandájának története	HU	W	imdbDisplay	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	2	Ned Kelly and His Gang	AU	W	imdbDisplay	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	3	Priča o Keljevoj bandi	RS	W	imdbDisplay	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	4	The Story of the Kelly Gang	GB	W	imdbDisplay	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	5	The Story of the Kelly Gang	W	W	original	W	1	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	6	The Story of the Kelly Gang	SG	en	imdbDisplay	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	7	The Story of the Kelly Gang	US	W	imdbDisplay	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	8	Ned Kelly and His Gang	W	W	W	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:
tt0000574	9	Die Geschichte der Kelly Ba...	DE	W	W	W	0	2022-04-20 08:49:49	vBp86r	2023-12-03 13:46:

stg\_imdb\_title\_akas 2 x

Apply

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas x

2

3 • use imdb\_movies\_datamart;

4

5 • select count(\*) from stg\_imdb\_name\_basics\_titles;

6

7 • select \* from stg\_imdb\_title\_akas;

8 • select count(\*) from stg\_imdb\_title\_akas;

**Result Grid** | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types |

count(*)
2612798

Result 3 x

Read Only

## Stg\_imdb\_title\_basics –

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas ×

```

5 •   select count(*) from stg_imdb_name_basics_titles;
6
7 •   select * from stg_imdb_title_akas;
8 •   select count(*) from stg_imdb_title_akas;
9
10 •  select * from stg_imdb_title_basics;
11 •  select count(*) from stg_imdb_title_basics;
```

**Result Grid** | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types |

tconst	titleType	primaryTitle	originalTitle	isAdult	startYear	endYear	runtimeMinutes	genres	
tt0000502	movie	Bohemios	Bohemios	0	1905	W	100		2
tt0000574	movie	The Story of the Kelly Gang	The Story of the Kelly Gang	0	1906	W	70	Action,Adventure,Biography	2
tt0000591	movie	The Prodigal Son	L'enfant prodigue	0	1907	W	90	Drama	2
tt0000615	movie	Robbery Under Arms	Robbery Under Arms	0	1907	W		Drama	2
tt0000630	movie	Hamlet	Amleto	0	1908	W		Drama	2
tt0000675	movie	Don Quijote	Don Quijote	0	1908	W		Drama	2
tt0000679	movie	The Fairylogue and Radio-Plays	The Fairylogue and Radio-Plays	0	1908	W	120	Adventure,Fantasy	2
tt0000793	movie	Andreas Hofer	Andreas Hofer	0	1909	W		Drama	2
tt0000814	movie	La bocana de Mar Chica	La bocana de Mar Chica	0	1909	W			2
tt0000838	movie	A Cultura do Cacau	A Cultura do Cacau	0	1909	W			2
tt0000842	movie	De Garraf a Barcelona	De Garraf a Barcelona	0	1909	W			2

stg\_imdb\_title\_basics 5 ×

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas ×

```

5 •   select count(*) from stg_imdb_name_basics_titles;
6
7 •   select * from stg_imdb_title_akas;
8 •   select count(*) from stg_imdb_title_akas;
9
10 •  select * from stg_imdb_title_basics;
11 •  select count(*) from stg_imdb_title_basics;
```

**Result Grid** | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types |

count(*)
60743

Result 6 ×

Read Only

## Stg\_imdb\_name\_basics\_titles\_genres –

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas ×

```

8 •   select count(*) from stg_imdb_title_akas;
9
10 •  select * from stg_imdb_title_basics;
11 •  select count(*) from stg_imdb_title_basics;
12
13 •  select * from stg_imdb_title_basics_genres;
14 •  select count(*) from stg_imdb_title_basics_genres;

```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types | Apply

sitbg_id	tconst	genres	Process_ID	DI_Create_DT
1	tt0000574	Action	0x1Hn	2023-12-03 13:50:38
2	tt0000574	Adventure	0x1Hn	2023-12-03 13:50:38
3	tt0000574	Biography	0x1Hn	2023-12-03 13:50:38
4	tt0000591	Drama	0x1Hn	2023-12-03 13:50:38
5	tt0000615	Drama	0x1Hn	2023-12-03 13:50:38
6	tt0000630	Drama	0x1Hn	2023-12-03 13:50:38
7	tt0000675	Drama	0x1Hn	2023-12-03 13:50:38
8	tt0000679	Adventure	0x1Hn	2023-12-03 13:50:38
9	tt0000679	Fantasy	0x1Hn	2023-12-03 13:50:38
10	tt0000793	Drama	0x1Hn	2023-12-03 13:50:38
11	tt0000886	Drama	0x1Hn	2023-12-03 13:50:38
12	tt0000941	Drama	0x1Hn	2023-12-03 13:50:38

stg\_imdb\_title\_genres 7 ×

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas ×

```

8 •  select count(*) from stg_imdb_title_akas;
9
10 • select * from stg_imdb_title_basics;
11 • select count(*) from stg_imdb_title_basics;
12
13 • select * from stg_imdb_title_basics_genres;
14 • select count(*) from stg_imdb_title_basics_genres;

```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types | Read Only

count(*)
819212

Result 8 ×

**Stg\_imdb\_title\_crew –**

stg\_imdb\_title\_crew 9 x

tconst	directors	writers	Source_Create_DT	Process_ID	DI_Create_DT
tt0000502	nm0063413	nm0063413,nm0657268,nm0675388	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000574	nm0846879	nm0846879	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000591	nm0141150	nm0141150	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000615	nm0533958	nm0092809,nm0533958	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000630	nm0143333	nm0000636	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000675	nm0194088	nm0148859	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000679	nm0091767,nm0877783	nm0000875,nm0877783	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000793	nm0081288,nm0691994	\N	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000814	nm0185426	\N	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000838	nm0017074	\N	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000842	nm0005717	\N	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01
tt0000846	nm0023107	\N	2022-04-20 07:20:36	ERhljO	2023-12-03 13:56:01

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas x

stg\_imdb\_title\_crew 10 x

count(*)
607423

## Stg\_imdb\_title\_crew\_directors –

14 • select count(\*) from stg\_imdb\_title\_genres;

15

16 • select \* from stg\_imdb\_title\_crew;

17 • select count(\*) from stg\_imdb\_title\_crew;

18

19 • select \* from stg\_imdb\_title\_crew\_directors;

20 • select count(\*) from stg\_imdb\_title\_crew\_directors;

21

sitcd_id	tconst	directors	Process_ID	DI_Create_DT
1	tt0000502	nm0063413	92lFGI	2023-12-03 13:53:38
2	tt0000574	nm0846879	92lFGI	2023-12-03 13:53:38
3	tt0000591	nm0141150	92lFGI	2023-12-03 13:53:38
4	tt0000615	nm0533958	92lFGI	2023-12-03 13:53:38
5	tt0000630	nm0143333	92lFGI	2023-12-03 13:53:38
6	tt0000675	nm0194088	92lFGI	2023-12-03 13:53:38
7	tt0000679	nm0091767	92lFGI	2023-12-03 13:53:38
8	tt0000679	nm0877783	92lFGI	2023-12-03 13:53:38
9	tt0000793	nm0081288	92lFGI	2023-12-03 13:53:38
10	tt0000793	nm0691994	92lFGI	2023-12-03 13:53:38
11	tt0000814	nm0185426	92lFGI	2023-12-03 13:53:38
12	tt0000838	nm0017074	92lFGI	2023-12-03 13:53:38

stg\_imdb\_title\_crew\_directors 11 X

Output:

15

16 • select \* from stg\_imdb\_title\_crew;

17 • select count(\*) from stg\_imdb\_title\_crew;

18

19 • select \* from stg\_imdb\_title\_crew\_directors;

20 • select count(\*) from stg\_imdb\_title\_crew\_directors;

21

count(*)
608498

Result 12 X

Read Only

## Stg\_imdb\_title\_crew\_writers-

```
17 • select count(*) from stg_imdb_title_crew;
18
19 • select * from stg_imdb_title_crew_directors;
20 • select count(*) from stg_imdb_title_crew_directors;
21
22 • select * from stg_imdb_title_crew_writers;
23 • select count(*) from stg_imdb_title_crew_writers;
24
```

Result Grid | Filter Rows:  | Edit:    | Export/Import:   | Wrap Cell Content:  | Fetch rows:

sitcw_id	tconst	writers	Process_ID	DI_Create_DT
1	tt0000502	nm0063413	p23x8S	2023-12-03 13:54:51
2	tt0000502	nm0657268	p23x8S	2023-12-03 13:54:51
3	tt0000502	nm0675388	p23x8S	2023-12-03 13:54:51
4	tt0000574	nm0846879	p23x8S	2023-12-03 13:54:51
5	tt0000591	nm0141150	p23x8S	2023-12-03 13:54:51
6	tt0000615	nm0092809	p23x8S	2023-12-03 13:54:51
7	tt0000615	nm0533958	p23x8S	2023-12-03 13:54:51
8	tt0000630	nm0000636	p23x8S	2023-12-03 13:54:51
9	tt0000675	nm0148859	p23x8S	2023-12-03 13:54:51
10	tt0000679	nm0000875	p23x8S	2023-12-03 13:54:51
11	tt0000679	nm0877783	p23x8S	2023-12-03 13:54:51
12	tt0000850	nm0063413	p23x8S	2023-12-03 13:54:51

stg\_imdb\_title\_crew\_writers 13 ×

```
18
19 • select * from stg_imdb_title_crew_directors;
20 • select count(*) from stg_imdb_title_crew_directors;
21
22 • select * from stg_imdb_title_crew_writers;
23 • select count(*) from stg_imdb_title_crew_writers;
```

Result Grid	Filter Rows:	Export:	Wrap Cell Content:	
count(*) 293830				Result Grid

Result 14 ×

## **Stg\_imdb\_title\_principals –**

```

4* etl_job_log drop index and schema s_movies SQL File 8* SQL File 10* dim_imdb_date dim_imdb_title fct_movie imdb_title_akas x
  | Limit to 1000 rows | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types | Apply |
20 • select count(*) from stg_imdb_title_crew_directors;
21
22 • select * from stg_imdb_title_crew_writers;
23 • select count(*) from stg_imdb_title_crew_writers;
24
25 • select * from stg_imdb_title_principals;
26 • select count(*) from stg_imdb_title_principals;
27

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types | Apply |
tconst ordering nconst category job characters Source_Create_DT Process_ID DI_Create_DT
tt0000502 1 nm0215752 actor \N \N 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000502 2 nm0252720 actor \N \N 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000502 3 nm0063413 director \N \N 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000502 4 nm0657268 writer \N \N 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000502 5 nm0675388 writer \N \N 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000574 1 nm0846887 actress \N ["Kate Kelly"] 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000574 2 nm0846894 actor \N ["School Master"] 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000574 3 nm1431224 actor \N ["Joe Byrne"] 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000574 4 nm3002376 actor \N ["Steve Hart"] 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000574 5 nm0846879 director \N \N 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000574 6 nm0317210 producer pro... \N 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41
tt0000574 7 nm0425854 producer pro... \N 2022-04-20 07:21:50 LIJaUL 2023-12-03 13:57:41

```

```

e 4* etl_job_log drop index and schema s_movies SQL File 8* SQL File 10* dim_imdb_date dim_imdb_title fct_movie imdb_title_akas x
  | Limit to 1000 rows | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types | Apply |
21
22 • select * from stg_imdb_title_crew_writers;
23 • select count(*) from stg_imdb_title_crew_writers;
24
25 • select * from stg_imdb_title_principals;
26 • select count(*) from stg_imdb_title_principals;
27

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types | Apply |
count(*)
4283620

```

**Stg\_imdb\_title\_ratings –**

etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas

```

22 • select * from stg_imdb_title_crew_writers;
23 • select count(*) from stg_imdb_title_crew_writers;
24
25 • select * from stg_imdb_title_principals;
26 • select count(*) from stg_imdb_title_principals;
27
28 • select * from stg_imdb_title_ratings;
29

```

**Result Grid** | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types

tconst	averageRating	numVotes	Source_Create_DT	Process_ID	DI_Create_DT
tt0000502	4.5	14	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000574	6	759	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000591	4.6	17	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000615	4.5	23	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000630	3.8	24	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000675	4.9	19	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000679	5.3	65	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000793	4.6	20	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000862	5.1	16	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000886	4.7	37	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0000941	4.8	23	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01
tt0001010	4.6	14	2022-04-20 07:20:30	A4Yls3	2023-12-03 14:06:01

stg\_imdb\_title\_ratings 17 x

4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas

```

24
25 • select * from stg_imdb_title_principals;
26 • select count(*) from stg_imdb_title_principals;
27
28 • select * from stg_imdb_title_ratings;
29 • select count(*) from stg_imdb_title_ratings;
30

```

**Result Grid** | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types

count(*)
277171

Result 18 x | Read Only

## Stg\_newdata\_name\_basics –

```
etl_job_log    drop index and schema    s_movies    SQL File 8*    SQL File 10*    dim_imdb_date    dim_imdb_title    fct_movie    imdb_title_akas    4*
```

Limit to 1000 rows

```
26 •   select count(*) from stg_imdb_title_principals;
27
28 •   select * from stg_imdb_title_ratings;
29 •   select count(*) from stg_imdb_title_ratings;
30
31 •   select * from stg_newdata_name_basics;
32 •   select count(*) from stg_newdata_name_basics;
```

---

Result Grid | Filter Rows: \_\_\_\_\_ | Export: \_\_\_\_\_ | Wrap Cell Content:

count(*)
7

Result Grid

Form Editor

Field Types

## **Stg\_newdata\_title\_basics –**

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas x

29 • select count(\*) from stg\_imdb\_title\_ratings;  
30  
31 • select \* from stg\_newdata\_name\_basics;  
32 • select count(\*) from stg\_newdata\_name\_basics;  
33  
34 • select \* from stg\_newdata\_title\_basics;  
35 • select count(\*) from stg\_newdata\_title\_basics;

---

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Result Grid | Form Editor | Field Types

sntb_id	tconst	titleType	primaryTitle	originalTitle	isAdult	startYear	endYear	runtimeMinutes
1	tt0120338	movie	Titanic (1997)	Titanic	0	1997	N	194
2	tt0499549	movie	Avatar (2009)	Avatar	0	2009	N	162
3	tt0848228	movie	The Avengers (2012)	The Avengers	0	2012	N	143
4	tt10872600	movie	Spider-Man: No Way Home (2021)	Spider-Man: No Way Home	0	2021	N	148
5	tt1825683	movie	Black Panther (2018)	Black Panther	0	2018	N	134
6	tt2395427	movie	Avengers: Age of Ultron (2015)	Avengers: Age of Ultron	0	2015	N	141
7	tt2488496	movie	Star Wars: Episode VII - The Force Awakens (2015)	Star Wars: Episode VII - The Force Awakens	0	2015	N	138
8	tt4154756	movie	Avengers: Infinity War (2018)	Avengers: Infinity War	0	2018	N	149
9	tt4154796	movie	Avengers: Endgame (2019)	Avengers: Endgame	0	2019	N	181
10	tt0120338	movie	Titanic (1997) remake	Titanic	0	1997	N	194
*	NULL	HULL	NULL	HULL	NULL	NULL	NULL	NULL

---

e 4\* etl\_job\_log drop index and schema s\_movies SQL File 8\* SQL File 10\* dim\_imdb\_date dim\_imdb\_title fct\_movie imdb\_title\_akas x

29 • select count(\*) from stg\_imdb\_title\_ratings;  
30  
31 • select \* from stg\_newdata\_name\_basics;  
32 • select count(\*) from stg\_newdata\_name\_basics;  
33  
34 • select \* from stg\_newdata\_title\_basics;  
35 • select count(\*) from stg\_newdata\_title\_basics;

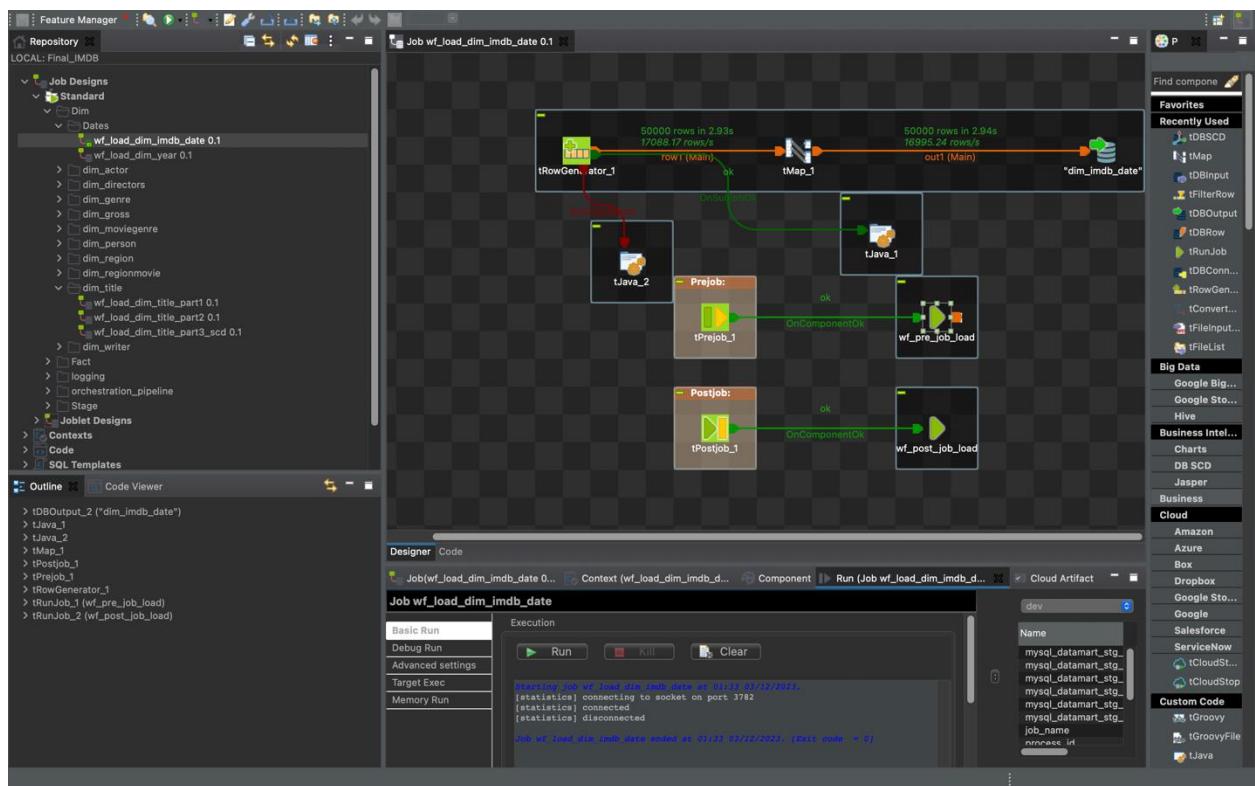
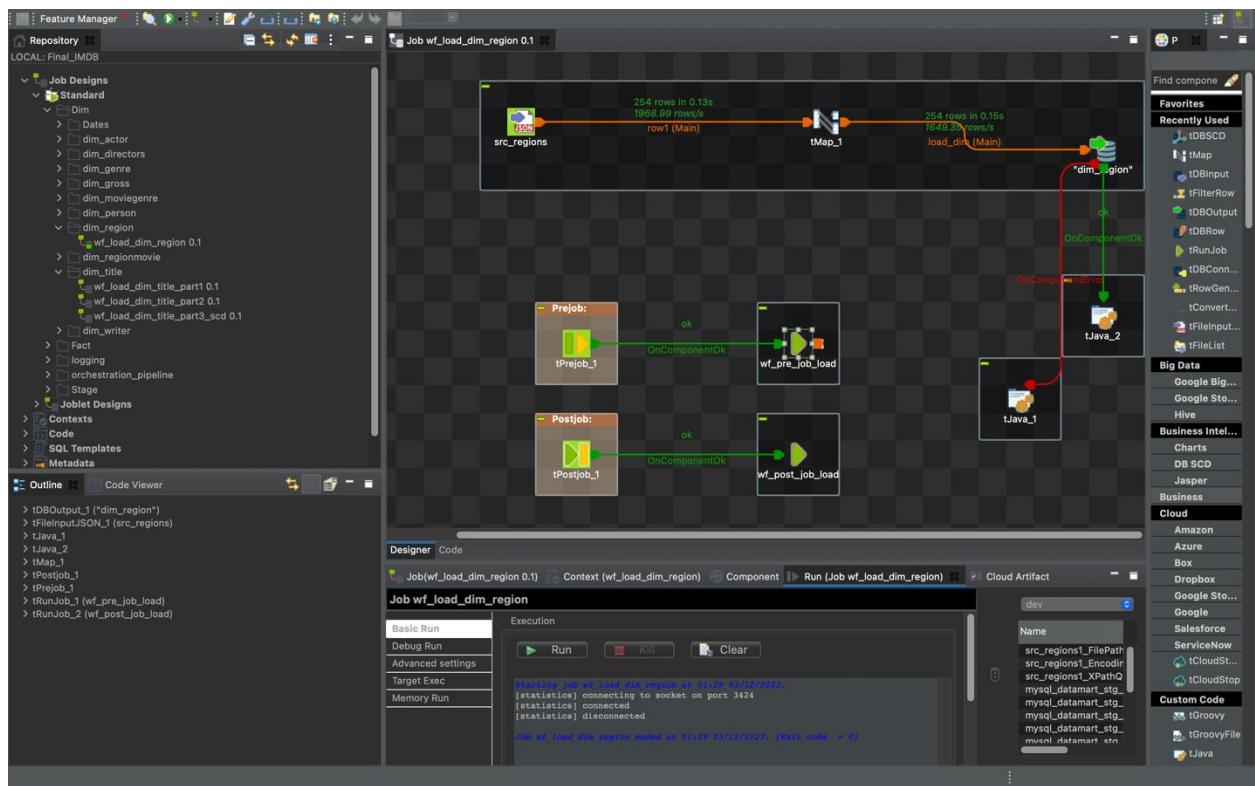
---

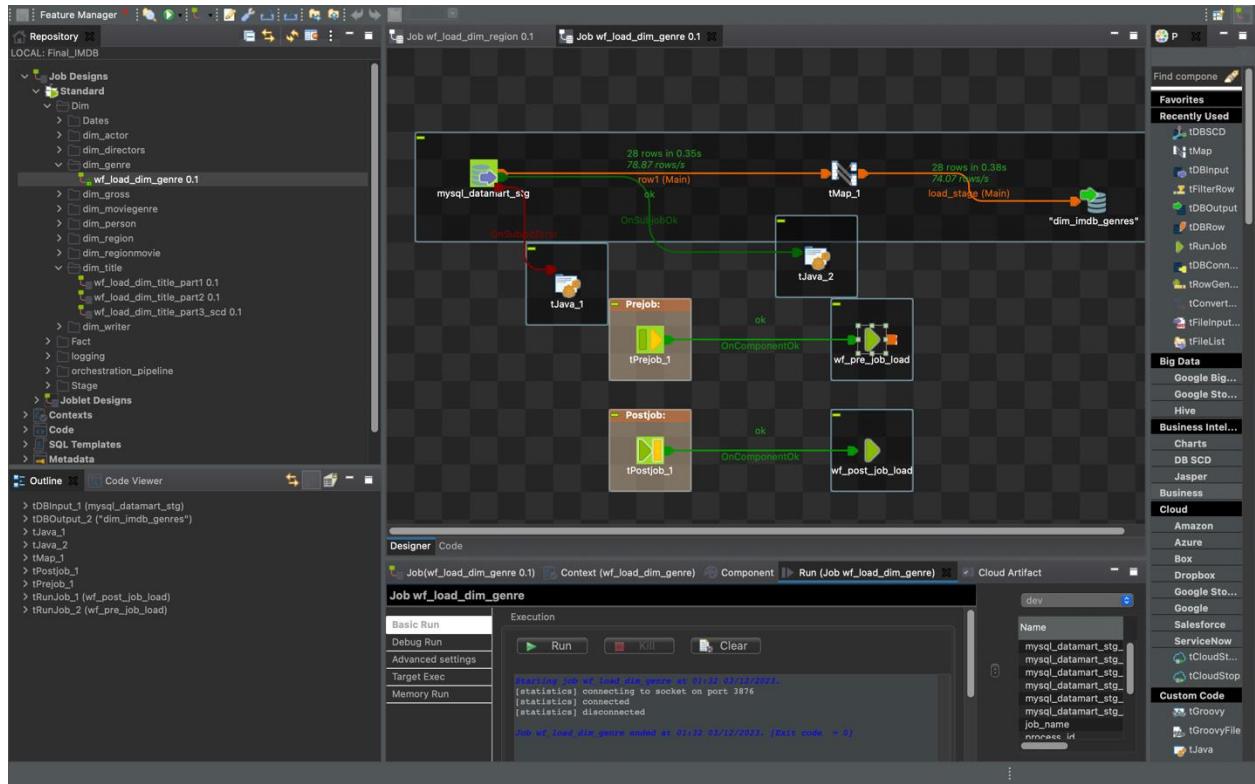
Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Result Grid | Form Editor | Field Types

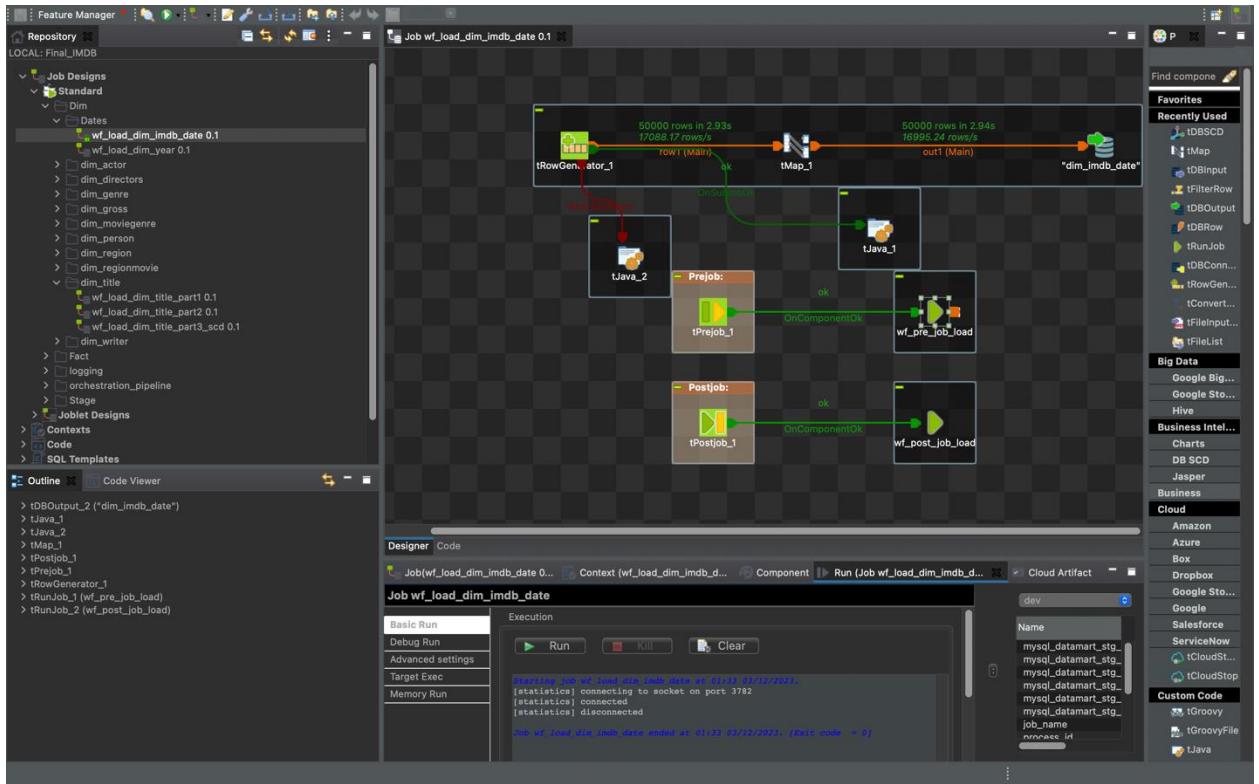
count(*)
10

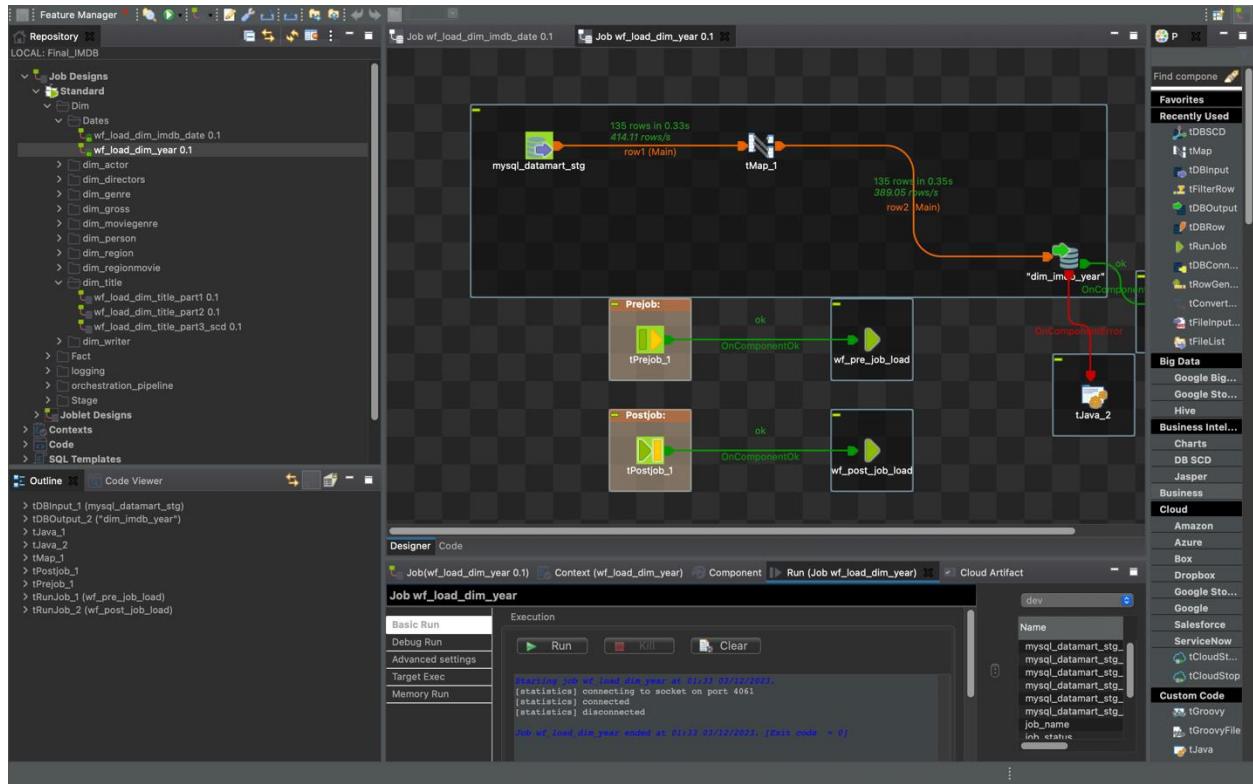
Result 22 | Dead Only

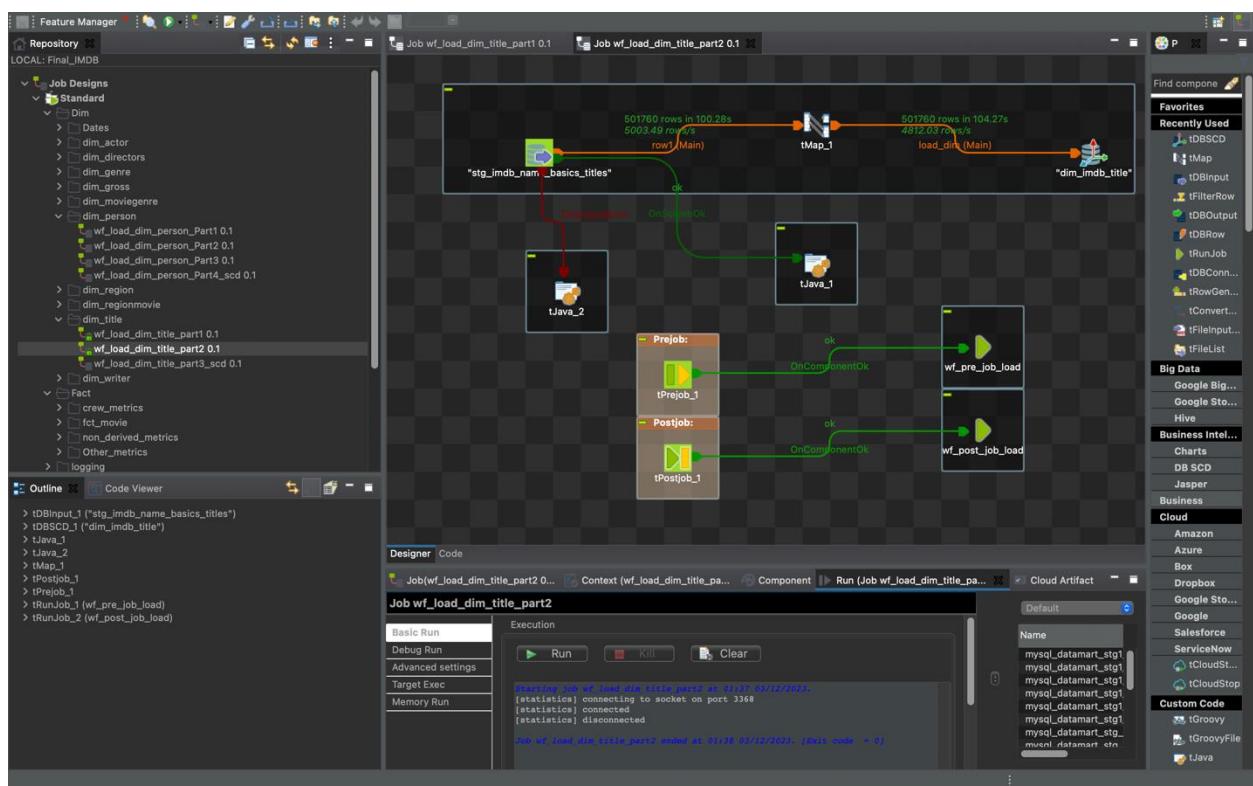
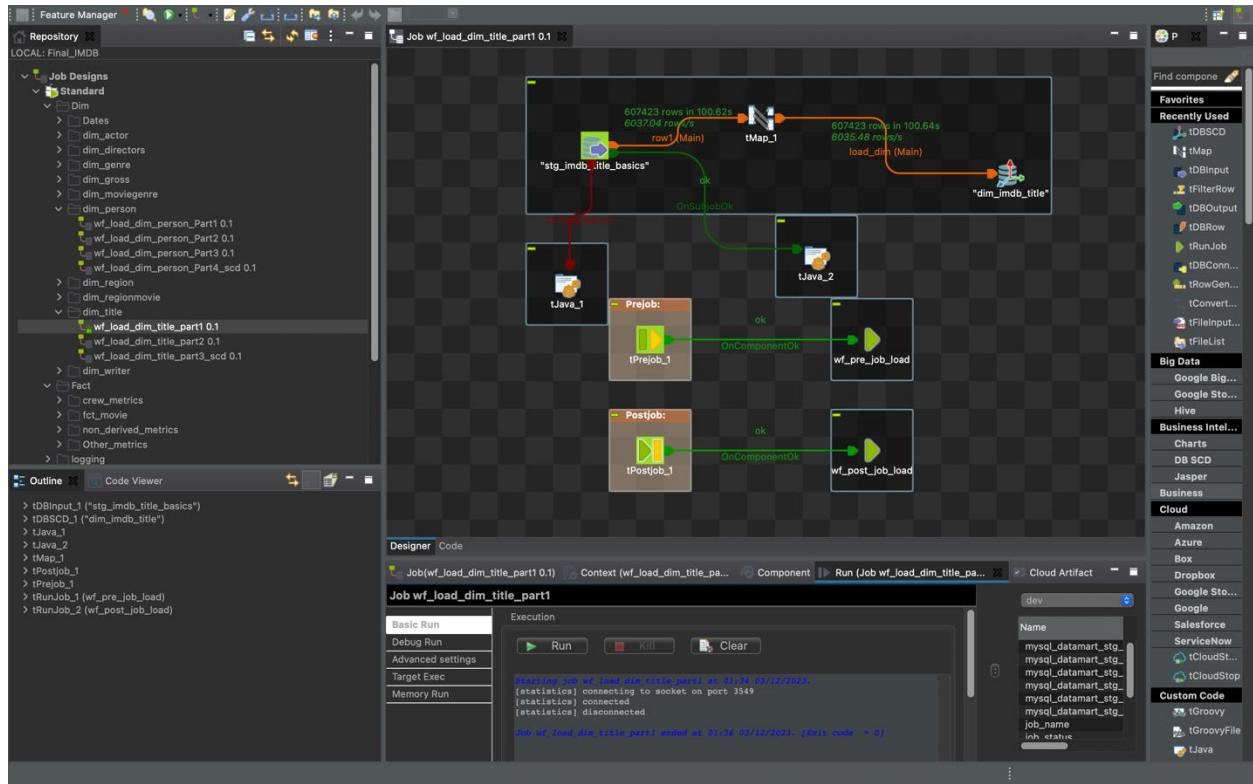
## DIM Loading:

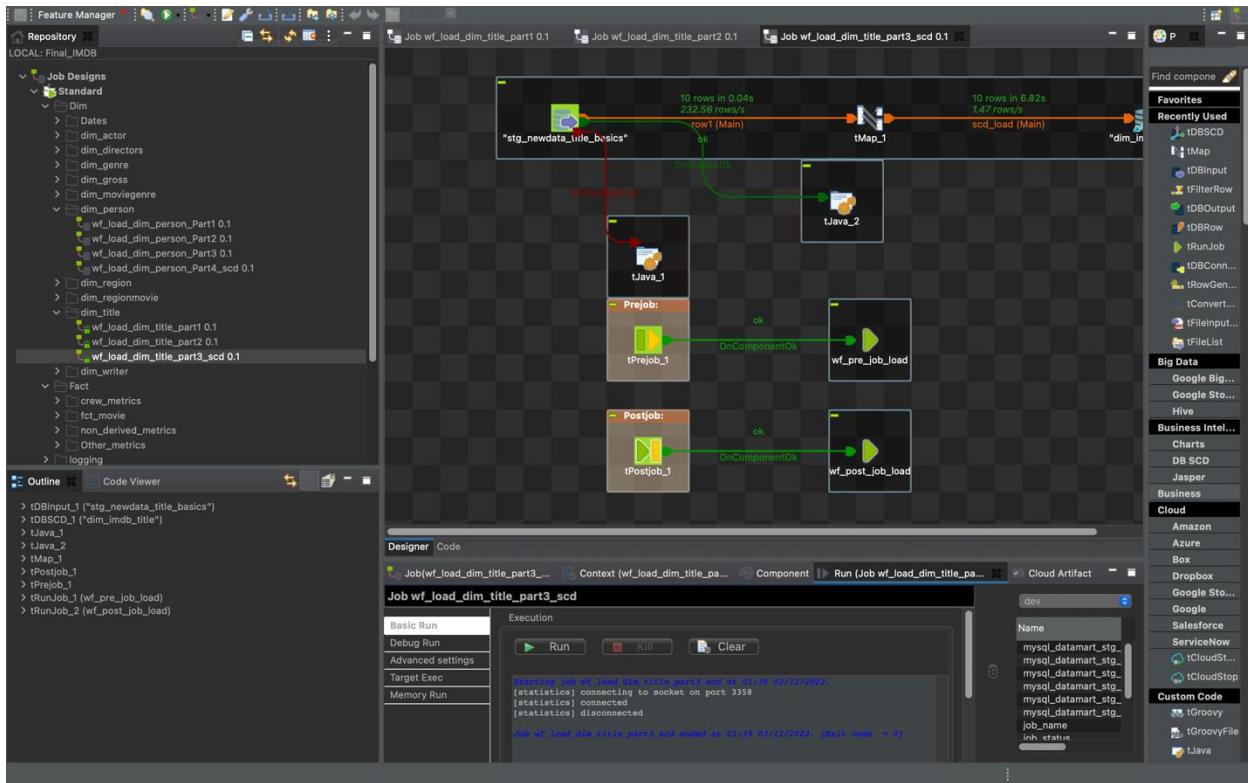


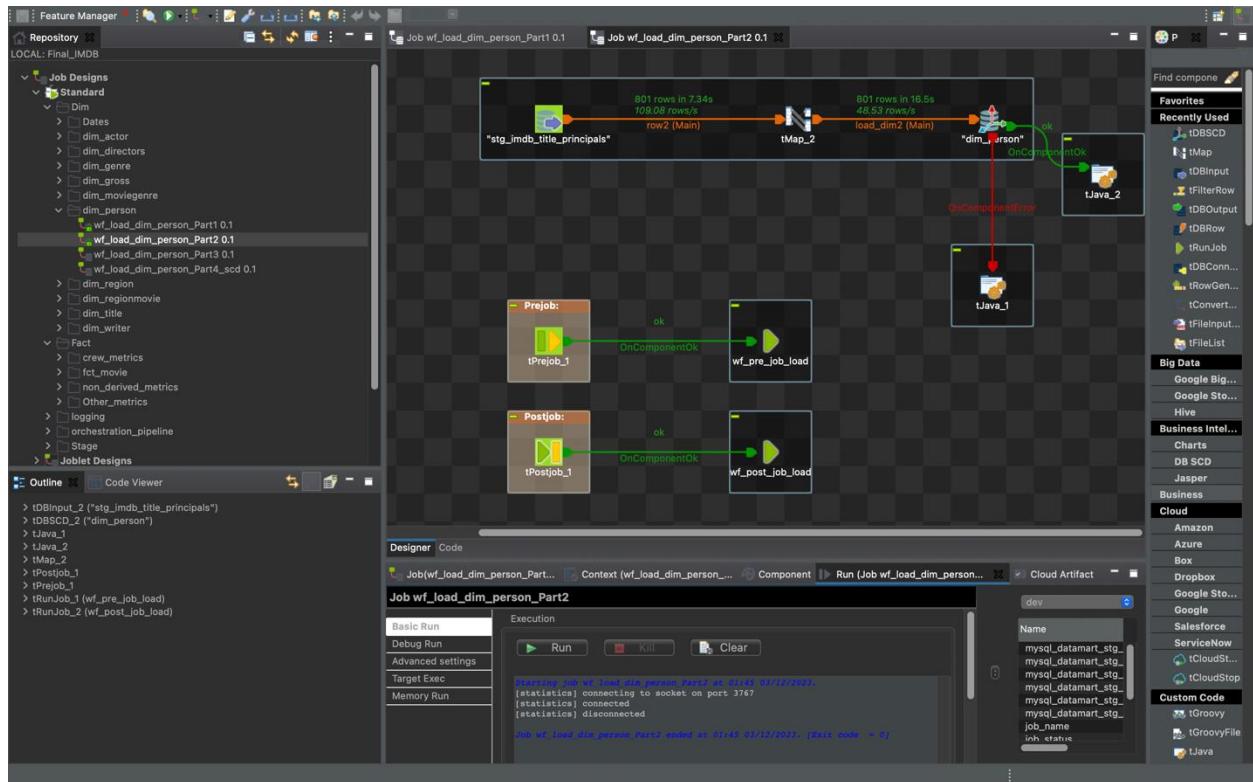
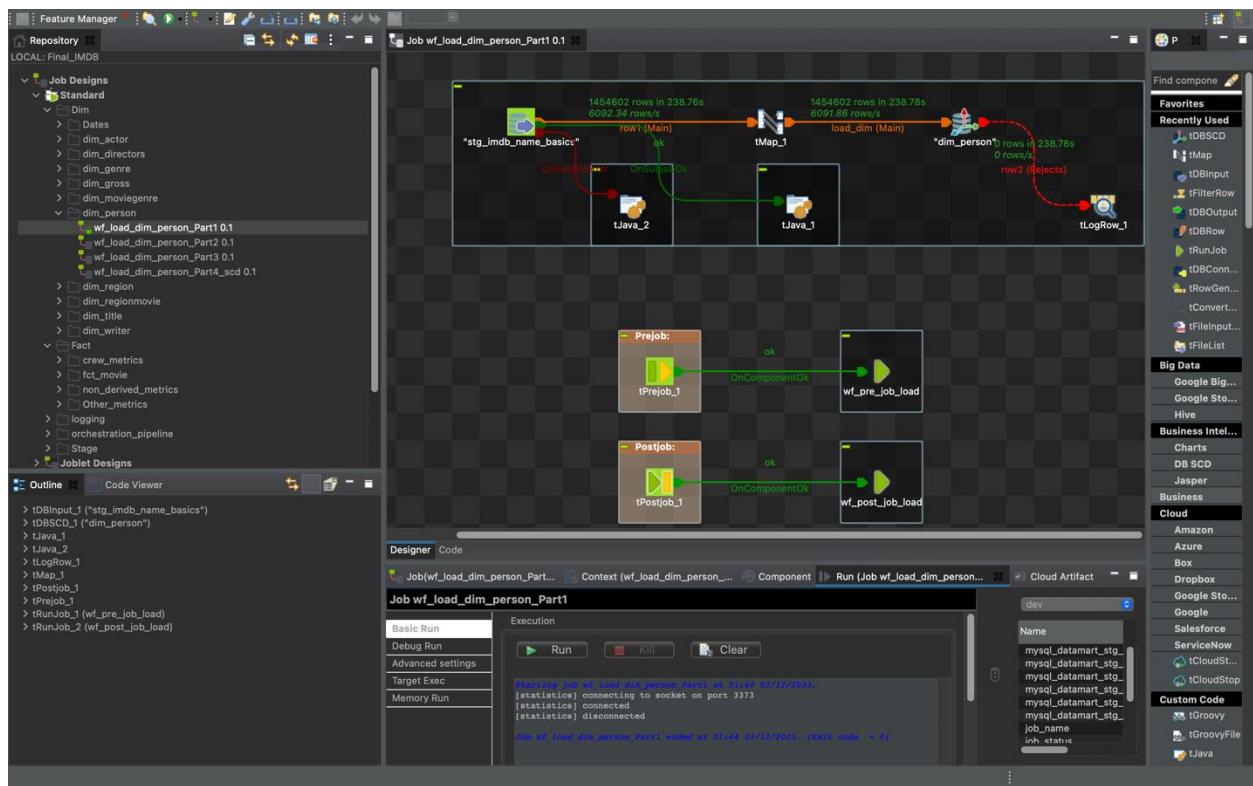


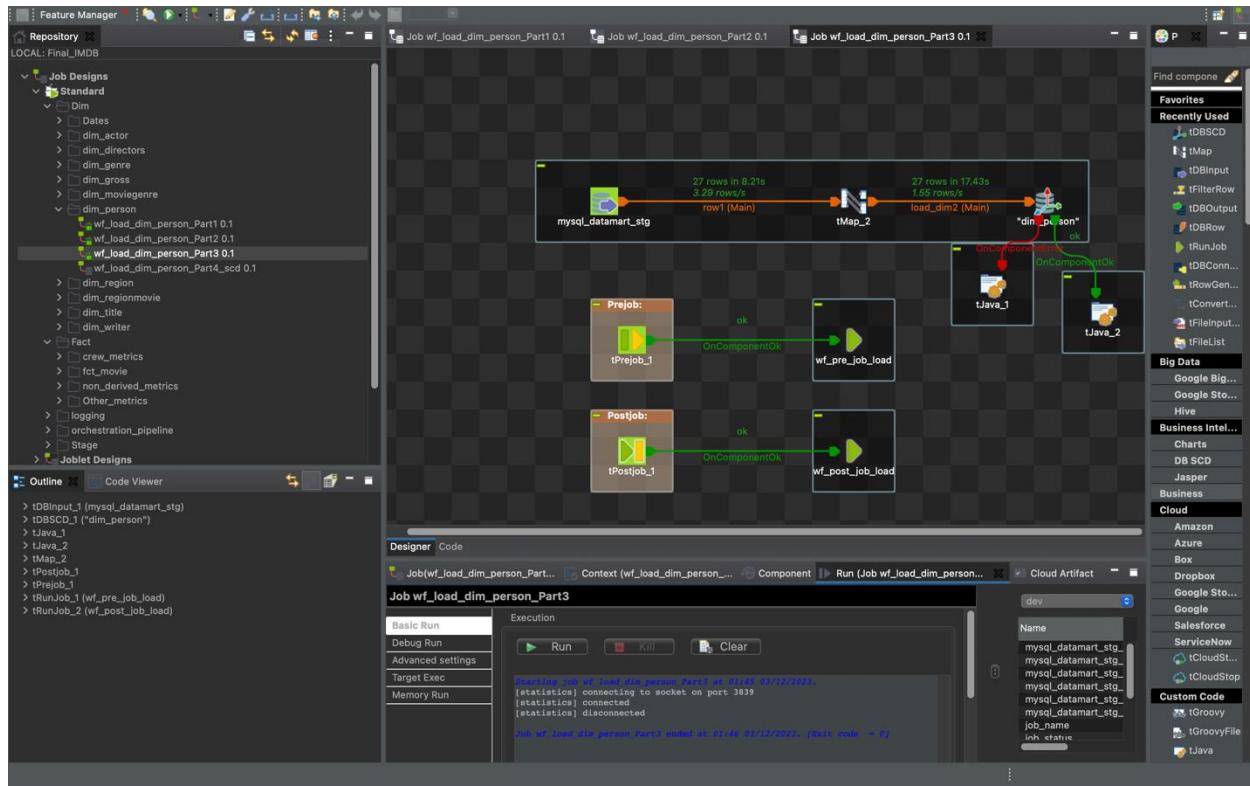


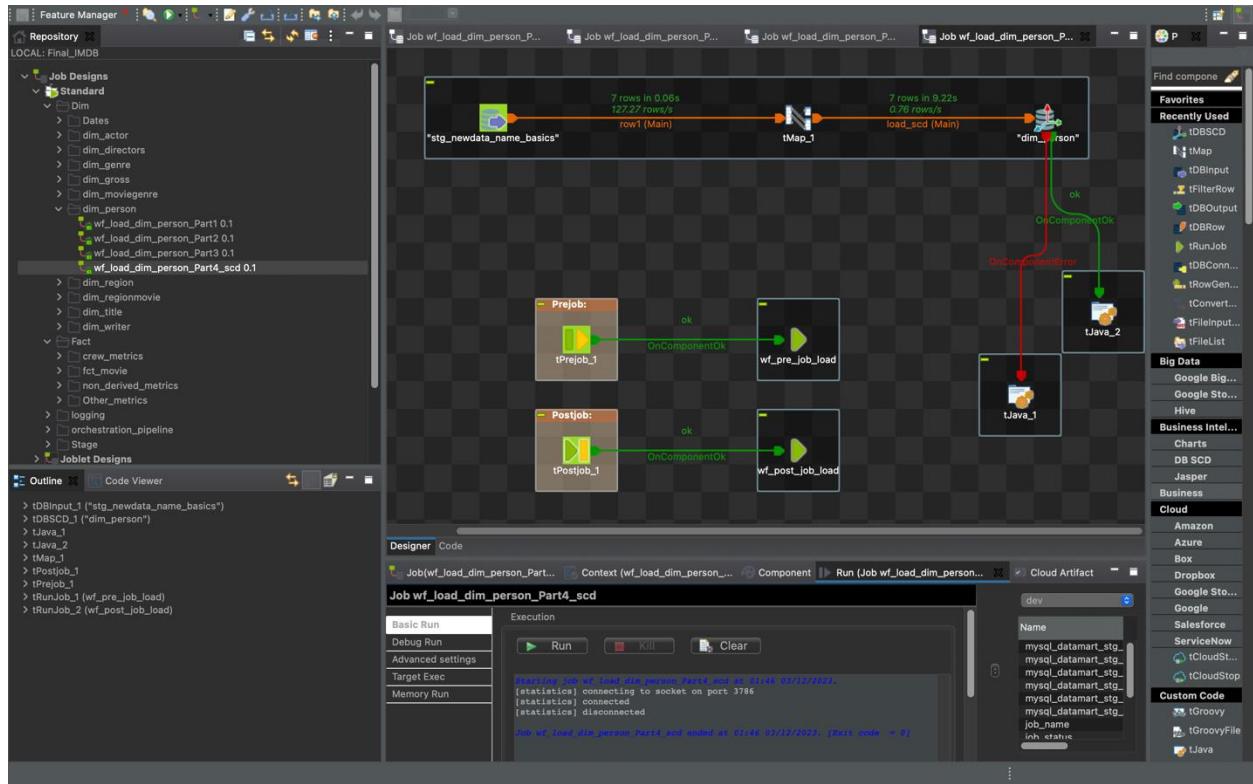


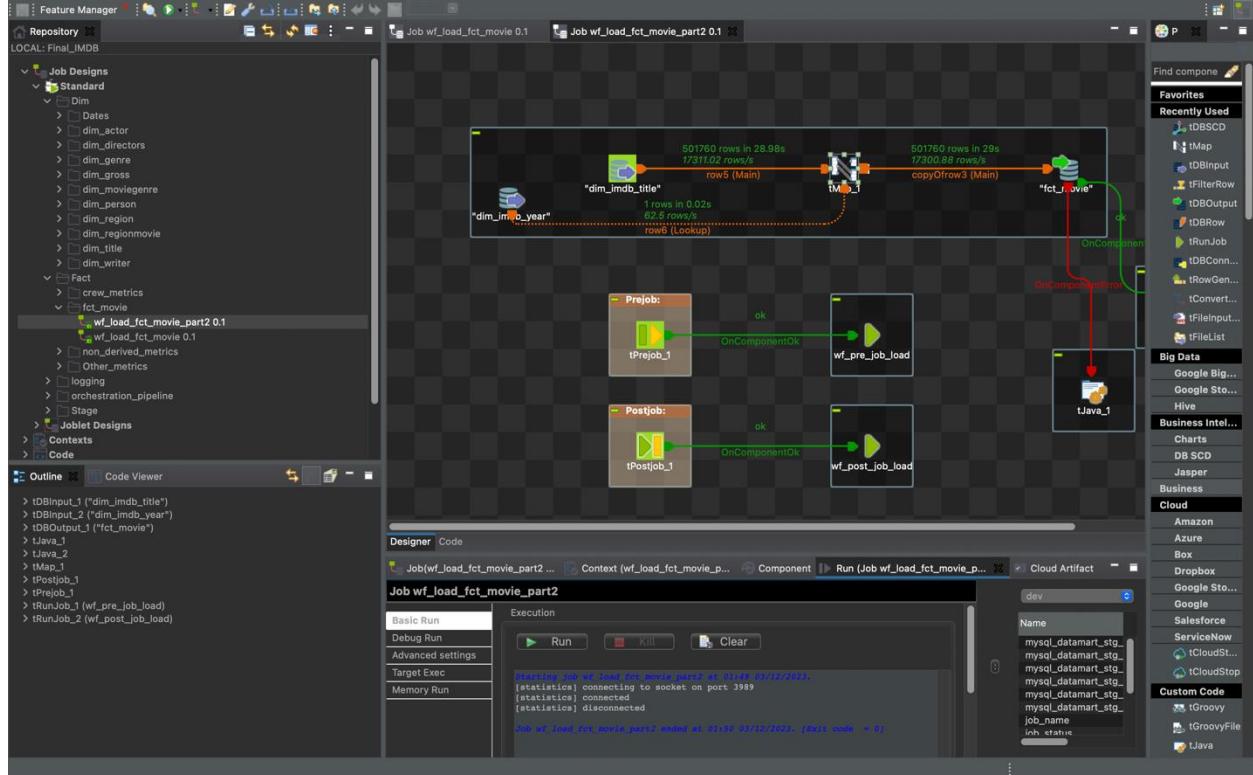
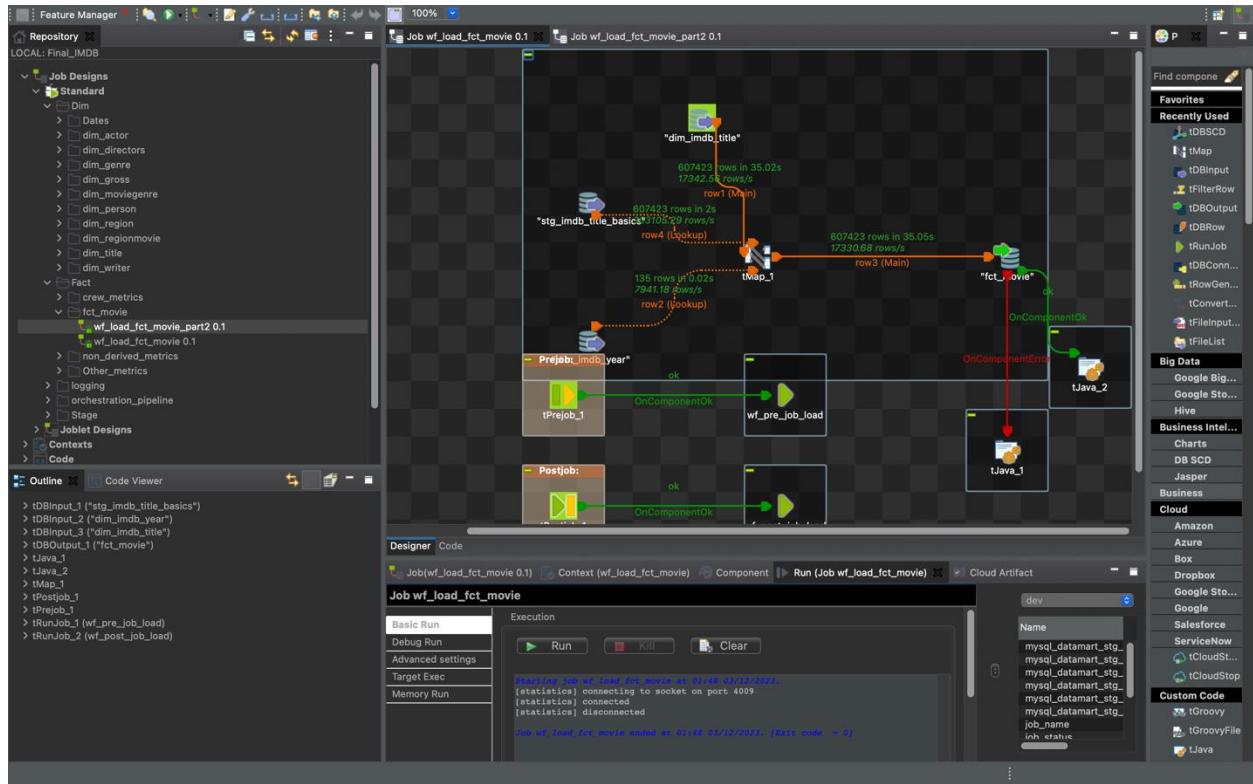


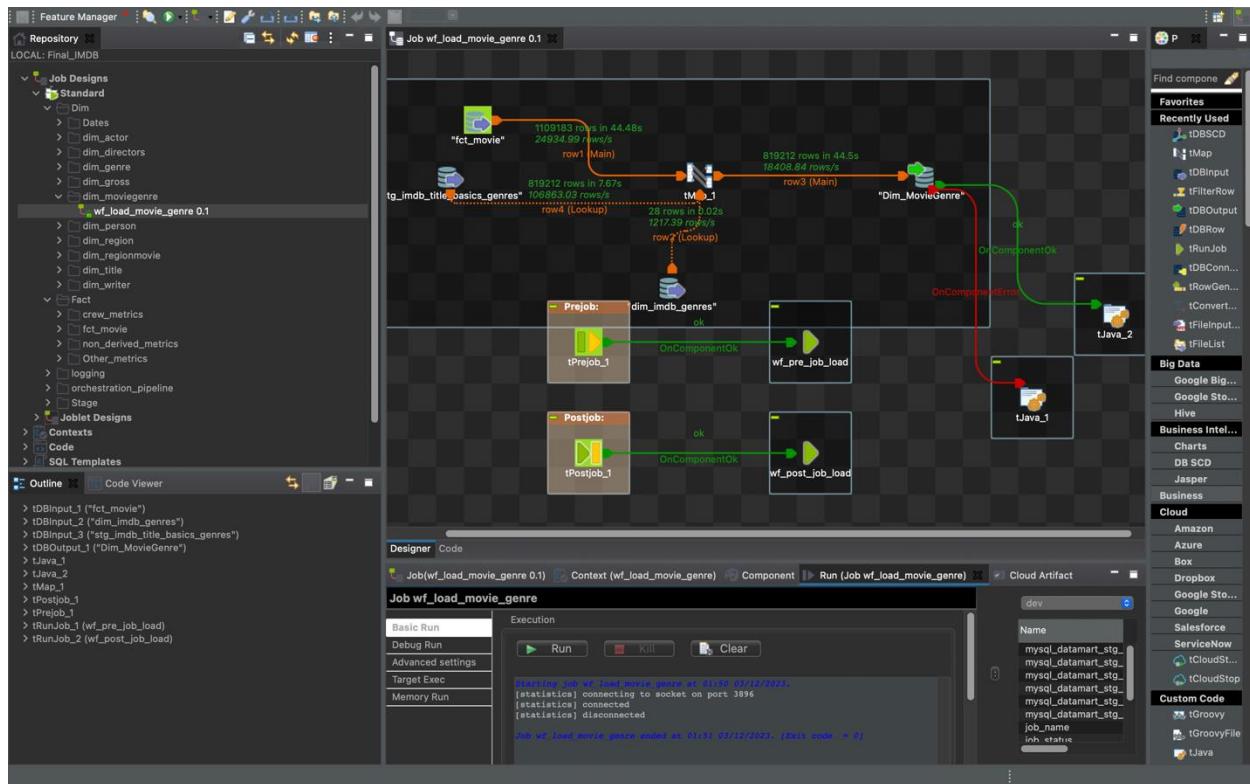




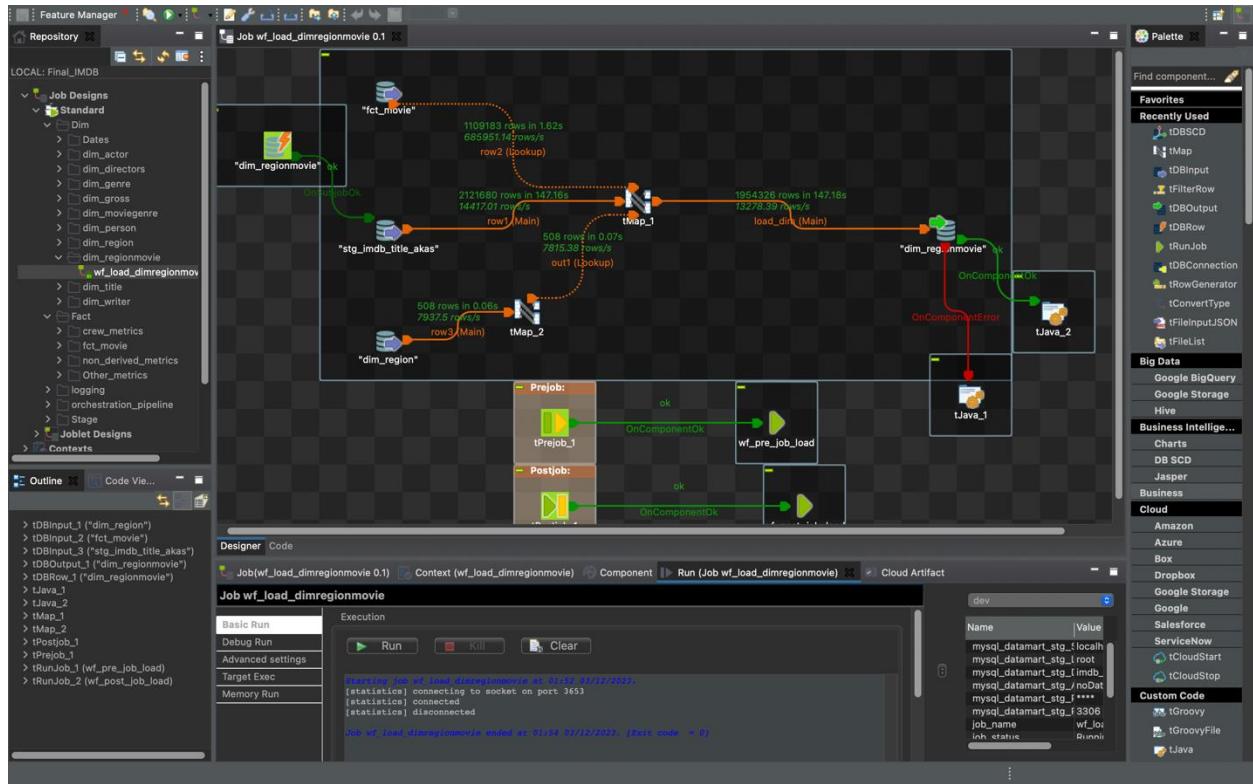


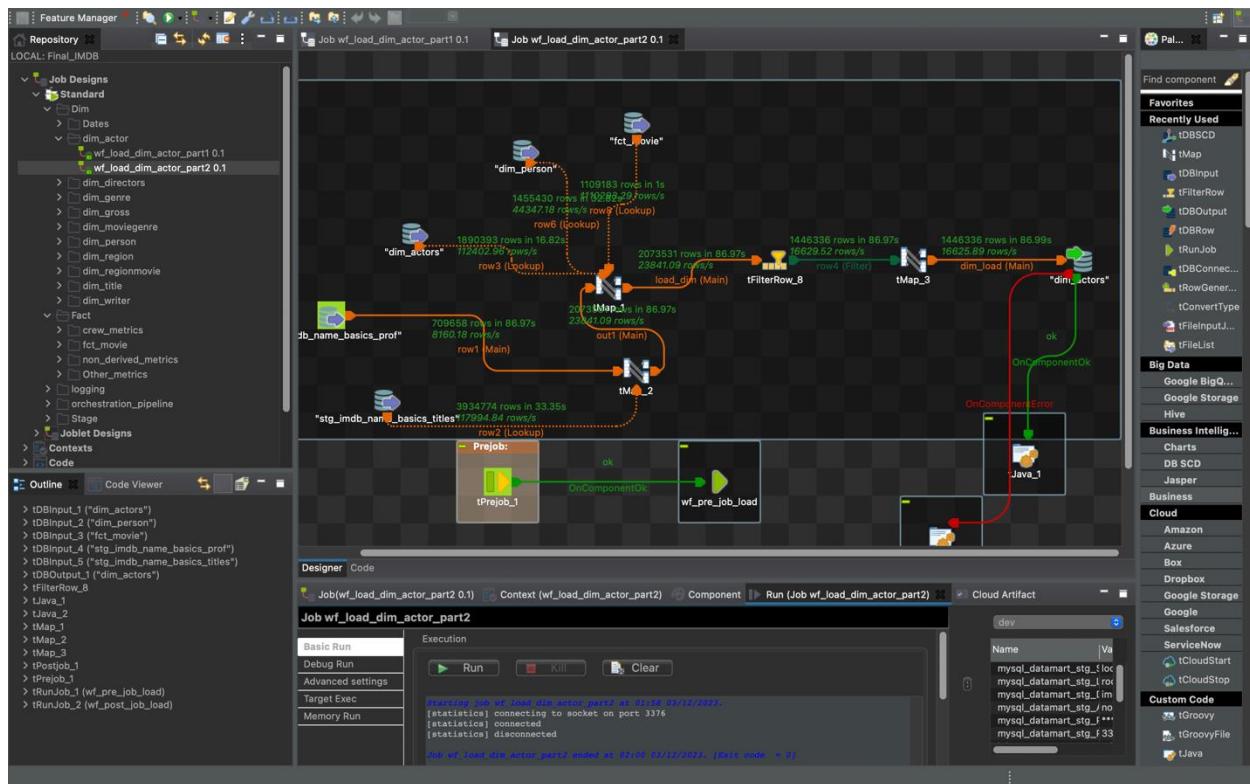
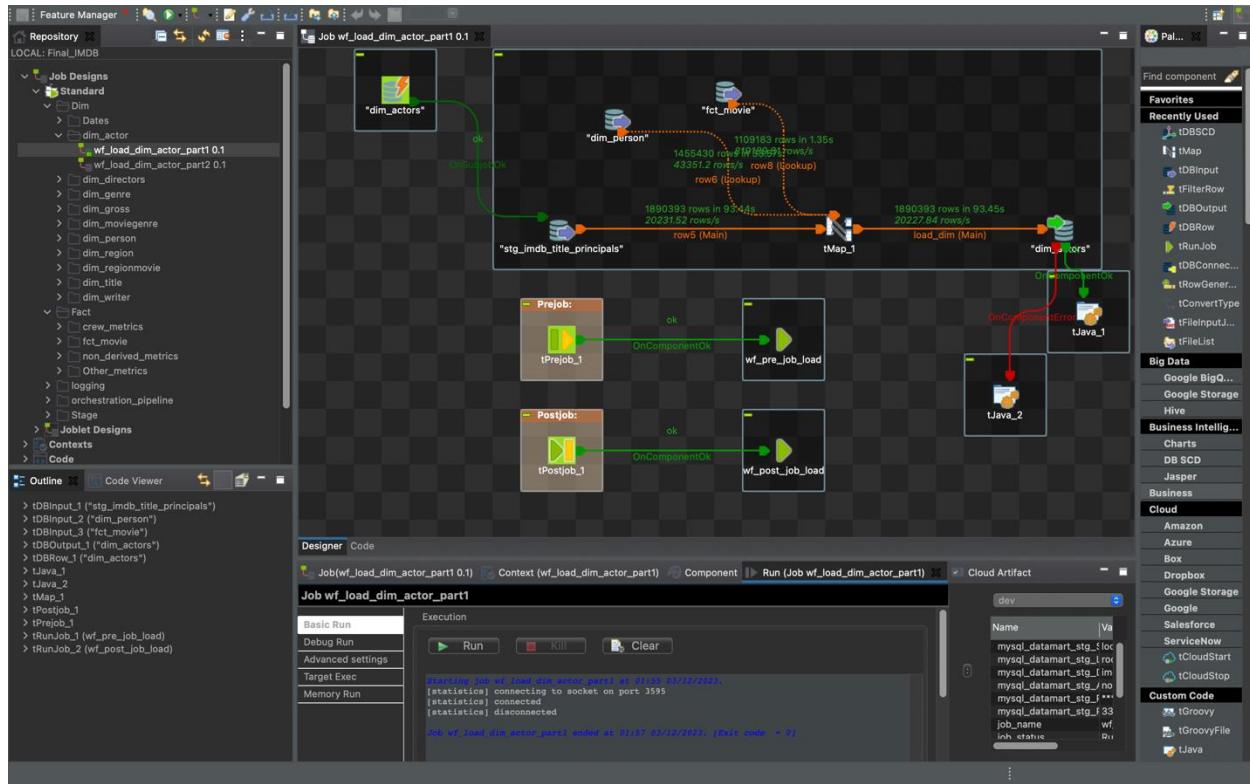




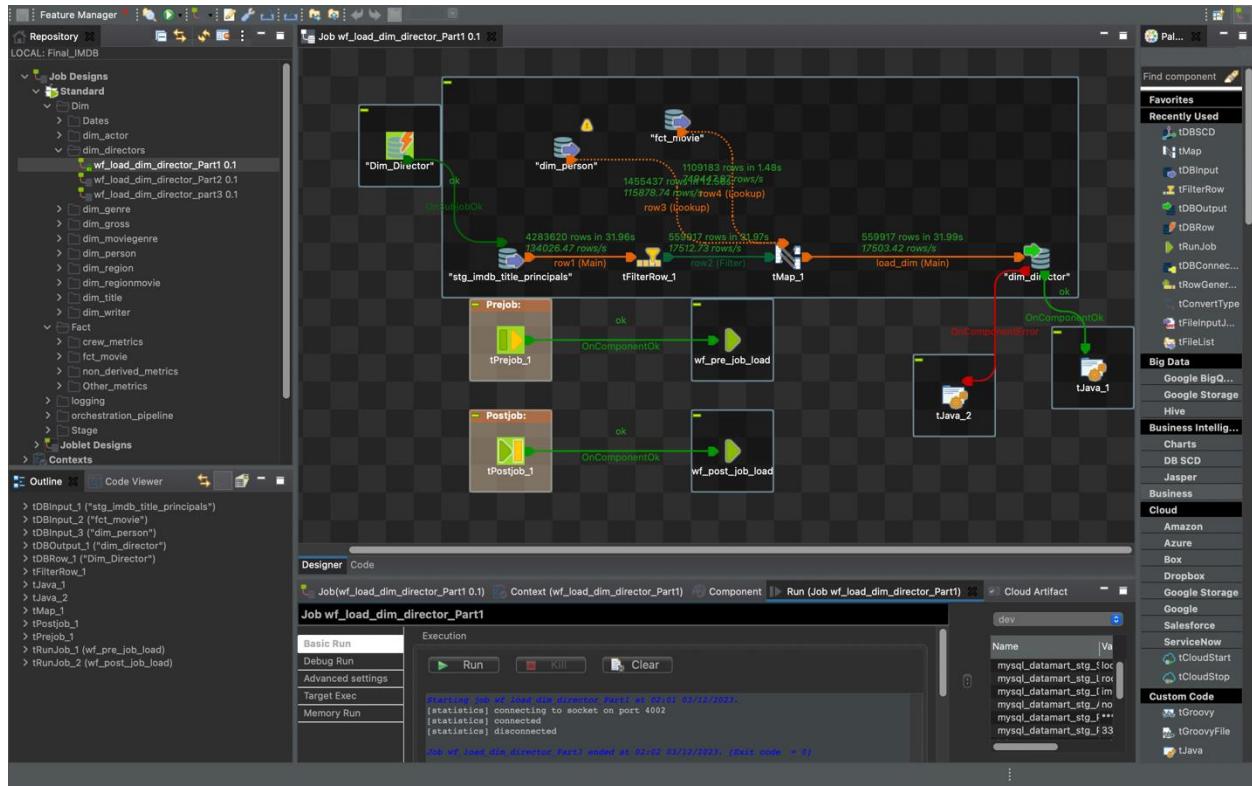


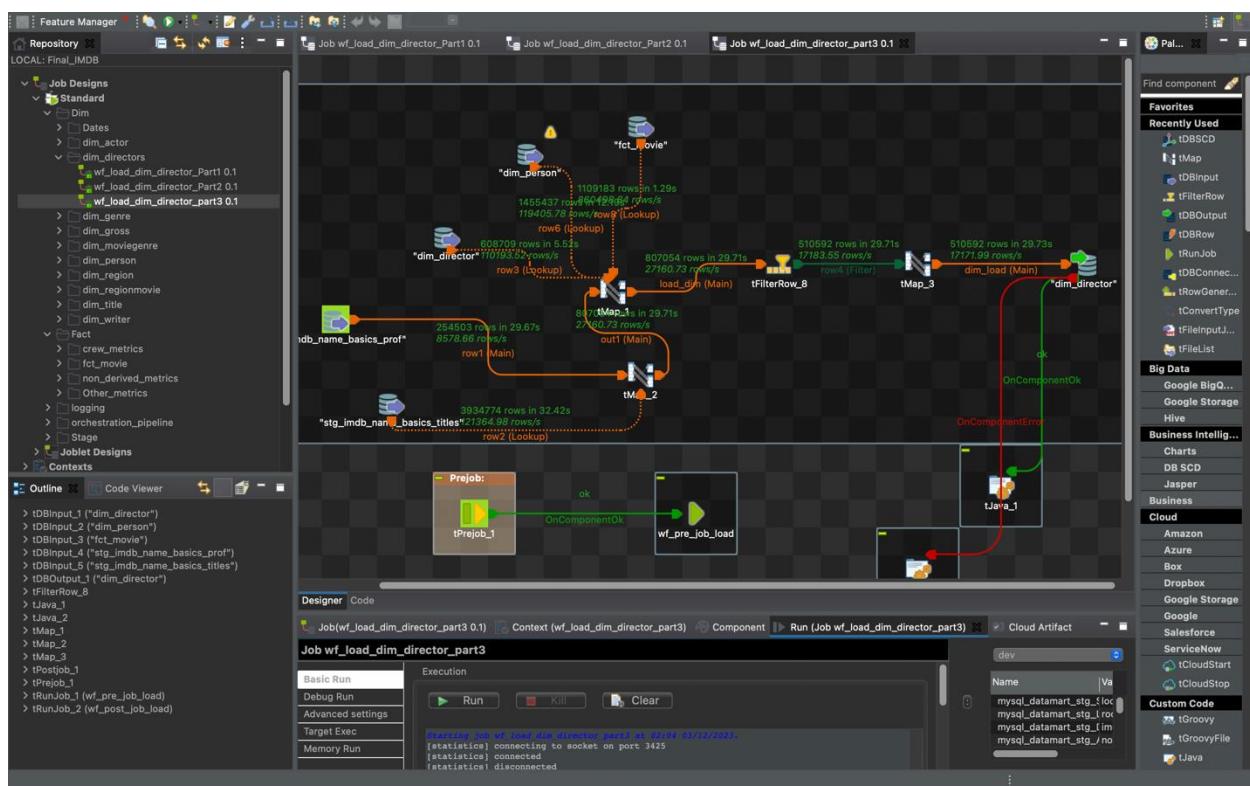
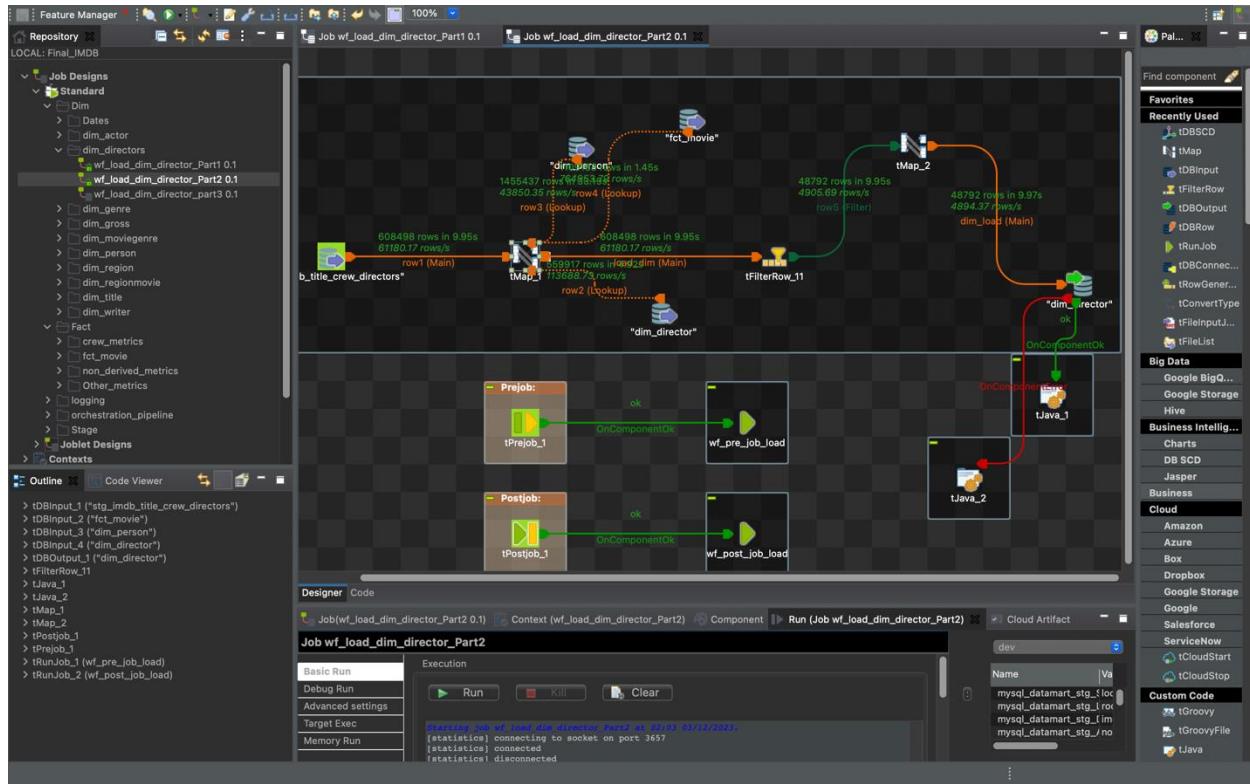
## Dim\_Imdb\_actor:



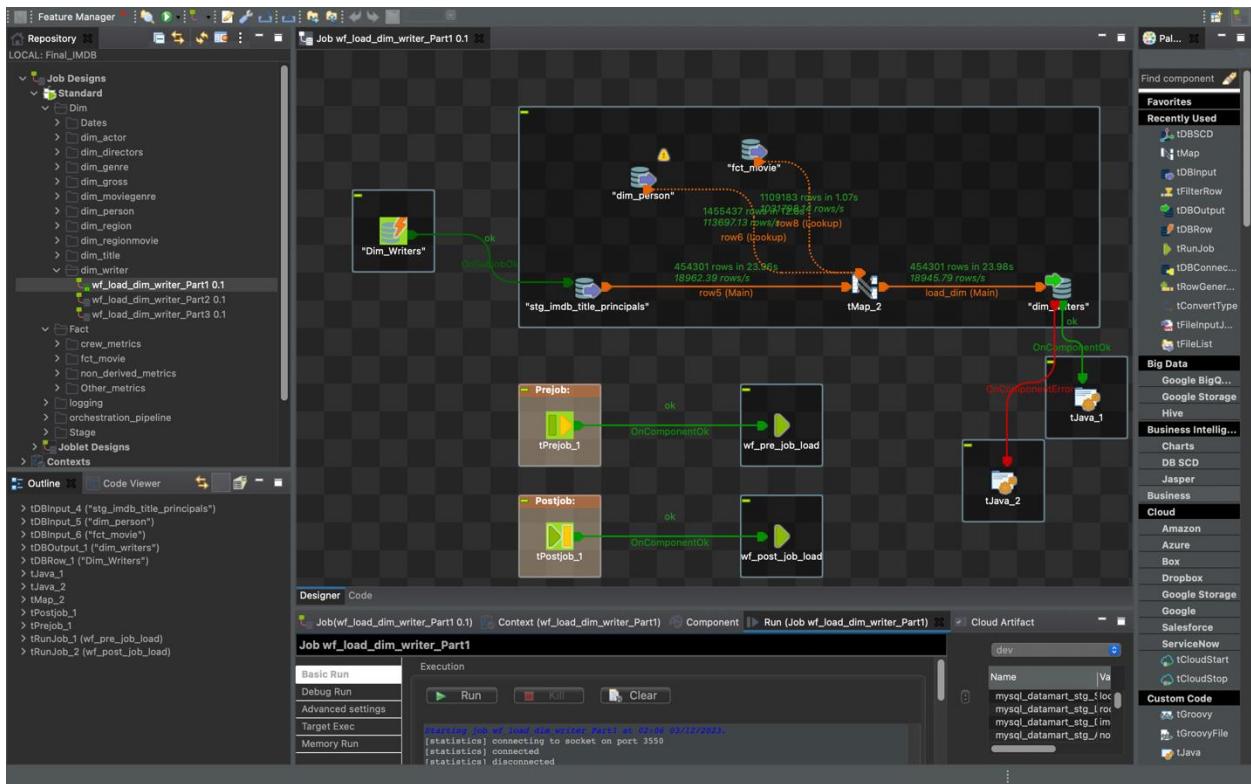


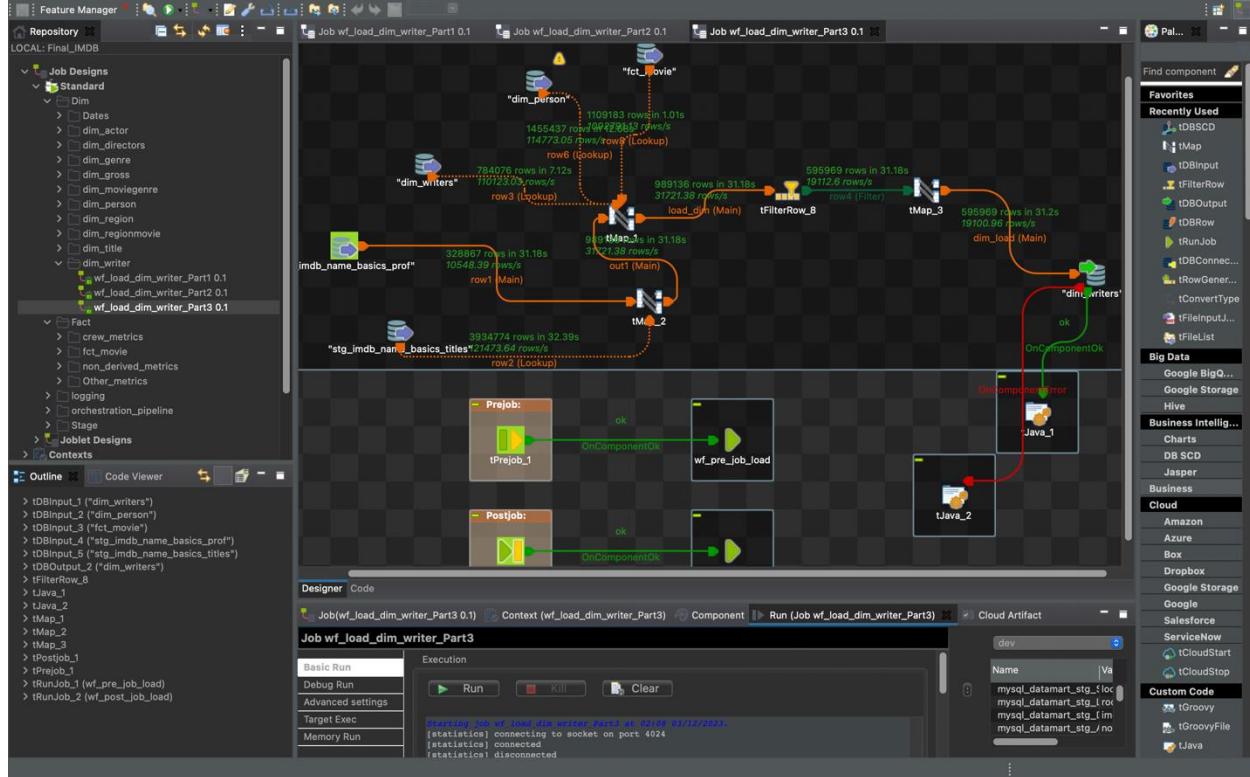
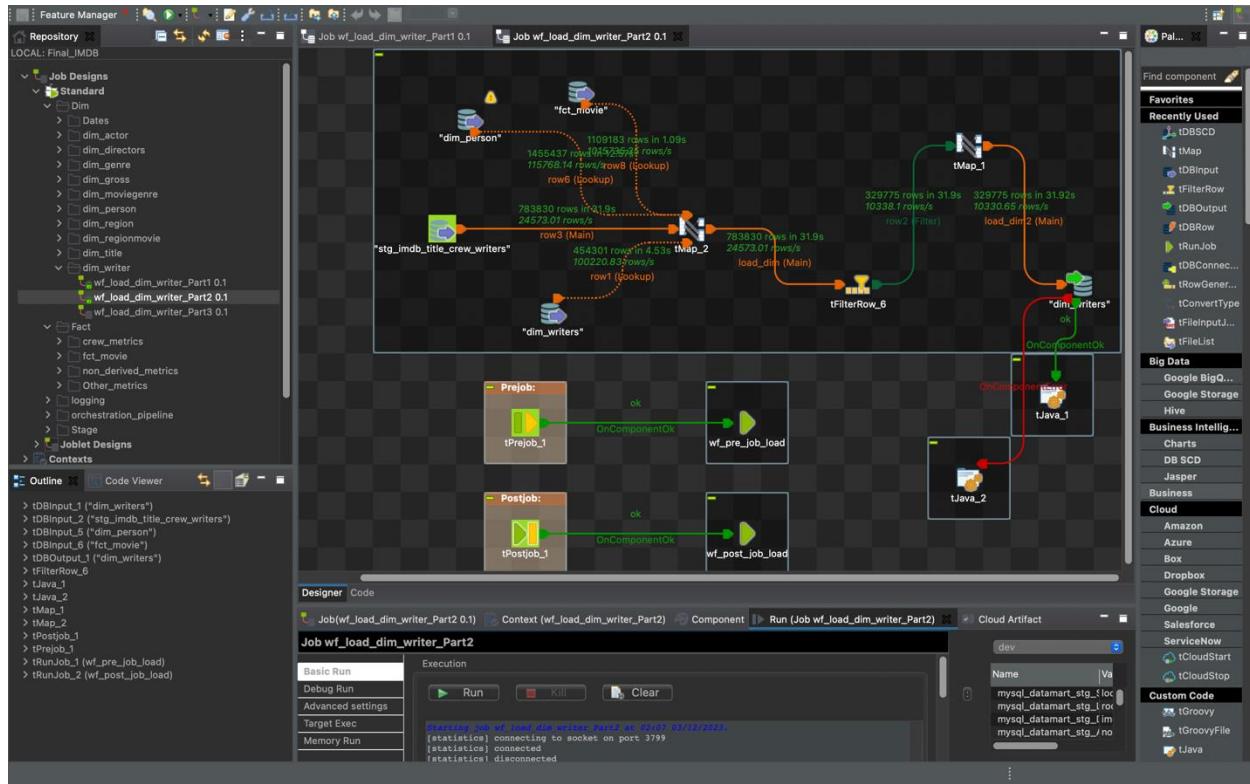
## Dim\_Imdb\_director :

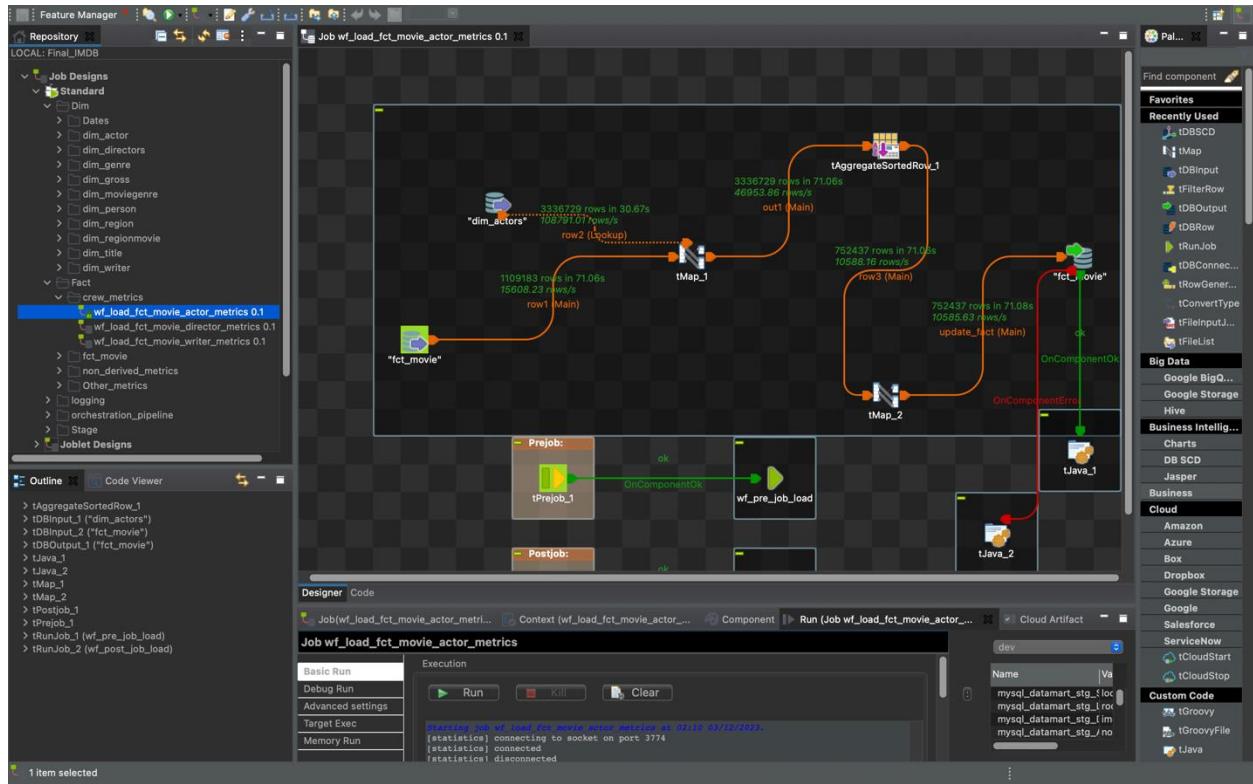


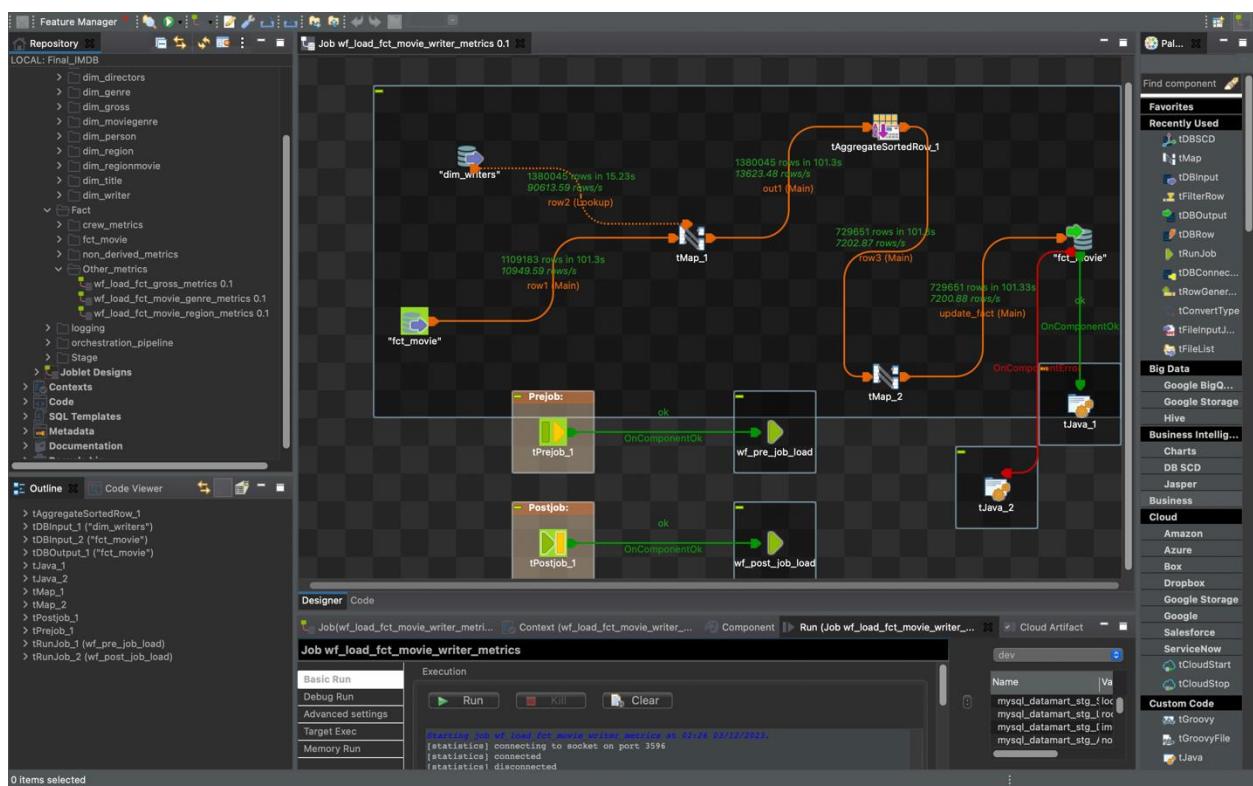
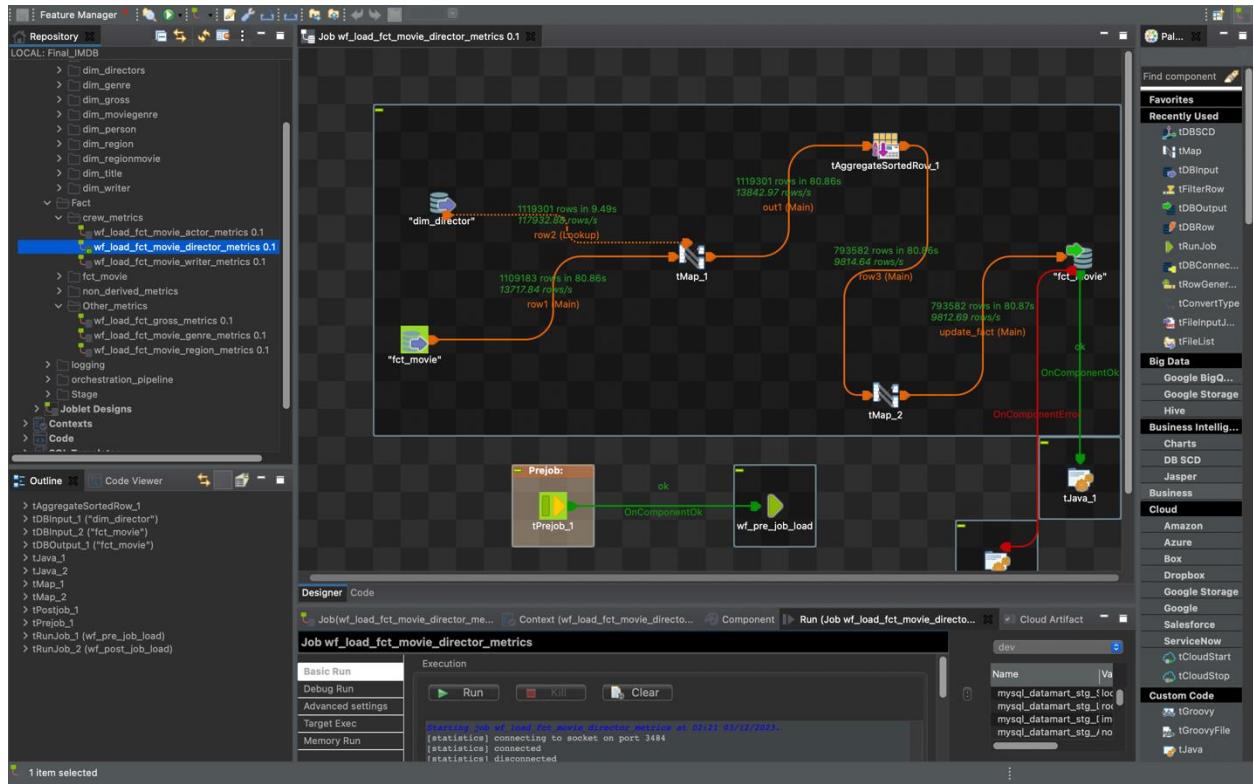


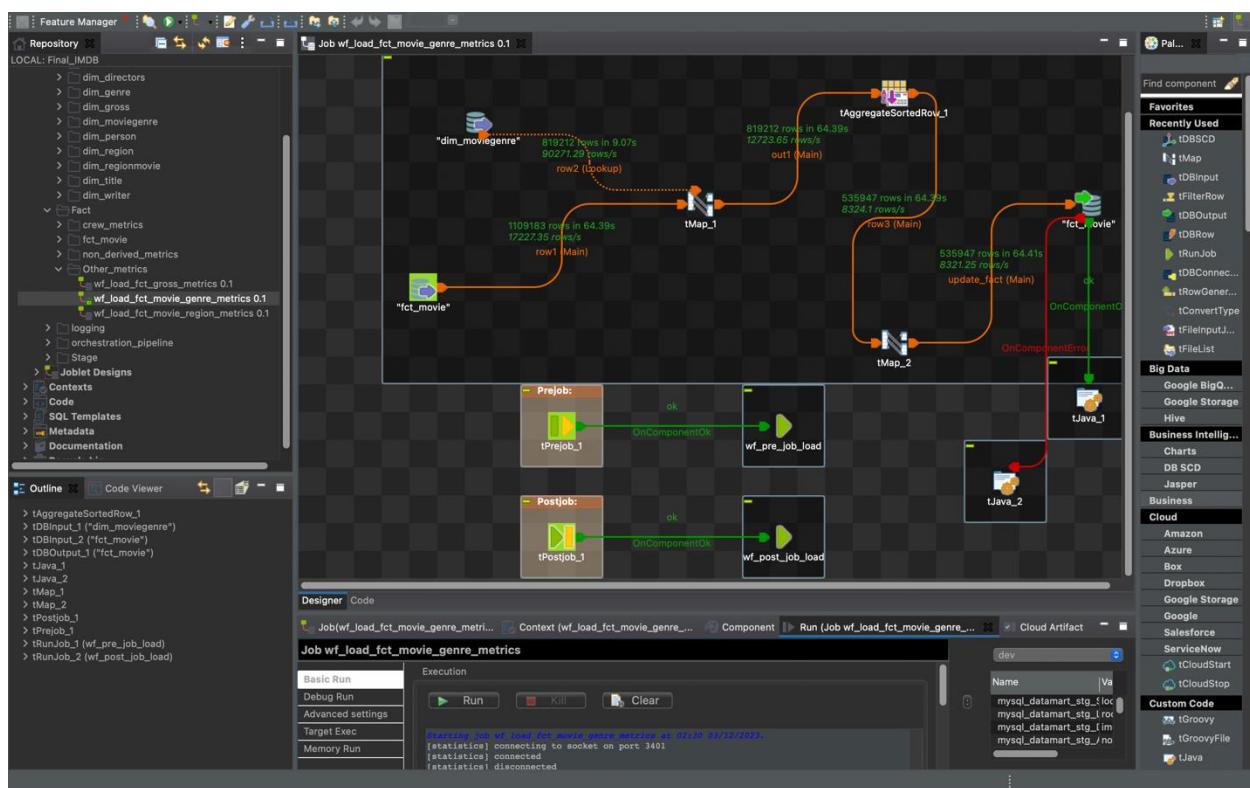
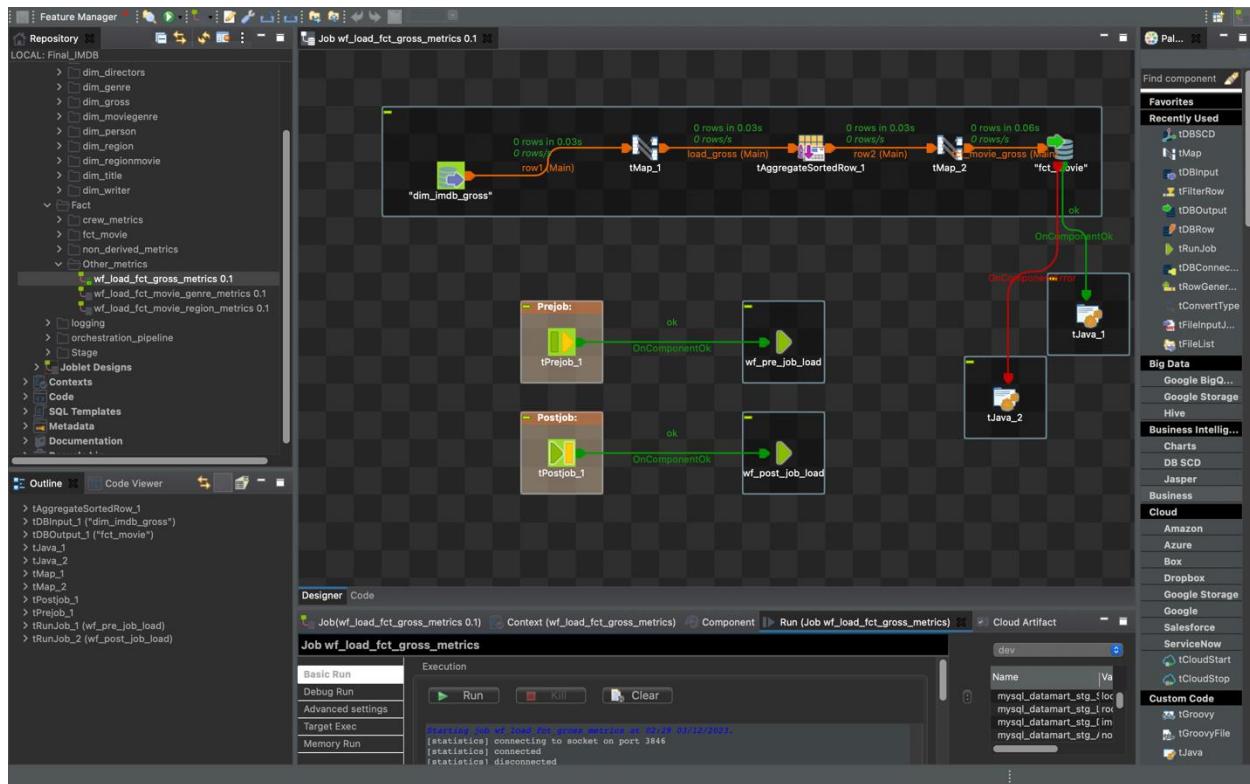
## Dim\_Imdb\_writer :

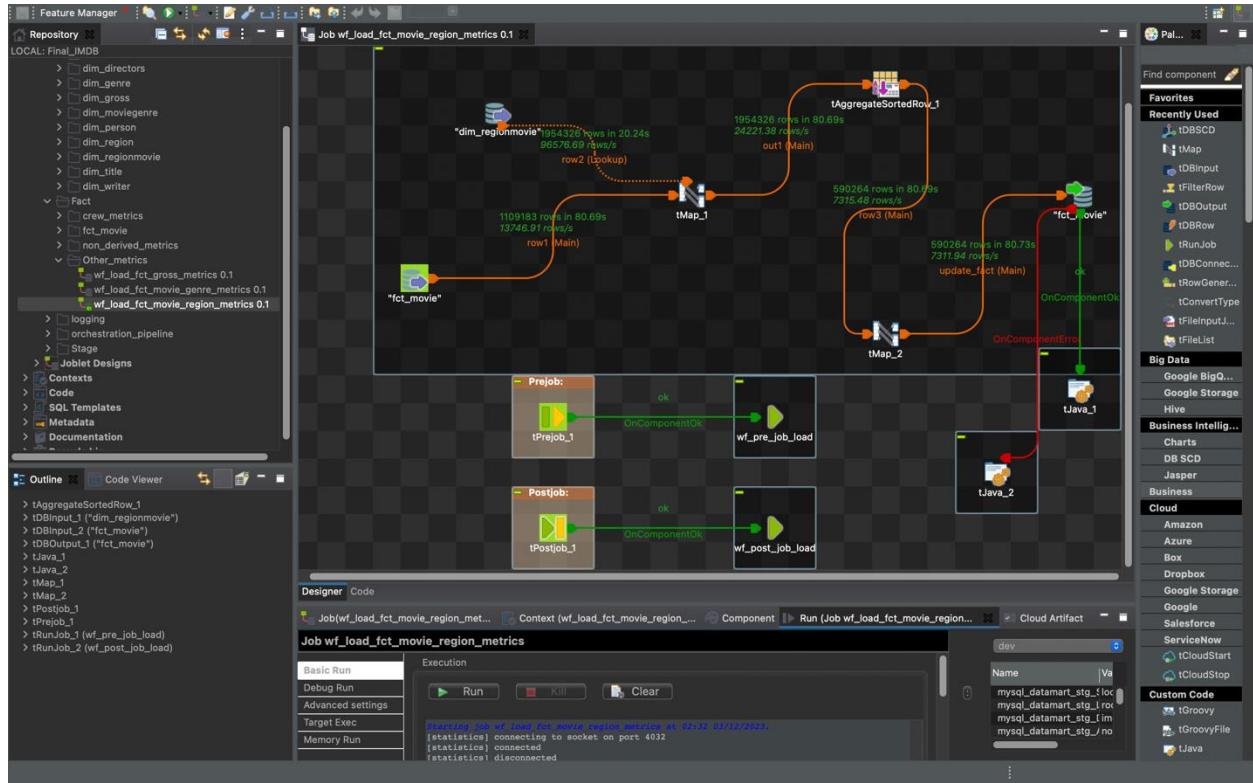


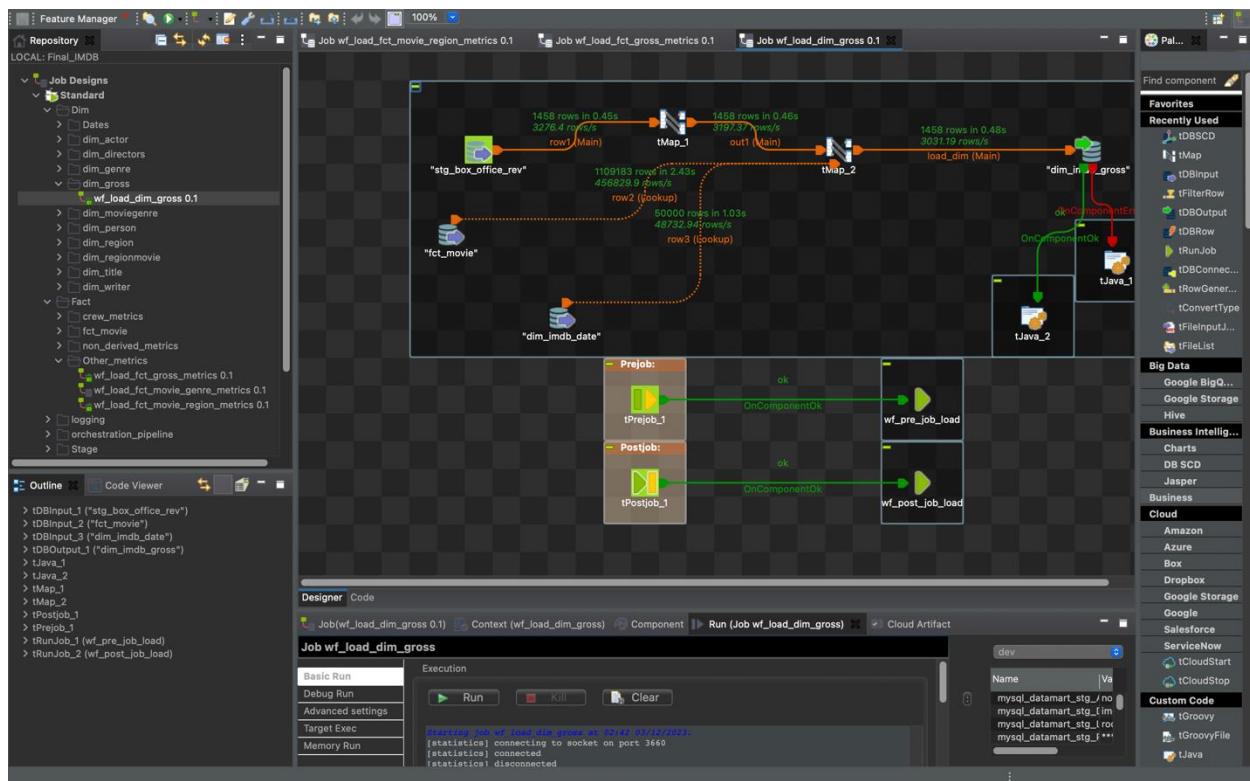
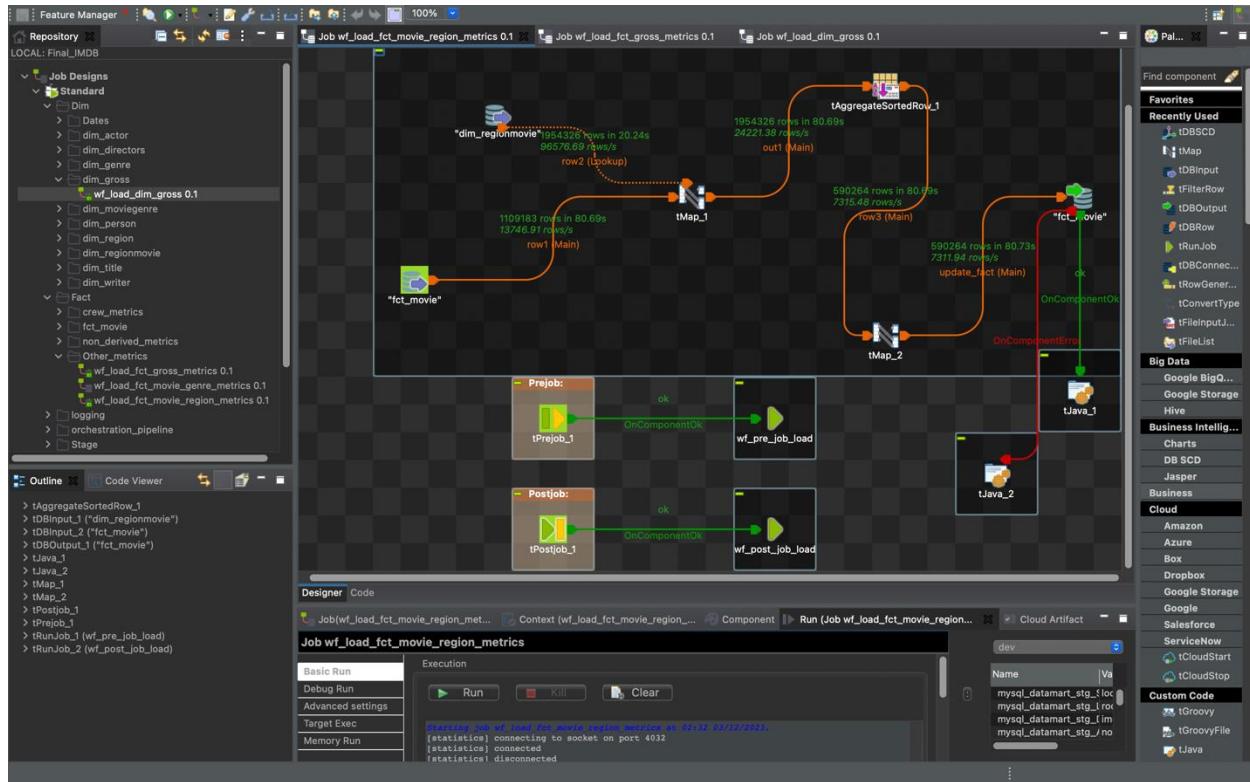


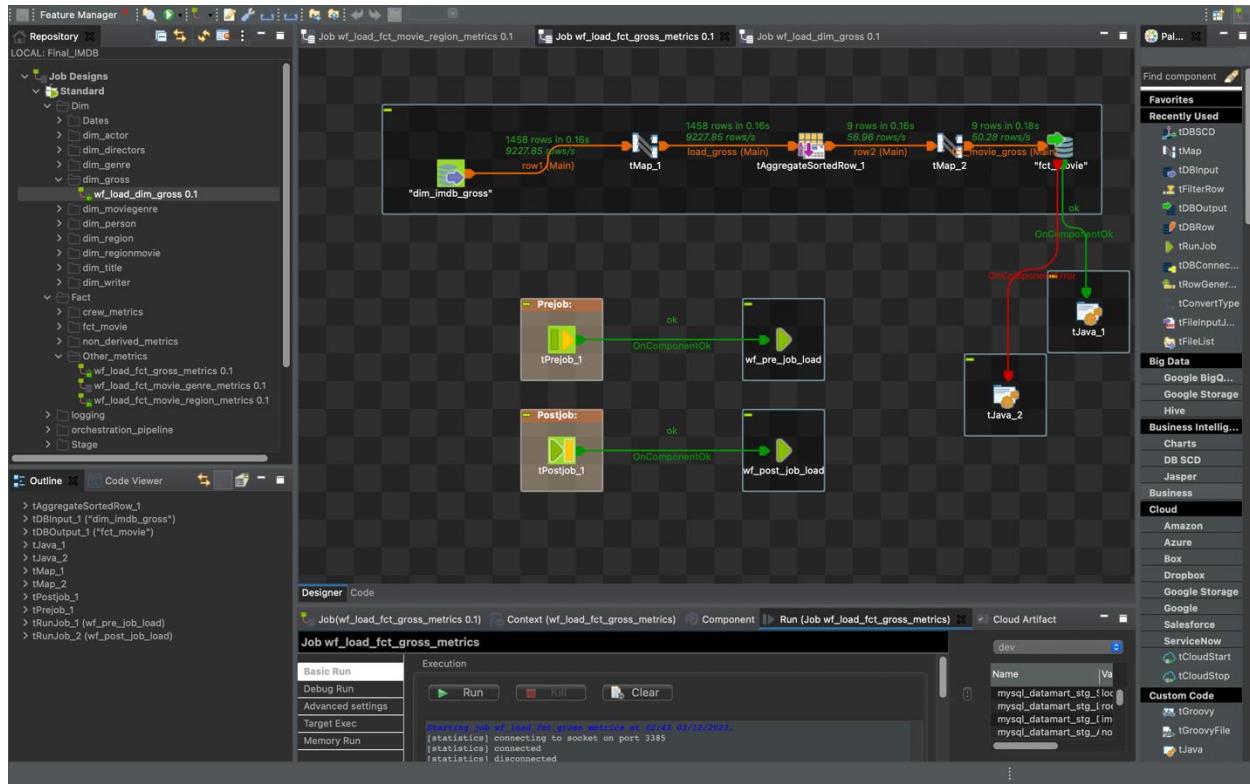












## Dim\_region query -

```

15 •   select * from fct_movie where tconst = "tt00000574";
16
17 •   use `imdb_movies_datamart`;
18 •   Select * from dim_region;
19 •   Select count(*) from dim_region;
20
21
22 •   select * from dim_imdb_title;
23 •   select * from dim_imdb_genres;
24 •   select * from dim_imdb_year;
25 •   select * from dim_person;
26 •   select * from dim_region;
27 •   select * from fct_movie;
28 •   select * from dim_imdb_gross;
29
30 •

```

The screenshot shows the MySQL Result Grid window. The query executed was 'Select count(\*) from dim\_region;'. The result is a single row with the column 'count(\*)' and the value '508'.

count(*)
508

```

15 • select * from fct_movie where tconst = "tt0000574";
16
17 • use `imdb_movies_datamart`;
18 • Select * from dim_region;
19
20
21 • select * from dim_imdb_title;
22 • select * from dim_imdb_genres;
23 • select * from dim_imdb_year;
24 • select * from dim_person;
25 • select * from dim_region;
26 • select * from fct_movie;
27 • select * from dim_imdb_gross;
28
29 • drop table imdb_movies_datamart.fct_movie;
30

```

100% 1:19

**Result Grid** Filter Rows: Search Edit: Export/Import: Result Grid Form Editor

RegionSK	name	code	region	sub_region	ProcessID	DI_Create_DT
1	Afghanistan	AF	Asia	Southern Asia	JK13Bk	2023-12-03 06:22:33
2	Aland Islands	AX	Europe	Northern Europe	JK13Bk	2023-12-03 06:22:33
3	Albania	AL	Europe	Southern Europe	JK13Bk	2023-12-03 06:22:33
4	Algeria	DZ	Africa	Northern Africa	JK13Bk	2023-12-03 06:22:33
5	American Samoa	AS	Oceania	Polynesia	JK13Bk	2023-12-03 06:22:33
6	Andorra	AD	Europe	Southern Europe	JK13Bk	2023-12-03 06:22:33
7	Angola	AO	Africa	Sub-Saharan Africa	JK13Bk	2023-12-03 06:22:33

dim\_region 17

Apply Revert

```

15 • select * from fct_movie where tconst = "tt0000574";
16
17 • use `imdb_movies_datamart`;
18 • Select * from dim_region;
19 • Select count(*) from dim_region;
20
21
22 • select * from dim_imdb_title;
23 • select * from dim_imdb_genres;
24 • select * from dim_imdb_year;
25 • select * from dim_person;
26 • select * from dim_region;
27 • select * from fct_movie;
28 • select * from dim_imdb_gross;
29
30

```

00% 1:23

**Result Grid** Filter Rows: Search Edit: Export/Import: Fetch rows: Result Grid Form Editor

TitleSK	tconst	primaryTitle	originalTitle	DI_Create_DT	Process_id	scd_start	scd_end	scd_version	scd_active
1	tt0000502	Bohemios	Bohemios	2023-12-03 06:34:53	zqrY0s	2023-12-03 06:34:52	9999-01-01 17:00:00	1	1
2	tt0000574	The Story of the Kelly Gang	The Story of the Kelly Gang	2023-12-03 06:34:53	zqrY0s	2023-12-03 06:34:52	9999-01-01 17:00:00	1	1
3	tt0000591	The Prodigal Son	L'enfant prodigue	2023-12-03 06:34:53	zqrY0s	2023-12-03 06:34:52	9999-01-01 17:00:00	1	1
4	tt0000615	Robbery Under Arms	Robbery Under Arms	2023-12-03 06:34:53	zqrY0s	2023-12-03 06:34:52	9999-01-01 17:00:00	1	1
5	tt0000630	Hamlet	Amleto	2023-12-03 06:34:53	zqrY0s	2023-12-03 06:34:52	9999-01-01 17:00:00	1	1
6	tt0000675	Don Quijote	Don Quijote	2023-12-03 06:34:53	zqrY0s	2023-12-03 06:34:52	9999-01-01 17:00:00	1	1
7	tt0000679	The Fairylogue and Radio-Plays	The Fairylogue and Radio-Plays	2023-12-03 06:34:53	zqrY0s	2023-12-03 06:34:52	9999-01-01 17:00:00	1	1

actions\_imdb\_title\_10

```
• select * from dim_imdb_title;
• select count(*) from dim_imdb_title;

• select * from dim_imdb_genres;
• select * from dim_imdb_year;
• select * from dim_person;
• select * from dim_region;
•
```

1:24

Result Grid | Filter Rows: Search Export: | Result Grid | Form Editor

	count(*)
1	1109193

```
25
26 • select * from dim_imdb_genres;
27 • select * from dim_imdb_genres;
28
29
30 • select * from dim_imdb_year;
31 • select * from dim_person;
32 • select * from dim_region;
33 • select * from fct_movie;
34 •
```

00% 1:27

Result Grid | Filter Rows: Search Edit: Export/Import: | Result Grid | Form Editor

genreSK	genres	Process_ID	DL_Create_DT
1	Action	RXHm3u	2023-12-03 06:32:24
2	Adventure	RXHm3u	2023-12-03 06:32:24
3	Biography	RXHm3u	2023-12-03 06:32:24
4	Drama	RXHm3u	2023-12-03 06:32:24
5	Fantasy	RXHm3u	2023-12-03 06:32:24
6	Comedy	RXHm3u	2023-12-03 06:32:24
7	War	RXHm3u	2023-12-03 06:32:24

```
26 • select * from dim_imdb_genres;
27 • select count(*) from dim_imdb_genres;
28
29
30 • select * from dim_imdb_year;
31 • select * from dim_person;
32 • select * from dim_region;
33 • select * from fct_movie;
34 •
```

0% 1:28

Result Grid | Filter Rows: Search Export: | Result Grid | Form Editor

count(*)
28

Result 22

```

28
29
30 • select * from dim_imdb_year;
31 • select count(*) from dim_imdb_year;
32
33
34 •

```

100% 1:31

**Result Grid** Filter Rows: Search Edit: Export/Import: Result Grid Form Editor

YearSK	Year	Decade	ProcessID	DL_Created_DT
1	1905	1900s	iQbRmh	2023-12-03 06:33:51
2	1906	1900s	iQbRmh	2023-12-03 06:33:51
3	1907	1900s	iQbRmh	2023-12-03 06:33:51
4	1908	1900s	iQbRmh	2023-12-03 06:33:51
5	1909	1900s	iQbRmh	2023-12-03 06:33:51
6	1910	1910s	iQbRmh	2023-12-03 06:33:51
7	1912	1910s	iQbRmh	2023-12-03 06:33:51

dim\_imdb\_year 24 | Apply | Revert

Action Output :

```

30 • select * from dim_imdb_year;
31 • select count(*) from dim_imdb_year;
32
33
34 •

```

100% 1:32

**Result Grid** Filter Rows: Export: Result Grid Form Editor

count(*)
135

Result 25 | Read Only

```

34 • select * from dim_person;
35 • select * from dim_person;
36
37
38 •

```

100% 1:35

**Result Grid** Filter Rows: Search Edit: Export/Import: Fetch rows: Result Grid Form Editor

PersonSK	nconst	Name	BirthYear	DeathYear	Process_id	DL_Create_DT	scd_start	scd_end	scd_version	scd_active
1	nm0000001	Fred Astaire	1899	1987	SmSAKm	2023-12-03 06:40:55	2023-12-03 06:40:54	9999-01-01 17:00:00	1	1
2	nm0000002	Lauren Bacall	1924	2014	SmSAKm	2023-12-03 06:40:55	2023-12-03 06:40:54	9999-01-01 17:00:00	1	1
3	nm0000003	Brigitte Bardot	1934	9999	SmSAKm	2023-12-03 06:40:55	2023-12-03 06:40:54	9999-01-01 17:00:00	1	1
4	nm0000004	John Belushi	1949	1982	SmSAKm	2023-12-03 06:40:55	2023-12-03 06:40:54	9999-01-01 17:00:00	1	1
5	nm0000005	Ingmar Bergman	1918	2007	SmSAKm	2023-12-03 06:40:55	2023-12-03 06:40:54	9999-01-01 17:00:00	1	1
6	nm0000006	Ingrid Bergman	1915	1982	SmSAKm	2023-12-03 06:40:55	2023-12-03 06:40:54	9999-01-01 17:00:00	1	1
7	nm0000007	Humphrey Bogart	1899	1957	SmSAKm	2023-12-03 06:40:55	2023-12-03 06:40:54	9999-01-01 17:00:00	1	1

dim\_person 26 | Apply | Revert

```
34 • select * from dim_person;
35 • select count(*) from dim_person;
36
37
38 •
```

100% 1:36

Result Grid Filter Rows: Search Export:

count(*)
1455437

Result Grid Form Editor

Result 27 Read Only

```
38 • select * from dim_region;
39 • select count(*) from dim_region;
40
41
42 •
```

100% 1:39

Result Grid Filter Rows: Search Edit: Export/Import: Form Editor Result Grid

	RegionSK	name	code	region	sub_region	ProcessID	DL_Create_DT
1	Afghanistan	AF	Asia	Southern Asia	JK13Bk	2023-12-03 06:22:33	
2	Aland Islands	AX	Europe	Northern Europe	JK13Bk	2023-12-03 06:22:33	
3	Albania	AL	Europe	Southern Europe	JK13Bk	2023-12-03 06:22:33	
4	Algeria	DZ	Africa	Northern Africa	JK13Bk	2023-12-03 06:22:33	
5	American Samoa	AS	Oceania	Polynesia	JK13Bk	2023-12-03 06:22:33	
6	Andorra	AD	Europe	Southern Europe	JK13Bk	2023-12-03 06:22:33	
7	Angola	AO	Africa	Sub-Saharan Africa	JK13Bk	2023-12-03 06:22:33	

dim\_region 28

Apply Revert

```
37
38 • select * from dim_region;
39 • select count(*) from dim_region;
40 |
41
42 •
```

100% 1:40

Result Grid Filter Rows: Search Export:

count(*)
508

Result 29

```
44
45 • select * from dim_imdb_gross;
46 • select count(*) from dim_imdb_gross;
47
48 • drop table imdb_movies_datamart.fct_movie;
49
50 • select count(*) from dim_imdb_year y
51 join fct_movie f
52
```

```
44
45 • select * from dim_imdb_gross;
46 • select count(*) from dim_imdb_gross;
47
48 • drop table imdb_movies_datamart.fct_movie;
49
50 • select count(*) from dim_imdb_year y
51 join fct_movie f
52
```

100% 1:47

Result Grid Filter Rows: Search Export:

count(*)
1458

Form Editor

```
41
42 • select * from fct_movie;
43 • select count(*) from fct_movie;
44
45 • select * from dim_imdb_gross;
46 •
```

100% 1:43

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows: Form Editor Result Grid

MoviesSK	tconst	ReleaseYearSK	isAdult	RuntimeMin	NoofDirectors	NoofActors	NoofWriters	TotalGross	AvgRating	NoofRegions	NoofGenres	NoofVotes	ProcessID	DL_Create
1	tt0000502	1	0	100	1	2	3	-1	4.5	1	-1	14	afOy8d	2023-12-03
2	tt0000574	2	0	70	1	7	2	-1	6	8	3	759	afOy8d	2023-12-03
3	tt0000591	3	0	90	1	5	1	-1	4.6	2	1	17	afOy8d	2023-12-03
4	tt0000615	3	0	-1	1	6	2	-1	4.5	1	1	23	afOy8d	2023-12-03
5	tt0000630	4	0	-1	1	1	1	-1	3.8	4	1	24	afOy8d	2023-12-03
6	tt0000675	4	0	-1	1	-1	2	-1	4.9	1	1	19	afOy8d	2023-12-03
7	tt0000679	4	0	120	2	5	4	-1	5.3	2	2	65	afOy8d	2023-12-03

```
41
42 •   select * from fct_movie;
43 •   select count(*) from fct_movie;
44 |
45 •   select * from dim_imdb_gross;
46 •
```

100% 1:44

Result Grid Filter Rows: Search Export:

count(*)
1109183

Result Grid Form Editor

```
45
Execute the selected portion of the script or genre;
everything, if there is no selection
46 •   select count(*) from dim_moviegenre;
47 |
48
49 •   select * from dim_imdb_gross;
50 •   select count(*) from dim_imdb_gross;
51
52 •
```

100% 1:47

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows:

MovieGenreSK	MovieSK	GenreSK	ProcessID	DL_Create_DT
1	2	1	A3XBtd	2023-12-03 06:50:50
2	2	2	A3XBtd	2023-12-03 06:50:50
3	2	3	A3XBtd	2023-12-03 06:50:50
4	3	4	A3XBtd	2023-12-03 06:50:50
5	4	4	A3XBtd	2023-12-03 06:50:50
6	5	4	A3XBtd	2023-12-03 06:50:50
7	6	4	A3XBtd	2023-12-03 06:50:50

dim\_moviegenre 34

Apply Revert

```
46 • select * from dim_moviegenre;
47 • select count(*) from dim_moviegenre;
48
49 • select * from dim_imdb_gross;
50 • select count(*) from dim_imdb_gross;
51
52 •
```

100% 1:48

Result Grid Filter Rows: Search Export:

count(*)
819212

Result Grid Form Editor

```
49
50 • select * from dim_imdb_date;
51 • select count(*) from dim_imdb_date;
52
53 • select * from dim_imdb_gross;
54 • select count(*) from dim_imdb_gross;
55
56 • drop table imdb_movies_datamart.fct_movie;
57
58 • select count(*) from dim_imdb_year y
59
```

100% 1:51

Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows: Result Grid Form Editor

DateSK	Date	Year	Month	Decade	Season	ProcessID	DL_Create_DT
1	1990-01-02	1990	Jan	1980s	Spring	gbPt5j	2023-12-03 06:33:02
2	1990-01-03	1990	Jan	1980s	Spring	gbPt5j	2023-12-03 06:33:02
3	1990-01-04	1990	Jan	1980s	Spring	gbPt5j	2023-12-03 06:33:02
4	1990-01-05	1990	Jan	1980s	Spring	gbPt5j	2023-12-03 06:33:02
5	1990-01-06	1990	Jan	1980s	Spring	gbPt5j	2023-12-03 06:33:02
6	1990-01-07	1990	Jan	1980s	Spring	gbPt5j	2023-12-03 06:33:02
7	1990-01-08	1990	Jan	1980s	Spring	gbPt5j	2023-12-03 06:33:02

Result Grid Form Editor

```
50 • select * from dim_imdb_date;
51 • select count(*) from dim_imdb_date;
52
53 • select * from dim_imdb_gross;
54 • select count(*) from dim_imdb_gross;
55
56 • drop table imdb_movies_datamart.fct_movie;
57
58 • select count(*) from dim_imdb_year y
59
```

100% 1:52

Result Grid Filter Rows: Search Export: Result Grid Form Editor

count(*)
50000

Result 37 Read Only

```

55
56 • select * from dim_regionmovie;
57 • select count(*) from dim_regionmovie;
58
100% 1:57 | Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows: Result Grid Form Editor
Result Grid
RegionMovieSK RegionSK MovieSK Process_id DI_Create_DT
1 464 1 7WQSux 2023-12-03 06:52:31
2 356 2 7WQSux 2023-12-03 06:52:31
3 268 2 7WQSux 2023-12-03 06:52:31
4 452 2 7WQSux 2023-12-03 06:52:31
5 489 2 7WQSux 2023-12-03 06:52:31
6 455 2 7WQSux 2023-12-03 06:52:31
7 490 2 7WQSux 2023-12-03 06:52:31
dim_regionmovie 38

```

```

55
56 • select * from dim_regionmovie;
57 • select count(*) from dim_regionmovie;
58
59 •
100% 1:58 | Result Grid Filter Rows: Search Export: Result Grid Form Editor
Result Grid
count(*)
1954326
Result 20

```

```

58
59 • select * from dim_writers;
60 • select count(*) from dim_writers;
61
62
100% 1:60 | Result Grid Filter Rows: Search Edit: Export/Import: Fetch rows: Result Grid Form Editor
Result Grid
WriterSK MoviesSK PersonSK nconst ProcessID DI_Create_DT
1 1 220366 nm0657268 8GuqtS 2023-12-03 07:06:51
2 1 225962 nm0675388 8GuqtS 2023-12-03 07:06:51
3 4 37655 nm0092809 8GuqtS 2023-12-03 07:06:51
4 5 634 nm0000636 8GuqtS 2023-12-03 07:06:51
5 6 56571 nm0148859 8GuqtS 2023-12-03 07:06:51
6 15 312658 nm0957652 8GuqtS 2023-12-03 07:06:51
7 20 634 nm0000636 8GuqtS 2023-12-03 07:06:51
dim_writers 40
Action Output

```

```
58
59 • select * from dim_writers;
60 • select count(*) from dim_writers;
61
62
```

Result Grid | Filter Rows: Search Export:

count(*)
1380045

Result Grid | Form Editor

```
61
62 • select * from etl_job;
63
64
65 • drop table imdb_movies_datamart.fct_movie;
66
```

Result Grid | Filter Rows: Search Edit: Export/Import:

job_id	job_name	table_name	schedule	created_by	created_date
1	wf_load_dim_imdb_date	dim_imdb_date	DAILY	root@localhost	2023-12-03 01:27:45
2	wf_load_dim_year	dim_imdb_year	DAILY	root@localhost	2023-12-03 01:27:45
3	wf_load_dim_actor_Part1	dim_actors	DAILY	root@localhost	2023-12-03 01:27:45
4	wf_load_dim_actor_Part2	dim_actors	DAILY	root@localhost	2023-12-03 01:27:45
5	wf_load_dim_director_Part1	dim_director	DAILY	root@localhost	2023-12-03 01:27:45
6	wf_load_dim_director_Part2	dim_director	DAILY	root@localhost	2023-12-03 01:27:45
7	wf_load_dim_director_Part3	dim_director	DAILY	root@localhost	2023-12-03 01:27:45

Result Grid | Form Editor

```
62 • select * from etl_job;
63 • select count(*) from etl_job;
64
65 • drop table imdb_movies_datamart.fct_movie;
66
```

Result Grid | Filter Rows: Search Export:

count(*)
47

Result Grid | Form Editor

```
64
65 • select * from etl_job_log;
66 • select count(*) from etl_job_log;
67
```

Result Grid | Filter Rows: Search Edit: Export/Import:

job_log_id	job_id	process_id	job_status	created_by	created_date
1	15	nWMF0	Running	asawari	2023-12-03 06:29:28
2	15	nWMF0	Successful	asawari	2023-12-03 06:29:28
3	8	RXHm3u	Running	asawari	2023-12-03 06:32:24
4	8	RXHm3u	Successful	asawari	2023-12-03 06:32:24
5	1	gbPt5j	Running	asawari	2023-12-03 06:33:02
6	1	gbPt5j	Successful	asawari	2023-12-03 06:33:05
7	2	iQbRmh	Running	asawari	2023-12-03 06:33:50

etl\_job\_log 44 | Apply | Revert

```
65 • select * from etl_job_log;
66 • select count(*) from etl_job_log;
67
68
```

100% 1:67

Result Grid Filter Rows: Search Export: Result Grid Form Editor

count(*)
70