

NORTHEASTERN UNIVERSITY

INFORMATION SYSTEMS FALL '23



IMDB Database Project

DESIGNING ADVANCED DATA ARCHITECTURES FOR
BUSINESS INTELLIGENCE

Guided By:

Prof. Naveen Kuragayala

Group 13:

Shreyasi Wakankar - 002771284

Muskan Srivastava - 002794929

Pramita Sandhyan - 002766881

Rhea Bajpai - 002702927

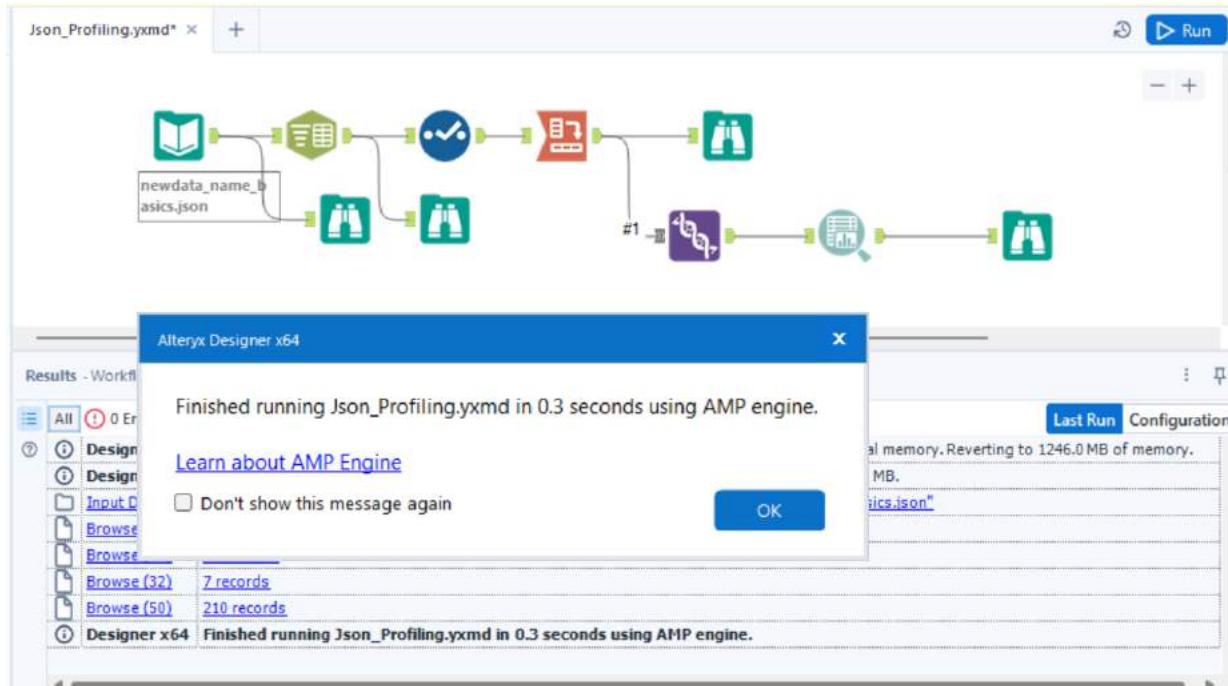
List of Tables:

r No.	Table Name	Table Type	Count
Source: MySQL Script			

1	stg_imdb_name_basics	Staging Table	14,54,602
2	stg_imdb_title_basics	Staging Table	6,07,423
3	stg_imdb_title_akas	Staging Table	26,12,798
4	stg_imdb_title_crew	Staging Table	6,07,423
5	stg_imdb_title_principal	Staging Table	42,83,620
6	stg_imdb_title_rating	Staging Table	2,77,171
Source: tsv File			
1	stg_movie	Staging Table	265
Source: JSON			
1	stg_newdata_name_baisc	Staging Table	7
2	stg_newdata_title_baisc	Staging Table	10
Dimensional Model Tables			
1	dim_person_profession	Dimension Table	2544773
2	dim_date	Dimension Table	10000
3	dim_person	Dimension Table	1454602
4	dim_person_knownfor	Dimension Table	3934774
5	dim_titleregion	Dimension Table	2612798
6	dim_crew	Dimension Table	4283620
7	dim_genre	Dimension Table	
8	im_movie_title_json_with_scd	Dimension Table	
9	dim_movie_title_scd	Dimension Table	10
10	dim_movie_title	Dimension Table	890688
11	fct_rating	Fact Table	277171
12	fct_IMDB	Fact Table	872504
13	fct_Movie_Earnings	Fact Table	265

Data Profiling in Alteryx:

1. newdata_name_basics.json



Observations:

3 of 3 Fields				<input checked="" type="checkbox"/>	<input type="checkbox"/>	210 records displayed, 3,192 bytes	X	✓	Search
Record	FieldName	Name	Value						
20	1	Uniques	7						
21	1	Unique Values	0						
22	nconst	Name	nconst						
23	nconst	Data Type	V_WString						
24	nconst	Size	2048						
25	nconst	Source	CrossTab:Header:Column_Name:nconst:Concat						
26	nconst	Description	[Null]						
27	nconst	OKs	7						
28	nconst	Nulls	0						
29	nconst	Non-Nulls	7						
30	nconst	Blanks	0						
31	nconst	Values with Leading Whitespace	0						
32	nconst	Values with Trailing Whitespace	0						
33	nconst	Values with Both Whitespace	0						
34	nconst	Average Length	9.0						
35	nconst	Longest Length	9						

Profile

[← FieldName](#)

Distinct Values

1	21	
DI_CreateDatetime	21	
DI_JobId	21	
DI_JobName	21	
birthYear	21	
deathYear	21	
knownForTitles	21	
nconst	21	
primaryName	21	
primaryProfession	21	

[.FieldName](#) [×](#)

Summary

Type	Records	Data Type Size
V_WString	210	1,073,741,823

● Ok	210	100.00%
Unique	10	4.76%
● Null	0	0.00%
Not Ok	0	0.00%
● Empty	0	0.00%

Length Statistics

Min	1
Max	17
Average	10.20
Shortest Value	1
Longest Value	primaryProfession
First Alphanumeric Value	1
Last Alphanumeric Value	primaryProfession
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

Frequent Values

Value by count

10 out of 10 values



Profile

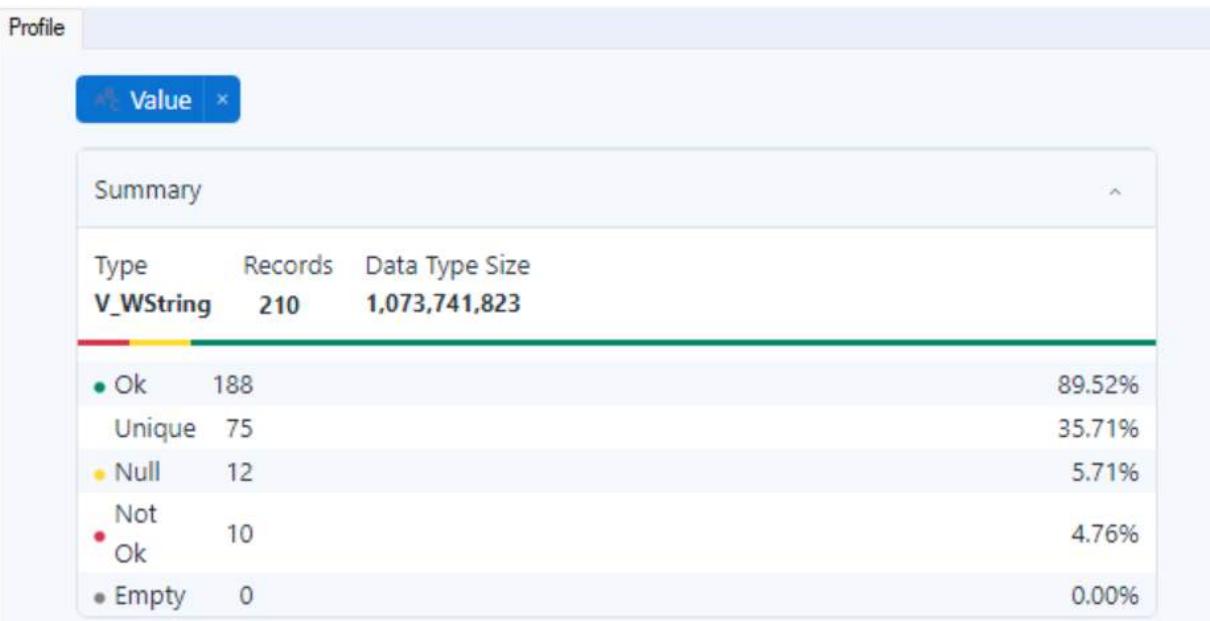
Name

Summary

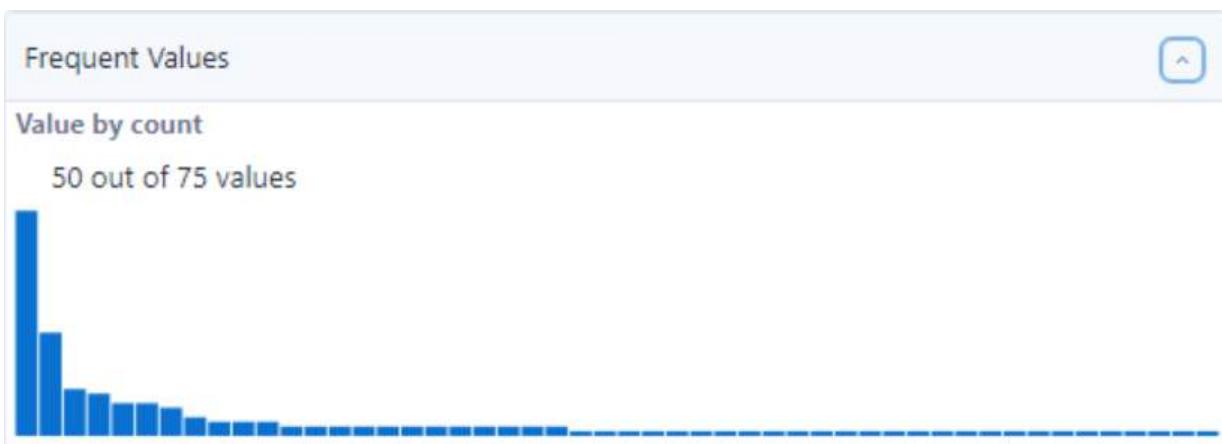
Type	Records	Data Type Size
V_WString	210	1,073,741,823

● Ok	210	100.00%
Unique	21	10.00%
■ Null	0	0.00%
Not Ok	0	0.00%
● Empty	0	0.00%

Length Statistics	
Min	3
Max	31
Average	12.40
Shortest Value	OKs
Longest Value	Values with Trailing Whitespace
First Alphanumeric Value	Average Length
Last Alphanumeric Value	Values with Trailing Whitespace
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0



Length Statistics	
Min	1
Max	280
Average	10.90
Shortest Value	1
Longest Value	tt0371746,tt4154796,tt1300854,tt0988045 tt0848228,tt080...
First Alphanumeric Value	2023-11-19 15:20:37
Last Alphanumeric Value	tt2395427,tt0458339,tt3498820,tt0848228
Blanks	0
Values with Leading Whitespace	3
Values with Trailing Whitespace	10



2. Newdata_title_basics.json:

newdata_name_basics_profiling.ymd x + ⌂ Run

Alteryx Designer x64

Finished running newdata_name_basics_profiling.ymd in 0.5 seconds using AMP engine.

[Learn about AMP Engine](#)

Don't show this message again

OK

Results - Workflow - Messages

All ① 0 Errors ④ 0 Conv Errors ③ 0 Warnings ② 3 Info ⑤ 5 Files Last Run Configuration

- ① Browse (42) 45 records
- ② Browse (32) 7 records
- ③ Browse (50) 210 records
- ④ Designer x64 Finished running newdata_name_basics_profiling.ymd in 0.5 seconds using AMP engine.

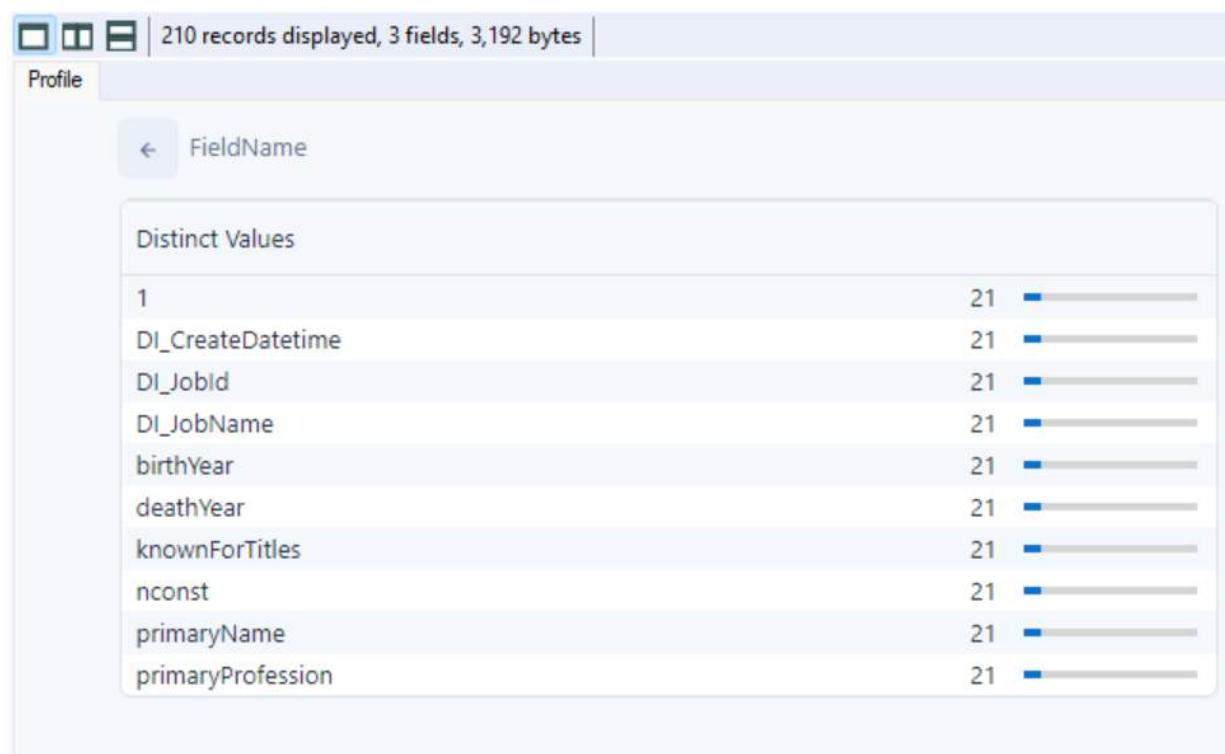
newdata_name_basics_profiling.yxml × + Run

Results - Browse (50) - Input

3 of 3 Fields 210 records displayed, 3,192 bytes

X ✓ Search Metadata Actions 000

Record	FieldName	Name	Value
22	nconst	Name	nconst
23	nconst	Data Type	V_WString
24	nconst	Size	2048
25	nconst	Source	CrossTab:Header:Column_Name:nconst:Concat
26	nconst	Description	[Null]
27	nconst	OKs	7
28	nconst	Nulls	0
29	nconst	Non-Nulls	7
30	nconst	Blanks	0
31	nconst	Values with Leading Whitespace	0
32	nconst	Values with Trailing Whitespace	0
33	nconst	Values with Both Whitespace	0
34	nconst	Average Length	9.0
35	nconst	Longest Length	9
36	nconst	Longest Value	nm0749263



FieldName x

Summary

Type	Records	Data Type Size
V_WString	210	1,073,741,823

• Ok	210	100.00%
Unique	10	4.76%
• Null	0	0.00%
Not	0	0.00%
• Ok	0	0.00%
• Empty	0	0.00%

Length Statistics

Min	1
Max	17
Average	10.20
Shortest Value	1
Longest Value	primaryProfession
First Alphanumeric Value	1
Last Alphanumeric Value	primaryProfession
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

Frequent Values

Value by count

10 out of 10 values



Profile	
Distinct Values	
Average Length	10
Blanks	10
Data Type	10
Description	10
Longest Length	10
Longest Value	10
Maximum	10
Minimum	10
Name	10
Non-Nulls	10
Nulls	10
OKs	10
Shortest (Non Blank) Length	10
Shortest Value	10
Size	10
Source	10

Summary			
Type	Records	Data Type	Size
V_WString	210	1,073,741,823	
● Ok	210		100.00%
Unique	21		10.00%
● Null	0		0.00%
Not Ok	0		0.00%
● Empty	0		0.00%

Length Statistics	
Min	3
Max	31
Average	12.40
Shortest Value	OKs
Longest Value	Values with Trailing Whitespace
First Alphanumeric Value	Average Length
Last Alphanumeric Value	Values with Trailing Whitespace
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

Profile

Value

Summary

Type	Records	Data Type	Size
V_WString	210	1,073,741,823	

Ok	188	89.52%
Unique	75	35.71%
Null	12	5.71%
Not Ok	10	4.76%
Empty	0	0.00%

Length Statistics

Min	1
Max	280
Average	10.90
Shortest Value	1
Longest Value	tt0371746,tt4154796,tt1300854,tt0988045 tt0848228,tt0800369,...
First Alphanumeric Value	2023-11-19 15:20:37
Last Alphanumeric Value	tt2395427,tt0458339,tt3498820,tt0848228
Blanks	0
Values with Leading Whitespace	3
Values with Trailing Whitespace	10

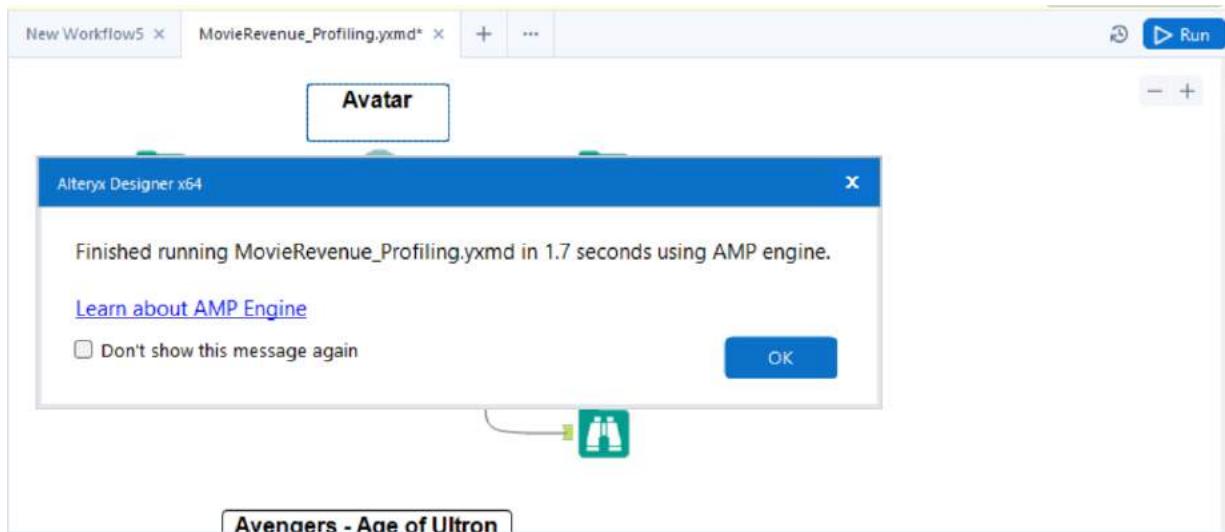
Frequent Values

Value by count

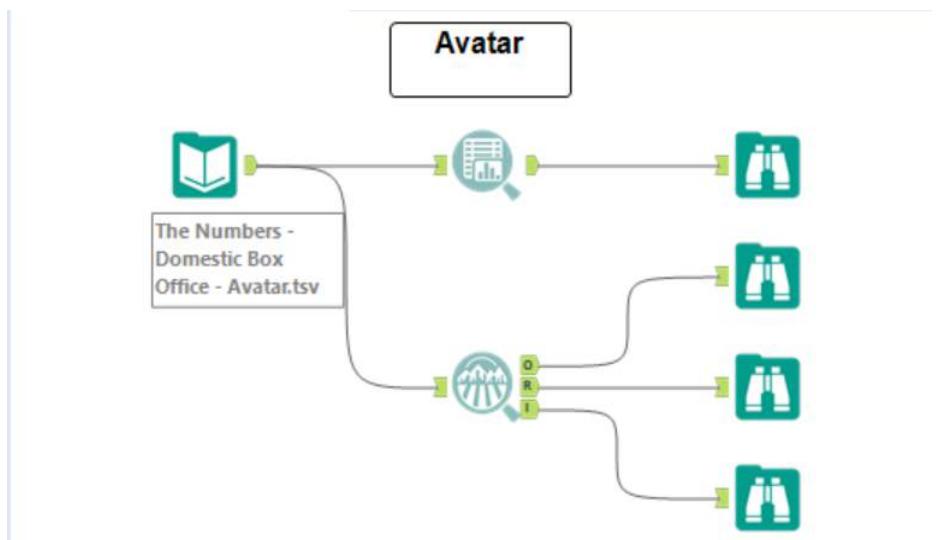
50 out of 75 values



3. Movie Revenue Profiling:



3.1 Avatar



Results - Browse (3) - Input

3 of 3 Fields 231 records displayed, 4,789 bytes X ✓

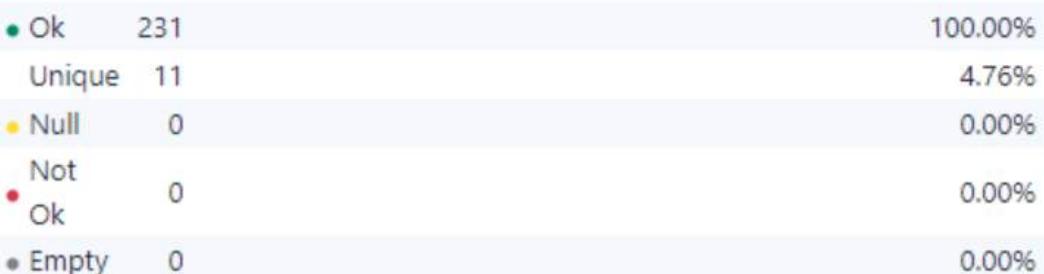
Record	FieldName	Name	Value	⋮
1	tconst	Name	tconst	
2	tconst	Data Type	V_String	
3	tconst	Size	254	
4	tconst	Source	File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...	
5	tconst	Description	[Null]	
6	tconst	OKs	318	
7	tconst	Nulls	0	
8	tconst	Non-Nulls	318	
9	tconst	Blanks	0	
10	tconst	Values with Leading Whitespace	0	
11	tconst	Values with Trailing Whitespace	0	
12	tconst	Values with Both Whitespace	0	
13	tconst	Average Length	9.0	
14	tconst	Longest Length	9	
15	tconst	Longest Value	tt0499549	
16	tconst	Shortest (Non-Blank) Length	0	

Profile

.FieldName

Summary

Type	Records	Data Type Size
V_WString	231	1,073,741,823



Profile

Name x

Summary ^

Type	Records	Data Type Size
V_WString	231	1,073,741,823

● Ok	231	100.00%
Unique	21	9.09%
■ Null	0	0.00%
Not	0	0.00%
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Value x

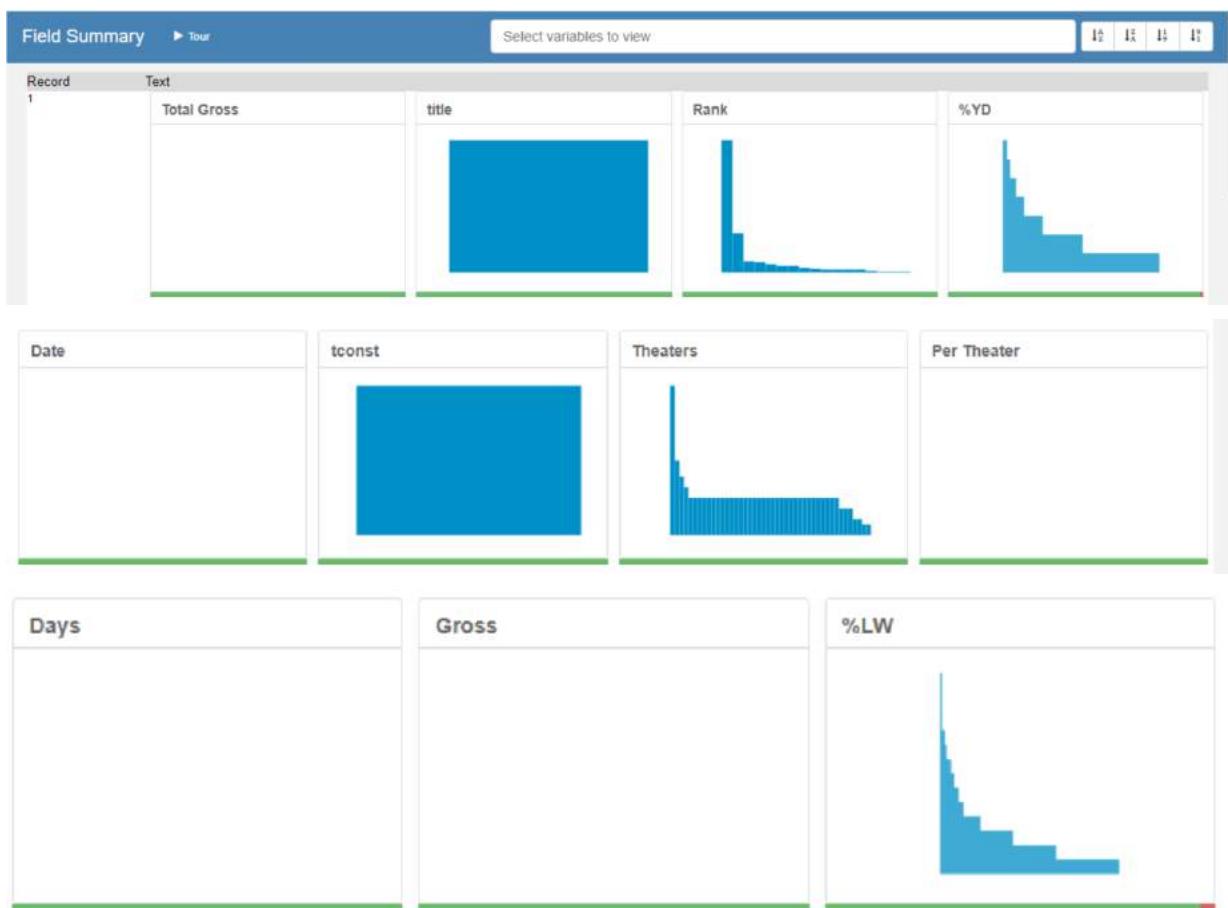
Summary ^

Type	Records	Data Type Size
V_WString	231	1,073,741,823

● Ok	209	90.48%
Unique	81	35.06%
■ Null	15	6.49%
Not	7	3.03%
● Ok	7	3.03%
● Empty	0	0.00%

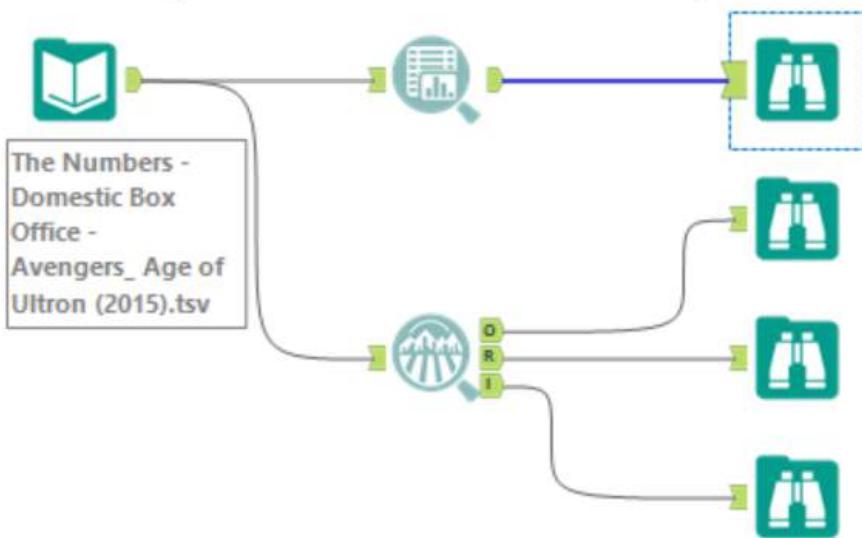
Report Profile | 1 of 1 Fields | Records 1 to 1 | String/Character Fields

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	318	\$26,752,099	\$109,497,762	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	17	1	10	1	176	176 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	0.6%	155	2	-34%	1	7	7 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	318	Jan 1, 2010	Dec 18, 2009	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	44	9	3,452	2	28	28 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	0.0%	293	\$60	\$7,750	1	3	3 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	0.0%	318	1	100	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining
Gross	0.0%	317	\$571	\$26,752,099	1	2	2 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	4.4%	116	-2%	-14%	1	14	14 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	Avatar	Avatar	318	318	
tconst	0.0%	1	tt0499549	tt0499549	318	318	



3.2 Avengers - Age of Ultron

Avengers - Age of Ultron



Record	FieldName	Name	Value
1	tconst	Name	tconst
2	tconst	Data Type	V_String
3	tconst	Size	254
4	tconst	Source	File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...
5	tconst	Description	[Null]
6	tconst	OKs	79
7	tconst	Nulls	0
8	tconst	Non-Nulls	79
9	tconst	Blanks	0
10	tconst	Values with Leading Whitespace	0
11	tconst	Values with Trailing Whitespace	0
12	tconst	Values with Both Whitespace	0
13	tconst	Average Length	9.0
14	tconst	Longest Length	9
15	tconst	Longest Value	tt2395427

Profile

.FieldName *

Summary ^

Type	Records	Data Type Size
V_WString	231	1,073,741,823

● Ok	231	100.00%
Unique	11	4.76%
● Null	0	0.00%
Not	0	0.00%
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Name *

Summary ^

Type	Records	Data Type Size
V_WString	231	1,073,741,823

● Ok	231	100.00%
Unique	21	9.09%
● Null	0	0.00%
Not	0	0.00%
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Value

Summary

Type Records Data Type Size
V_WString 231 1,073,741,823

● Ok	209	90.48%
Unique	80	34.63%
● Null	11	4.76%
Not Ok	11	4.76%
● Empty	0	0.00%

Report Profile

1 of 1 Fields ▾ ✓ | Records 1 to 1 |

String/Character Fields

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	79	\$27,600,000	\$140,951,167	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	15	P	10	1	1	14 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	2.5%	55	-1%	-33%	1	1	5 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	79	May 1, 2015	Apr 30, 2015	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	13	0	4,276	1	1	21 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	1.3%	78	\$944	\$19,744	1	1	2 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	1.3%	79	1	10	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Gross	0.0%	79	\$69,041	\$27,600,000	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	10.1%	39	-1%	-75%	1	1	8 This field has over 10% missing values. Consider imputing these values. Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	Avengers: Age of Ultron	Avengers: Age of Ultron	79	79	
tconst	0.0%	1	tt2395427	tt2395427	79	79	

Field Summary

► Tour:

Select variables to view

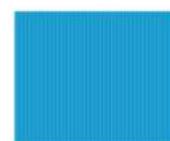
1 1 1 1 1

Record

Text

1

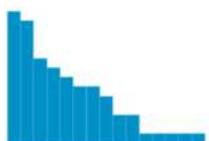
Total Gross



title



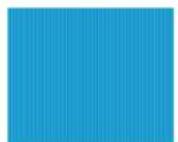
Rank



%YD



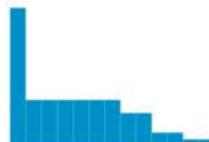
Date



tconst



Theaters

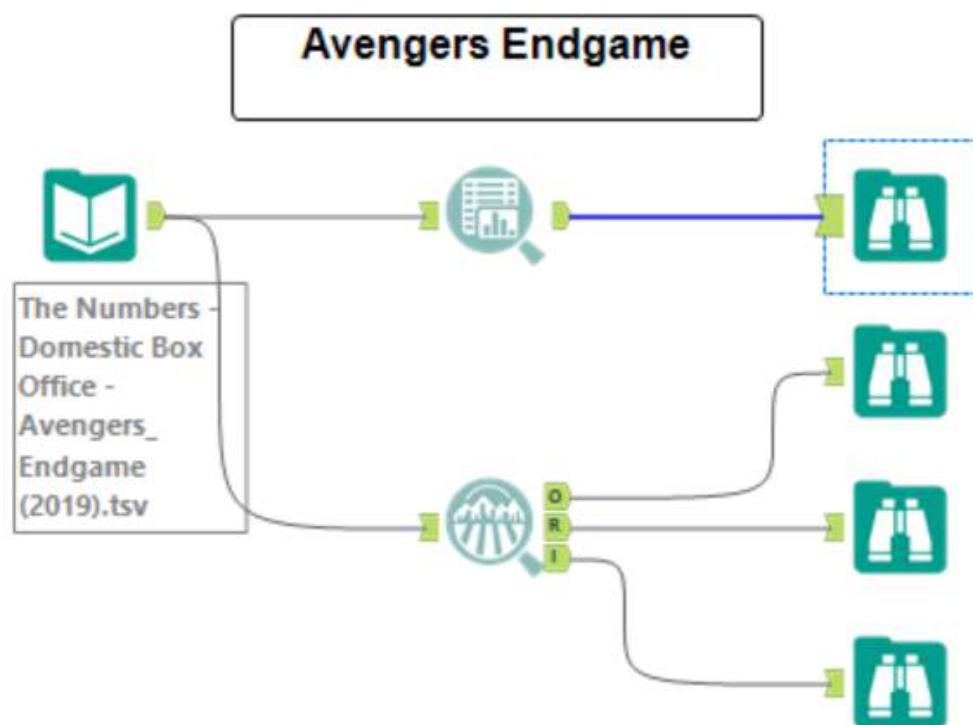


Per Theater





3.3 Avengers Endgame



3 of 3 Fields				<input checked="" type="checkbox"/>	<input type="checkbox"/>	231 records displayed, 4,366 bytes	<input type="button" value="X"/>	<input type="button" value="✓"/>	<input type="button" value="Q"/>	Search
Record	FieldName	Name	Value							
1	tconst	Name	tconst							
2	tconst	Data Type	V_String							
3	tconst	Size	254							
4	tconst	Source	File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...							
5	tconst	Description	[Null]							
6	tconst	OKs	141							
7	tconst	Nulls	0							
8	tconst	Non-Nulls	141							
9	tconst	Blanks	0							
10	tconst	Values with Leading Whitespace	0							
11	tconst	Values with Trailing Whitespace	0							
12	tconst	Values with Both Whitespace	0							
13	tconst	Average Length	9.0							
14	tconst	Longest Length	9							
15	tconst	Longest Value	tt4154796							
16	tconst	Shortest (Non Blank) Length	9							

Profile

Field Name

Summary

Type	Records	Data Type	Size
V_WString	231	1,073,741,823	

● Ok	231	100.00%
Unique	11	4.76%
■ Null	0	0.00%
Not		
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Name

Summary

Type	Records	Data Type	Size
V_WString	231	1,073,741,823	

● Ok	231	100.00%
Unique	21	9.09%
■ Null	0	0.00%
Not		
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Value

Summary

Type	Records	Data Type	Size
V_WString	231		1,073,741,823

● Ok	209	90.48%
Unique	81	35.06%
● Null	14	6.06%
Not Ok	8	3.46%
● Empty	0	0.00%

Report Profile

1 of 1 Fields ▾ ✓ | Records 1 to 1 |

String/Character Fields

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	141	\$60,000,000	\$157,461,641	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	14	P	12	1	35	35 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	1.4%	93	0.7	-31%	1	1	6 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	141	May 1, 2019	Apr 25, 2019	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	18	0	4,662	1	1	21 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	0.7%	131	\$91	\$33,776	1	1	3 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	0.0%	141	0	100	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Gross	0.0%	141	\$940	\$157,461,641	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	5.7%	62	-5%	-74%	1	1	8 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	Avengers: Endgame	Avengers: Endgame	141	141	
tconst	0.0%	1	tt4154796	tt4154796	141	141	

Field Summary

▶ Your

Select variables to view

1 1 1 1

Record Text

1

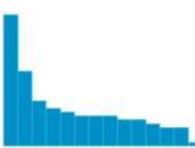
Total Gross



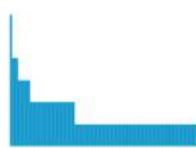
title



Rank



%YD



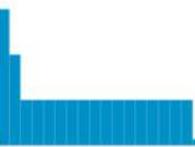
Date



tconst

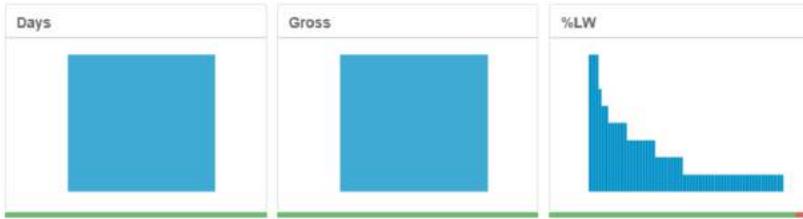


Theaters

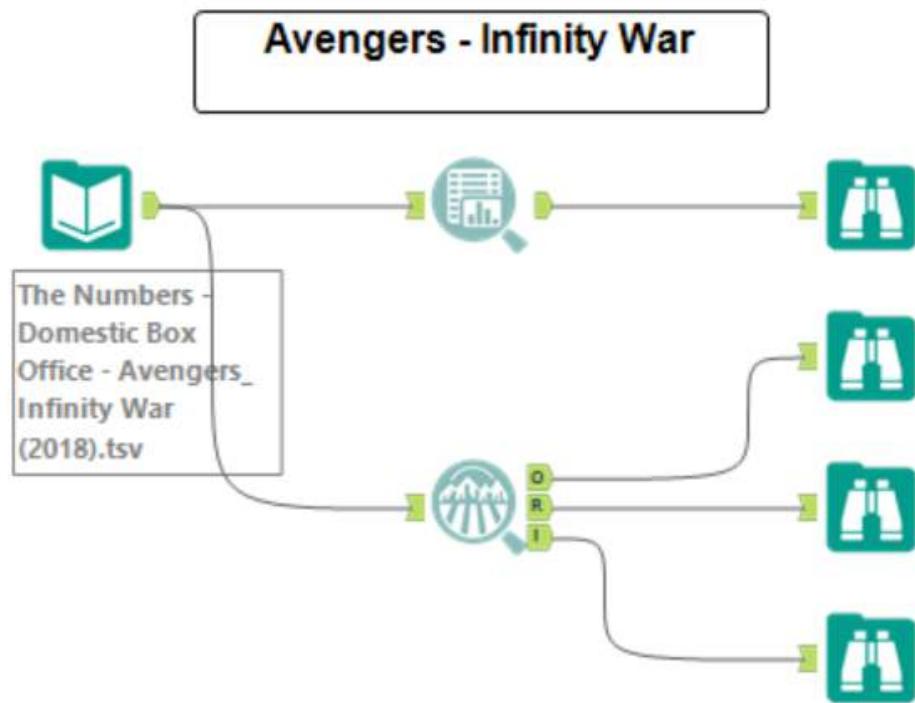


Per Theater





3.4 Avengers - Infinity War



3 of 3 Fields Cell Viewer 231 records displayed, 4,360 bytes Search Data

Record	FieldName	Name	Value
1	tconst	Name	tconst
2	tconst	Data Type	V_String
3	tconst	Size	254
4	tconst	Source	File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...
5	tconst	Description	[Null]
6	tconst	OKs	141
7	tconst	Nulls	0
8	tconst	Non-Nulls	141
9	tconst	Blanks	0
10	tconst	Values with Leading Whitespace	0
11	tconst	Values with Trailing Whitespace	0
12	tconst	Values with Both Whitespace	0
13	tconst	Average Length	9.0
14	tconst	Longest Length	9
15	tconst	Longest Value	tt4154756
16	tconst	Shortest (Non Blank) Length	9

Profile

.FieldName

Summary

Type	Records	Data Type	Size
V_WString	231	1,073,741,823	

● Ok	231	100.00%
Unique	11	4.76%
■ Null	0	0.00%
Not	0	0.00%
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Name

Summary

Type	Records	Data Type	Size
V_WString	231	1,073,741,823	

● Ok	231	100.00%
Unique	21	9.09%
■ Null	0	0.00%
Not	0	0.00%
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Value

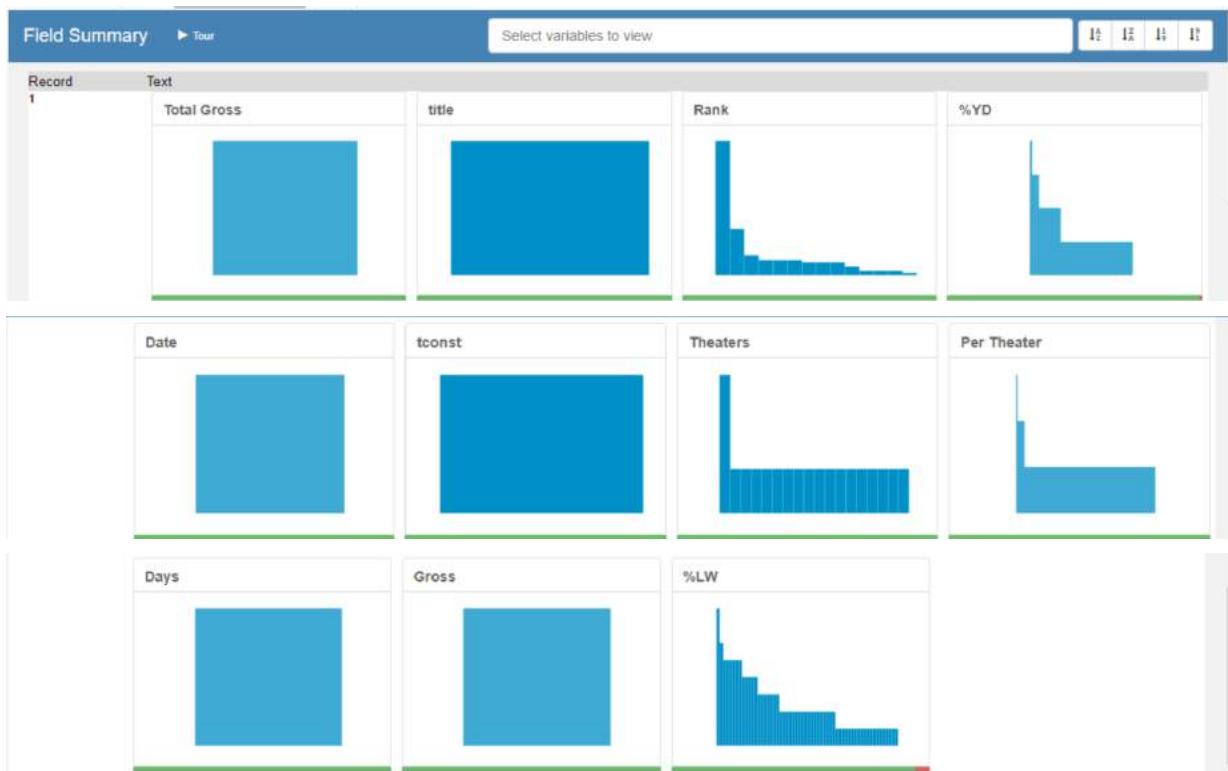
Summary

Type	Records	Data Type	Size
V_WString	231		1,073,741,823

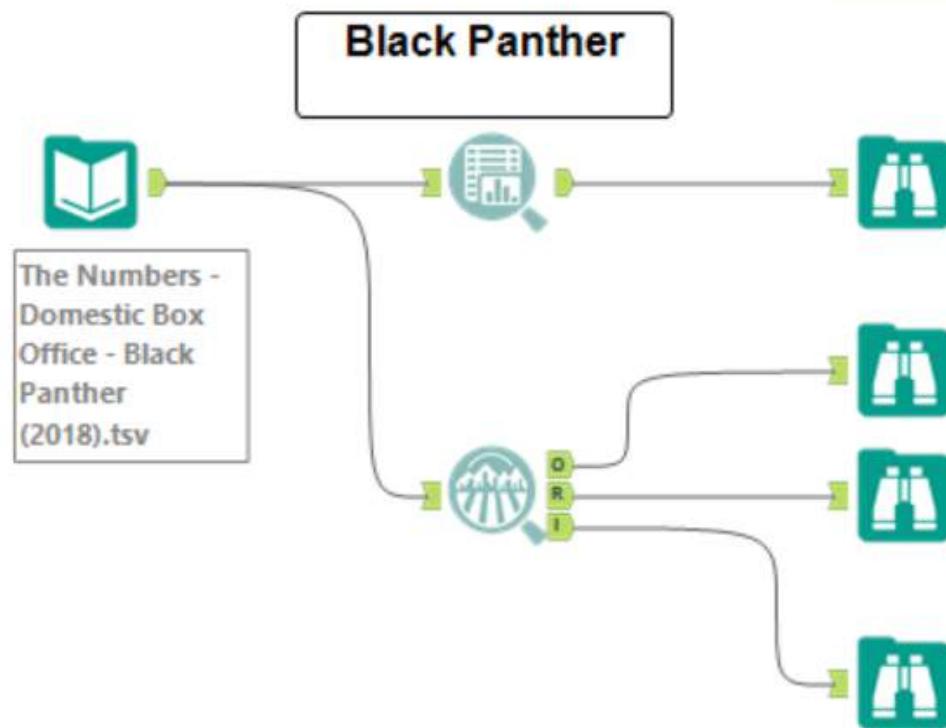
Ok: 209 (90.48%)
 Unique: 81 (35.06%)
 Null: 14 (6.06%)
 Not Ok: 8 (3.46%)
 Empty: 0 (0.00%)

Report Profile | 1 of 1 Fields | Records 1 to 1 | String/Character Fields

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	141	\$39,000,000	\$106,334,939	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	14	P	12	1	1	61 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	1.4%	100	1	-23%	1	1	4 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	141	May 1, 2018	Apr 26, 2018	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	18	92	4,474	7	7	22 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	0.0%	132	\$68	\$23,767	1	1	3 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	0.0%	141	0	100	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Gross	0.0%	141	\$9,993	\$106,334,939	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	5.7%	58	-6%	-70%	1	1	8 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	Avengers: Infinity War	Avengers: Infinity War	141	141	
tconst	0.0%	1	tt4154756	tt4154756	141	141	



3.5 Black Panther



3 of 3 Fields				X	✓	Search	Data
Record	FieldName	Name	Value				
1	tconst	Name	tconst				
2	tconst	Data Type	V_String				
3	tconst	Size	254				
4	tconst	Source	File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...				
5	tconst	Description	[Null]				
6	tconst	OKs	176				
7	tconst	Nulls	0				
8	tconst	Non-Nulls	176				
9	tconst	Blanks	0				
10	tconst	Values with Leading Whitespace	0				
11	tconst	Values with Trailing Whitespace	0				
12	tconst	Values with Both Whitespace	0				
13	tconst	Average Length	9.0				
14	tconst	Longest Length	9				
15	tconst	Longest Value	tt1825683				
16	tconst	Shortest (Non Blank) Length	9				

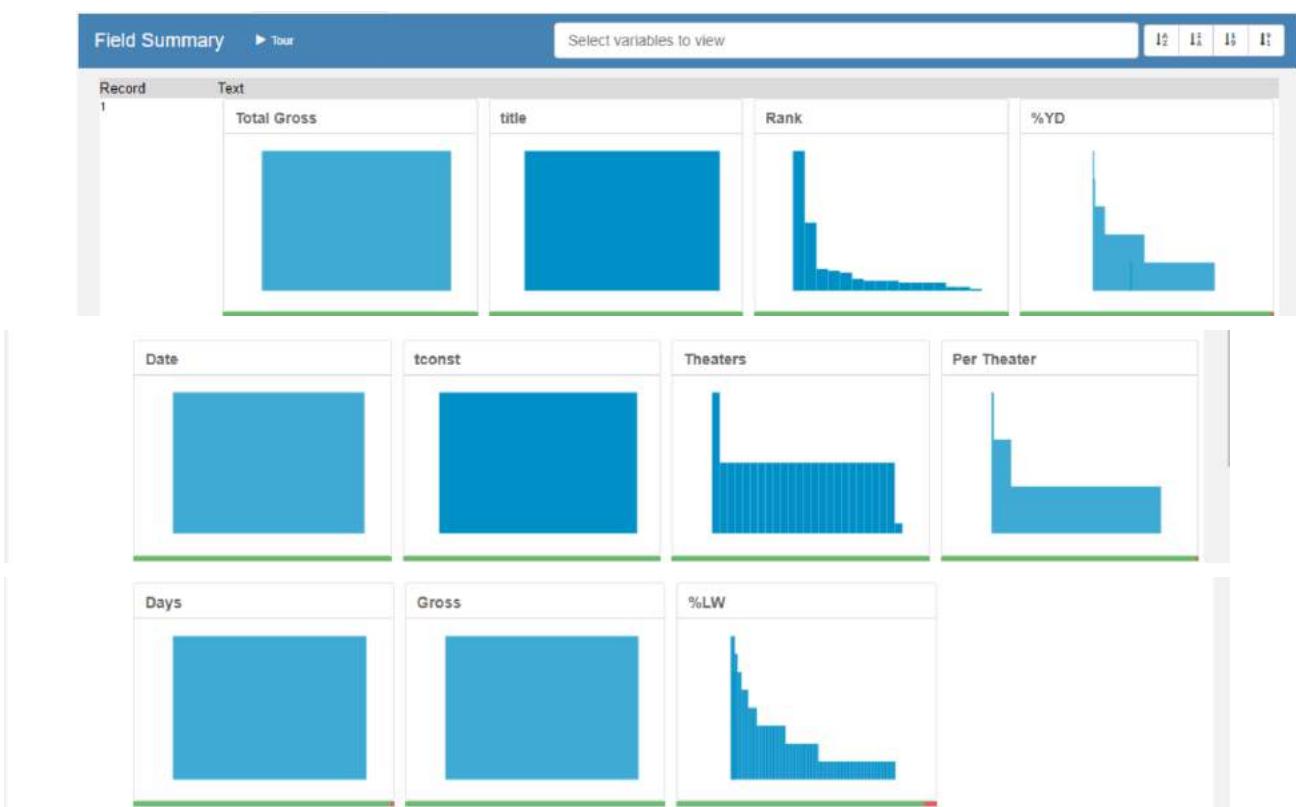
Summary			
Type	Records	Data Type Size	
V_WString	231	1,073,741,823	
● Ok	231	100.00%	
Unique	11	4.76%	
■ Null	0	0.00%	
Not	0	0.00%	
● Ok	0	0.00%	
● Empty	0	0.00%	

Summary			
Type	Records	Data Type Size	
V_WString	231	1,073,741,823	
● Ok	231	100.00%	
Unique	21	9.09%	
■ Null	0	0.00%	
Not	0	0.00%	
● Ok	0	0.00%	
● Empty	0	0.00%	

Summary			
Type	Records	Data Type Size	
V_WString	231	1,073,741,823	
● Ok	207	89.61%	
Unique	85	36.80%	
■ Null	16	6.93%	
Not	8	3.46%	
● Ok	0	0.00%	
● Empty	0	0.00%	

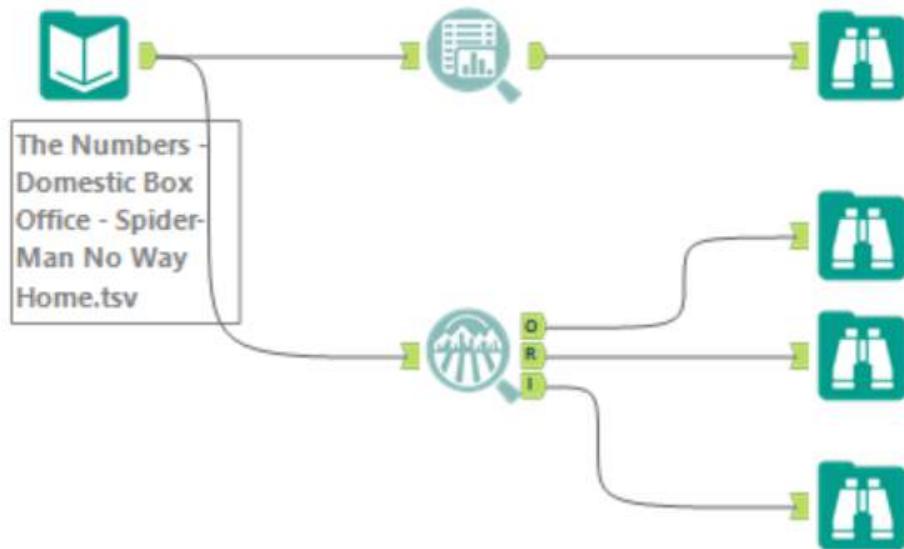
String/Character Fields							
Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	176	\$25,200,000	\$141,936,512	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	16	P	10	1	70	2 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	1.1%	114	-13%	-13%	1	5	3 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	176	Mar 1, 2018	Feb 15, 2018	1	1	4 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	25	0	4,020	1	14	5 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	0.6%	156	\$79	\$18,891	1	3	6 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	0.6%	176	1	100	1	1	7 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.

Gross	0.0%	176	\$780	\$25,200,000	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	4.5%	76	-62%		1	8 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	Black Panther	Black Panther	176	176
tconst	0.0%	1	tt1825683	tt1825683	176	176



3.6 Spider Man - No Way Home

Spider - Man No Way Home



3 of 3 Fields			
Record	FieldName	Name	Value
1	tcont	Name	tcont
2	tcont	Data Type	V_String
3	tcont	Size	254
4	tcont	Source	File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...
5	tcont	Description	[Null]
6	tcont	OKs	128
7	tcont	Nulls	0
8	tcont	Non-Nulls	128
9	tcont	Blanks	0
10	tcont	Values with Leading Whitespace	0
11	tcont	Values with Trailing Whitespace	0
12	tcont	Values with Both Whitespace	0
13	tcont	Average Length	10.0
14	tcont	Longest Length	10
15	tcont	Longest Value	tt10872600
16	tcont	Shortest (Non Blank) Length	10

Profile

FieldName

Summary			
Type	Records	Data Type	Size
V_WString	231	1,073,741,823	
• Ok	231		100.00%
Unique	11		4.76%
• Null	0		0.00%
Not	0		0.00%
• Ok	0		0.00%
• Empty	0		0.00%

Profile

Name

Summary			
Type	Records	Data Type	Size
V_WString	231	1,073,741,823	
• Ok	231		100.00%
Unique	21		9.09%
• Null	0		0.00%
Not	0		0.00%
• Ok	0		0.00%
• Empty	0		0.00%

Profile

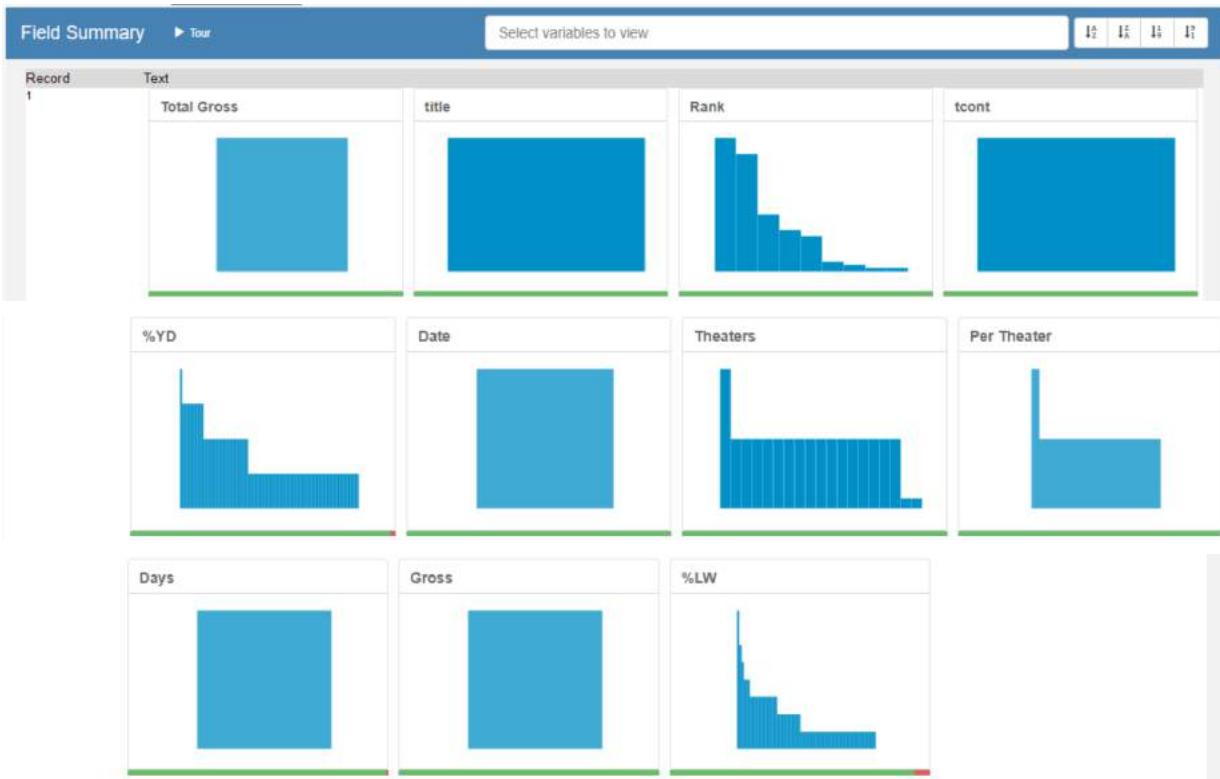
Value

Summary			
Type	Records	Data Type	Size
V_WString	231	1,073,741,823	
• Ok	209		90.48%
Unique	83		35.93%
• Null	14		6.06%
Not	8		3.46%
• Ok	0		0.00%
• Empty	0		0.00%

Report Profile | 1 of 1 Fields | Records 1 to 1 | ▶◀ ◀▶ ▶▶

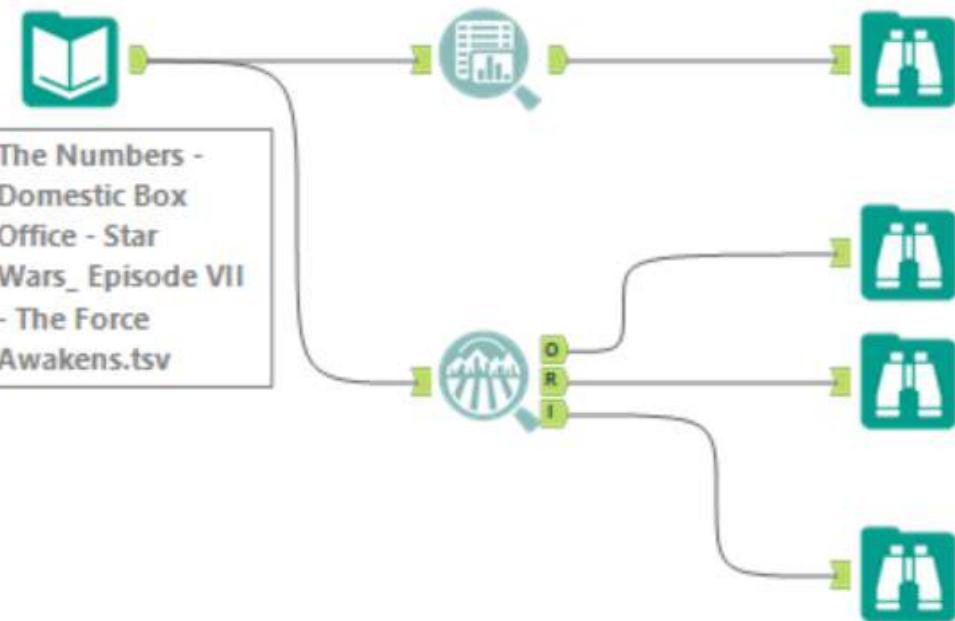
String/Character Fields

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	128	\$50,000,000	\$121,964,712	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	9	P	P	1	1	42 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	1.6%	84	1	-39%	1	1	4 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	128	Jan 1, 2022	Dec 16, 2021	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	19	387	3,767	1	1	14 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	0.0%	121	\$95	\$13,273	1	1	2 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	0.8%	128	1	100	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Gross	0.0%	128	\$98,951	\$121,964,712	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	6.2%	66	-6%	-84%	1	1	8 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	Spider-Man: No Way Home	Spider-Man: No Way Home	128	128	
tcont	0.0%	1	tt10872600	tt10872600	128	128	



3.7 Star Wars Episode VII - The Force Awakens

Star Wars Episode VII - The Force Awakens



3 of 3 Fields		<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	231 records displayed, 4,237 bytes	X	✓	Q Search	Data
Record	FieldName	Name		Value					
1	tconst	Name		tconst					
2	tconst	Data Type		V_String					
3	tconst	Size		254					
4	tconst	Source		File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...					
5	tconst	Description		[Null]					
6	tconst	OKs		120					
7	tconst	Nulls		0					
8	tconst	Non-Nulls		120					
9	tconst	Blanks		0					
10	tconst	Values with Leading Whitespace		0					
11	tconst	Values with Trailing Whitespace		0					
12	tconst	Values with Both Whitespace		0					
13	tconst	Average Length		9.0					
14	tconst	Longest Length		9					
15	tconst	Longest Value		tt2488496					
16	tconst	Shortest (Non Blank) Length		9					

Profile

.FieldName x

Summary ^

Type	Records	Data Type Size
V_WString	231	1,073,741,823

● Ok	231	100.00%
Unique	11	4.76%
■ Null	0	0.00%
Not		
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Name x

Summary ^

Type	Records	Data Type Size
V_WString	231	1,073,741,823

● Ok	231	100.00%
Unique	21	9.09%
■ Null	0	0.00%
Not		
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Value

Summary

Type Records Data Type Size
V_WString 231 1,073,741,823

● Ok	209	90.48%
Unique	81	35.06%
■ Null	14	6.06%
Not Ok	8	3.46%
● Empty	0	0.00%

1 of 1 Fields ▾

Records 1 to 1

String/Character Fields

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	120	\$57,000,000	\$119,119,282	1	25	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	17	P	11	1	25	2 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	1.7%	89	-7%	-43%	1	3	3 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	120	Jan 1, 2016	Dec 17, 2015	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	16	0	4,134	1	28	2 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	0.8%	113	\$72	\$28,615	1	2	2 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	0.0%	120	0	100	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.

Gross	0.0%	120	\$56,853	\$119,119,282	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	6.7%	62	-5%	-59%	1	8 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	Star Wars: Episode VII - The Force Awakens	Star Wars: Episode VII - The Force Awakens	120	120
tconst	0.0%	1	tt2488496	tt2488496	120	120

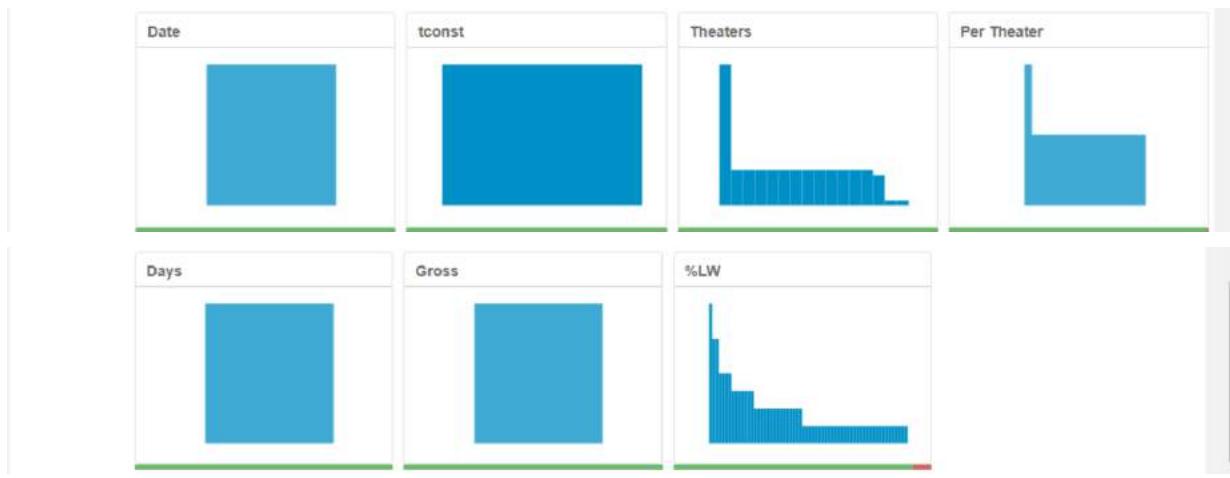
Field Summary

Select variables to view

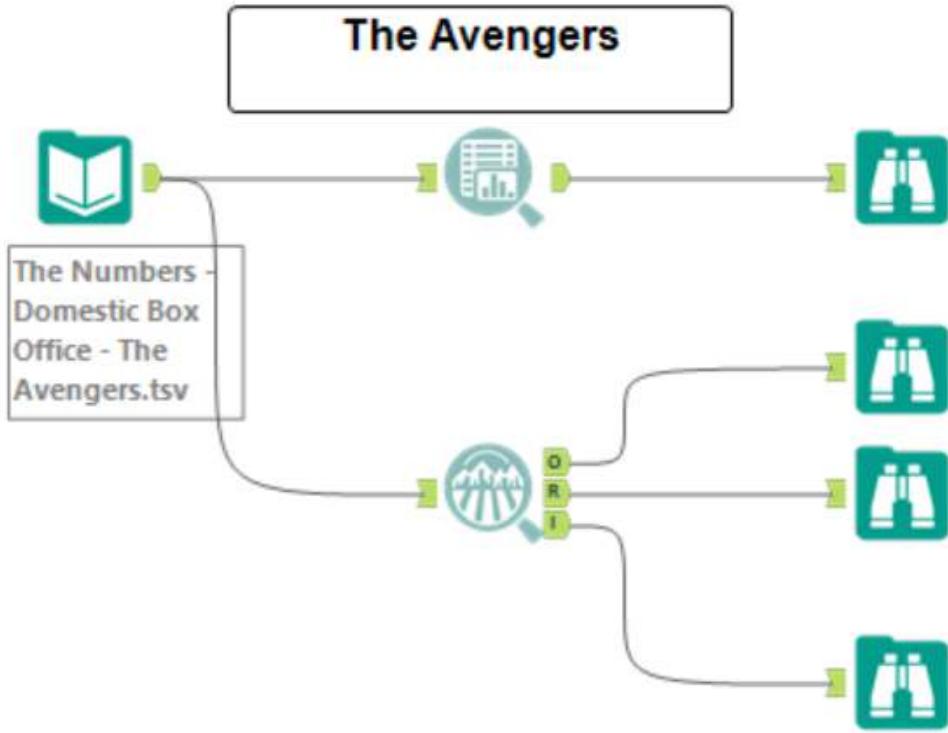
1A 1B 1C 1D

Record Text





3.8 The Avengers



3 of 3 Fields 231 records displayed, 4,812 bytes ✖️ ✓

Record	FieldName	Name	Value
1	tconst	Name	tconst
2	tconst	Data Type	V_String
3	tconst	Size	254
4	tconst	Source	File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...
5	tconst	Description	[Null]
6	tconst	OKs	97
7	tconst	Nulls	0
8	tconst	Non-Nulls	97
9	tconst	Blanks	0
10	tconst	Values with Leading Whitespace	0
11	tconst	Values with Trailing Whitespace	0
12	tconst	Values with Both Whitespace	0
13	tconst	Average Length	9.0
14	tconst	Longest Length	9
15	tconst	Longest Value	tt0848228
16	tconst	Shortest (Non Blank) Length	9

Profile

.FieldName ✖️

Summary			
Type	Records	Data Type	Size
V_WString	231	1,073,741,823	

Type	Count	Percentage
Ok	231	100.00%
Unique	11	4.76%
Null	0	0.00%
Not Ok	0	0.00%
Empty	0	0.00%

Profile

Name x

Summary

Type	Records	Data Type Size
V_WString	231	1,073,741,823

● Ok	231	100.00%
Unique	21	9.09%
■ Null	0	0.00%
Not	0	0.00%
● Ok	0	0.00%
● Empty	0	0.00%

Profile

Value x

Summary

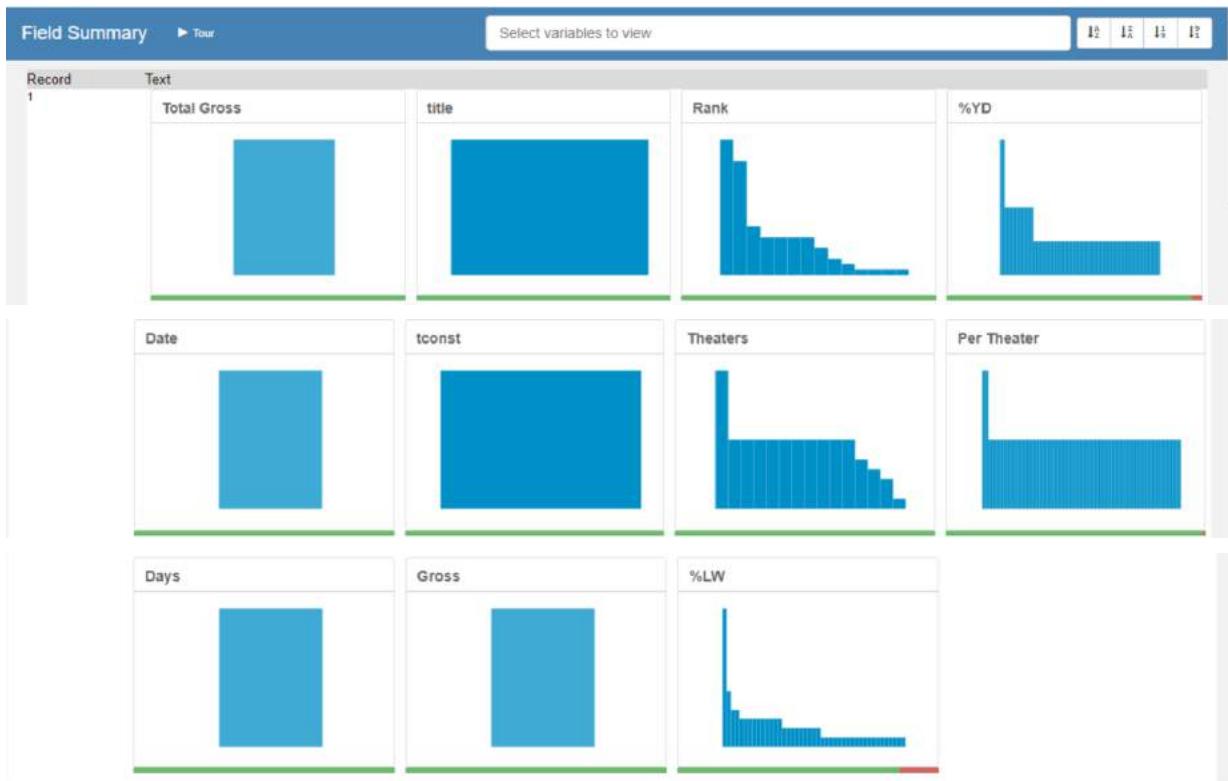
Type	Records	Data Type Size
V_WString	231	1,073,741,823

● Ok	209	90.48%
Unique	77	33.33%
■ Null	13	5.63%
Not	9	3.90%
● Ok	0	0.00%
● Empty	0	0.00%

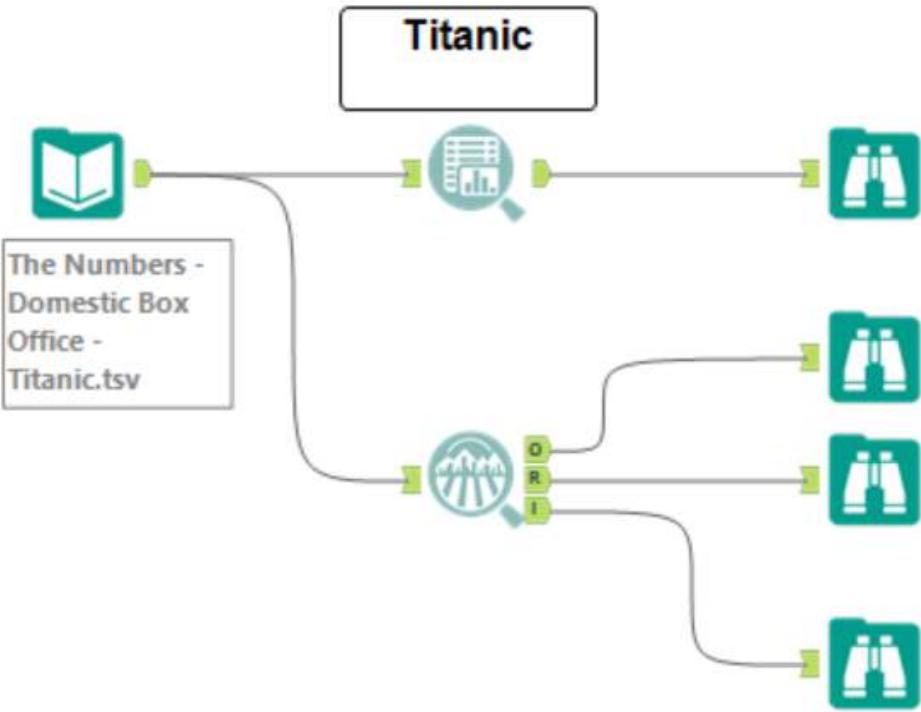
Report Profile
1 of 1 Fields | Records 1 to 1 |

String/Character Fields

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	97	\$18,700,000	\$150,371,975	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	14	P	11	1	25	25 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	4.1%	77	-6%	-14%	1	4	4 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	97	May 3, 2012	May 10, 2012	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	15	0	4,349	1	14	14 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	1.0%	94	\$56	\$18,582	1	2	2 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	0.0%	97	0	120	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Gross	0.0%	97	\$98,364	\$18,700,000	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	15.5%	43	n/c	-64%	1	15	15 This field has over 10% missing values. Consider imputing these values. Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	The Avengers	The Avengers	97	97	
tconst	0.0%	1	tt0848228	tt0848228	97	97	



3.9 Titanic



3 of 3 Fields | 231 records displayed, 4,570 bytes X ✓

Record	FieldName	Name	Value
1	tconst	Name	tconst
2	tconst	Data Type	V_String
3	tconst	Size	254
4	tconst	Source	File: D:\dadabi\Final Project\IMDB\IMDB\TheNu...
5	tconst	Description	[Null]
6	tconst	OKs	265
7	tconst	Nulls	0
8	tconst	Non-Nulls	265
9	tconst	Blanks	0
10	tconst	Values with Leading Whitespace	0
11	tconst	Values with Trailing Whitespace	0
12	tconst	Values with Both Whitespace	0
13	tconst	Average Length	9.0
14	tconst	Longest Length	9
15	tconst	Longest Value	tt0120338
16	tconst	Shortest (Non Blank) Length	9

Profile x

FieldName

Summary			
Type	Records	Data Type	Size
V_WString	231	1,073,741,823	
<ul style="list-style-type: none"> ● Ok 231 100.00% Unique 11 4.76% ■ Null 0 0.00% Not 0 0.00% ● Ok 0 0.00% ● Empty 0 0.00% 			

Profile x

Name

Summary			
Type	Records	Data Type	Size
V_WString	231	1,073,741,823	
<ul style="list-style-type: none"> ● Ok 231 100.00% Unique 21 9.09% ■ Null 0 0.00% Not 0 0.00% ● Ok 0 0.00% ● Empty 0 0.00% 			

Profile x

Value

Summary			
Type	Records	Data Type	Size
V_WString	231	1,073,741,823	
<ul style="list-style-type: none"> ● Ok 209 90.48% Unique 84 36.36% ■ Null 15 6.49% Not 7 3.03% ● Ok 0 0.00% ● Empty 0 0.00% 			

String/Character Fields							
Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	265	\$8,658,814	\$104,380,438	1	1	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	14	1	11	1	97	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%YD	1.5%	143	-7%	0.23	1	7	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Date	0.0%	265	Jan 1, 1998	Dec 19, 1997	1	1	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Theaters	0.0%	36	70	2,674	4	16	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Per Theater	0.0%	228	\$88	\$3,238	1	3	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Days	0.0%	265	1	5,221	1	1	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Gross	0.0%	264	\$8,128	\$10,672,013	1	2	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
%LW	9.4%	102	0.4	0.17	1	25	Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
title	0.0%	1	Titanic	Titanic	265	265	
tconst	0.0%	1	tt0120338	tt0120338	265	265	

Field Summary ▶ Tour Select variables to view

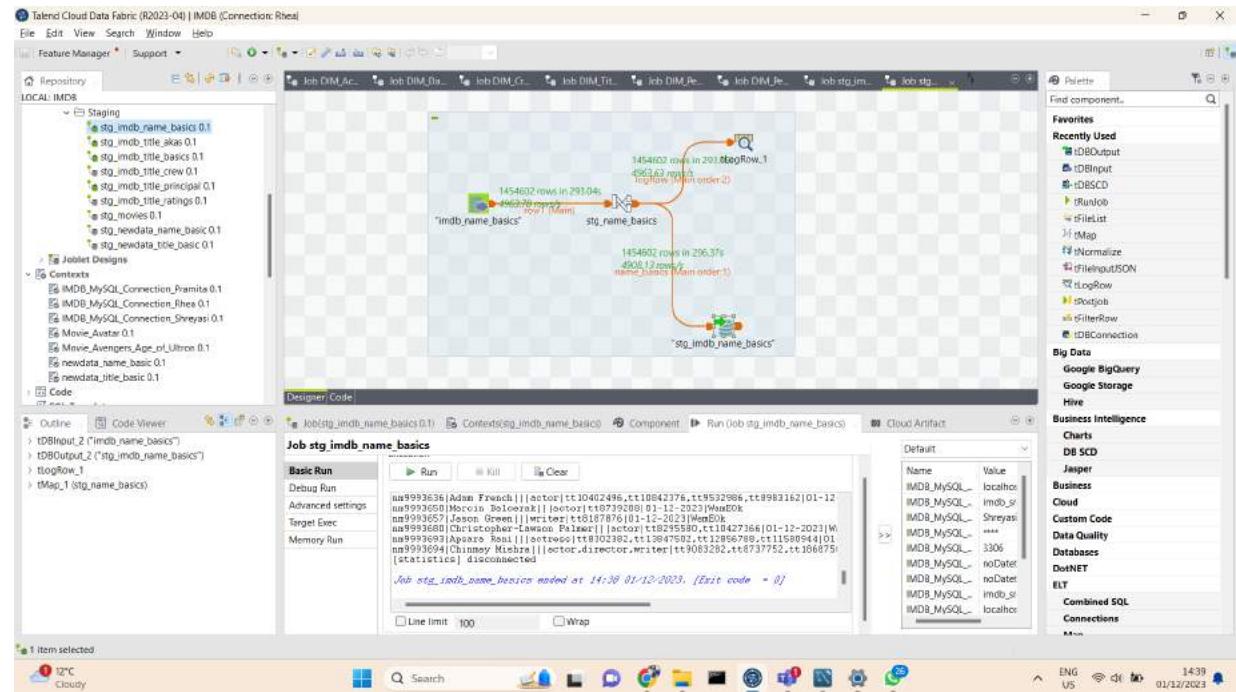
Record 1

Text	title	Rank	%YD
Total Gross			
Date	tconst	Theaters	Per Theater
Days	Gross	%LW	

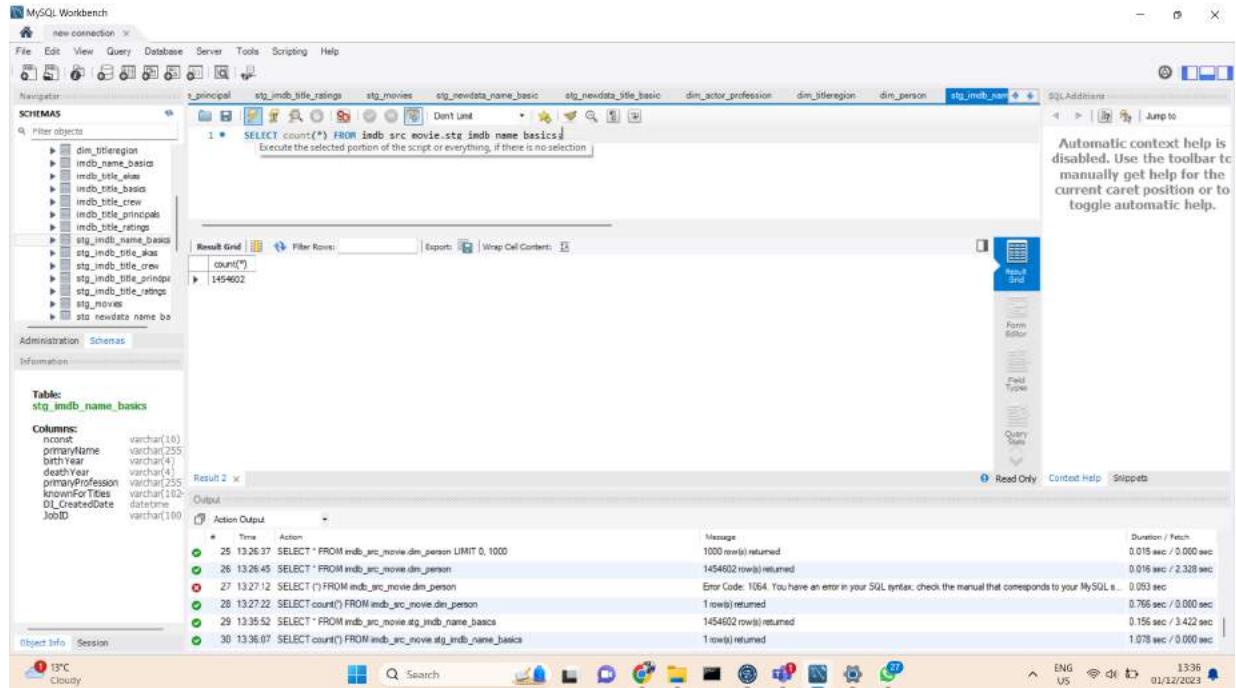
Jobs and Screenshots of Tables:

Stage Tables:

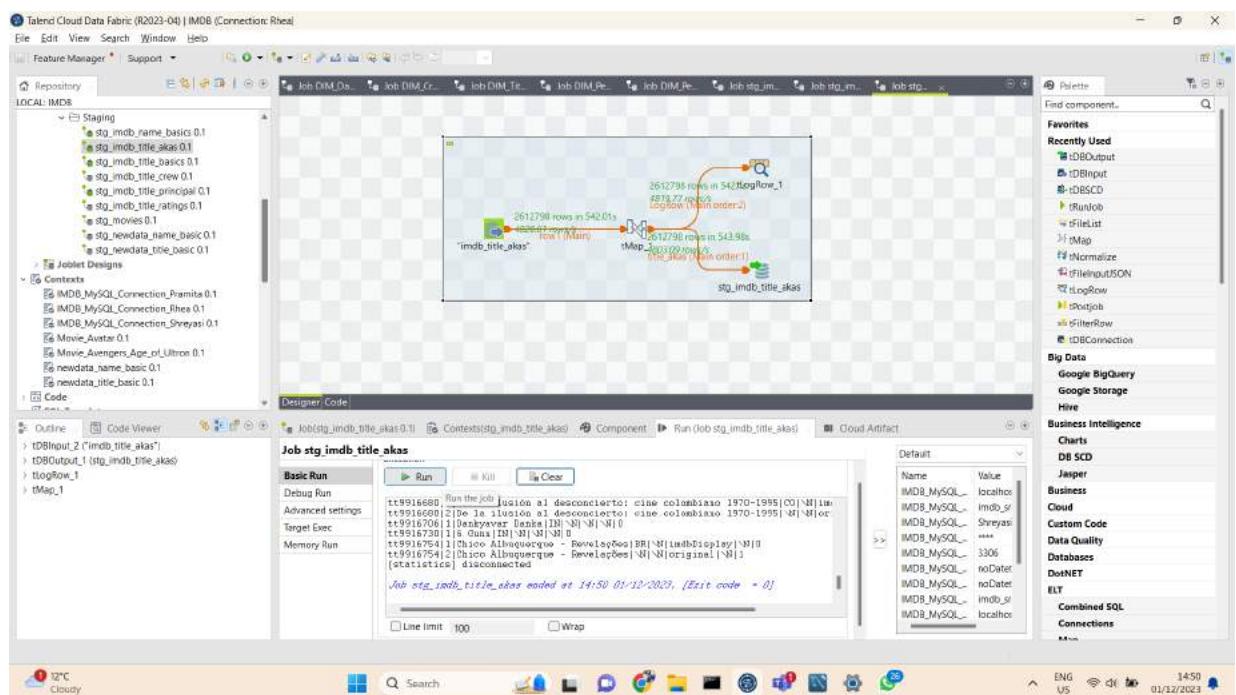
stg_imdb_name_basics:



The screenshot shows the MySQL Workbench interface. The left sidebar lists schemas and tables, including 'stg_imdb_name_basics'. The main area displays the structure and data of the 'stg_imdb_name_basics' table. The bottom pane shows the execution history of SQL queries, indicating successful runs for various selects against the 'imdb_src_movie' schema.



stg imdb title akas



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: stg_imdb_title_akas

Table: stg_imdb_title_akas

Columns:

- ctid: varchar(10)
- ordering: int
- title: varchar(104)
- region: varchar(255)
- language: varchar(255)
- types: varchar(104)
- attributes: varchar(104)
- isOriginalTitle: varchar(255)
- DL_CreatedDate: datetime
- pid: varchar(109)

Result Grid

ctid	ordering	title	region	language	types	attributes	isOriginalTitle	DL_CreatedDate	pid
t0000574-002	1	Bohemian	W	original	W	1		2023-11-29 13:20:17	985d9H
t0000574-002	2	Bohemian	ES	W	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000574-001	1	Kelly Annémalé Nérénée	HU	W	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000574-002	2	Ned Kelly and His Gang	AU	W	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000574-003	3	Prix à Kelly-o bandi	RS	W	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000574-004	4	The Story of the Kelly Gang	GB	W	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000574-005	5	The Story of the Kelly Gang	W	W	original	W	1	2023-11-29 13:20:17	985d9H
t0000574-006	6	The Story of the Kelly Gang	SG	en	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000574-007	7	The Story of the Kelly Gang	US	W	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000574-008	8	Ned Kelly and His Gang	W	W	W	0	2023-11-29 13:20:17	985d9H	
t0000574-009	9	Die Geschichte der Kelly Br... The Story of the Kelly Gang	DE	W	W	0	2023-11-29 13:20:17	985d9H	
t0000574-010	10	The Story of the Kelly Gang	AU	W	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000591-001	1	L'enfant prodige	FR	W	imdbDisplay	W	0	2023-11-29 13:20:17	985d9H
t0000591-002	2	L'enfant prodige	W	W	original	W	1	2023-11-29 13:20:17	985d9H
t0000591-003	3	The Prodigal Son	US	W	W	0	2023-11-29 13:20:17	985d9H	

Action Output

#	Time	Action	Message	Duration / Fetch
11	03:52:52	SELECT * FROM `imdb_src_movie.dim_actor_profession` LIMIT 0, 1000	1000 rows(s) returned	0.063 sec / 0.015 sec
12	03:53:01	SELECT * FROM `imdb_src_movie.dim_actor_profession` LIMIT 0, 1000	1000 rows(s) returned	0.000 sec / 0.000 sec
13	03:53:25	SELECT * FROM `imdb_src_movie.dim_actor_knowfor` LIMIT 0, 1000	0 rows(s) returned	0.000 sec / 0.000 sec
14	03:53:33	SELECT * FROM `imdb_src_movie.dim_actor_profession` LIMIT 0, 1000	1000 rows(s) returned	0.016 sec / 0.047 sec
15	03:54:50	SELECT * FROM `imdb_src_movie.stg_imdb_name_basics` LIMIT 0, 1000	1000 rows(s) returned	0.078 sec / 0.000 sec
16	03:56:24	SELECT * FROM `imdb_src_movie.stg_imdb_title_akas` LIMIT 0, 1000	1000 rows(s) returned	0.906 sec / 0.000 sec

Result 2

count(*)
261298

Action Output

#	Time	Action	Message	Duration / Fetch
27	13:27:12	SELECT (*) FROM `imdb_src_movie.dim_person`	Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL ...	0.093 sec
28	13:27:22	SELECT count(*) FROM `imdb_src_movie.dim_person`	1row(s) returned	0.796 sec / 0.000 sec
29	13:35:51	SELECT * FROM `imdb_src_movie.stg_imdb_name_basics`	1454602 row(s) returned	0.19 sec / 3.422 sec
30	13:36:07	SELECT count(*) FROM `imdb_src_movie.stg_imdb_name_basics`	1row(s) returned	1.078 sec / 0.000 sec
31	13:36:47	SELECT * FROM `imdb_src_movie.stg_imdb_title_akas`	261298 row(s) returned	0.203 sec / 5.375 sec
32	13:37:07	SELECT count(*) FROM `imdb_src_movie.stg_imdb_title_akas`	1row(s) returned	1.031 sec / 0.000 sec

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: stg_imdb_title_akas

Table: stg_imdb_title_akas

Columns:

- ctid: varchar(10)
- ordering: int
- title: varchar(104)
- region: varchar(255)
- language: varchar(255)
- types: varchar(104)
- attributes: varchar(104)
- isOriginalTitle: varchar(255)
- DL_CreatedDate: datetime
- pid: varchar(109)

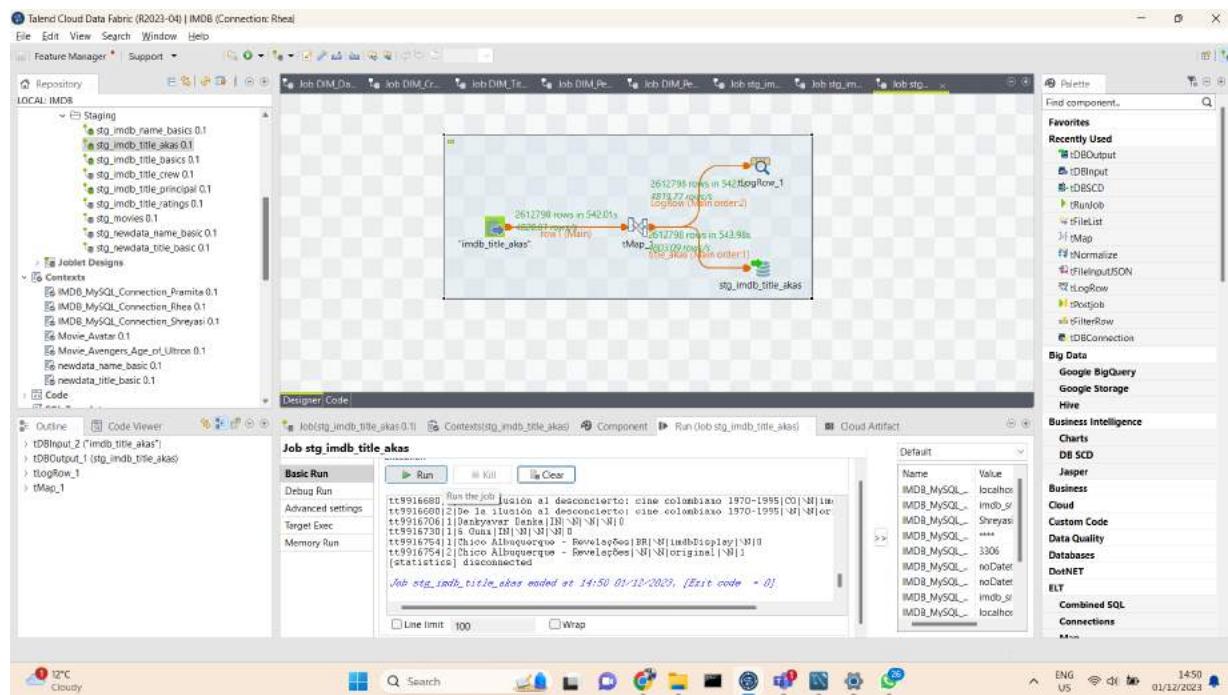
Result Grid

count(*)
261298

Action Output

#	Time	Action	Message	Duration / Fetch
27	13:27:12	SELECT (*) FROM `imdb_src_movie.dim_person`	Error Code: 1064. You have an error in your SQL syntax; check the manual that corresponds to your MySQL ...	0.093 sec
28	13:27:22	SELECT count(*) FROM `imdb_src_movie.dim_person`	1row(s) returned	0.796 sec / 0.000 sec
29	13:35:51	SELECT * FROM `imdb_src_movie.stg_imdb_name_basics`	1454602 row(s) returned	0.19 sec / 3.422 sec
30	13:36:07	SELECT count(*) FROM `imdb_src_movie.stg_imdb_name_basics`	1row(s) returned	1.078 sec / 0.000 sec
31	13:36:47	SELECT * FROM `imdb_src_movie.stg_imdb_title_akas`	261298 row(s) returned	0.203 sec / 5.375 sec
32	13:37:07	SELECT count(*) FROM `imdb_src_movie.stg_imdb_title_akas`	1row(s) returned	1.031 sec / 0.000 sec

stg_imdb_title_basics



The screenshot shows the MySQL Workbench interface. The left sidebar shows the 'SCHEMAS' section with tables like 'imdb_title_principals', 'stg_imdb_title_akas', and 'stg_imdb_title_basics'. The central area shows the results of a query on the 'stg_imdb_title_basics' table:

```
SELECT * FROM imdb_src_movie.stg_imdb_title_basics;
```

The results grid displays the following data:

tsn	titleType	primaryTitle	originalTitle	isAdult	startYear	endYear	runtimeMinutes	genres	DL_CreatedDate
t5000502	movie	Bohemian		0	1905	W	100	W	2023-12-01 14:29
t5000574	movie	The Story of the Kelly Gang		0	1906	W	70	Action,Adventure,Biography	2023-12-01 14:29
t5000575	movie	The Prodigal Son	L'enfant prodige	0	1907	W	90	Drama	2023-12-01 14:29
t5000615	movie	Rebel Under Arms	Rebel Under Arms	0	1907	W	W	Drama	2023-12-01 14:29
t5000658	movie	Aneto	Aneto	0	1908	W	W	Drama	2023-12-01 14:29
t5000675	movie	Don Quijote	Don Quijote	0	1908	W	W	Drama	2023-12-01 14:29
t5000679	movie	The Fairylogue and Radio-Plays	The Fairylogue and Radio-Plays	0	1908	W	120	Adventure,Fantasy	2023-12-01 14:29
t5000793	movie	Andreas Hefer	Andreas Hefer	0	1909	W	W	Drama	2023-12-01 14:29
t5000814	movie	La bocana de Mar Chica	La bocana de Mar Chica	0	1909	W	W	W	2023-12-01 14:29

Object Info

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: rinnopal, stg_imdb_title_ratings, stg_movies, stg_newdata_name_basic, stg_newdata_title_basic, dim_date, dim_crew, dim_person, dim_tierregion, stg_imdb_title

SQL Editor:

```
1 *  SELECT count(*) FROM imbd_src_movie.stg_imdb_title_basics;
```

Result Grid:

count(*)
607423

Information:

Table: stg_imdb_title_basics

Columns:

- title: varchar(10)
- titleType: varchar(255)
- primaryTitle: varchar(1024)
- originalTitle: varchar(1024)
- startYear: tinyint
- endYear: tinyint
- runtimesMinutes: varchar(10)
- genres: varchar(255)
- DT_CreatedDate: datetime
- JobID: varchar(100)

Action Output:

#	Time	Action	Message	Duration / Fetch
49	13:47:00	SELECT * FROM imbd_src_movie.dim_person	1454602 rows(s) returned	0.031 sec / 1.563 sec
50	13:47:05	SELECT count(*) FROM imbd_src_movie.dim_person	1 rows(s) returned	0.266 sec / 0.000 sec
51	13:47:45	SELECT * FROM imbd_src_movie.dim_tierregion	261298 rows(s) returned	0.140 sec / 4.344 sec
52	13:48:05	SELECT count(*) FROM imbd_src_movie.dim_tierregion	1 rows(s) returned	0.397 sec / 0.000 sec
53	14:51:51	SELECT * FROM imbd_src_movie.stg_imdb_title_basics	607423 rows(s) returned	0.093 sec / 2.157 sec
54	14:52:43	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_basics	1 rows(s) returned	0.312 sec / 0.000 sec

Object Info Session

12°C Cloudy ENG US 1452 01/12/2023

stg_imdb_title_crew

Talend Cloud Data Fabric (R2023-04) | IMDB (Connection: Rhee)

File Edit View Search Window Help

Feature Manager Support

Repository LOCAL: IMDB

Staging:

- stg_imdb_name_basics_0.1
- stg_imdb_title_akas_0.1
- stg_imdb_title_basics_0.1
- stg_imdb_title_crew_0.1
- stg_imdb_title_principal_0.1
- stg_imdb_title_ratings_0.1
- stg_movies_0.1
- stg_newdata_name_basic_0.1
- stg_newdata_title_basic_0.1

Joblet Designs

Contexts:

- IMDB_MySQL_Connection_Pramita_0.1
- IMDB_MySQL_Connection_Rhee_0.1
- IMDB_MySQL_Connection_Shreyasi_0.1
- Movie_Avatar_0.1
- Movie_Avengers_Age_of_Ultron_0.1
- newdata_name_basic_0.1
- newdata_title_basic_0.1

Code

Job: Job(stg_imdb_title_crew_0.1) Context: Contexting_imdb_title_crew Component Run Job stg_imdb_title_crew Job stg_imdb_title_crew

Designer/Code

Job stg_imdb_title_crew

Basic Run

Run Kill Clear

```
1st91620|MI|Mar73|17/09|01-12-2023|2yvwkk
1st91622|MI|Mar27|24/09|awz724811|nx5274499|nx9272491|01-12-2023|2yvwkk
1st91660|OM|652213|nm652213|nx10538576|01-12-2023|2yvwkk
1st916706|OM|764440|nx7931903|01-12-2023|2yvwkk
1st916730|OM|10538612|nx10538612|01-12-2023|2yvwkk
[statistics] disconnected
Job stg_imdb_title_crew ended at 14:58 01/12/2023, [Exit code = 0]
```

Default

Name	Value
IMDB_MySQL_Connection_Pramita	localhost
IMDB_MySQL_Connection_Rhee	imdb_sf
IMDB_MySQL_Connection_Shreyasi	****
IMDB_MySQL_Connection	3306
IMDB_MySQL_Connection	noDate
IMDB_MySQL_Connection	noDate
IMDB_MySQL_Connection	imdb_sf
IMDB_MySQL_Connection	localhost

Find component...

Favorites Recently Used

- tDBOutput
- tDBInput
- tDBSCD
- tRunJob
- tFileList
- tMap
- tNormalize
- tFilePutJSON
- tLogRow
- tPostJob
- tFilterRow
- tDBConnection

Big Data

- Google BigQuery
- Google Storage
- Hive

Business Intelligence

- Charts
- DB SCD
- Jasper
- Business
- Cloud
- Custom Code
- Data Quality
- Databases
- DotNET
- ELT
- Combined SQL
- Connections

12°C Cloudy ENG US 1593 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: lim_actor_infoinfo dim_actor_profession dim_actor_profession dim_actor_knowfor dim_actor_profession stg_imdb_name_basics stg_imdb_title_aliases stg_imdb_title

Navigator: lim_actor_infoinfo dim_actor_profession dim_actor_profession dim_actor_knowfor dim_actor_profession dim_actor_profession stg_imdb_name_basics stg_imdb_title_aliases stg_imdb_title

Result Grid: Filter Rows: Export: Wrap Cell Content: Fetch rows: Result Grid Form Editor Field Types Query Stats

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

1 * SELECT * FROM imdb_src_movie.stg_imdb_title_crew;

icon	directors	writers	DL_CreatedDate	JobID
...	nm0063413	nm0063413,nm0657258,nm0675388	2023-11-30 01:18:58	b7y4
...	nm0095974	nm0095974	2023-11-30 01:18:58	b7y4
...	nm009981	nm009981	2023-11-30 01:18:58	b7y4
...	nm0096115	nm0096115,nm533998	2023-11-30 01:18:58	b7y4
...	nm006620	nm006620	2023-11-30 01:18:58	b7y4
...	nm006675	nm006675	2023-11-30 01:18:58	b7y4
...	nm006679	nm006679,nm0877783	2023-11-30 01:18:58	b7y4
...	nm000793	nm000793	2023-11-30 01:18:58	b7y4
...	nm000814	nm000814	2023-11-30 01:18:58	b7y4
...	nm000838	nm000838	2023-11-30 01:18:58	b7y4
...	nm000842	nm000842	2023-11-30 01:18:58	b7y4
...	nm000846	nm000846	2023-11-30 01:18:58	b7y4
...	nm000850	nm000850,nm055220,nm063413	2023-11-30 01:18:58	b7y4
...	nm000859	nm000859,nm005717	2023-11-30 01:18:58	b7y4
...	nm000862	nm000862,nm0876467	2023-11-30 01:18:58	b7y4

Object Info Session

13 °C Mostly clear 03:57 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: otp_newdata_name_basic stg_newdata_title_basic dim_actor_profession dim_stateregion dim_person stg_imdb_name_basics stg_imdb_title_aliases stg_imdb_title

Navigator: otp_newdata_name_basic stg_newdata_title_basic dim_actor_profession dim_stateregion dim_person stg_imdb_name_basics stg_imdb_title_aliases stg_imdb_title

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

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

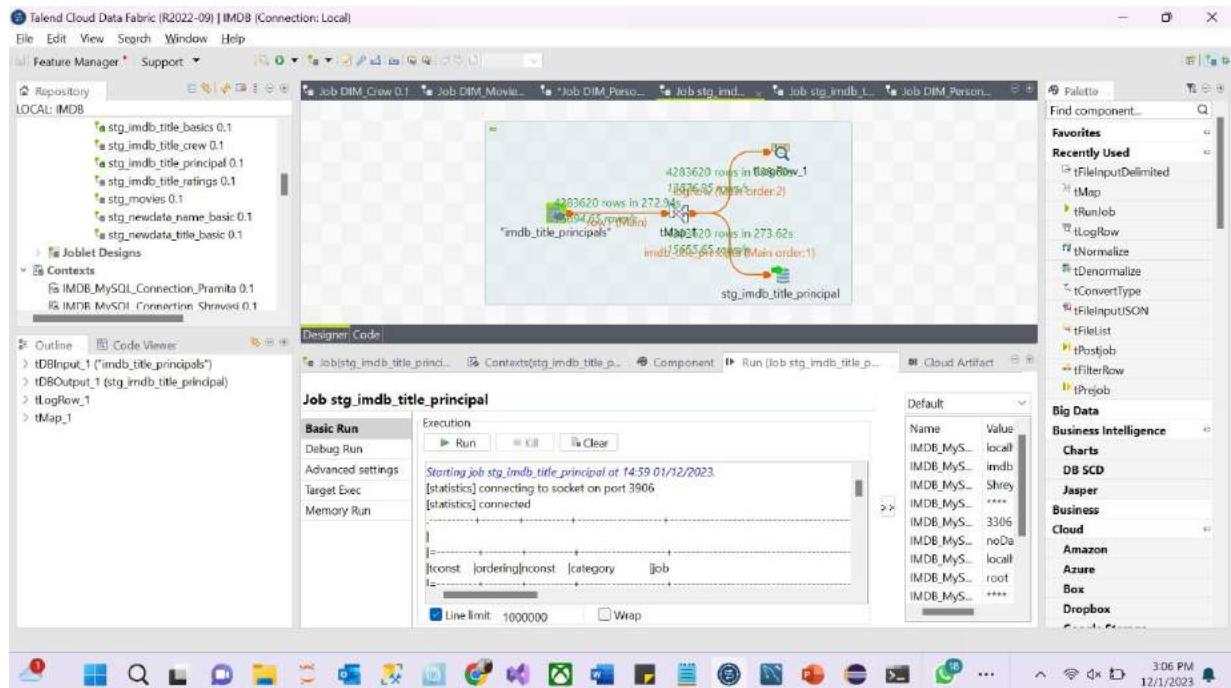
1 * SELECT count(*) FROM imdb_src_movie.stg_imdb_title_crew;

count(*)
607423

Object Info Session

13 °C Cloudy 13:37 01/12/2023

stg_imdb_title_principal



The screenshot shows the MySQL Workbench interface. The left sidebar lists databases and tables, including 'imdb_src_movie.stg_imdb_title_principal'. A query editor window is open with the following SQL statement:

```
1 *  SELECT * FROM imdb_src_movie.stg_imdb_title_principal;
```

The results grid shows 10 rows of data:

	id	name	category	job	characters	DL_CreatedDate	JobId
1	t0000502	nm0215752	actor	actress	W	2023-11-30 15:08:09	lyprf
2	t0000502	nm0252720	actor	actress	W	2023-11-30 15:08:09	lyprf
3	t0000502	nm0063413	director	actress	W	2023-11-30 15:08:09	lyprf
4	t0000502	nm0657268	writer	actress	W	2023-11-30 15:08:09	lyprf
5	t0000502	nm0675388	writer	actress	W	2023-11-30 15:08:09	lyprf
6	t0000574	nm0846887	actress	actress	W ["Vale Kelly"]	2023-11-30 15:08:09	lyprf
7	t0000574	nm0846994	actor	actress	W ["School Master"]	2023-11-30 15:08:09	lyprf
8	t0000574	nm091224	actor	actress	W ["Zoe Byrne"]	2023-11-30 15:08:09	lyprf
9	t0000574	nm091224	director	actress	W ["Steve Hart"]	2023-11-30 15:08:09	lyprf
10	t0000574	nm0986879	producer	actress	W	2023-11-30 15:08:09	lyprf
11	t0000574	nm1317210	producer	actress	W	2023-11-30 15:08:09	lyprf
12	t0000574	nm425854	producer	actress	W	2023-11-30 15:08:09	lyprf
13	t0000574	nm0499111	producer	actress	W	2023-11-30 15:08:09	lyprf
14	t0000574	nm2401839	composer	actress	W	2023-11-30 15:08:09	lyprf
15	t0000574	nm0675329	cinematographer	actress	W	2023-11-30 15:08:09	lyprf

The bottom status bar indicates the session is mostly clear and shows system information like battery level, signal strength, and date/time.

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: basic stg_newdata_side_basic dim_actor_profession dim_stfregion dim_person stg_imdb_name_basics stg_imdb_title_akas stg_imdb_title_crew stg_imdb_title_principals

Query:

```
1 *  SELECT count(*) FROM imbd_src_movie.stg_imdb_title_principal
```

Result Grid:

count(*)
4283620

Information:

Table: stg_imdb_title_principal

Columns:

- idname: varchar(10)
- ordering: int
- const: varchar(10)
- category: varchar(255)
- join: varchar(1024)
- characters: varchar(1024)
- DLOreatedDate: datetime
- Jold: varchar(100)

Action Output:

#	Time	Action	Message	Duration / Fetch
31	13:36:47	SELECT * FROM imbd_src_movie.stg_imdb_title_akas	2612798 rows(s) returned	0.203 sec / 5.375 sec
32	13:37:07	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_akas	1 rows(s) returned	0.031 sec / 0.000 sec
33	13:37:28	SELECT * FROM imbd_src_movie.stg_imdb_title_crew	60742 rows(s) returned	0.157 sec / 0.937 sec
34	13:37:49	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_crew	1 rows(s) returned	0.156 sec / 0.000 sec
35	13:38:20	SELECT * FROM imbd_src_movie.stg_imdb_title_principal	4283620 rows(s) returned	0.079 sec / 7.812 sec
36	13:40:16	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_principal	1 rows(s) returned	1.359 sec / 0.000 sec

Object Info Session

13°C Cloudy ENG US 13:40 01/12/2023

stg_imdb_title_ratings

Talend Cloud Data Fabric (R2023-04) | IMDB (Connection: Rhei)

File Edit View Search Window Help

Feature Manager Support 100% | Job: stg_imdb_title_ratings.0.1

Repository LOCAL: IMDB

Job Designs Standard IMDB_Jobs Load_Dimensions Load_Fact Staging stg_imdb_name_basics.0.1 stg_imdb_name_basics.0.1 stg_imdb_title_akas.0.1 stg_imdb_title_basics.0.1 stg_imdb_title_crew.0.1 stg_imdb_title_principal.0.1 stg_imdb_title_ratings.0.1 stg_movies.0.1 stg_newdata_name_basic.0.1 stg_newdata_title_basic.0.1

Job: stg_imdb_title_ratings.0.1

Code SQL Templates Metadata

Outline Code Viewer

Job: stg_imdb_title_ratings.0.1 Contexts(stg_imdb_title_ratings) Component Run (Job stg_imdb_title_ratings) Cloud Artifact

Basic Run Execution Run Kill Clear

```
tDBInput_2 ("imdb_title_ratings")
tMap_1
    377171 rows in 24.60s
    377171 rows in 24.7s
    logRow (MainOrder2)
    277171 rows in 24.7s
    1723241 rows(s)
    logRow (MainOrder2)

    377171 rows in 25.45s
    10890.33 rows(s)
    imbd_title_ratings (Main Order1)

    "stg_imdb_title_ratings"
```

Advanced settings Target Exec Memory Run

Component

Default

Name	Value
IMDB_MySQL_localhost	localhost
IMDB_MySQL_imdb_sf	imdb_sf
IMDB_MySQL_Shreyasi	Shreyasi
IMDB_MySQL_	****
IMDB_MySQL_3306	3306
IMDB_MySQL_noDate	noDate
IMDB_MySQL_noDate	noDate
IMDB_MySQL_BLT	BLT

Notification To take full advantage of your Studio, we recommend you get available updates. Learn more

10°C Light rain ENG US 17:00 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas

Table: stg_imdb_title_ratings

Columns:

- iconID: varchar(10)
- averageRating: double(22,0)
- numVotes: int
- DL_CreatedDate: datetime
- JobID: varchar(100)

Result Grid

iconID	averageRating	numVotes	DL_CreatedDate	JobID
t00000002	4	14	2022-11-30 15:31:28	q0Q9q9
t00000074	6	759	2022-11-30 15:31:28	q0Q9q9
t00000091	9	17	2022-11-30 15:31:28	q0Q9q9
t00000615	4	23	2022-11-30 15:31:28	d00099
t00000620	4	24	2022-11-30 15:31:28	d00099
t00000675	5	19	2022-11-30 15:31:28	q0Q9q9
t00000679	5	45	2022-11-30 15:31:28	q0Q9q9
t00000793	5	20	2022-11-30 15:31:28	q0Q9q9
t00000962	5	16	2022-11-30 15:31:28	q0Q9q9
t00000986	5	37	2022-11-30 15:31:28	q0Q9q9
t00000941	5	23	2022-11-30 15:31:28	q0Q9q9
t0001010	5	14	2022-11-30 15:31:28	q0Q9q9
t0001028	4	17	2022-11-30 15:31:28	q0Q9q9
t0001038	4	21	2022-11-30 15:31:28	q0Q9q9
t0001049	4	16	2022-11-30 15:31:28	q0Q9q9

Action Output

#	Time	Action	Message	Duration / Fetch
14	03:53:53	SELECT * FROM imbd_src_movie.dim_actor_profession LIMIT 0, 1000	1000 rows(s) returned	0.016 sec / 0.047 sec
15	03:54:50	SELECT * FROM imbd_src_movie.stg_imdb_name_basics LIMIT 0, 1000	1000 rows(s) returned	0.079 sec / 0.000 sec
16	03:56:24	SELECT * FROM imbd_src_movie.stg_imdb_aka LIMIT 0, 1000	1000 rows(s) returned	0.908 sec / 0.000 sec
17	03:57:01	SELECT * FROM imbd_src_movie.stg_imdb_title_crew LIMIT 0, 1000	1000 rows(s) returned	0.171 sec / 0.016 sec
18	03:57:45	SELECT * FROM imbd_src_movie.stg_imdb_title_principal LIMIT 0, 1000	1000 rows(s) returned	0.129 sec / 0.031 sec
19	03:58:17	SELECT * FROM imbd_src_movie.stg_imdb_title_ratings LIMIT 0, 1000	1000 rows(s) returned	0.105 sec / 0.016 sec

Output

Result Grid

count(*)
277171

Action Output

#	Time	Action	Message	Duration / Fetch
33	13:37:28	SELECT * FROM imbd_src_movie.stg_imdb_title_crew	607423 rows(s) returned	0.157 sec / 0.937 sec
34	13:37:49	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_crew	1 row(s) returned	0.156 sec / 0.000 sec
35	13:38:45	SELECT * FROM imbd_src_movie.stg_imdb_title_principal	4283520 rows(s) returned	0.079 sec / 7.812 sec
36	13:40:16	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_principal	1 row(s) returned	1.359 sec / 0.000 sec
37	13:40:48	SELECT * FROM imbd_src_movie.stg_imdb_title_ratings	277171 rows(s) returned	0.187 sec / 0.484 sec
38	13:40:47	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_ratings	1 row(s) returned	0.015 sec / 0.000 sec

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas

Table: stg_imdb_title_ratings

Columns:

- iconID: varchar(10)
- averageRating: double(22,0)
- numVotes: int
- DL_CreatedDate: datetime
- JobID: varchar(100)

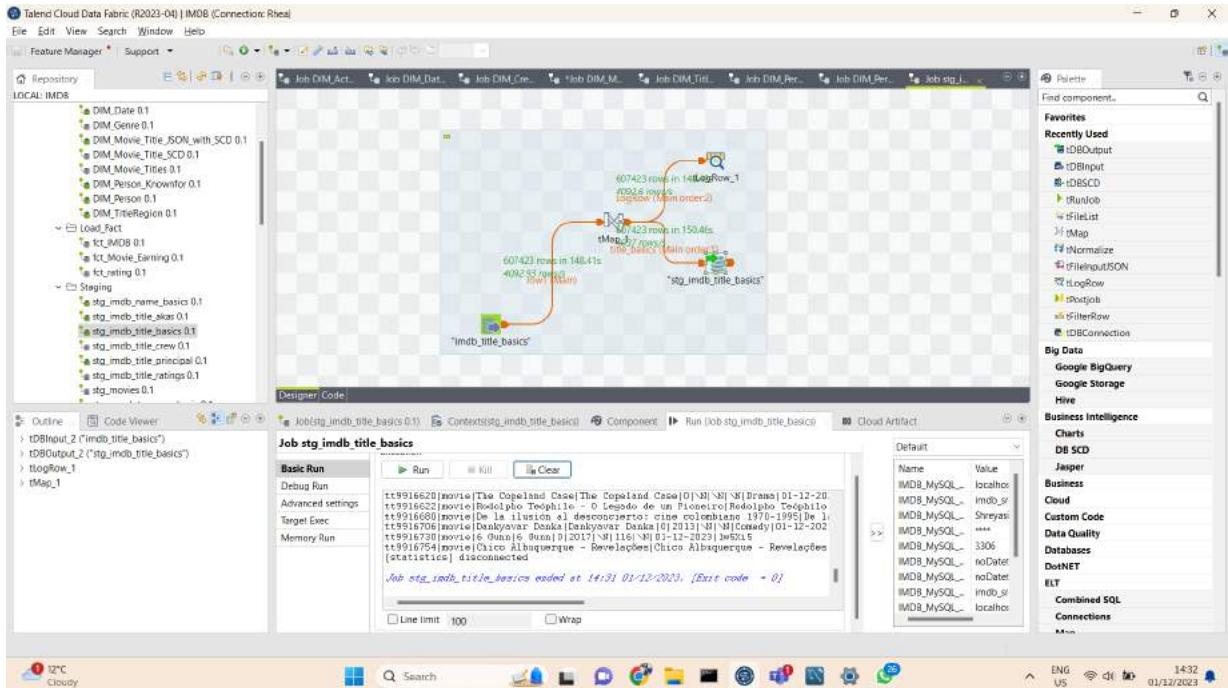
Result Grid

count(*)
277171

Action Output

#	Time	Action	Message	Duration / Fetch
33	13:37:28	SELECT * FROM imbd_src_movie.stg_imdb_title_crew	607423 rows(s) returned	0.157 sec / 0.937 sec
34	13:37:49	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_crew	1 row(s) returned	0.156 sec / 0.000 sec
35	13:38:45	SELECT * FROM imbd_src_movie.stg_imdb_title_principal	4283520 rows(s) returned	0.079 sec / 7.812 sec
36	13:40:16	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_principal	1 row(s) returned	1.359 sec / 0.000 sec
37	13:40:48	SELECT * FROM imbd_src_movie.stg_imdb_title_ratings	277171 rows(s) returned	0.187 sec / 0.484 sec
38	13:40:47	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_ratings	1 row(s) returned	0.015 sec / 0.000 sec

stg_title_basics



The screenshot shows the MySQL Workbench interface. The top navigation bar includes tabs for 'Local instance MySQL80 (imdb_tr.x)', 'Local instance MySQL80 (imdb..)', 'File', 'Edit', 'View', 'Query', 'Database', 'Server', 'Tools', 'Scripting', and 'Help'. Below the navigation is a toolbar with various icons for database management. The 'Navigator' pane on the left lists 'SCHEMAS' and 'Tables' (including dim_crew, dim_movie_title_my, dim_name_basics, dim_title_akas, dim_title_basics, dim_title_crew, dim_title_principals, dim_title_ratings, stg_imdb_name_basics, stg_imdb_title_akas, stg_imdb_title_basics, stg_imdb_title_crew, stg_imdb_title_princ, stg_imdb_title_rating). The main area displays a query result grid for the 'stg_imdb_title_basics' table, showing columns like tconst, titleType, primaryTitle, originalTitle, isAdult, startYear, endYear, runtimeMinutes, genres, and DI_CreatedDate. The grid contains several movie entries. At the bottom, the 'Object Info' and 'Session' panes are visible, along with a system tray at the bottom right.

MySQL Workbench

Local Instance MySQL80 (imdb_src) Local instance MySQL80 (imdb_...)

File Edit View Query Database Server Tools Scripting Help

Navigator Schemas

SCHEMAS

Tables

- dim_crew
- dim_movie_title_my...
- imdb_name_basics
- imdb_title_aliases
- imdb_title_basics
- imdb_title_crew
- imdb_title_principals
- imdb_title_ratings
- stg_imdb_name_bas...
- stg_imdb_title_aliases
- stg_imdb_title_basics
- stg_imdb_title_crew
- stg_imdb_title_principals
- stg_imdb_title_ratings

Query 1 pet_care_connect - Schema dim_movie_title_mysql dim_crew stg_imdb_title_basics

```
1 •   SELECT count(*) FROM imdb_src.movie.stg_imdb_title_basics;
```

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

	count(*)
	507423

Administration Schemas Information

Table: stg_imdb_title_basics

Columns:

Column	Type	PK	Nullable
const	varchar(255)	PK	NO
titleType	varchar(255)		NO
primaryTitle	varchar(255)		NO
originalTitle	varchar(255)		NO

Action Output

Time	Action	Message	Duration / Fetch
6 14:12:23	SELECT * FROM imdb_src.movie.stg_imdb_title_basics LIMIT 0, 500	500 row(s) returned	0.016 sec / 0.000 sec
7 14:12:45	SELECT count(*) FROM imdb_src.movie.stg_imdb_title_basics LIMIT 0, 500	1 row(s) returned	0.375 sec / 0.000 sec

Object Info Session

23:22 PM 12/1/2023

stg_movies

Talend Cloud Data Fabric (R2022-09) | IMDB (Connection: Local)

File Edit View Search Window Help

Feature Manager Support

Repository Joblet Designers Contexts

IMDB MySQL Connection Pramita 0.1

LOCAL: IMDB

- stg_imdb_title_basics 0.1
- stg_imdb_title_crew 0.1
- stg_imdb_title_principal 0.1
- stg_imdb_title_principals 0.1
- stg_imdb_title_ratings 0.1
- stg_movies 0.1
- stg_newdata_name_basic 0.1
- stg_newdata_title_basic 0.1

Joblet Designers

Job stg_movies 0.1

Designer Code

Job(stg_movies 0.1) Contexts(stg_movies) Component Run (Job stg_movies) Cloud Artifact

Job stg_movies

Basic Run

Execution

Run Kill Clear

tt0120338|Titanic|2017-12-10|{320,646}-24%|-83%|75|275|{659,328,801}|7,297
tt0120338|Titanic|2017-12-11|{57,419,-64%}-81%|75|599|{659,336,220}|7,298
tt0120338|Titanic|2017-12-12|{10,392,0,4}-74%|75|139|{659,346,612}|7,299
tt0120338|Titanic|2017-12-13|{31,1,555,0,3,-6,99}-75|181|{659,360,167}|7,300
tt0120338|Titanic|2017-12-14|{3,777,-72%}-86%|75|500|{659,363,944}|7,301
[statistics] disconnected

Job stg_movies ended at 17:32:30/11/2023. [Exit code = 0]

Default

Name	Value
strFilePath	C:\Us...
IMDB.MyS...	noDa...
IMDB.MyS...	3306
IMDB.MyS...	Shrey...
IMDB.MyS...	imdb
IMDB.MyS...	local
IMDB.MyS...	local
IMDB.MyS...	root
IMDB.MyS...	***
IMDB.MyS...	noDa...

Favorites Recently Used

- tFileInputDelimited
- tMap
- tRunJob
- tLogRow
- tNormalize
- tDeNormalize
- tConvertType
- tFileInputJSON
- tFileList
- tPostjob
- tFilterRow
- tPrejob

Big Data Business Intelligence

- Charts
- DB SCD
- Jasper
- Business
- Cloud
- Amazon
- Azure
- Box
- Dropbox

5:46 PM 11/30/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: Filter objects

Table: stg_movies

Columns:

- title: varchar(100)
- date: varchar(100)
- rank: varchar(100)
- gross: varchar(100)
- id: varchar(100)
- l_w: varchar(100)
- theaters: varchar(100)
- per_theater: varchar(100)
- total_gross: varchar(100)
- days: varchar(100)
- dl_createddate: datetime
- jobid: varchar(100)

Action Output:

- 15 03:54:50 SELECT * FROM imbd_src_movie.stg_imbd_name_basics LIMIT 0, 1000 1000 rows(s) returned 0.078 sec / 0.000 sec
- 16 03:56:24 SELECT * FROM imbd_src_movie.stg_imbd_title_akas LIMIT 0, 1000 1000 rows(s) returned 0.095 sec / 0.000 sec
- 17 03:57:01 SELECT * FROM imbd_src_movie.stg_imbd_title_crew LIMIT 0, 1000 1000 rows(s) returned 0.171 sec / 0.016 sec
- 18 03:57:45 SELECT * FROM imbd_src_movie.stg_imbd_title_principal LIMIT 0, 1000 1000 rows(s) returned 0.125 sec / 0.031 sec
- 19 03:58:17 SELECT * FROM imbd_src_movie.stg_imbd_title_ratings LIMIT 0, 1000 1000 rows(s) returned 0.109 sec / 0.016 sec
- 20 03:58:45 SELECT * FROM imbd_src_movie.stg_movies LIMIT 0, 1000 265 rows(s) returned 0.203 sec / 0.000 sec

Result Grid | Filter Rows | Export | Wrap Cell Content | Result Grid | Form Editor | Field Types | Query Stats | Read Only | Context Help | Snippets

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Result Grid | Filter Rows | Export | Wrap Cell Content | Result Grid | Form Editor | Field Types | Query Stats | Read Only | Context Help | Snippets

Object Info Session

13°C Mostly clear ENG US 03:59 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: Filter objects

Table: stg_movies

Columns:

- title: varchar(100)
- date: varchar(100)
- rank: varchar(100)
- gross: varchar(100)
- id: varchar(100)
- l_w: varchar(100)
- theaters: varchar(100)
- per_theater: varchar(100)
- total_gross: varchar(100)
- days: varchar(100)
- dl_createddate: datetime
- jobid: varchar(100)

Action Output:

- 35 13:38:20 SELECT * FROM imbd_src_movie.stg_imbd_title_principal 4293620 rows(s) returned 0.079 sec / 7.812 sec
- 36 13:40:16 SELECT count(*) FROM imbd_src_movie.stg_imbd_title_principal 1 row(s) returned 1.359 sec / 0.000 sec
- 37 13:40:37 SELECT * FROM imbd_src_movie.stg_imbd_title_rating 277171 rows(s) returned 0.187 sec / 0.494 sec
- 38 13:40:47 SELECT count(*) FROM imbd_src_movie.stg_imbd_title_ratings 1 row(s) returned 0.015 sec / 0.000 sec
- 39 13:41:07 SELECT * FROM imbd_src_movie.stg_movies 265 rows(s) returned 0.109 sec / 0.000 sec
- 40 13:41:22 SELECT count(*) FROM imbd_src_movie.stg_movies 1 row(s) returned 0.000 sec / 0.000 sec

Result Grid | Filter Rows | Export | Wrap Cell Content | Result Grid | Form Editor | Field Types | Query Stats | Read Only | Context Help | Snippets

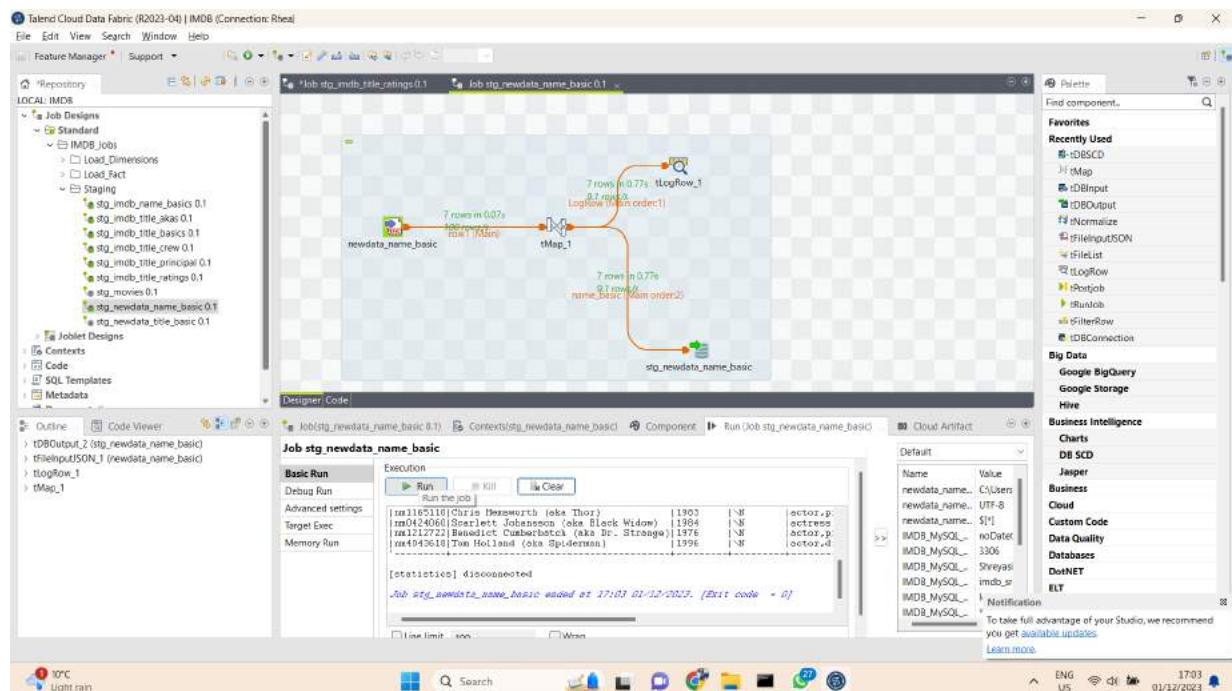
Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Result Grid | Filter Rows | Export | Wrap Cell Content | Result Grid | Form Editor | Field Types | Query Stats | Read Only | Context Help | Snippets

Object Info Session

13°C Cloudy ENG US 13:41 01/12/2023

stg_newdata_name_basic



MySQL Workbench

new connection

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Table: stg_newdata_name_basic

Columns:

nconst	primaryName	birthYear	deathYear	primaryProfession	knownForTitles	DL_CreatedDate	JobID
nn0000375	Robert Downey Jr. (aka Iron Man)	1965		actor,producer,soundtrack	tt0371746,tt1415476,tt1300854,tt0988845	2023-11-30 15:55:23	ff7c7u
nn026265	Chris Evans (aka Captain America)	1981		actor,producer,soundtrack	tt2394577,tt10483379,tt3498802,tt0946228	2023-11-30 15:55:23	ff7c7u
nn0749263	Mark Ruffalo (aka Hulk)	1967		actor,producer,soundtrack	tt0848228,tt1100009,tt1895567,tt0947926	2023-11-30 15:55:23	ff7c7u
nn1165110	Chris Hemsworth (aka Thor)	1983		actor,producer,soundtrack	tt0848228,tt0803069,tt1753888,tt3501632	2023-11-30 15:55:23	ff7c7u
nn0474060	Scarlett Johansson (aka Black Widow)	1984		actress,soundtrack,producer	tt108799,tt335266,tt0848228,tt162346	2023-11-30 15:55:23	ff7c7u
nn1212722	Benedict Cumberbatch (aka Dr. Strange)	1976		actor,producer,soundtrack	tt108101,tt2094970,tt1211837,tt1322269	2023-11-30 15:55:23	ff7c7u
nn0403618	Tom Holland (aka Spiderman)	1996		actor,director,producer	tt1649419,tt6320628,tt2250912,tt3498820	2023-11-30 15:55:23	ff7c7u

Result Grid

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Read Only

Object Info Session

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas

Table: stg_newdata_name_basic

Columns:

- name varchar(0)
- primaryName varchar(30)
- birthYear int
- deathYear varchar(2)
- deceasedProfession varchar(30)
- knownForTitles varchar(30)
- DL_CreatedDate datetime
- JobID varchar(100)

Result Grid | Filter Rows | Export | Wrap Cell Content | Result 2 | Action Output | Read Only | Context Help | Snippets

Action Output

#	Time	Action	Message	Duration / Fetch
37	13:40:36	SELECT * FROM imbd_src_movie.stg_imdb_title_ratings	27717 row(s) returned	0.167 sec / 0.484 sec
38	13:40:47	SELECT count(*) FROM imbd_src_movie.stg_imdb_title_ratings	1 row(s) returned	0.015 sec / 0.000 sec
39	13:41:07	SELECT * FROM imbd_src_movie.stg_movies	265 row(s) returned	0.109 sec / 0.000 sec
40	13:41:22	SELECT count(*) FROM imbd_src_movie.stg_movies	1 row(s) returned	0.000 sec / 0.000 sec
41	13:41:43	SELECT * FROM imbd_src_movie.stg_newdata_name_basic	7 row(s) returned	0.110 sec / 0.000 sec
42	13:41:52	SELECT count(*) FROM imbd_src_movie.stg_newdata_name_basic	1 row(s) returned	0.000 sec / 0.000 sec

Object Info Session

13°C Cloudy ENG US 13:41 01/12/2023

stg_newdata_title_basic

Talend Cloud Data Fabric (R2023-04) | IMDB (Connection: Rhei)

File Edit View Search Window Help

Repository

LOCAL: IMDB

Job Designs

Standard

IMDB_Jobs

- Load_Dimensions
- Load_Fact
- Staging
 - stg_imdb_name_basic_0.1
 - stg_imdb_title_akas_0.1
 - stg_imdb_title_basic_0.1
 - stg_imdb_title_crew_0.1
 - stg_imdb_title_principal_0.1
 - stg_imdb_title_ratings_0.1
 - stg_movies_0.1
 - stg_newdata_name_basic_0.1
 - stg_newdata_title_basic_0.1

JobListings

Contexts

Code

SQL Templates

Metadata

Job: Job stg_newdata_title_basic_0.1

Job: Job stg_newdata_name_basic_0.1

Job: Job stg_newdata_title_basic_0.1

Designer/Code

Execution

Basic Run

Run Kill Clear

```
ttt2495496 [movie] Star Wars: Episode VII - The Force Awakens (2015)
ttt454756 [movie] Avengers: Infinity War (2018)
ttt4154796 [movie] Avengers: Endgame (2019)
ttt0120338 [movie] Titanic (1997) remake
```

[statistics] disconnected

Job stg_newdata_title_basic ended at 17:04 01/12/2023. [Exit code = 0]

Time limit: 600 Wrapping

Palette

Find component...

Favorites

Recently Used

- tBSCD
- tMap
- tBInput
- tBOutput
- tNormalize
- tFileInputJSON
- tFileList
- tLogRow
- tPortobj
- tRunobj
- tFilterRow
- tDBConnection

Big Data

Google BigQuery

Google Storage

Hive

Business Intelligence

Charts

DB SCDF

Jasper

Business

Cloud

Custom Code

Data Quality

Databases

DotNET

ELT

Notification

To take full advantage of your Studio, we recommend you get available updates.

Learn more

10°C Light rain ENG US 17:04 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: stg_imdb_name_basic, stg_imdb_title_akas, stg_imdb_title_crew, stg_imdb_title_principal, stg_imdb_title_ratings, stg_movies, stg_newdata_name_basic, stg_newdata_title_basic

Table: stg_newdata_title_basic

Columns:

- iconId: varchar(10)
- titleType: varchar(5)
- primaryTitle: varchar(49)
- originalTitle: varchar(42)
- isAdult: int
- startYear: int
- endYear: int
- runtimeMinutes: int
- genres: varchar(24)
- DT_CreatedDate: datetime
- JOBID: varchar(100)

Result Grid

iconId	titleType	primaryTitle	originalTitle	isAdult	startYear	endYear	runtimeMinutes	genres
t0120338	movie	Titanic (1997)	Titanic	0	1997	W	194	Drama,Romance
t0499549	movie	Avatar (2009)	Avatar	0	2009	W	162	Action,Adventure,Fant
t0548228	movie	The Avengers (2012)	The Avengers	0	2012	W	143	Action,Adventure,Sci F
t110872600	movie	Spider-Man: No Way Home (2021)	Spider-Man: No Way Home	0	2021	W	148	Action,Adventure,Fant
t118250853	movie	Black Panther (2018)	Black Panther	0	2018	W	134	Action,Adventure,Sci F
t22991427	movie	Avengers: Age of Ultron (2015)	Avengers: Age of Ultron	0	2015	W	141	Action,Adventure,Sci F
t2488496	movie	Star Wars: Episode VII - The Force Awakens (2015)	Star Wars: Episode VII - The Force Awakens	0	2015	W	138	Action,Adventure,Sci F
t1454756	movie	Avengers: Infinity War (2018)	Avengers: Infinity War	0	2018	W	149	Action,Adventure,Sci F
t1454796	movie	Avengers: Endgame (2019)	Avengers: Endgame	0	2019	W	181	Action,Adventure,Dram
t0120338	movie	Titanic (1997) remake	Titanic	0	1997	W	194	Drama,Romance

Output

Action Output

- 17 03:57:01 Time Action SELECT * FROM `imdb_src_movie`.`stg_imdb_title_crew` LIMIT 0, 1000 1000 row(s) returned 0.171 sec / 0.016 sec
- 18 03:57:01 Time Action SELECT * FROM `imdb_src_movie`.`stg_imdb_title_principal` LIMIT 0, 1000 1000 row(s) returned 0.125 sec / 0.011 sec
- 19 03:58:17 Time Action SELECT * FROM `imdb_src_movie`.`stg_imdb_title_ratings` LIMIT 0, 1000 1000 row(s) returned 0.109 sec / 0.016 sec
- 20 03:58:46 Time Action SELECT * FROM `imdb_src_movie`.`stg_movies` LIMIT 0, 1000 269 row(s) returned 0.203 sec / 0.000 sec
- 21 03:59:26 Time Action SELECT * FROM `imdb_src_movie`.`stg_newdata_name_basic` LIMIT 0, 1000 7 row(s) returned 0.141 sec / 0.000 sec
- 22 03:59:53 Time Action SELECT * FROM `imdb_src_movie`.`stg_newdata_title_basic` LIMIT 0, 1000 10 row(s) returned 0.125 sec / 0.000 sec

Object Info Session

13°C Mostly clear 04:00 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: stg_imdb_name_basic, stg_imdb_title_akas, stg_imdb_title_crew, stg_imdb_title_principal, stg_imdb_title_ratings, stg_movies, stg_newdata_name_basic, stg_newdata_title_basic

Table: stg_newdata_title_basic

Columns:

- iconId: varchar(10)
- titleType: varchar(5)
- primaryTitle: varchar(49)
- originalTitle: varchar(42)
- isAdult: int
- startYear: int
- endYear: int
- runtimeMinutes: int
- genres: varchar(24)
- DT_CreatedDate: datetime
- JOBID: varchar(100)

Result Grid

count(*)
10

Output

Action Output

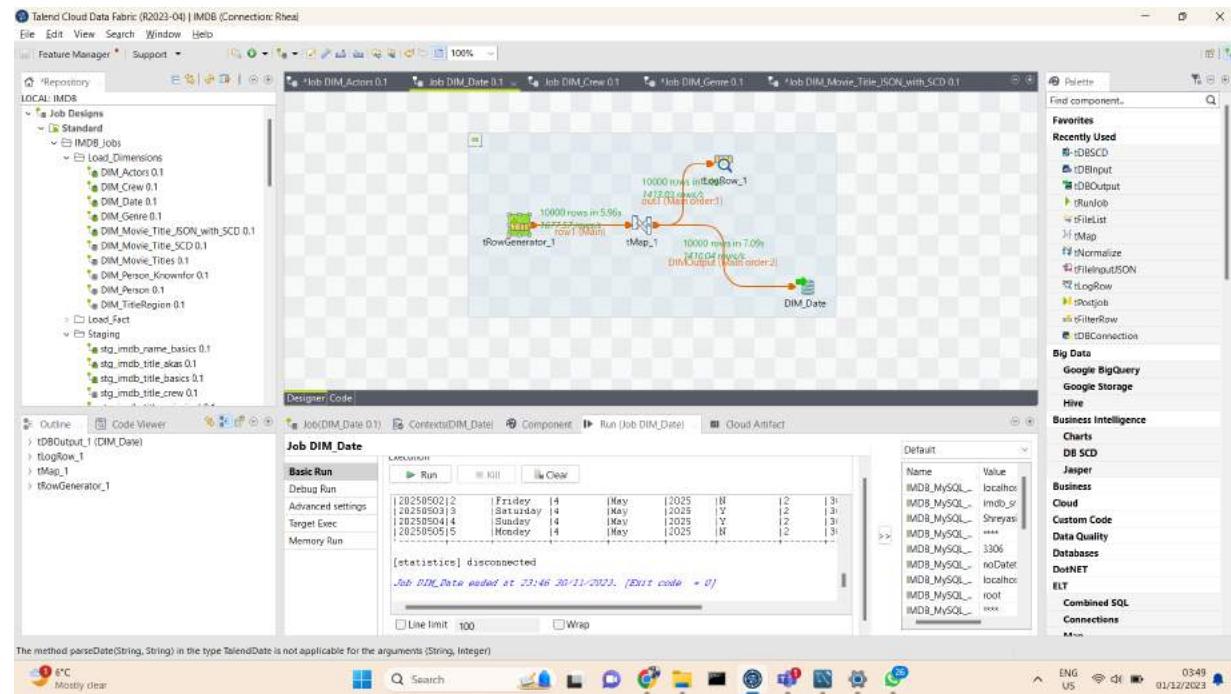
- 39 13:41:07 Time Action SELECT * FROM `imdb_src_movie`.`stg_movies` 265 row(s) returned 0.109 sec / 0.000 sec
- 40 13:41:22 Time Action SELECT count(*) FROM `imdb_src_movie`.`stg_movies` 1 row(s) returned 0.000 sec / 0.000 sec
- 41 13:41:42 Time Action SELECT * FROM `imdb_src_movie`.`stg_newdata_name_basic` 7 row(s) returned 0.110 sec / 0.000 sec
- 42 13:41:52 Time Action SELECT count(*) FROM `imdb_src_movie`.`stg_newdata_name_basic` 1 row(s) returned 0.000 sec / 0.000 sec
- 43 13:42:07 Time Action SELECT * FROM `imdb_src_movie`.`stg_newdata_title_basic` 10 row(s) returned 0.281 sec / 0.000 sec
- 44 13:42:15 Time Action SELECT count(*) FROM `imdb_src_movie`.`stg_newdata_title_basic` 1 row(s) returned 0.000 sec / 0.000 sec

Object Info Session

13°C Cloudy 13:42 01/12/2023

Dim Tables:

DIM Date:



The screenshot shows the MySQL Workbench interface with the "dim_date" table selected. The table structure is displayed with columns: dateRank, Day_Num, Date_Str, Month_Num, Month_Short, Year_Num, Is_Weekend, Qtr_Num, DI_CreatedDate, DI_Processed, and dt. The data grid shows rows from 1997-12-29 to 1998-01-02. Below the table, the "Output" tab displays the results of a series of actions, including dropping tables, selecting from them, and then selecting from the dim_date table itself. The duration for each action is shown in the "Duration / Fetch" column.

MySQL Workbench

Schemas: stg_imdb_title_akas, stg_imdb_title_crew, stg_imdb_title_principal, stg_imdb_title_ratings, stg_movies, stg_newdata_name_basic, stg_newdata_title_basic, dim_date

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Table: dim_date

Result Grid:

```
SELECT count(*) FROM imbd_src_movie.dim_date
Execute the selected portion of the script or everything, if there is no selection
```

Result 2:

Action	Time	Action	Message	Duration / Fetch
41	13:41:43	SELECT * FROM imbd_src_movie.stg_newdata_name_basic;	7 row(s) returned	0.110 sec / 0.000 sec
42	13:41:52	SELECT count(*) FROM imbd_src_movie.stg_newdata_name_basic;	1 row(s) returned	0.000 sec / 0.000 sec
43	13:42:07	SELECT * FROM imbd_src_movie.stg_newdata_title_basic;	10 row(s) returned	0.281 sec / 0.000 sec
44	13:42:15	SELECT count(*) FROM imbd_src_movie.stg_newdata_title_basic;	1 row(s) returned	0.000 sec / 0.000 sec
45	13:44:01	SELECT * FROM imbd_src_movie.dim_date;	10000 row(s) returned	0.125 sec / 0.109 sec
46	13:44:09	SELECT count(*) FROM imbd_src_movie.dim_date;	1 row(s) returned	0.000 sec / 0.000 sec

Object Info | Session

EUR/GBP 0.82% 13:44 01/12/2023

DIM Crew:

Talend Cloud Data Fabric (R2023-04) | IMDB (Connection: Rhei)

File Edit View Search Window Help

Job DIM_Crew 0.1

Job DIM_Actors 0.1 Job DIM_Date 0.1 Job DIM_Crew 0.1 Job DIM_Movie_Title_JSON_with_SCD 0.1

Repository Feature Manager Support

LOCAL: IMDB

Job Designs Standard IMDB_Jobs Load_Dimensions Load_Facts Staging

Job DIM_Crew

tDBInput_2 ("stg_imdb_title_principal") tDBOutput_2 (DIM_Crew) tMap_1

4283620 rows in 1860.8s 2309.92 MB/s 4283620 rows in 1863s 2309.92 MB/s (Main) tMap_1 DIM_Crew

Starting job DIM_Crew at 00:04 01-12-2023. [statistics] connecting to socket on port 3875 [statistics] disconnected [statistics] disconnected

Job DIM_Crew ended at 00:28 01-12-2023. [Exit code = 0]

The method parseDate(String, String) in the type TalendDate is not applicable for the arguments (String, Integer)

ENG US 03:50 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: food_inspection, imdb_src_movie, stg_imdb_name_basics, dim_actors, dim_date, dim_crew

Query 1: SELECT * FROM imdb_src_movie.dim_crew;

CrewID	const	name	category	ct_Createdate	JobID
1	nm000592	nm0215752	actor	2023-12-01 09:04:40	6936Q
2	nm000592	nm000592	producer	2023-12-01 09:04:40	6936Q
3	nm000592	nm0005413	director	2023-12-01 09:04:40	6936Q
4	nm000592	nm0657288	writer	2023-12-01 09:04:40	6936Q
5	nm000592	nm0675388	writer	2023-12-01 09:04:40	6936Q
6	nm000592	nm0846887	inches	2023-12-01 09:04:40	6936Q
7	nm000592	nm0846894	actor	2023-12-01 09:04:40	6936Q
8	nm000592	nm1412124	actor	2023-12-01 09:04:40	6936Q
9	nm000592	nm3002376	actor	2023-12-01 09:04:40	6936Q
10	nm000592	nm0846879	director	2023-12-01 09:04:40	6936Q
11	nm000592	nm0317210	producer	2023-12-01 09:04:40	6936Q
12	nm000592	nm0426854	producer	2023-12-01 09:04:40	6936Q
13	nm000592	nm0846911	producer	2023-12-01 09:04:40	6936Q
14	nm000592	nm2418394	composer	2023-12-01 09:04:40	6936Q
15	nm000592	nm0675239	cinemat...	2023-12-01 09:04:40	6936Q

dim_crew 1 x

Action Output:

Time	Action	Message	Duration / Fetch
3:22:52	DROP TABLE `imdb_src_movie`.`dim_actors`	0 row(s) affected.	0.204 sec
4:22:13:00	DROP TABLE `imdb_src_movie`.`dim_crew`	0 row(s) affected.	6.054 sec
5:22:52:01	SELECT * FROM `imdb_src_movie.stg_imdb_name_basics` LIMIT 0, 1000	1000 row(s) returned.	0.047 sec / 0.078 sec
6:03:48:53	SELECT * FROM `imdb_src_movie.dim_actors` LIMIT 0, 1000	1000 row(s) returned.	0.125 sec / 0.078 sec
7:03:49:57	SELECT * FROM `imdb_src_movie.dim_date` LIMIT 0, 1000	1000 row(s) returned.	0.047 sec / 0.031 sec
8:03:50:56	SELECT * FROM `imdb_src_movie.dim_crew` LIMIT 0, 1000	1000 row(s) returned.	0.078 sec / 0.000 sec

Object Info Session

12°C Mostly clear ENG US 03:51 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: food_inspection, imdb_src_movie, stg_imdb_title_crew, stg_imdb_title_principal, stg_imdb_title_ratings, stg_movies, stg_newdata_name_basic, stg_newdata_title_basic, dim_date, dim_crew

Query 1: SELECT count(*) FROM `imdb_src_movie.dim_crew`;

count(*)
4283620

Result 2 x

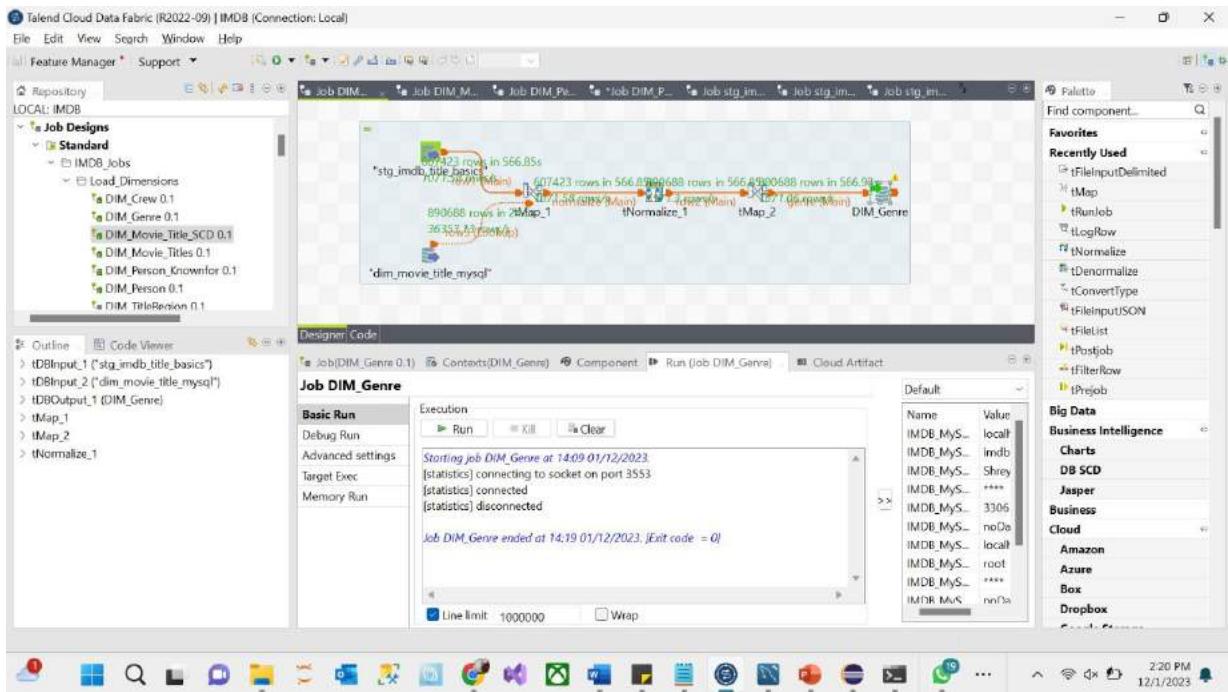
Action Output:

#	Time	Action	Message	Duration / Fetch
43	13:42:07	SELECT * FROM `imdb_src_movie.stg_newdata_title_basic`	10 row(s) returned.	0.281 sec / 0.000 sec
44	13:42:15	SELECT count(*) FROM `imdb_src_movie.stg_newdata_title_basic`	1 row(s) returned.	0.000 sec / 0.000 sec
45	13:44:01	SELECT * FROM `imdb_src_movie.dim_date`	10000 row(s) returned.	0.125 sec / 0.109 sec
46	13:44:01	SELECT count(*) FROM `imdb_src_movie.dim_date`	1 row(s) returned.	0.000 sec / 0.000 sec
47	13:45:05	SELECT * FROM `imdb_src_movie.dim_crew`	4283620 row(s) returned.	0.156 sec / 0.406 sec
48	13:46:19	SELECT count(*) FROM `imdb_src_movie.dim_crew`	1 row(s) returned.	1.157 sec / 0.000 sec

Object Info Session

12°C Cloudy ENG US 13:46 01/12/2023

DIM Genre:



The screenshot shows the MySQL Workbench interface. The left sidebar shows the 'Schemas' tree, with 'imdb_src_movie' selected. Under 'Tables', 'dim_genre' is shown. The 'Result Grid' pane displays the data for the 'dim_genre' table, which includes columns: GenresK, MovieTitle_SK, tconst, genres, DT_CreatedDate, and JobId. The data shows 17 rows of genre information. The 'Information' pane shows the table structure for 'dim_genre'. The bottom pane shows the 'Action Output' history, listing two recent actions: a 'DROP TABLE' command for 'imdb_src_movie.dim_title_principals' and a 'SELECT' query for the 'dim_genre' table.

MySQL Workbench

Local Instance MySQL80 (imdb_src) Local instance MySQL80 (imdb_src)

File Edit View Query Database Server Tools Scripting Help

Navigator: dim_genre

Query 1: SELECT COUNT(*) FROM imdb_src_movie.dim_genre

Result Grid:

COUNT(*)
890688

Administration Schemas Information

Table: dim_genre

Columns:

- GenreID int PK
- MovieTitle_SK int
- Title varchar
- Genres varchar
- DL_CreatedDate datetime
- InIt varchar

Action Output:

- 10 14:21:02 SELECT * FROM imdb_src_movie.dim_genre LIMIT 0, 500 500 row(s) returned 0.000 sec / 0.000 sec
- 11 14:21:26 SELECT COUNT(*) FROM imdb_src_movie.dim_genre LIMIT 0, 500 1 row(s) returned 0.297 sec / 0.000 sec

Object Info Session

Result 2:

Output: Read Only

Action Output:

- Time Action Message Duration / Fetch
- 10 14:21:02 SELECT * FROM imdb_src_movie.dim_genre LIMIT 0, 500 500 row(s) returned 0.000 sec / 0.000 sec
- 11 14:21:26 SELECT COUNT(*) FROM imdb_src_movie.dim_genre LIMIT 0, 500 1 row(s) returned 0.297 sec / 0.000 sec

2:21 PM 12/1/2023

DIM Movie Title SCD:

Talend Cloud Data Fabric (R2022-09) | IMDB (Connection: Local)

File Edit View Search Window Help

Feature Manager Support

Job Designs Standard IMDB.Jobs Load Dimensions

- IMDB.Crew 0.1
- IMDB.Genre 0.1
- IMDB.Movie_Title_SCD 0.1
- IMDB.Titles 0.1
- IMB_Person_Knownor 0.1
- IMB_Person 0.1
- IMB.TitleRepon 0.1

Designer Code

Job DIM_Movie_Title_SCD

Basic Run Execution

Run

Starting job DIM_Movie_Title_SCD at 17:31 01/12/2023.

[statistics] connecting to socket on port 3713

[statistics] connected

[statistics] disconnected

Job DIM_Movie_Title_SCD ended at 17:31 01/12/2023. [exit code = 0]

Code Viewer

```
tDBInput_2("stg_newdata_title_basic")
tBSCD_1(DIM_Movie_Title_SCD)
tMap_2
```

Palettes

Favorites Recently Used

- tFileInputDelimited
- tMap
- tRunJob
- tLogRow
- tNormalize
- tDeNormalize
- tConvertType
- tFileInputJSON
- tPostjob
- tFilterRow
- tPrejob

Big Data Business Intelligence

- Charts
- DB SCD
- Jasper
- Business
- Cloud
- Amazon
- Azure
- Box
- Dropbox

5:45 PM 12/1/2023

MySQL Workbench

Local instance MySQL80 (imdb_sr) Local instance MySQL80 (imdb_sr)

File Edit View Query Database Server Tools Scripting Help

Navigator: t1_person_knowfor dim_titlerregion fct_rating dim_date fct_movie_earning dim_person fct_imdb fct_rating fct_movie_earning fct_rating dim_movie_title

SCHEMAS: imdb_movies, imdb_src_movie

Tables: dim_crew, dim_date, dim_genre, dim_movie_title, dim_movie_title_my, dim_person, dim_person_knownfor, dim_person_profess, dim_titlerregion, fc_rating, imbd_title_ratings, imbd_title_principals

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

1 • SELECT * FROM imbd_src_movie.dim_movie_title;

tconst	titleType	primaryTitle	originalTitle	isAdult	runtmeMinutes	DI_CreatedDate	JobID	scd_sb
tt120338	movie	Titanic (1997)	Titanic	0	194	2023-12-01 14:24:13	rgo7sz	1997-0-
tt0499549	movie	Avatar (2009)	Avatar	0	162	2023-12-01 14:24:13	rgo7sz	2009-0-
tt0848228	movie	The Avengers (2012)	The Avengers	0	143	2023-12-01 14:24:13	rgo7sz	2012-0-
tt10872600	movie	Spider-Man: No Way Home (2021)	Spider-Man: No Way Home	0	148	2023-12-01 14:24:13	rgo7sz	2021-0-
tt1825683	movie	Black Panther (2018)	Black Panther	0	134	2023-12-01 14:24:13	rgo7sz	2018-0-
tt2395427	movie	Avengers: Age of Ultron (2015)	Avengers: Age of Ultron	0	141	2023-12-01 14:24:13	rgo7sz	2015-0-
tt2488496	movie	Star Wars: Episode VII - The Force Awakens (2015)	Star Wars: Episode VII - The Force Awakens	0	138	2023-12-01 14:24:13	rgo7sz	2015-0-

Object Info Session

Table: dim_movie_title

Columns: tconst, titleType, primaryTitle, originalTitle, isAdult, runtmeMinutes, DI_CreatedDate, JobID, scd_sb

Action Output: Time Action Message Duration / Fetch

- 990 17:27:18 SELECT * FROM imbd_src_movie.fct_rating LIMIT 0, 500 500 row(s) returned 0.000 sec / 0.000 sec
- 991 17:30:20 SELECT * FROM imbd_src_movie.dim_movie_title LIMIT 0, 500 10 row(s) returned 0.031 sec / 0.000 sec

5:30 PM 12/1/2023

MySQL Workbench

Local instance MySQL80 (imdb_sr) Local instance MySQL80 (imdb_sr)

File Edit View Query Database Server Tools Scripting Help

Navigator: t1_person_knowfor dim_titlerregion fct_rating dim_date fct_movie_earning dim_person fct_imdb fct_rating fct_movie_earning fct_rating dim_movie_title

SCHEMAS: imdb_movies, imdb_src_movie

Tables: dim_crew, dim_date, dim_genre, dim_movie_title, dim_movie_title_my, dim_person, dim_person_knownfor, dim_person_profess, dim_titlerregion, fc_rating, imbd_title_ratings, imbd_title_principals

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

1 • SELECT count(*) FROM imbd_src_movie.dim_movie_title;

count(*)
10

Object Info Session

Table: dim_movie_title

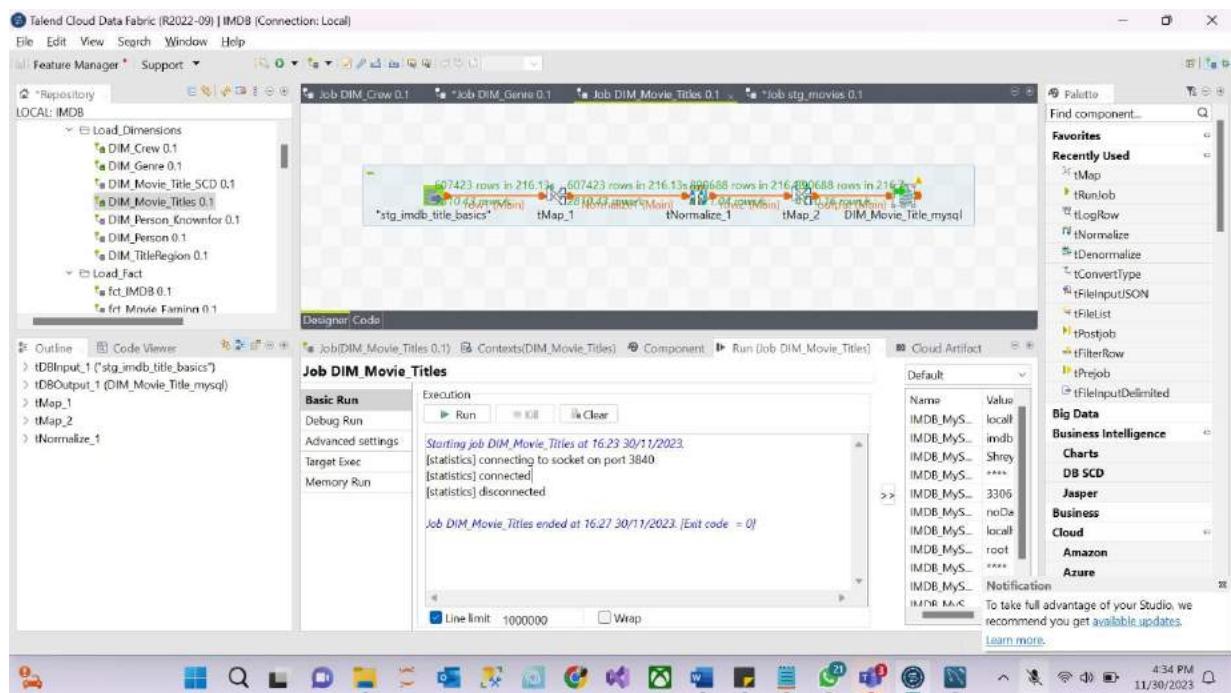
Columns: tconst, titleType, primaryTitle, originalTitle, isAdult, runtmeMinutes, DI_CreatedDate, JobID, scd_sb

Action Output: Time Action Message Duration / Fetch

- 991 17:30:20 SELECT * FROM imbd_src_movie.dim_movie_title LIMIT 0, 500 10 row(s) returned 0.031 sec / 0.000 sec
- 992 17:31:36 SELECT count(*) FROM imbd_src_movie.dim_movie_title LIMIT 0, 500 1 row(s) returned 0.000 sec / 0.000 sec

5:31 PM 12/1/2023

DIM Movie Titles:



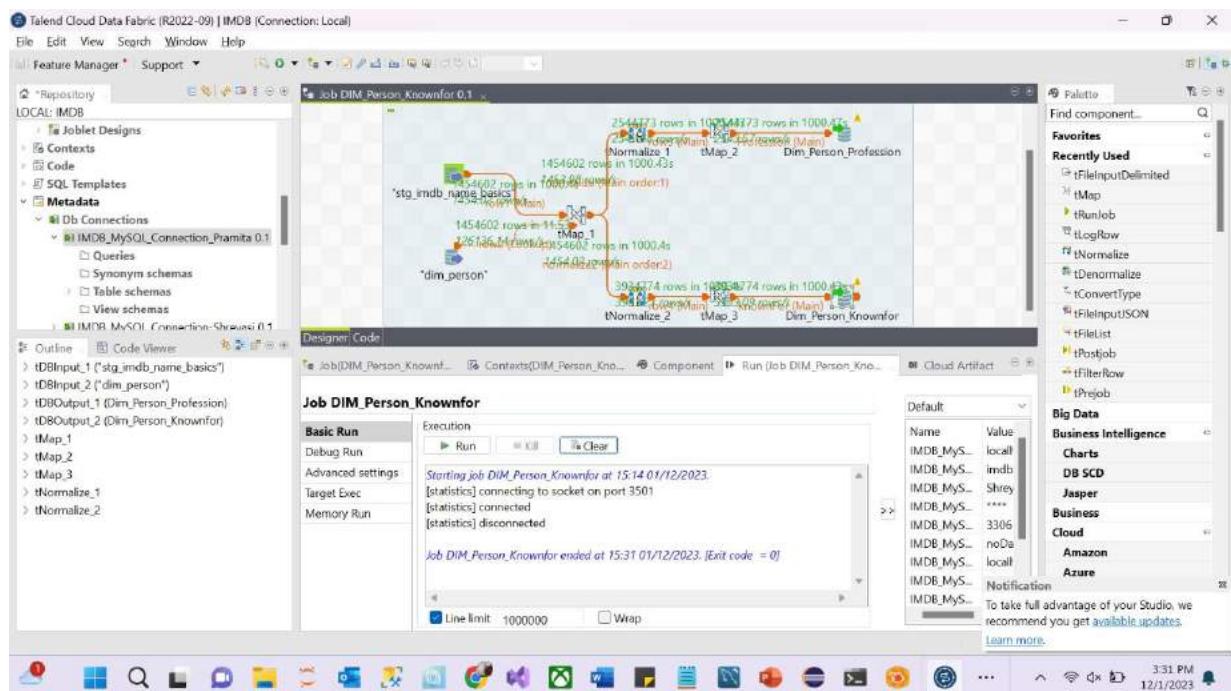
The screenshot shows the MySQL Workbench interface with the following details:

- Title Bar:** MySQL Workbench
- File Bar:** File, Edit, View, Query, Database, Server, Tools, Scripting, Help
- Navigator:** Local instance MySQL80 imdb_src.x, Local instance MySQL80 (imd... x)
- Query Editor:**
 - Schema: dim_movie_title_mysql
 - Query: SELECT * FROM imdb_src.movie.dim_movie_title_mysql;
 - Result Grid:

MovieTitle_SK	tconst	titleType	primaryTitle	originalTitle	isAdult	scd_start	endYear	runtimeMinutes	genres	DI_
2	tt0000502	movie	Bohemios	Bohemios	0	1905-01-01 00:00:00.0	1905	100		202
4	tt0000574	movie	The Story of the Kelly Gang	The Story of the Kelly Gang	0	1906-01-01 00:00:00.0	1906	70	Action	202
5	tt0000574	movie	The Story of the Kelly Gang	The Story of the Kelly Gang	0	1906-01-01 00:00:00.0	1906	70	Adventure	202
6	tt0000574	movie	The Story of the Kelly Gang	The Story of the Kelly Gang	0	1906-01-01 00:00:00.0	1906	70	Biography	202
8	tt0000591	movie	The Prodigal Son	L'enfant prodige	0	1907-01-01 00:00:00.0	1907	90	Drama	202
10	tt0000615	movie	Robbery Under Arms	Robbery Under Arms	0	1907-01-01 00:00:00.0	1907	100	Drama	202
12	tt0000630	movie	Hamlet	Hamlet	0	1908-01-01 00:00:00.0	1908	100	Drama	202
 - Output:**
 - Action Output:

Time	Action	Message	Duration / Fetch
1 23:43:26	Apply changes to Pet_Care_Connect	Changes applied	0.000 sec / 0.000 sec
2 14:07:21	SELECT * FROM imdb_src.movie.dim_movie_title_mysql LIMIT 0, 500	500 row(s) returned	0.000 sec / 0.000 sec
 - System Bar:** 2:07 PM, 12/1/2023

DIM Person Known for:



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: Local instance MySQL80 imbd_sr.x Local instance MySQL80 (im... x

SCHEMAS: dim_crew dim_genre dim_movie_title dim_movie_title_myr dim_person dim_person_knowfor dim_person_profess imbd_title_ratings imbd_title_principals imbd_title_crew imbd_title_basics imbd_title_akas stg_imdb_name_basics stg_imdb_title_aliases

Table: dim_person_knownfor

Columns: KnownFor_TitleSK int! PersonSK int! nconst varc knownForTitles varc nt_createdate date!

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

	KnownFor_TitleSK	PersonSK	nconst	knownForTitles	nt_createdate	JobID
4	1	nm0000001	t0072308	2023-12-01 15:14:58	BopmRU	
5	1	nm0000001	t0055137	2023-12-01 15:14:58	BopmRU	
6	1	nm0000001	t0050419	2023-12-01 15:14:58	BopmRU	
7	1	nm0000001	t0031983	2023-12-01 15:14:58	BopmRU	
10	2	nm0000002	t0038355	2023-12-01 15:14:58	BopmRU	
11	2	nm0000002	t0117057	2023-12-01 15:14:58	BopmRU	
12	2	nm0000002	t0071877	2023-12-01 15:14:58	BopmRU	
13	2	nm0000002	t0033782	2023-12-01 15:14:58	BopmRU	
...						

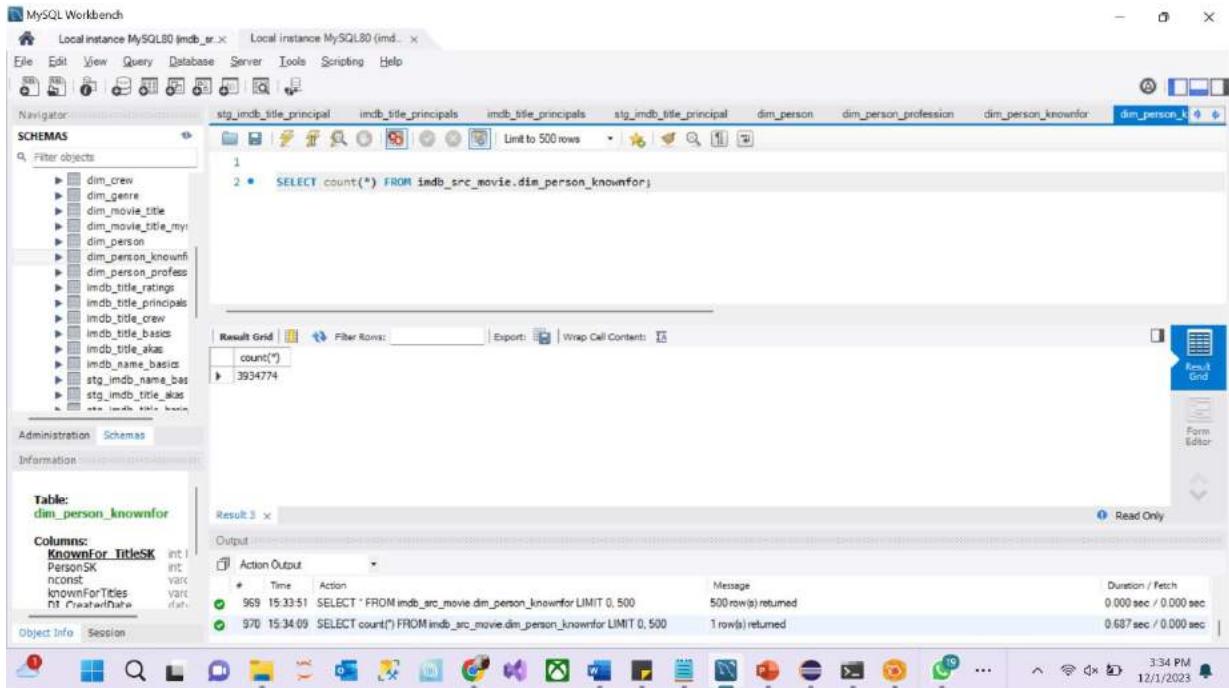
Action Output: Time Action Message Duration / Fetch

- 968 15:33:00 SELECT * FROM imbd_src_movie.dim_person_knownfor LIMIT 0, 500 500 row(s) returned 0.000 sec / 0.000 sec
- 969 15:33:51 SELECT * FROM imbd_src_movie.dim_person_knownfor LIMIT 0, 500 500 row(s) returned 0.000 sec / 0.000 sec

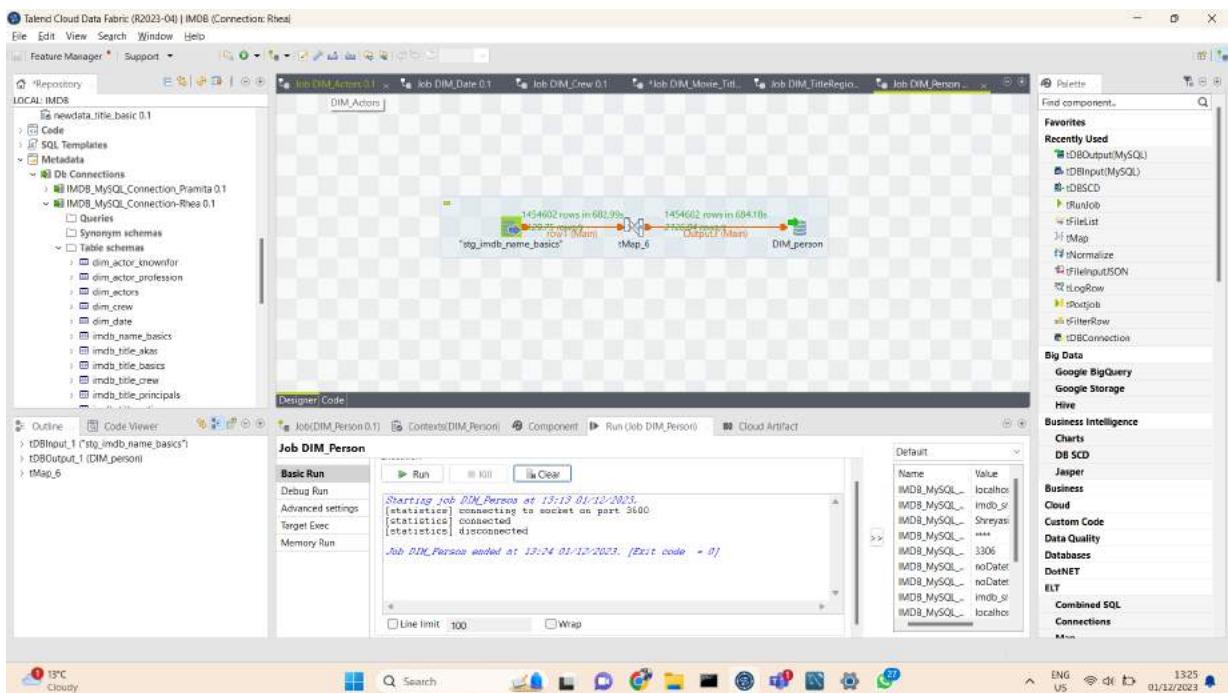
Result Grid Form Editor

Object Info Session

3:33 PM 12/1/2023



DIM Person:



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas

- food_inspection
- imdb_src_movie
 - Tables
 - dim_actor_knowsfor
 - dim_date
 - dim_crew
 - dim_actors
 - dim_actor_profession
 - dim_person
 - dim_titreregion
 - imdb_name_basic
 - imdb_title_ids
 - imdb_title_basic
 - imdb_title_crew

Administration Schemas Information

Table: dim_person

Columns:

PersonSK	const	primaryName	birthYear	deathYear	DL_CreatedDate	JobID
rm00000001		Fred Astaire	1899-01-01 00:00:00	1987-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000002		Lauren Bacall	1924-01-01 00:00:00	2014-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000003		Brigitte Bardot	1934-01-01 00:00:00	1982-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000004		John Belushi	1949-01-01 00:00:00	2007-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000005		Ingrid Bergman	1915-01-01 00:00:00	1982-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000006		Humphrey Bogart	1899-01-01 00:00:00	1957-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000008		Marlon Brando	1924-01-01 00:00:00	2004-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000009		Richard Burton	1925-01-01 00:00:00	1984-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000010		James Cagney	1899-01-01 00:00:00	1986-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000011		Gary Cooper	1901-01-01 00:00:00	1961-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000012		Bette Davis	1908-01-01 00:00:00	1989-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000013		Doris Day	1922-01-01 00:00:00	2019-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000014		Olivia de Havilland	1916-01-01 00:00:00	2020-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000015		James Dean	1931-01-01 00:00:00	1955-01-01 00:00:00	2023-12-01 13:13:36	789fFt

Result Grid | Filter Rows | Edit | Export/Import | Wrap Cell Content | Patch rows | Result Grid | Form Editor | Field Types | Query Stats | Output | Action Output | Read Only | Context Help | Snippets

Action Output

#	Time	Action	Message	Duration / Fetch
21	03:59:26	SELECT * FROM imdb_src_movie.stg_newdata_name_basic LIMIT 0, 1000	7 row(s) returned	0.141 sec / 0.000 sec
22	03:59:53	SELECT * FROM imdb_src_movie.stg_newdata_idh_basic LIMIT 0, 1000	10 row(s) returned	0.125 sec / 0.000 sec
23	04:39:09	SELECT * FROM imdb_src_movie.dim_actor_profession LIMIT 0, 1000	1000 row(s) returned	0.063 sec / 0.000 sec
24	04:39:29	SELECT * FROM imdb_src_movie.dim_titreregion LIMIT 0, 1000	1000 row(s) returned	0.047 sec / 0.000 sec
25	13:26:37	SELECT * FROM imdb_src_movie.dim_person LIMIT 0, 1000	1000 row(s) returned	0.015 sec / 0.000 sec
26	13:26:45	SELECT * FROM imdb_src_movie.dim_person	1454602 row(s) returned	0.016 sec / 2.338 sec
27	13:27:12	SELECT (*) FROM imdb_src_movie.dim_person	Error Code: 1064 You have an error in your SQL syntax; check the manual that corresponds to your MySQL s...	0.093 sec
28	13:27:22	SELECT count(*) FROM imdb_src_movie.dim_person	1 row(s) returned	0.766 sec / 0.000 sec

Object Info Session

13°C Cloudy 13:26 ENG US 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas

- food_inspection
- imdb_src_movie
 - Tables
 - dim_actor_knowsfor
 - dim_date
 - dim_crew
 - dim_actors
 - dim_actor_profession
 - dim_person
 - dim_titreregion
 - imdb_name_basic
 - imdb_title_ids
 - imdb_title_basic
 - imdb_title_crew

Administration Schemas Information

Table: dim_person

Columns:

PersonSK	const	primaryName	birthYear	deathYear	DL_CreatedDate	JobID
rm00000001		Fred Astaire	1899-01-01 00:00:00	1987-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000002		Lauren Bacall	1924-01-01 00:00:00	2014-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000003		Brigitte Bardot	1934-01-01 00:00:00	1982-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000004		John Belushi	1949-01-01 00:00:00	2007-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000005		Ingrid Bergman	1915-01-01 00:00:00	1982-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000006		Humphrey Bogart	1899-01-01 00:00:00	1957-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000008		Marlon Brando	1924-01-01 00:00:00	2004-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000009		Richard Burton	1925-01-01 00:00:00	1984-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000010		James Cagney	1899-01-01 00:00:00	1986-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000011		Gary Cooper	1901-01-01 00:00:00	1961-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000012		Bette Davis	1908-01-01 00:00:00	1989-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000013		Doris Day	1922-01-01 00:00:00	2019-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000014		Olivia de Havilland	1916-01-01 00:00:00	2020-01-01 00:00:00	2023-12-01 13:13:36	789fFt
rm00000015		James Dean	1931-01-01 00:00:00	1955-01-01 00:00:00	2023-12-01 13:13:36	789fFt

Result Grid | Filter Rows | Edit | Export/Import | Wrap Cell Content | Patch rows | Result Grid | Form Editor | Field Types | Query Stats | Output | Action Output | Read Only | Context Help | Snippets

Action Output

#	Time	Action	Message	Duration / Fetch
21	03:59:26	SELECT * FROM imdb_src_movie.stg_newdata_name_basic LIMIT 0, 1000	7 row(s) returned	0.141 sec / 0.000 sec
22	03:59:53	SELECT * FROM imdb_src_movie.stg_newdata_idh_basic LIMIT 0, 1000	10 row(s) returned	0.125 sec / 0.000 sec
23	04:39:09	SELECT * FROM imdb_src_movie.dim_actor_profession LIMIT 0, 1000	1000 row(s) returned	0.063 sec / 0.000 sec
24	04:39:29	SELECT * FROM imdb_src_movie.dim_titreregion LIMIT 0, 1000	1000 row(s) returned	0.047 sec / 0.000 sec
25	13:26:37	SELECT * FROM imdb_src_movie.dim_person LIMIT 0, 1000	1000 row(s) returned	0.015 sec / 0.000 sec
26	13:26:45	SELECT * FROM imdb_src_movie.dim_person	1454602 row(s) returned	0.016 sec / 2.338 sec
27	13:27:12	SELECT (*) FROM imdb_src_movie.dim_person	Error Code: 1064 You have an error in your SQL syntax; check the manual that corresponds to your MySQL s...	0.093 sec
28	13:27:22	SELECT count(*) FROM imdb_src_movie.dim_person	1 row(s) returned	0.766 sec / 0.000 sec

Object Info Session

13°C Cloudy 13:26 ENG US 01/12/2023

DIM Person Profession:

The screenshot shows the MySQL Workbench interface with the 'dim_person_profession' table selected in the schema list. The table has the following columns:

ProfessionSK	PersonSK	nconst	primaryProfession	DL_CreatedDate	JobID
1	1	nm0000001	soundtrack	2023-12-01 15:14:58	BopmRU
2	1	nm0000001	actor	2023-12-01 15:14:58	BopmRU
3	1	nm0000001	miscellaneous	2023-12-01 15:14:58	BopmRU
8	2	nm0000002	actress	2023-12-01 15:14:58	BopmRU
9	2	nm0000002	soundtrack	2023-12-01 15:14:58	BopmRU
14	3	nm0000003	actress	2023-12-01 15:14:58	BopmRU
15	3	nm0000003	soundtrack	2023-12-01 15:14:58	BopmRU
16	3	nm0000003	music_department	2023-12-01 15:14:58	BopmRU
17	4	nm0000004	actor	2023-12-01 15:14:58	BopmRU

Output window:

```

Action Output
Time Action
965 15:10:33 SELECT * FROM imdb_src_movie.dim_person LIMIT 0, 500
966 15:32:42 SELECT * FROM imdb_src_movie.dim_person_profession LIMIT 0, 500

```

The screenshot shows the MySQL Workbench interface with the 'dim_person_knownfor' table selected in the schema list. The table has the following columns:

KnownFor TitleSK	PersonSK	nconst	knownForTitles	DL_CreatedDate
count(*)				
2544773				

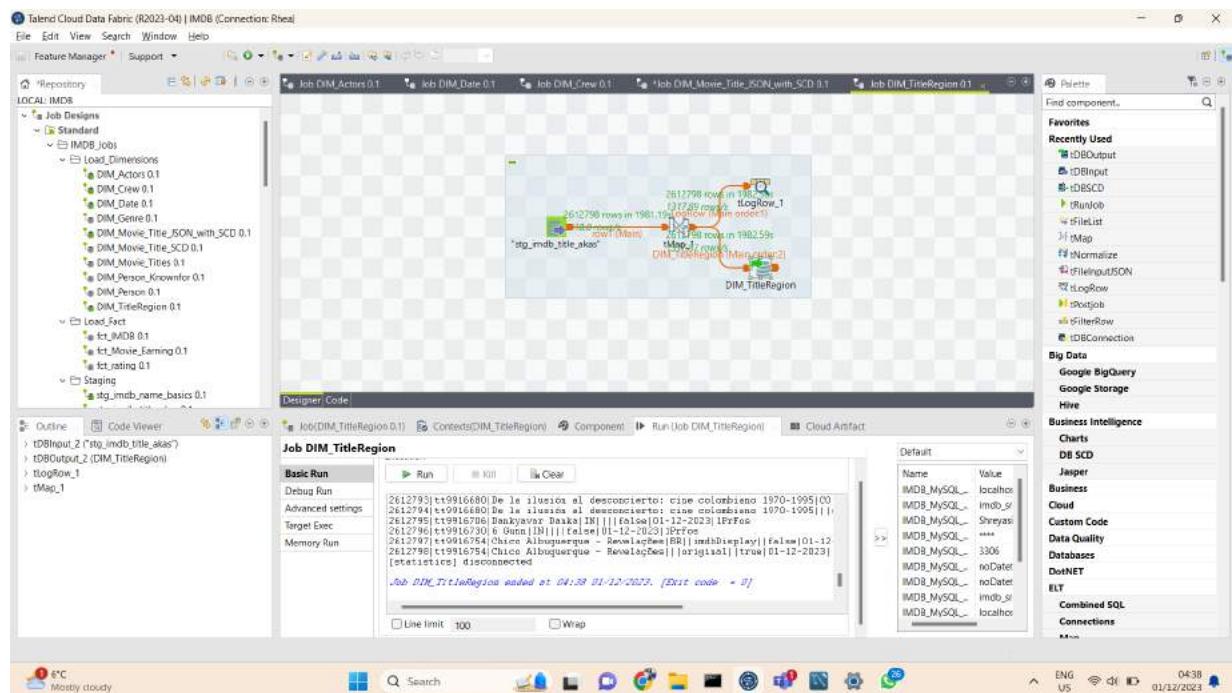
Output window:

```

Action Output
Time Action
970 15:34:09 SELECT count(*) FROM imdb_src_movie.dim_person_knownfor LIMIT 0, 500
971 15:34:28 SELECT count(*) FROM imdb_src_movie.dim_person_profession LIMIT 0, 500

```

DIM Title Region:



MySQL Workbench

Schemas

- food_inspection
- imdb_src_movie
 - Tables
 - dim_actor_knowfor
 - dim_date
 - dim_crew
 - dim_actors
 - dim_actor_profession
 - dim_name_basics
 - dim_title_akas
 - dim_title_basics
 - dim_title_crew
 - dim_title_principals

Table: dim_titleregion

Columns:

dimTitleRegionSK	dimTitleID	dimTitle	region	language	types	attributes	isOriginalTitle	DL_CreatedDate	pid
1	H0000502	Bohemian	ES	es	original	0	1	2023-12-01 04:05:26	IPFos
2	H0000502	Bohemian	ES	es	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
3	H0000524	Kelly bandásnak története	HU	hu	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
4	H0000524	Ned Kelly and His Gang	AU	en	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
5	H0000524	Prix de Kelly-nélki	RS	es	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
6	H0000524	The Story of the Kelly Gang	GB	en	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
7	H0000524	The Story of the Kelly Gang	GB	en	original	1	1	2023-12-01 04:05:27	IPFos
8	H0000524	The Story of the Kelly Gang	SG	en	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
9	H0000524	The Story of the Kelly Gang	US	en	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
10	H0000524	Ned Kelly and His Gang	US	en	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
11	H0000524	Die Geschichte der Kelly Br... DE	DE	de	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
12	H0000524	The Story of the Kelly Gang	AU	en	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
13	H0000591	L'enfant prodige	FR	fr	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
14	H0000591	L'enfant prodige	FR	fr	original	1	1	2023-12-01 04:05:27	IPFos
15	H0000591	The Bruyère Fan	FR	fr	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
16	H0000591	The Bruyère Fan	FR	fr	original	1	1	2023-12-01 04:05:27	IPFos
17	H0000591	The Bruyère Fan	FR	fr	imdbDisplay	0	0	2023-12-01 04:05:27	IPFos
18	H0000591	The Bruyère Fan	FR	fr	original	1	1	2023-12-01 04:05:27	IPFos
19	03:58:17	SELECT * FROM imbd_src_movie.dim_title_ratings LIMIT 0, 1000						1000 rows(s) returned	0.109 sec / 0.016 sec
20	03:58:46	SELECT * FROM imbd_src_movie.dim_movies LIMIT 0, 1000						265 rows(s) returned	0.203 sec / 0.000 sec
21	03:59:26	SELECT * FROM imbd_src_movie.dim_newdata_name_basic LIMIT 0, 1000						7 rows(s) returned	0.141 sec / 0.000 sec
22	03:59:53	SELECT * FROM imbd_src_movie.dim_newdata_type_basic LIMIT 0, 1000						10 rows(s) returned	0.125 sec / 0.000 sec
23	04:39:29	SELECT * FROM imbd_src_movie.dim_actor_profession LIMIT 0, 1000						1000 rows(s) returned	0.063 sec / 0.015 sec
24	04:39:29	SELECT * FROM imbd_src_movie.dim_dimregion LIMIT 0, 1000						1000 rows(s) returned	0.047 sec / 0.000 sec

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas

- stg_imdb_file_principal
- stg_imdb_site_ratings
- stg_movies
- stg_newdata_name_basic
- stg_newdata_title_basic
- dim_date
- dim_crew
- dim_person
- dim_titleregion

Automatic context help is disabled. Use the toolbar to manually get help for the current caret position or to toggle automatic help.

Table: dim_titleregion

Columns:

titleregionSK	int PK
id	varchar(10)
title	varchar(1024)
region	varchar(255)
language	varchar(255)
type	varchar(1024)
attributes	varchar(1024)
isOriginalTitle	varchar(255)
DL_CreatedDate	datetime
pid	varchar(109)

Result Grid | Filter Rows | Export | Wrap Cell Content | Result 2 | Read Only | Context Help | Snippets

Action Output

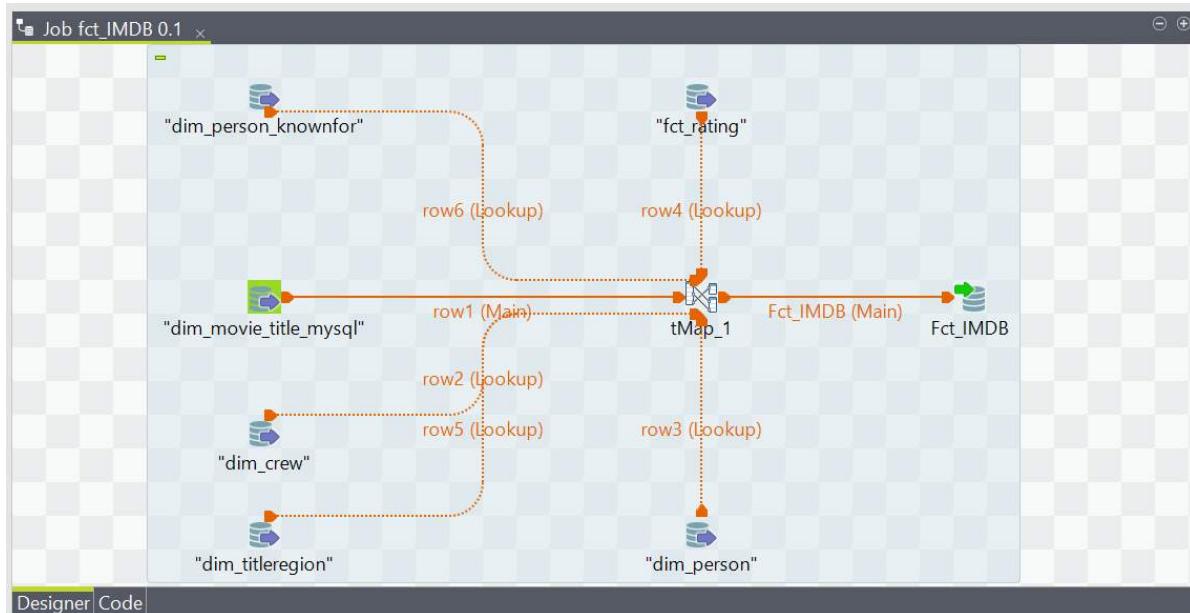
Time Action Message Duration / Fetch

- 47 13:45:05 SELECT * FROM imbd_src_movie.dim_crew 4283620 rows(s) returned 0.156 sec / 4.406 sec
- 48 13:46:19 SELECT count(*) FROM imbd_src_movie.dim_crew 1 row(s) returned 1.157 sec / 0.000 sec
- 49 13:47:00 SELECT * FROM imbd_src_movie.dim_person 1454602 rows(s) returned 0.031 sec / 1.563 sec
- 50 13:47:09 SELECT count(*) FROM imbd_src_movie.dim_person 1 row(s) returned 0.266 sec / 0.000 sec
- 51 13:47:46 SELECT * FROM imbd_src_movie.dim_titleregion 2612798 rows(s) returned 0.140 sec / 4.344 sec
- 52 13:48:05 SELECT count(*) FROM imbd_src_movie.dim_titleregion 1 row(s) returned 0.597 sec / 0.000 sec

Rain coming In about 2.5 hours

Fact Tables:

Fact IMDB:



```
1 •  select * from fct_imdb;
```

	IMDB_SK	CrewSK	PersonSK	titleRegionSK	MovieTitle_SK	RatingSK	KnownFor_TitleSK	DI_CreatedDate	JobID
▶	1	1	78737	2	1	1	273677	2023-12-02 23:45:08	932oOL
2	2	90925	2		1	1	315725	2023-12-02 23:45:08	932oOL
3	3	27130	2		1	1	0	2023-12-02 23:45:08	932oOL
4	4	220366	2		1	1	765605	2023-12-02 23:45:08	932oOL
5	5	225962	2		1	1	785128	2023-12-02 23:45:08	932oOL
6	6	278664	12		2	2	968730	2023-12-02 23:45:08	932oOL
7	7	278665	12		2	2	968731	2023-12-02 23:45:08	932oOL
8	8	616427	12		2	2	1713276	2023-12-02 23:45:08	932oOL
9	9	836037	12		2	2	2394060	2023-12-02 23:45:08	932oOL
10	10	278660	12		2	2	968717	2023-12-02 23:45:08	932oOL
11	11	111817	12		2	2	388148	2023-12-02 23:45:08	932oOL

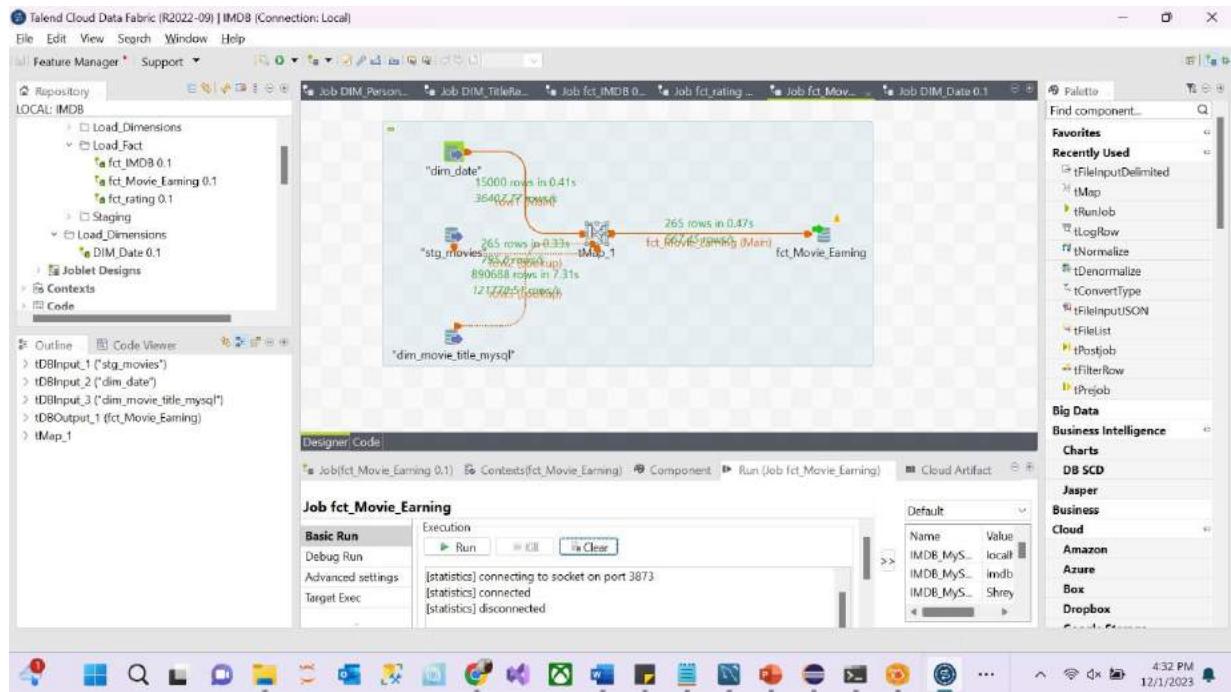
fct_imdb 1 ×

Apply

```
1 •  select COUNT(*) from fct_imdb;
```

	Count(*)
▶	4299743

Fact Movie Earnings:



The screenshot shows the MySQL Workbench interface. The left sidebar displays the database schema, including tables like "stg_imdb_title_principal", "dim_person", "dim_person_profession", etc., and the "imdb_src_movie.fct_movie_earning" table. The main area shows the result grid of a query: "SELECT * FROM imdb_src_movie.fct_movie_earning;". The results are as follows:

	MovieEarningsSK	MovieTitle_SK	Rank	Gross	Theaters	Per_Theater	Total_Gross	Days	DL_CreatedDate	JobID
2909	19971219	223874	1	\$8,658,814	2,674	\$3,238	\$8,658,814	1	2023-12-01 16:32:23	EuZaBZ
2910	19971220	223874	1	\$10,672,013	2,674	\$3,991	\$19,330,827	2	2023-12-01 16:32:23	EuZaBZ
2911	19971221	223874	1	\$9,307,304	2,674	\$3,481	\$28,638,131	3	2023-12-01 16:32:23	EuZaBZ
2912	19971222	223874	1	\$5,578,212	2,674	\$2,086	\$54,216,343	4	2023-12-01 16:32:23	EuZaBZ
2913	19971223	223874	1	\$6,003,119	2,674	\$2,245	\$40,219,462	5	2023-12-01 16:32:23	EuZaBZ
2914	19971224	223874	1	\$3,571,345	2,674	\$1,336	\$43,790,807	6	2023-12-01 16:32:23	EuZaBZ
2915	19971225	223874	1	\$9,178,529	2,674	\$3,433	\$52,969,336	7	2023-12-01 16:32:23	EuZaBZ
2916	19971226	223874	1	\$12,122,298	2,711	\$4,472	\$85,091,634	8	2023-12-01 16:32:23	EuZaBZ
2917	19971227	223874	1	\$11,464,166	2,711	\$4,200	\$11,464,166	9	2023-12-01 16:32:23	EuZaBZ

The bottom status bar shows the connection details: Local instance MySQL80 imdb_src_x, Local instance MySQL80 (imdb_src_x), and the timestamp 12/1/2023 4:32 PM.

MySQL Workbench

Local Instance MySQL80 (imdb_src) Local instance MySQL80 (imdb_src)

File Edit View Query Database Server Tools Scripting Help

Navigator: `sig_imdb_title_principal dim_person dim_person_profession dim_person_knowfor dim_person_knowfor dim_littleregion fct_rating dim_date fact_movie_earning`

SCHEMAS: `imdb_movies imdb_src_movie`

Tables: `dim_crew dim_date dim_genre dim_movie_title dim_movie_title_myself dim_person dim_person_knowfor dim_person_profession dim_littleregion fct_rating imdb_title_ratings imdb_title_principals`

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

265

No object selected

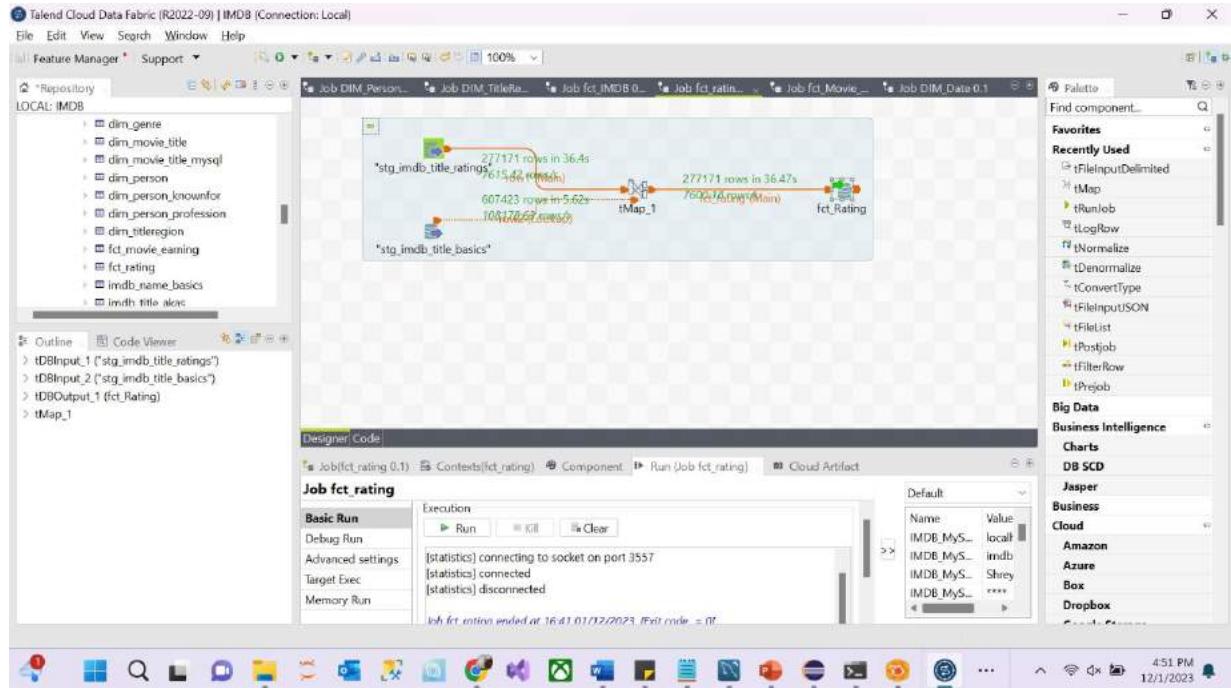
Result: 3 × Output: Action Output

Time	Action	Message	Duration / Fetch
980 16:32:29	SELECT * FROM imdb_src_movie.fct_movie_earning LIMIT 0, 500	265 row(s) returned	0.000 sec / 0.000 sec
981 16:32:57	SELECT count(*) FROM imdb_src_movie.fct_movie_earning LIMIT 0, 500	1 row(s) returned	0.000 sec / 0.000 sec

Object Info Session

432 PM 12/1/2023

Fact Rating:



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Local instance MySQL80 (imdb_src) Local instance MySQL80 (imdb_...)

Navigator: als imbd_title_principals stg_imdb_title_principal dim_person dim_person_profession dim_person_knownfor dim_person_knownfor dim_person_knownfor dim_titlerregion fct_rating

SCHEMAS

Administration Schemas Information

Table: fct_rating

Columns:

RatingSK	tconst	averageRating	numVotes	runtimeMinutes	DT_CreatedDate	JobID
1	tt00000502	4	14	100	2023-12-01 16:14:51	aF5Wxx
2	tt00000574	6	759	70	2023-12-01 16:14:51	aF5Wxx
3	tt00000591	5	17	90	2023-12-01 16:14:51	aF5Wxx
4	tt00000615	4	23	100	2023-12-01 16:14:51	aF5Wxx
5	tt00000630	4	24	100	2023-12-01 16:14:51	aF5Wxx
6	tt00000675	5	19	100	2023-12-01 16:14:51	aF5Wxx
7	tt00000679	5	65	120	2023-12-01 16:14:51	aF5Wxx
8	tt00000793	5	20	100	2023-12-01 16:14:51	aF5Wxx
9	tt00000847	5	14	100	2023-12-01 16:14:51	aF5Wxx

Action Output

Time	Action	Message	Duration / Fetch
973 15:53:51	SELECT count(*) FROM imbd_src_movie.dim_titlerregion LIMIT 0, 500	1 row(s) returned	0.657 sec / 0.000 sec
974 16:25:08	SELECT * FROM imbd_src_movie.fct_rating LIMIT 0, 500	500 row(s) returned	0.000 sec / 0.000 sec

Result Grid Form Editor

Output

Object Info Session

4:25 PM 12/1/2023

MySQL Workbench

Local Instance MySQL80 (imdb_sr) Local instance MySQL80 (imdb_sr)

File Edit View Query Database Server Tools Scripting Help

Navigator: als imbd_title_principals stg_imdb_title_principal dim_person dim_person_profession dim_person_knownfor dim_person_knownfor dim_titterregion lct_rating

SCHEMAS: Filter objects: dim_person_profession dim_titterregion fc_rating imbd_title_ratings imbd_title_principals imbd_title_crew imbd_title_basics imbd_name_basics stg_imdb_name_basics stg_imdb_title_aliases stg_imdb_title_basics stg_imdb_title_crew stg_imdb_title_principals stg_imdb_title_ratings

1 • SELECT count(*) FROM imdb_src_movie.fct_rating;

Execute the selected portion of the script or everything, if there is no selection.

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

count(*)
277171

Administration Schemas Information

Table: fct_rating

Columns:

RatingSK	int PK
content	varchar(double;
averageRating	double;
numVotes	int;
runtimeMinutes	varchar(datetime;
DT_CreatedDate	

Object Info Session

Action Output: Read Only

Time Action Message Duration / Fetch

974 16:25:08 SELECT * FROM imdb_src_movie.fct_rating LIMIT 0, 500 500 row(s) returned 0.000 sec / 0.000 sec

975 16:25:26 SELECT count(*) FROM imdb_src_movie.fct_rating LIMIT 0, 500 1 row(s) returned 0.032 sec / 0.000 sec

4:25 PM 12/1/2023

SQL Queries -

Trend Analysis:

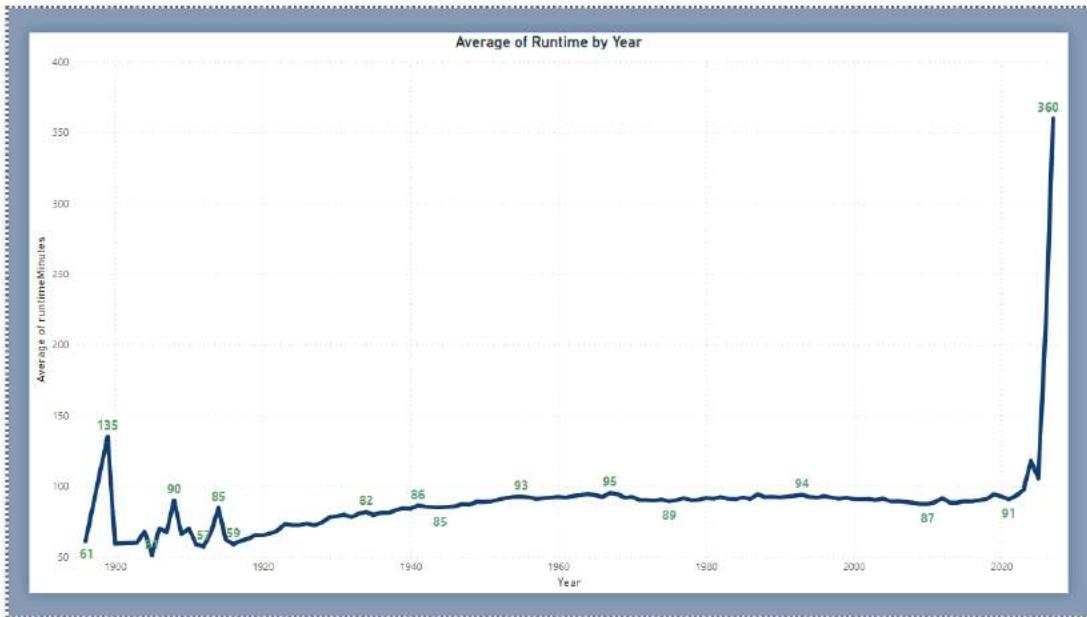
1. Average movie runtime change over the years:
 - a. For every year what is the average run-time and plot it on the graph.

```
select startYr, avg(runtimeMinutes) as AvgRuntime from dim_movie_title_mysql
group by startYr
order by startYr;
```

```
2 • select startYr, avg(runtimeMinutes) as AvgRuntime from dim_movie_title_mysql
3   group by startYr
4   order by startYr;
5
```

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

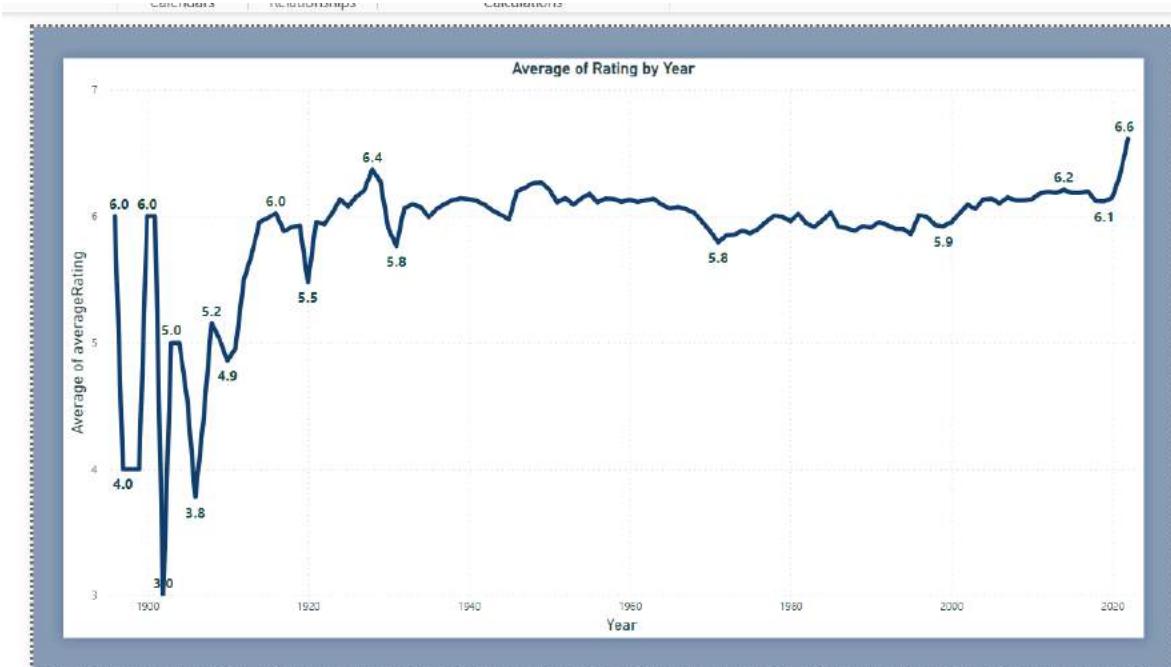
startYear	AvgRuntime
1896-01-01 00:00:00	61
1897-01-01 00:00:00	NULL
1898-01-01 00:00:00	NULL
1899-01-01 00:00:00	135
1900-01-01 00:00:00	59.5
1901-01-01 00:00:00	NULL
1902-01-01 00:00:00	NULL
1903-01-01 00:00:00	60
1904-01-01 00:00:00	68
1905-01-01 00:00:00	51
1906-01-01 00:00:00	70
1907-01-01 00:00:00	67.5
1908-01-01 00:00:00	90
1909-01-01 00:00:00	66.25
1910-01-01 00:00:00	69.9333333333334
1911-01-01 00:00:00	58.91304347826087
1912-01-01 00:00:00	57.285714285714285
1913-01-01 00:00:00	67.33333333333333
1914-01-01 00:00:00	84.84065934065934



2. Is there a correlation between the average movie rating and the year of release?
 a. The goal is to identify in which year we had good movies.

```
SELECT AVG(a.averageRating) AS avgRating, s.startYear
FROM fct_rating a
JOIN dim_movie_title_mysql s ON a.tconst = s.tconst
GROUP BY s.startYear
ORDER BY s.startYear;
```

avgRating	startYear
6.0000	1896-01-01 00:00:00
4.0000	1897-01-01 00:00:00
4.0000	1899-01-01 00:00:00
6.0000	1900-01-01 00:00:00
6.0000	1901-01-01 00:00:00
3.0000	1902-01-01 00:00:00
5.0000	1903-01-01 00:00:00
5.0000	1904-01-01 00:00:00
4.5556	1905-01-01 00:00:00
3.7778	1906-01-01 00:00:00
4.3846	1907-01-01 00:00:00
5.1538	1908-01-01 00:00:00
5.0286	1909-01-01 00:00:00
4.8571	1910-01-01 00:00:00
4.9565	1911-01-01 00:00:00
5.4884	1912-01-01 00:00:00
5.7011	1913-01-01 00:00:00
5.9532	1914-01-01 00:00:00
5.9854	1915-01-01 00:00:00
...	1916-01-01 00:00:00



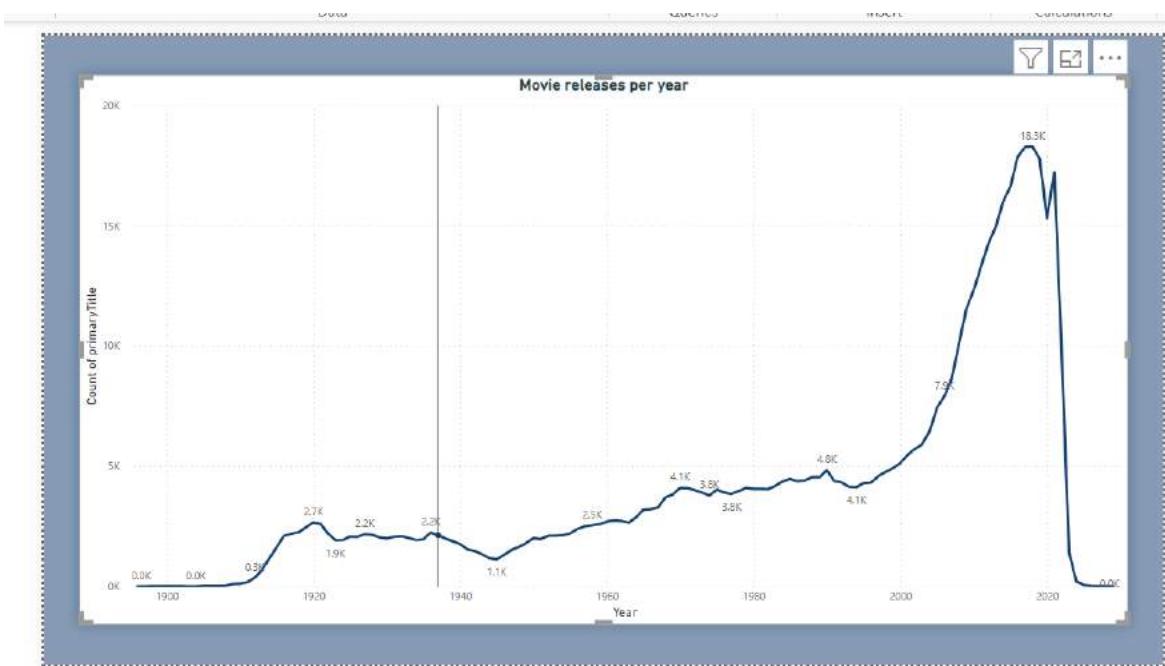
3. Visualization of the distribution of movie releases per year
a. No. of movies released every year

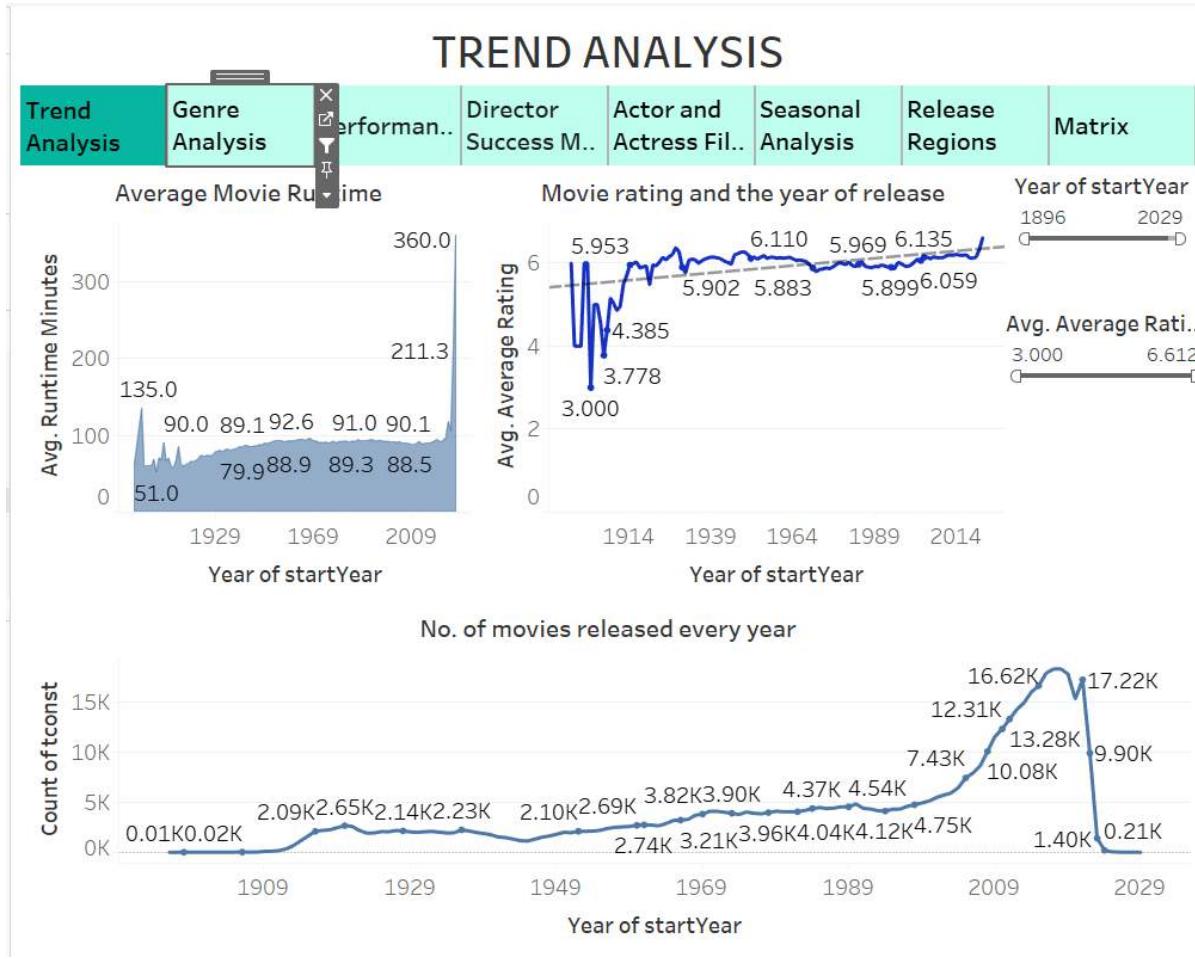
```
SELECT COUNT(tconst) AS movieCount, startYear  
FROM dim_movie_title_mysql  
GROUP BY startYear;
```

```
11
12 •  SELECT COUNT(tconst) AS movieCount, startYear
13   FROM dim_movie_title_mysql
14   GROUP BY startYear;
15
```

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

	movieCount	startYear
▶	18	1905-01-01 00:00:00
	23	1906-01-01 00:00:00
	15	1907-01-01 00:00:00
	21	1908-01-01 00:00:00
	92	1909-01-01 00:00:00
	105	1910-01-01 00:00:00
	347	1912-01-01 00:00:00
	165	1911-01-01 00:00:00
	673	1913-01-01 00:00:00
	1633	1915-01-01 00:00:00
	1156	1914-01-01 00:00:00
	2456	1919-01-01 00:00:00
	2093	1916-01-01 00:00:00
	2188	1917-01-01 00:00:00
	2231	1936-01-01 00:00:00
	2064	1925-01-01 00:00:00
	2244	1918-01-01 00:00:00
	2652	1920-01-01 00:00:00
	2192	1922-01-01 00:00:00





Genre Analysis:

1. Which genres have seen the most significant ratings in popularity over the past decade?
 a. The goal is to identify the popularity of the genres.

```

SELECT
  g.genres,avg(r.averageRating) as avg_rating
FROM
  dim_genre g
JOIN
  dim_movie_title_mysql t ON g.tconst = t.tconst
JOIN
  fct_imdb i ON t.MovieTitle_SK = i.MovieTitle_SK
JOIN
  fct_rating r ON r.RatingSK = i.RatingSK
WHERE
  t.startYear BETWEEN '2014-01-01' AND '2023-12-31'
  
```

```

GROUP BY g.genres
ORDER BY
  avg_rating DESC;
  
```

56

57 • SELECT

58 g.genres,avg(r.averageRating) as avg_rating

59 FROM

60 dim_genre g

61 JOIN

62 dim_movie_title_mysql t ON g.tconst = t.tconst

63 JOIN

64 fct_imdb i ON t.MovieTitle_SK = i.MovieTitle_SK

65 JOIN

66 fct_rating r ON r.RatingSK = i.RatingSK

67 WHERE

68 t.startYear BETWEEN '2014-01-01' AND '2023-12-31'

69

70 GROUP BY g.genres

71 ORDER BY

72 avg_rating DESC;

73

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

genres	avg_rating
Talk-Show	7.5714
News	7.2368
Short	7.2000
Documentary	7.1882
Biography	6.9849
Music	6.9549
History	6.8990
Sport	6.7898
Game-Show	6.5000
Musical	6.4798
...	6.4755

Result 24 ×

Output

Average of averageRating by genres

Genre	Average Rating
Talk-Show	7.38
News	7.24
Short	7.20
Documentary	7.19
Biography	6.99
Music	6.95
History	6.90
Sport	6.79
Game-Show	6.50
Musical	6.48
War	6.44
Animation	6.25
Drama	6.22
Family	6.12
Romance	6.02
Adventure	5.97
Crime	5.96
Reality-TV	5.85
Adult	5.82
Comedy	5.81
Fantasy	5.76
Action	5.70
Mystery	5.63
Western	5.54
Thriller	5.43
Sci-Fi	5.26
Horror	4.84

2. Show the top 5 genres as compared to gross earnings for the 9 box office movies.

```
SELECT
g.genres,
SUM(e.Gross) AS total_earnings
FROM
```

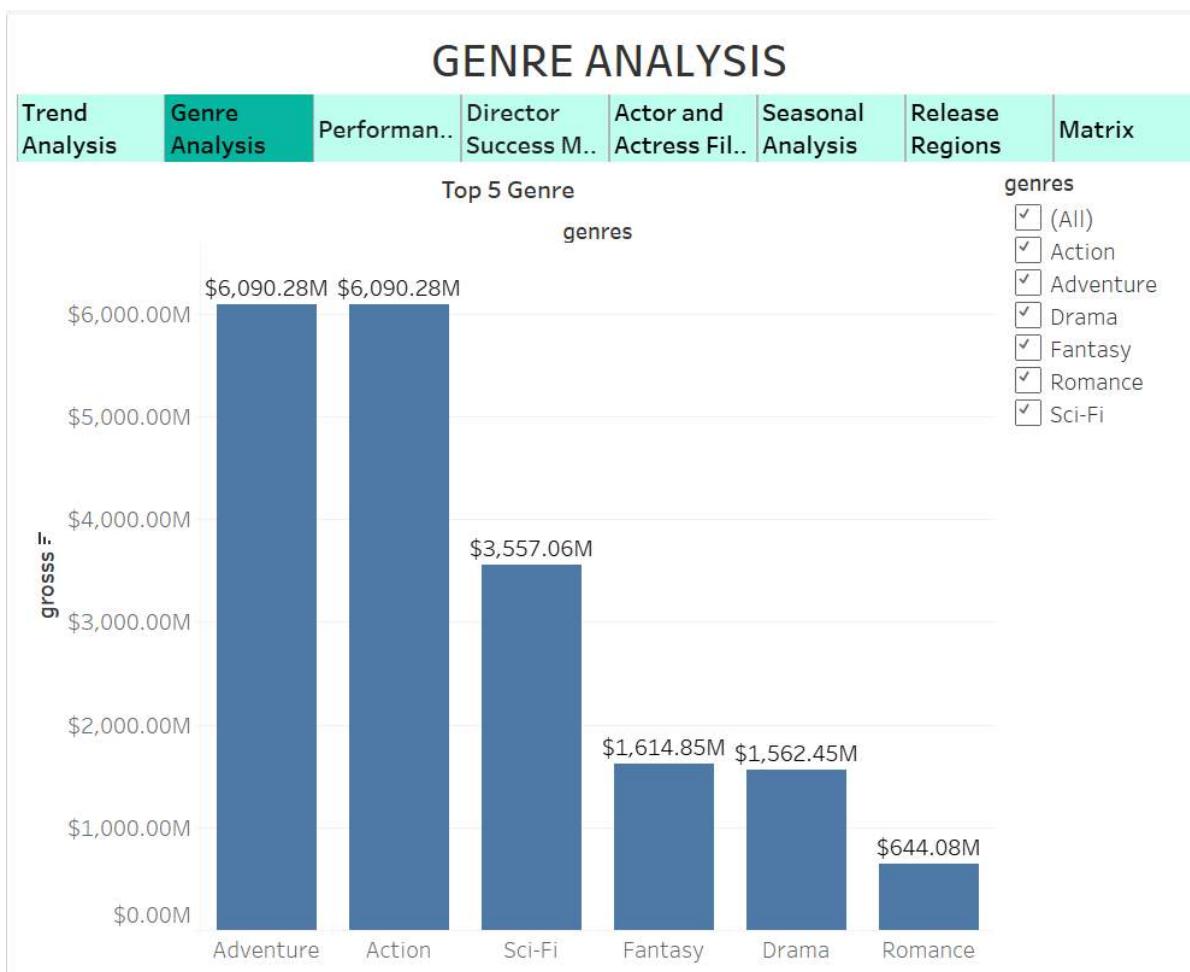
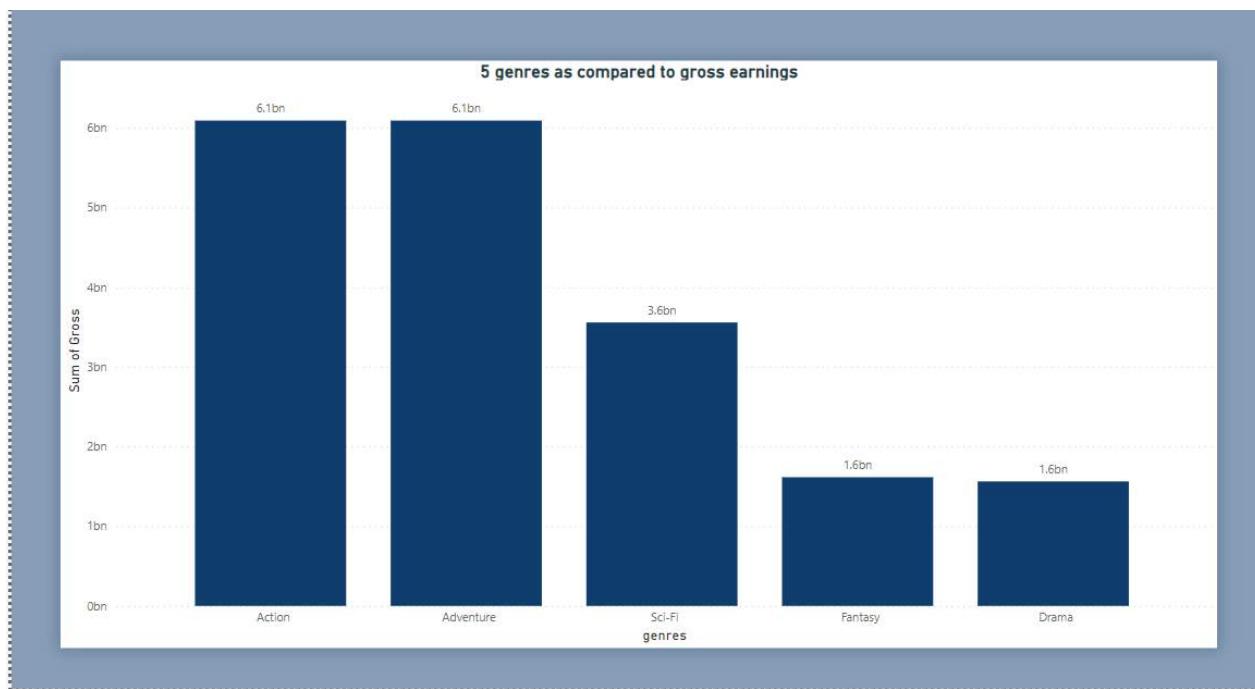
```

fct_movie_earning e
JOIN
    dim_movie_title_mysql m ON e.tconst = m.tconst
JOIN
    dim_genre g ON e.tconst = g.tconst
GROUP BY
    g.genres
ORDER BY
    total_earnings DESC
LIMIT 5;

```

The screenshot shows a MySQL query editor interface. At the top, there is a toolbar with various icons for file operations, search, and navigation. Below the toolbar, the SQL query is displayed in a code editor with line numbers from 43 to 54. The code uses standard SQL syntax to join three tables (fct_movie_earning, dim_movie_title_mysql, and dim_genre) based on their primary keys (tconst). It groups the results by genre and orders them by total earnings in descending order, limiting the output to the top 5 entries.

	genres	total_earnings
▶	Action	6090283475
	Adventure	6090283475
	Sci-Fi	3557064362
	Fantasy	1614846113
	Drama	1562451638



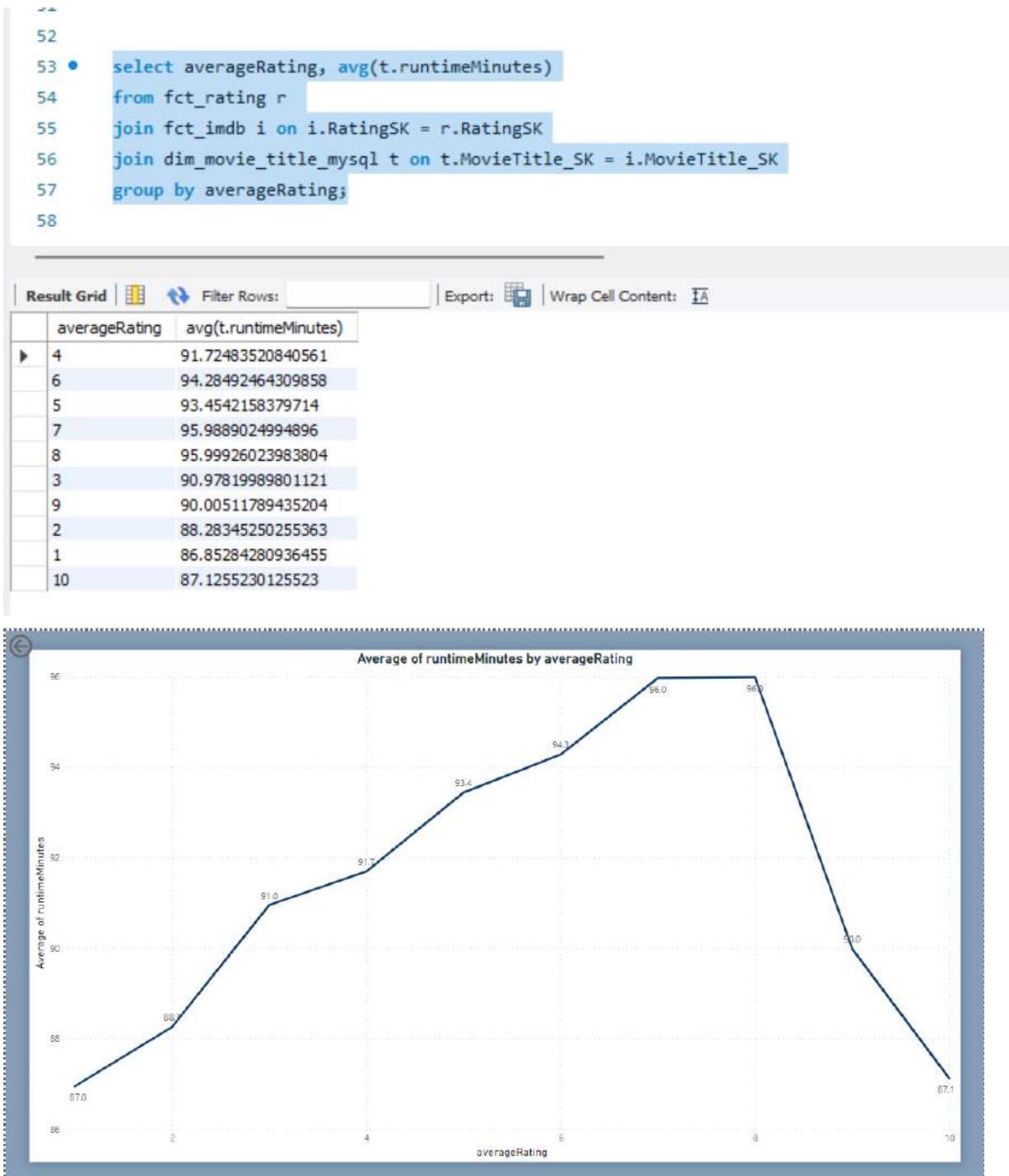
Performance Metrics:

1. Correlation between the movie's runtime and its average rating
 - a. The goal is to understand the impact on the ratings due to the movie's runtime.

```

select averageRating, avg(t.runtimeMinutes)
from fct_rating r
join fct_imdb i on i.RatingSK = r.RatingSK
join dim_movie_title_mysql t on t.MovieTitle_SK = i.MovieTitle_SK
group by averageRating;

```



2. Correlation between the movie's runtime and its average gross

- a. The goal is to understand the impact on the gross earnings due to the movie's runtime.

```

select primaryTitle, sum(Gross), runTimeMinutes
from dim_movie_title_mysql t

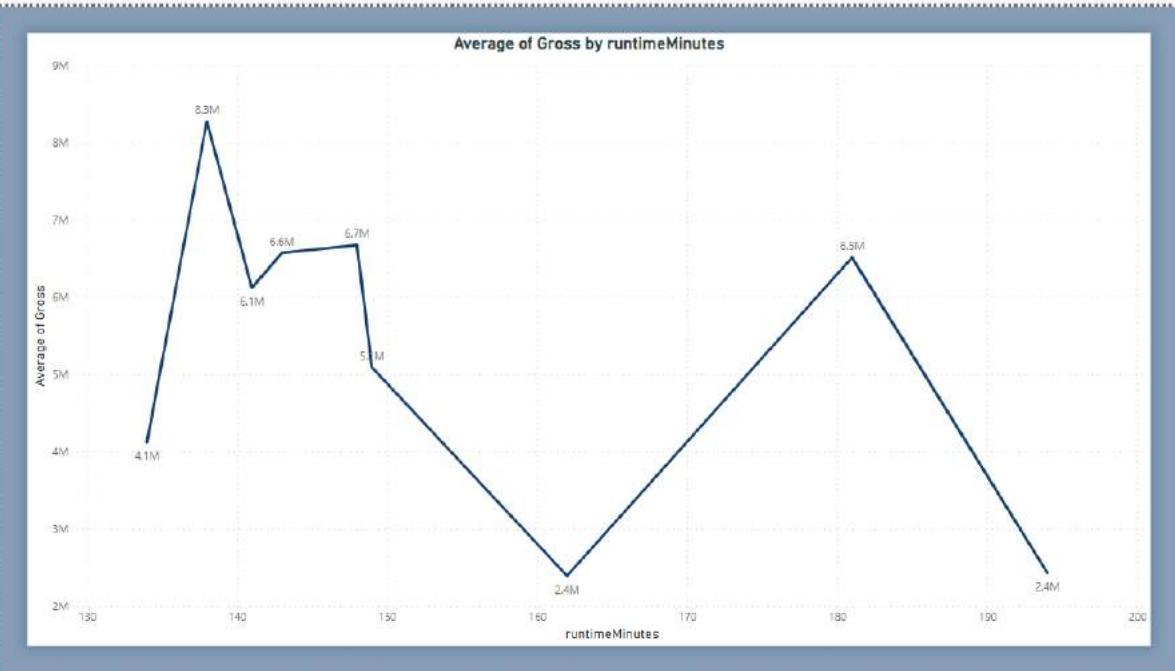
```

```
join fct_movie_earning e on t.MovieTitle_SK = e.MovieTitle_SK  
group by primaryTitle,runTimeMinutes;
```

```
52  
53 • select primaryTitle, sum(Gross),runTimeMinutes  
54 from dim_movie_title_mysql t  
55 join fct_movie_earning e on t.MovieTitle_SK = e.MovieTitle_SK  
56 group by primaryTitle,runTimeMinutes;  
57
```

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

primaryTitle	sum(Gross)	runTimeMinutes
Avatar	760508625	162
Avengers: Age of Ultron	483598418	141
Avengers: Endgame	918373000	181
Avengers: Infinity War	717815482	149
Black Panther	725259566	134
Spider-Man: No Way Home	854337488	148
Star Wars: Episode VII - The Force Awakens	992642689	138
The Avengers	637748207	143
Titanic	644078638	194



3. Relationship between the number of votes and the average rating of movies.

a. The goal is to identify the correlation between votes and rating

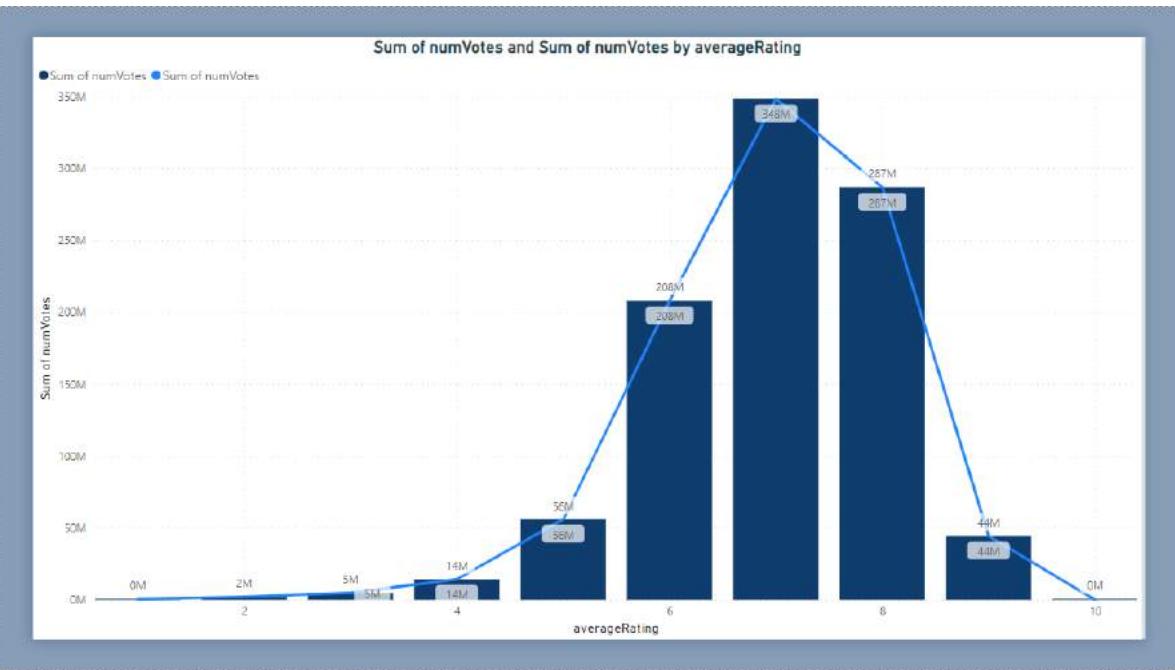
```
select averageRating,sum(numVotes) from fct_rating
```

```
group by averageRating;
```

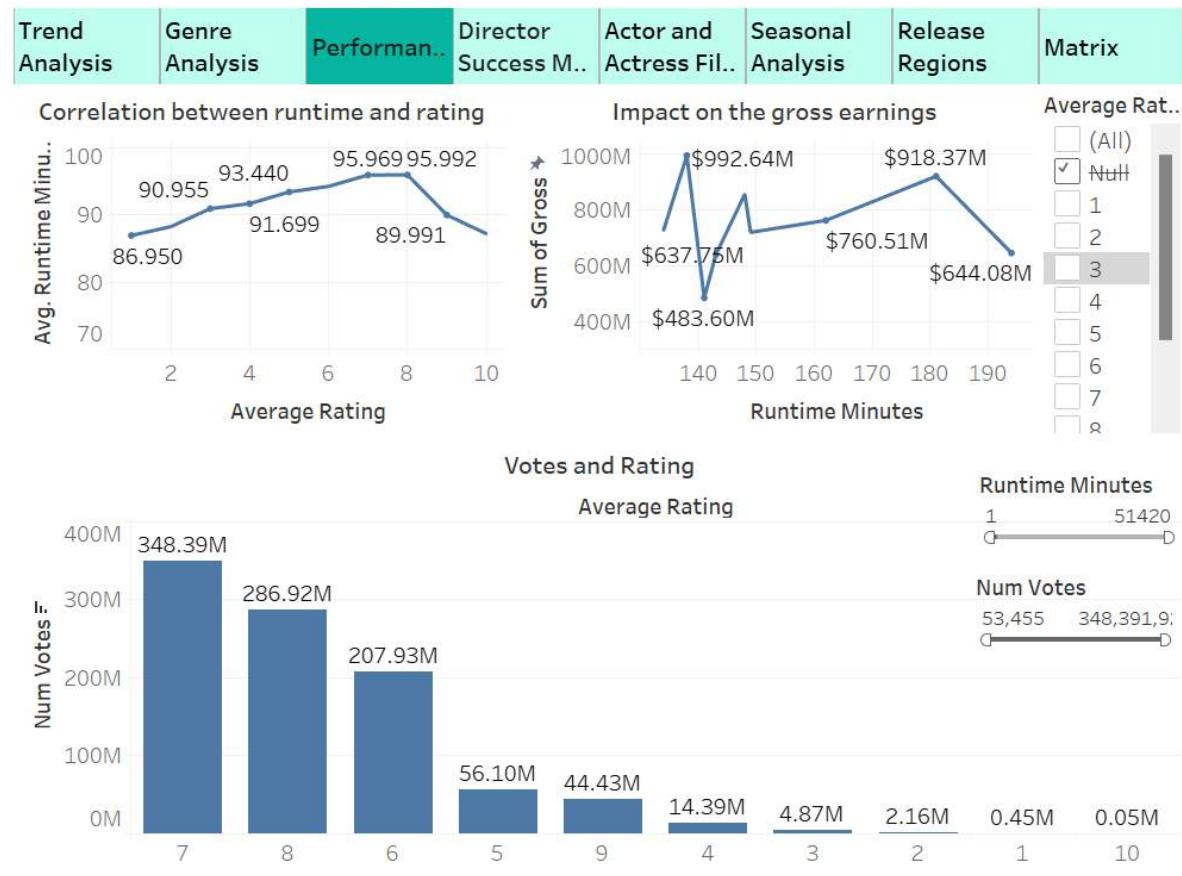
```
58 • select averageRating,sum(numVotes) from fct_rating  
59 group by averageRating;  
60
```

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

	averageRating	sum(numVotes)
▶	4	14389566
	6	207928332
	5	56097207
	7	348392088
	8	286926542
	3	4872904
	9	44427452
	2	2159728
	1	451294
	10	53463



PERFORMANCE METRICS

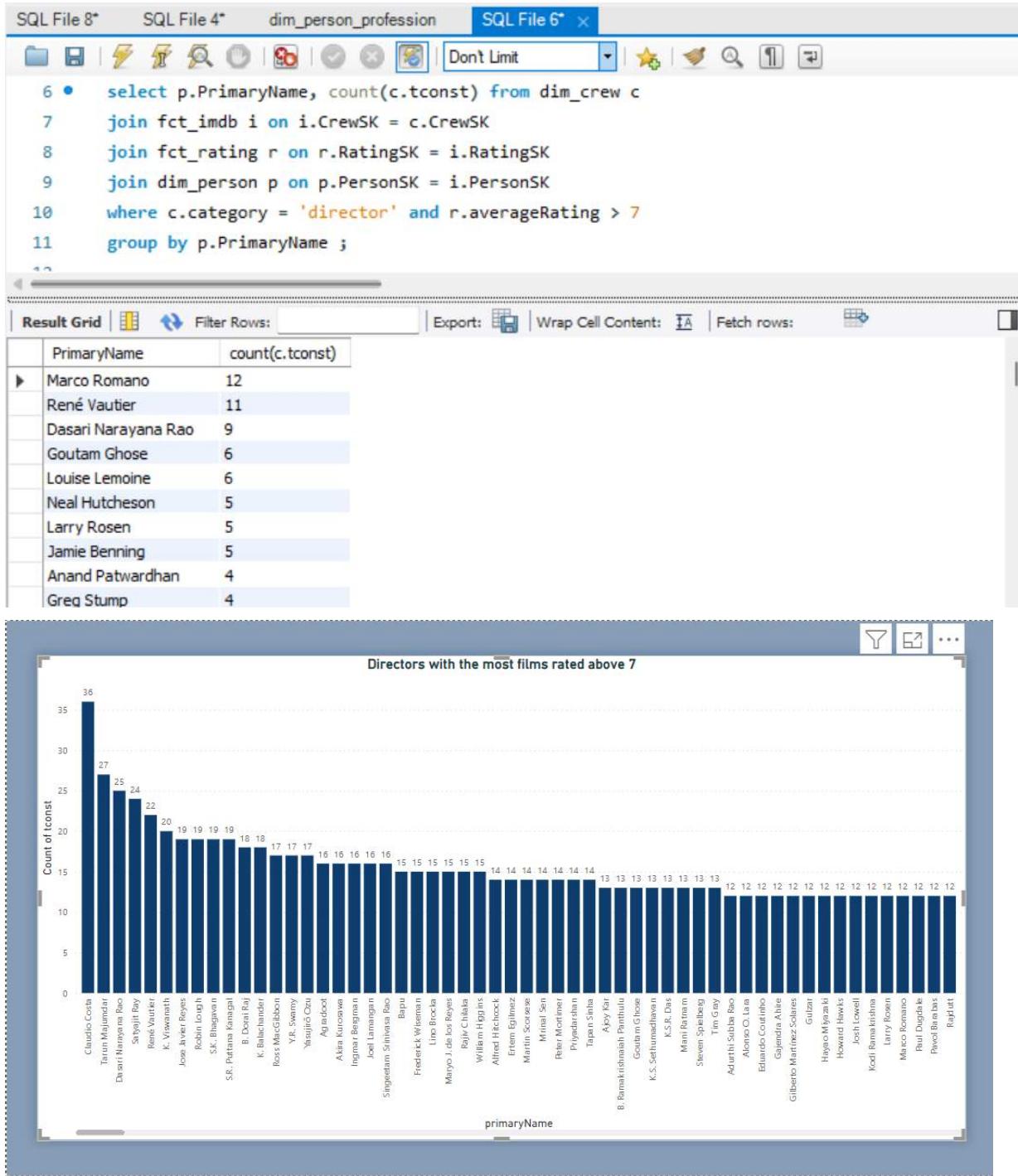


Director Success Metrics:

1. Identify directors with the most films rated above 7. Sort them in descending order
 - a. The goal is to identify which directors have made films that were received well by the audience

```

select p.PrimaryName, count(c.tconst) from dim_crew c
join fct_imdb i on i.CrewSK = c.CrewSK
join fct_rating r on r.RatingSK = i.RatingSK
join dim_person p on p.PersonSK = i.PersonSK
where c.category = 'director' and r.averageRating > 7
group by p.PrimaryName ;
  
```



2. Determine the directors who have directed the highest number of films overall and their respective gross earnings trends.
- a. The goal is to identify if there are any notable directors who consistently contribute to high-grossing movies

```

SELECT
  COUNT(DISTINCT(e.tconst)) AS movie_count,
  p.PrimaryName,
  c.category,
  SUM(e.Gross) AS total_gross_earnings
FROM
  dim_crew c

```

```

JOIN
    dim_person p ON p.nconst = c.nconst
JOIN
    fct_movie_earning e ON e.tconst = c.tconst
WHERE
    c.category = 'director'
GROUP BY
    p.PrimaryName, c.category
ORDER BY
    total_gross_earnings DESC;

```

Screenshot of a database query interface showing the results of the provided SQL query.

The interface includes a toolbar with various icons for file operations, search, and navigation. The query editor shows the following code:

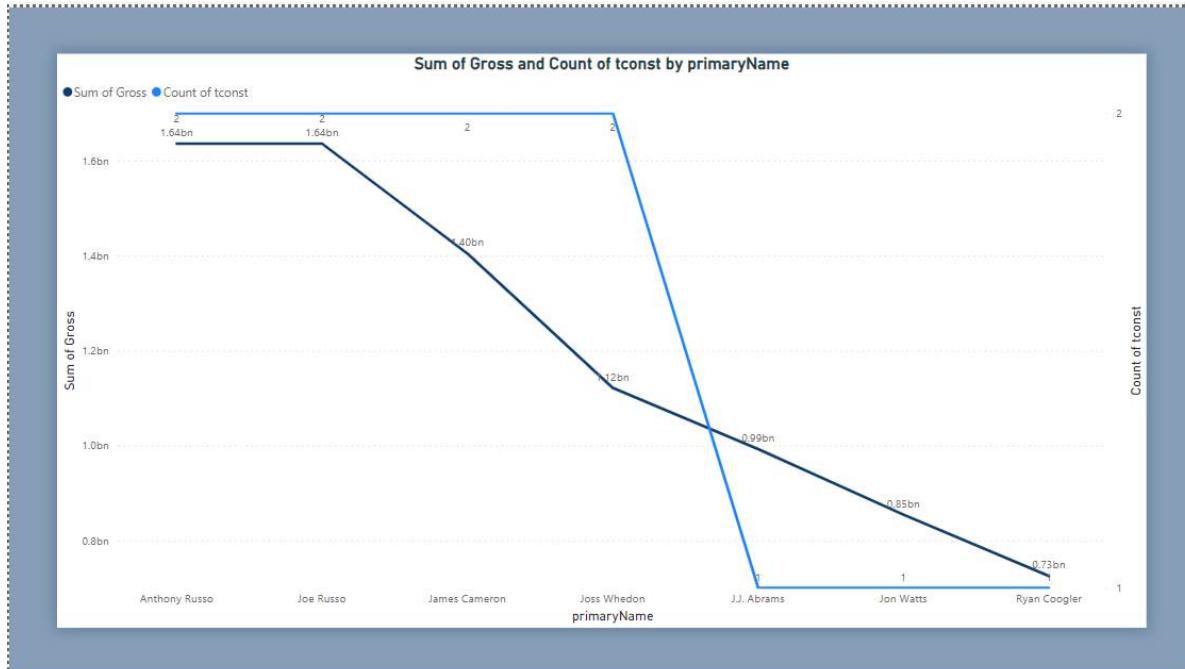
```

43 • SELECT
44     COUNT(DISTINCT(e.tconst)) AS movie_count,
45     p.PrimaryName,
46     c.category,
47     SUM(e.Gross) AS total_gross_earnings
48 FROM
49     dim_crew c
50 JOIN
51     dim_person p ON p.nconst = c.nconst
52 JOIN
53     fct_movie_earning e ON e.tconst = c.tconst
54 WHERE
55     c.category = 'director'

```

The results grid displays the following data:

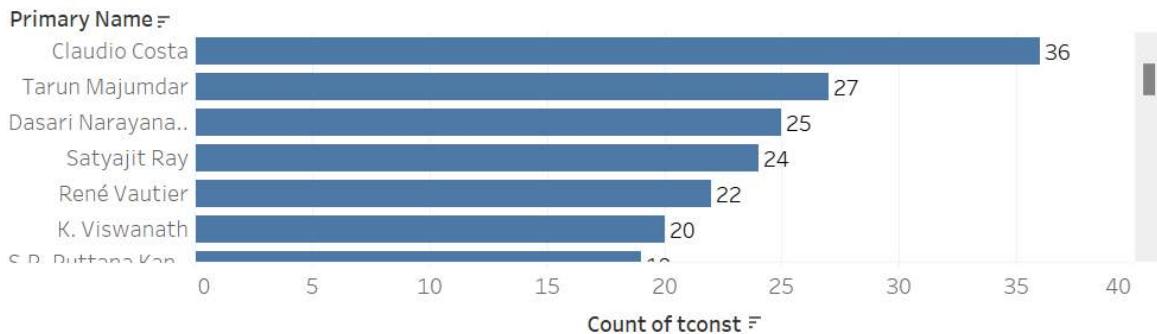
	movie_count	PrimaryName	category	total_gross_earnings
▶	2	Anthony Russo	director	1636188482
	2	Joe Russo	director	1636188482
	2	James Cameron	director	1404587263
	2	Joss Whedon	director	1121346625
	1	J.J. Abrams	director	992642689
	1	Jon Watts	director	854337488



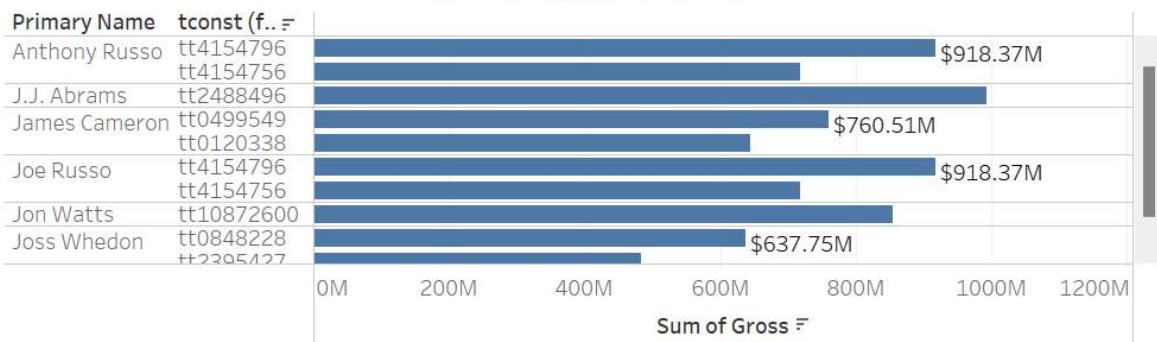
DIRECTOR SUCCESS METRICS

Trend Analysis	Genre Analysis	Performance Analysis	Director Success Metrics	Actor and Actress Film Records	Seasonal Analysis	Release Regions	Matrix
----------------	----------------	----------------------	--------------------------	--------------------------------	-------------------	-----------------	--------

Directors with the most rated films



Directors with high grossing movies



Actor and Actress Film Records:

1. List the top 10 actors/actresses with the most films rated between 4 and 7.

a. The goal is to identify the popular actors/actresses

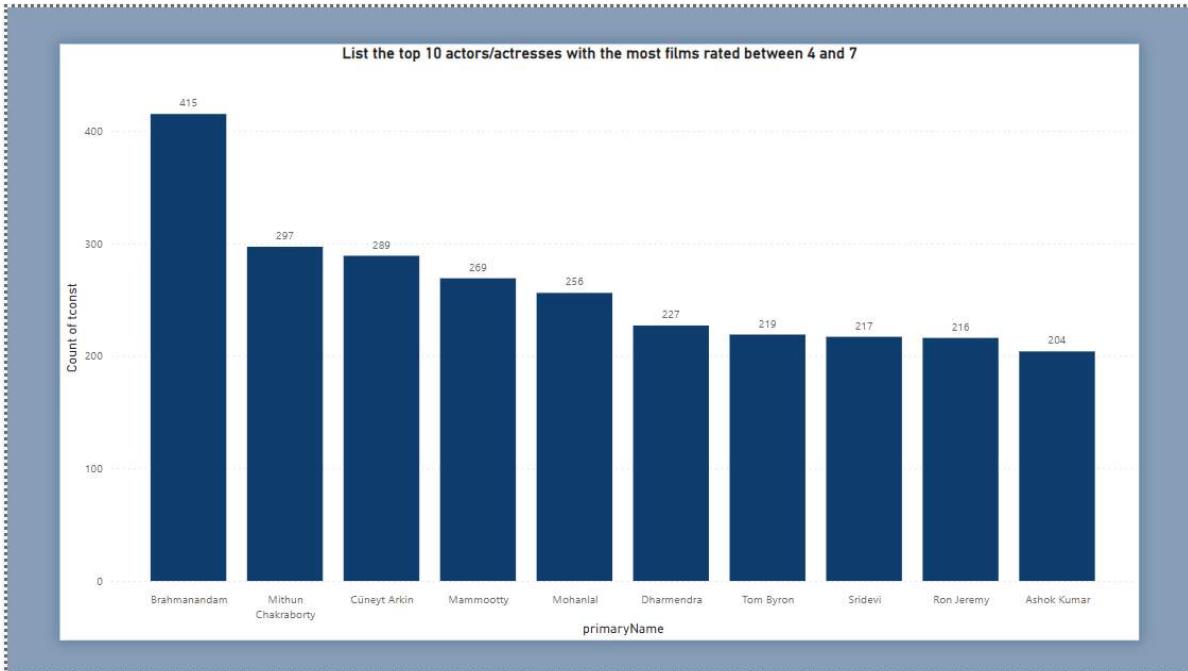
SELECT

```
da.primaryName AS actor_name,
COUNT(DISTINCT dc.tconst) AS total_films_between_4_and_7
FROM
dim_person da
JOIN
dim_crew dc ON da.nconst = dc.nconst
JOIN
fct_rating fr ON dc.tconst = fr.tconst
WHERE
fr.averageRating BETWEEN 4 AND 7 and dc.category = "actor" or dc.category = "actress"
GROUP BY
da.primaryName
ORDER BY
total_films_between_4_and_7 DESC
Limit 10;
```

The screenshot shows a database query editor with the following interface elements:

- Toolbar:** Includes icons for file operations (New, Open, Save, Print, Copy, Paste), search, and various database functions.
- Query Editor:** Displays the SQL query numbered from 1 to 10. The WHERE clause is currently collapsed.
- Result Grid:** A table showing the results of the query. It has two columns: "actor_name" and "total_films_between_4_and_7".
- Grid Headers:** "Result Grid", "Filter Rows:", "Export:", and "Wrap Cell Content:".
- Data:** The result grid contains the following data:

actor_name	total_films_between_4_and_7
Brahmanandam	415
Mithun Chakraborty	297
Cüneyt Arkin	289
Mammootty	269
Mohanlal	256
Sridevi	232
Dharmendra	227
Tom Byron	219
Ron Jeremy	216
Aachi Manorama	213



2. Compare the top 5 actors and actresses based on movie ratings.

a. The goal is to compare the top 5 actor and actresses

```
select primaryName,avg(r.averageRating) from dim_person_profession pp
join dim_person p on p.nconst = pp.nconst
join fct_imdb i on i.PersonSK = p.PersonSK
join dim_crew c on c.CrewSK = i.CrewSK
join fct_rating r on r.RatingSK = i.RatingSK
where c.category IN ('actor','actress')
group by primaryName;
```

```

112
113 -- Compare the top 5 actors and actresses based on movie ratings
114 • select primaryName,avg(r.averageRating) from dim_person_profession pp
115 join dim_person p on p.nconst = pp.nconst
116 join fct_imdb i on i.PersonSK = p.PersonSK
117 join dim_crew c on c.CrewSK = i.CrewSK
118 join fct_rating r on r.RatingSK = i.RatingSK
119 where c.category IN ('actor','actress')
120 group by primaryName;
121
122 • SELECT
123

```

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

primaryName	avg(r.averageRating)
Srdjan Jablanovic	10.0000
Pierre Mirel	10.0000
Paul Roraback	10.0000
Donovan Bailey	10.0000
Sophia Laia	10.0000
Jugoslav Jospovic	10.0000
Davi Freitas	10.0000
Orin Friesen	10.0000
Sam DeRose	10.0000
Brad	10.0000
Margareta Spănu-Cemărтан	10.0000
Hercules Nikolopoulos	10.0000
Raj Tirandasu	10.0000
Ana Blake	10.0000
Taylor Van Doorne	10.0000
Martin Nadin	10.0000
Deeksha Welekar	10.0000
Pulakesh Bhattacharya	10.0000
Ahmed Al-Ansari	10.0000

Result 3 ×

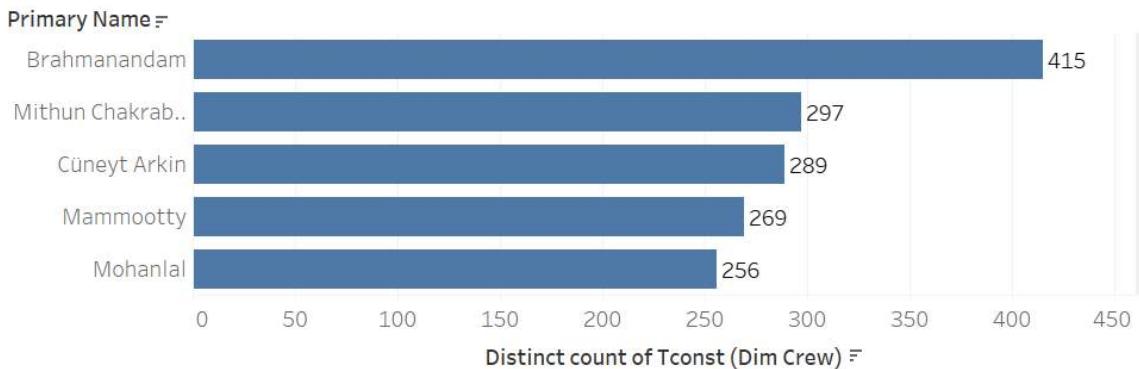
Output: Action Output

#	Time	Action	Message
13	02:54:56	select primaryName,avg(r.averageRating) from dim_person_profession pp join dim_person p on p.nconst = pp.nconst join fct_imdb i on i.Pe...	Error Code: 2013. Lost connection to MySQL server during query
14	03:06:07	select primaryName,avg(r.averageRating) from dim_person_profession pp join dim_person p on p.nconst = pp.nconst join fct_imdb i on i.Pe...	379252 row(s) returned

ACTOR AND ACTRESS FILM RECORDS

Trend Analysis	Genre Analysis	Performance	Director Success M.	Actor and Actress Fil..	Seasonal Analysis	Release Regions	Matrix
----------------	----------------	-------------	---------------------	-------------------------	-------------------	-----------------	--------

Top 10 Actors

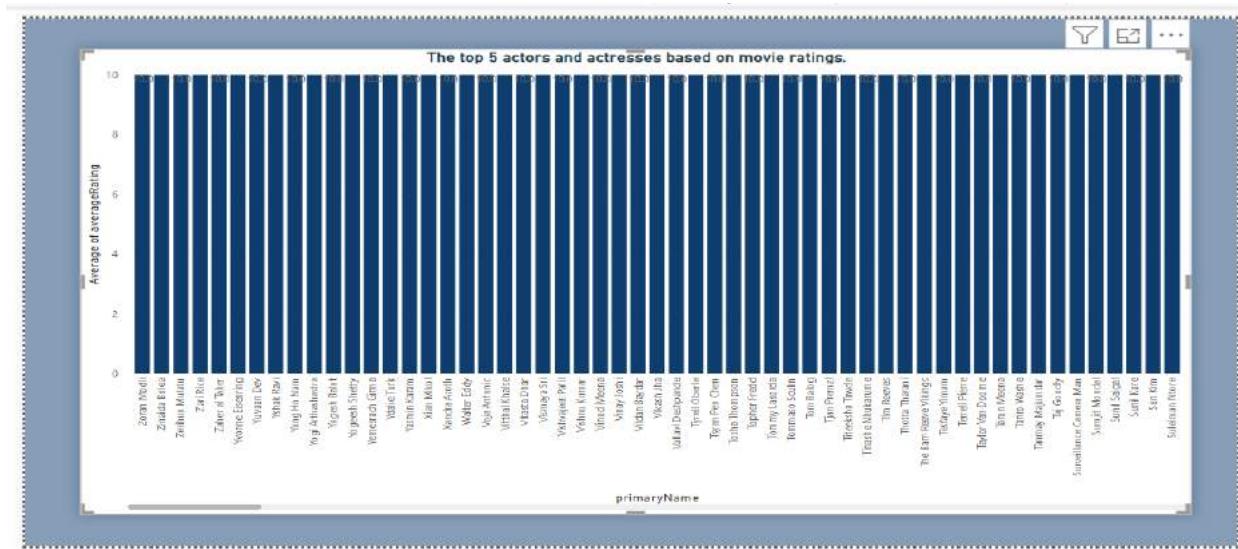


Top 5 by rating

Primary Name	Avg. ...	Category
Zoran Modli	10	(All)
Zinaida Bolea	10	Null
Zerihun Mulatu	10	actor
Zari Rice	10	actress
Zaher al Taher	10	archive_fo...
Yvonne Eisenring	10	archive_so...
Yuvaan Dev	10	cinematogr...
Yshak Ravi	10	composer
Yong Ho Nam	10	director
Yogi Arthashastra	10	editor
Yogesh Balot	10	

Category

- (All)
- Null
- actor
- actress
- archive_fo...
- archive_so...
- cinematogr...
- composer
- director
- editor



Seasonal Analysis:

1. How does movie performance vary across different seasons of the year? (based on gross earnings)
 - a. The goal is to identify which season has an impact on the revenue.

```

SELECT
CASE
    WHEN d.Month_Num IN (0, 1, 2, 3) THEN 'Spring'
    WHEN d.Month_Num IN (4, 5, 6, 7) THEN 'Summer'
    WHEN d.Month_Num IN (8, 9, 10, 11) THEN 'Fall'
END AS season,
SUM(e.Gross) AS total_gross_earnings
FROM
fct_movie_earning e
JOIN
dim_date d ON e.dateSK = d.dateSK
GROUP BY
season
ORDER BY
MIN(d.Month_Num);
  
```

299 • SELECT

300 CASE

301 WHEN d.Month_Num IN (0, 1, 2, 3) THEN 'Spring'

302 WHEN d.Month_Num IN (4, 5, 6, 7) THEN 'Summer'

303 WHEN d.Month_Num IN (8, 9, 10, 11) THEN 'Fall'

304 END AS season,

305 SUM(e.Gross) AS total_gross_earnings

306 FROM

307 fct_movie_earning e

308 JOIN

309 dim_date d ON e.dateSK = d.dateSK

310 GROUP BY

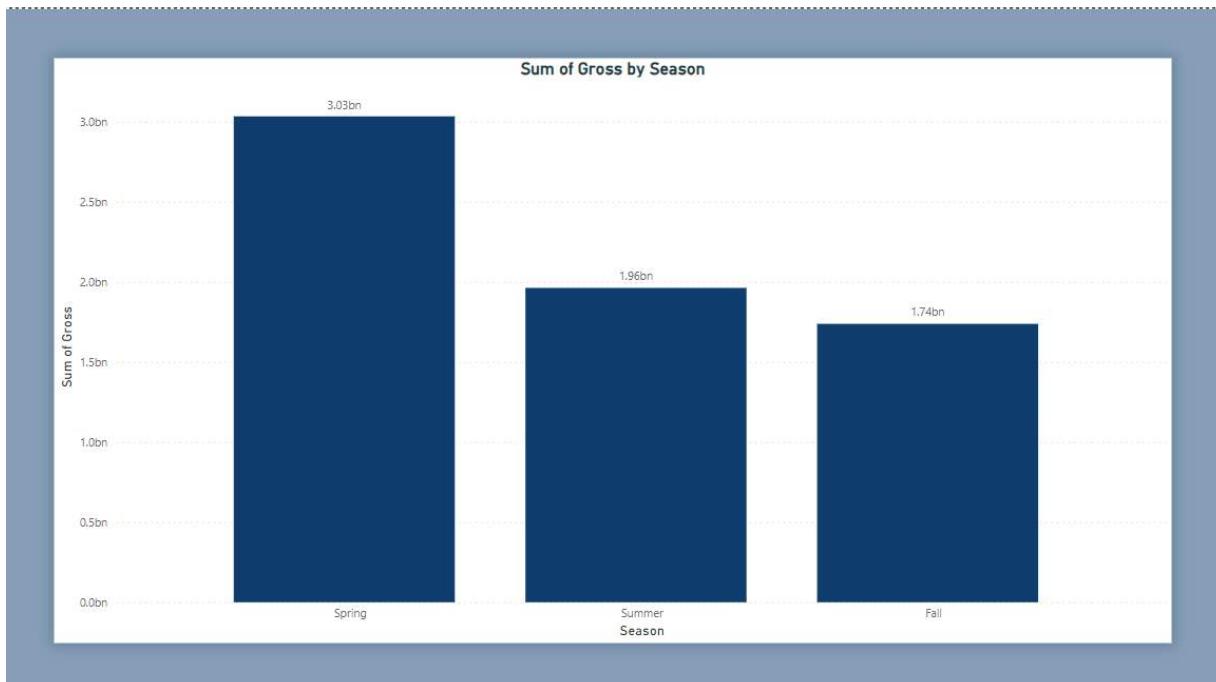
311 season

312 ORDER BY

313 MIN(d.Month_Num);

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

	season	total_gross_earnings
▶	Spring	3033650624
	Summer	1962604141
	Fall	1738107348



2. Top 3 movies based on the season (spring, summer, fall)
- Spring - Jan to April
 - Summer - May to August
 - Fall - September to December

```

SELECT
    e.primaryTitle,
    e.originalTitle,
    e.startYear,
    e.runtimeMinutes,
    e.Gross,
    e.season
FROM (
    SELECT
        t.primaryTitle,
        t.originalTitle,
        t.startYear,
        t.runtimeMinutes,
        e.Gross,
        CASE
            WHEN d.Month_Num BETWEEN 1 AND 4 THEN 'Spring'
            WHEN d.Month_Num BETWEEN 5 AND 8 THEN 'Summer'
            WHEN d.Month_Num BETWEEN 9 AND 12 THEN 'Fall'
            ELSE 'Unknown'
        END AS season,
        ROW_NUMBER() OVER (PARTITION BY
            CASE
                WHEN d.Month_Num BETWEEN 1 AND 4 THEN 'Spring'
                WHEN d.Month_Num BETWEEN 5 AND 8 THEN 'Summer'
                WHEN d.Month_Num BETWEEN 9 AND 12 THEN 'Fall'
                ELSE 'Unknown'
            END
            ORDER BY e.Gross DESC) AS season_rank
    FROM
        fct_movie_earning e
    JOIN
        dim_date d ON e.dateSK = d.dateSK
    JOIN
        dim_movie_title_mysql t ON e.tconst = t.tconst
    WHERE
        d.Month_Num BETWEEN 1 AND 12
) e
WHERE
    e.season_rank <= 3
ORDER BY
    e.season, e.Gross DESC;

```

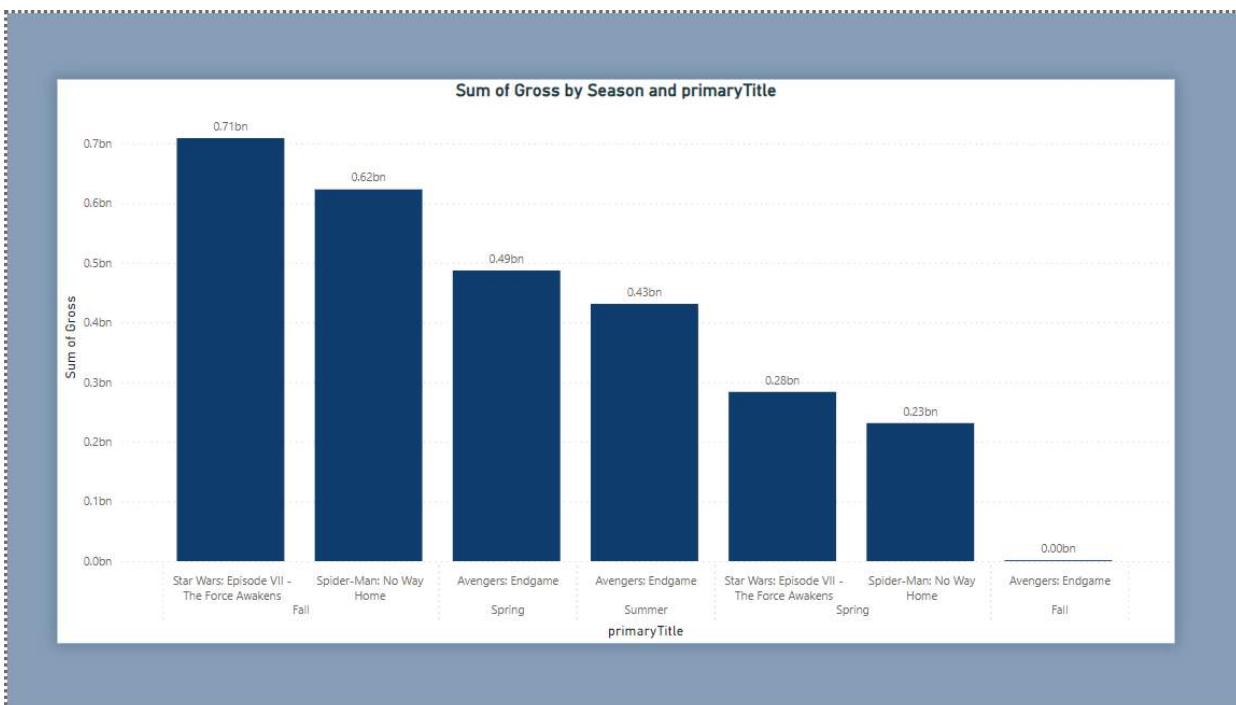
```

307 WHERE
308   d.Month_Num BETWEEN 1 AND 12
309 ) e
310 WHERE
311   e.season_rank <= 3
312 ORDER BY
313   e.season, e.Gross DESC;
314

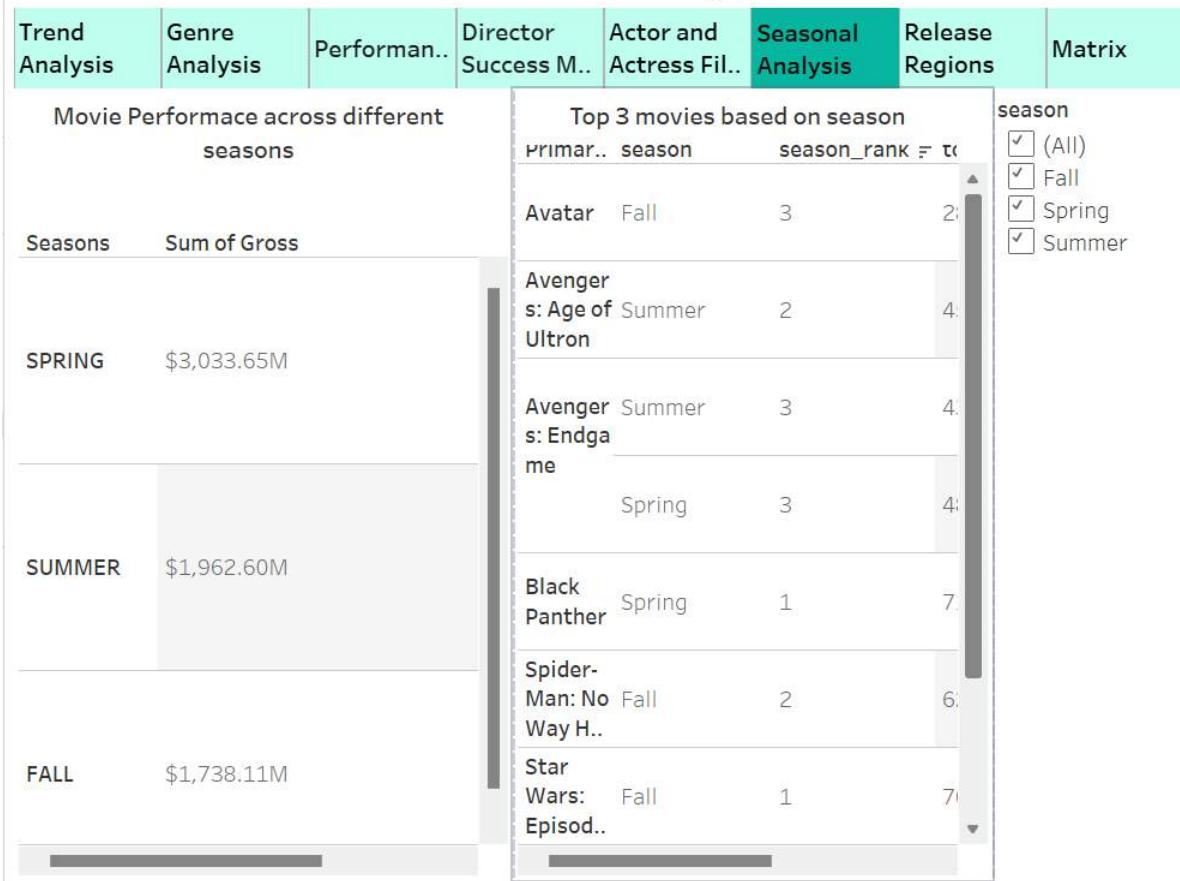
```

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

primaryTitle	originalTitle	startYear	runtimeMinutes	Gross	season
Spider-Man: No Way Home	Spider-Man: No Way Home	2021-01-01 00:00:00	148	121964712	Fall
Star Wars: Episode VII - The Force Awakens	Star Wars: Episode VII - The Force Awakens	2015-01-01 00:00:00	138	119119282	Fall
Spider-Man: No Way Home	Spider-Man: No Way Home	2021-01-01 00:00:00	148	73941279	Fall
Avengers: Endgame	Avengers: Endgame	2019-01-01 00:00:00	181	157461641	Spring
Avengers: Endgame	Avengers: Endgame	2019-01-01 00:00:00	181	109264122	Spring
Avengers: Infinity War	Avengers: Infinity War	2018-01-01 00:00:00	149	106334939	Spring
The Avengers	The Avengers	2012-01-01 00:00:00	143	9017913	Summer
The Avengers	The Avengers	2012-01-01 00:00:00	143	5826863	Summer
The Avengers	The Avengers	2012-01-01 00:00:00	143	5641642	Summer



Seasonal Analysis



Release Regions:

1. Identify movies that have had the widest release across multiple regions.
- a. The goal is to identify which movie was released in most regions.

```

SELECT
    tr.titleId,
    MAX(tr.title) AS movie_name,
    COUNT(DISTINCT tr.region) AS region_count
FROM
    dim_titleregion tr
GROUP BY
    tr.titleId
ORDER BY
    region_count DESC
LIMIT 7;
  
```

```

265   FROM
266     dim_titleregion tr
267   GROUP BY
268     tr.titleId
269   ORDER BY
270     region_count DESC
271   LIMIT 7;
272

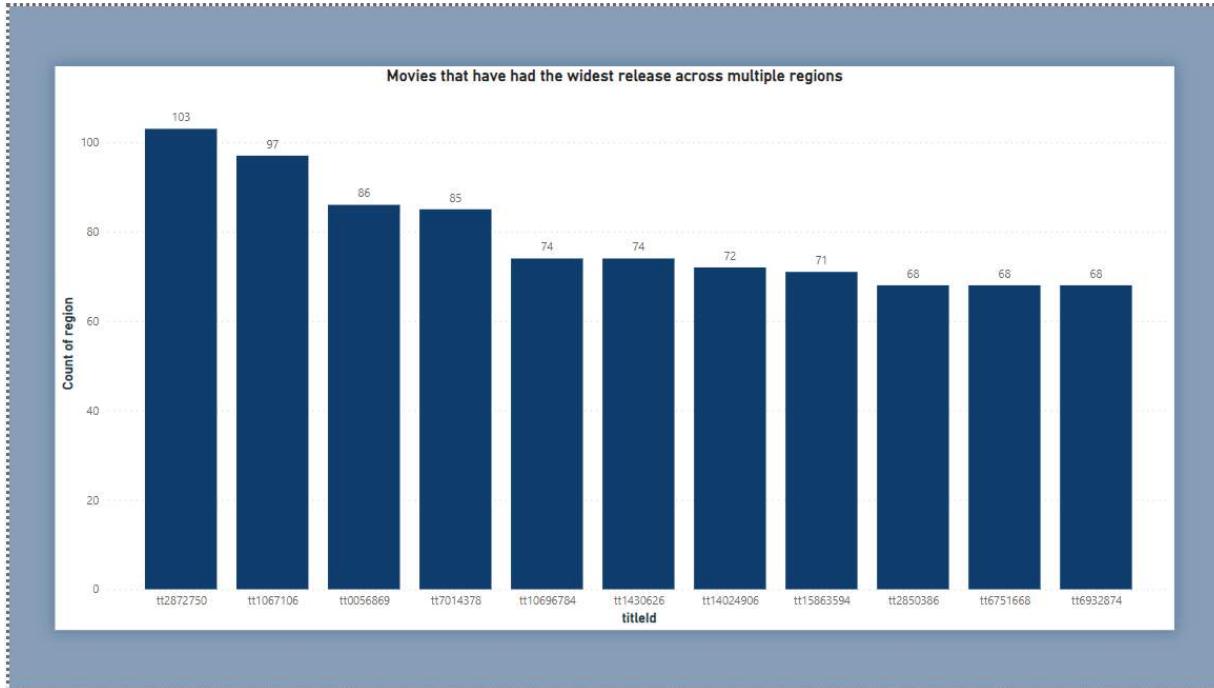
```

Result Grid | Filter Rows: _____ | Export: | Wrap Cell Content: | Fetch rows: |

	titleId	movie_name	region_count
▶	tt2872750	笑笑羊大電影	103
	tt1067106	聖誕夜怪譚	97
	tt0056869	鳥	86
	tt7014378	神之山嶺	85
	tt1430626	海賊天團	74
	tt10696784	復仇之淵	74
	tt14024906	Paranoid: Her sey sil bastan	72

RELEASE REGIONS

Trend Analysis	Genre Analysis	Performance	Director Success M..	Actor and Actress Fil..	Seasonal Analysis	Release Regions	Matrix
Widest release across multiple regions							
Title Id		Max. Title		Distinct count of Region			
tt7014378		Le Sommet des dieux		1		Abc	
tt6932874		Босс-молокосос 2		1		Abc	
tt6751668		Parazit		1		Abc	
tt4302938		Kubo in dve struni		1		Abc	
tt2872750		Xiao Yang Xiao En		1		Abc	
tt2850386		The Croods 2: A New Age		1		Abc	
tt2386490		How to Train Your Dragon: The Hid..	1			Abc	
tt2262227		Книга жизни		1		Abc	
tt1667889		Ice Age: Th4w		1		Abc	
tt15863594		Just Short of Perfect		1		Abc	
tt1430626		Pirates!		1		Abc	
tt14024906		Paranoid: Everything Wipe from th..	1			Abc	
tt10696784		Vực Sâu Thủ Hận		1		Abc	
tt1067106		Navidad divertida de Scrooge		1		..	



Matrix:

Create a dashboard for the 9 box office movies that shows movie-wise:

1. No. of directors

SELECT

```
me.tconst,
me.primaryTitle AS movie_title,
GROUP_CONCAT(DISTINCT dp.primaryName) AS director_names,
COUNT(DISTINCT c.nconst) AS num_directors
```

FROM

dim_crew c

JOIN

fct_movie_earning fe ON c.tconst = fe.tconst

JOIN

dim_movie_title_mysql me ON fe.tconst = me.tconst

JOIN

dim_person dp ON c.nconst = dp.nconst

where

c.category = "Director"

GROUP BY

me.tconst, me.primaryTitle;

```

534     fct_movie_earning fe ON c.tconst = fe.tconst
535     JOIN
536         dim_movie_title_mysql me ON fe.tconst = me.tconst
537     JOIN
538         dim_person dp ON c.nconst = dp.nconst
539     where
540         c.category = "Director"
541     GROUP BY
542         me.tconst, me.primaryTitle;
543

```

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

	tconst	movie_title	director_names	num_directors
▶	tt0120338	Titanic	James Cameron	1
	tt0499549	Avatar	James Cameron	1
	tt0848228	The Avengers	Joss Whedon	1
	tt10872600	Spider-Man: No Way Home	Jon Watts	1
	tt1825683	Black Panther	Ryan Coogler	1
	tt2395427	Avengers: Age of Ultron	Joss Whedon	1
	tt2488496	Star Wars: Episode VII - The Force Awakens	J.J. Abrams	1
	tt4154756	Avengers: Infinity War	Anthony Russo, Joe Russo	2
	tt4154796	Avengers: Endgame	Anthony Russo, Joe Russo	2

2. No. of actors/actresses

```

SELECT
me.tconst,
me.primaryTitle AS movie_title,
GROUP_CONCAT(DISTINCT dp.primaryName) AS actor_names,
COUNT(DISTINCT c.nconst) AS num_actors
FROM
dim_crew c
JOIN
fct_movie_earning fe ON c.tconst = fe.tconst
JOIN
dim_movie_title_mysql me ON fe.tconst = me.tconst
JOIN
dim_person dp ON c.nconst = dp.nconst
where
c.category = "actor" or c.category = "actress"
GROUP BY
me.tconst, me.primaryTitle;

```

```

526 •      SELECT
527      me.tconst,
528      me.primaryTitle AS movie_title,
529      GROUP_CONCAT(DISTINCT dp.primaryName) AS actors_names,
530      COUNT(DISTINCT c.nconst) AS num_actors
531  FROM
532      dim_crew c
533  JOIN

```

Result Grid | Filter Rows: | Exports: | Wrap Cell Content: |

tconst	movie_title	actors_names	num_actors
tt0120338	Titanic	Billy Zane,Kate Winslet,Kathy Bates,Leonardo DiCaprio	4
tt0499549	Avatar	Michelle Rodriguez,Sam Worthington,Sigourney Weaver,Zoe Saldana	4
tt0848228	The Avengers	Chris Evans,Jeremy Renner,Robert Downey Jr.,Scarlett Johansson	4
tt10872600	Spider-Man: No Way Home	Benedict Cumberbatch,Jacob Batalon,Tom Holland,Zendaya	4
tt1825683	Black Panther	Chadwick Boseman,Danai Gurira,Lupita Nyong'o,Michael B. Jordan	4
tt2395427	Avengers: Age of Ultron	Chris Evans,Chris Hemsworth,Mark Ruffalo,Robert Downey Jr.	4
tt2488496	Star Wars: Episode VII - The Force Awakens	Daisy Ridley,Domhnall Gleeson,John Boyega,Oscar Isaac	4
tt4154756	Avengers: Infinity War	Chris Evans,Chris Hemsworth,Mark Ruffalo,Robert Downey Jr.	4
tt4154796	Avengers: Endgame	Chris Evans,Chris Hemsworth,Mark Ruffalo,Robert Downey Jr.	4

3. No. of writers

```

SELECT
me.tconst,
me.primaryTitle AS movie_title,
GROUP_CONCAT(DISTINCT dp.primaryName) AS writer_names,
COUNT(DISTINCT c.nconst) AS num_writers
FROM
dim_crew c
JOIN
fct_movie_earning fe ON c.tconst = fe.tconst
JOIN
dim_movie_title_mysql me ON fe.tconst = me.tconst
JOIN
dim_person dp ON c.nconst = dp.nconst
where
c.category = "writer"
GROUP BY
me.tconst, me.primaryTitle;

```

```

553 JOIN
554     fct_movie_earning fe ON c.tconst = fe.tconst
555 JOIN
556     dim_movie_title_mysql me ON fe.tconst = me.tconst
557 JOIN
558     dim_person dp ON c.nconst = dp.nconst
559     where
560         c.category = "writer"
561 GROUP BY
562     me.tconst, me.primaryTitle;

```

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

	tconst	movie_title	writer_names	num_writers
▶	tt0848228	The Avengers	Zak Penn	1
	tt10872600	Spider-Man: No Way Home	Chris McKenna,Erik Sommers,Stan Lee,Steve Ditko	4
	tt1825683	Black Panther	Jack Kirby,Joe Robert Cole,Stan Lee	3
	tt2395427	Avengers: Age of Ultron	Jack Kirby,Jim Starlin,Joe Simon,Stan Lee	4
	tt2488496	Star Wars: Episode VII - The Force Awakens	George Lucas,Lawrence Kasdan,Michael Arndt	3
	tt4154756	Avengers: Infinity War	Christopher Markus,Jack Kirby,Stan Lee,Stephe...	4
	tt4154796	Avengers: Endgame	Christopher Markus,Jack Kirby,Stan Lee,Stephe...	4

4. Total worldwide gross earnings

SELECT

```

COUNT(DISTINCT fr.tconst) AS movie_count,
GROUP_CONCAT(DISTINCT me.primaryTitle ORDER BY me.primaryTitle ASC) AS
    movie_names,
SUM(fr.Gross) AS total_gross_earnings
FROM
    dim_crew c
JOIN
    fct_movie_earning fr ON c.tconst = fr.tconst
JOIN
    dim_movie_title_mysql me ON fr.tconst = me.tconst;

```

```

373 • SELECT
374     COUNT(DISTINCT fr.tconst) AS movie_count,
375     GROUP_CONCAT(DISTINCT me.primaryTitle ORDER BY me.primaryTitle ASC) AS movie_names,
376     SUM(fr.Gross) AS total_gross_earnings
377     FROM
378     dim_crew c
379     JOIN

```

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

	movie_count	movie_names	total_gross_earnings
▶	9	Avatar,Avengers: Age of Ultron,Avengers: Endgame,Avengers: Infinity War,Black Panther,Spider-Man: No Way Home,S...	67343621130

SELECT

```

me.tconst,
me.primaryTitle AS movie_title,
SUM(fr.Gross) AS total_gross_earnings
FROM
    dim_crew c
JOIN
    fct_movie_earning fr ON c.tconst = fr.tconst
JOIN

```

```

dim_movie_title_mysql me ON fr.tconst = me.tconst
GROUP BY
me.tconst, me.primaryTitle;

```

The screenshot shows a MySQL query editor with the following code:

```

450
451 • SELECT
452     me.tconst,
453     me.primaryTitle AS movie_title,
454     SUM(fr.Gross) AS total_gross_earnings
455 FROM
456     dim_crew c
457 JOIN
458     fct_movie_earning fr ON c.tconst = fr.tconst

```

Below the code is a result grid displaying the following data:

	tconst	movie_title	total_gross_earnings
▶	tt0120338	Titanic	6440786380
	tt0499549	Avatar	7605086250
	tt0848228	The Avengers	6377482070
	tt10872600	Spider-Man: No Way Home	8543374880
	tt1825683	Black Panther	7252595660
	tt2395427	Avengers: Age of Ultron	4835984180
	tt2488496	Star Wars: Episode VII - The Force Awakens	9926426890
	tt4154756	Avengers: Infinity War	7178154820
	tt4154796	Avengers: Endgame	9183730000

5. Average ratings

```

SELECT
    me.tconst,
    me.primaryTitle AS movie_title,
    AVG(rt.averageRating) AS average_rating
FROM
    dim_crew c
JOIN
    fct_movie_earning fe ON c.tconst = fe.tconst
JOIN
    dim_movie_title_mysql me ON fe.tconst = me.tconst
JOIN
    fct_rating rt ON fe.tconst = rt.tconst
GROUP BY
    me.tconst, me.primaryTitle;

```

```

466     me.tconst,
467     me.primaryTitle AS movie_title,
468     AVG(rt.averageRating) AS average_rating
469 FROM
470     dim_crew c
471 JOIN
472     fct_movie_earning fe ON c.tconst = fe.tconst
473 JOIN
474     dim_movie_title_mysql me ON fe.tconst = me.tconst

```

Result Grid			
	tconst	movie_title	average_rating
▶	tt0120338	Titanic	8.0000
	tt0499549	Avatar	8.0000
	tt0848228	The Avengers	8.0000
	tt10872600	Spider-Man: No Way Home	8.0000
	tt1825683	Black Panther	7.0000
	tt2395427	Avengers: Age of Ultron	7.0000
	tt2488496	Star Wars: Episode VII - The Force Awakens	8.0000
	tt4154756	Avengers: Infinity War	8.0000
	tt4154796	Avengers: Endgame	8.0000

6. No. of regions, it was released

```

SELECT
me.tconst,
me.primaryTitle AS movie_title,
COUNT(DISTINCT tr.region) AS num_regions_released
FROM
dim_crew c
JOIN
fct_movie_earning fe ON c.tconst = fe.tconst
JOIN
dim_movie_title_mysql me ON fe.tconst = me.tconst
JOIN
dim_titleregion tr ON me.tconst = tr.titleId
GROUP BY
me.tconst, me.primaryTitle;

```

```

479
480 •   SELECT
481     me.tconst,
482     me.primaryTitle AS movie_title,
483     COUNT(DISTINCT tr.region) AS num_regions_released

```

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

	tconst	movie_title	num_regions_released
▶	tt0120338	Titanic	48
	tt0499549	Avatar	44
	tt0848228	The Avengers	50
	tt10872600	Spider-Man: No Way Home	53
	tt1825683	Black Panther	42
	tt2395427	Avengers: Age of Ultron	50
	tt2488496	Star Wars: Episode VII - The Force Awakens	51
	tt4154756	Avengers: Infinity War	46
	tt4154796	Avengers: Endgame	49

7. No. of genres

```

SELECT
    me.tconst,
    me.primaryTitle AS movie_title,
    COUNT(DISTINCT dg.genres) AS num_genres
FROM
    dim_crew c
JOIN
    fct_movie_earning fe ON c.tconst = fe.tconst
JOIN
    dim_movie_title_mysql me ON fe.tconst = me.tconst
JOIN
    dim_genre dg ON me.tconst = dg.tconst
GROUP BY
    me.tconst, me.primaryTitle;

```

```

500      dim_crew c
501      JOIN
502          fct_movie_earning fe ON c.tconst = fe.tconst
503      JOIN
504          dim_movie_title_mysql me ON fe.tconst = me.tconst
505      JOIN
506          dim_genre dg ON me.tconst = dg.tconst
507      GROUP BY
508          me.tconst, me.primaryTitle;

```

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

	tconst	movie_title	num_genres
▶	tt0120338	Titanic	2
	tt0499549	Avatar	3
	tt0848228	The Avengers	3
	tt10872600	Spider-Man: No Way Home	3
	tt1825683	Black Panther	3
	tt2395427	Avengers: Age of Ultron	3
	tt2488496	Star Wars: Episode VII - The Force Awakens	3
	tt4154756	Avengers: Infinity War	3
	tt4154796	Avengers: Endgame	3

Dashboard for the 9 box office movies

primaryTitle	Count of genres	Average of averageRating	Sum of Gross	Count of region	Actor/Actress	Director	Writer	Count of region	Count of titleId
■ Titanic	2	8.00	644078638	1	4	1		1	1
■ The Avengers	3	8.00	637748207	1	4	1	1	1	1
■ Star Wars: Episode VII - The Force Awakens	3	8.00	992642689	1	4	1	3	1	1
■ Spider-Man: No Way Home	3	8.00	854337488	1	4	1	4	1	1
■ Black Panther	3	7.00	725259566	1	4	1	3	1	1
■ Avengers: Infinity War	3	8.00	717815482	1	4	2	4	1	1
■ Avengers: Endgame	3	8.00	918373000	1	4	2	4	1	1
■ Avengers: Age of Ultron	3	7.00	483598418	1	4	1	4	1	1
■ Avatar	3	8.00	760508625	1	4	1		1	1

MATRIX

Trend Analysis	Genre Analysis	Performance	Director Success M.	Actor and Actress Fil...	Seasonal Analysis	Release Regions	Matrix
Movie-Wise Director, Actor, Actress, Writer, Gross Earnings, Average Rating and Region							
<code>tconst (</code>							
<code>fct_movie... =</code>							
<code>tt2488496</code>	Primary Title	Primary Name		Sum of Gross	Average		Category
	Star Wars: Episode VII - The Force Awakens	Daisy Ridley	\$992.64M	8			<input type="checkbox"/> (All)
		Domhnall Gleeson	\$992.64M	8			<input checked="" type="checkbox"/> actor
		John Boyega	\$992.64M	8			<input checked="" type="checkbox"/> actress
		Oscar Isaac	\$992.64M	8			<input type="checkbox"/> archive_foo...
<code>tt4154796</code>	Avengers: Endgame	Chris Evans	\$918.37M	8			<input type="checkbox"/> archive_sou...
		Chris Hemsworth	\$918.37M	8			<input type="checkbox"/> cinematogr...
		Mark Ruffalo	\$918.37M	8			<input type="checkbox"/> composer
		Robert Downey Jr.	\$918.37M	8			<input type="checkbox"/> director
<code>tt10872600</code>	Spider-Man: No Way Home	Benedict Cumberbatch	\$854.34M	8			<input type="checkbox"/> editor
		Jacob Batalon	\$854.34M	8			<input type="checkbox"/> producer
		Tom Holland	\$854.34M	8			<input type="checkbox"/> production...
		Zendaya	\$854.34M	8			<input type="checkbox"/> self
<code>tt0499549</code>	Avatar	Michelle Rodriguez	\$760.51M	8			<input type="checkbox"/> writer
		Sam Worthington	\$760.51M	8			
		Sigourney Weaver	\$760.51M	8			
		Zoe Saldana	\$760.51M	8			
<code>tt1825683</code>	Black Panther	Chadwick Boseman	\$725.26M	7			
		Danai Gurira	\$725.26M	7			
		Lupita Nyong'o	\$725.26M	7			
		Michael B. Jordan	\$725.26M	7			
<code>tt4154756</code>	Avengers: Infinity War	Chris Evans	\$717.82M	8			
		Chris Hemsworth	\$717.82M	8			
		Mark Ruffalo	\$717.82M	8			