

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

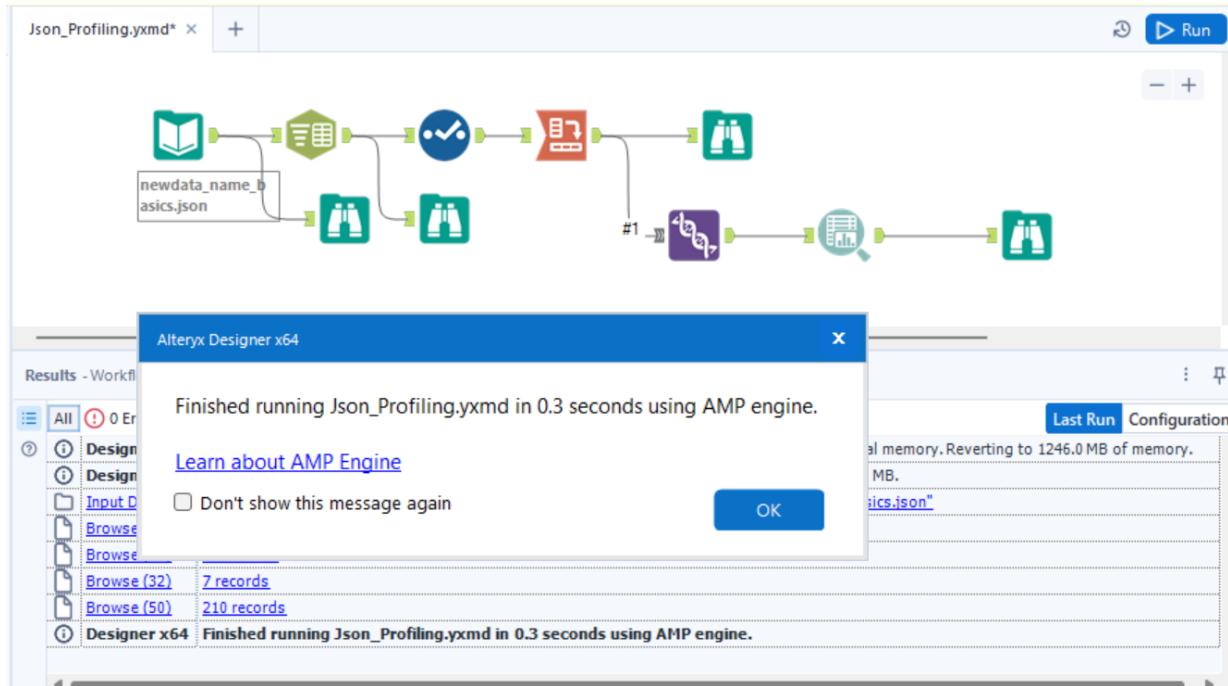
List of Tables:

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

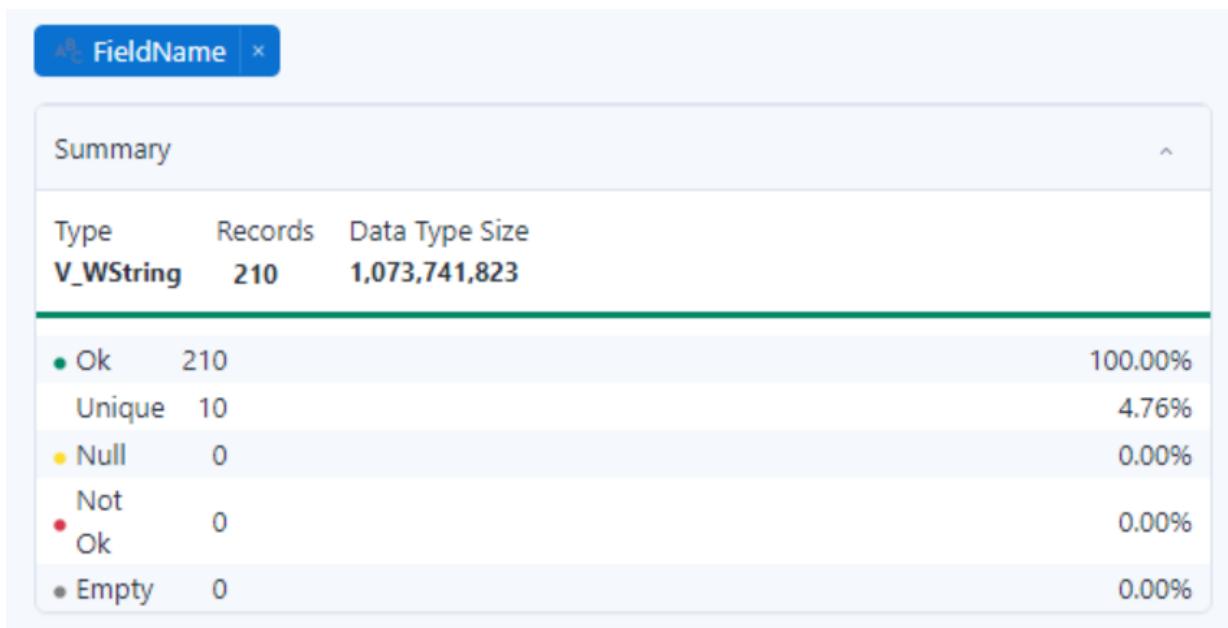
Data Profiling in Alteryx:

1. newdata_name_basics.json



Observations:

3 of 3 Fields		<input checked="" type="checkbox"/>	210 records displayed, 3,192 bytes	<input type="button" value="X"/>	<input type="button" value="✓"/>	<input type="button" value="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			



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

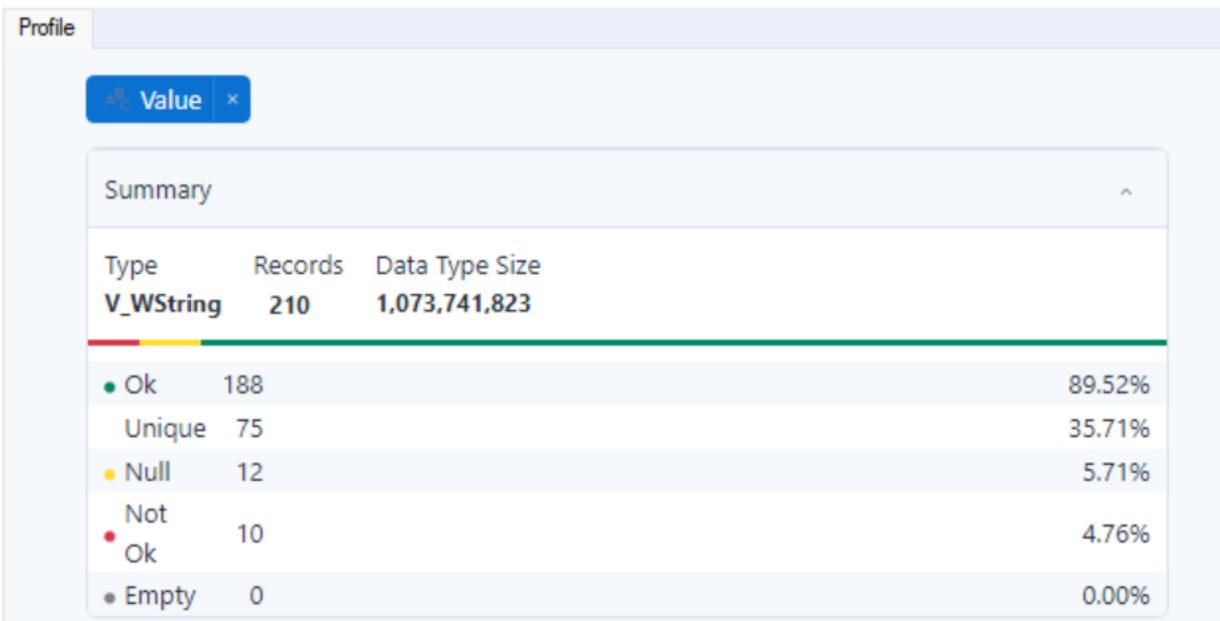
Name

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	0 Errors	0 Conv Errors	0 Warnings	3 Info	5 Files
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.			

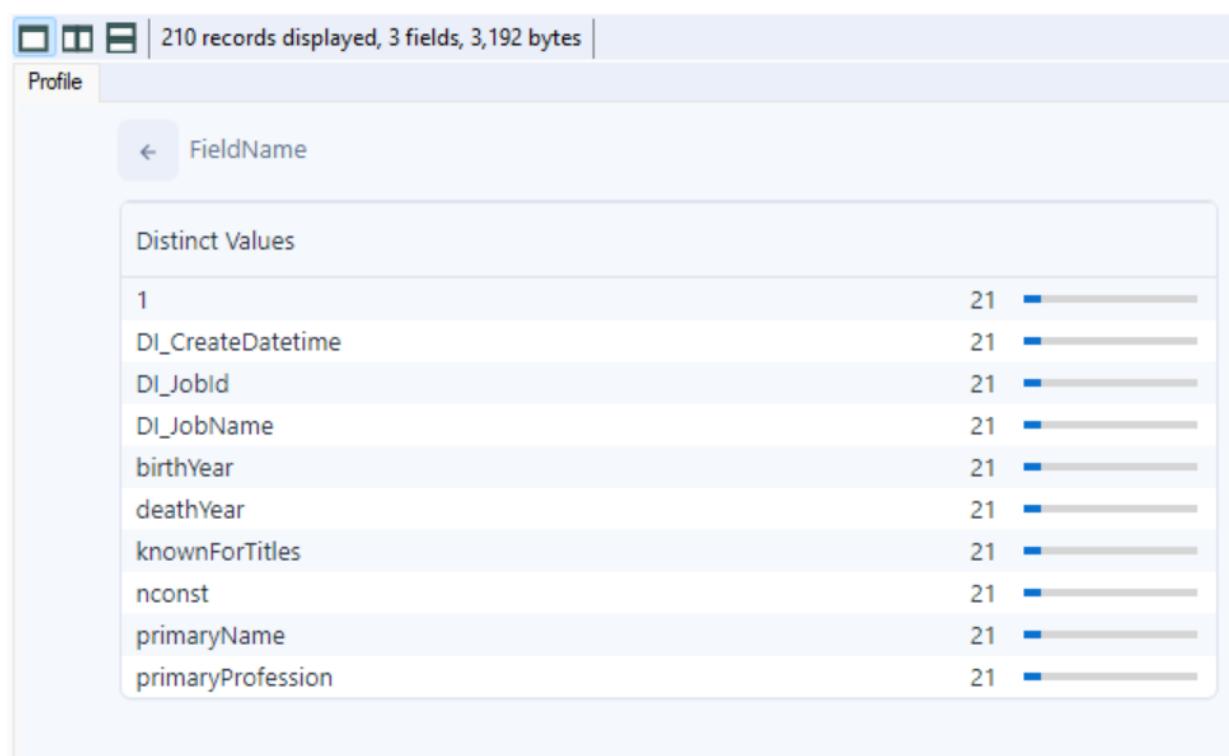
newdata_name_basics_profiling.ymd x + Run

Results - Browse (50) - Input

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

Record FieldName Name Value

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	Ok's	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

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

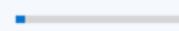
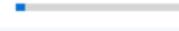
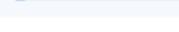
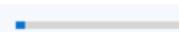
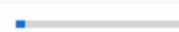
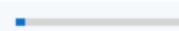
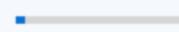
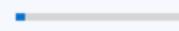
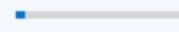
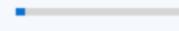
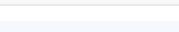
Min	1
Max	17
Average	10.20
Shortest Value	1
Longest Value	primaryProfession
First Alphanumeric Value	1
Last Alphanumeric Value	primaryProfession
Blanks	0
Values with Leading Whitespace	0
Values with Trailing Whitespace	0

Frequent Values

Value by count

10 out of 10 values



Profile	
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

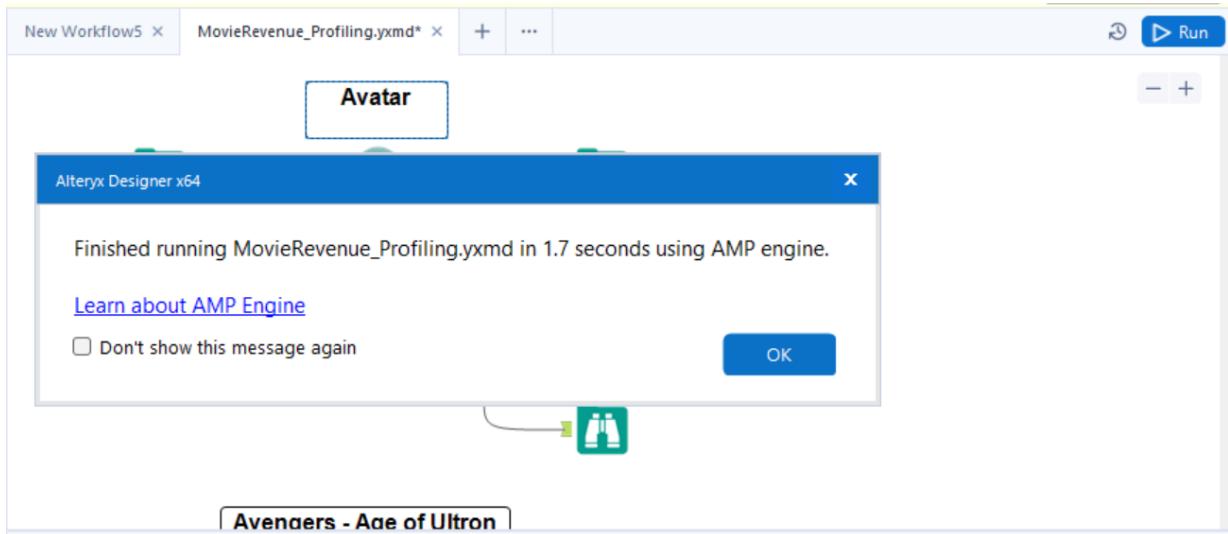
Frequent Values

Value by count

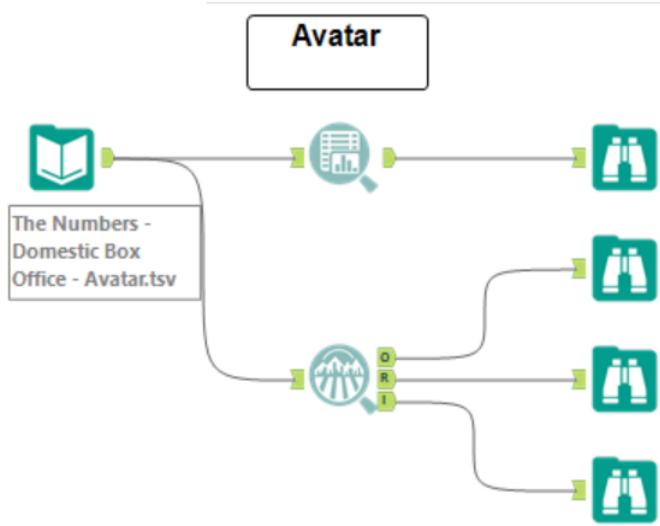
50 out of 75 values



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	9

Profile

Field**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 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

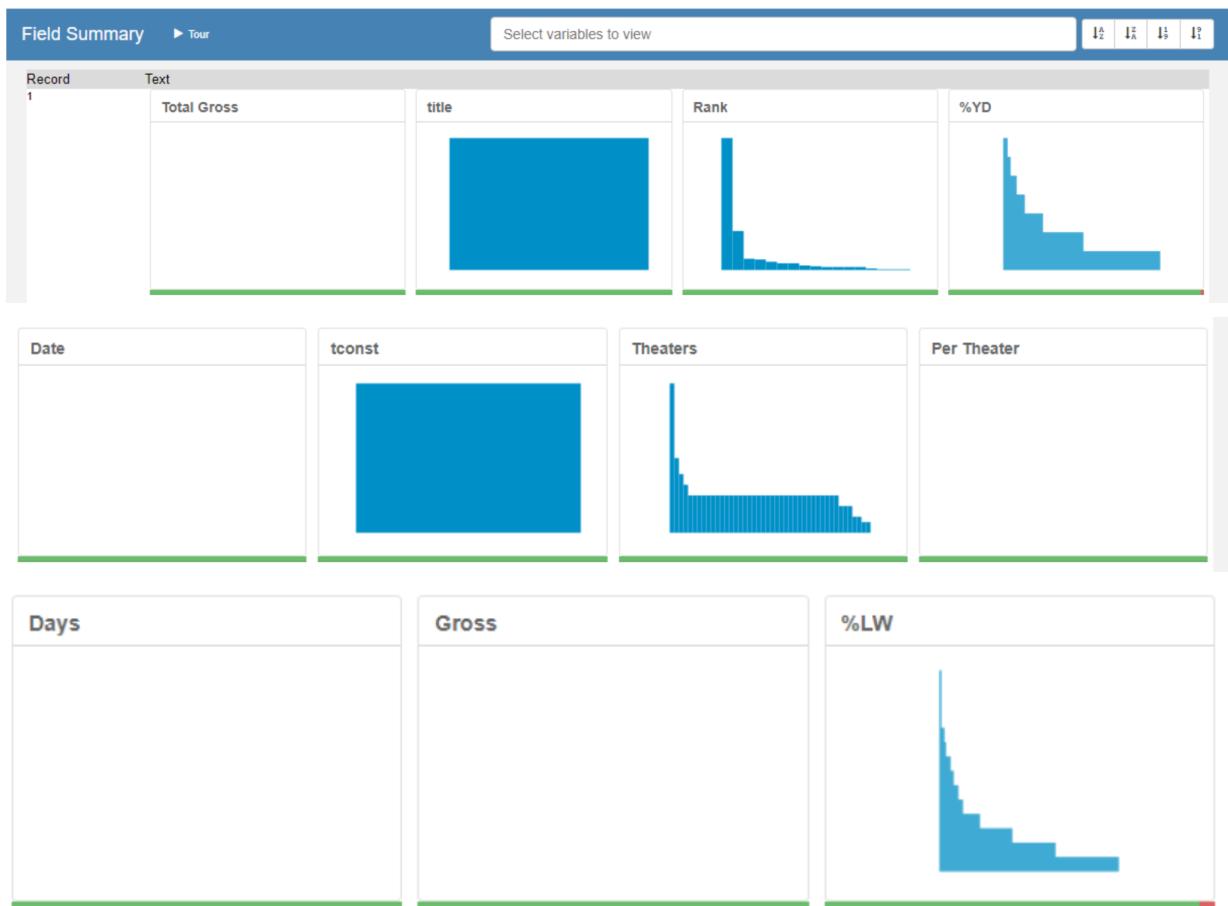
Type	Records	Data Type Size
V_WString	231	1,073,741,823

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

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

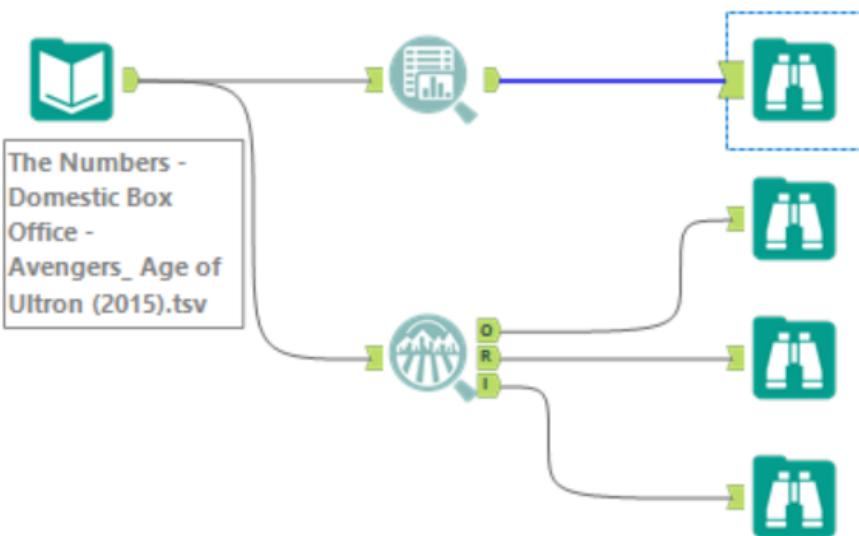
String/Character Fields

Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	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	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	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	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	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	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	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

ABC 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

ABC 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

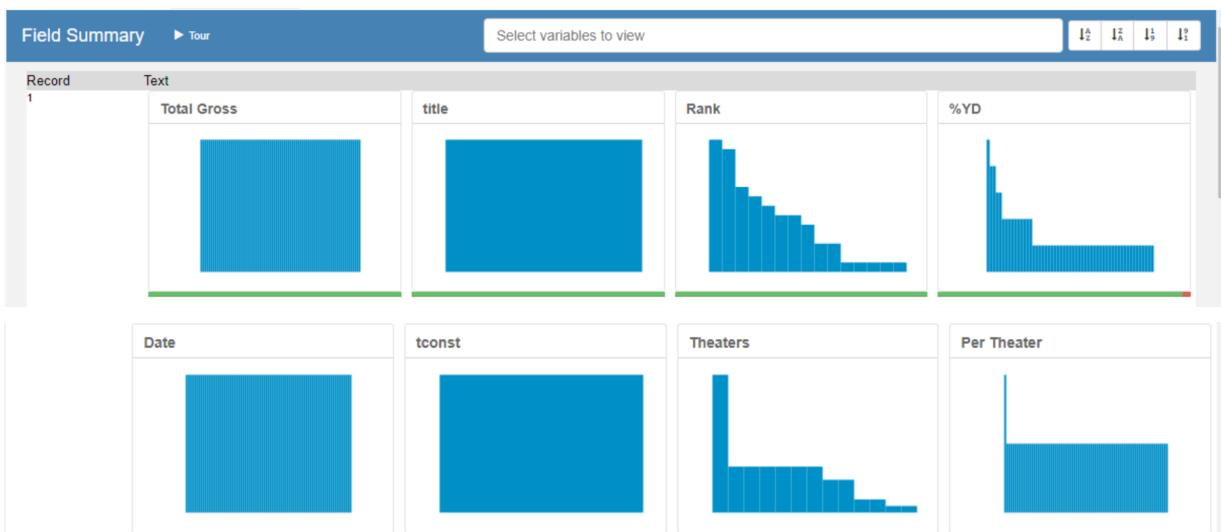
Value

Summary

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

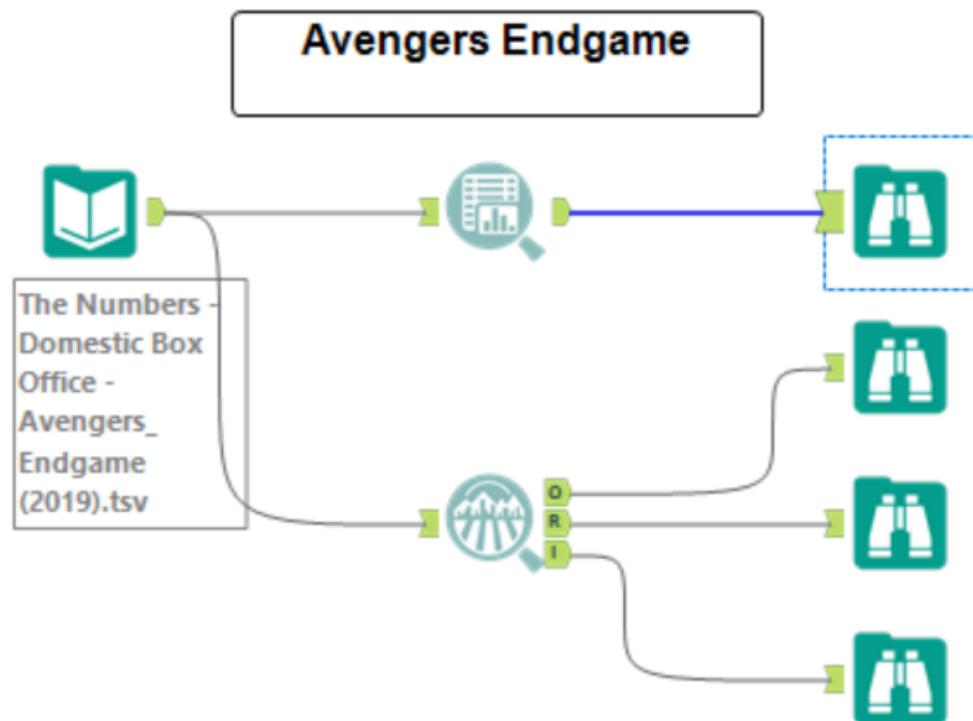
● Ok	209	90.48%
Unique	80	34.63%
● Null	11	4.76%
Not Ok	11	4.76%
● Empty	0	0.00%

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	\$69,841	\$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	





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="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

A B C 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

A B C 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%

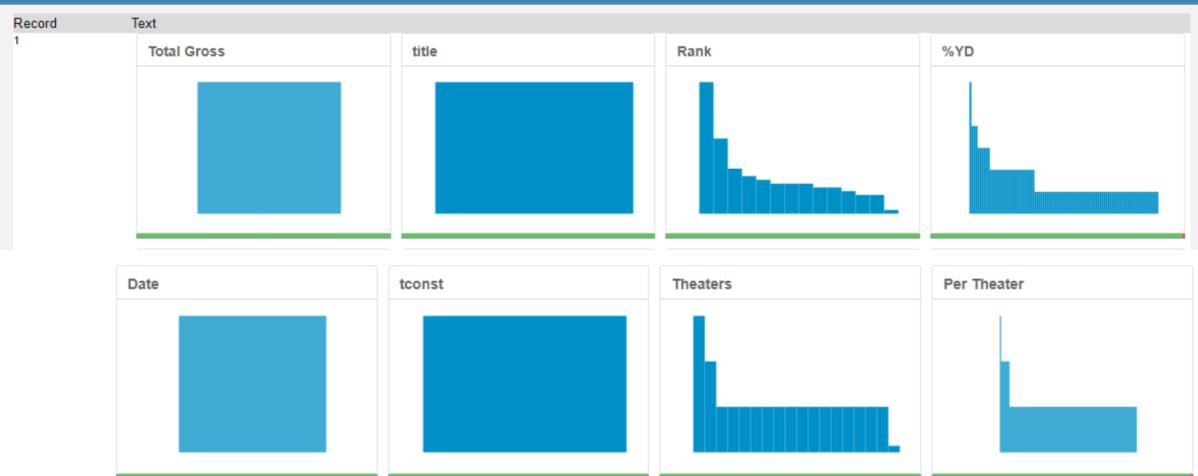
String/Character Fields						
Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count
Total Gross	0.0%	141	\$60,000,000	\$157,461,641	1	1
Rank	0.0%	14	P	12	1	35
%YD	1.4%	93	0.7	-31%	1	6
Date	0.0%	141	May 1, 2019	Apr 25, 2019	1	1
Theaters	0.0%	18	0	4,662	1	21
Per Theater	0.7%	131	\$91	\$33,776	1	3
Days	0.0%	141	0	100	1	1
Gross	0.0%	141	\$940	\$157,461,641	1	1
%LW	5.7%	62	-5%	-74%	1	8
title	0.0%	1	Avengers: Endgame	Avengers: Endgame	141	141
tconst	0.0%	1	tt4154796	tt4154796	141	141

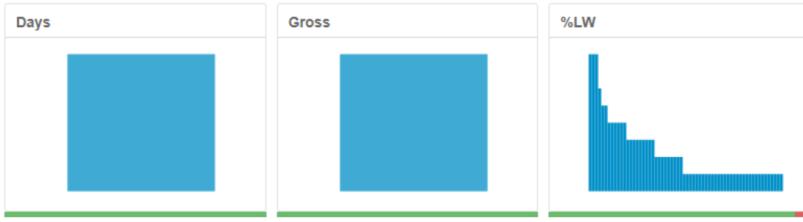
Field Summary

► Tour

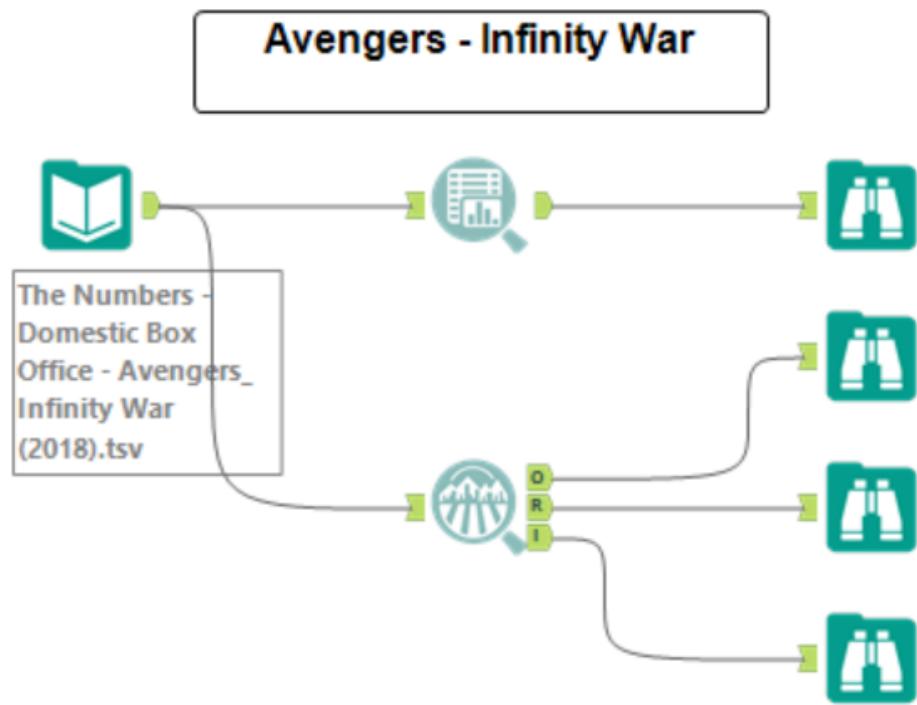
Select variables to view

1 1 1 1



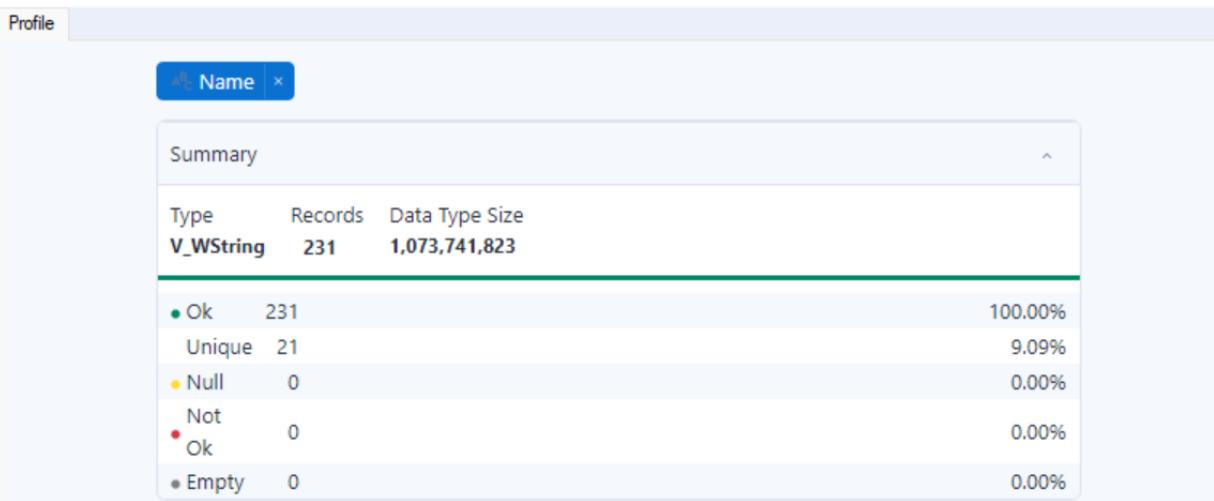
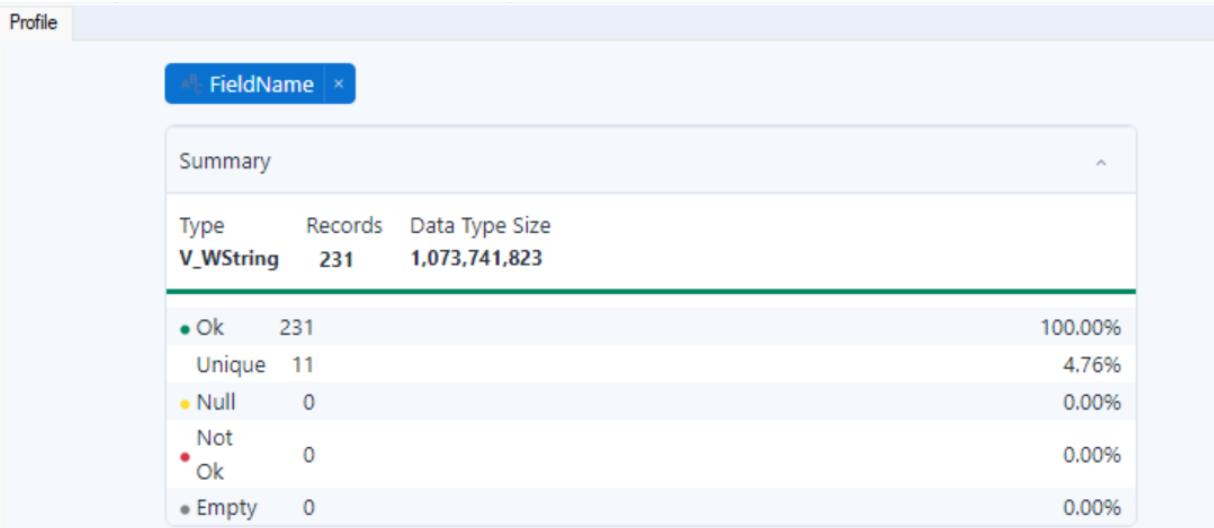


3.4 Avengers - Infinity War



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

Value

Summary

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

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

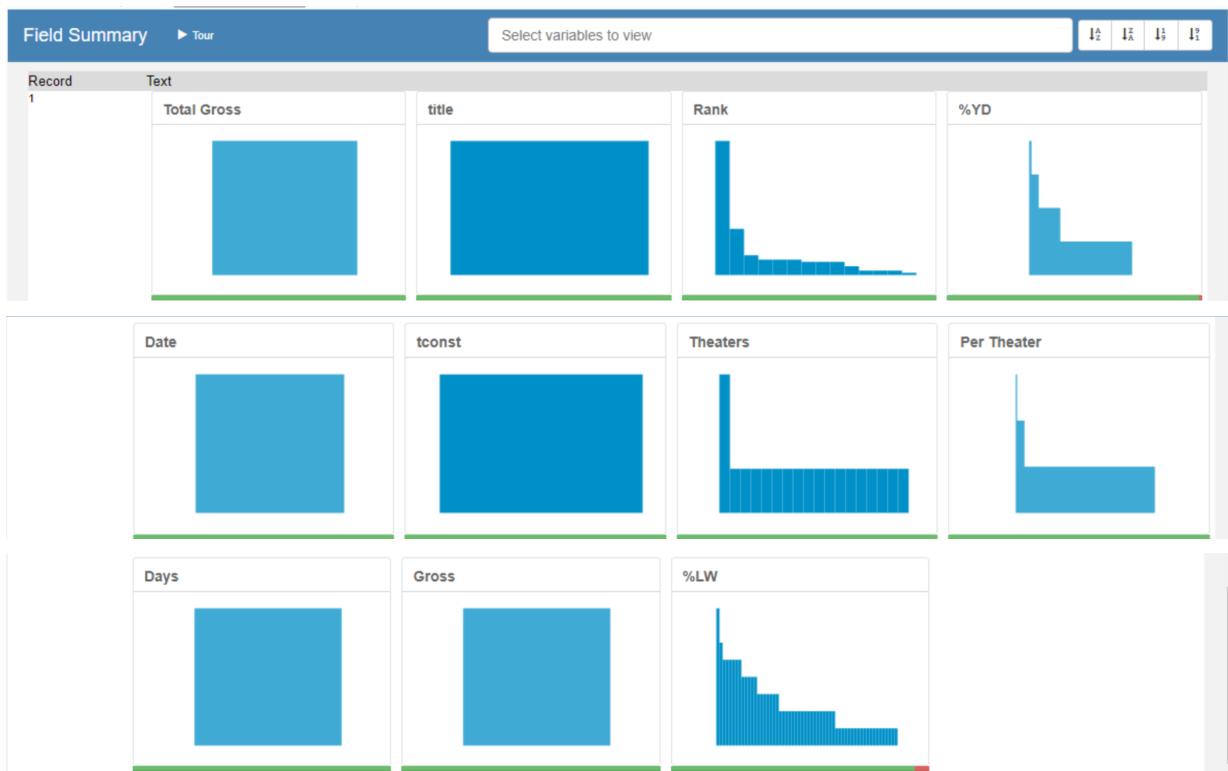
Report Profile

1 of 1 Fields | Records 1 to 1 |

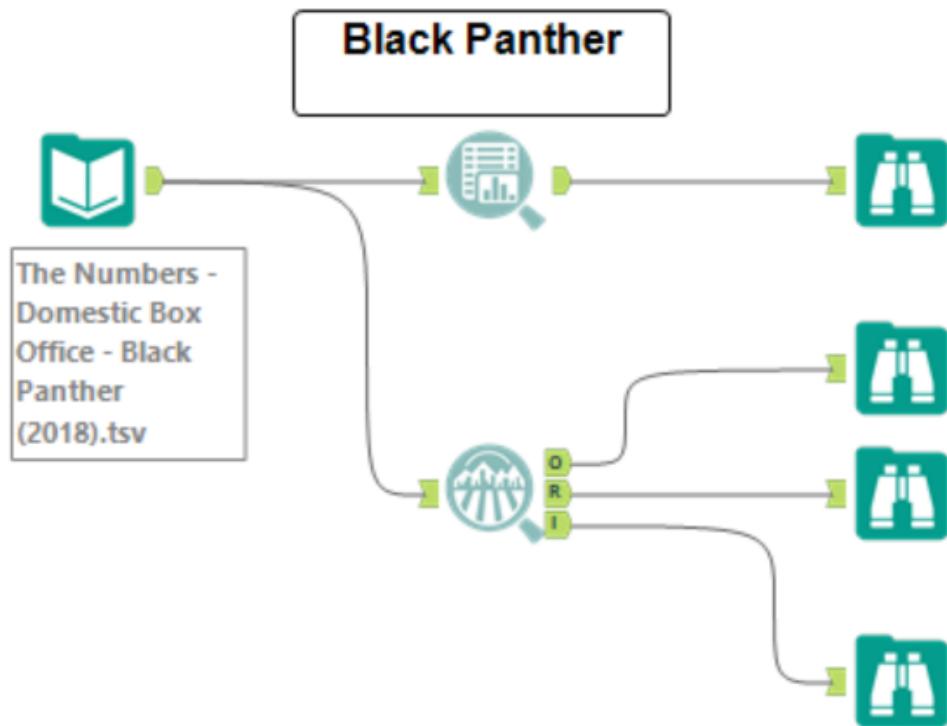
String/Character Fields

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	61	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	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%	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	22	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	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%	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.

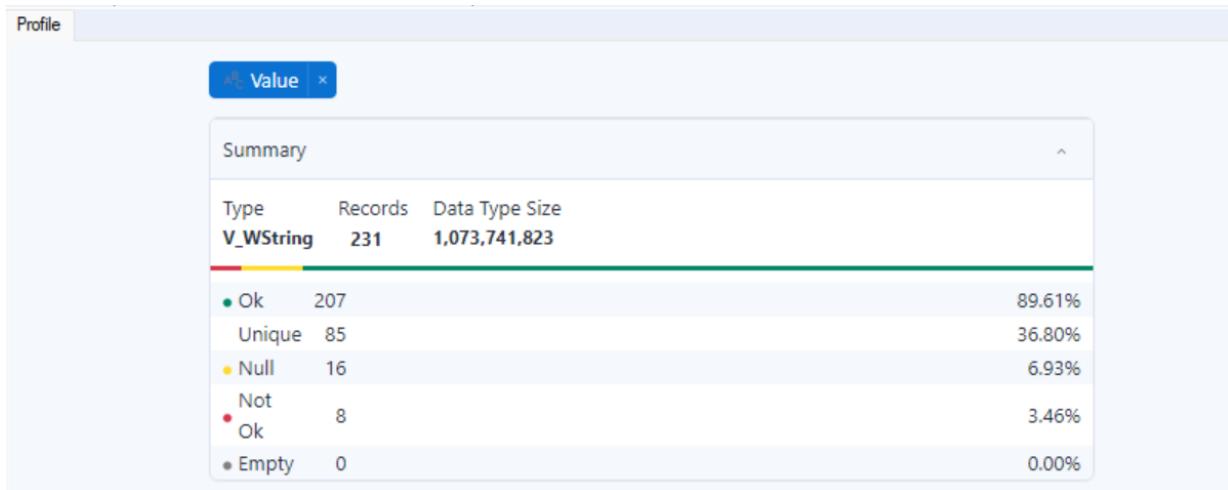
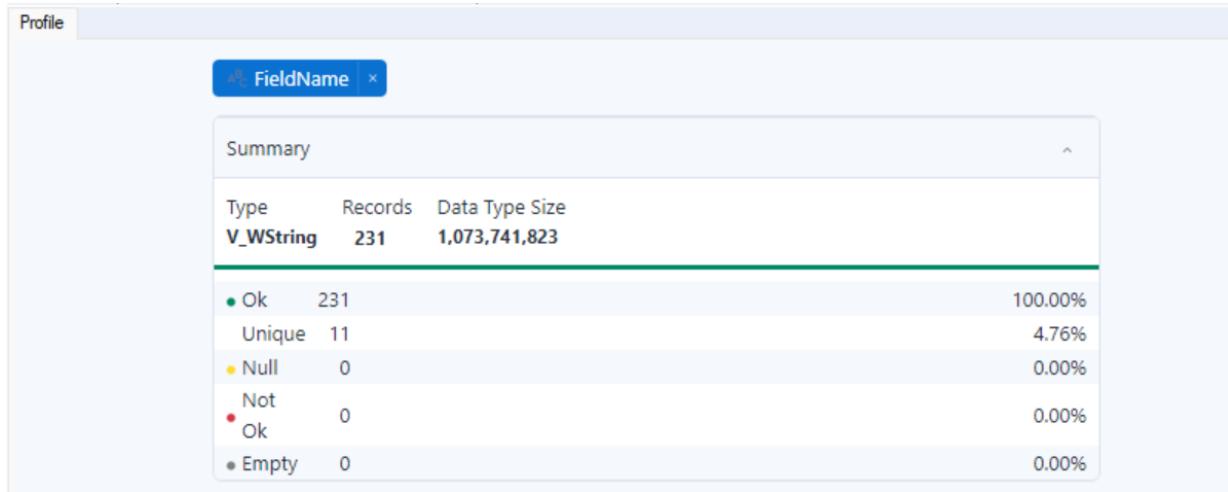
Gross	0.0%	141	\$9,993	\$106,334,939	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.	
%LW	5.7%	58	-6%	-70%	1	8 Some values of this field have a small number of value counts. If Appropriate, consider combining some value 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				

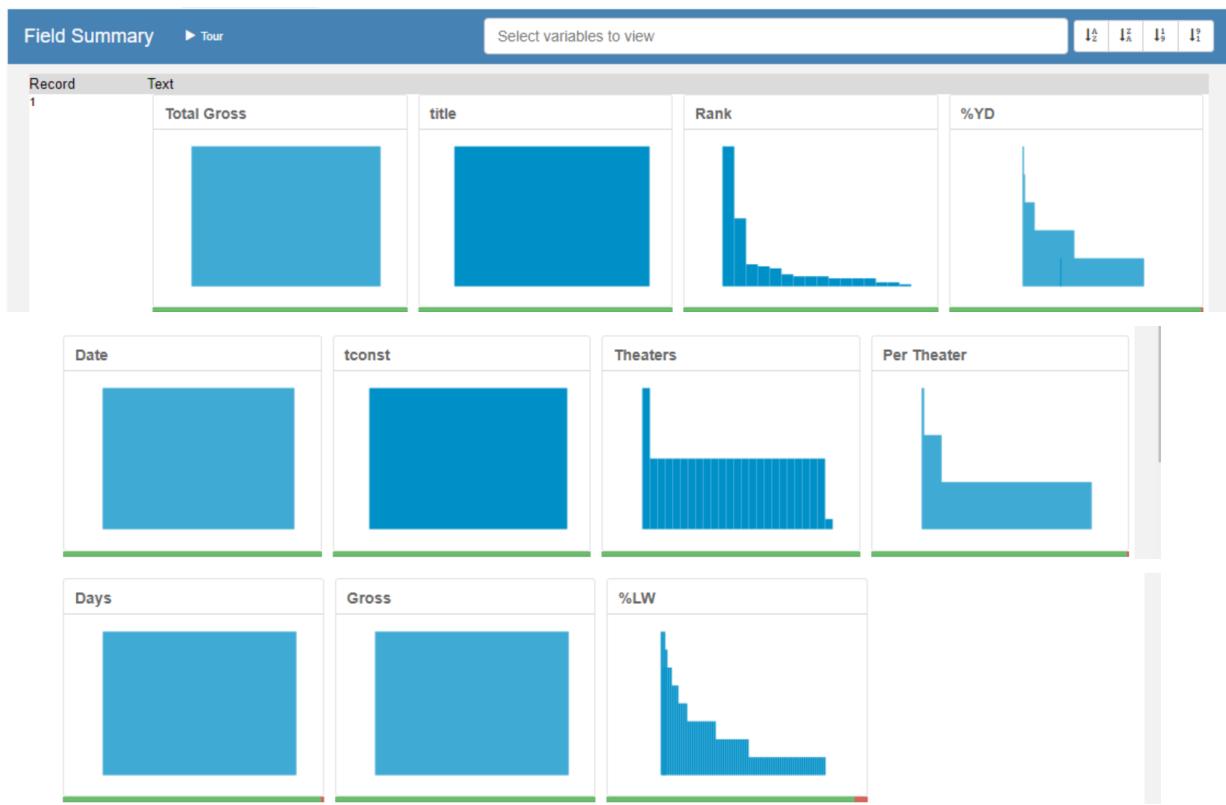


Report Profile

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

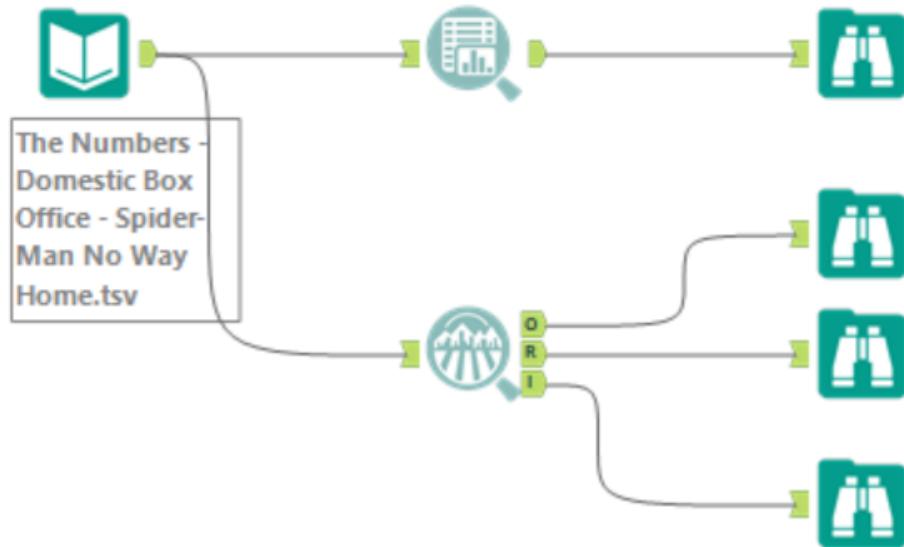
Name	% Missing	Unique Values	Shortest Value	Longest Value	Min Value Count	Max Value Count	Remarks
Total Gross	0.0%	176	\$25,200,000	\$141,936,512	1	1	1 Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.
Rank	0.0%	16	P	10	1	70	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 <input checked="" type="checkbox"/> <input type="button" value="X"/> 231 records displayed, 4,233 bytes <input type="button" value="X"/> <input type="button" value="✓"/> <input type="button" value="Search"/>			
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 Field Name 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	0		0.00%
● 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%

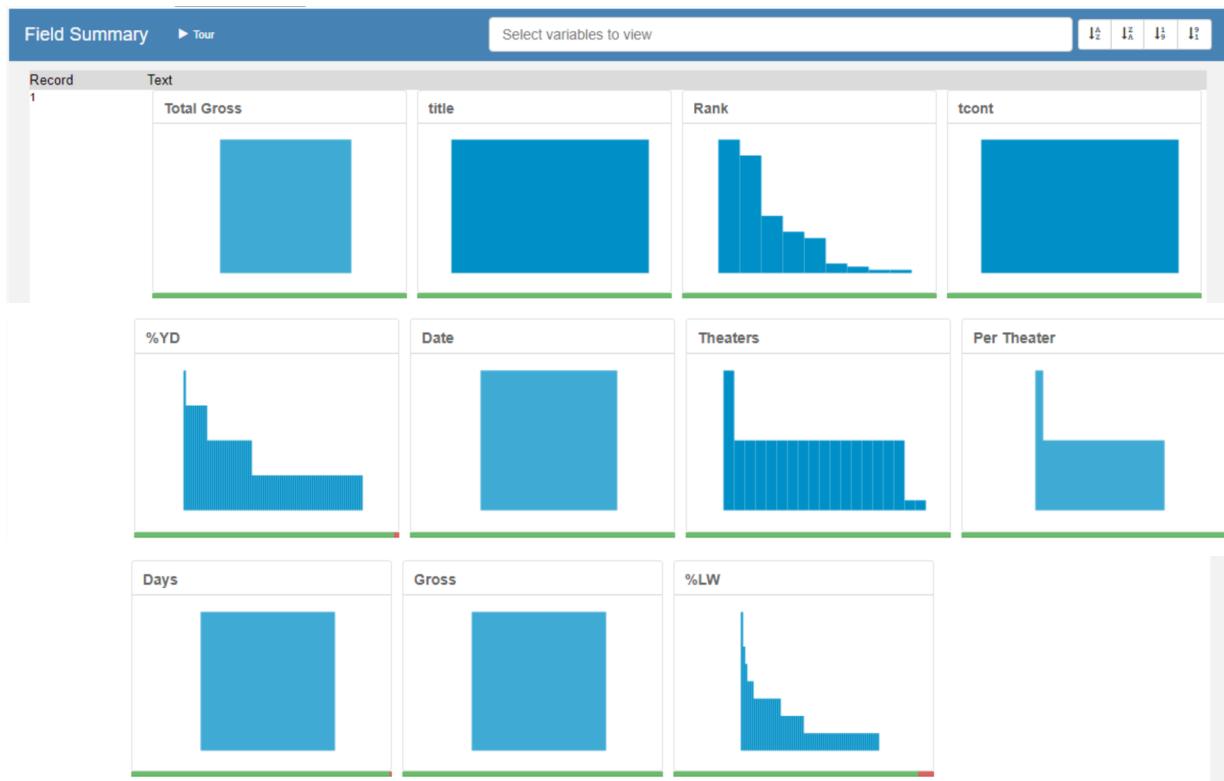
Profile Value x

Summary			
Type	Records	Data Type	Size
V_WString	231	1,073,741,823	
● Ok	209		90.48%
Unique	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 | ◀ ▶

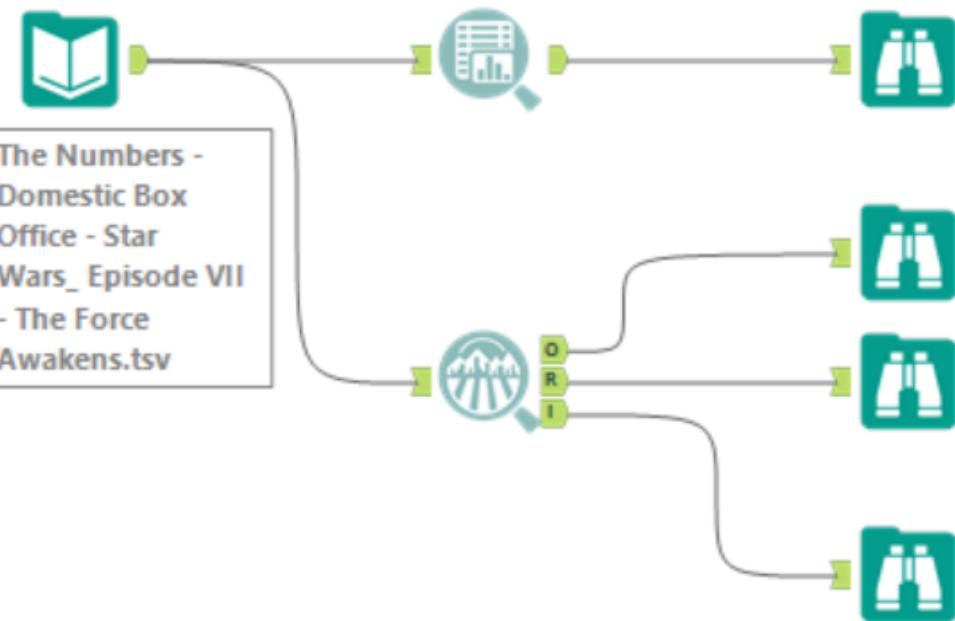
String/Character Fields

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



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

A 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

A 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	\$28,815	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

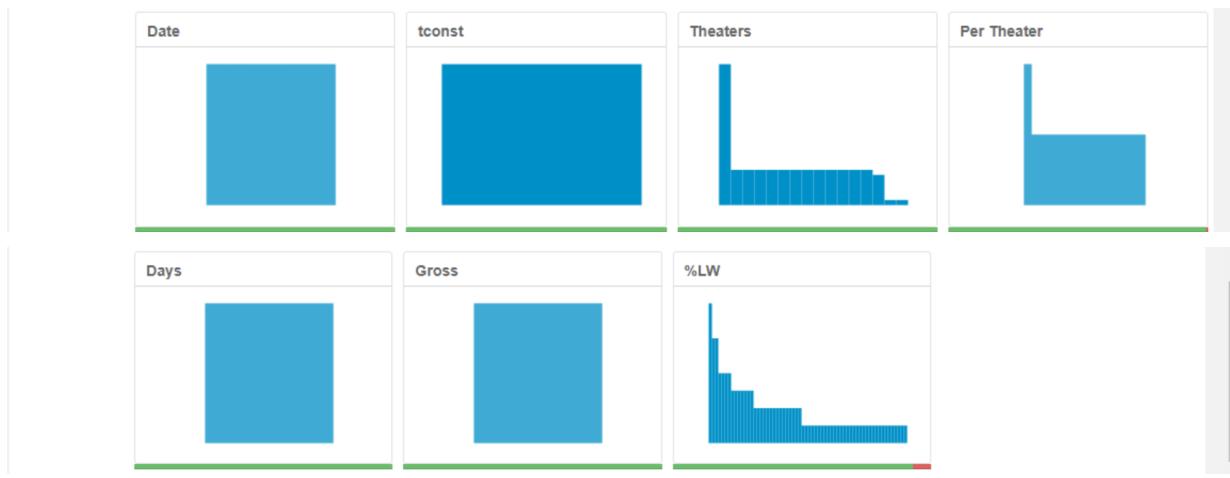
Select variables to view

I₂ I₅ I₉ I₁

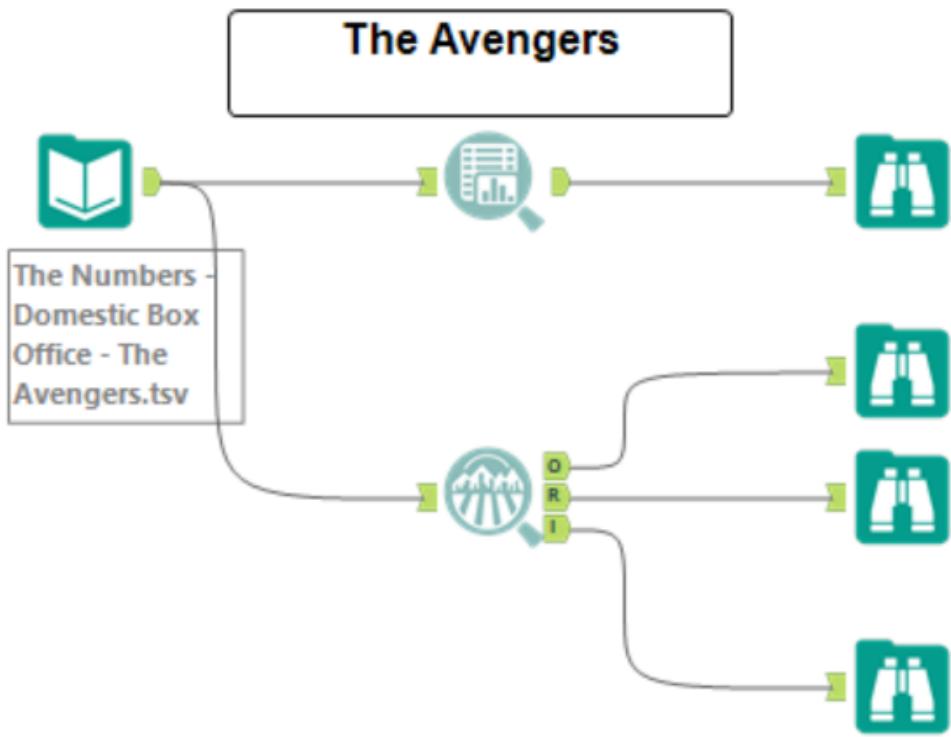
Record

Text





3.8 The Avengers



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

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	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

Field**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 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

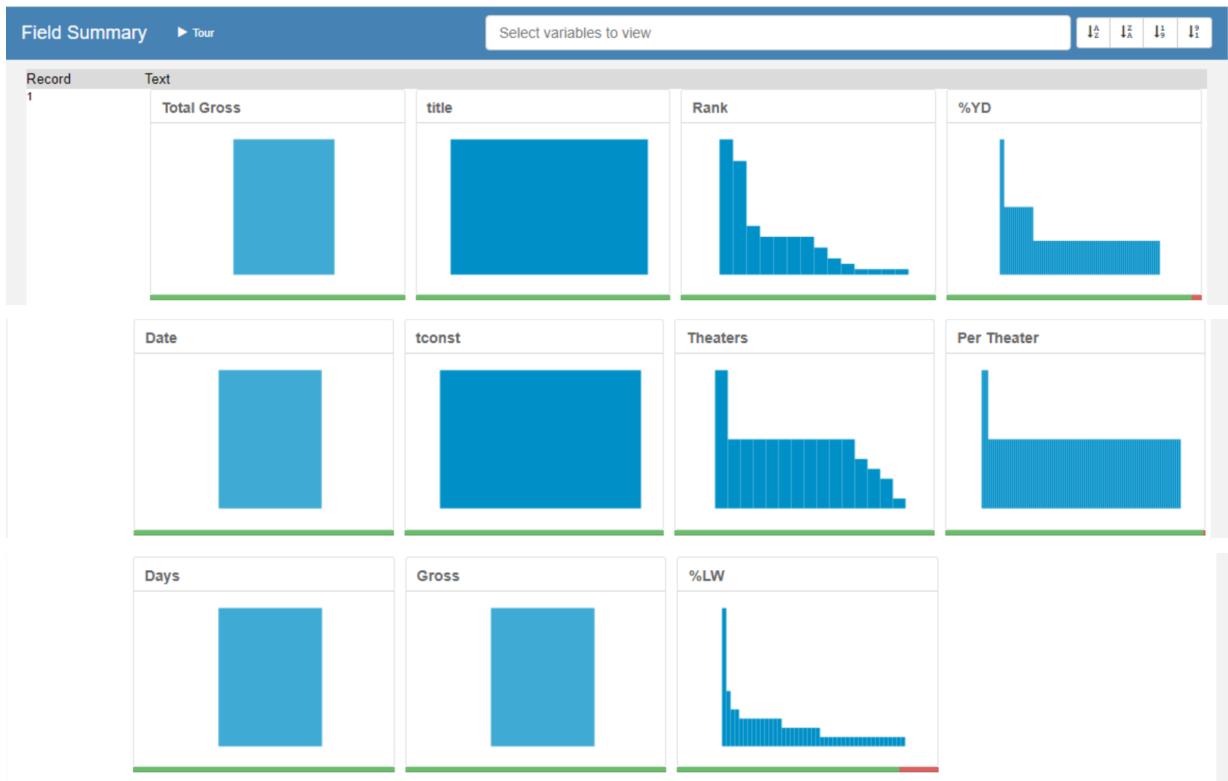
Summary ^

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

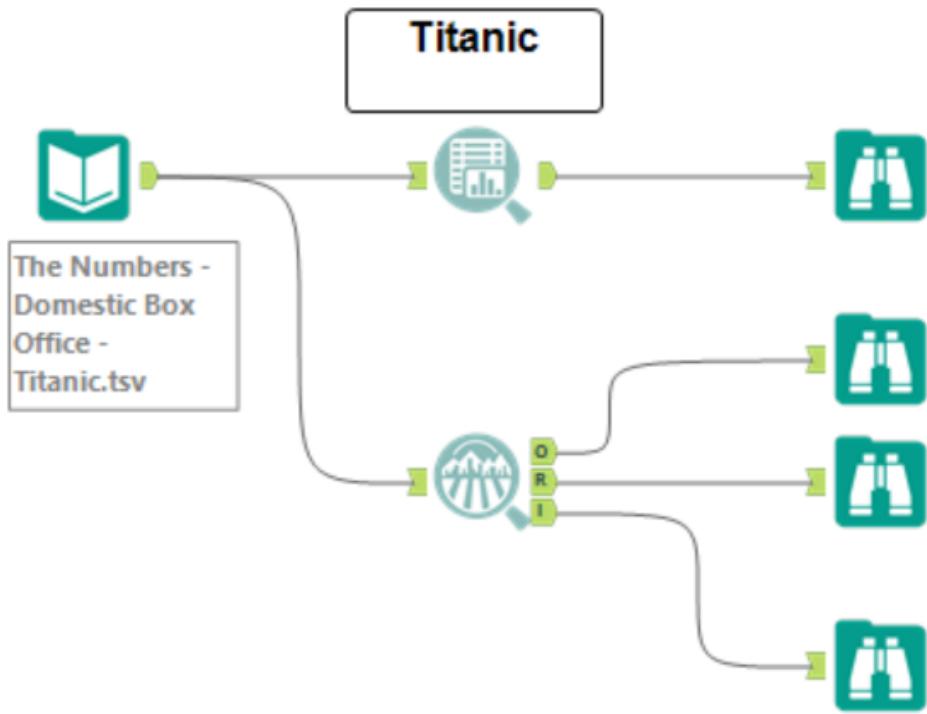
● Ok	209	90.48%
Unique	77	33.33%
● Null	13	5.63%
Not	9	3.90%
● Ok	0	0.00%
● Empty	0	0.00%

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	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%	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	1	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



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	
● 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	
● 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	
● Ok	209	90.48%	
Unique	84	36.36%	
● Null	15	6.49%	
Not	7	3.03%	
● Ok	0	0.00%	
● Empty	0	0.00%	

Report		Profile		Records 1 to 1				
1 of 1 Fields								
1		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	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	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	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	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	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	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	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	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	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

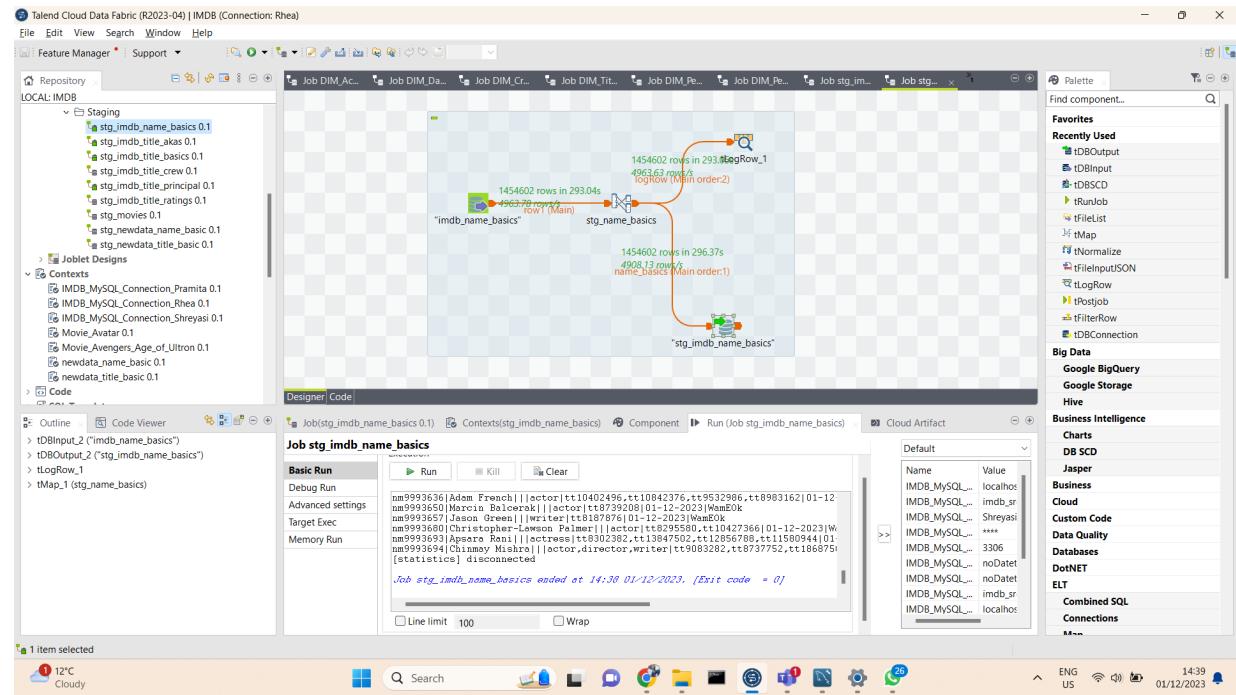
Total Gross	title	Rank	%YD
Date	tconst	Theaters	Per Theater
Days	Gross	%LW	

12 13 14 15

Jobs and Screenshots of Tables:

Stage Tables:

stg_imdb_name_basics:



MySQL Workbench

Schemas: stg_imdb_name_basics

Table: stg_imdb_name_basics

Columns:

- name: varchar(10)
- primaryName: varchar(255)
- birthYear: varchar(4)
- deathYear: varchar(4)
- profession: varchar(255)
- knownForTitles: varchar(1024)
- DL_CreatedDate: datetime
- JobID: varchar(100)

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

Action Output | Message | Duration / Fetch | Time | Action | 13:26.37 | SELECT * FROM imdb_src.movie.stg_imdb_name_basics | 0.015 sec / 0.000 sec | 13:26.45 | SELECT * FROM imdb_src.movie.dim_person | 0.016 sec / 2.328 sec | 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 server version for the right syntax to use near 'dim_person' at line 1 | 0.093 sec | 13:27.22 | SELECT count(*) FROM imdb_src.movie.dim_person | 0.766 sec / 0.000 sec | 13:35.52 | SELECT * FROM imdb_src.movie.stg_imdb_name_basics | 0.156 sec / 3.422 sec | 13:36.07 | SELECT count(*) FROM imdb_src.movie.stg_imdb_name_basics | 1.078 sec / 0.000 sec

Object Info | Session | 13°C Cloudy | ENG US | 13:36 | 01/12/2023

stg_imdb_title_akas

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

Job stg_imdb_title_akas

Designer | Code

Job stg_imdb_title_akas

Basic Run | Run | Kill | Clear

```
tt9316600 Run the job
tt9316600 [2] De la ilusión al desconcierto: cine colombiano 1970-1995[0][N]imdb_src.movie.dim_person
tt9316600 [1] Dankyavar Bankai[N][N][N]
tt9316730 [1] Gunn[N][N][N]
tt9316754 [1] Chico Albuquerque - Revelaciones[BR][S]imdb_src.movie.dim_person
tt9316754 [1] Chico Albuquerque - Revelaciones[BR][S]imdb_src.movie.dim_person
[statistics] disconnected
```

Job stg_imdb_title_akas ended at 14:50 01/12/2023. [Exit code = 0]

Outline | Code Viewer | Find component... | Favorites | Recently Used | Big Data | Business Intelligence | Jasper | Business | Cloud | Custom Code | Data Quality | Databases | DotNET | ELT | Combined SQL | Connections

12°C Cloudy | ENG US | 14:50 | 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_imdb_title_akas

SCHEMAS: Filter objects

Table: stg_imdb_title_akas

Columns:

- titleId: varchar(10)
- ordering: int
- title: varchar(1024)
- region: varchar(255)
- language: varchar(255)
- types: varchar(255)
- attributes: varchar(1024)
- isOriginalTitle: varchar(255)
- DT_CreatedDate: datetime
- pid: varchar(100)

sb_title_akas x

Output:

Action Output:

#	Time	Action	Message	Duration / Fetch
11	03:52:52	SELECT * FROM `imdb_src_movie.dim_actor_profession` LIMIT 0, 1000	1000 row(s) returned	0.063 sec / 0.015 sec
12	03:53:01	SELECT * FROM `imdb_src_movie.dim_actor_profession` LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec
13	03:53:25	SELECT * FROM `imdb_src_movie.dim_actor_knowfor` LIMIT 0, 1000	0 row(s) returned	0.000 sec / 0.000 sec
14	03:53:53	SELECT * FROM `imdb_src_movie.dim_actor_profession` LIMIT 0, 1000	1000 row(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 row(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 row(s) returned	0.906 sec / 0.000 sec

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

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

Read Only Context Help Snippets

Object Info Session

6°C Mostly clear

03:56 ENG US 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_imdb_title_akas

SCHEMAS: Filter objects

Table: stg_imdb_title_akas

Columns:

- titleId: varchar(10)
- ordering: int
- title: varchar(1024)
- region: varchar(255)
- language: varchar(255)
- types: varchar(255)
- attributes: varchar(1024)
- isOriginalTitle: varchar(255)
- DT_CreatedDate: datetime
- pid: varchar(100)

sb_title_akas x

Output:

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 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
29	13:35:50	SELECT * FROM `imdb_src_movie.stg_imdb_name_basics`	1454602 row(s) returned	0.156 sec / 3.422 sec
30	13:36:07	SELECT count(*) FROM `imdb_src_movie.stg_imdb_name_basics`	1 row(s) returned	1.078 sec / 0.000 sec
31	13:36:47	SELECT * FROM `imdb_src_movie.stg_imdb_title_akas`	2612798 row(s) returned	0.203 sec / 5.375 sec
32	13:37:07	SELECT count(*) FROM `imdb_src_movie.stg_imdb_title_akas`	1 row(s) returned	1.031 sec / 0.000 sec

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

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

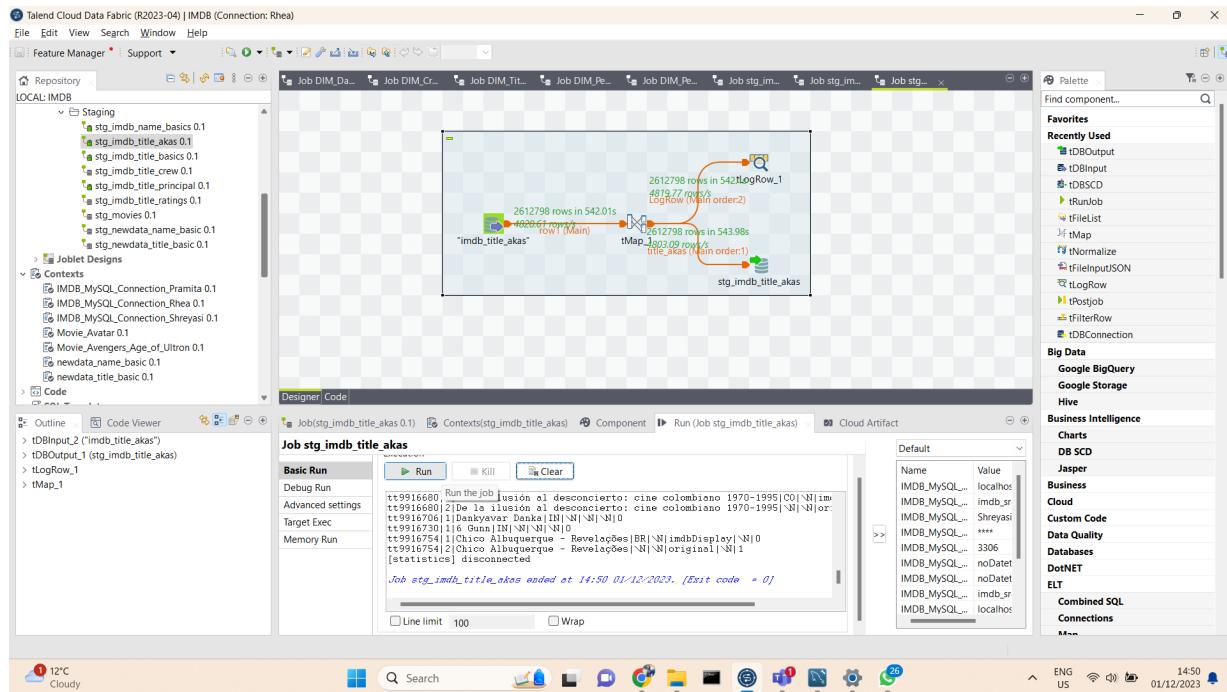
Read Only Context Help Snippets

Object Info Session

13°C Cloudy

13:37 ENG US 01/12/2023

stg_imdb_title_basics



The screenshot shows the MySQL Workbench interface. The left sidebar displays the database schema, including tables like 'imdb_title_principals', 'stg_imdb_title_ratings', and 'stg_imdb_title_basics'. The main area shows the results of a SQL query: 'SELECT * FROM imdb_src_movie.stg_imdb_title_basics;'. The results grid lists various movies with their titles, years, genres, and other metadata. The bottom panel shows the execution history of the query.

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: principal stg_imdb_title_ratings stg_movies stg_newdata_name_basic stg_newdata_title_basic dim_date dim_crew dim_person dim_titleregion stg_imdb_title

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: □

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

Result 2 x

Action Output

Time Action Message Duration / Fetch

- 49 13:47:00 SELECT * FROM imdb_src_movie.dim_person 1454602 row(s) returned 0.031 sec / 1563 sec
- 50 13:47:09 SELECT count(*) FROM imdb_src_movie.dim_person 1 row(s) returned 0.266 sec / 0.000 sec
- 51 13:47:46 SELECT * FROM imdb_src_movie.dim_titleregion 2612798 row(s) returned 0.140 sec / 4.344 sec
- 52 13:48:05 SELECT count(*) FROM imdb_src_movie.dim_titleregion 1 row(s) returned 0.937 sec / 0.000 sec
- 53 14:51:51 SELECT * FROM imdb_src_movie.stg_imdb_title_basics 607423 row(s) returned 0.093 sec / 2.157 sec
- 54 14:52:43 SELECT count(*) FROM imdb_src_movie.stg_imdb_title_basics 1 row(s) returned 0.312 sec / 0.000 sec

Object Info Session

12°C Cloudy ENG US 14:52 01/12/2023

stg_imdb_title_crew

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

File Edit View Search Window Help

Repository LOCAL: IMDB

Job: Job_stg_imdb_title_crew_0.1

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

Contexts: IMDB(MySQL_Connection_Pramita 0.1, IMDB(MySQL_Connection_Rhea 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(stg_imdb_title_crew 0.1), Contexts(stg_imdb_title_crew), Component, Run (Job_stg_imdb_title_crew), Cloud Artifact

Designer | Code

Job stg_imdb_title_crew

Basic Run: Run, Kill, Clear

Advanced settings: Target Exec: Memory Run, Line limit: 100, Wrap

Component: tDBInput_2 ("imdb_title_crew"), tLogRow_1, tMap_1, tDBOutput_2 (stg_imdb_title_crew), stg_imdb_title_crew

Cloud Artifact: Default

Name	Value
IMDB(MySQL_	localhos
IMDB(MySQL_	imdb_sr
IMDB(MySQL_	Shreyasi
IMDB(MySQL_	****
IMDB(MySQL_	3306
IMDB(MySQL_	noDate
IMDB(MySQL_	noDate
IMDB(MySQL_	imdb_sr
IMDB(MySQL_	localhos

12°C Cloudy ENG US 15:03 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: lim_actor_knowfor dim_actor_profession dim_actor_profession dim_actor_knowfor dim_actor_profession stg_imdb_name_basics stg_imdb_title_akas stg_imdb_title

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

Table: stg_imdb_title_crew

Columns:

- tcost varchar(10)
- directors varchar(4000)
- writers varchar(4000)
- DL_CreatedDate datetime
- JobID varchar(100)

Object Info Session

6°C Mostly clear

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

1 • SELECT * FROM imdb_src_movie.stg_imdb_title_crew

JobID	tcost	directors	writers	DL_CreatedDate	JobID
b7fj74	t0000592 nm0653413,nm0657268,nm0675388	2023-11-30 01:18:58			t0000574 nm06534579
b7fj74	t0000591 nm0141150	2023-11-30 01:18:58			t0000615 nm0533958
b7fj74	t0000630 nm0143333	2023-11-30 01:18:58			t0000675 nm0194088
b7fj74	t000079 nm0911767,nm0877783	2023-11-30 01:18:58			t000092 nm0081288,nm0691994
b7fj74	t0000814 nm0185426	2023-11-30 01:18:58			t0000838 nm0017074
b7fj74	t0000842 nm0005717	2023-11-30 01:18:58			t0000846 nm0023107
b7fj74	t0000850 nm0595220,nm0063413	2023-11-30 01:18:58			t0000859 nm0005717
b7fj74	t0000862 nm0879467	2023-11-30 01:18:58			t0000862 nm0957652

Result Grid Form Editor Field Types Query Stats

Action Output

#	Time	Action	Message	Duration / Fetch
12	03:53:01	SELECT * FROM imdb_src_movie.dim_actor_profession LIMIT 0, 1000	1000 row(s) returned	0.000 sec / 0.000 sec
13	03:53:25	SELECT * FROM imdb_src_movie.dim_actor_knowfor LIMIT 0, 1000	0 row(s) returned	0.000 sec / 0.000 sec
14	03:53:53	SELECT * FROM imdb_src_movie.dim_actor_profession LIMIT 0, 1000	1000 row(s) returned	0.016 sec / 0.047 sec
15	03:54:40	SELECT * FROM imdb_src_movie.stg_imdb_name_basics LIMIT 0, 1000	1000 row(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 row(s) returned	0.906 sec / 0.000 sec
17	03:57:01	SELECT * FROM imdb_src_movie.stg_imdb_title_crew LIMIT 0, 1000	1000 row(s) returned	0.171 sec / 0.016 sec

Read Only Context Help Snippets

Output

ENG US 03:57 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_newdata_name_basic stg_newdata_title_basic dim_actor_profession dim_titlerregion dim_person stg_imdb_name_basics stg_imdb_title_akas stg_imdb_title

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

Table: stg_imdb_title_crew

Columns:

- tcost varchar(10)
- directors varchar(4000)
- writers varchar(4000)
- DL_CreatedDate datetime
- JobID varchar(100)

Object Info Session

13°C Cloudy

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

1 • SELECT count(*) FROM imdb_src_movie.stg_imdb_title_crew

count(*)
607423

Result Grid Form Editor Field Types Query Stats

Action Output

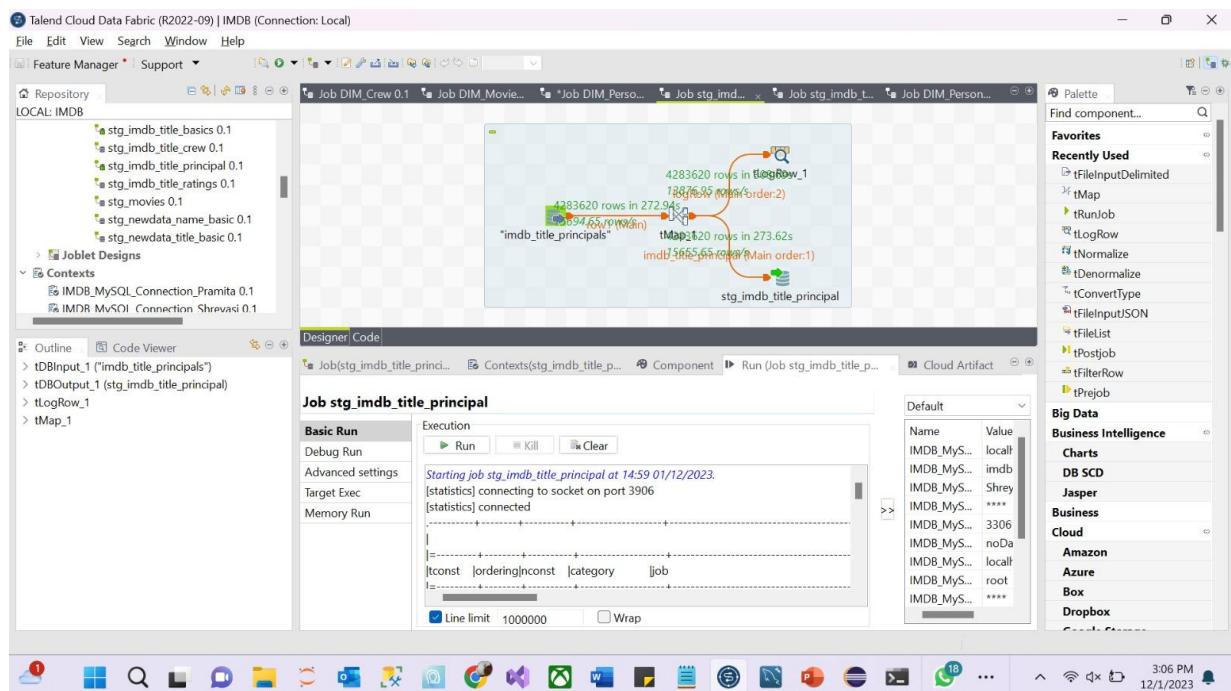
#	Time	Action	Message	Duration / Fetch
29	13:35:52	SELECT * FROM imdb_src_movie.stg_imdb_name_basics	1454602 row(s) returned	0.156 sec / 3.422 sec
30	13:36:07	SELECT count(*) FROM imdb_src_movie.stg_imdb_name_basics	1 row(s) returned	1.078 sec / 0.000 sec
31	13:36:47	SELECT * FROM imdb_src_movie.stg_imdb_title_akas	2612798 row(s) returned	0.203 sec / 5.375 sec
32	13:37:07	SELECT count(*) FROM imdb_src_movie.stg_imdb_title_akas	1 row(s) returned	1.031 sec / 0.000 sec
33	13:37:28	SELECT * FROM imdb_src_movie.stg_imdb_title_crew	607423 row(s) returned	0.157 sec / 0.937 sec
34	13:37:49	SELECT count(*) FROM imdb_src_movie.stg_imdb_title_crew	1 row(s) returned	0.156 sec / 0.000 sec

Result Grid Form Editor Field Types Query Stats

Output

ENG US 13:37 01/12/2023

stg_imdb_title_principal



The screenshot shows MySQL Workbench. The left sidebar shows the 'Schemas' tree with the 'imdb_src_movie.stg_imdb_title_principal' table selected. The main area displays the results of a SELECT query: 'SELECT * FROM imdb_src_movie.stg_imdb_title_principal;'. The results grid shows 10 rows of data, each with columns: tconst, ordering, nconst, category, job, characters, DL_CreatedDate, and JobId. The 'Output' pane at the bottom shows the execution history with 18 actions, including various SELECT statements and their execution times (e.g., 0.000 sec / 0.000 sec, 0.016 sec / 0.047 sec, etc.). The system status bar at the bottom indicates 'ENG US 03:58 01/12/2023'.

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Schemas: basic stg_newdata_title_basic dim_actor_profession dim_tilregion dim_person stg_imdb_name_basics stg_imdb_title_akas stg_imdb_title_crew stg_imdb_title_principal

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 2 | Read Only | Context Help | Snippets

Action Output | Duration / Fetch

Object Info Session

13°C Cloudy 13:40 01/12/2023

stg_imdb_title_ratings

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

File Edit View Search Window Help

Repository LOCAL:IMDB Job Designs Standard IMDB_Jobs Load_Dimensions Load_Fact Staging stg_imdb_name_basics 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 JOblet Designs Contexts Code SQL Templates Metadata

Job stg_imdb_title_ratings 0.1

Find component... Palette

Recently Used tBSCD tMap tDBInput tDBOutput tNormalize tFileInputJSON tFileList tLogRow tPostgres tRunJob tFilterRow tDBConnection

Big Data Google BigQuery Google Storage Hive Business Intelligence Charts DB SCDF Jasper Business Cloud Custom Code Data Quality Databases DotNET ELT

10°C Light rain 17:00 01/12/2023

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

- Navigator:** Shows tables like `t1_actor_profession`, `dim_actor_knowfor`, etc., and schemas like `imdb_title_ratings`.
- Schemas:** Shows the schema for `imdb_title_ratings` with columns: `tconst`, `averageRating`, `numVotes`, `D1_CreatedDate`, and `JobID`.
- Table:** `stg_imdb_title_ratings` is selected.
- Columns:** `tconst`, `averageRating`, `numVotes`, `D1_CreatedDate`, and `JobID` are listed.
- Output:** The results grid shows data from the query: `SELECT * FROM imdb_src_movie.stg_imdb_title_ratings;` with 1000 rows returned.
- Action Output:** The execution history shows the following queries and their results:

Time	Action	Message	Duration / Fetch
14:03:53.93	SELECT * FROM `imdb_src_movie.dim_actor_profession` LIMIT 0, 1000	1000 row(s) returned	0.016 sec / 0.047 sec
14:03:54.50	SELECT * FROM `imdb_src_movie.stg_imdb_name_basics` LIMIT 0, 1000	1000 row(s) returned	0.075 sec / 0.000 sec
14:03:56.24	SELECT * FROM `imdb_src_movie.stg_imdb_title_akas` LIMIT 0, 1000	1000 row(s) returned	0.095 sec / 0.000 sec
14:03:57.01	SELECT * FROM `imdb_src_movie.stg_imdb_title_crew` LIMIT 0, 1000	1000 row(s) returned	0.171 sec / 0.016 sec
14:03:57.45	SELECT * FROM `imdb_src_movie.stg_imdb_title_principal` LIMIT 0, 1000	1000 row(s) returned	0.125 sec / 0.031 sec
14:03:58.17	SELECT * FROM `imdb_src_movie.stg_imdb_title_ratings` LIMIT 0, 1000	1000 row(s) returned	0.105 sec / 0.016 sec

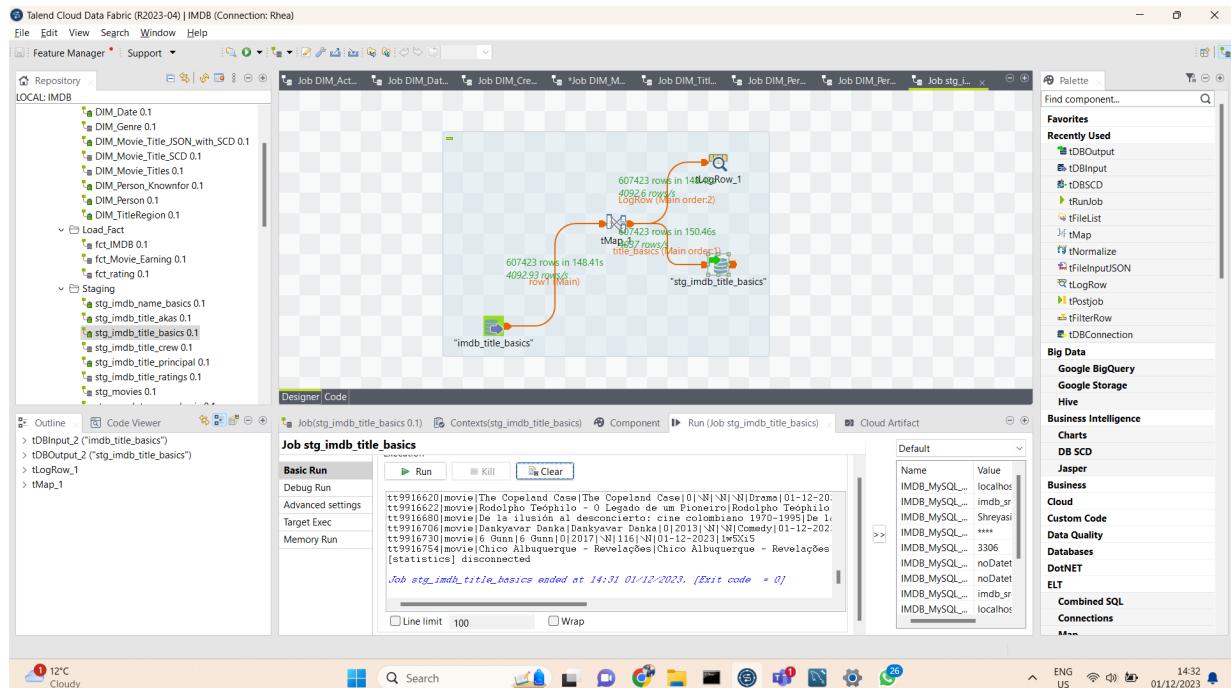
- SQL Additions:** A toolbar with various icons for managing the connection and session.
- Information:** Shows the current session information.
- Object Info:** Shows the session status as "Mostly clear".

The screenshot shows the MySQL Workbench interface. A query is being run against the 'imdb_src_movie' database:

```
SELECT count(*) FROM imdb_src_movie.stg_imdb_title_ratings;
```

The result grid shows a single row with the value 277171. The 'Information' pane on the left displays details for the 'stg_imdb_title_ratings' table, including columns like 'JobID', 'TitleID', 'averageRating', 'numVotes', 'DL_CreatedDate', and 'JobID'. The 'Object Info' and 'Session' tabs are also visible at the bottom.

stg_title_basics



MySQL Workbench

Query 1

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

Result Grid

tconst	titleType	primaryTitle	originalTitle	isAdult	startYear	endYear	runtimeMinutes	genres	DL_CreatedDate
tt0000502	movie	Bohemios	Bohemios	0	1905	N	100	W	2023-11-25 04:54
tt0000574	movie	The Story of the Kelly Gang	The Story of the Kelly Gang	0	1906	N	70	Action,Adventure,Biography	2023-11-25 04:54
tt0000591	movie	The Prodigal Son	L'enfant prodige	0	1907	N	90	Drama	2023-11-25 04:54
tt0000615	movie	Robbery Under Arms	Robbery Under Arms	0	1907	N	N	Drama	2023-11-25 04:54
tt0000630	movie	Hamlet	Amleto	0	1908	N	N	Drama	2023-11-25 04:54
tt0000675	movie	Don Quijote	Don Quijote	0	1908	N	N	Drama	2023-11-25 04:54
tt0000679	movie	The Fairylogue and Radio-Plays	The Fairylogue and Radio-Plays	0	1908	N	120	Adventure,Fantasy	2023-11-25 04:54
tt0000680	movie	Andréenne Weber	Andréenne Weber	0	1908	N	N	N	2023-11-25 04:54

Object Info

Session

Action Output

#	Time	Action	Message	Duration / Fetch
5	14:08:18	SELECT count(*) FROM `imdb_src_movie.dim_movie_title_mysql` LIMIT 0, 500	1 row(s) returned	0.250 sec / 0.000 sec
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

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: pet_care_connect - Schema: dim_movie_title_mysql dim_crew stg_imdb_title_basics

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

Result Grid:

count(*)
607423

Result 2 (Read Only):

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

Information:

Table: stg_imdb_title_basics

Columns:

- tconst varchar(255) PK
- primaryTitle varchar(255)
- originalTitle varchar(255)

Object Info Session

stg_movies

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

File Edit View Search Window Help

Feature Manager Support 100% |

Repository LOCAL: IMDB

Job DIM_Crew 0.1 Job DIM_Genre 0.1 Job DIM_Movie_Titles 0.1 Job stg_movies 0.1

Designer | Code

Job stg_movies

Basic Run Execution

Run Kill Clear

tt0120338|Titanic|2017-12-10|-\$20,646|-24%-83%|75|\$275|\$659,328,801|7,297
 tt0120338|Titanic|2017-12-11|-\$7,419|-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|-\$13,555|0.3|-69%|75|\$181|\$659,360,167|7,300
 tt0120338|Titanic|2017-12-14|-\$3,777|-72%-86%|75|\$50|\$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...	localt
IMDB.MyS...	****
IMDB.MyS...	localt
IMDB.MyS...	root
IMDB.MyS...	****
IMDB.MyS...	nnDn

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

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: n_actor_knowsfor dim_actor_profession stg_imdb_name_basics stg_imdb_title_akas stg_imdb_title_crew stg_imdb_title_principal stg_imdb_title_ratings stg_movies SQLAdditions

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

Table: stg_movies

Columns:

- tconst: varchar(100)
- title: varchar(100)
- Date: varchar(100)
- Rank: varchar(100)
- Gross: varchar(100)
- VD: varchar(100)
- LW: varchar(100)
- Theaters: varchar(100)
- Per_Theater: varchar(100)
- Total_Gross: varchar(100)
- Days: varchar(100)
- DL_CreatedDate: datetime
- JobID: varchar(100)

Object Info Session

6°C Mostly clear

Result Grid | Filter Rows: Limit to 1000 rows | Export: Wrap Cell Content:

1 • SELECT * FROM imbd_src_movie.stg_movies;

tconst	title	Date	Rank	Gross	VD	LW	Theaters	Per_Theater	Total_Gross	Days	DL_CreatedDate	JobID
t0120338	Titanic	Dec 19, 1997	1	\$8,458,214	2,674	\$3,238	\$8,458,214	\$8,458,214	\$8,458,214	1	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 20, 1997	1	\$10,472,013	0.23	6,274	\$10,472,013	\$10,472,013	\$10,472,013	2	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 21, 1997	1	\$6,307,304	-13%	2,674	\$3,481	\$20,638,131	\$20,638,131	3	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 22, 1997	1	\$5,578,212	-40%	2,674	\$2,086	\$34,216,943	\$4,216,943	4	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 23, 1997	1	\$6,003,119	0.00	2,674	\$2,245	\$40,219,462	\$5,219,462	5	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 24, 1997	1	\$3,571,345	-41%	2,674	\$1,336	\$43,790,807	\$3,790,807	6	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 25, 1997	1	\$9,178,529	1.57	2,674	\$3,493	\$52,969,336	\$7,969,336	7	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 26, 1997	1	\$12,122,294	0.32	0.4	2,711	\$4,472	\$65,091,634	8	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 27, 1997	1	\$12,466,455	0.03	0.17	2,711	\$4,598	\$77,558,089	9	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 29, 1997	1	\$7,942,920	-27%	0.42	2,711	\$2,930	\$96,367,529	11	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 30, 1997	1	\$8,012,909	0.01	0.33	2,711	\$2,956	\$104,380,438	12	2023-11-30 15:41:32	410w11
t0120338	Titanic	Dec 31, 1997	1	\$8,213,735	0.03	1.3	2,711	\$3,030	\$112,594,173	13	2023-11-30 15:41:32	410w11
t0120338	Titanic	Jan 1, 1998	1	\$11,558,520	0.41	0.26	2,711	\$4,264	\$124,152,693	14	2023-11-30 15:41:32	410w11
t0120338	Titanic	Jan 2, 1998	1	\$11,698,861	0.01	-3%	2,727	\$4,290	\$135,851,554	15	2023-11-30 15:41:32	410w11

Action Output

- Time Action Message Duration / Fetch
- 03:54:50 SELECT * FROM imbd_src_movie.stg_imdb_name_basics LIMIT 0, 1000 row(s) returned 0.078 sec / 0.000 sec
- 03:56:24 SELECT * FROM imbd_src_movie.stg_imdb_title_akas LIMIT 0, 1000 row(s) returned 0.906 sec / 0.000 sec
- 03:57:01 SELECT * FROM imbd_src_movie.stg_imdb_title_crew LIMIT 0, 1000 row(s) returned 0.171 sec / 0.016 sec
- 03:57:45 SELECT * FROM imbd_src_movie.stg_imdb_title_principal LIMIT 0, 1000 row(s) returned 0.125 sec / 0.031 sec
- 03:58:17 SELECT * FROM imbd_src_movie.stg_imdb_title_ratings LIMIT 0, 1000 row(s) returned 0.109 sec / 0.016 sec
- 03:58:46 SELECT * FROM imbd_src_movie.stg_movies LIMIT 0, 1000 row(s) returned 0.203 sec / 0.000 sec

Result Grid | Filter Rows: Limit to 1000 rows | Export: Wrap Cell Content:

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

count(*)
265

Action Output

- # Time Action Message Duration / Fetch
- 35 13:38:20 SELECT * FROM imbd_src_movie.stg_imdb_title_principal 42033620 row(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 13:40:37 SELECT * FROM imbd_src_movie.stg_imdb_title_ratings 277171 row(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
- 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

Result Grid | Filter Rows: Limit to 1000 rows | Export: Wrap Cell Content:

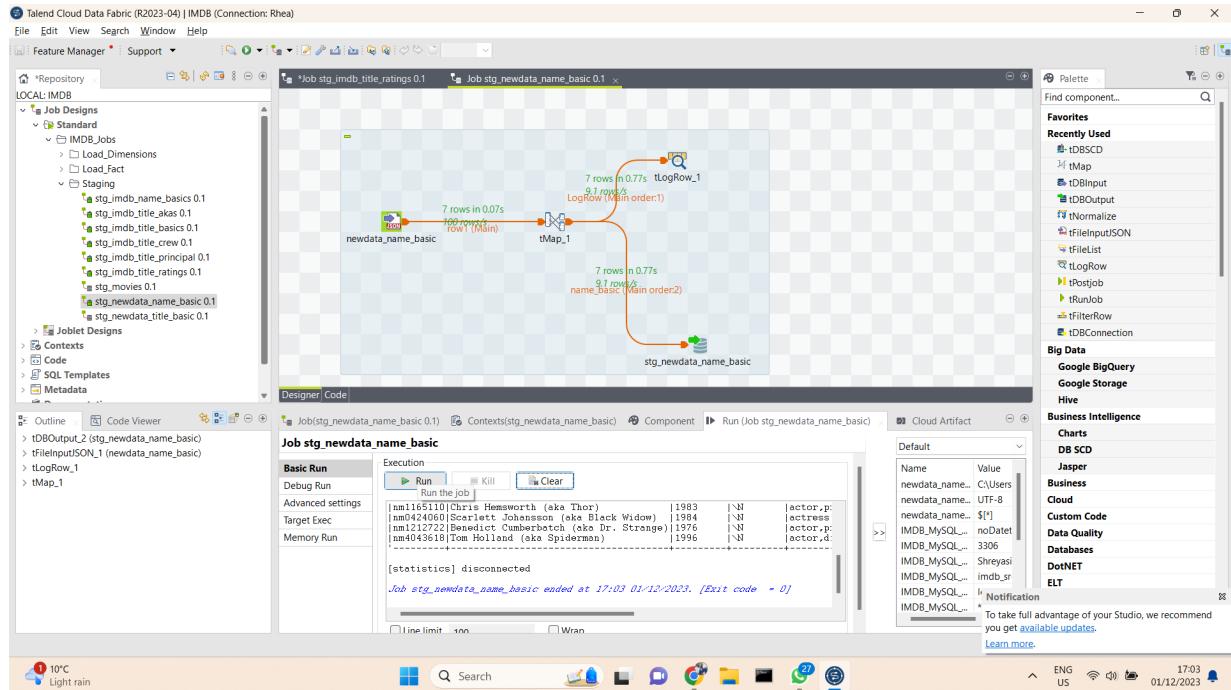
1 • SELECT count(*) FROM imbd_src_movie.stg_movies;

count(*)
265

Action Output

- # Time Action Message Duration / Fetch
- 35 13:38:20 SELECT * FROM imbd_src_movie.stg_imdb_title_principal 42033620 row(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 13:40:37 SELECT * FROM imbd_src_movie.stg_imdb_title_ratings 277171 row(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
- 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

stg_newdata_name_basic



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

Filter objects

dim_actor_profession stg_imdb_name_basics stg_imdb_title_akas stg_imdb_title_crew stg_imdb_title_principal stg_imdb_title_ratings stg_movies stg_newdata_name_basic

1 • SELECT * FROM imdb_src_movie.stg_newdata_name_basic;

Result Grid

nconst	primaryName	birthYear	deathYear	primaryProfession	knownForTitles	DL_CreatedDate	JobID
nm0000375	Robert Downey Jr. (aka IRON MAN)	1965	N	actor,producer,soundtrack	tt0371746,tt1415796,tt1300854,tt0988045	2023-11-30 15:55:23	ifrc7u
nm0262635	Chris Evans (aka Captain America)	1981	N	actor,producer,soundtrack	tt2395427,tt0458339,tt3498820,tt0848228	2023-11-30 15:55:23	ifrc7u
nm0749263	Mark Ruffalo (aka Hulk)	1967	N	actor,producer,soundtrack	tt0848228,tt1100089,tt1895587,tt0842926	2023-11-30 15:55:23	ifrc7u
nm1165110	Chris Hemsworth (aka Thor)	1983	N	actor,producer,soundtrack	tt0848228,tt0800369,tt1735898,tt3501632	2023-11-30 15:55:23	ifrc7u
nm0424060	Scarlett Johansson (aka Black Widow)	1984	N	actress,soundtrack,producer	tt1798709,tt0335268,tt0848228,tt162346	2023-11-30 15:55:23	ifrc7u
nm1212722	Benedict Cumberbatch (aka Dr. Strange)	1976	N	actor,producer,soundtrack	tt1408101,tt2084970,tt1211837,tt1322269	2023-11-30 15:55:23	ifrc7u
nm043618	Tom Holland (aka Spiderman)	1996	N	actor,director,producer	tt1649419,tt6320629,tt2250912,tt3498820	2023-11-30 15:55:23	ifrc7u

SQL Additions

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

Table: stg_newdata_name_basic

Columns:

- nconst: varchar(9)
- primaryName: varchar(38)
- birthYear: int
- deathYear: varchar(2)
- primaryProfession: varchar(27)
- knownForTitles: varchar(39)
- DL_CreatedDate: datetime
- JobID: varchar(100)

Object Info Session

6°C Mostly clear

03:59 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: tileregion dim_person stg_imdb_name_basics stg_imdb_title_akas stg_imdb_title_crew stg_imdb_title_principal stg_imdb_title_ratings stg_movies stg_newdata_...

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

Table: stg_newdata_name_basic

Columns:

- name: varchar(9)
- primaryName: varchar(38)
- birthYear: int
- deathYear: varchar(2)
- profession: varchar(100)
- knownForTitles: varchar(39)
- DL_CreatedDate: datetime
- JobID: varchar(100)

Result Grid

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

Result 2

Action Output

Time	Action	Message	Duration / Fetch
37 13:40:36	SELECT * FROM imdb_src_movie stg_imdb_title_ratings	27711 row(s) returned	0.187 sec / 0.484 sec
38 13:40:47	SELECT count(*) FROM imdb_src_movie stg_imdb_title_ratings	1 row(s) returned	0.015 sec / 0.000 sec
39 13:41:07	SELECT * FROM imdb_src_movie stg_movies	269 row(s) returned	0.109 sec / 0.000 sec
40 13:41:22	SELECT count(*) FROM imdb_src_movie stg_movies	1 row(s) returned	0.000 sec / 0.000 sec
41 13:41:43	SELECT * FROM imdb_src_movie stg_newdata_name_basic	7 row(s) returned	0.110 sec / 0.000 sec
42 13:41:52	SELECT count(*) FROM imdb_src_movie stg_newdata_name_basic	1 row(s) returned	0.000 sec / 0.000 sec

Object Info Session

13°C Cloudy 13:41 ENG US 01/12/2023

stg_newdata_title_basic

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

File Edit View Search Window Help

Feature Manager Support

Repository LOCAL: IMDB

Job Designs Standard IMDB_Jobs Load_Dimensions Load_Fact Staging stg_imdb_name_basics_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 JOblet Designs Contexts Code SQL Templates Metadata

Job stg_imdb_title_ratings_0.1 Job stg_newdata_name_basic_0.1 Job stg_newdata_title_basic_0.1

Designer Code

Job stg_newdata_title_basic

Basic Run Execution

```
tLogRow_1
  10 rows in 0.71s
  LogRow (Main order:1)
  10.12 rows/s
  10 rows in 0.74s
  LogRow (Main order:2)
  13.53 rows/s
  stg_newdata_title_basic
```

Advanced settings Target Exec Memory Run

Cloud Artifact Default

Name	Value
newdata_title_...	C:/Users
newdata_title_...	UTF-8
newdata_title_...	\$data/*
IMDB_MySQL_...	noDatet
IMDB_MySQL_...	3306
IMDB_MySQL_...	Shreyasi
IMDB_MySQL_...	imdb_sr
IMDB_MySQL_...	b_Notification

[statistics] disconnected

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

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

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_imdb_name_basics 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

Limit to 1000 rows SQLAdditions

1 • SELECT * FROM imbd_src_movie.stg_newdata_title_basic;

tconst	titleType	primaryTitle	originalTitle	isAdult	startYear	endYear	runtimeMinutes	genres
t0120338	movie	Titanic (1997)	Titanic	0	1997	W	194	Drama,Romance
t049549	movie	Avatar (2009)	Avatar	0	2009	W	162	Action,Adventure,Fant
t0848228	movie	The Avengers (2012)	The Avengers	0	2012	W	143	Action,Adventure,Sci-F
t10872600	movie	Spider-Man: No Way Home (2021)	Spider-Man: No Way Home	0	2021	W	148	Action,Adventure,Fant
t1825683	movie	Black Panther (2018)	Black Panther	0	2018	W	134	Action,Adventure,Sci-F
t2395427	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
t4154756	movie	Avengers: Infinity War (2018)	Avengers: Infinity War	0	2018	W	149	Action,Adventure,Sci-F
t4154796	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

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

Table: stg_newdata_title_basic

Object Info Session

6°C Mostly clear 04:00 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_imdb_name_basics 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

SQLAdditions

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

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

count(*)
10

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

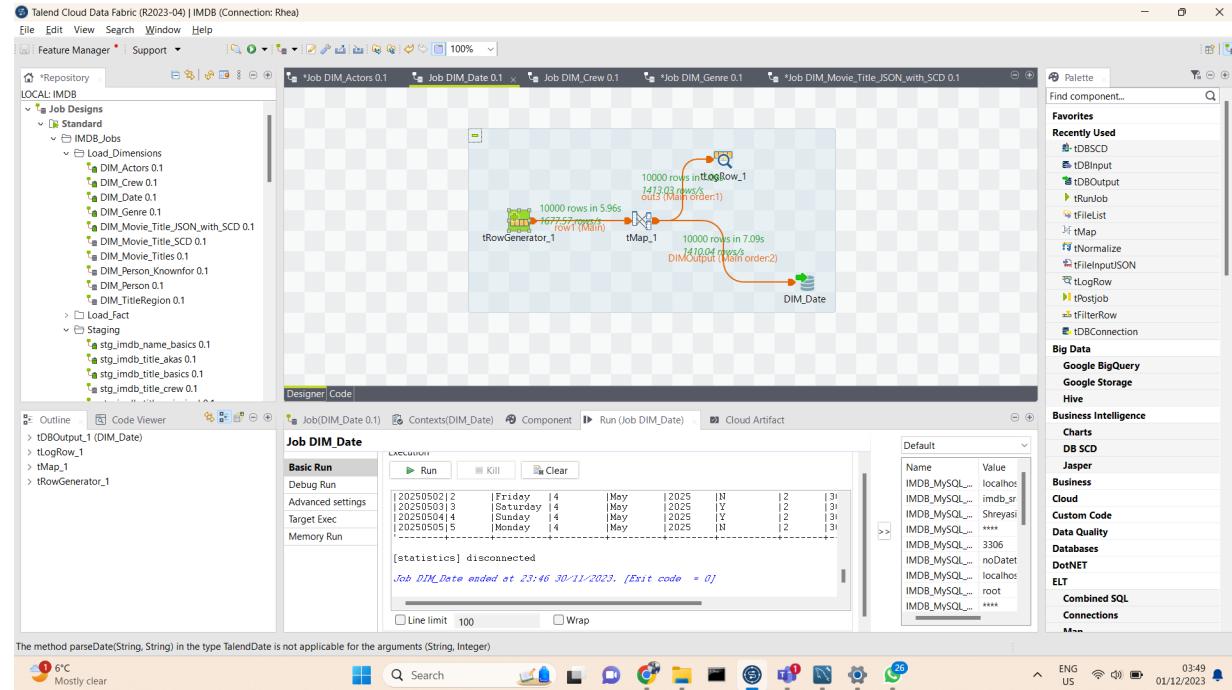
Table: stg_newdata_title_basic

Object Info Session

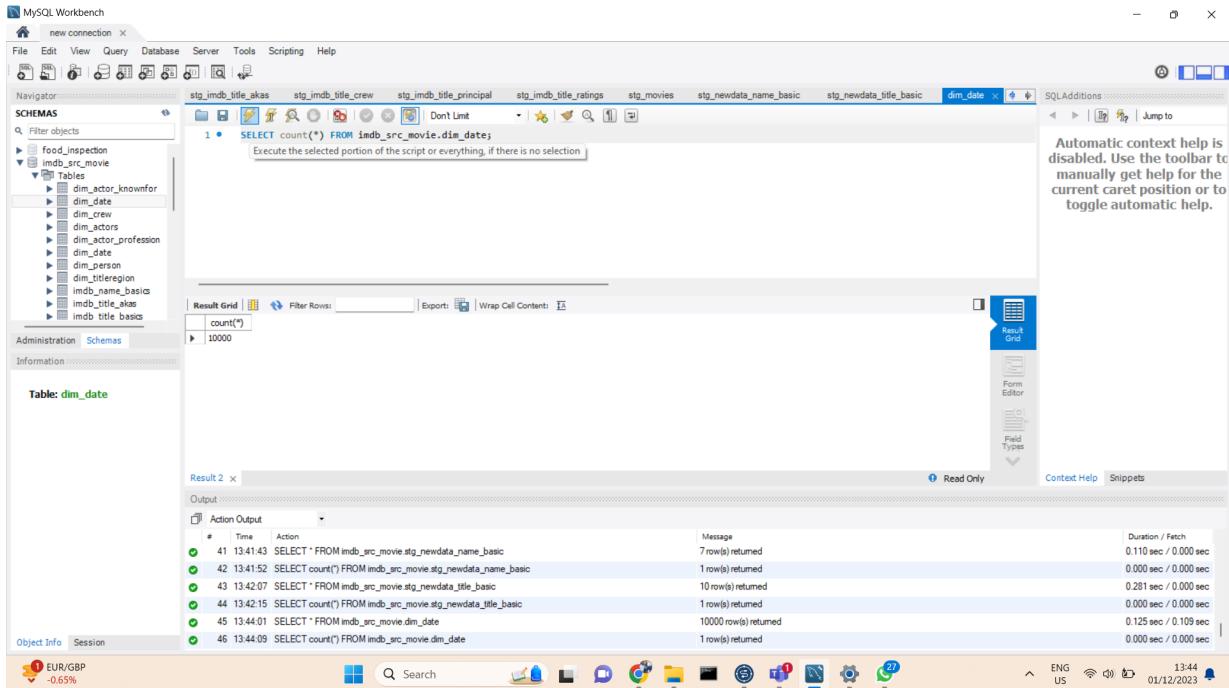
13°C Cloudy 13:42 01/12/2023

Dim Tables:

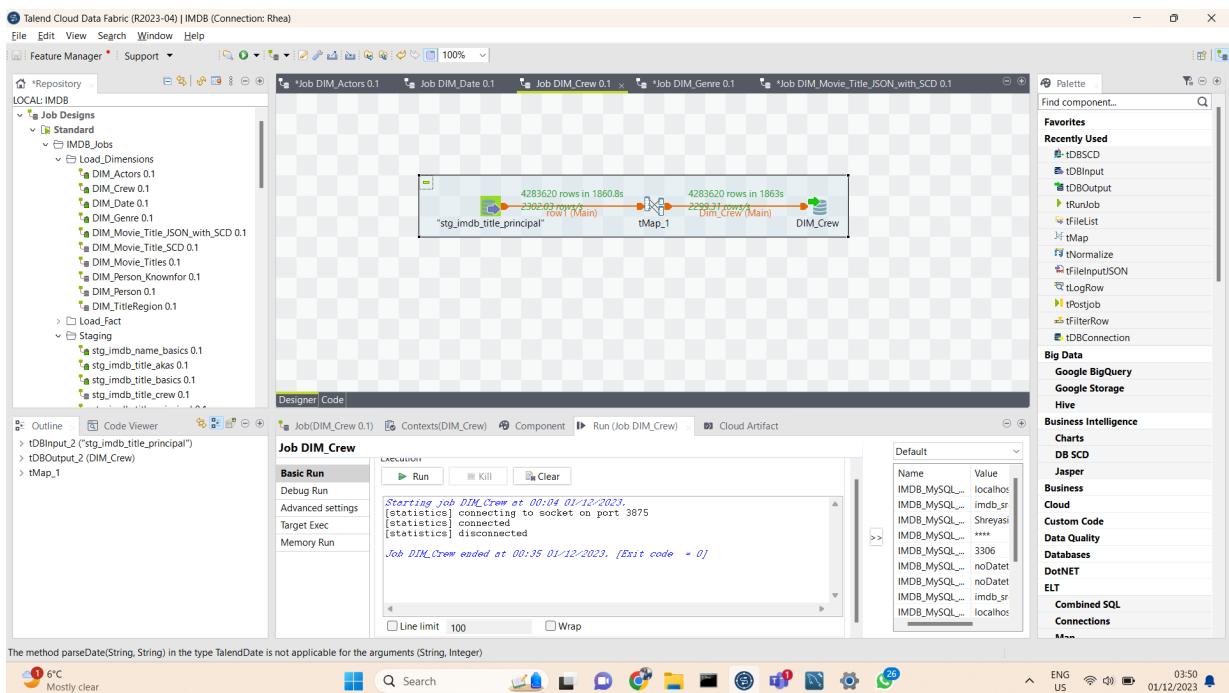
DIM Date:



The screenshot shows the MySQL Workbench interface. The left sidebar shows the 'dim_date' table structure. The main area displays the 'dim_date' table data grid with columns: dateSK, Day_Num, Month_Str, Month_Num, Month_Yr, Year_Num, Is_Weekend, Qtr_Num, DI_CreatedDate, DI_Processeddt, and dt. The bottom pane shows the 'Action Output' log with entries such as 'DROP TABLE', 'SELECT', and 'DROP TABLE' commands along with their execution times and affected row counts.



DIM Crew:



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: Schemas: food_inspection, imdb_src_movie, dim_date, dim_crew, dim_actors, dim_actor_profession, dim_actor_knowfor, dim_titelregion, dim_titlerregion, dim_name_basics, dim_title_akas

Query 1: SELECT * FROM imdb_src_movie.dim_crew

Result Grid: CrewID, name, category, DL_CreatedDate, jobID

CrewID	name	category	DL_CreatedDate	jobID	
1	tt0000002	nm015752	2023-10-01 00:04:49	ISK8Q	
2	tt0000002	nm0252720	2023-10-01 00:04:49	ISK8Q	
3	tt0000002	nm0065413	2023-10-01 00:04:49	ISK8Q	
4	tt0000002	nm0957268	2023-10-01 00:04:49	ISK8Q	
5	tt0000002	nm0675388	2023-10-01 00:04:49	ISK8Q	
6	tt00000574	nm0846807	actress	2023-10-01 00:04:49	ISK8Q
7	tt00000574	nm0846804	actor	2023-10-01 00:04:49	ISK8Q
8	tt00000574	nm3002376	actor	2023-10-01 00:04:49	ISK8Q
10	tt00000574	nm1431224	actor	2023-10-01 00:04:49	ISK8Q
11	tt00000574	nm017210	producer	2023-10-01 00:04:49	ISK8Q
12	tt00000574	nm0425854	producer	2023-10-01 00:04:49	ISK8Q
13	tt00000574	nm0846911	producer	2023-10-01 00:04:49	ISK8Q
14	tt00000574	nm2421834	composer	2023-10-01 00:04:49	ISK8Q
15	tt00000574	nm0675239	cinemat...	2023-10-01 00:04:49	ISK8Q

dim_crew 1

Output:

Action Output:

#	Time	Action	Message	Duration / Fetch
3	22:12:52	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	6.094 sec
5	22:52:01	SELECT * FROM `imdb_src_movie`.`dim_name_basics` LIMIT 0, 1000	1000 row(s) returned	0.047 sec / 0.000 sec
6	03:48:53	SELECT * FROM `imdb_src_movie`.`dim_date` LIMIT 0, 1000	1000 row(s) returned	0.125 sec / 0.078 sec
7	03:49:57	SELECT * FROM `imdb_src_movie`.`dim_crew` 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

6°C Mostly clear

03:51 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: Schemas: food_inspection, imdb_src_movie, dim_date, dim_crew, dim_actors, dim_actor_profession, dim_actor_knowfor, dim_titelregion, dim_titlerregion, dim_name_basics, dim_title_akas

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

Result Grid: count(*)

count(*)
4283620

Table: dim_crew

Output:

Action Output:

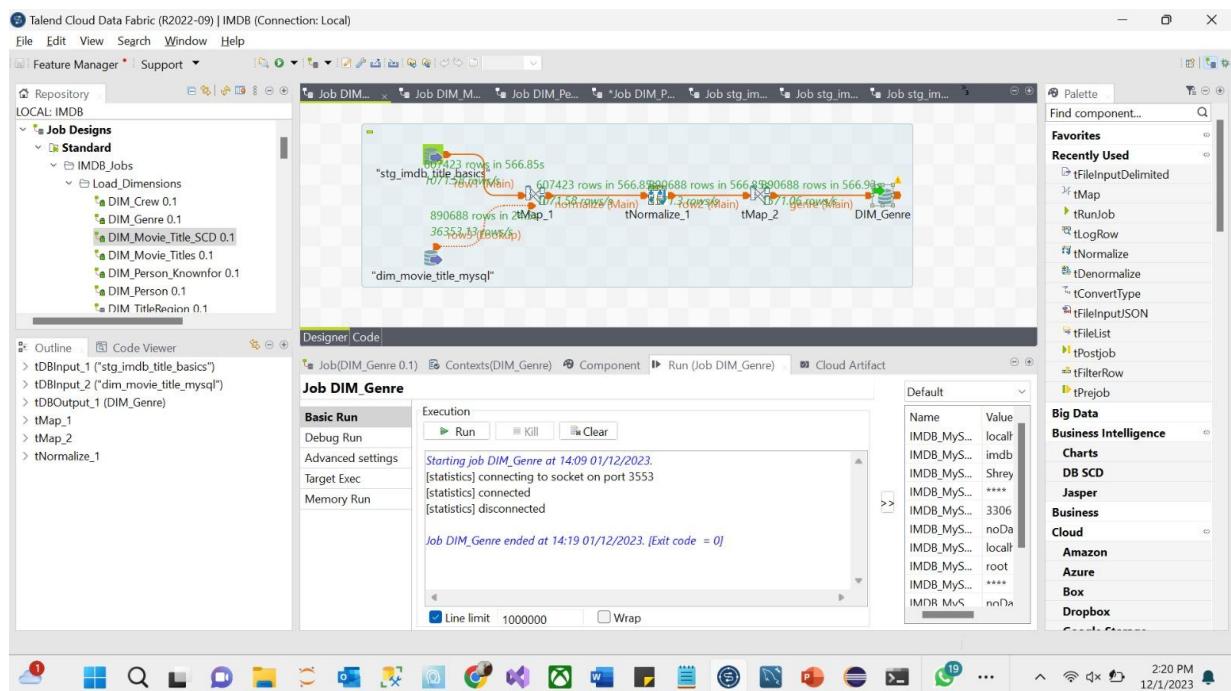
#	Time	Action	Message	Duration / Fetch
43	13:42:07	SELECT * FROM `imdb_src_movie`.`stg_newdata_title_basic`	10 row(s) returned	0.281 sec / 0.000 sec
44	13:42:15	SELECT count(*) FROM `imdb_src_movie`.`stg_newdata_title_basic`	1 row(s) returned	0.000 sec / 0.000 sec
45	13:44:01	SELECT * FROM `imdb_src_movie`.`dim_date`	10000 row(s) returned	0.125 sec / 0.109 sec
46	13:44:05	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 / 4.405 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

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. The central area shows the 'dim_genre' table in the 'Query 1' tab. The table has columns: GenreSK, MovieTitle_SK, tconst, genres, DL_CreatedDate, and JobId. The data grid shows rows for various genres. Below the table, the 'Action History' tab shows a log of recent actions, including a drop of the 'dim_genre_principals' table and a select query for the 'dim_genre' table.

MySQL Workbench

Local instance MySQL80 (imdb_src.x) Local instance MySQL80 (imd...)

File Edit View Query Database Server Tools Scripting Help

Navigator: petcare_connect - Schema: dim_movie_title_mysql dim_crew stg_imdb_title_basics imdb_title_principals dim_genre

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

Result Grid: COUNT(*)

COUNT(*)	890688
----------	--------

Administration Schemas Information

Table: dim_genre

Columns:

GenreSK	int PK
MovieTitle_SK	int
tnconst	varchar
genres	varchar
DL_CreatedDate	datetime
InhId	varchar

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

Object Info Session

Result Grid Form Editor

Result 2 Read Only

Output

Action Output

Message Duration / Fetch

Time Action # 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

Repository Job DIM_Movie_Title_SCD

LOCAL: IMDB

Job Designs Standard

IMDB.Jobs

Load Dimensions

- DIM_Crew 0.1
- DIM_Genre 0.1
- DIM_Movie_Title_SCD 0.1
- DIM_Movie_Titles 0.1
- DIM_Person_Knownfor 0.1
- DIM_Person 0.1
- DIM_TitleReonion 0.1

Designer Code

Job DIM_Movie_Title_SCD

Basic Run Execution

Run Kill Clear

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]

Default

Name	Value
IMDB.MyS...	localf
IMDB.MyS...	imdb
IMDB.MyS...	Shrey
IMDB.MyS...	****
IMDB.MyS...	3306
IMDB.MyS...	noDa
IMDB.MyS...	localf
IMDB.MyS...	root
IMDB.MyS...	****

Palette

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

5:45 PM 12/1/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: dim_movie_title

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

tconst	titleType	primaryTitle	originalTitle	isAdult	runtimeMinutes	DI_CreatedDate	JobID	scd_st
tt0120338	movie	Titanic (1997)	Titanic	0	194	2023-12-01 14:24:13	rgq7bz	1997-01-01
tt0499549	movie	Avatar (2009)	Avatar	0	162	2023-12-01 14:24:13	rgq7bz	2009-01-01
tt0848228	movie	The Avengers (2012)	The Avengers	0	143	2023-12-01 14:24:13	rgq7bz	2012-01-01
tt10872600	movie	Spider-Man: No Way Home (2021)	Spider-Man: No Way Home	0	148	2023-12-01 14:24:13	rgq7bz	2021-01-01
tt1825683	movie	Black Panther (2018)	Black Panther	0	134	2023-12-01 14:24:13	rgq7bz	2018-01-01
tt2395427	movie	Avengers: Age of Ultron (2015)	Avengers: Age of Ultron	0	141	2023-12-01 14:24:13	rgq7bz	2015-01-01
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	rgq7bz	2015-01-01

Table: dim_movie_title

Columns:

- tconst: varchar
- titleType: varchar
- primaryTitle: varchar
- originalTitle: varchar
- isAdult: tinyint

Action Output

#	Time	Action	Message	Duration / Fetch
990	17:27:18	SELECT * FROM imdb_src_movie.fct_rating LIMIT 0, 500	500 row(s) returned	0.000 sec / 0.000 sec
991	17:30:20	SELECT * FROM imdb_src_movie.dim_movie_title LIMIT 0, 500	10 row(s) returned	0.031 sec / 0.000 sec

Object Info Session

5:30 PM 12/1/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: dim_movie_title

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

count(*)
10

Table: dim_movie_title

Columns:

- tconst: varchar
- titleType: varchar
- primaryTitle: varchar
- originalTitle: varchar
- isAdult: tinyint

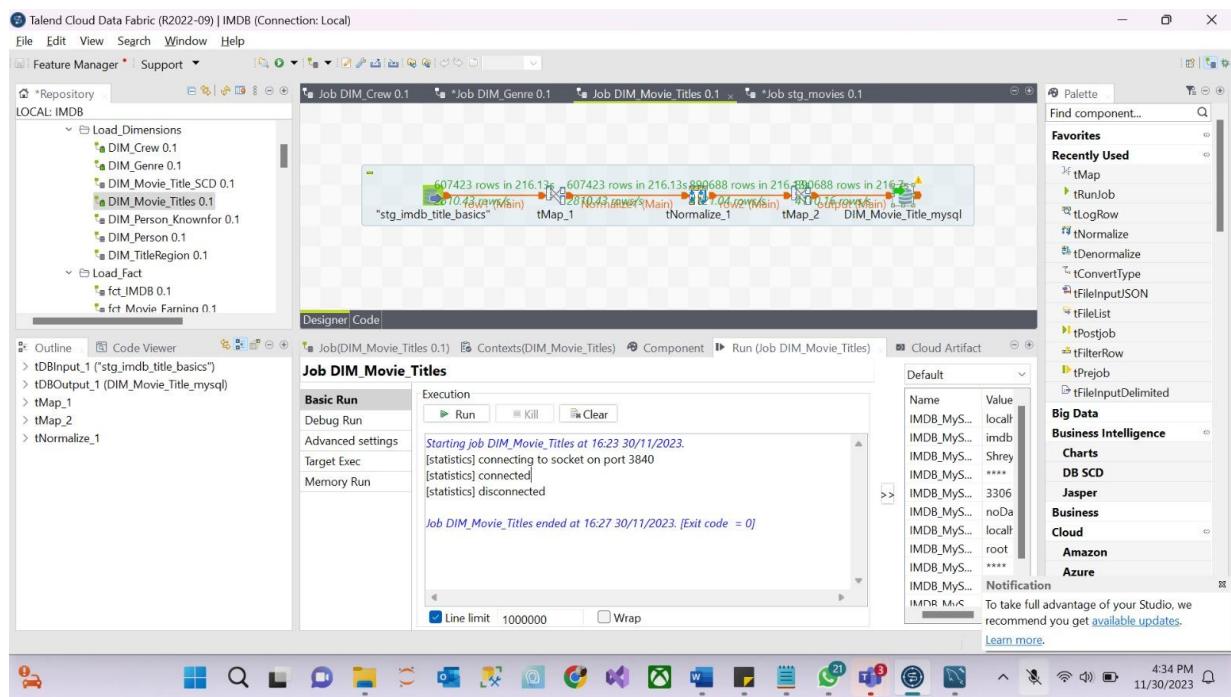
Action Output

#	Time	Action	Message	Duration / Fetch
991	17:30:20	SELECT * FROM imdb_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 imdb_src_movie.dim_movie_title LIMIT 0, 500	1 row(s) returned	0.000 sec / 0.000 sec

Object Info Session

5:31 PM 12/1/2023

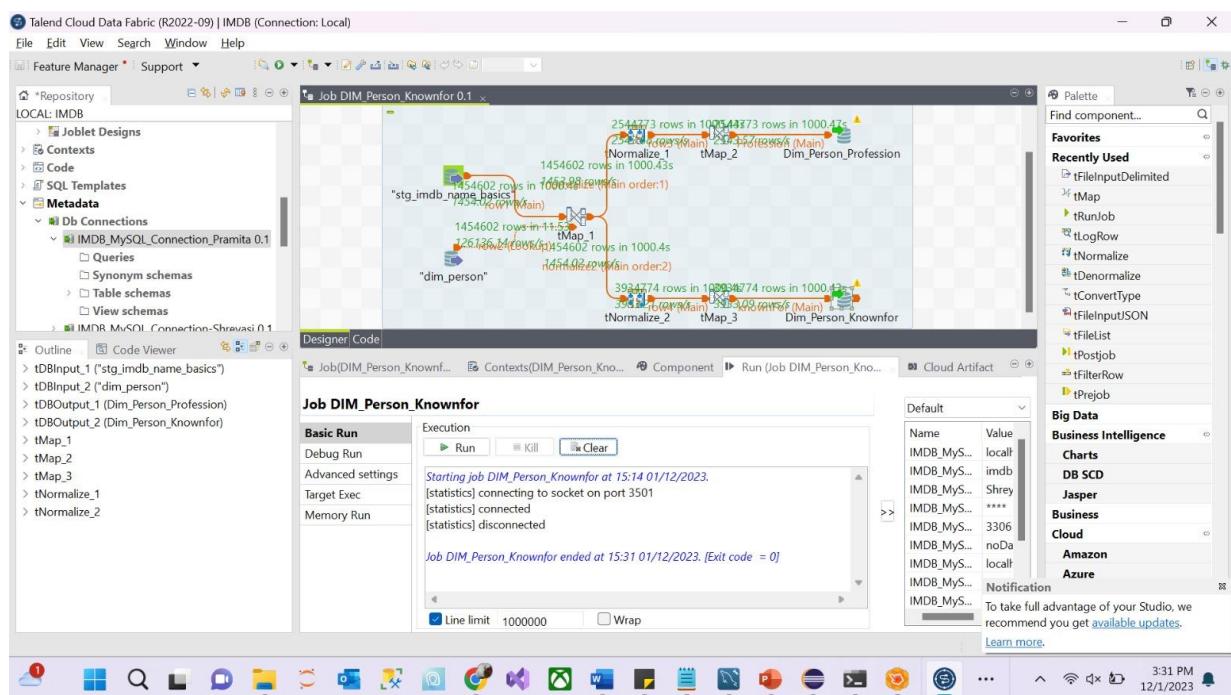
DIM Movie Titles:



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

- Title Bar:** MySQL Workbench
- Toolbar:** File, Edit, View, Query, Database, Server, Tools, Scripting, Help.
- Navigator:** Schemas (cosmos, directors, first_schema, food_inspection, imdb_movies, imdb_src_movie), Tables (dim_crew, dim_movie_title_my, dim_name_basics, dim_title_akas, dim_title_principals, dim_title_crew, dim_title_ratings).
- Query Editor:** Query 1: SELECT * FROM imdb_src_movie.dim_movie_title_mysql; Result Grid (14 rows, columns: MovieTitle_SK, tconst, titleType, primaryTitle, originalTitle, isAdult, scd_start, endYear, runtimeMinutes, genres, DI_...).
- Output:** Action Output (1. 23:43:26 Apply changes to Pet_Care_Connect, 2. 14:07:21 SELECT * FROM imdb_src_movie.dim_movie_title_mysql LIMIT 0..500).
- Bottom:** Windows taskbar with various icons.

DIM Person Known for:



MySQL Workbench

Local instance MySQL80 (imdb_src) Local instance MySQL80 (imdb_sr...)

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_imdb_title_principal imdb_title_principals imdb_title_principals stg_imdb_title_principal dim_person dim_person_profession dim_person_knownfor dim_person_k...

SCHEMAS

Filter objects

- dim_crew
- dim_genre
- dim_movie_title
- dim_movie_title_my...
- dim_person
- dim_person_knownfor
- dim_person_profess
- imdb_title_ratings
- imdb_title_principals
- imdb_title_crew
- imdb_title_basics
- imdb_title_akas
- imdb_name_basics
- stg_imdb_name_basics
- stg_imdb_title_akas
- stg_imdb_title_basics

Administration Schemas Information

Table: dim_person_knownfor

Columns:

KnownFor TitleSK	PersonSK	nconst	knownForTitle	DL_CreatedDate	JobID
4	1	nm0000001	t0072308	2023-12-01 15:14:58	BopmRU
5	1	nm0000001	t0055317	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	t0037382	2023-12-01 15:14:58	BopmRU

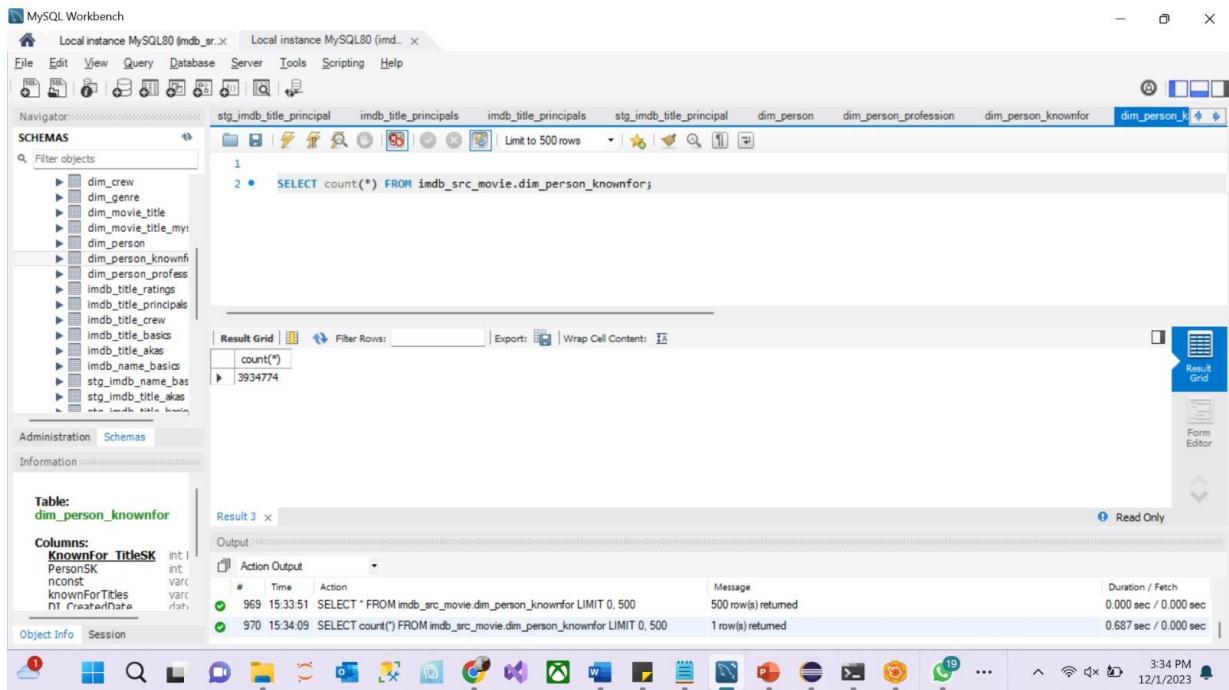
Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor

Action Output

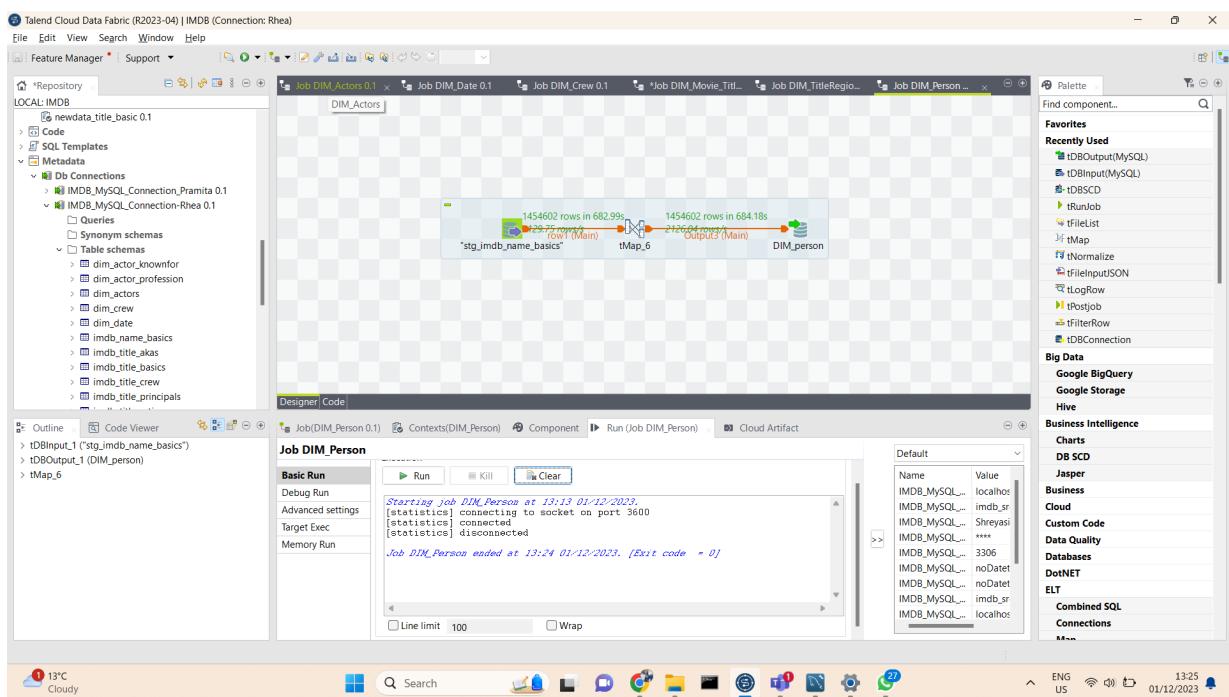
#	Time	Action	Message	Duration / Fetch
968	15:33:00	SELECT * FROM imdb_src.movie.dim_person_knownfor LIMIT 0, 500	500 row(s) returned	0.000 sec / 0.000 sec
969	15:33:51	SELECT * FROM imdb_src.movie.dim_person_knownfor LIMIT 0, 500	500 row(s) returned	0.000 sec / 0.000 sec

Object Info Session

3:33 PM 12/1/2023



DIM Person:



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_imdb_title_principal stg_imdb_title_ratings stg_movies stg_newdata_name_basic stg_newdata_title_basic dim_actor_profession dim_titleregion dim_person SQLAdditions

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

Table: dim_person

Columns:

PersonSK	ncnst	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	789vFt	
rm00000002	Lauren Bacall	1924-01-01 00:00:00	2014-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000003	Brigitte Bardot	1934-01-01 00:00:00	2023-12-01 13:13:36	789vFt		
rm00000004	John Belushi	1949-01-01 00:00:00	1982-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000005	Ingrid Bergman	1918-01-01 00:00:00	2007-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000006	Ingrid Bergman	1915-01-01 00:00:00	1982-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000007	Humphrey Bogart	1899-01-01 00:00:00	1957-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000008	Marlon Brando	1924-01-01 00:00:00	2004-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000009	Richard Burton	1925-01-01 00:00:00	1984-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000010	James Cagney	1899-01-01 00:00:00	1986-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000011	Gary Cooper	1901-01-01 00:00:00	1961-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000012	Bette Davis	1908-01-01 00:00:00	1989-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000013	Doris Day	1922-01-01 00:00:00	2019-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000014	Olivia de Havilland	1916-01-01 00:00:00	2020-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000015	Thelma Ritter	1911-01-01 00:00:00	1975-12-21 18:18:48	789vFt		

Result Grid | Filter Rows: | Edits: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types | Query Stats | 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_title_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 / 0.105 sec
24	04:39:29	SELECT * FROM imdb_src_movie dim_titleregion 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 / 2.328 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

Output

Object Info Session

13°C Cloudy 13:26 01/12/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_imdb_title_principal stg_imdb_title_ratings stg_movies stg_newdata_name_basic stg_newdata_title_basic dim_actor_profession dim_titleregion dim_person SQLAdditions

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

Table: dim_person

Columns:

PersonSK	ncnst	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	789vFt	
rm00000002	Lauren Bacall	1924-01-01 00:00:00	2014-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000003	Brigitte Bardot	1934-01-01 00:00:00	2023-12-01 13:13:36	789vFt		
rm00000004	John Belushi	1949-01-01 00:00:00	1982-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000005	Ingrid Bergman	1918-01-01 00:00:00	2007-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000006	Ingrid Bergman	1915-01-01 00:00:00	1982-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000007	Humphrey Bogart	1899-01-01 00:00:00	1957-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000008	Marlon Brando	1924-01-01 00:00:00	2004-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000009	Richard Burton	1925-01-01 00:00:00	1984-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000010	James Cagney	1899-01-01 00:00:00	1986-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000011	Gary Cooper	1901-01-01 00:00:00	1961-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000012	Bette Davis	1908-01-01 00:00:00	1989-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000013	Doris Day	1922-01-01 00:00:00	2019-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000014	Olivia de Havilland	1916-01-01 00:00:00	2020-01-01 00:00:00	2023-12-01 13:13:36	789vFt	
rm00000015	Thelma Ritter	1911-01-01 00:00:00	1975-12-21 18:18:48	789vFt		

Result Grid | Filter Rows: | Edits: | Export/Import: | Wrap Cell Content: | Fetch rows: | Result Grid | Form Editor | Field Types | Query Stats | 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_title_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 / 0.105 sec
24	04:39:29	SELECT * FROM imdb_src_movie dim_titleregion 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 / 2.328 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

Output

Object Info Session

13°C Cloudy 13:27 01/12/2023

DIM Person Profession:

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

- Navigator:** Shows the schema structure, including tables like dim_crew, dim_genre, dim_movie_title, dim_movie_title_my, dim_person, dim_person_profession, dim_person_knownfor, dim_person_ratings, dim_title_principals, dim_title_crew, dim_title_basics, dim_name_basics, stg_imdb_name_basics, and stg_imdb_title_akas.
- Table:** dim_person_profession
- Columns:**

Column Name	Type	Properties
ProfessionSK	int	PK
PersonSK	int	
nconst	varchar	
primaryProfession	varchar	
DL_CreatedDate	datetime	
- Result Grid:** Displays 16 rows of data from the dim_person_profession table. The columns are ProfessionSK, PersonSK, nconst, primaryProfession, DL_CreatedDate, and JobID.
- Action Output:** Shows two recent actions: a SELECT query on dim_person_profession and another on dim_person_profession.

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

- Navigator:** Shows the schema structure, including tables like dim_crew, dim_genre, dim_movie_title, dim_movie_title_my, dim_person, dim_person_knownfor, dim_person_profession, dim_title_principals, dim_title_crew, dim_title_basics, dim_name_basics, stg_imdb_name_basics, and stg_imdb_title_akas.
- Table:** dim_person_knownfor
- Columns:**

Column Name	Type	Properties
KnownFor TitleSK	int	PK
PersonSK	int	
nconst	varchar	
KnownForTitles	varchar	
DL_CreatedDate	datetime	
- Result Grid:** Displays 1 row of data from the dim_person_knownfor table, showing the count(*) result.
- Action Output:** Shows two recent actions: a SELECT query on dim_person_knownfor and another on dim_person_profession.

DIM Title Region:

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

Job DIM_TitleRegion 0.1

Designer | Code

Job DIM_TitleRegion

Basic Run

- Run
- Kill
- Clear

2612793|tt9916680|De la ilusión al desconcierto: cine colombiano 1970-1995|CO
2612795|tt9916706|Denkyeyer Denke|[IN]|||false|01-12-2023|1PrFos
2612796|tt9916730|6 Guan|[IN]|||false|01-12-2023|1PrFos
2612797|tt9916754|Chico Albuquerque - Revelaciones|BR||imdbDisplay||false|01-12-2023|1PrFos
2612798|tt9916754|Chico Albuquerque - Revelaciones|||original||true|01-12-2023|1PrFos

Job DIM_TitleRegion ended at 04:38 01/12/2023. [Exit code = 0]

Default

Name	Value
IMDB MySQL_	localhost
IMDB MySQL_	imdb_sr
IMDB MySQL_	Shreyasi
IMDB MySQL_	****
IMDB MySQL_	3306
IMDB MySQL_	noDate
IMDB MySQL_	noDate
IMDB MySQL_	imdb_sr
IMDB MySQL_	localhost

Line limit: 100 Wrap

6°C Mostly cloudy 04:38 01/12/2023 ENG US

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator

Table: dim_titleregion

Columns:

dimTitleRegionSK	int PK
titleID	varchar(10)
title	varchar(1024)
region	varchar(255)
language	varchar(255)
types	varchar(255)
attributes	varchar(1024)
DL_CreatedDate	datetime
pid	varchar(100)

Action Output

#	Time	Action	Message	Duration / Fetch
19	03:58:17	SELECT * FROM imbd_src_movie.stg_imdb_title_ratings LIMIT 0, 1000	1000 row(s) returned	0.109 sec / 0.016 sec
20	03:58:40	SELECT * FROM imbd_src_movie.stg_movies LIMIT 0, 1000	265 row(s) returned	0.203 sec / 0.000 sec
21	03:59:26	SELECT * FROM imbd_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 imbd_src_movie.stg_newdata_title_basic LIMIT 0, 1000	10 row(s) returned	0.125 sec / 0.000 sec
23	04:39:00	SELECT * FROM imbd_src_movie.dim_actor_profession LIMIT 0, 1000	1000 row(s) returned	0.063 sec / 0.015 sec
24	04:39:29	SELECT * FROM imbd_src_movie.dim_titleRegion LIMIT 0, 1000	1000 row(s) returned	0.047 sec / 0.000 sec

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/Import: Wrap Cell Content: Fetch rows: Result Grid Form Editor Field Types Query Stats Context Help Snippets

6°C Mostly cloudy 04:39 01/12/2023 ENG US

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: stg_imdb_title_principal, stg_imdb_title_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.

Result Grid | Filter Rows: Export: Wrap Cell Content: Result Grid | Form Editor | Field Types | Context Help | Snippets | Read Only | Object Info | Session | Rain coming In about 2.5 hours | 13:48 | 01/12/2023

Table: dim_titleregion

Columns:

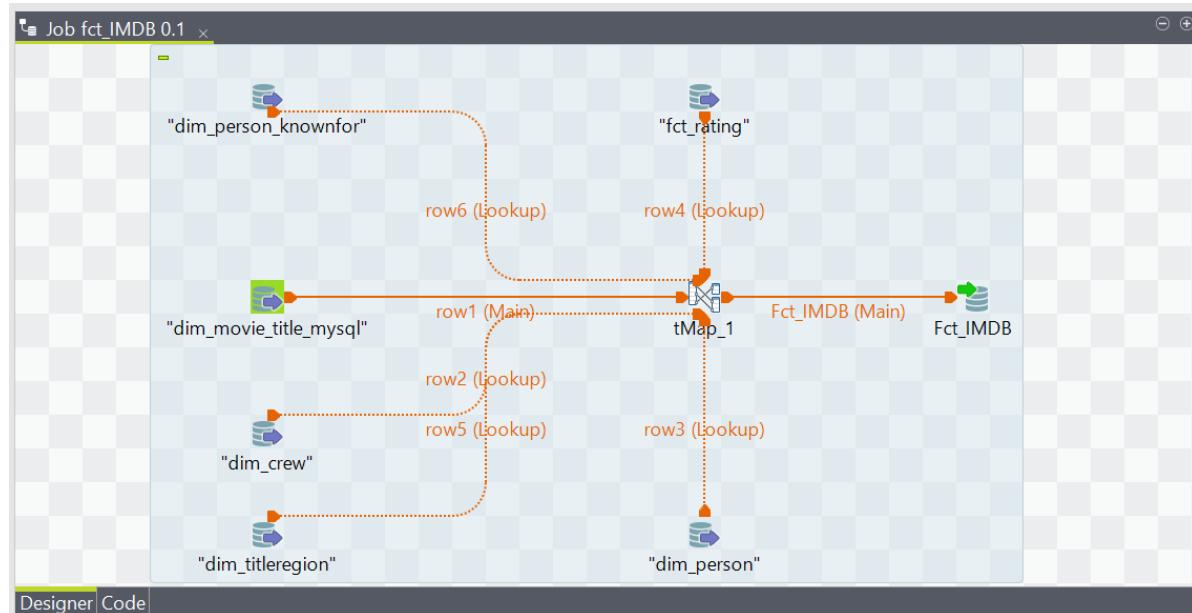
- titleRegionSK int PK
- id varchar(10)
- title varchar(1024)
- region varchar(255)
- language varchar(255)
- country varchar(1024)
- attributes varchar(1024)
- isOriginalTitle varchar(255)
- DJ_CreatedDate datetime
- pid varchar(100)

Action Output

#	Time	Action	Message	Duration / Fetch
47	13:45:05	SELECT * FROM imdb_src_movie.dim_crew	4283620 row(s) returned	0.156 sec / 4.406 sec
48	13:46:19	SELECT count(*) FROM imdb_src_movie.dim_crew	1 row(s) returned	1.157 sec / 0.000 sec
49	13:47:00	SELECT * FROM imdb_src_movie.dim_person	1454602 row(s) returned	0.031 sec / 1.563 sec
50	13:47:09	SELECT count(*) FROM imdb_src_movie.dim_person	1 row(s) returned	0.266 sec / 0.000 sec
51	13:47:46	SELECT * FROM imdb_src_movie.dim_titleregion	2612798 row(s) returned	0.140 sec / 4.344 sec
52	13:48:05	SELECT count(*) FROM imdb_src_movie.dim_titleregion	1 row(s) returned	0.937 sec / 0.000 sec

Fact Tables:

Fact IMDB:



```
1 •  select * from fct_imdb;
```

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

	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

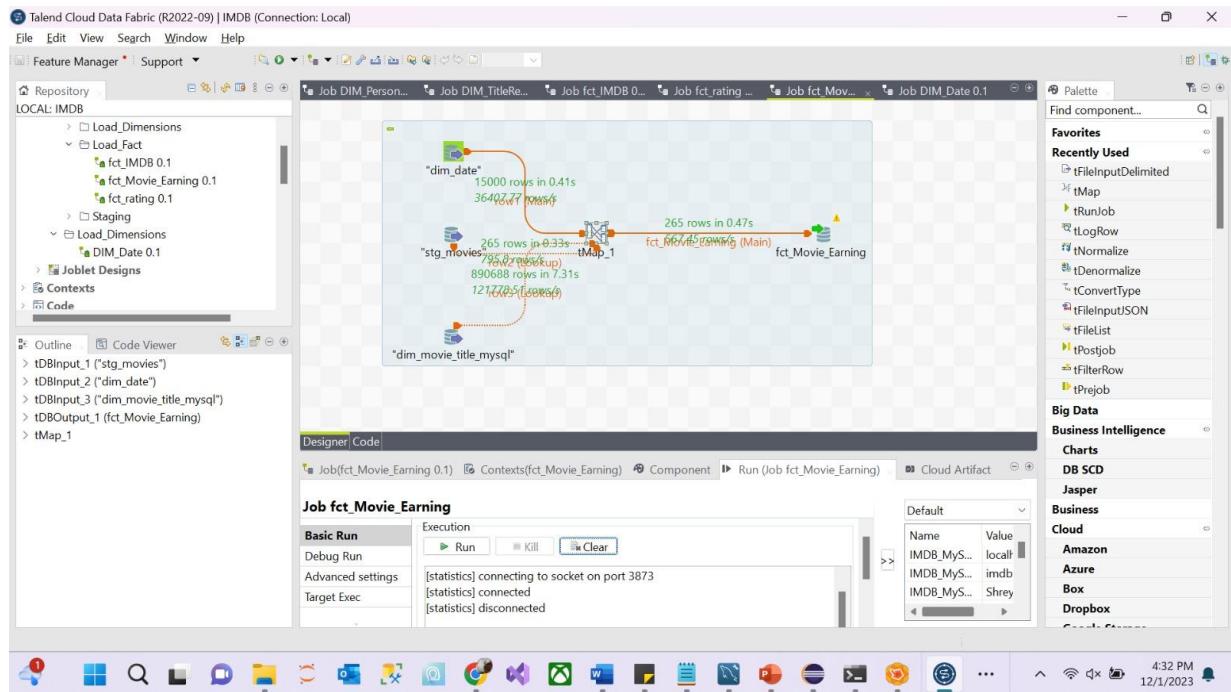
fct_imdb 1 x Apply

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

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

COUNT(*)
▶ 4299743

Fact Movie Earnings:



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator

SCHEMAS

imdb_movies
imdb_src_movie

Tables

stg_imdb_title_principal dim_person dim_person_profession dim_person_knownfor dim_person_knownfor dim_titlerregion fct_rating dim_date

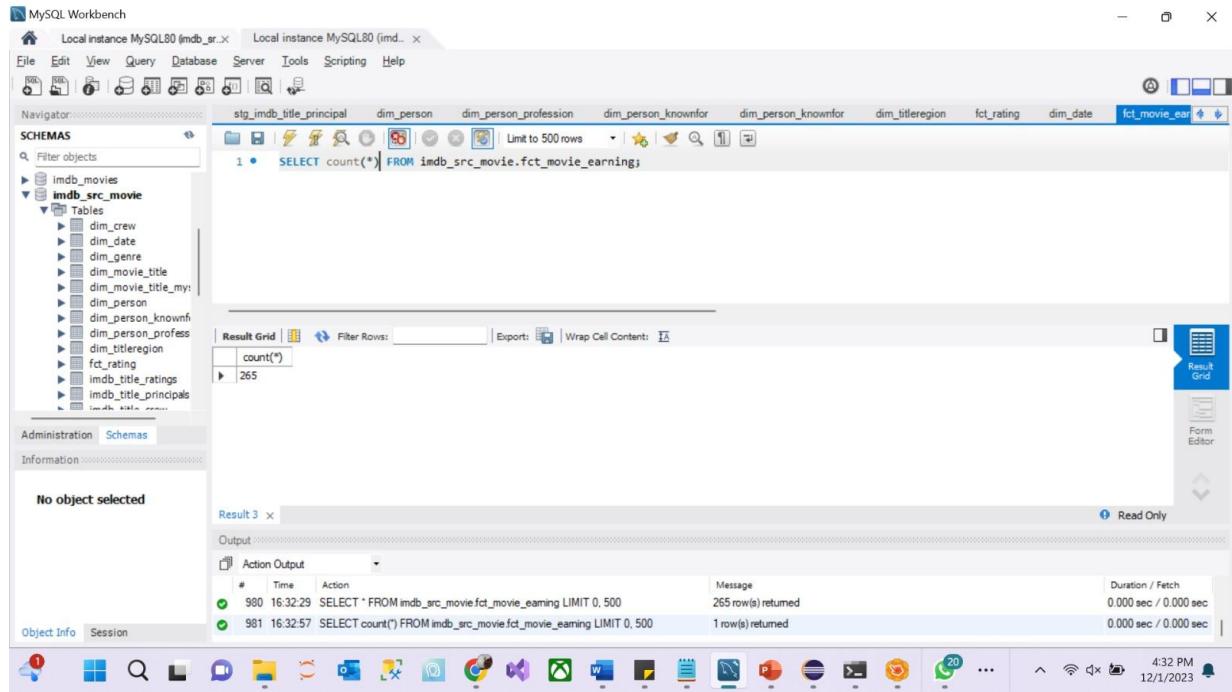
Result Grid

1 • SELECT * FROM imdb_src_movie.fct_movie_earning;

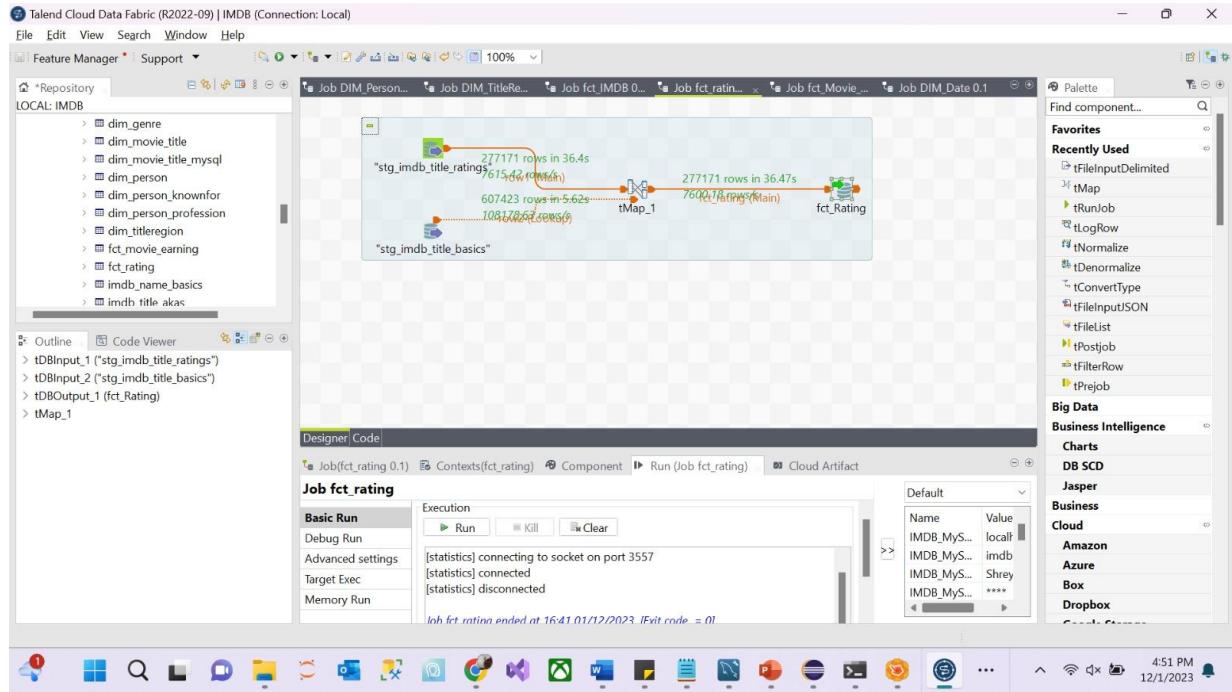
MovieEarningsSK	dateSK	MovieTitle_SK	Rank	Gross	Theaters	Per_Theater	Total_Gross	Days	DL_CreatedDate	JobID
2909	19971219	223874	1	\$8,658,814	2,674	\$3,238	\$8,658,814	1	2023-12-01 16:32:23	EuZaBZ
2910	19971220	223874	1	\$10,672,013	2,674	\$3,991	\$19,330,827	2	2023-12-01 16:32:23	EuZaBZ
2911	19971221	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	\$34,216,343	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	\$65,091,634	8	2023-12-01 16:32:23	EuZaBZ

Object Info Session

4:32 PM 12/1/2023



Fact Rating:



MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: fct_rating

SCHEMAS

- dim_person_profession
- dim_titleregion
- fct_rating
- imdb_title_ratings
- imdb_title_principals
- imdb_title_crew
- imdb_title_basics
- imdb_title_akas
- imdb_name_basics
- stg_imdb_title_basics
- stg_imdb_title_akas
- stg_imdb_title_basic
- stg_imdb_title_crew
- stg_imdb_title_princ
- stg_imdb_title_rating

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	HULL	2023-12-01 16:14:51	af5Wxx
5	tt00000630	4	24	HULL	2023-12-01 16:14:51	af5Wxx
6	tt00000675	5	19	HULL	2023-12-01 16:14:51	af5Wxx
7	tt00000679	5	65	120	2023-12-01 16:14:51	af5Wxx
8	tt00000793	5	20	HULL	2023-12-01 16:14:51	af5Wxx
9	tt00000827	6	14	HULL	2023-12-01 16:14:51	af5Wxx

Result Grid

Filter Rows:

Action Output

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

Object Info Session

4:25 PM 12/1/2023

MySQL Workbench

File Edit View Query Database Server Tools Scripting Help

Navigator: als imdb_title_principals stg_imdb_title_principal dim_person dim_person_profession dim_person_knownfor dim_person_knownfor dim_titlerregion fct_rating

SCHEMAS: dim_person_profession dim_titlerregion fct_rating imdb_title_ratings imdb_title_principals imdb_title_crew imdb_title_basics imdb_title_akas imdb_name_basics stg_imdb_name_basics stg_imdb_title_akas stg_imdb_title_basic 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: Result Grid Form Editor

Table: fct_rating

Columns:

RatingSK	int - PK
const	varchar
averageRating	double
numVotes	int
runtimeMinutes	varchar
CreatedDate	datetime

Action Output

#	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

Object Info Session

Result 2 x Read Only

Output:

Action Output

Message

Duration / Fetch

12/1/2023 4:25 PM

SQL Queries -

Trend Analysis:

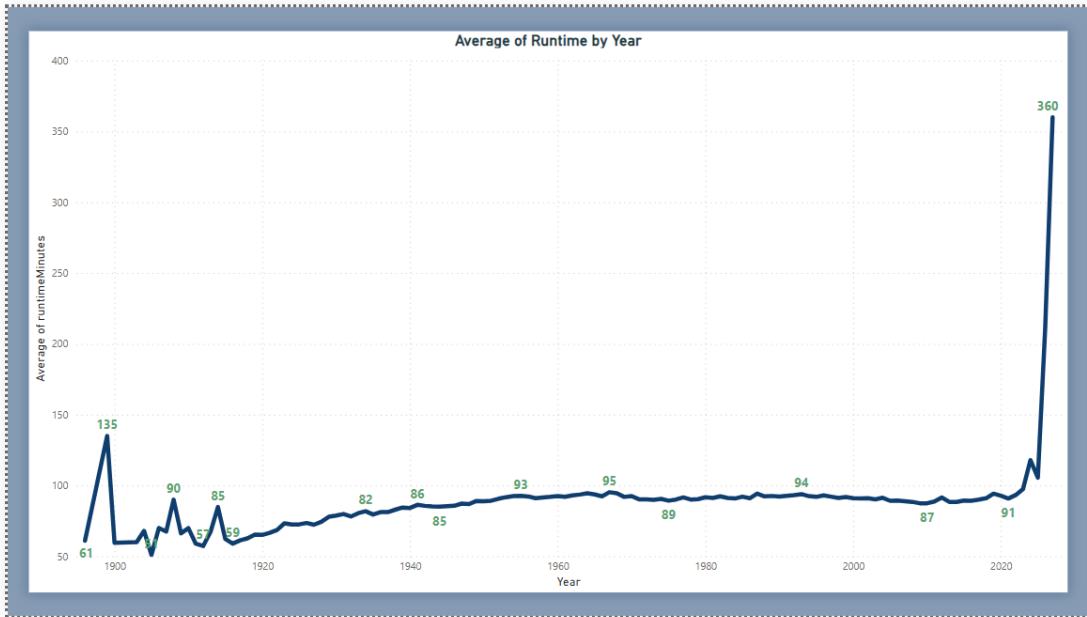
1. Average movie runtime change over the years:
 - a. For every year what is the average run-time and plot it on the graph.

```
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: Result Grid Form Editor

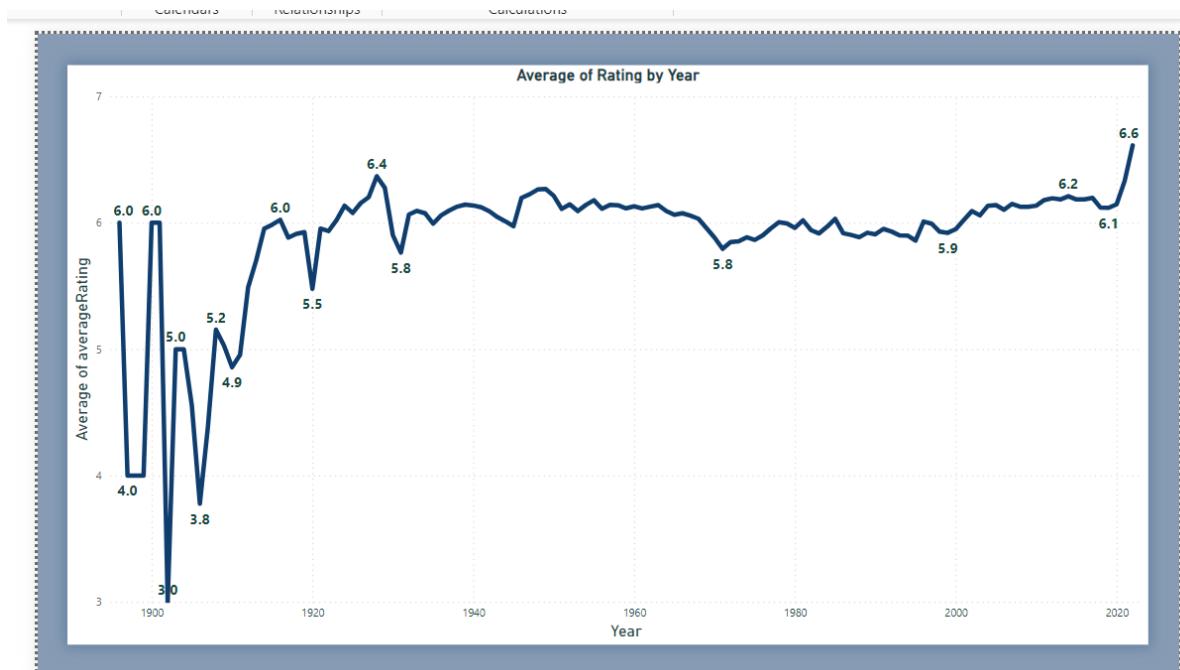
startYear	AvgRuntime
1896-01-01 00:00:00	61
1897-01-01 00:00:00	NULL
1898-01-01 00:00:00	NULL
1899-01-01 00:00:00	135
1900-01-01 00:00:00	59.5
1901-01-01 00:00:00	NULL
1902-01-01 00:00:00	NULL
1903-01-01 00:00:00	60
1904-01-01 00:00:00	68
1905-01-01 00:00:00	51
1906-01-01 00:00:00	70
1907-01-01 00:00:00	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
6.0000	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

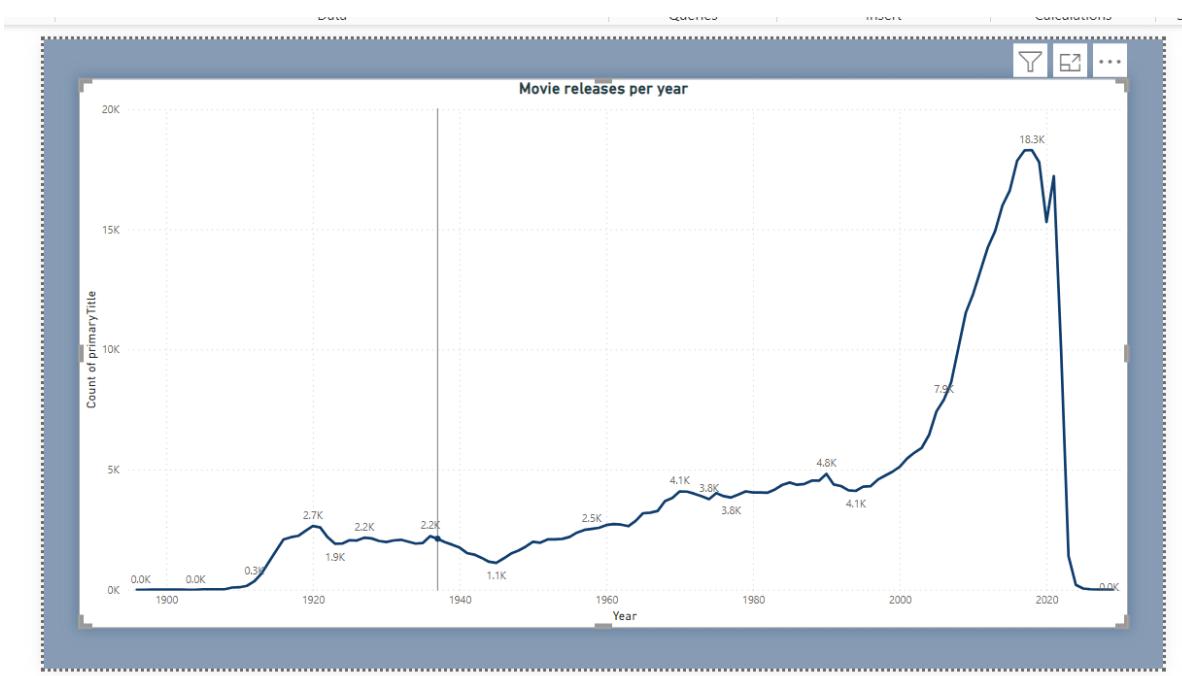
12 • `SELECT COUNT(tconst) AS movieCount, startYear
FROM dim_movie_title_mysql
GROUP BY startYear;`

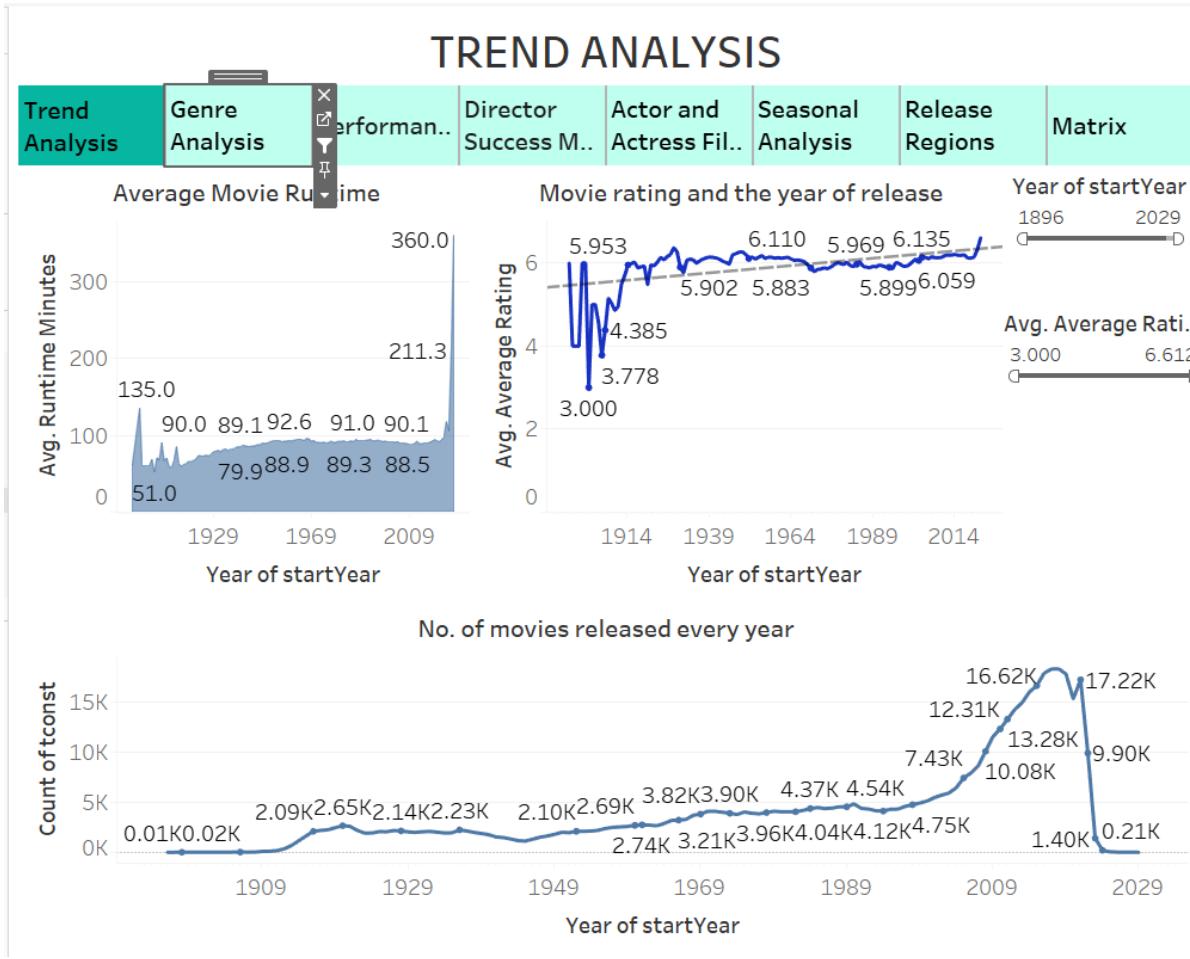
13

14

15

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: Result 24 ×

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
...	6.4765

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
```

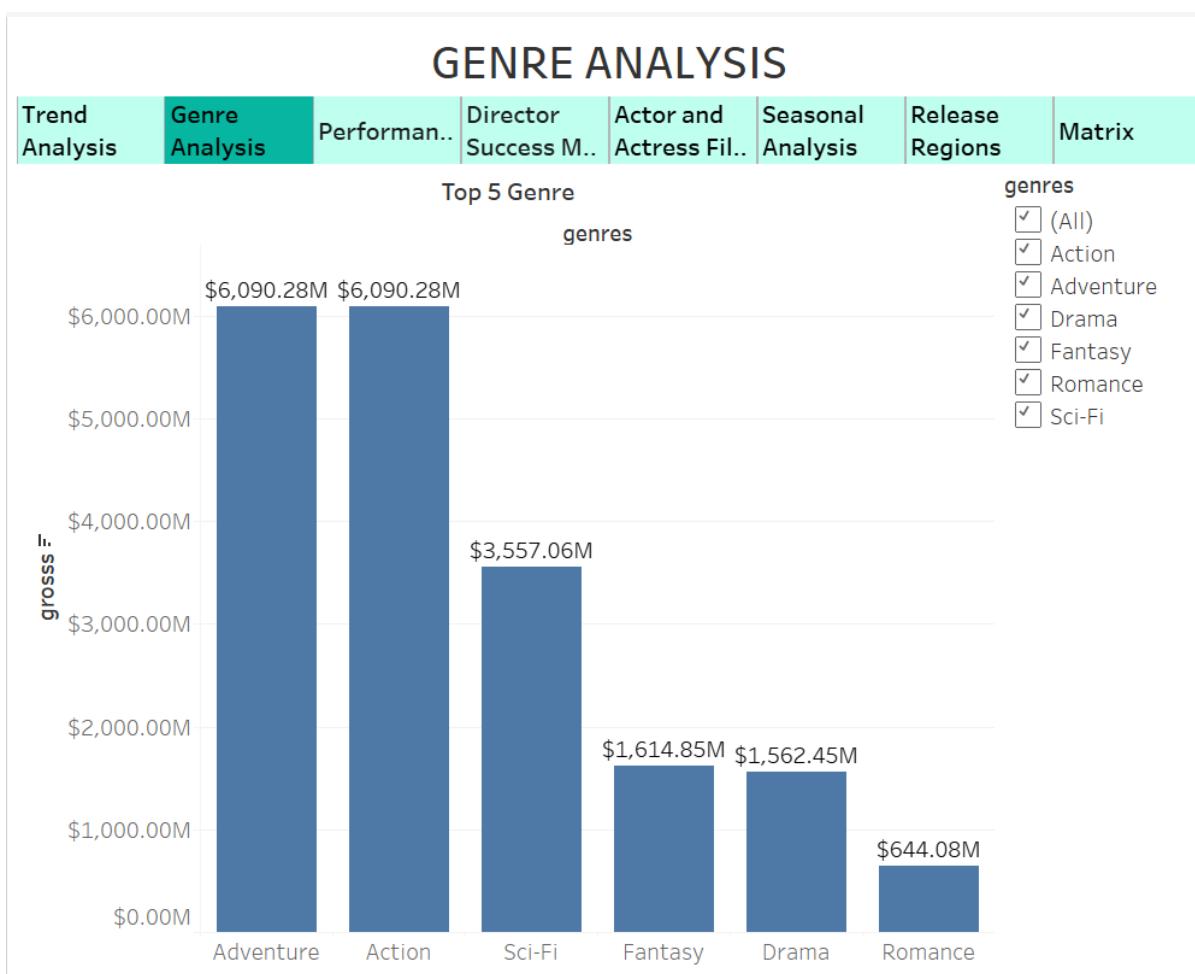
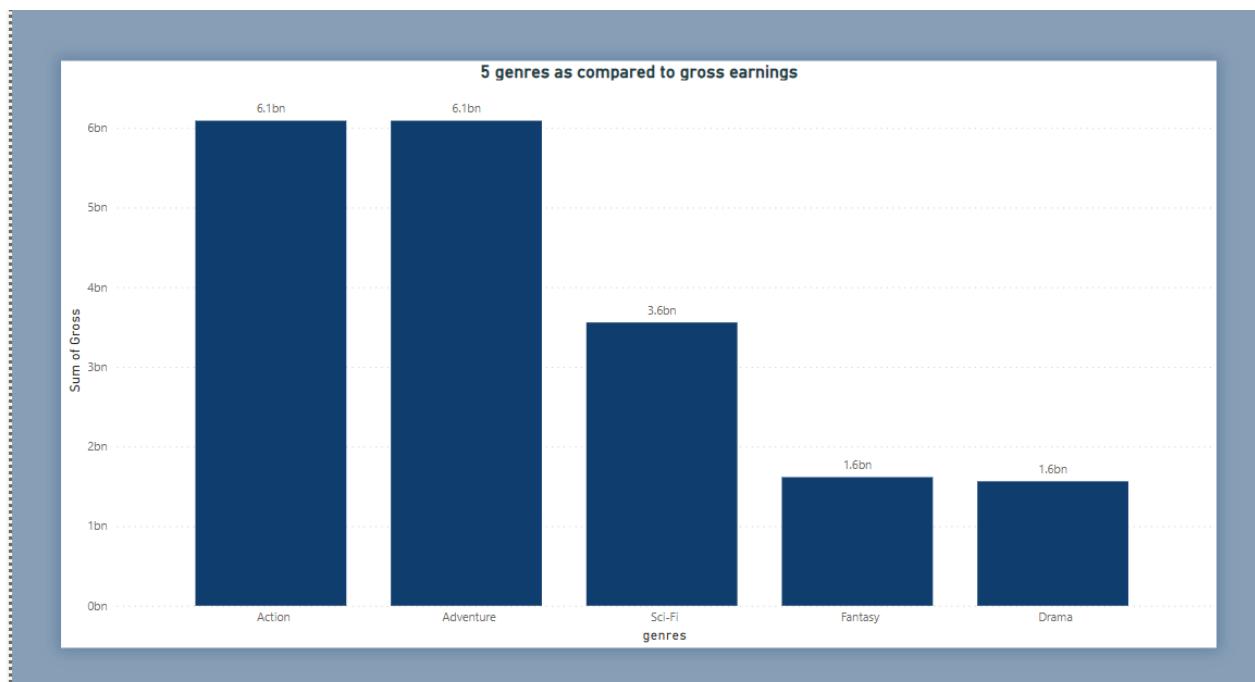
```

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 Workbench interface. At the top is a toolbar with various icons for file operations, search, and navigation. Below the toolbar is a query editor window containing the SQL code provided above. The code is numbered from 43 to 54. The result grid below the editor displays the output of the query, which consists of two columns: 'genres' and 'total_earnings'. The data rows are:

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



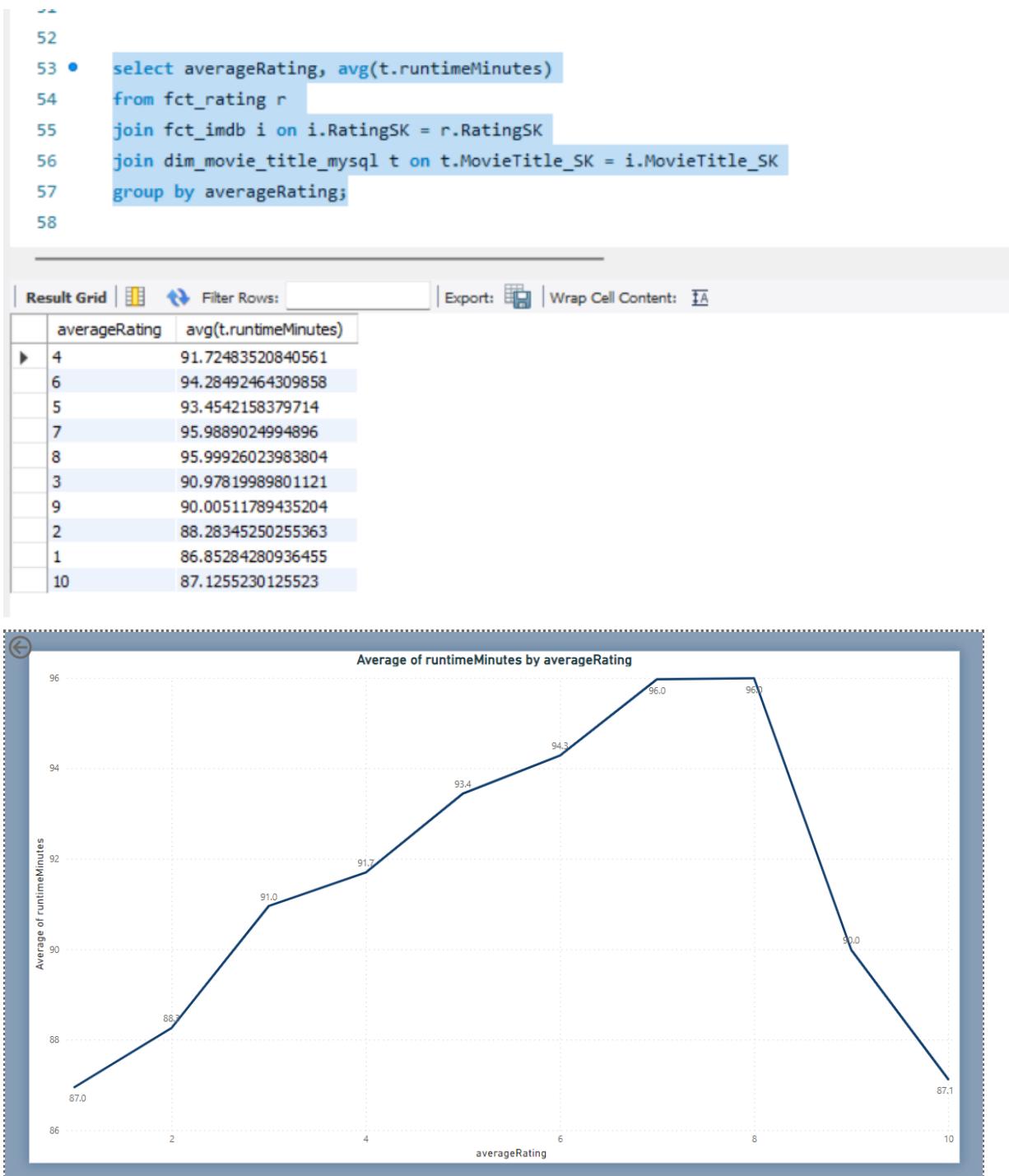
Performance Metrics:

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

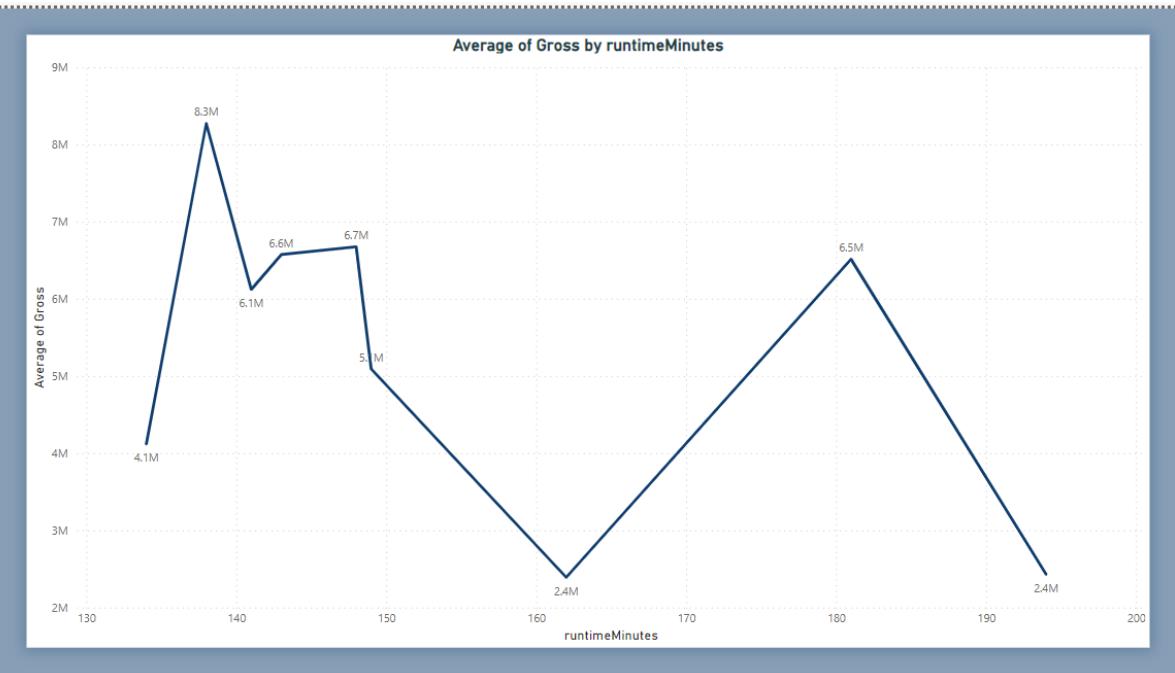
```

```
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

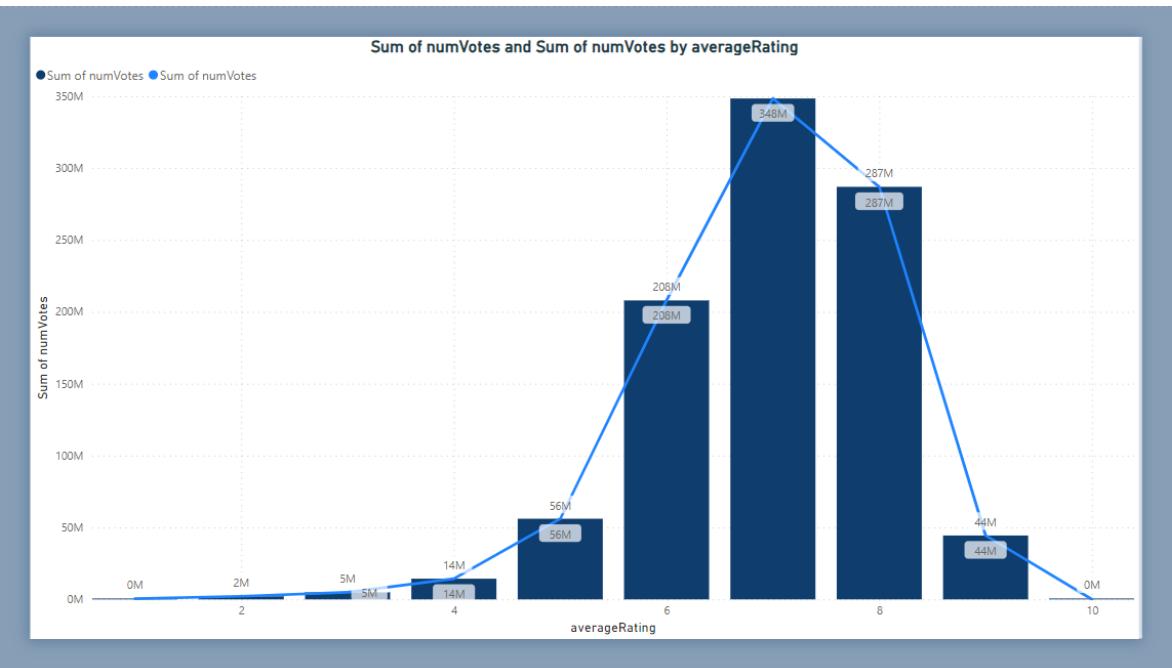
```
select averageRating,sum(numVotes) from fct_rating
```

```
group by averageRating;
```

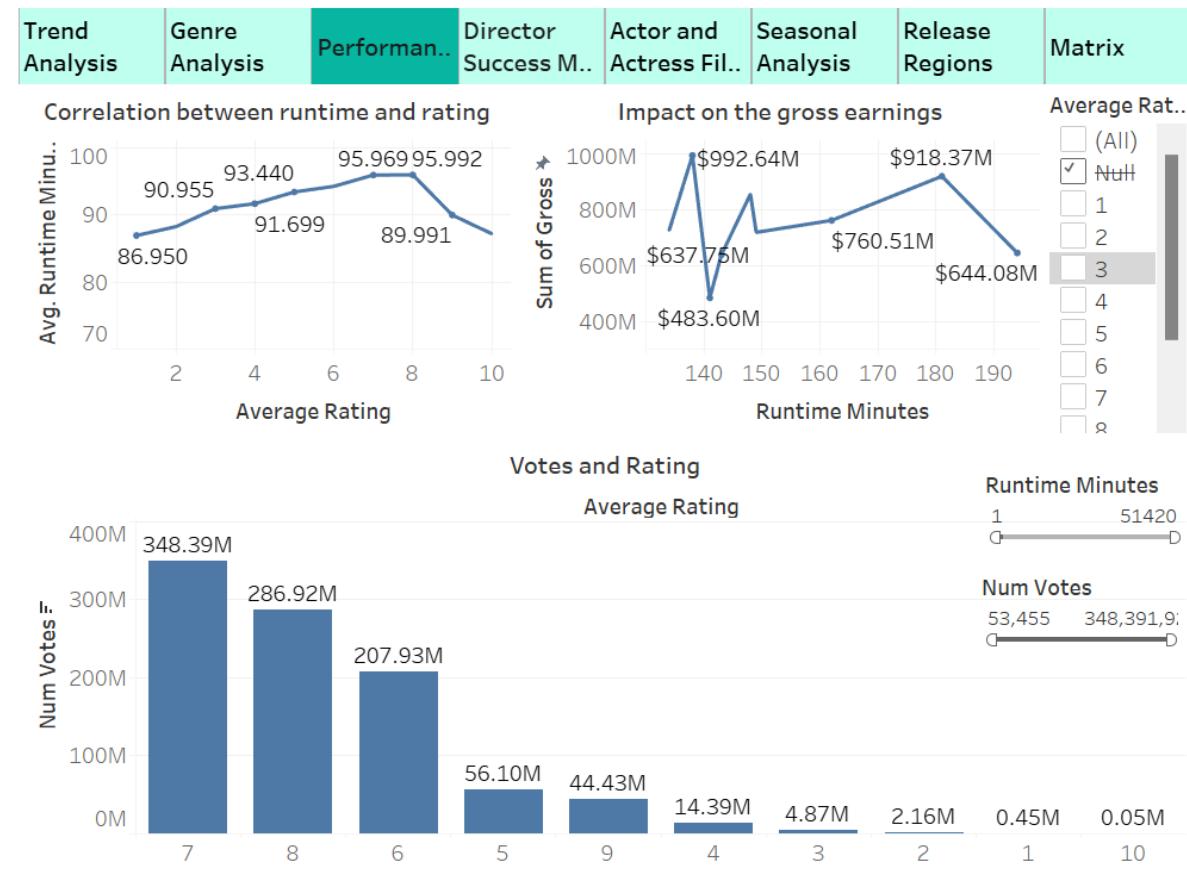
```
58 • select averageRating, sum(numVotes) from fct_rating  
59 group by averageRating;  
60
```

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



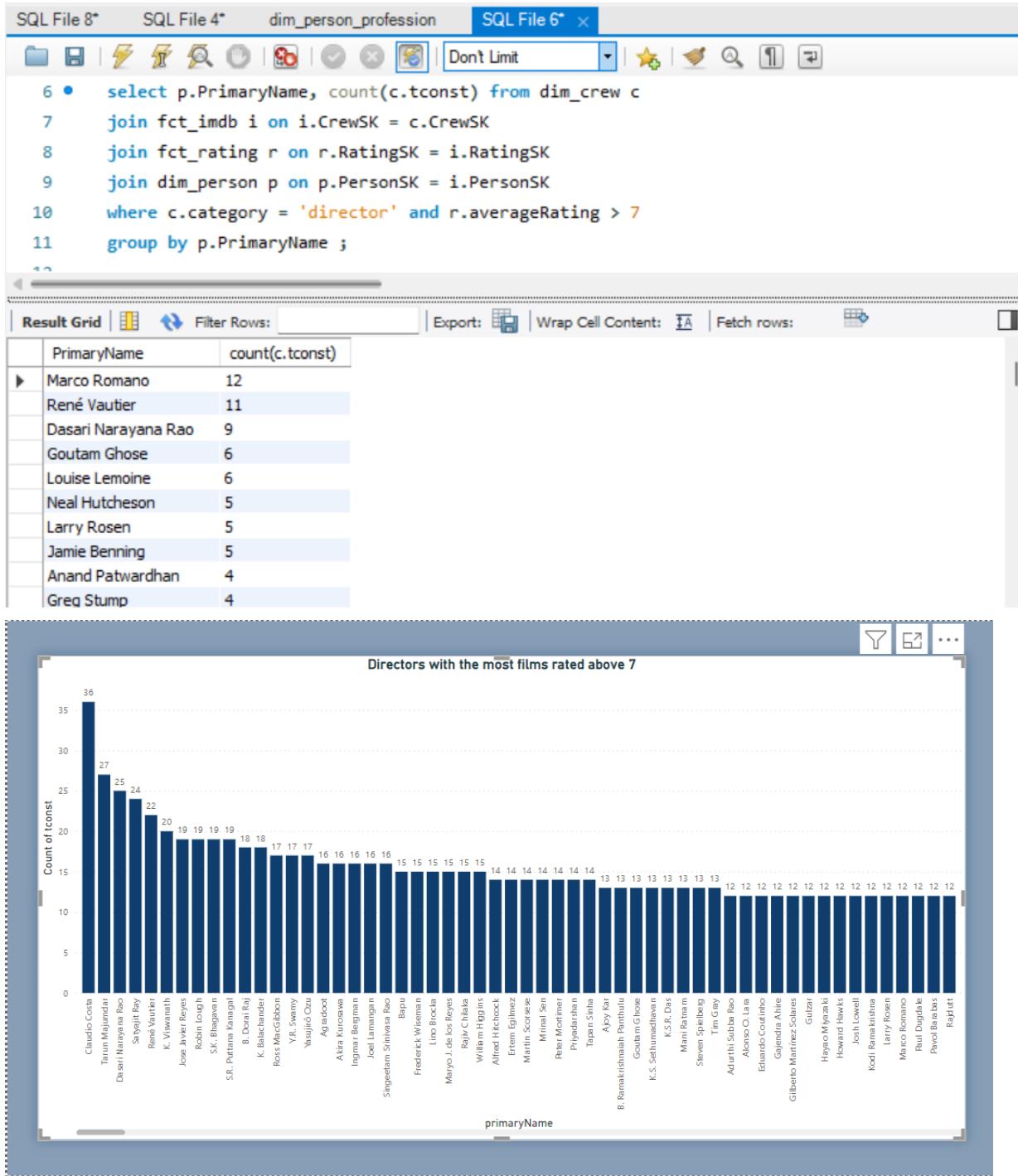
PERFORMANCE METRICS



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 ;
```



```

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 execution of the provided SQL code and its results.

The query (labeled 43) is:

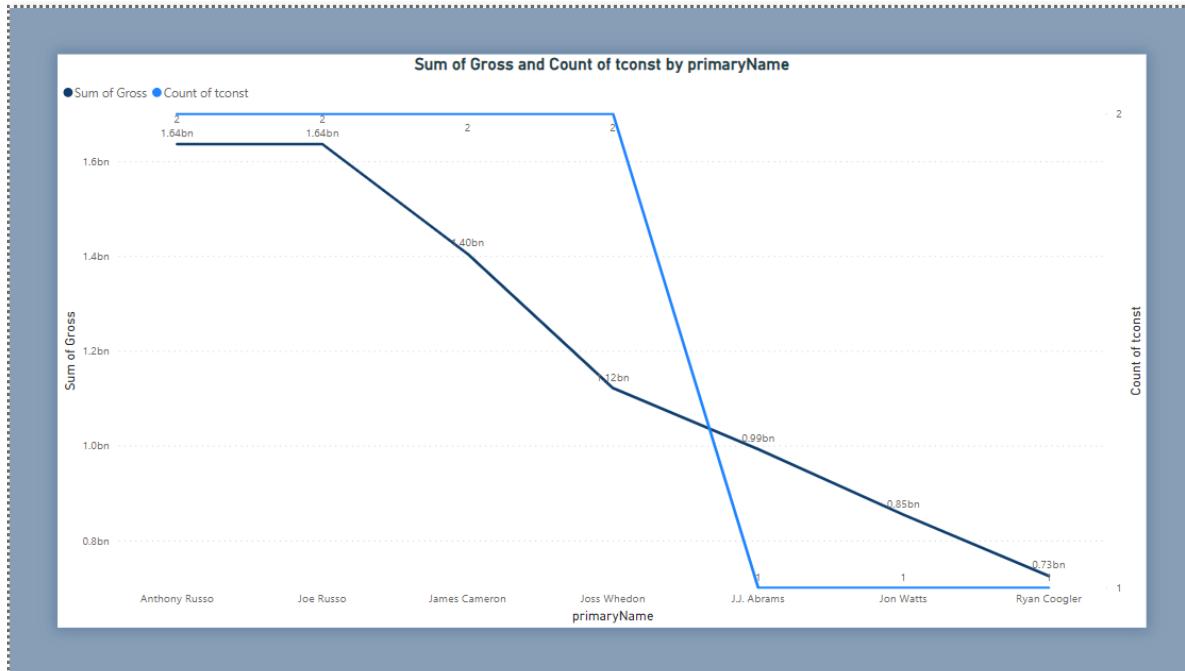
```

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 Result Grid displays the following data:

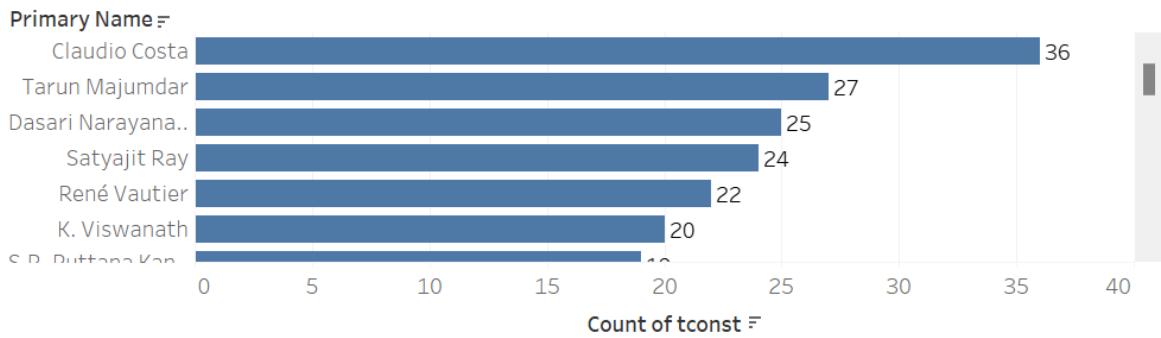
	movie_count	PrimaryName	category	total_gross_earnings
▶	2	Anthony Russo	director	1636188482
	2	Joe Russo	director	1636188482
	2	James Cameron	director	1404587263
	2	Joss Whedon	director	1121346625
	1	J.J. Abrams	director	992642689
	1	Jon Watts	director	854337488



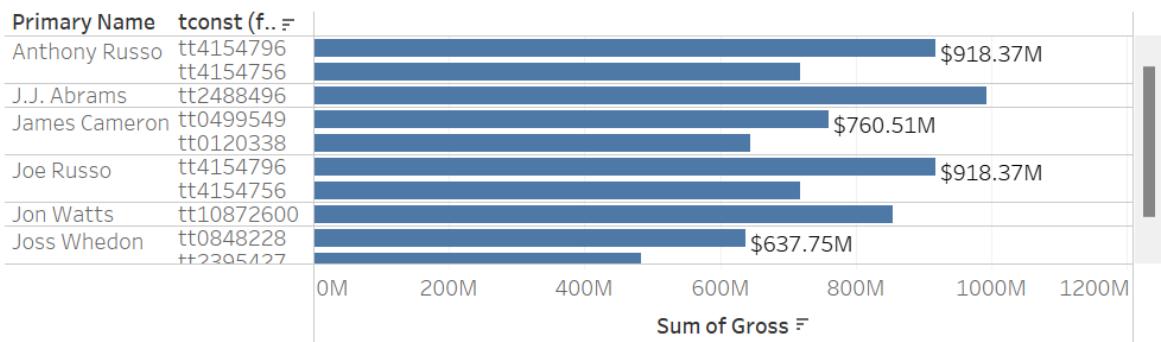
DIRECTOR SUCCESS METRICS

Trend Analysis	Genre Analysis	Performance	Director Success M..	Actor and Actress Fil..	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.

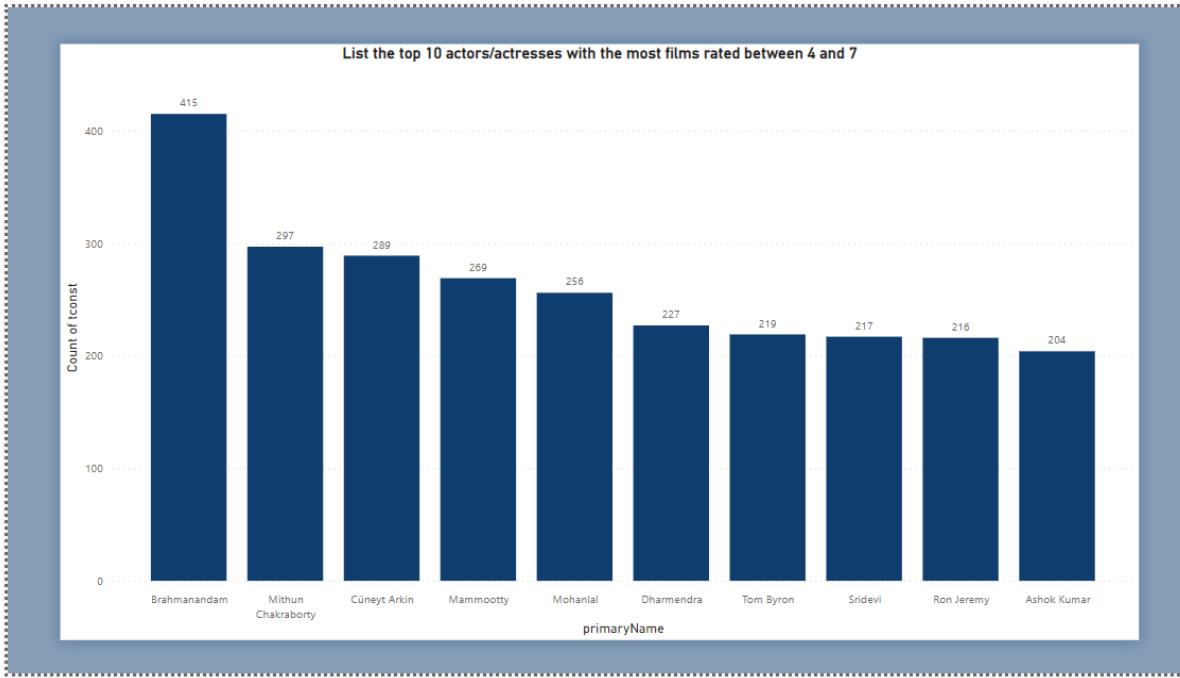
a. The goal is to identify the popular actors/actresses

```
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 code with line numbers 1 through 10. Line 1 is marked with a blue dot indicating the current line.
- Result Grid:** A table showing the results of the query. The columns are labeled "actor_name" and "total_films_between_4_and_7". The data is as follows:

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

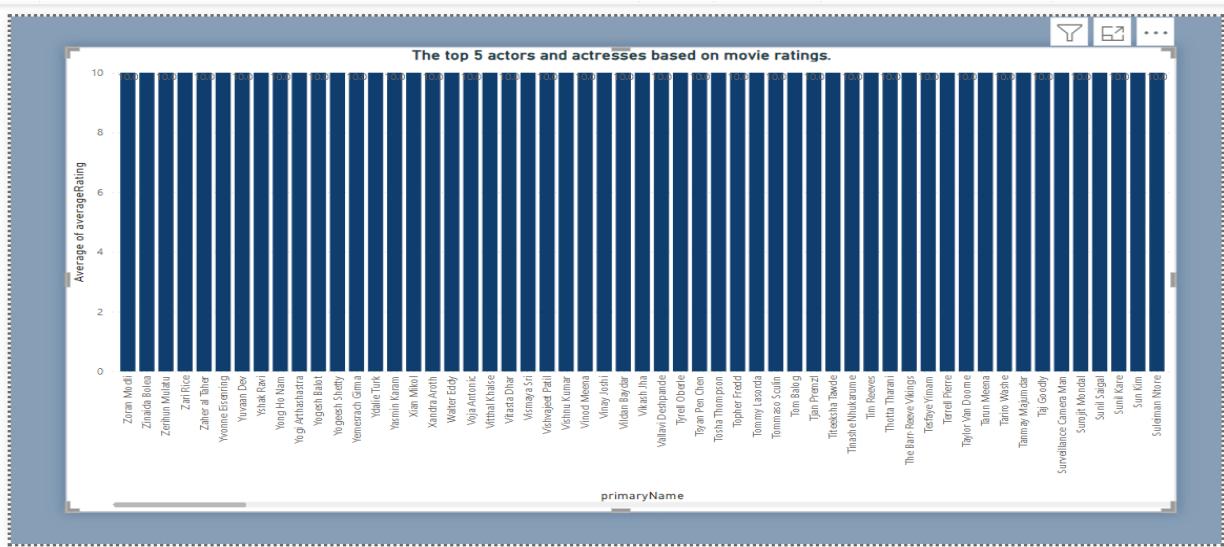
Result 3 ×

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																																											
Primary Name																																											
<table border="1"> <thead> <tr> <th>Actor</th> <th>Distinct count of Tconst (Dim Crew)</th> </tr> </thead> <tbody> <tr><td>Brahmanandam</td><td>415</td></tr> <tr><td>Mithun Chakrab..</td><td>297</td></tr> <tr><td>Cüneyt Arkin</td><td>289</td></tr> <tr><td>Mammootty</td><td>269</td></tr> <tr><td>Mohanlal</td><td>256</td></tr> </tbody> </table>								Actor	Distinct count of Tconst (Dim Crew)	Brahmanandam	415	Mithun Chakrab..	297	Cüneyt Arkin	289	Mammootty	269	Mohanlal	256																								
Actor	Distinct count of Tconst (Dim Crew)																																										
Brahmanandam	415																																										
Mithun Chakrab..	297																																										
Cüneyt Arkin	289																																										
Mammootty	269																																										
Mohanlal	256																																										
Top 5 by rating																																											
<table border="1"> <thead> <tr> <th>Primary Name</th> <th>Avg. ...</th> <th>Category</th> </tr> </thead> <tbody> <tr><td>Zoran Modli</td><td>10</td><td>(All)</td></tr> <tr><td>Zinaida Bolea</td><td>10</td><td>Null</td></tr> <tr><td>Zerihun Mulatu</td><td>10</td><td>Actor</td></tr> <tr><td>Zari Rice</td><td>10</td><td>Actress</td></tr> <tr><td>Zaher al Taher</td><td>10</td><td>archive_fo...</td></tr> <tr><td>Yvonne Eisenring</td><td>10</td><td>archive_so...</td></tr> <tr><td>Yuvaan Dev</td><td>10</td><td>cinematogr...</td></tr> <tr><td>Yshak Ravi</td><td>10</td><td>composer</td></tr> <tr><td>Yong Ho Nam</td><td>10</td><td>director</td></tr> <tr><td>Yogi Arthashastra</td><td>10</td><td>editor</td></tr> <tr><td>Yogesh Balot</td><td>10</td><td>None</td></tr> </tbody> </table>								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	None
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	None																																									



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

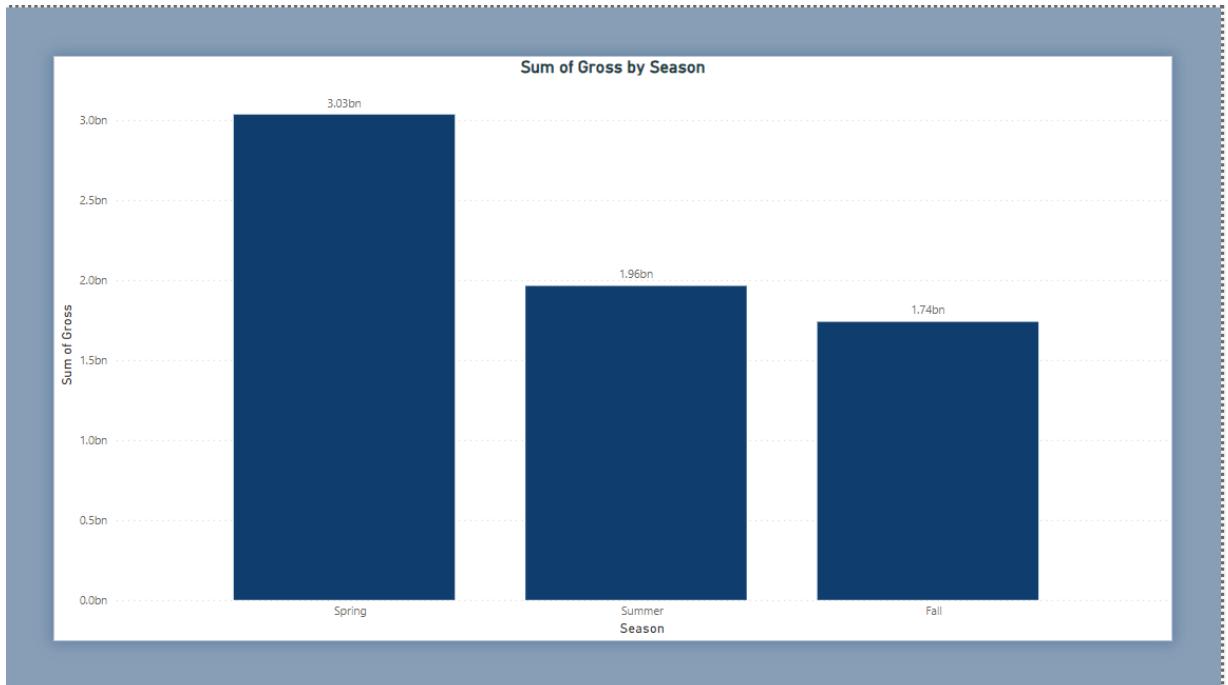
311 season

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;

```

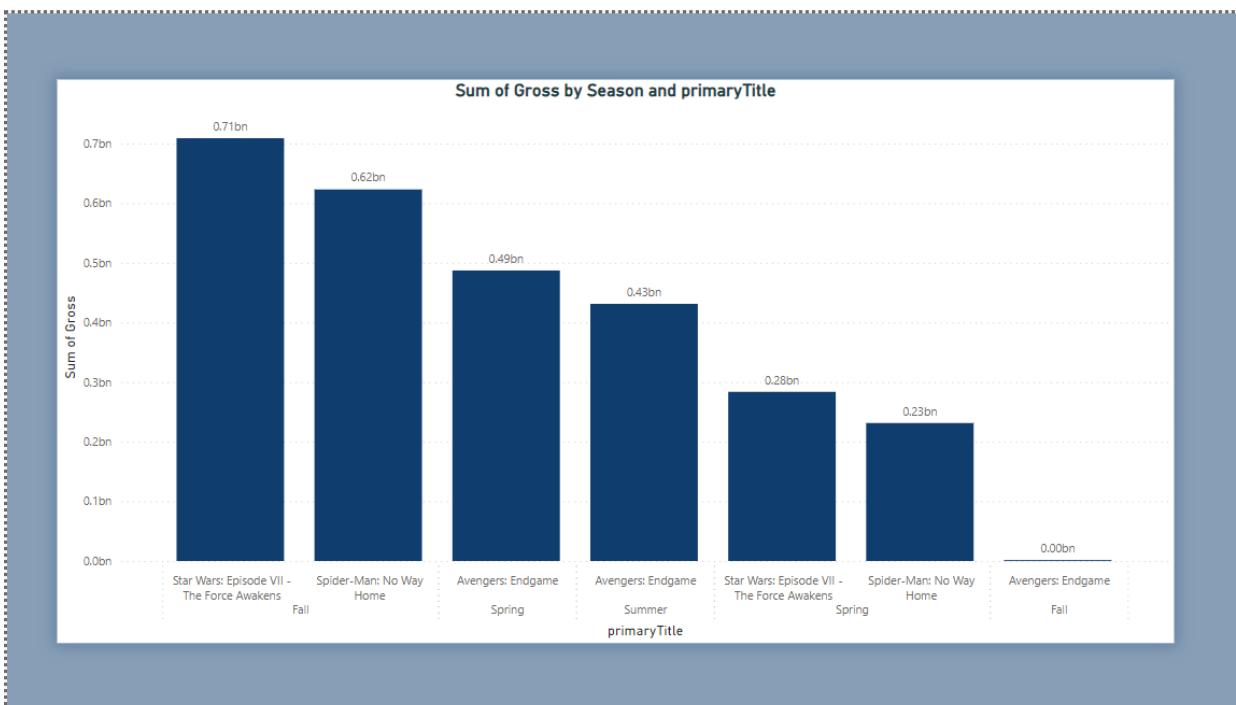
```

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

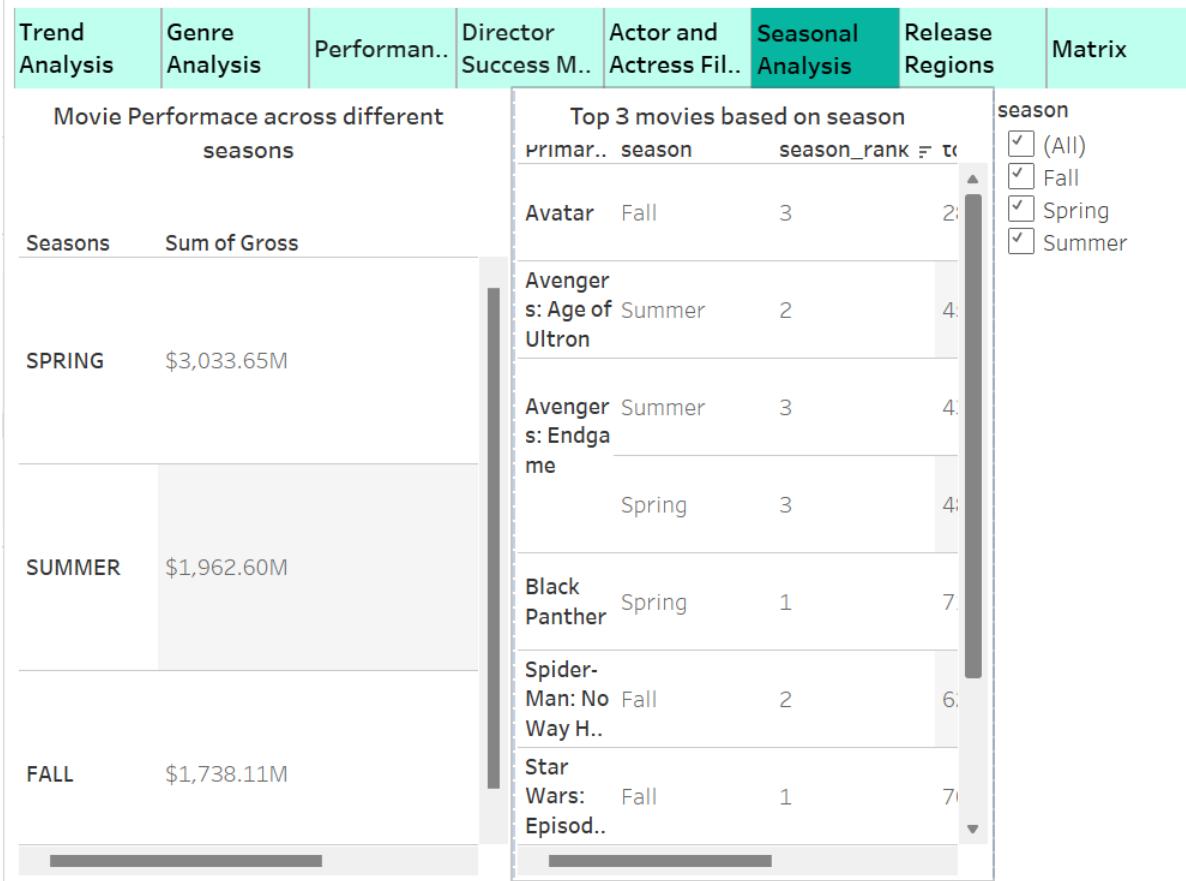
```

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: |

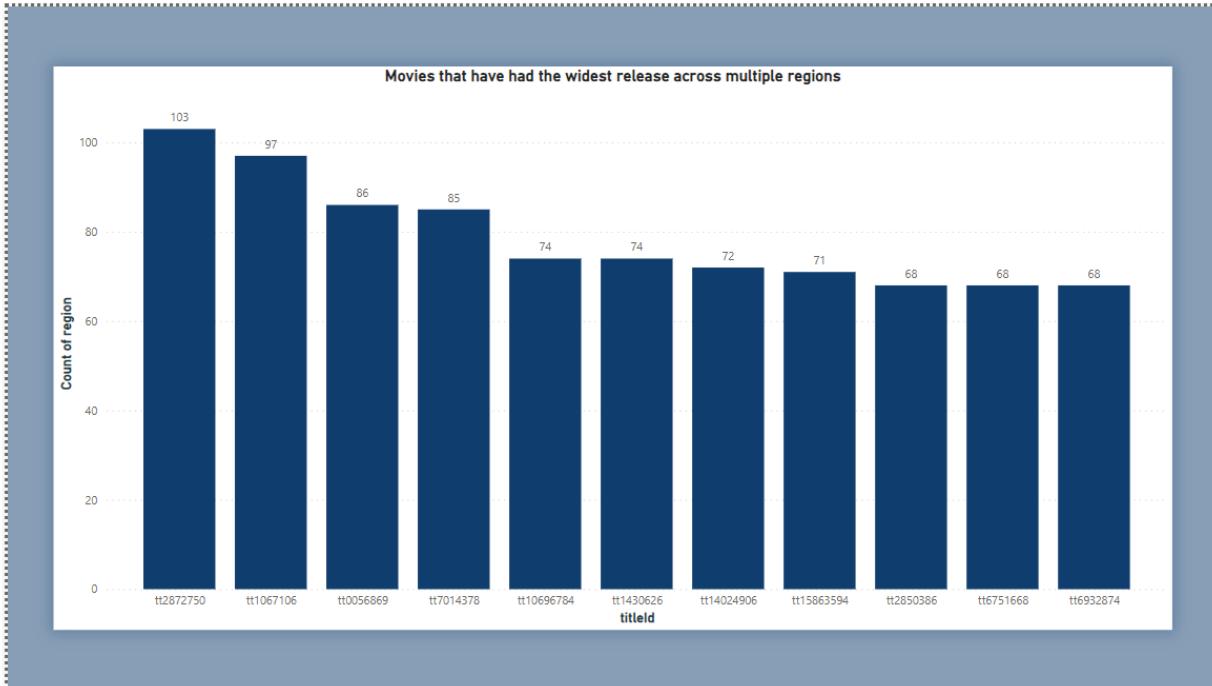
	titleId	movie_name	region_count
▶	tt2872750	笑笑羊大電影	103
	tt1067106	聖誕夜怪譚	97
	tt0056869	鳥	86
	tt7014378	神之山嶺	85
	tt1430626	海賊天團	74
	tt10696784	復仇之淵	74
	tt14024906	Paranoid: Hersey sil bastan	72

RELEASE REGIONS

Trend Analysis | Genre Analysis | Performance Analysis | Director Success Metrics | Actor and Actress Filmography | 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 Hidden World	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 the Past	1	Abc
tt10696784	Vực Sâu Thủ Hân	1	Abc
tt1067106	Navidad divertida de Scrooge	1	Abc



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 [REDACTED]
536         dim_movie_title_mysql me ON fe.tconst = me.tconst
537     JOIN [REDACTED]
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: Export: 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 with 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

```

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