

Data Flow pipelines and Load

Stage Data Loading

1] title.crew dataset

The screenshot shows a Jupyter Notebook interface with the following details:

- Header: 02-16 8:17pm, 2025-02-16 8:33pm, 2025-02-16 8:59pm, 2025-02-20 11:47am, Midterm Project - IMDB, 2025-03-09 7:22pm.
- User: ACCOUNTADMIN, COMPUTE_WH (X-Small).
- Code Versions: Code Versions, Share, +, -.
- SQL Query:

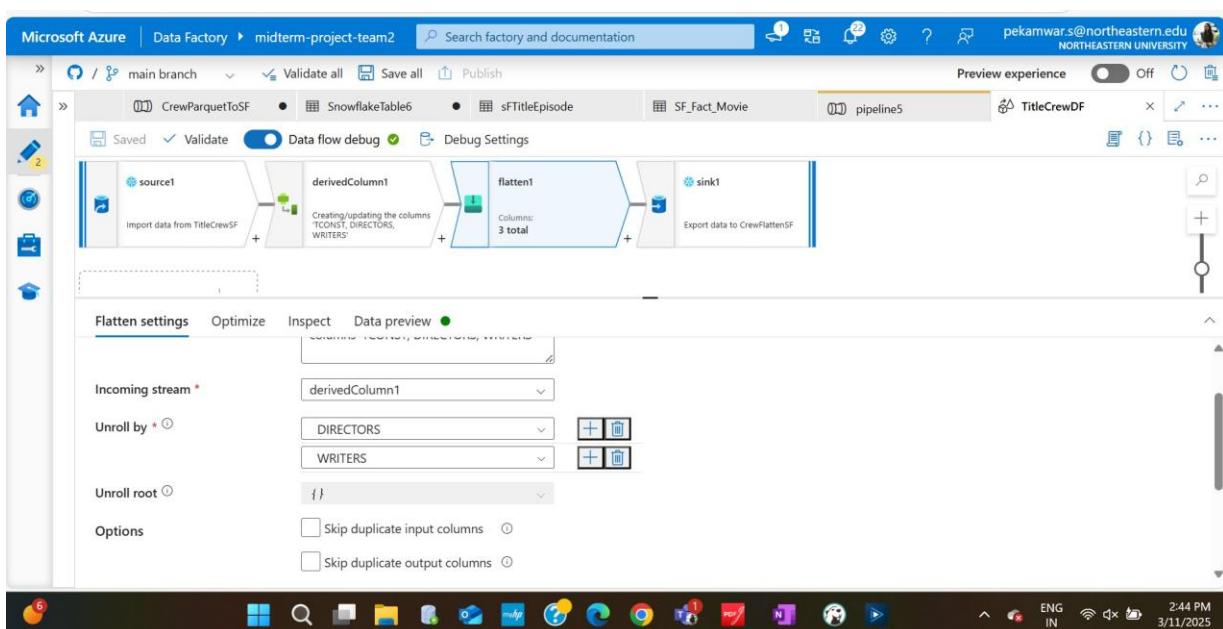
```
1 select * from title_crew where tconst = 'tt10930678';
2 select * from title_crew_flatten where tconst = 'tt10930678';
3
4
5 select * from title_basics;
```

- Results: Shows a table with columns TCONST, DIRECTORS, and WRITERS. The data is as follows:

	TCONST	DIRECTORS	WRITERS
1	tt042444	nm0059223,nm0083290,nm0123364,nm0201234	nm0894640
2	tt042446	nm10145312	nm10145312
3	tt042448	nm0702142	nm5107308
4	tt042450	nm0442687	NA
5	tt042452	nm0776255	nm1075143,nm2360862,nm2917457,nm10145312
6	tt042454	nm8412331	nm8412331
7	tt042456	nm4458749	nm4458749

- Query Details: Query duration 3.1s, Rows 115M, Query ID 01baf1f-0001-a4af-005...
- Ask Copilot button.

As the data in writers and directors column has values in the form of string, the data is flattened using adf pipeline



02-16 8:17pm 2025-02-16 8:33pm 2025-02-16 8:59pm 2025-02-20 11:47am Midterm Project - IMDB 2025-03-09 7:22pm + ▾

IMDB_DB.IMDB_SCHEMA Settings

```
1 select * from title_crew where tconst = 'tt10930678';
2 select * from title_crew_flatten where tconst = 'tt10930678';
3
4
5 select * from title_basics;
```

Code Versions Q

Results Chart

	A TCONST	A DIRECTORS	A WRITERS
1	tt1093067	nm0284159	nm0001468
2	tt1093067	nm0284159	nm0611848
3	tt1093067	nm0284159	nm1669351
4	tt1093067	nm0284159	nm0750523
5	tt1093067	nm0284159	nm2227704
6	tt1093067	nm0284159	nm0996024
7	tt1093067	nm0284159	nm0780980

Query Details ...
 Query duration 5.0s
 Rows 27.6M
 Query ID 01baf122-0001-a4b0-0...
 Show more Ask Copilot

2] title.episode

02-16 8:17pm 2025-02-16 8:33pm 2025-02-16 8:59pm 2025-02-20 11:47am Midterm Project - IMDB 2025-03-09 7:22pm + ▾

IMDB_DB.IMDB_SCHEMA Settings

```
8 select * from title_episode;
9 describe table title_episode;
10
11 select * from title_principals;
```

Code Versions Q

Results Chart

	A TCONST	A PARENTTCONST	# SEASONNUMBI	# EPISODENUMBI	A DI_JOB_ID
1	tt12250694	tt8859520	0	0	24e7bfaa-a0d1-4f5e-9f36-b51ff096e4a0
2	tt12250696	tt8859520	0	0	24e7bfaa-a0d1-4f5e-9f36-b51ff096e4a0
3	tt12250698	tt8859520	0	0	24e7bfaa-a0d1-4f5e-9f36-b51ff096e4a0
4	tt1225070	tt0106086	1	9	24e7bfaa-a0d1-4f5e-9f36-b51ff096e4a0
5	tt12250700	tt4151188	0	0	24e7bfaa-a0d1-4f5e-9f36-b51ff096e4a0
6	tt12250702	tt8859520	0	0	24e7bfaa-a0d1-4f5e-9f36-b51ff096e4a0

Query Details ...
 Query duration 80ms
 Rows 26.5M
 Query ID 01baf125-0001-a521-0...
 Show more Ask Copilot

3] title.principals

https://app.snowflake.com/east-us-2.azure/nua03035/worLh40kjUD#query

02-16 8:17pm 2025-02-16 8:33pm 2025-02-16 8:59pm 2025-02-20 11:47am Midterm Project - IMDB 2025-03-09 7:22pm +

ACCOUNTADMIN COMPUTE_WH (X-Small) Share

IMDB_DB.IMDB_SCHEMA Settings

```

11 | select * from title_principals;
12 |
13 | select * from title_akas;
14 | select * from title_akas_flatten where attributes not like 'NA';
15 |

```

Results Chart

T	CATEGORY	JOB	CHARACTERS	DI_JOB_ID	DI_LOAD_DT
1	director	NA	NA	e9b40ce0-fbb4-45e1-e	2025-03-10
2	self	NA	["Self - Host"]	e9b40ce0-fbb4-45e1-e	2025-03-10
3	actress	NA	NA	e9b40ce0-fbb4-45e1-e	2025-03-10
4	self	NA	["Presenter"]	e9b40ce0-fbb4-45e1-e	2025-03-10
5	actor	NA	["DueÃ±o Casa"]	e9b40ce0-fbb4-45e1-e	2025-03-10
6	director	director	NA	e9b40ce0-fbb4-45e1-e	2025-03-10

Query Details

- Query duration: 48ms
- Rows: 91.3M
- Query ID: 01baf125-0001-a51f-00...

Show more Ask Copilot

4] title.akas

https://app.snowflake.com/east-us-2.azure/nua03035/worLh40kjUD#query

02-16 8:17pm 2025-02-16 8:33pm 2025-02-16 8:59pm 2025-02-20 11:47am Midterm Project - IMDB 2025-03-09 7:22pm +

ACCOUNTADMIN COMPUTE_WH (X-Small) Share

IMDB_DB.IMDB_SCHEMA Settings

```

12 |
13 | select * from title_akas;
14 | select * from title_akas_flatten where attributes not like 'NA';
15 |
16 | select * from date_dim;

```

Results Chart

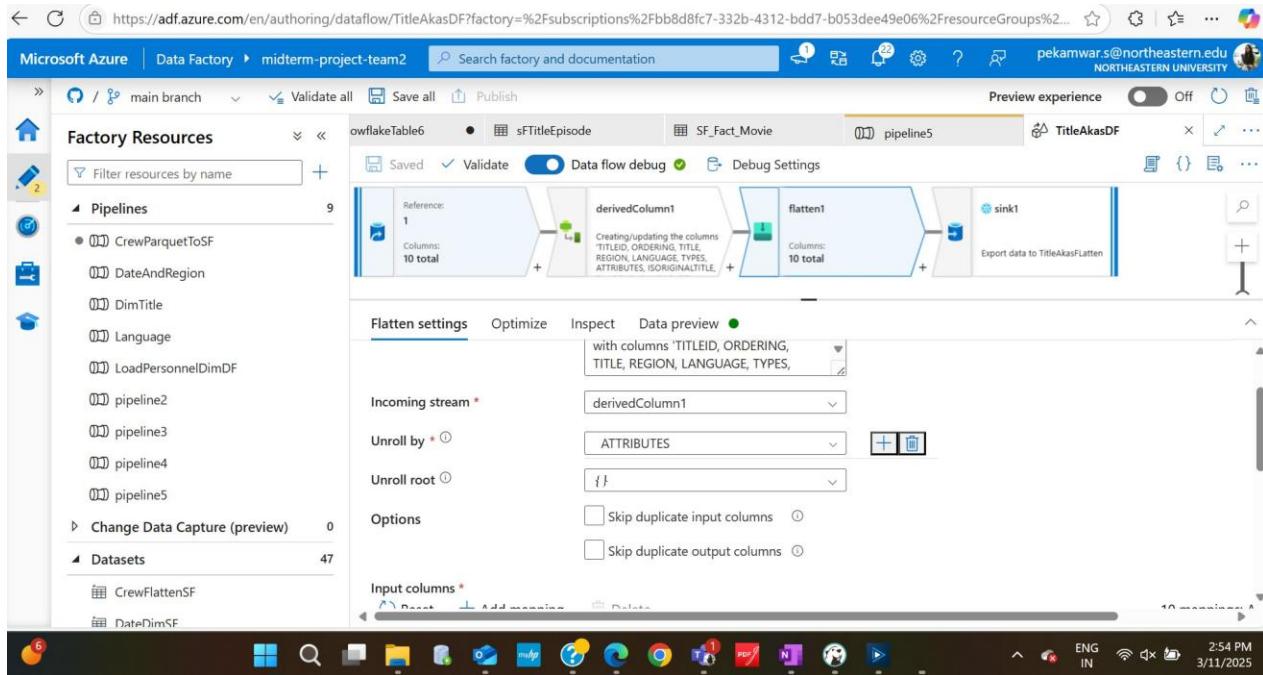
ION	LANGUAC	TYPE	ATTRIBUTI	O1 ISORIGINALTIT	DI_JOB_ID	DI_LOAD_DT
1	ja	NA	NA	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
2	es	NA	NA	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
3	n	Unknown	original	NA	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
4	pt	NA	NA	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
5	hi	NA	NA	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
6	fr	NA	NA	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10

Query Details

- Query duration: 15s
- Rows: 51.5M
- Query ID: 01baf127-0001-a4b0-0...

Show more Ask Copilot

Here the attributes column is flattened and separated into different columns



Query Results:

```

12
13 select * from title_akas;
14 select * from title_akas_flatten where attributes not like 'NA';
15
16 select * from date_dim;

```

LANGUAGE	TYP	ATTRIBUTES	ISORIGINALTITLE	DI_JOB_ID	DI_LOAD_DT
Unknown	NA	alternative transliteration	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
Unknown	NA	promotional title	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
Unknown	NA	transliterated title	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
en	NA	informal English title	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
cmm	NA	literal title	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10
cmm	NA	literal title	FALSE	33bea6c4-b9bc-420a-b5fe-3ecbe0920ca2	2025-03-10

Query Details:

- Query duration: 45ms
- Rows: 296.4K
- Query ID: 01baf12b-0001-a4af-00...

5] title.ratings

The screenshot shows the Snowflake web interface. The left sidebar displays the database structure under the IMDB_DB schema, including tables like BRIDGE_MOVIE, DATE_DIM, LANGUAGE_DIM, NAME_BASIC, NAME_BASIC_FLATTEN, and TITLE_RATINGS. The main pane shows a query script:

```
--TABLES FROM DBEAVER  
37 | SELECT * FROM TITLE_RATINGS;  
38 | SELECT * FROM NAME_BASIC;  
39 | SELECT * FROM NAME_BASIC_FLATTEN;
```

The results pane displays the following data:

	A TCONST	# AVERAGERATING	# NUMVOTES
1	TT0580671	8	9
2	TT2342059	7.9	26
3	TT11213434	9.7	31
4	TT0454230	6.1	60
5	TT6464426	7.6	111
6	TT1524275	8.2	6
7	TT8488538	7.2	78
8	TT1563143	7.2	19

Query Details:
Query duration: 43ms
Rows: 4.6M
Query ID: 01baf16d-0001-a51f-00...
Show more

Ask Copilot

6] name.basics

The screenshot shows the Snowflake web interface. The left sidebar displays the database structure under the IMDB_DB schema, including tables like BRIDGE_MOVIE, DATE_DIM, LANGUAGE_DIM, NAME_BASIC, NAME_BASIC_FLATTEN, and TITLE_RATINGS. The main pane shows a query script:

```
GRANT ROLE IMDB_Role TO ROLE ACCOUNTADMIN;  
--TABLES FROM DBEAVER  
37 | SELECT * FROM TITLE_RATINGS;
```

The results pane displays the following data:

	A NCONST	A PRIMARYNAME	# BIRTHYEAR	# DEATHYEAR	A PRIMARYPROFESSION	A KNOWNFORTITLES
1	nm10667873	Reine Fassi	0	0	NA	tt4964258
2	nm10667874	Roger Emmenecker	0	0	NA	tt4964258
3	nm10667875	Alain Saustier	0	0	NA	tt4964258
4	nm10667876	Jacques-Alain de Sedouy	1935	2018	NA	tt4964258
5	nm10667877	Lucien Zinger	0	0	NA	tt4964258
6	nm10667878	Manon Pignot	0	0	NA	tt4964258
7	nm10667879	Shelton Sallie	0	0	actor	tt4467154
8	nm10667888	Jan Verbraecken	1923	1983	actor,soundtrack	tt3459472,tt0948548,tt

Query Details:
Query duration: 6.4s
Rows: 14.2M
Query ID: 01baf175-0001-a51f-00...
Show more

Ask Copilot

Factory Resources

- dataflow4
- date_dimDL
- LanguageDimDF
- LoadFactMovie
- LoadGenreDim
- LoadPersonnelDim
- NameBasicsFlatten**
- RegionDF
- title_ratingsDL
- TitleAkasDF
- TitleBasicsDF
- TitleCrewDF
- TitleDimDF

Flatten settings

Output stream name *: flatten1

Description: Unrolling arrays from derivedColumn1 with columns 'NCONST, PRIMARYNAME, BIRTHYEAR, DEATHYEAR, PRIMARYPROFESSION,'

Incoming stream *: derivedColumn1

Unroll by *: PRIMARYPROFESSION, KNOWNFORTITLES

Unroll root: { }

Options: Skip duplicate input columns

IMDB Script

```

38 | SELECT * FROM NAME_BASICS;
39 | SELECT * FROM NAME_BASICS_FLATTEN;
40 | SELECT * FROM DATE_DIM;
41 | SELECT * FROM LANGUAGE_DIM;
42 | SELECT * FROM BRIDGE_MOVIE;
43 |

```

Results

NCONST	PRIMARYNAME	BIRTHYEAR	DEATHYEAR	PRIMARYPROFESSION	KNOWNFORTITLES
nm13598923	Katarina Fiallos	0	0	actor	tt28437533
nm13598923	Katarina Fiallos	0	0	actress	tt27155819
nm13598923	Katarina Fiallos	0	0	actress	tt28437533
nm13598924	Jonah Brown	0	0	NA	NA
nm13598925	Ratul Ahmed	0	0	NA	NA
nm13598926	Hannah	0	0	actress	tt21524284
nm13598927	Bob Jutson	0	0	actor	tt1260682
nm13598928	Sierra Segalowitz	0	0	actress	tt18183178

Query Details

- Query duration: 38ms
- Rows: 38,7M
- Query ID: 01baef180-0001-a51f-00...

7] title.basics

The screenshot shows a database interface with a query editor and a results table. The query is:

```
4
5   select * from title_basics;
6   select * from title_basics_flatten;
7
8   select * from title_episode;
```

The results table has columns: TCONST, TITLETYPE, PRIMARYTITLE, ORIGINALTITLE, ISADULT, and STAR. The data includes:

TCONST	TITLETYPE	PRIMARYTITLE	ORIGINALTITLE	ISADULT	STAR
tt17592272	tvEpisode	Episode #1.29	Episode #1.29	FALSE	
tt7079292	tvMovie	Helping Hands	Helping Hands	FALSE	
tt277626	tvEpisode	My Dad's Got Fans?!	My Dad's Got Fans?!	FALSE	
tt13563072	tvEpisode	The Sandys	The Sandys	FALSE	
tt30789616	tvEpisode	The Agony And Ecstasy Of 2023	The Agony And Ecstasy Of 2023	FALSE	
tt13178382	tvEpisode	Armed and Delicious	Armed and Delicious	FALSE	

Query Details: Query duration 6.5s, Rows 11.5M, Query ID 01baf132-0001-a521-00...

Here the flattening of columns on Genres, OriginalTitle and PrimaryTitle into a new table

The screenshot shows a database interface with a query editor and a results table. The query is:

```
4
5   select * from title_basics;
6   select * from title_basics_flatten;
7
8   select * from title_episode;
```

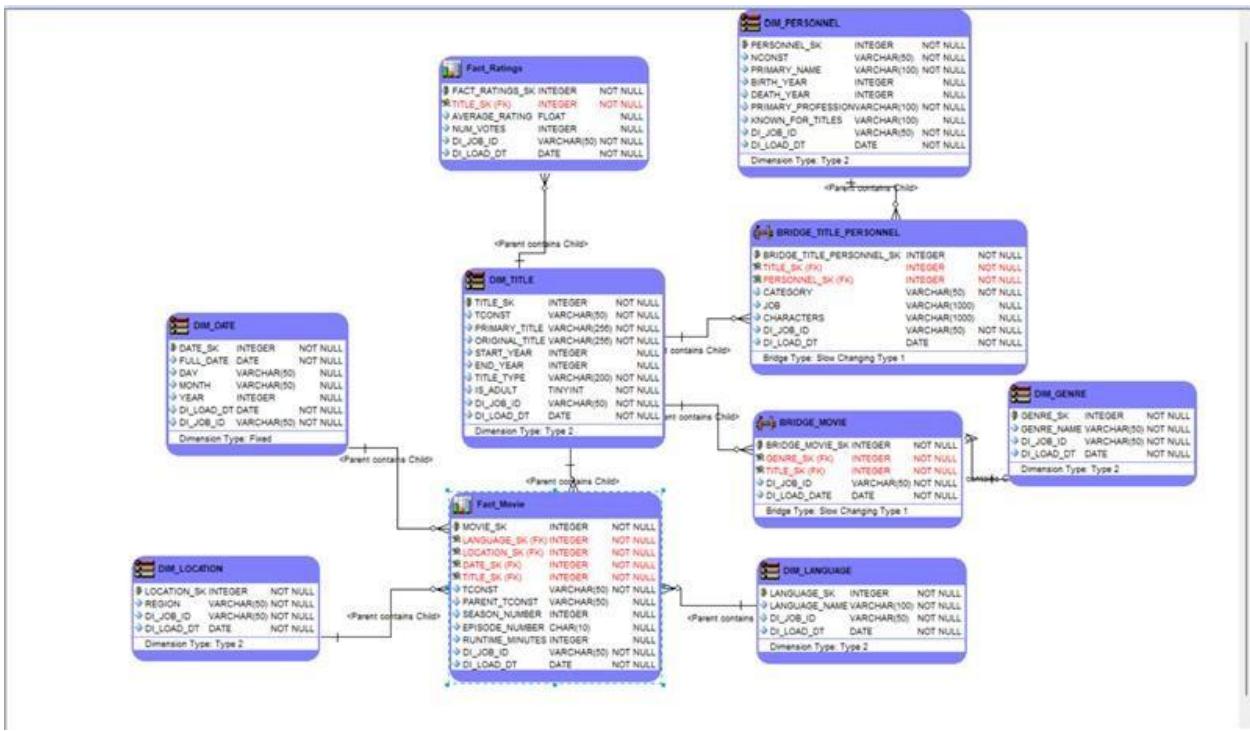
The results table has columns: TCONST, TITLETYPE, PRIMARYTITLE, ORIGINALTITLE, ISADULT, and STAR. The data includes:

TCONST	TITLETYPE	PRIMARYTITLE	ORIGINALTITLE	ISADULT	STAR
tt4833930	tvEpisode	True Cheatin' Whore	True Cheatin' Whore	TRUE	
tt31242955	tvEpisode	Episode #1.8	Episode #1.8	FALSE	
tt14042862	tvEpisode	Episode #1.50	Episode #1.50	FALSE	
tt29931681	tvEpisode	Skye's Bathroom Video	Skye's Bathroom Video	TRUE	
tt1986887	movie	Rain or Shine	La pluie et le beau temp	FALSE	
tt6459118	tvEpisode	Capones Vag	Capones Vag	TRUE	

Query Details: Query duration 7.3s, Rows 18.3M, Query ID 01baf138-0001-a521-00...

Loading Dims and Facts Tables :

ER Studio Diagram :



1] Date_Dim

https://app.snowflake.com/east-us-2.azure/nua03035/worLh40kjUD#query

```

15
16 | select * from date_dim;
17 | select * from genre_dim;
18
19 DESCRIBE TABLE genre_dim;

```

	DATE	DAY_OF_WEEK	MONTH_NAME	YEAR_NUM	DI_JOB_ID	DI_LOAD_DT
1	2021-01-01	Fri	January	2021	2f755f14-f74b-4612-9a47-c886123b373c	2025-03-11
2	2021-01-02	Sat	January	2021	c690f2ca-a330-4f8b-af91-416eae6de57d	2025-03-11
3	2021-01-03	Sun	January	2021	776302a3-0568-4399-89dd-034d9c988631	2025-03-11
4	2021-01-04	Mon	January	2021	77bd42ca-df6c-4c4a-97e4-4df6c4ef8ae9	2025-03-11
5	2021-01-05	Tue	January	2021	2d026e9e-ec98-48d6-9a6e-9e609a6b173a	2025-03-11
6	2021-01-06	Wed	January	2021	4c327f0e-d1bd-4325-ad18-b91a3f5b788b	2025-03-11
7	2021-01-07	Thu	January	2021	4d82a5c5-ed5e-4c92-9815-bdaa7622628c	2025-03-11

Query Details
Query duration 76ms
Rows 1.8K
Query ID 01baf13c-0001-a4b0-0...

2] Genre _Dim

```

15
16 | select * from date_dim;
17 | select * from genre_dim;
18
19 DESCRIBE TABLE genre_dim;

```

	GENRE_SK	GENRE	DI_JOB_ID	DI_LOAD_DT
1	8847361	Comedy	cc9265bc-c3ce-4431-95ad-6296f8542e78	2025-03-11
2	8847362	Drama	68b283e5-1505-47e9-86f0-61ac3caa89b3	2025-03-11
3	8847363	Drama	9cf3ada4-869f-4138-b5a9-984ee28b9fb3	2025-03-11
4	8847364	Romance	4d239280-a4f1-4025-a864-1da3fa86820f	2025-03-11
5	8847365	Adult	26b0f967-6fb0-459d-9df0-de1d283df778	2025-03-11
6	8847366	Crime	a833082c-d8a3-4583-8b77-49c736337493	2025-03-11
7	8847367	Drama	538ead83-5030-48d6-a0e9-fa5e741d294e	2025-03-11

Query Details
Query duration 55ms
Rows 18.3M
Query ID 01baf13d-0001-a51f-00...

3] Region_Dim

02-16 8:17pm 2025-02-16 8:33pm 2025-02-16 8:59pm 2025-02-20 11:47am Midterm Project - IMDB 2025-03-09 7:22pm + ⚙

IMDB_DB.IMDB_SCHEMA Settings

```

24 GRANT INSERT ON ALL TABLES IN SCHEMA IMDB_SCHEMA TO ROLE IMDB_OWNER;
25
26 ┌─| SELECT * FROM REGION_DIM;
27
28 ┌─| SELECT * FROM GENRE_DIM;

```

Code Versions Share

Results Chart

	REGION_CODE	REGION_NAME	DI_JOB_ID	DI_LOAD_DT	REGION_DIM_SK
1	AF	Afghanistan	96698ed8-f75a-4b70-b562-3ce755840dcc	2025-03-11	1
2	AX	Åland Islands	cbccfbbe-8d88-4008-9c3c-20406c41fe06	2025-03-11	2
3	AL	Albania	69c74cae-33f4-44d7-9019-274b5f4a3065	2025-03-11	3
4	DZ	Algeria	eb1cb253-df29-45a6-ad51-61a4b01e51a3	2025-03-11	4
5	AS	American Samoa	d383a1f1-60f2-4b25-8595-66baa93f494c	2025-03-11	5
6	AD	Andorra	bdaee182-47c8-4603-ba95-d2d2b10c4462	2025-03-11	6
7	AO	Angola	fb69649b-376c-477d-b419-33b901ea11c5	2025-03-11	7

Query Details

- Query duration 117ms
- Rows 228
- Query ID Q1baf13d-0001-a4b0-0...

Show more

REGION_CODE Ask Copilot

REGION_CODE

Ask Copilot

3:09 PM 3/11/2025 ENG IN

4] Language_Dim

Midterm-Proj Dataset Analy IMDB Script ! announcement Commits - Sh Data Profiling Data Load Fl > +

Sample queries on TPC-H ... 2025-02-06 12:12pm 2025-02-20 4:26pm IMDB Script

IMDB_DB.IMDB_SCHEMA Settings

```

38 SELECT * FROM NAME_BASICS;
39 SELECT * FROM NAME_BASICS_FLATTEN;
40 SELECT * FROM DATE_DIM;
41 ┌─| SELECT * FROM LANGUAGE_DIM;
42 SELECT * FROM BRIDGE_MOVIE;
43

```

Code Versions Share

Results Chart

	LANGUAGE_CODE	LANGUAGE_NAME	DI_JOB_ID	DI_LOAD_DT	LANGUAGE_DIM_SK
1	AF	Afrikaans	8eed8e4d-67d8-44bb-881c-ffc8729243b	2025-03-11	1
2	AX	Åland Islands	ad9c8ab5-d662-4379-8296-5f4d1dfdf3ae	2025-03-11	2
3	AL	Albanian	ecfe3a73-a99d-4019-81a4-ae7ede20aa1f	2025-03-11	3
4	DZ	Algerian Arabic	687126db-f9eb-4220-a47f-09aff7273669	2025-03-11	4
5	AS	American Sign Language	a4db349a-40dc-4aac-9fb5-ce3e5bd69497	2025-03-11	5
6	AD	Andorran Catalan	5c033310-a295-4d7e-ae9d-6361875cb50b	2025-03-11	6
7	AO	Angolan Portuguese	b9064141-4b30-4c3e-88d7-e5b3c070f780	2025-03-11	7
8	AI	Anguillian English	e7b7e37e-f9b8-4001-bc18-f35f0c88369c	2025-03-11	8

Query Details

- Query duration 121ms
- Rows 86
- Query ID Q1baf181-0001-a4b0-0...

Show more

LANGUAGE_CODE

100% filled Ask Copilot

LANGUAGE_CODE

Ask Copilot

4:17 PM 3/11/2025 ENG IN

5] Personnel_Dim

02-16 8:17pm 2025-02-16 8:33pm 2025-02-16 8:59pm 2025-02-20 11:47am Midterm Project - IMDB 2025-03-09 7:22pm + ▾

IMDB_DB.IMDB_SCHEMA Settings ▾

```

30
31
> 32   SELECT * FROM REGION_DIM;
33   SELECT * FROM FACT_MOVIE;
34

```

↳ Results ▾ Chart

	# PERSONNEL_SK	▲ NCONST	▲ PRIMARY_NAME	# BIRTH_YEAR	# DEATH_YEAR	▲ PRIM▲
1	23693553	nm1591987	Amaury Gutiérrez	1963	0	soundtrac
2	23693554	nm1591987	Amaury Gutiérrez	1963	0	soundtrac
3	23693555	nm1591987	Amaury Gutiérrez	1963	0	soundtrac
4	23693556	nm1591987	Amaury Gutiérrez	1963	0	soundtrac
5	23693557	nm15919870	Jari Murto	0	0	NA
6	23693558	nm15919871	Jyri Komulainen	0	0	NA

Query Details ...
 Query duration 18s
 Rows 38.7M
 Query ID 01baf13e-0001-a4af-00...
 Show more ▾

Ask Copilot

6] Title_Dim

midtermproject Midterm Project - IMDB dang737 INFO6105 Shrya-Pei Google Drive Data Profiling Share results Power BI +

damsgspring25 mypd 2025-02-17 8:57pm 2025-02-13 2:24pm 2025-02-23 4:58pm 2025-02-06 6:18pm midtermproject +

IMDB_DB.IMDB_SCHEMA Settings ▾

```

57 ORDER BY DI_LOAD_DT DESC
58 LIMIT 10;
59
60 DESC TABLE IMDB_SCHEMA.BRIDGE_TITLE_PERSONNEL;
61
62 select * from BRIDGE_TITLE_PERSONNEL;
63
64
65 select * from dim_title;
66

```

↳ Results ▾ Chart

# TITLE_SK	▲ TCONST	▲ PRIMARY_TITLE	▲ ORIGINAL_TITLE	# START_YEAR	# END_YEAR	▲ TITLE_TYPE	0 IS_A
1	4209535	tt4092626	Die Laubenpieper	1963	0	tvSeries	FALSE
2	4209536	tt3570924	Episode #1.11	1955	0	tvEpisode	FALSE
3	4209537	tt33086649	Episode #2.8	1962	0	tvEpisode	FALSE
4	4209538	tt33086649	Episode #2.8	1962	0	tvEpisode	FALSE
5	4209539	tt33086649	Episode #2.8	1962	0	tvEpisode	FALSE
6	4209540	tt12920810	Prisoners of the Body	2020	0	short	FALSE
7	4209541	tt9508736	Episode #1.209	1996	0	tvEpisode	FALSE
8	4209542	tt1881991	Episode #1.21	2011	0	tvEpisode	FALSE
9	4209543	tt6411566	The Letter	2017	0	short	FALSE
10	4209544	tt6411566	The Letter	2017	0	short	FALSE

Query Details ...
 Query duration 23ms
 Rows 18.3M
 Query ID 01baf303-0001-a4af-00...
 Show more ▾

Ask Copilot

FACTS :

1] Fact_Movie

The screenshot shows the Snowflake web interface. The URL is https://app.snowflake.com/east-us-2.azure/nua03035/worLh40kjUD#query. The sidebar shows various database and schema options. The main area displays a query in the SQL editor:

```
31
32  SELECT * FROM REGION_DIM;
33  | SELECT * FROM FACT_MOVIE;
34
35
```

The results pane shows a table with the following data:

#	MOVIE_S	TITLE_SK	TCONST	PARENT_TCONST	SEASON_NUMBI	EPISODE_NUMBI	#
1	0	13302861	tt0031458	tt32857063	0	0	
2	0	13302861	tt0041951	tt0041038	1	9	
3	0	13302861	tt0042816	tt0989125	1	17	
4	0	13302861	tt0042889	tt0989125	0	0	
5	0	13302861	tt0043426	tt0040051	3	42	
6	0	13302861	tt0043631	tt0989125	2	16	

On the right, the Query Details panel shows:

- Query duration: 1.1s
- Rows: 574.3K
- Query ID: 01baf2e7-0001-a61b-0...

Buttons for "Ask Copilot" and "Show more" are also present.

2] Fact_Ratings

The screenshot shows the Snowflake web interface. The URL is https://app.snowflake.com/east-us-2.azure/aza59135/w2AKgaEUw5k3#query. The sidebar shows various database and schema options. The main area displays a query in the SQL editor:

```
40
41  SELECT * FROM DATE_DIM;
42  SELECT * FROM LANGUAGE_DIM;
43  SELECT * FROM BRIDGE_MOVIE;
44  | SELECT * FROM DIM_TITLE;
45  | SELECT * FROM FACT_RATINGS;
```

The results pane shows a table with the following data:

#	FACT_RATINGS_S	TITLE_SK	AVERAGE_RATING	NUM_VOTE	DI_JOB_ID	DI_LOAD_D
1	1	13302845	8	9	e2af572e-1442-49e9-9c12-581be3b1b291	2025-03-12
2	2	13302846	8	9	5a11d6e3-8ad8-42eb-bf23-29e9406e3bc2	2025-03-12
3	3	13302847	8	9	59ac98c6-5f08-426a-b508-368acf9124d	2025-03-12
4	4	13302848	8	9	9e2f86a-1f12-49d1-b6df-4db9e9e23681	2025-03-12
5	5	13302849	8	9	cab5ed53-56a2-4e1c-9ba7-fe9b9586706e	2025-03-12
6	6	13302850	8	9	1515bc1f-ab3b-4b41-b535-4f97dfa4d9ae	2025-03-12
7	7	13302851	8	9	f02ab2d6-4305-4fc4-b8eb-ca4f649c163f	2025-03-12
8	8	13302852	8	9	7edc77ef-8a97-4f2e-abb1-47edf5f99f4d	2025-03-12

On the right, the Query Details panel shows:

- Query duration: 707ms
- Rows: 2.5K
- Query ID: 01baf267-0001-a4b0-0...

Buttons for "Ask Copilot" and "Show more" are also present.

BRIDGE TABLES :

1] Bridge Movie

Dim_Title and Dim_Genre

2]Bridge Title Personnel

Dim_Title and Dim_Personnel

The screenshot shows a data exploration interface with a query editor at the top containing the following SQL code:

```
> 61
62 | select * from BRIDGE_TITLE_PERSONNEL;
63
64
```

The main area displays a table titled "Results" with the following columns and data:

	# TITLE_SK	# PERSONNEL_SK	CATEGORY	JOB	CHARACTER	DI_JOB_ID	DI_LOAD_DT
1	12783863	5446486	director	director	NA	21063d81-07cf-41ab-9e00-46c89b768ae1	2025-03-11
2	12783863	5446487	director	director	NA	46bfa9b0-e161-4a8a-84f4-aaca9baa86a9	2025-03-11
3	12783863	5446488	director	director	NA	ef497e44-9494-44b7-becb-f3e3709af71	2025-03-11
4	12783863	5446489	director	director	NA	99b8586e-d2a3-455c-beb0-5e84f4cb83d5	2025-03-11
5	12783863	5446490	director	director	NA	7ca4d8d3-280d-41ed-9d8d-e6d4ae4c5a45	2025-03-11
6	12783864	5446479	director	director	NA	2e2bb51c-f19a-435e-b441-67a1f01e065d	2025-03-11
7	12783864	5446480	director	director	NA	d90645b3-0655-47e1-9292-2b60165abd7	2025-03-11
8	12783864	5446481	director	director	NA	7d69444b-e61d-4151-9b7a-70ac7a69db0e	2025-03-11
9	12783864	5446482	director	director	NA	d6c667db-80f3-4429-ba39-687860fa566e	2025-03-11
10	12783864	5446483	director	director	NA	4e2a434e-c968-4d0a-8b06-4ab3377f5936	2025-03-11

On the right side of the interface, there are "Query Details" showing a duration of 189ms and 1.38 rows. A "Show more" button is available. Below the details, there is a "Ask Copilot" button and a note about results being limited to under 1M. The bottom of the screen shows a Windows taskbar with various icons.

Loading Fact_Movie into Power BI

Navigator

Display Options ▾

- Get data ▾
- IMDB_DB [3]
- IMDB_SCHEMA [19]
 - DATE_DIM
 - DIM_TITLE
 - FACT_MOVIE**
 - GENRE_DIM
 - LANGUAGE_DIM
 - PERSONNEL_DIM
 - REGION_DIM
 - T_8241_249C7B01EED54B8692EF3E...
 - T_8405_73E56619155E4BE89CFFED...
 - T_8422_AE076D41C4D14F8391C95...
 - T_8830_150CA712A647446C88BE59...
 - TITLE_AKAS
 - TITLE_AKAS_FLATTEN
 - TITLE_BASICS
 - TITLE_BASICS_FLATTEN
 - TITLE_CREW

Select Related Tables

Load Transform Data Cancel

FACT_MOVIE

PARENT_TCONST	SEASON_NUMBER	EPISODE_NUMBER	RUNTIME_MINUTES	DLJ
tt0793361	2	3	63	8
tt0793361	2	3	63	8
tt0793361	2	3	63	8
tt0797756	1	6	84	8
tt0797756	1	6	84	8
tt0212699	8	4	30	8
tt0212699	8	4	30	8
tt0069572	1	0	90	8
tt0069572	1	0	90	8
tt0806910	1	158	90	8
tt0806910	1	158	90	8
tt0806910	1	158	90	8
tt0149504	1	0	92	8
tt0167732	1	30	80	8
tt0167732	1	30	80	8
tt0252792	1	2	58	8
tt9141320	1	6	27	8
tt9141320	1	6	27	8
tt9141320	1	6	27	8
tt0115122	3	1	90	8
tt0115122	3	1	90	8
tt0270780	2	5	55	8

Visualizations >< Data

Copilot

Add data fields here

Drill through

Cross-report

Keep all filters

Update available (click to download)

ENG IN 10:25 PM 3/11/2025

GIT SCREENSHOTS

COMMITS:

github.com/Shriya-Pekamwar/DaDaBi_MidTerm_Project

Shriya-Pekamwar / DaDaBi_MidTerm_Project

Type ⌘ to search

Code Issues Pull requests Actions Projects Wiki Security Insights

DaDaBi_MidTerm_Project Public

main 5 Branches 0 Tags Go to file Add file Code

Shriya-Pekamwar Updating pipeline: pipeline5 895c373 · 14 minutes ago 198 Commits

dataflow Updating dataflow: FactRatings 50 minutes ago

dataset Updating SnowflakeV2Table: SnowflakeTable2 18 minutes ago

factory Merge branch 'main' into Shriya yesterday

linkedService Updating linkedService: MidtermSnowflake 19 hours ago

pipeline Updating pipeline: pipeline5 14 minutes ago

README.md Create README.md 3 days ago

publish_config.json Update publish_config.json 3 days ago

About

No description, website, or topics provided.

Readme Activity 0 stars 1 watching 0 forks Report repository

Releases

No releases published Create a new release

Packages

README

11:36 PM 3/11/2025

github.com/Shriya-Pekamwar/DaDaBi_MidTerm_Project/commits/main/

Commits

main All users All time

Commits on Mar 11, 2025

Updating pipeline: pipeline5
Shriya-Pekamwar committed 15 minutes ago 895c373

Updating pipeline: Language
Shriya-Pekamwar committed 15 minutes ago da7b056

Updating SnowflakeV2Table: SnowflakeTable2
Shriya-Pekamwar committed 19 minutes ago 20d4a42

Updating pipeline: DimTitle
Shriya-Pekamwar committed 22 minutes ago fe0c78c

Updating SnowflakeV2Table: DIM_TITLE_SF
Shriya-Pekamwar committed 30 minutes ago 4e5e5c3

Updating dataflow: LanguageDimDF
Shriya-Pekamwar committed 33 minutes ago a5b6e1d

Updating dataflow: FactRatings
DhanvardiniRajendran committed 51 minutes ago f83eccf

Updating dataflow: LoadFactMovie

11:37 PM 3/11/2025

Respective Git Contribution

This screenshot shows the GitHub repository page for `Shriya-Pekamwar/DaDaBi_MidTerm_Project`. The main contributions listed are:

- Updating pipeline: pipeline5 (895c373 · 16 minutes ago)
- Updating dataflow: FactRatings (52 minutes ago)
- Updating SnowflakeV2Table: SnowflakeTable2 (20 minutes ago)
- Merge branch 'main' into Shriya (yesterday)
- Updating linkedService: MidtermSnowflake (19 hours ago)
- Updating pipeline: pipeline5 (16 minutes ago)
- Create Readme.md (3 days ago)
- Update publish_config.json (3 days ago)

The repository has 198 commits in total. The sidebar includes sections for About (no description), Activity (0 stars, 1 watching, 0 forks), Releases (no releases published), and Packages (no packages published). The status bar at the bottom shows the date as 3/11/2025 and the time as 11:38 PM.

This screenshot shows the GitHub repository page for `Shriya-Pekamwar/DaDaBi_MidTerm_Project`. The main contributions listed are:

- Adding dataflow: dataflow2 (c106b68 · yesterday)
- Adding dataflow: dataflow2 (yesterday)
- Adding dataflow: dataflow2 (yesterday)
- Adding factory: Midterm-Project-Team2 (3 days ago)
- Adding linkedService: IMDBSnowflake (2 days ago)
- Adding dataflow: date_dimDL (yesterday)
- Create Readme.md (3 days ago)
- Update publish_config.json (3 days ago)

This branch is 167 commits behind `main`. The repository has 31 commits in total. The sidebar includes sections for About (no description), Activity (0 stars, 1 watching, 0 forks), Releases (no releases published), and Packages (no packages published). The status bar at the bottom shows the date as 3/11/2025 and the time as 11:38 PM.

About

No description, website, or topics provided.

Readme
Activity
0 stars
1 watching
0 forks

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)

Commits

Author	Commit Message	Time Ago
Komal1799	Updating SnowflakeV2Table: Title_Akas_SF	b68152e - yesterday
dataset	Adding SnowflakeV2Table: Title_Akas_SF	2 days ago
factory	Adding factory: Midterm-Project-Team2	3 days ago
linkedService	Adding linkedService: SnowflakeMidterm	2 days ago
pipeline	Adding pipeline: pipeline1	2 days ago
Readme.md	Create Readme.md	3 days ago
publish_config.json	Create publish_config.json	3 days ago

About

No description, website, or topics provided.

Readme
Activity
0 stars
1 watching
0 forks

Releases

No releases published
[Create a new release](#)

Packages

No packages published
[Publish your first package](#)

Commits

Author	Commit Message	Time Ago
Shriya-Pekamwar	Adding pipeline: pipeline1	70fa41c - 20 hours ago
dataflow	Merge branch 'main' into Shriya	20 hours ago
dataset	Merge branch 'main' into Shriya	20 hours ago
factory	Merge branch 'main' into Shriya	20 hours ago
linkedService	Merge branch 'main' into Shriya	20 hours ago
pipeline	Adding pipeline: pipeline1	20 hours ago
Readme.md	Create Readme.md	3 days ago
publish_config.json	Update publish_config.json	3 days ago