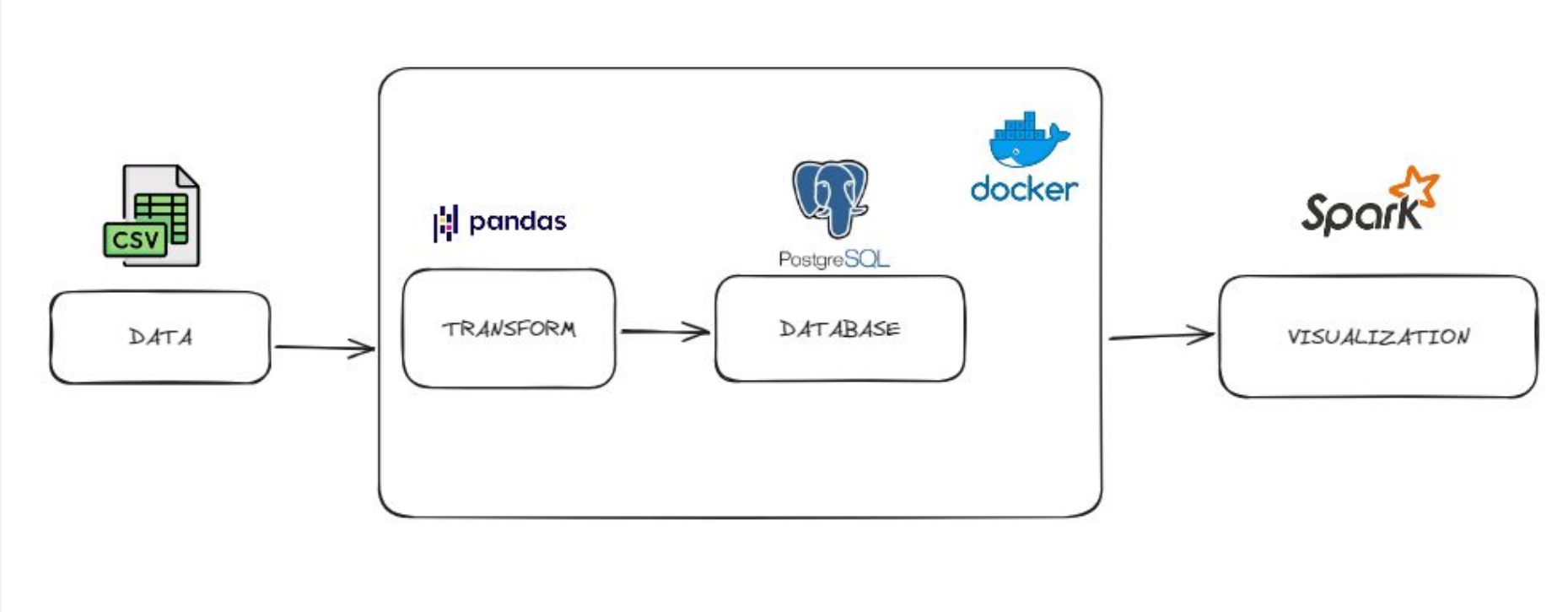
MOVIES ANALYSIS

Satriyo Wisnu Dwi Putro

Problem

Film identik dengan rating atau penilaian dari penonton, baik atau buruknya sebuah film dapat dilihat dengan rating, dengan menganalisa film dengan rating tertinggi diberbagai tahun, kita dapat menyimpulkan lebih banyak film lama atau terbaru yang memiliki rating tinggi.

Pipeline


CSV
DATA
pandas
TRANSFORM
PostgreSQL
DATABASE
docker
Spark
VISUALIZATION

Extract

	Movie	Released_year	Rating	Director
0	The Shawshank Redemption	1994	9.3	Frank Darabont
1	The Godfather	1972	9.2	Francis Ford Coppola
2	The Godfather: Part II	1974	9.0	Francis Ford Coppola
3	The Dark Knight	2008	9.0	Christopher Nolan
4	12 Angry Men	1957	8.9	Sidney Lumet

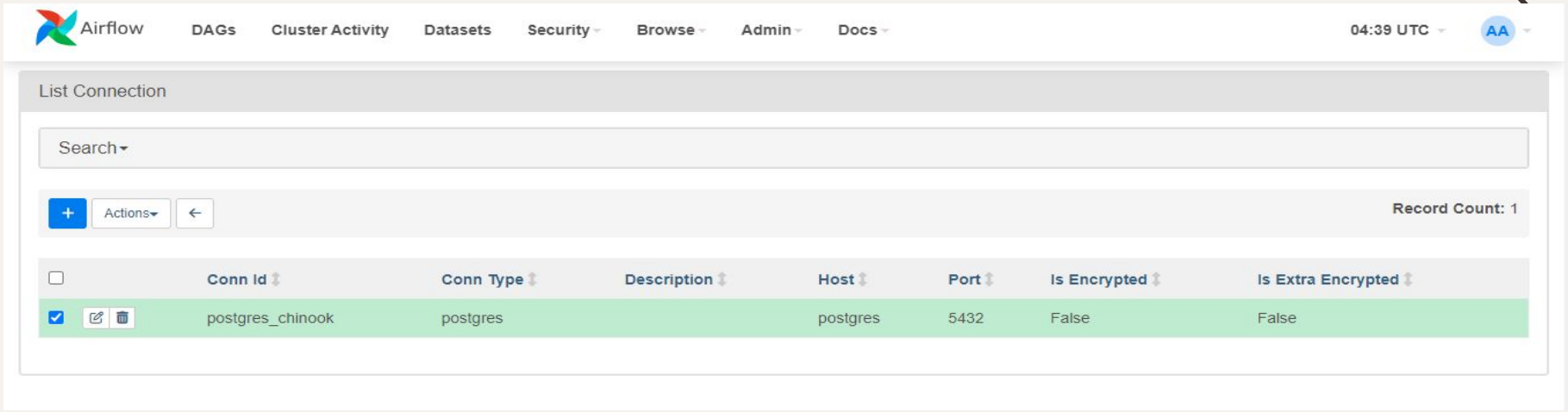
Movie: Nama Film
Released_year: Tahun Rilis Film
Rating: Nilai Film
Director: Sutradara

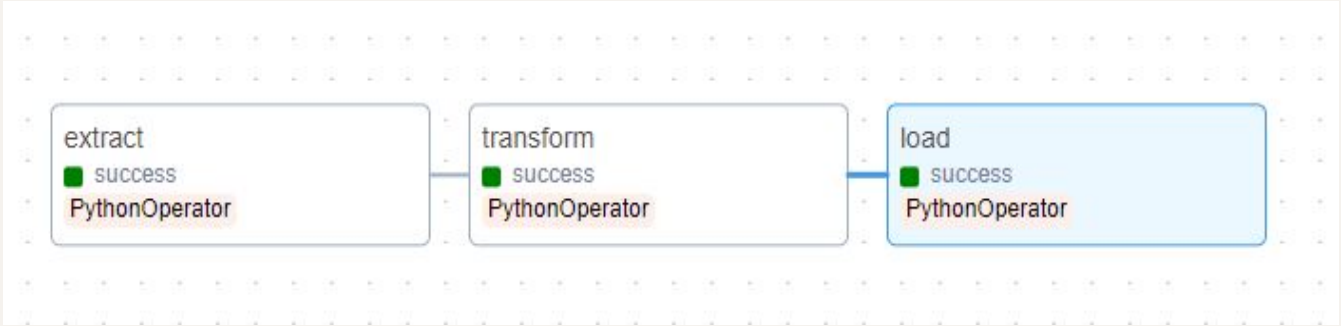
Transform

```python
def transform(**kwargs):
    ti = kwargs['ti']
    data_str = ti.xcom_pull(task_ids='extract')
    df = pd.read_json(data_str)

    def classify_year(released_year):
        if 1900 <= released_year <= 1999:
            return 'Film Lama'
        elif 2000 <= released_year <= 2012:
            return 'Film Menengah'
        elif 2013 <= released_year <= 2024:
            return 'Film Baru'

    df['group_year'] = df['Released_year'].apply(classify_year)
    return df.to_json()
```

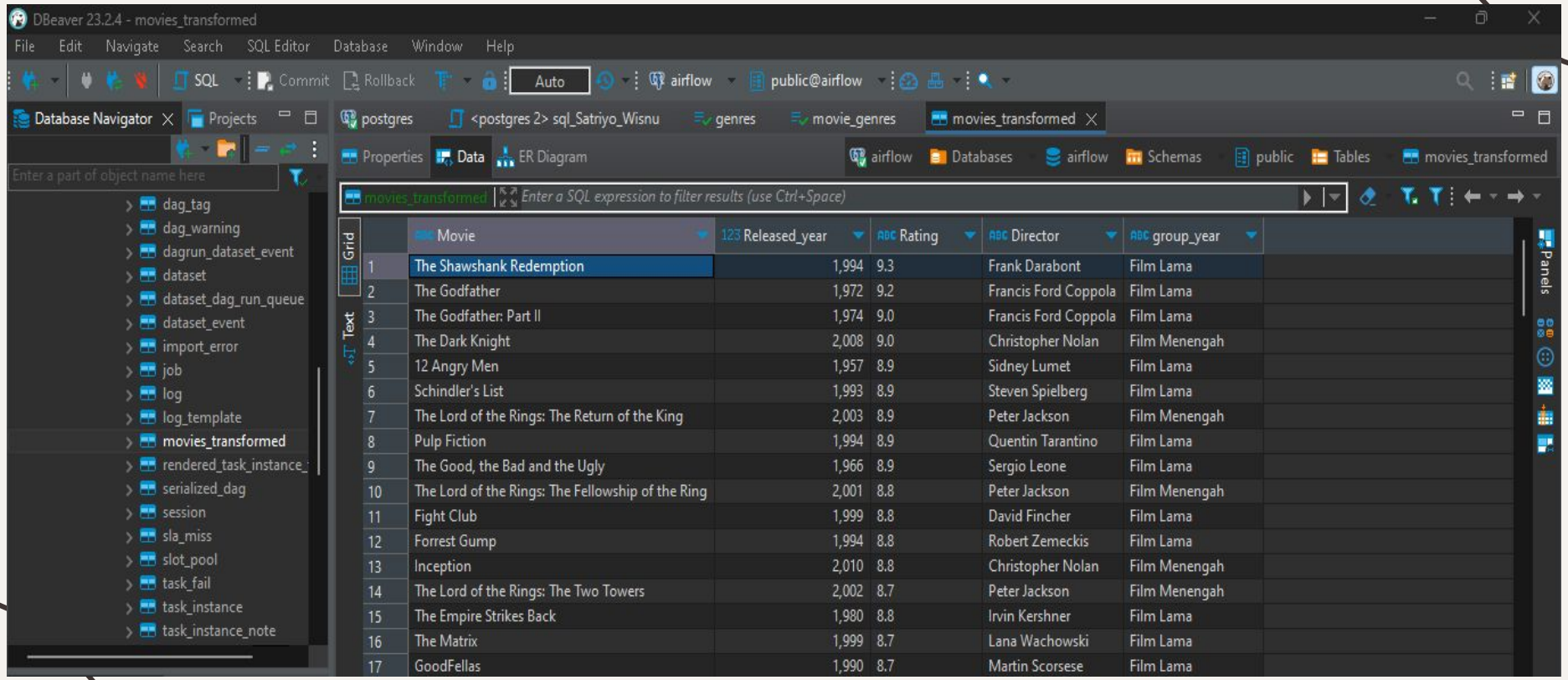
Memisahkan data menjadi ‘Film Lama’, ‘Film Menengah’, ‘Film Baru’ dan membuatnya dalam bentuk kolom baru bernama group_year

Load


Airflow
DAGs
Cluster Activity
Datasets
Security
Browse
Admin
Docs
04:39 UTC
AA
List Connection
Search
Actions
Record Count: 1
Conn Id
Conn Type
Description
Host
Port
Is Encrypted
Is Extra Encrypted
postgres_chinook
postgres
postgres
5432
False
False

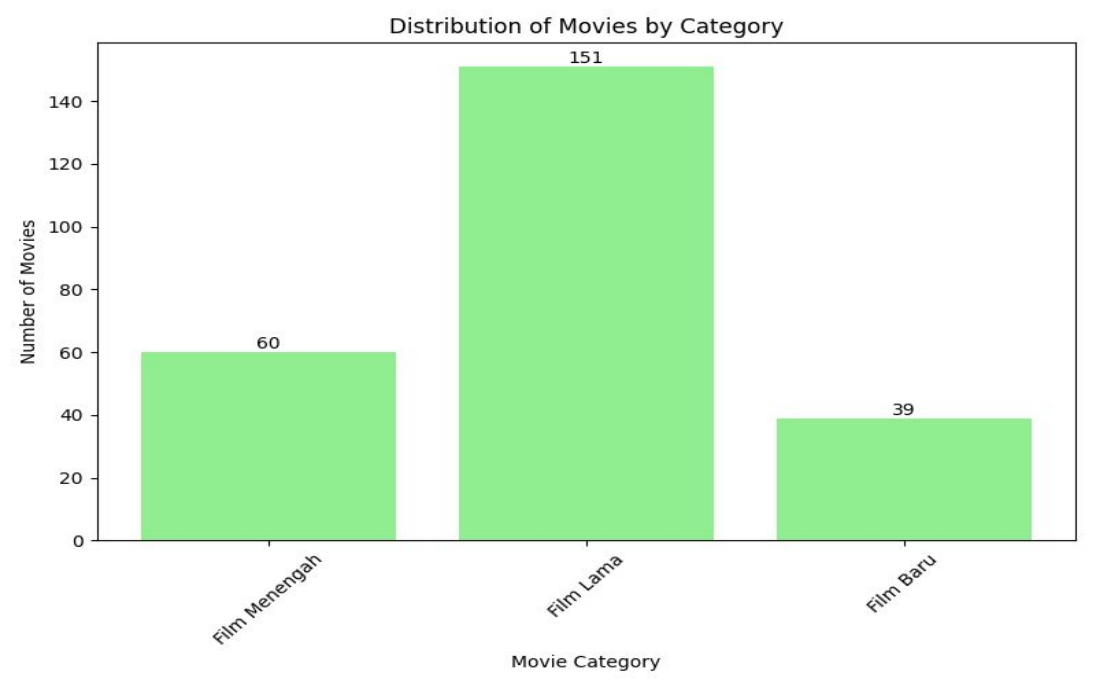

extract
success
PythonOperator
transform
success
PythonOperator
load
success
PythonOperator

Load



	Movie	Released_year	Rating	Director	group_year
1	The Shawshank Redemption	1,994	9.3	Frank Darabont	Film Lama
2	The Godfather	1,972	9.2	Francis Ford Coppola	Film Lama
3	The Godfather: Part II	1,974	9.0	Francis Ford Coppola	Film Lama
4	The Dark Knight	2,008	9.0	Christopher Nolan	Film Menengah
5	12 Angry Men	1,957	8.9	Sidney Lumet	Film Lama
6	Schindler's List	1,993	8.9	Steven Spielberg	Film Lama
7	The Lord of the Rings: The Return of the King	2,003	8.9	Peter Jackson	Film Menengah
8	Pulp Fiction	1,994	8.9	Quentin Tarantino	Film Lama
9	The Good, the Bad and the Ugly	1,966	8.9	Sergio Leone	Film Lama
10	The Lord of the Rings: The Fellowship of the Ring	2,001	8.8	Peter Jackson	Film Menengah
11	Fight Club	1,999	8.8	David Fincher	Film Lama
12	Forrest Gump	1,994	8.8	Robert Zemeckis	Film Lama
13	Inception	2,010	8.8	Christopher Nolan	Film Menengah
14	The Lord of the Rings: The Two Towers	2,002	8.7	Peter Jackson	Film Menengah
15	The Empire Strikes Back	1,980	8.8	Irvin Kershner	Film Lama
16	The Matrix	1,999	8.7	Lana Wachowski	Film Lama
17	GoodFellas	1,990	8.7	Martin Scorsese	Film Lama

Visualization



Dari hasil analisa menunjukkan Film Lama (90’s) paling banyak mendapatkan rating tinggi dibandingkan film-film terbaru

THANKYOU