```python
In [3]: import pandas as pd
```

```python
In [6]: movies = pd.read_csv(r"C:\Users\Rachitha\Desktop\Movie-Recommender-System\Dataset\movies.csv")
        ratings = pd.read_csv(r"C:\Users\Rachitha\Desktop\Movie-Recommender-System\Dataset\ratings.csv")
        movies.head()
```

Out[6]:

	movieId	title	genres
0	1	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
1	2	Jumanji (1995)	Adventure|Children|Fantasy
2	3	Grumpier Old Men (1995)	Comedy|Romance
3	4	Waiting to Exhale (1995)	Comedy|Drama|Romance
4	5	Father of the Bride Part II (1995)	Comedy

```python
In [7]: ratings.head()
```

Out[7]:

	userId	movieId	rating	timestamp
0	1	1	4.0	964982703
1	1	3	4.0	964981247
2	1	6	4.0	964982224
3	1	47	5.0	964983815
4	1	50	5.0	964982931

```python
In [8]: print("Movies Missing Values : \n " ,movies.isnull().sum())
        print("\nRatings Missing Values : \n" ,ratings.isnull().sum())
```

```
Movies Missing Values :
  movieId    0
title      0
genres     0
dtype: int64

Ratings Missing Values :
  userId     0
movieId    0
rating     0
timestamp  0
dtype: int64
```

```python
In [9]: movies['genres'].head()
```

```
Out[9]: 0    Adventure|Animation|Children|Comedy|Fantasy
        1                  Adventure|Children|Fantasy
        2                             Comedy|Romance
        3                       Comedy|Drama|Romance
        4                                     Comedy
        Name: genres, dtype: object
```

```python
In [10]: genre_list = movies['genres'].str.split('|')
         unique_genres = set(g for sublist in genre_list.dropna() for g in sublist)
         print("Unique Genres :",unique_genres)
```

```
Unique Genres : {'Comedy', 'IMAX', 'Romance', 'Sci-Fi', 'Thriller', 'Action', 'Film-Noir', 'Musical', 'Crime',
'Children', 'Adventure', 'Western', 'Drama', 'War', 'Mystery', '(no genres listed)', 'Documentary', 'Horror',
'Fantasy', 'Animation'}
```

```python
In [11]: data = ratings.merge(movies,on='movieId')
         data.head()
```

Out[11]:

	userId	movieId	rating	timestamp	title	genres
0	1	1	4.0	964982703	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
1	1	3	4.0	964981247	Grumpier Old Men (1995)	Comedy|Romance
2	1	6	4.0	964982224	Heat (1995)	Action|Crime|Thriller
3	1	47	5.0	964983815	Seven (a.k.a. Se7en) (1995)	Mystery|Thriller
4	1	50	5.0	964982931	Usual Suspects, The (1995)	Crime|Mystery|Thriller

```python
In [12]: rating_counts = data['title'].value_counts()
         popular_movies = rating_counts[rating_counts >= 10].index
         filtered_data = data[data['title'].isin(popular_movies)]
         filtered_data.head()
```

Out[12]:

	userId	movieId	rating	timestamp	title	genres
0	1	1	4.0	964982703	Toy Story (1995)	Adventure|Animation|Children|Comedy|Fantasy
1	1	3	4.0	964981247	Grumpier Old Men (1995)	Comedy|Romance
2	1	6	4.0	964982224	Heat (1995)	Action|Crime|Thriller
3	1	47	5.0	964983815	Seven (a.k.a. Se7en) (1995)	Mystery|Thriller
4	1	50	5.0	964982931	Usual Suspects, The (1995)	Crime|Mystery|Thriller