

```
In [1]: import numpy as np
import pandas as pd
#Importing pandas library as pd
import pandas as pd
#reading the excel sheet to python
dataset = pd.read_excel("Movies_Database.xlsx")
#checking the content in dataset
dataset
#showing first 5 rows of the dataset
dataset.head()
#check the tail:
dataset.tail(5)
#showing the column header
dataset.columns
dataset.index
dataset.values
type(dataset)
dataset.shape

Out[1]: (14, 16)

In [2]: dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14 entries, 0 to 13
Data columns (total 16 columns):
#   Column              Non-Null Count  Dtype
---  --
0   movie_title          14 non-null    object
1   num_critc_for_reviews 14 non-null    int64
2   duration             11 non-null    float64
3   DIRECTOR_facebook_likes 12 non-null    object
4   actor_3_facebook_likes 14 non-null    int64
5   ACTOR_1_facebook_likes 14 non-null    int64
6   gross               14 non-null    int64
7   num_voted_users      13 non-null    float64
8   Cast_Total_facebook_likes 12 non-null    float64
9   facenumber_in_poster 9 non-null     float64
10  num_user_for_reviews 14 non-null    int64
11  budget              14 non-null    int64
12  title_year          14 non-null    int64
13  ACTOR_2_facebook_likes 13 non-null    float64
14  imdb_score          14 non-null    float64
15  title_year.1        7 non-null     float64
dtypes: float64(7), int64(7), object(2)
memory usage: 1.9+ KB

In [3]: dataset['movie_title']

Out[3]:
0   Avatar
1   Pirates of the Caribbean: At World's End?
2   Spectre
3   The Dark Knight Rises?
4   John Carter?
5   Spider-Man 3?
6   Tangled?
7   Avengers: Age of Ultron?
8   Avengers: Age of Ultron?
9   Harry Potter and the Half-Blood Prince?
10  Batman v Superman: Dawn of Justice?
11  Superman Returns?
12  Quantum of Solace?
13  Pirates of the Caribbean: Dead Man's Chest?
Name: movie_title, dtype: object

In [4]: dataset['movie_title'] = dataset['movie_title'].str.replace(r'(?i)', '')
print(dataset)

0   movie_title  num_critc_for_reviews \
1   Pirates of the Caribbean: At World's End  392
2   Spectre  602
3   The Dark Knight Rises  813
4   John Carter  462
5   Spider-Man 3  392
6   Tangled  324
7   Avengers: Age of Ultron  635
8   Avengers: Age of Ultron  635
9   Harry Potter and the Half-Blood Prince  375
10  Batman v Superman: Dawn of Justice  673
11  Superman Returns  434
12  Quantum of Solace  403
13  Pirates of the Caribbean: Dead Man's Chest  313

duration DIRECTOR_facebook_likes actor_3_facebook_likes \
0   178.0 10 855
1   NaN 563 1000
2   148.0 20 161
3   NaN 22000 23000
4   132.0 475 530
5   156.0 23 4900
6   NaN 15 284
7   141.0 10 19900
8   141.0 10 19900
9   153.0 282 10000
10  183.0 NaN 2000
11  160.0 NaN 903
12  106.0 395 393
13  151.0 563 1000

ACTOR_1_facebook_likes gross num_voted_users \
0   1000 760505847 886204.0
1   40000 309404152 471220.0
2   11000 200074175 275868.0
3   27000 448130642 1144337.0
4   640 73058679 212204.0
5   24000 336530303 383056.0
6   799 200807262 294810.0
7   26000 458991599 462669.0
8   26000 458991599 462669.0
9   25000 301956980 321795.0
10  15000 330249062 NaN
11  18000 200069408 240396.0
12  451 168368427 330784.0
13  40000 423032628 522040.0

Cast_Total_facebook_likes facenumber_in_poster num_user_for_reviews \
0   4834.0 NaN 3054
1   48350.0 NaN 1238
2   11700.0 1.0 994
3   106759.0 NaN 2701
4   1873.0 1.0 738
5   46055.0 NaN 1992
6   NaN 1.0 387
7   92000.0 4.0 1117
8   92000.0 4.0 1117
9   58753.0 3.0 673
10  24450.0 NaN 3018
11  NaN 2.0 2367
12  2023.0 1.0 1243
13  48486.0 2.0 1832

budget title_year ACTOR_2_facebook_likes imdb_score title_year.1
0  237000000 2009 936.0 7.9 2009.0
1  300000000 2007 5000.0 7.1 NaN
2  245000000 2015 393.0 6.8 2015.0
3  250000000 2012 23000.0 8.5 NaN
4  263700000 2012 632.0 6.6 NaN
5  258000000 2007 11000.0 6.2 2007.0
6  260000000 2010 553.0 7.8 NaN
7  250000000 2015 21000.0 7.5 NaN
8  250000000 2015 21000.0 7.5 2015.0
9  250000000 2009 11000.0 7.9 NaN
10 260000000 2016 2016.0 6.8 2016.0
11 209000000 2006 10000.0 6.1 NaN
12 209000000 2008 412.0 6.7 2008.0
13 225000000 2006 5000.0 7.3 2006.0

C:\Users\ssaji\AppData\Local\Temp\ipykernel_16548\3964326471.py:1: FutureWarning: The default value of regex will change from True to False in a future version.
dataset['movie_title'] = dataset['movie_title'].str.replace(r'(?i)', '')

In [5]: dataset

Out[5]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 NaN
1 Pirates of the Caribbean: At World's End 302 NaN 563 1000 40000 309404152 471220.0 48350.0 NaN
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 NaN 22000 23000 27000 448130642 1144337.0 106759.0 NaN
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 46055.0 NaN
6 Tangled 324 NaN 15 284 799 200807262 294810.0 NaN 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 NaN 2000 15000 330249062 NaN 24450.0 NaN
11 Superman Returns 434 169.0 NaN 903 18000 200069408 240396.0 NaN 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 313 151.0 563 1000 40000 423032628 522040.0 48486.0 2.0

In [6]: dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14 entries, 0 to 13
Data columns (total 16 columns):
#   Column              Non-Null Count  Dtype
---  --
0   movie_title          14 non-null    object
1   num_critc_for_reviews 14 non-null    int64
2   duration             11 non-null    float64
3   DIRECTOR_facebook_likes 12 non-null    object
4   actor_3_facebook_likes 14 non-null    int64
5   ACTOR_1_facebook_likes 14 non-null    int64
6   gross               14 non-null    int64
7   num_voted_users      13 non-null    float64
8   Cast_Total_facebook_likes 12 non-null    float64
9   facenumber_in_poster 9 non-null     float64
10  num_user_for_reviews 14 non-null    int64
11  budget              14 non-null    int64
12  title_year          14 non-null    int64
13  ACTOR_2_facebook_likes 13 non-null    float64
14  imdb_score          14 non-null    float64
15  title_year.1        7 non-null     float64
dtypes: float64(7), int64(7), object(2)
memory usage: 1.9+ KB

In [7]: dataset.shape

Out[7]: (14, 16)

In [8]: pd.isnull(np.nan)

Out[8]: True

In [9]: pd.isnull(dataset)

Out[9]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 False False False False False False False False False False True
1 False False True False False False False False False False True
2 False False False False False False False False False False False
3 False False True False False False False False False False True
4 False False False False False False False False False False False
5 False False False False False False False False False False True
6 False False True False False False False False False False True
7 False False False False False False False False False False False
8 False False False False False False False False False False False
9 False False False False False False False False False False False
10 False False False True True False False True True True
11 False False False True False False False True True True
12 False False False False False False False False False False False
13 False False False False False False False False False False False

In [10]: dataset.isnull().sum()

Out[10]:
movie_title 0
num_critc_for_reviews 0
duration 3
DIRECTOR_facebook_likes 2
actor_3_facebook_likes 0
ACTOR_1_facebook_likes 0
gross 0
num_voted_users 1
Cast_Total_facebook_likes 2
facenumber_in_poster 5
num_user_for_reviews 0
budget 0
title_year 0
ACTOR_2_facebook_likes 1
imdb_score 0
title_year.1 7
dtype: int64

In [11]: #filling notnull values with 0
dataset.fillna(0)

Out[11]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 0.0
1 Pirates of the Caribbean: At World's End 302 0.0 563 1000 40000 309404152 471220.0 48350.0 0.0
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 0.0 22000 23000 27000 448130642 1144337.0 106759.0 0.0
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 0.0 0.0
6 Tangled 324 0.0 15 284 799 200807262 294810.0 0.0 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 0 2000 15000 330249062 0.0 24450.0 0.0
11 Superman Returns 434 169.0 0 903 18000 200069408 240396.0 0.0 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 313 151.0 563 1000 40000 423032628 522040.0 48486.0 2.0

In [12]: dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14 entries, 0 to 13
Data columns (total 16 columns):
#   Column              Non-Null Count  Dtype
---  --
0   movie_title          14 non-null    object
1   num_critc_for_reviews 14 non-null    int64
2   duration             11 non-null    float64
3   DIRECTOR_facebook_likes 12 non-null    object
4   actor_3_facebook_likes 14 non-null    int64
5   ACTOR_1_facebook_likes 14 non-null    int64
6   gross               14 non-null    int64
7   num_voted_users      13 non-null    float64
8   Cast_Total_facebook_likes 12 non-null    float64
9   facenumber_in_poster 9 non-null     float64
10  num_user_for_reviews 14 non-null    int64
11  budget              14 non-null    int64
12  title_year          14 non-null    int64
13  ACTOR_2_facebook_likes 13 non-null    float64
14  imdb_score          14 non-null    float64
15  title_year.1        7 non-null     float64
dtypes: float64(7), int64(7), object(2)
memory usage: 1.9+ KB

In [13]: dataset['DIRECTOR_facebook_likes'][4] =int(dataset['DIRECTOR_facebook_likes'][4].replace('',''))

C:\Users\ssaji\AppData\Local\Temp\ipykernel_16548\3128792715.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
dataset['DIRECTOR_facebook_likes'][4] =int(dataset['DIRECTOR_facebook_likes'][4].replace('',''))

In [14]: type(dataset['DIRECTOR_facebook_likes'][1])

Out[14]: int

In [15]: dataset.fillna(0)

Out[15]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 0.0
1 Pirates of the Caribbean: At World's End 302 0.0 563 1000 40000 309404152 471220.0 48350.0 0.0
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 0.0 22000 23000 27000 448130642 1144337.0 106759.0 0.0
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 46055.0 0.0
6 Tangled 324 0.0 15 284 799 200807262 294810.0 0.0 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 0 2000 15000 330249062 0.0 24450.0 0.0
11 Superman Returns 434 169.0 0 903 18000 200069408 240396.0 0.0 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 313 151.0 563 1000 40000 423032628 522040.0 48486.0 2.0

In [16]: dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14 entries, 0 to 13
Data columns (total 16 columns):
#   Column              Non-Null Count  Dtype
---  --
0   movie_title          14 non-null    object
1   num_critc_for_reviews 14 non-null    int64
2   duration             11 non-null    float64
3   DIRECTOR_facebook_likes 14 non-null    int32
4   actor_3_facebook_likes 14 non-null    int64
5   ACTOR_1_facebook_likes 14 non-null    int64
6   gross               14 non-null    int64
7   num_voted_users      13 non-null    float64
8   Cast_Total_facebook_likes 14 non-null    float64
9   facenumber_in_poster 14 non-null    int64
10  num_user_for_reviews 14 non-null    int64
11  budget              14 non-null    int64
12  title_year          14 non-null    int64
13  ACTOR_2_facebook_likes 14 non-null    float64
14  imdb_score          14 non-null    float64
15  title_year.1        7 non-null     float64
dtypes: float64(7), int32(1), int64(7), object(1)
memory usage: 1.8+ KB

In [17]: dataset['DIRECTOR_facebook_likes'] = dataset['DIRECTOR_facebook_likes'].fillna(0).astype(int)

In [18]: dataset.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14 entries, 0 to 13
Data columns (total 16 columns):
#   Column              Non-Null Count  Dtype
---  --
0   movie_title          14 non-null    object
1   num_critc_for_reviews 14 non-null    int64
2   duration             11 non-null    float64
3   DIRECTOR_facebook_likes 14 non-null    int32
4   actor_3_facebook_likes 14 non-null    int64
5   ACTOR_1_facebook_likes 14 non-null    int64
6   gross               14 non-null    int64
7   num_voted_users      13 non-null    float64
8   Cast_Total_facebook_likes 14 non-null    float64
9   facenumber_in_poster 14 non-null    int64
10  num_user_for_reviews 14 non-null    int64
11  budget              14 non-null    int64
12  title_year          14 non-null    int64
13  ACTOR_2_facebook_likes 14 non-null    float64
14  imdb_score          14 non-null    float64
15  title_year.1        7 non-null     float64
dtypes: float64(7), int32(1), int64(7), object(1)
memory usage: 1.8+ KB

In [19]: dataset

Out[19]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 0.0
1 Pirates of the Caribbean: At World's End 302 0.0 563 1000 40000 309404152 471220.0 48350.0 0.0
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 0.0 22000 23000 27000 448130642 1144337.0 106759.0 0.0
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 46055.0 0.0
6 Tangled 324 0.0 15 284 799 200807262 294810.0 0.0 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 0 2000 15000 330249062 0.0 24450.0 0.0
11 Superman Returns 434 169.0 0 903 18000 200069408 240396.0 0.0 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 313 151.0 563 1000 40000 423032628 522040.0 48486.0 2.0

In [20]: dataset

In [21]: dataset

Out[21]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 0.0
1 Pirates of the Caribbean: At World's End 302 0.0 563 1000 40000 309404152 471220.0 48350.0 0.0
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 0.0 22000 23000 27000 448130642 1144337.0 106759.0 0.0
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 46055.0 0.0
6 Tangled 324 0.0 15 284 799 200807262 294810.0 0.0 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 0 2000 15000 330249062 0.0 24450.0 0.0
11 Superman Returns 434 169.0 0 903 18000 200069408 240396.0 0.0 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 313 151.0 563 1000 40000 423032628 522040.0 48486.0 2.0

In [22]: dataset

In [23]: dataset

Out[23]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 0.0
1 Pirates of the Caribbean: At World's End 302 0.0 563 1000 40000 309404152 471220.0 48350.0 0.0
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 0.0 22000 23000 27000 448130642 1144337.0 106759.0 0.0
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 46055.0 0.0
6 Tangled 324 0.0 15 284 799 200807262 294810.0 0.0 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 0 2000 15000 330249062 0.0 24450.0 0.0
11 Superman Returns 434 169.0 0 903 18000 200069408 240396.0 0.0 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 313 151.0 563 1000 40000 423032628 522040.0 48486.0 2.0

In [24]: dataset

In [25]: dataset

Out[25]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 0.0
1 Pirates of the Caribbean: At World's End 302 0.0 563 1000 40000 309404152 471220.0 48350.0 0.0
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 0.0 22000 23000 27000 448130642 1144337.0 106759.0 0.0
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 46055.0 0.0
6 Tangled 324 0.0 15 284 799 200807262 294810.0 0.0 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 0 2000 15000 330249062 0.0 24450.0 0.0
11 Superman Returns 434 169.0 0 903 18000 200069408 240396.0 0.0 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 313 151.0 563 1000 40000 423032628 522040.0 48486.0 2.0

In [26]: dataset

In [27]: dataset

Out[27]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 0.0
1 Pirates of the Caribbean: At World's End 302 0.0 563 1000 40000 309404152 471220.0 48350.0 0.0
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 0.0 22000 23000 27000 448130642 1144337.0 106759.0 0.0
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 46055.0 0.0
6 Tangled 324 0.0 15 284 799 200807262 294810.0 0.0 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 0 2000 15000 330249062 0.0 24450.0 0.0
11 Superman Returns 434 169.0 0 903 18000 200069408 240396.0 0.0 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 313 151.0 563 1000 40000 423032628 522040.0 48486.0 2.0

In [28]: dataset

In [29]: dataset

Out[29]:
movie_title num_critc_for_reviews duration DIRECTOR_facebook_likes actor_3_facebook_likes ACTOR_1_facebook_likes gross num_voted_users Cast_Total_facebook_likes facenumber_in_poster num...
0 Avatar 723 178.0 10 855 1000 760505847 886204.0 4834.0 0.0
1 Pirates of the Caribbean: At World's End 302 0.0 563 1000 40000 309404152 471220.0 48350.0 0.0
2 Spectre 602 148.0 20 161 11000 200074175 275868.0 11700.0 1.0
3 The Dark Knight Rises 813 0.0 22000 23000 27000 448130642 1144337.0 106759.0 0.0
4 John Carter 462 132.0 475 530 640 73058679 212204.0 1873.0 1.0
5 Spider-Man 3 392 156.0 23 4000 24000 336530303 383056.0 46055.0 0.0
6 Tangled 324 0.0 15 284 799 200807262 294810.0 0.0 1.0
7 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
8 Avengers: Age of Ultron 635 141.0 10 19000 26000 458991599 462669.0 92000.0 4.0
9 Harry Potter and the Half-Blood Prince 375 153.0 282 10000 25000 301956980 321795.0 58753.0 3.0
10 Batman v Superman: Dawn of Justice 673 183.0 0 2000 15000 330249062 0.0 24450.0 0.0
11 Superman Returns 434 169.0 0 903 18000 200069408 240396.0 0.0 2.0
12 Quantum of Solace 403 106.0 395 393 451 168368427 330784.0 2023.0 1.0
13 Pirates of the Caribbean: Dead Man's Chest 31
```