

```

# Import Libraries
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from scipy.stats import mode
sns.set_style('whitegrid')

# Read the data
df = pd.read_csv(r"C:\Users\Lenovo\Desktop\IMDB_Movies.csv")

df.head()

```

	color	director_name	num_critic_for_reviews	duration	director_facebook_likes	actor_3_facebook_likes	actor_2_name	actor_1_facebook_likes	gross	genres	...	num_users
0	Color	James Cameron	723.0	178.0	0.0	855.0	Joel David Moore	1000.0	760505847.0	Action Adventure Fantasy Sci-Fi	...	
1	Color	Gore Verbinski	302.0	169.0	563.0	1000.0	Orlando Bloom	40000.0	309404152.0	Action Adventure Fantasy	...	
2	Color	Sam Mendes	602.0	148.0	0.0	161.0	Rory Kinnear	11000.0	200074175.0	Action Adventure Thriller	...	
3	Color	Christopher Nolan	813.0	164.0	22000.0	23000.0	Christian Bale	27000.0	448130642.0	Action Thriller	...	
4	NaN	Doug Walker	NaN	NaN	131.0	NaN	Rob Walker	131.0	NaN	Documentary	...	

5 rows x 28 columns

```

df.shape
(5043, 28)

df.columns
Index(['color', 'director_name', 'num_critic_for_reviews', 'duration',
      'director_facebook_likes', 'actor_3_facebook_likes', 'actor_2_name',
      'actor_1_facebook_likes', 'gross', 'genres', 'actor_1_name',
      'movie_title', 'num_voted_users', 'cast_total_facebook_likes',
      'actor_3_name', 'facenumber_in_poster', 'plot_keywords',
      'movie_imdb_link', 'num_user_for_reviews', 'language', 'country',
      'content_rating', 'budget', 'title_year', 'actor_2_facebook_likes',
      'imdb_score', 'aspect_ratio', 'movie_facebook_likes'],
      dtype='object')

# Extracting necessary columns for analysis
movies = df[['color', 'director_name', 'duration', 'actor_1_name', 'genres', 'movie_title', 'num_voted_users', 'language', 'country', 'budget', 'gross', 'title_year', 'imdb_score']]
movies.head()

```

	color	director_name	duration	actor_1_name	genres	movie_title	num_voted_users	language	country	budget	gross	title_year	imdb_score
0	Color	James Cameron	178.0	CCH Pounder	Action Adventure Fantasy Sci-Fi	Avatar	886204	English	USA	237000000.0	760505847.0	2009.0	7.9
1	Color	Gore Verbinski	169.0	Johnny Depp	Action Adventure Fantasy	Pirates of the Caribbean: At World's End	471220	English	USA	300000000.0	309404152.0	2007.0	7.1
2	Color	Sam Mendes	148.0	Christoph Waltz	Action Adventure Thriller	Spectre	275868	English	UK	245000000.0	200074175.0	2015.0	6.8
3	Color	Christopher Nolan	164.0	Tom Hardy	Action Thriller	The Dark Knight Rises	1144337	English	USA	250000000.0	448130642.0	2012.0	8.5
4	NaN	Doug Walker	NaN	Doug Walker	Documentary	Star Wars: Episode VII - The Force Awakens ...	8	NaN	NaN	NaN	NaN	NaN	7.1

```

movies.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 5043 entries, 0 to 5042
Data columns (total 13 columns):
 #   Column        Non-Null Count  Dtype
---  ---
 0   color         5024 non-null   object
 1   director_name 4939 non-null   object
 2   duration      5028 non-null   float64
 3   actor_1_name  5036 non-null   object
 4   genres        5043 non-null   object
 5   movie_title   5043 non-null   object
 6   num_voted_users 5043 non-null   int64
 7   language     5031 non-null   object
 8   country       5038 non-null   object
 9   budget       4551 non-null   float64
10  gross        4159 non-null   float64
11  title_year   4935 non-null   float64
12  imdb_score   5043 non-null   float64
dtypes: float64(5), int64(1), object(7)
memory usage: 512.3+ KB

#Calculating sum of null values in each column
movies.isnull().sum()
color              0
director_name      0
duration           0
actor_1_name       0
genres             0
movie_title        0
num_voted_users    0
language           0
country            0
budget             0
gross              0
title_year         0
imdb_score         0
dtype: int64

# Dropping null values from movies table
movies=movies.dropna()
movies

```

	color	director_name	duration	actor_1_name	genres	movie_title	num_voted_users	language	country	budget	gross	title_year	imdb_score
0	Color	James Cameron	178.0	CCH Pounder	Action Adventure Fantasy Sci-Fi	Avatar	886204	English	USA	237000000.0	760505847.0	2009.0	7.9
1	Color	Gore Verbinski	169.0	Johnny Depp	Action Adventure Fantasy	Pirates of the Caribbean: At World's End	471220	English	USA	300000000.0	309404152.0	2007.0	7.1
2	Color	Sam Mendes	148.0	Christoph Waltz	Action Adventure Thriller	Spectre	275868	English	UK	245000000.0	200074175.0	2015.0	6.8
3	Color	Christopher Nolan	164.0	Tom Hardy	Action Thriller	The Dark Knight Rises	1144337	English	USA	250000000.0	448130642.0	2012.0	8.5
5	Color	Andrew Stanton	132.0	Daryl Sabara	Action Adventure Sci-Fi	John Carter	212204	English	USA	263700000.0	73058679.0	2012.0	6.6
...
5033	Color	Shane Carruth	77.0	Shane Carruth	Drama Sci-Fi Thriller	Primer	72639	English	USA	7000.0	424760.0	2004.0	7.0
5034	Color	Neill Dela Liana	80.0	Ian Gamazon	Thriller	Cavite	589	English	Philippines	7000.0	70071.0	2005.0	6.3
5035	Color	Robert Rodriguez	81.0	Carlos Gallardo	Action Crime Drama Romance Thriller	El Mariachi	52055	Spanish	USA	7000.0	2040920.0	1992.0	6.9
5037	Color	Edward Burns	95.0	Kerry Bishé	Comedy Drama	Newlyweds	1338	English	USA	9000.0	4584.0	2011.0	6.4
5042	Color	Jon Gunn	90.0	John August	Documentary	My Date with Drew	4285	English	USA	1100.0	85222.0		