

```
import pandas as pd
import numpy as np
import seaborn as sns

import matplotlib.pyplot as plt
import matplotlib.mlab as mlab
import matplotlib
plt.style.use('ggplot')
from matplotlib.pyplot import figure

%matplotlib inline
matplotlib.rcParams['figure.figsize'] = (12,8)

pd.options.mode.chained_assignment = None
path = "/movies.csv"
df = pd.read_csv(path)
```

df

	name	rating	genre	year	released	score	votes	director	writer	star	country	budget
0	The Shining	R	Drama	1980	June 13, 1980 (United States)	8.4	927000.0	Stanley Kubrick	Stephen King	Jack Nicholson	United Kingdom	19000000.0
1	The Blue Lagoon	R	Adventure	1980	July 2, 1980 (United States)	5.8	65000.0	Randal Kleiser	Henry De Vere Stacpoole	Brooke Shields	United States	4500000.0
2	Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	June 20, 1980 (United States)	8.7	1200000.0	Irvin Kershner	Leigh Brackett	Mark Hamill	United States	18000000.0
3	Airplane!	PG	Comedy	1980	July 2, 1980 (United States)	7.7	221000.0	Jim Abrahams	Jim Abrahams	Robert Hays	United States	3500000.0
4	Caddyshack	R	Comedy	1980	July 25, 1980 (United States)	7.3	108000.0	Harold Ramis	Brian Doyle-Murray	Chevy Chase	United States	6000000.0
...
7663	More to Life	NaN	Drama	2020	October 23, 2020 (United States)	3.1	18.0	Joseph Ebanks	Joseph Ebanks	Shannon Bond	United States	7000.0
7664	Dream Round	NaN	Comedy	2020	February 7, 2020 (United States)	4.7	36.0	Dusty Dukatz	Lisa Huston	Michael Saquella	United States	NaN
7665	Saving Mbango	NaN	Drama	2020	April 27, 2020 (Cameroon)	5.7	29.0	Nkanya Nkwai	Lynno Lvert	Onyama Laura	United States	58750.0
7666	It's Just Us	NaN	Drama	2020	October 1, 2020 (United States)	NaN	NaN	James Randall	James Randall	Christina Roz	United States	15000.0
7667	Tee em el	NaN	Horror	2020	August 19, 2020 (United States)	5.7	7.0	Pereko Mosia	Pereko Mosia	Siyabonga Mabaso	South Africa	NaN

7668 rows × 15 columns

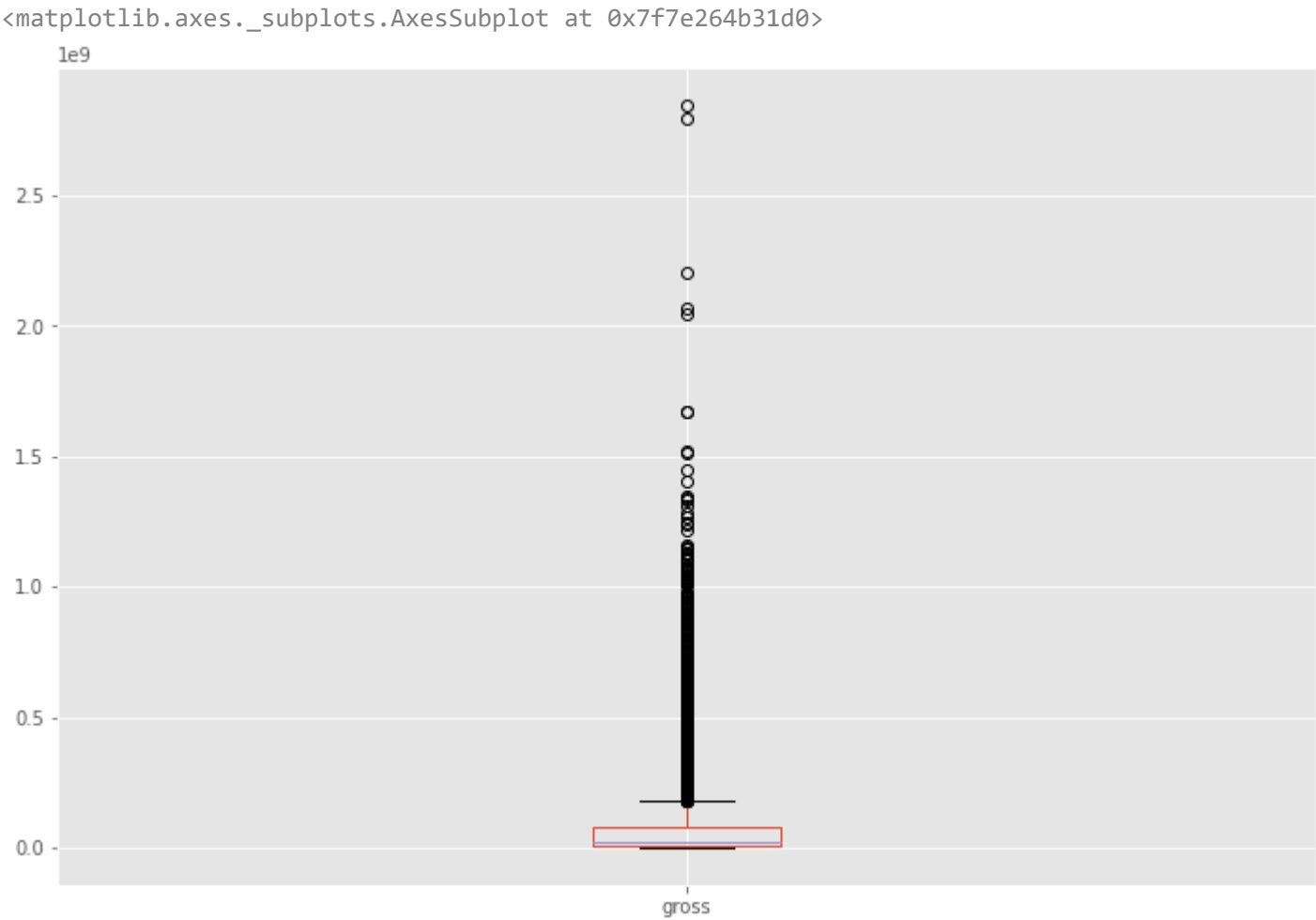
```
for col in df.columns:
    pct_missing = np.mean(df[col].isnull())
    print('{} - {}'.format(col, round(pct_missing*100)))

name - 0%
rating - 1%
genre - 0%
year - 0%
released - 0%
score - 0%
votes - 0%
director - 0%
writer - 0%
star - 0%
country - 0%
budget - 28%
gross - 2%
company - 0%
runtime - 0%

print(df.dtypes)

name          object
rating        object
genre         object
year          int64
released      object
score         float64
votes         float64
director      object
writer        object
star          object
country       object
budget        float64
gross         float64
company       object
runtime       float64
dtype: object

df.boxplot(column=['gross'])
```



df.drop_duplicates()

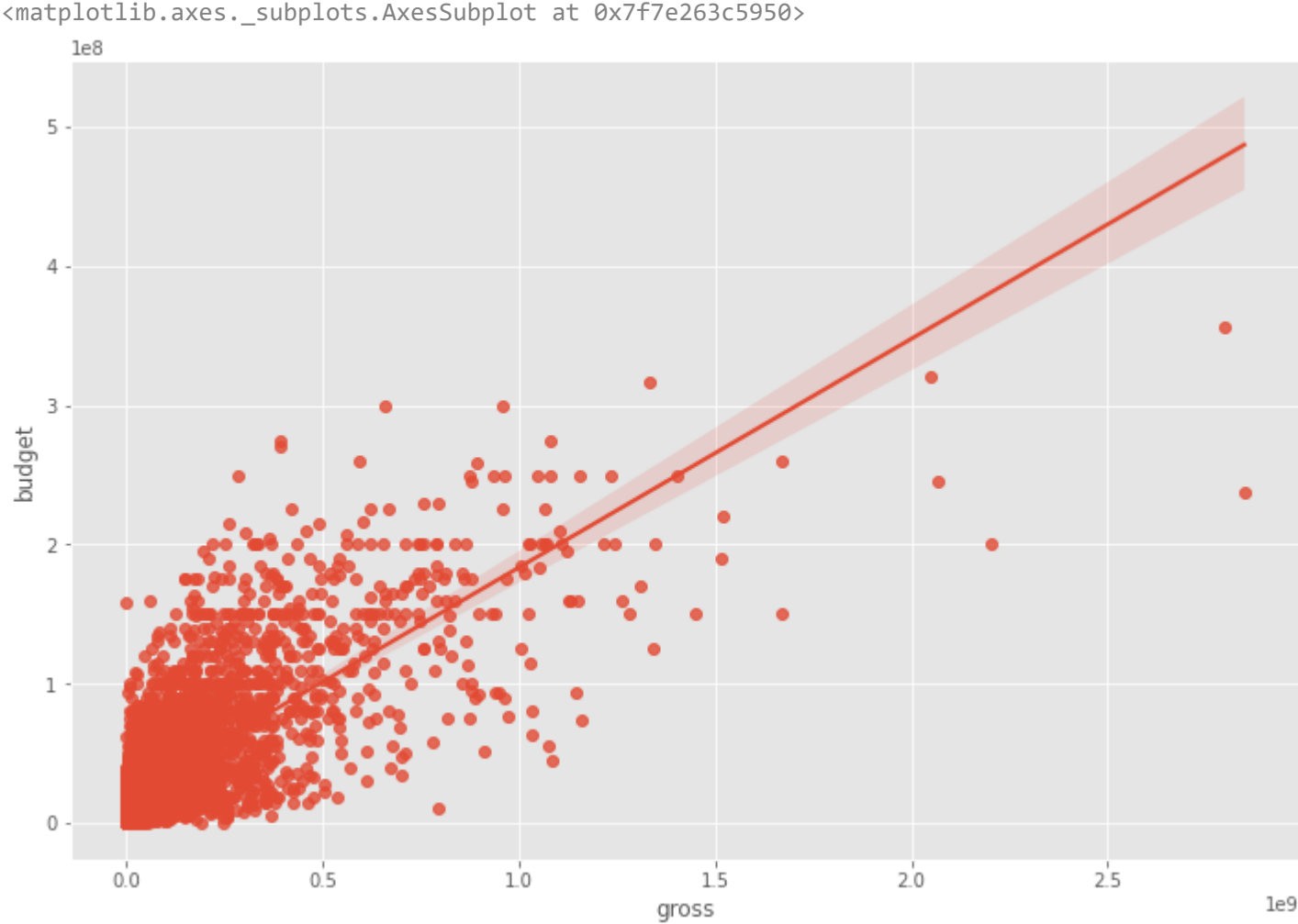
	name	rating	genre	year	released	score	votes	director	writer	star	country	budget
0	The Shining	R	Drama	1980	June 13, 1980 (United States)	8.4	927000.0	Stanley Kubrick	Stephen King	Jack Nicholson	United Kingdom	19000000.0
1	The Blue Lagoon	R	Adventure	1980	July 2, 1980 (United States)	5.8	65000.0	Randal Kleiser	Henry De Vere Stacpoole	Brooke Shields	United States	4500000.0
2	Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	June 20, 1980 (United States)	8.7	1200000.0	Irvin Kershner	Leigh Brackett	Mark Hamill	United States	18000000.0
3	Airplane!	PG	Comedy	1980	July 2, 1980 (United States)	7.7	221000.0	Jim Abrahams	Jim Abrahams	Robert Hays	United States	3500000.0
4	Caddyshack	R	Comedy	1980	July 25, 1980 (United States)	7.3	108000.0	Harold Ramis	Brian Doyle-Murray	Chevy Chase	United States	6000000.0
...
7663	More to Life	NaN	Drama	2020	October 23, 2020 (United States)	3.1	18.0	Joseph Ebanks	Joseph Ebanks	Shannon Bond	United States	7000.0
7664	Dream Round	NaN	Comedy	2020	February 7, 2020 (United States)	4.7	36.0	Dusty Dukat	Lisa Huston	Michael Saquella	United States	NaN
7665	Saving Mbango	NaN	Drama	2020	April 27, 2020 (Cameroon)	5.7	29.0	Nkanya Nkwai	Lynno Lvert	Onyama Laura	United States	58750.0
7666	It's Just Us	NaN	Drama	2020	October 1, 2020 (United States)	NaN	NaN	James Randall	James Randall	Christina Roz	United States	15000.0
7667	Tee em el	NaN	Horror	2020	August 19, 2020 (United States)	5.7	7.0	Pereko Mosia	Pereko Mosia	Siyabonga Mabaso	South Africa	NaN

7668 rows × 15 columns

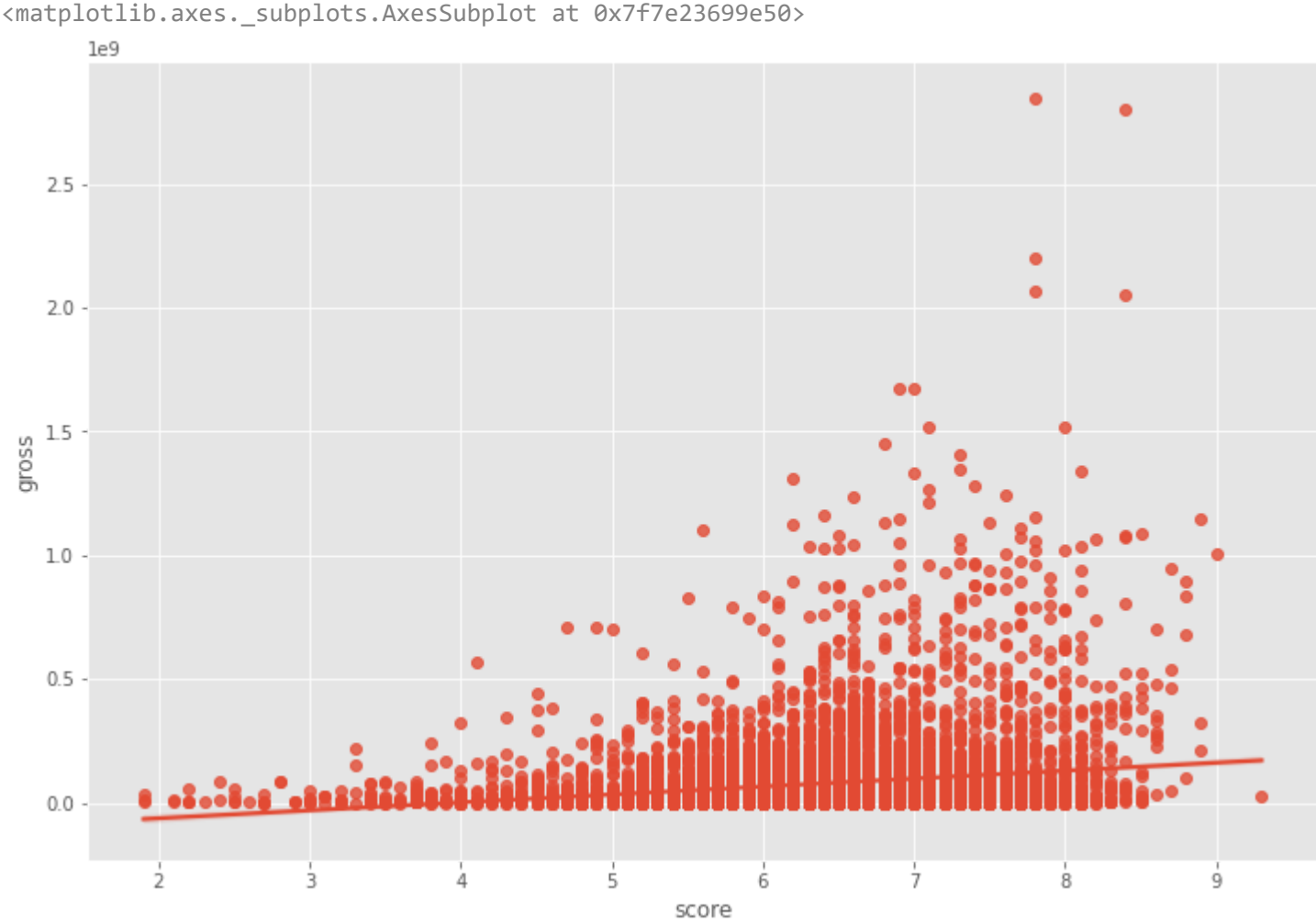
df.sort_values(by=['gross'], inplace=False, ascending=False)

	name	rating	genre	year	released	score	votes	director	writer	star	country	budget	
5445	Avatar	PG-13	Action	2009	December 18, 2009 (United States)	7.8	1100000.0	James Cameron	James Cameron	Sam Worthington	United States	237000000.0	2.
7445	Avengers: Endgame	PG-13	Action	2019	April 26, 2019 (United States)	8.4	903000.0	Anthony Russo	Christopher Markus	Robert Downey Jr.	United States	356000000.0	2.
3045	Titanic	PG-13	Drama	1997	December 19, 1997 (United States)	7.8	1100000.0	James Cameron	James Cameron	Leonardo DiCaprio	United States	200000000.0	2.
6663	Star Wars: Episode VII - The Force	PG-13	Action	2015	December 18, 2015 (United States)	7.8	876000.0	J.J. Abrams	Lawrence Kasdan	Daisy Ridley	United States	245000000.0	2.

sns.regplot(x="gross", y="budget", data=df)



sns.regplot(x="score", y="gross", data=df)



df.corr(method = 'pearson')

	year	score	votes	budget	gross	runtime
year	1.000000	0.097995	0.222945	0.329321	0.257486	0.120811
score	0.097995	1.000000	0.409182	0.076254	0.186258	0.399451
votes	0.222945	0.409182	1.000000	0.442429	0.630757	0.309212
budget	0.329321	0.076254	0.442429	1.000000	0.740395	0.320447
gross	0.257486	0.186258	0.630757	0.740395	1.000000	0.245216
runtime	0.120811	0.399451	0.309212	0.320447	0.245216	1.000000

df.corr(method = 'kendall')

	year	score	votes	budget	gross	runtime
year	1.000000	0.067652	0.331465	0.224120	0.200618	0.097184
score	0.067652	1.000000	0.300115	-0.000566	0.086046	0.283611
votes	0.331465	0.300115	1.000000	0.353702	0.548899	0.198240
budget	0.224120	-0.000566	0.353702	1.000000	0.512637	0.235483
gross	0.200618	0.086046	0.548899	0.512637	1.000000	0.168933
runtime	0.097184	0.283611	0.198240	0.235483	0.168933	1.000000

df.corr(method = 'spearman')

	year	score	votes	budget	gross	runtime
year	1.000000	0.099045	0.469829	0.317336	0.293084	0.142977
score	0.099045	1.000000	0.428138	-0.001403	0.126116	0.399857
votes	0.469829	0.428138	1.000000	0.502466	0.742050	0.290159
budget	0.317336	-0.001403	0.502466	1.000000	0.693670	0.336370
gross	0.293084	0.126116	0.742050	0.693670	1.000000	0.246243
runtime	0.142977	0.399857	0.290159	0.336370	0.246243	1.000000

```
correlation_matrix = df.corr()

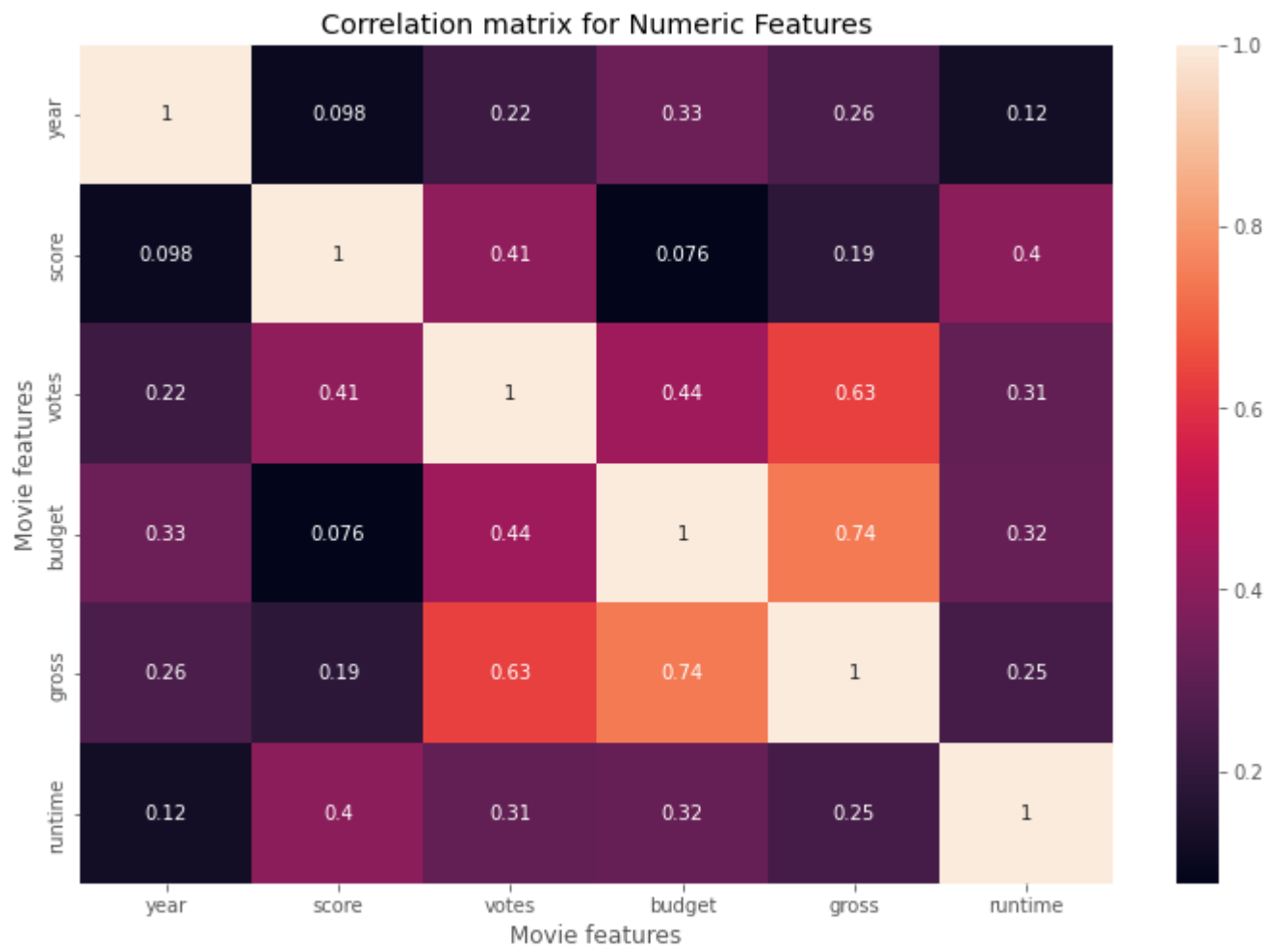
sns.heatmap(correlation_matrix, annot = True)

plt.title("Correlation matrix for Numeric Features")

plt.xlabel("Movie features")

plt.ylabel("Movie features")

plt.show()
```



```
df.apply(lambda x: x.factorize()[0]).corr(method='pearson')
```

	name	rating	genre	year	released	score	votes	director	writer	star	country
name	1.000000	0.143938	0.036367	0.965761	0.959015	-0.046733	0.287776	0.745905	0.805211	0.731565	0.142828
rating	0.143938	1.000000	-0.086723	0.156713	0.146606	0.012595	0.099972	0.085520	0.103623	0.093116	0.000494
genre	0.036367	-0.086723	1.000000	0.037184	0.035940	-0.002437	0.023285	0.047288	0.033688	0.038649	-0.015795
year	0.965761	0.156713	0.037184	1.000000	0.993190	-0.044981	0.312401	0.770497	0.824770	0.756400	0.140216
released	0.959015	0.146606	0.035940	0.993190	1.000000	-0.045761	0.299905	0.770876	0.819617	0.754468	0.148468
score	-0.046733	0.012595	-0.002437	-0.044981	-0.045761	1.000000	-0.009749	-0.022687	-0.034685	-0.009896	0.023097
votes	0.287776	0.099972	0.023285	0.312401	0.299905	-0.009749	1.000000	0.192220	0.224122	0.179601	-0.045914
director	0.745905	0.085520	0.047288	0.770497	0.770876	-0.022687	0.192220	1.000000	0.748340	0.682385	0.155471
writer	0.805211	0.103623	0.033688	0.824770	0.819617	-0.034685	0.224122	0.748340	1.000000	0.675685	0.157202
star	0.731565	0.093116	0.038649	0.756400	0.754468	-0.009896	0.179601	0.682385	0.675685	1.000000	0.182045
country	0.142828	0.000494	-0.015795	0.140216	0.148468	0.023097	-0.045914	0.155471	0.157202	0.182045	1.000000
budget	0.277488	0.193353	0.073008	0.300621	0.285691	-0.012642	0.398519	0.106617	0.187238	0.107991	-0.082082
gross	0.947324	0.158582	0.038616	0.980873	0.976423	-0.047041	0.286180	0.750911	0.805576	0.735680	0.133982
company	0.591667	-0.028035	0.009566	0.601571	0.607954	-0.028432	0.008900	0.552258	0.546151	0.527116	0.226346
runtime	0.048955	0.032741	0.001462	0.050647	0.048235	0.026436	0.106024	-0.011070	0.032264	0.035392	0.124154

```
correlation_matrix = df.apply(lambda x: x.factorize()[0]).corr(method='pearson')

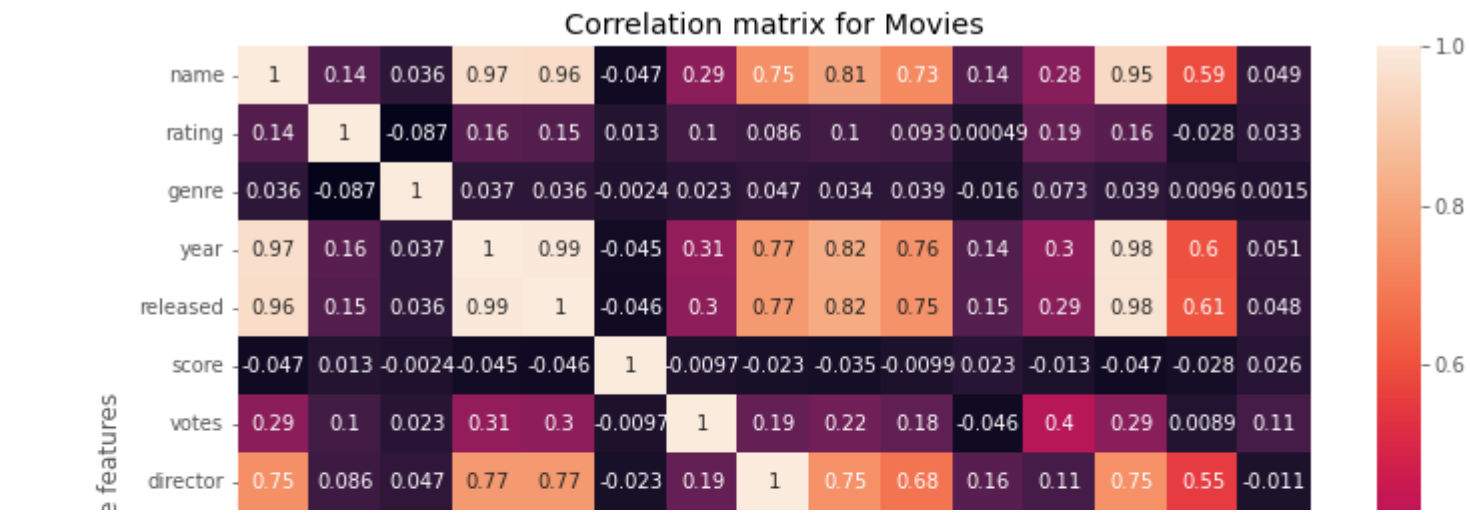
sns.heatmap(correlation_matrix, annot = True)

plt.title("Correlation matrix for Movies")

plt.xlabel("Movie features")

plt.ylabel("Movie features")

plt.show()
```



```
correlation_mat = df.apply(lambda x: x.factorize()[0]).corr()
```

```
corr_pairs = correlation_mat.unstack()
```

```
print(corr_pairs)
```

```
name      name      1.000000
          rating    0.143938
          genre     0.036367
          year      0.965761
          released  0.959015
          ...
runtime   country   0.124154
          budget    0.112097
          gross     0.042978
          company   0.005137
          runtime   1.000000
Length: 225, dtype: float64
```

```
sorted_pairs = corr_pairs.sort_values(kind="quicksort")
```

```
print(sorted_pairs)
```

```
budget     company  -0.092249
company    budget  -0.092249
genre      rating  -0.086723
rating     genre   -0.086723
budget     country -0.082082
          ...
year       year     1.000000
genre      genre    1.000000
rating     rating    1.000000
company    company   1.000000
runtime    runtime   1.000000
Length: 225, dtype: float64
```

```
strong_pairs = sorted_pairs[abs(sorted_pairs) > 0.5]
```

```
print(strong_pairs)
```

```
star      company   0.527116
company    star     0.527116
          writer    0.546151
writer     company   0.546151
director   company   0.552258
          ...
year       year     1.000000
genre      genre    1.000000
rating     rating    1.000000
company    company   1.000000
runtime    runtime   1.000000
Length: 71, dtype: float64
```

```
CompanyGrossSum = df.groupby('company')[["gross"]].sum()
```

```
CompanyGrossSumSorted = CompanyGrossSum.sort_values('gross', ascending = False)[:15]
```

```
CompanyGrossSumSorted = CompanyGrossSumSorted['gross'].astype('int64')
```

```
CompanyGrossSumSorted
```

```
company
Warner Bros.      56491421806
Universal Pictures 52514188890
Columbia Pictures 43008941346
Paramount Pictures 40493607415
Twentieth Century Fox 40257053857
Walt Disney Pictures 36327887792
New Line Cinema   19883797684
Marvel Studios    15065592411
DreamWorks Animation 11873612858
Touchstone Pictures 11795832638
Dreamworks Pictures 11635441081
Metro-Goldwyn-Mayer (MGM) 9230230105
Summit Entertainment 8373718838
Pixar Animation Studios 7886344526
Fox 2000 Pictures  7443502667
Name: gross, dtype: int64
```

```
df['Year'] = df['released'].astype(str).str[:4]
df
```

	name	rating	genre	year	released	score	votes	director	writer	star	country	budget
0	The Shining	R	Drama	1980	June 13, 1980 (United States)	8.4	927000.0	Stanley Kubrick	Stephen King	Jack Nicholson	United Kingdom	19000000.0
1	The Blue Lagoon	R	Adventure	1980	July 2, 1980 (United States)	5.8	65000.0	Randal Kleiser	Henry De Vere Stacpoole	Brooke Shields	United States	4500000.0
2	Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	June 20, 1980 (United States)	8.7	1200000.0	Irvin Kershner	Leigh Brackett	Mark Hamill	United States	18000000.0
3	Airplane!	PG	Comedy	1980	July 2, 1980 (United States)	7.7	221000.0	Jim Abrahams	Jim Abrahams	Robert Hays	United States	3500000.0
4	Caddyshack	R	Comedy	1980	July 25, 1980 (United States)	7.3	108000.0	Harold Ramis	Brian Doyle-Murray	Chevy Chase	United States	6000000.0
...
7663	More to Life	NaN	Drama	2020	October 23, 2020 (United States)	3.1	18.0	Joseph Ebanks	Joseph Ebanks	Shannon Bond	United States	7000.0
7664	Dream Round	NaN	Comedy	2020	February 7, 2020 (United States)	4.7	36.0	Dusty Dukatz	Lisa Huston	Michael Saquella	United States	NaN

df.groupby(['company', 'year'])["gross"].sum()

gross		
	company	year
	"DIA" Productions GmbH & Co. KG	2003
	"Weathering With You" Film Partners	2019
	.406 Production	1996
	1+2 Seisaku linkai	2000
	10 West Studios	2010

	i am OTHER	2015
	i5 Films	2001
	iDeal Partners Film Fund	2013
	micro_scope	2010
	thefyzz	2017

4536 rows × 1 columns

CompanyGrossSum = df.groupby(['company', 'year'])["gross"].sum()

CompanyGrossSumSorted = CompanyGrossSum.sort_values(['gross','company','year'], ascending = False)[:15]

CompanyGrossSumSorted = CompanyGrossSumSorted['gross'].astype('int64')

CompanyGrossSumSorted

company	year	
Walt Disney Pictures	2019	5773131804
Marvel Studios	2018	4018631866
Universal Pictures	2015	3834354888
Twentieth Century Fox	2009	3793491246
Walt Disney Pictures	2017	3789382071
Paramount Pictures	2011	3565705182
Warner Bros.	2010	3300479986
	2011	3223799224
Walt Disney Pictures	2010	3104474158
Paramount Pictures	2014	3071298586
Columbia Pictures	2006	2934631933
	2019	2932757449
Marvel Studios	2019	2797501328
Warner Bros.	2018	2774168962
Columbia Pictures	2011	2738363306
Name: gross, dtype: int64		

CompanyGrossSum = df.groupby(['company'])["gross"].sum()

CompanyGrossSumSorted = CompanyGrossSum.sort_values(['gross','company'], ascending = False)[:15]

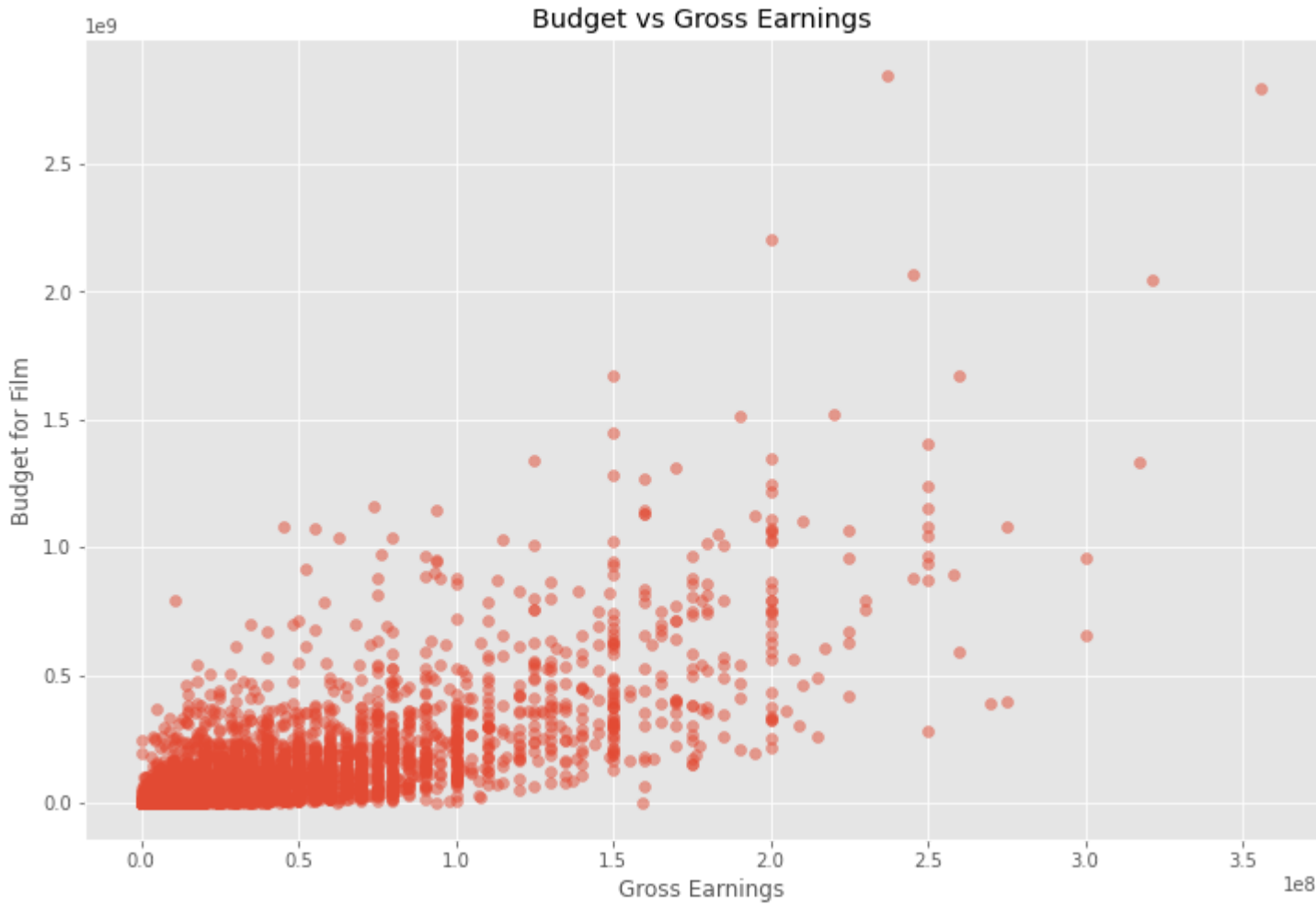
CompanyGrossSumSorted = CompanyGrossSumSorted['gross'].astype('int64')

CompanyGrossSumSorted

company	
Warner Bros.	56491421806
Universal Pictures	52514188890
Columbia Pictures	43008941346
Paramount Pictures	40493607415
Twentieth Century Fox	40257053857
Walt Disney Pictures	36327887792
New Line Cinema	19883797684
Marvel Studios	15065592411
DreamWorks Animation	11873612858
Touchstone Pictures	11795832638
Dreamworks Pictures	11635441081
Metro-Goldwyn-Mayer (MGM)	9230230105
Summit Entertainment	8373718838
Pixar Animation Studios	7886344526
Fox 2000 Pictures	7443502667
Name: gross, dtype: int64	

plt.scatter(x=df['budget'], y=df['gross'], alpha=0.5)
plt.title('Budget vs Gross Earnings')


```
plt.xlabel('Gross Earnings')
plt.ylabel('Budget for Film')
plt.show()
```



df

	name	rating	genre	year	released	score	votes	director	writer	star	country	budget
0	The Shining	R	Drama	1980	June 13, 1980 (United States)	8.4	927000.0	Stanley Kubrick	Stephen King	Jack Nicholson	United Kingdom	19000000.0
1	The Blue Lagoon	R	Adventure	1980	July 2, 1980 (United States)	5.8	65000.0	Randal Kleiser	Henry De Vere Stacpoole	Brooke Shields	United States	4500000.0
2	Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	June 20, 1980 (United States)	8.7	1200000.0	Irvin Kershner	Leigh Brackett	Mark Hamill	United States	18000000.0
3	Airplane!	PG	Comedy	1980	July 2, 1980 (United States)	7.7	221000.0	Jim Abrahams	Jim Abrahams	Robert Hays	United States	3500000.0
4	Caddyshack	R	Comedy	1980	July 25, 1980 (United States)	7.3	108000.0	Harold Ramis	Brian Doyle-Murray	Chevy Chase	United States	6000000.0
...
7663	More to Life	NaN	Drama	2020	October 23, 2020 (United States)	3.1	18.0	Joseph Ebanks	Joseph Ebanks	Shannon Bond	United States	7000.0
7664	Dream Round	NaN	Comedy	2020	February 7, 2020 (United States)	4.7	36.0	Dusty Dukatz	Lisa Huston	Michael Saquella	United States	NaN
7665	Saving Mbango	NaN	Drama	2020	April 27, 2020 (Cameroon)	5.7	29.0	Nkanya Nkwai	Lynno Lvert	Onyama Laura	United States	58750.0
7666	It's Just Us	NaN	Drama	2020	October 1, 2020 (United States)	NaN	NaN	James Randall	James Randall	Christina Roz	United States	15000.0
7667	Tee em el	NaN	Horror	2020	August 19, 2020 (United States)	5.7	7.0	Pereko Mosia	Pereko Mosia	Siyabonga Mabaso	South Africa	NaN

7668 rows × 16 columns

```
df_numerized = df
```

```
for col_name in df_numerized.columns:
    if(df_numerized[col_name].dtype == 'object'):
        df_numerized[col_name]= df_numerized[col_name].astype('category')
        df_numerized[col_name] = df_numerized[col_name].cat.codes
```

```
df_numerized
```

	name	rating	genre	year	released	score	votes	director	writer	star	country	budget	gross	company
0	6587	6	6	1980	1705	8.4	927000.0	2589	4014	1047	54	19000000.0	46998772.0	2319
1	5573	6	1	1980	1492	5.8	65000.0	2269	1632	327	55	4500000.0	58853106.0	731
2	5142	4	0	1980	1771	8.7	1200000.0	1111	2567	1745	55	18000000.0	538375067.0	1540
3	286	4	4	1980	1492	7.7	221000.0	1301	2000	2246	55	3500000.0	83453539.0	1812
4	1027	6	4	1980	1543	7.3	108000.0	1054	521	410	55	6000000.0	30846344.0	1771

```
df_numerized.corr(method='pearson')
```

	name	rating	genre	year	released	score	votes	director	writer	star	country	budget	gross	company
name	1.000000	-0.008069	0.016355	0.011453	-0.011311	0.017097	0.013088	0.009079	0.009081	0.006472	-0.010737	0		
rating	-0.008069	1.000000	0.072423	0.008779	0.016613	-0.001314	0.033225	0.019483	-0.005921	0.013405	0.081244	-0		
genre	0.016355	0.072423	1.000000	-0.081261	0.029822	0.027965	-0.145307	-0.015258	0.006567	-0.005477	-0.037615	-0		
year	0.011453	0.008779	-0.081261	1.000000	-0.000695	0.097995	0.222945	-0.020795	-0.008656	-0.027242	-0.070938	0		
released	-0.011311	0.016613	0.029822	-0.000695	1.000000	0.042788	0.016097	-0.001478	-0.002404	0.015777	-0.020427	0		
score	0.017097	-0.001314	0.027965	0.097995	0.042788	1.000000	0.409182	0.009559	0.019416	-0.001609	-0.133348	0		
votes	0.013088	0.033225	-0.145307	0.222945	0.016097	0.409182	1.000000	0.000260	0.000892	-0.019282	0.073625	0		
director	0.009079	0.019483	-0.015258	-0.020795	-0.001478	0.009559	0.000260	1.000000	0.299067	0.039234	0.017490	-0		
writer	0.009081	-0.005921	0.006567	-0.008656	-0.002404	0.019416	0.000892	0.299067	1.000000	0.027245	0.015343	-0		
star	0.006472	0.013405	-0.005477	-0.027242	0.015777	-0.001609	-0.019282	0.039234	0.027245	1.000000	-0.012998	-0		
country	-0.010737	0.081244	-0.037615	-0.070938	-0.020427	-0.133348	0.073625	0.017490	0.015343	-0.012998	1.000000	0		
budget	0.023970	-0.176002	-0.356564	0.329321	0.014683	0.076254	0.442429	-0.012272	-0.039451	-0.019589	0.054063	1		
gross	0.005533	-0.107339	-0.235650	0.257486	0.001659	0.186258	0.630757	-0.014441	-0.023519	-0.002717	0.092129	0		
company	0.009211	-0.032943	-0.071067	-0.010431	-0.010474	0.001030	0.133204	0.004404	0.005646	0.012442	0.095548	0		
runtime	0.010392	0.062145	-0.052711	0.120811	0.000868	0.399451	0.309212	0.017624	-0.003511	0.010174	-0.078412	0		
Year	-0.011725	0.013475	0.028397	-0.001562	0.993694	0.040993	0.017337	-0.000105	-0.002892	0.015406	-0.022277	0		

```
correlation_matrix = df_numerized.corr(method='pearson')
```

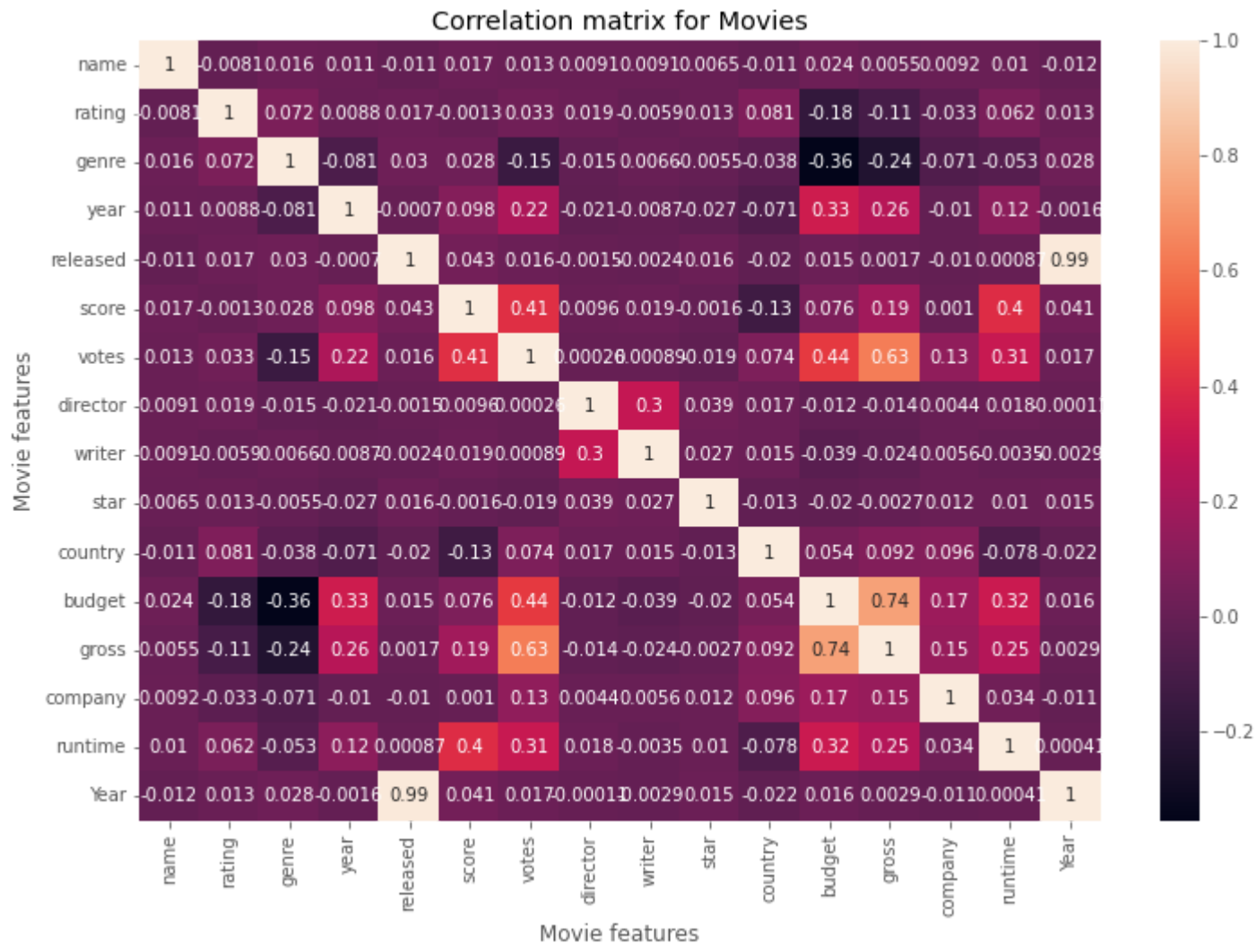
```
sns.heatmap(correlation_matrix, annot = True)
```

```
plt.title("Correlation matrix for Movies")
```

```
plt.xlabel("Movie features")
```

```
plt.ylabel("Movie features")
```

```
plt.show()
```



```
for col_name in df.columns:
    if(df[col_name].dtype == 'object'):
        df[col_name]= df[col_name].astype('category')
        df[col_name] = df[col_name].cat.codes
```

```
sns.swarmplot(x="rating", y="gross", data=df)
```



```
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 53.2% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 48.4% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 60.9% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 80.6% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 84.4% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 88.2% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 94.4% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 11.1% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
/usr/local/lib/python3.7/dist-packages/seaborn/categorical.py:1296: UserWarning: 76.9% of the points cannot be placed; you m
warnings.warn(msg, UserWarning)
<matplotlib.axes._subplots.AxesSubplot at 0x7f7e218e6cd0>
```

