In [1]:

```python
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib
import matplotlib.pyplot as plt
plt.style.use('ggplot')
from matplotlib.pyplot import figure

%matplotlib inline
matplotlib.rcParams['figure.figsize'] = (12,8) # adjust the configuration of
```

In [2]:

```python
# Read in the data
df = pd.read_csv(r"C:\Users\tamil\Desktop\Tamil\DATA ANALYST\Python\Movies\m
```

In [3]:

```python
df.head()
```

Out[3]:

	name	rating	genre	year	released	score	votes	director	
0	The Shining	R	Drama	1980	1980-06-13	8.4	927000	Stanley Kubrick	S
1	The Blue Lagoon	R	Adventure	1980	1980-07-02	5.8	65000	Randal Kleiser	H St
2	Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	1980-06-20	8.7	1200000	Irvin Kershner	E
3	Airplane!	PG	Comedy	1980	1980-07-02	7.7	221000	Jim Abrahams	Ab
4	Caddyshack	R	Comedy	1980	1980-07-25	7.3	108000	Harold Ramis	

In [9]:

```python
# Find missing data's
for col in df.columns:
    pct_missing = np.mean(df[col].isnull())
    print('{} - {}%'.format(col,pct_missing))
```

```
name - 0.0%
rating - 0.0%
genre - 0.0%
year - 0.0%
released - 0.0%
score - 0.0%
votes - 0.0%
director - 0.0%
writer - 0.0%
star - 0.0%
country - 0.0%
budget - 0.0%
gross - 0.0%
company - 0.0%
runtime - 0.0%
```

In [5]:

```
# data types for our columns
df.dtypes
```

Out[5]:

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

In [8]:

```
# Replace null values
#df['budget'] = pd.to_numeric(df['budget'], errors='coerce')  # Converts nor
df['rating'] = df['rating'].fillna(0).astype(object)

#df['gross'] = pd.to_numeric(df['gross'], errors='coerce')  # Converts non-r
df['company'] = df['company'].fillna(0).astype(object)
```

In [7]:

```
df
```

Out[7]:

	name	rating	genre	year	released	score	votes	director
0	The Shining	R	Drama	1980	1980-06-13	8.4	927000	Stanley Kubrick
1	The Blue Lagoon	R	Adventure	1980	1980-07-02	5.8	65000	Randal Kleiser
2	Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	1980-06-20	8.7	1200000	Irvin Kershner
3	Airplane!	PG	Comedy	1980	1980-07-02	7.7	221000	Jim Abrahams
4	Caddyshack	R	Comedy	1980	1980-07-25	7.3	108000	Harold Ramis
...	...	...	...	...	...	...	...	...
1103	True Colors	R	Drama	1991	1991-03-15	6.3	5000	Herbert Ross
1104	A Kiss Before Dying	R	Crime	1991	1991-04-26	5.7	5100	James Dearden
1105	Bingo	PG	Adventure	1991	1991-08-09	4.9	2700	Matthew Robbins
1106	Body Parts	R	Horror	1991	1991-08-02	5.6	4100	Eric Red
1107	One Good Cop	R	Action	1991	1991-05-03	5.9	3900	Heywood Gould

1108 rows × 15 columns

In [10]:

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

Out[10]:

	name	rating	genre	year	released	score	votes	director
107	E.T. the Extra-Terrestrial	PG	Family	1982	1982-06-11	7.8	381000	Steven Spielberg
2	Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	1980-06-20	8.7	1200000	Irvin Kershner
1002	Terminator 2: Judgment Day	R	Action	1991	1991-07-03	8.5	1000000	James Cameron
896	Ghost	PG-13	Drama	1990	1990-07-13	7.1	203000	Jerry Zucker
895	Home Alone	PG	Comedy	1990	1990-11-16	7.6	501000	Chris Columbus
...	...	...	...	...	...	...	...	...
1049	The Lovers on the Bridge	R	Drama	1991	1999-07-02	7.6	13000	Leos Carax
256	My Brother's Wedding	Not Rated	Drama	1983	March 1985	7.2	826	Charles Burnett
403	Smooth Talk	PG-13	Drama	1985	1985-11-15	6.5	2200	Joyce Chopra
423	Crimewave	PG-13	Comedy	1985	1986-04-25	5.7	5300	Sam Raimi
154	Parasite	R	Horror	1982	1982-03-12	3.9	2300	Charles Band

1108 rows × 15 columns

In [11]:

```
pd.set_option('display.max_rows', None)
```

In [12]:

```
# drop duplicates
df.drop_duplicates()
```

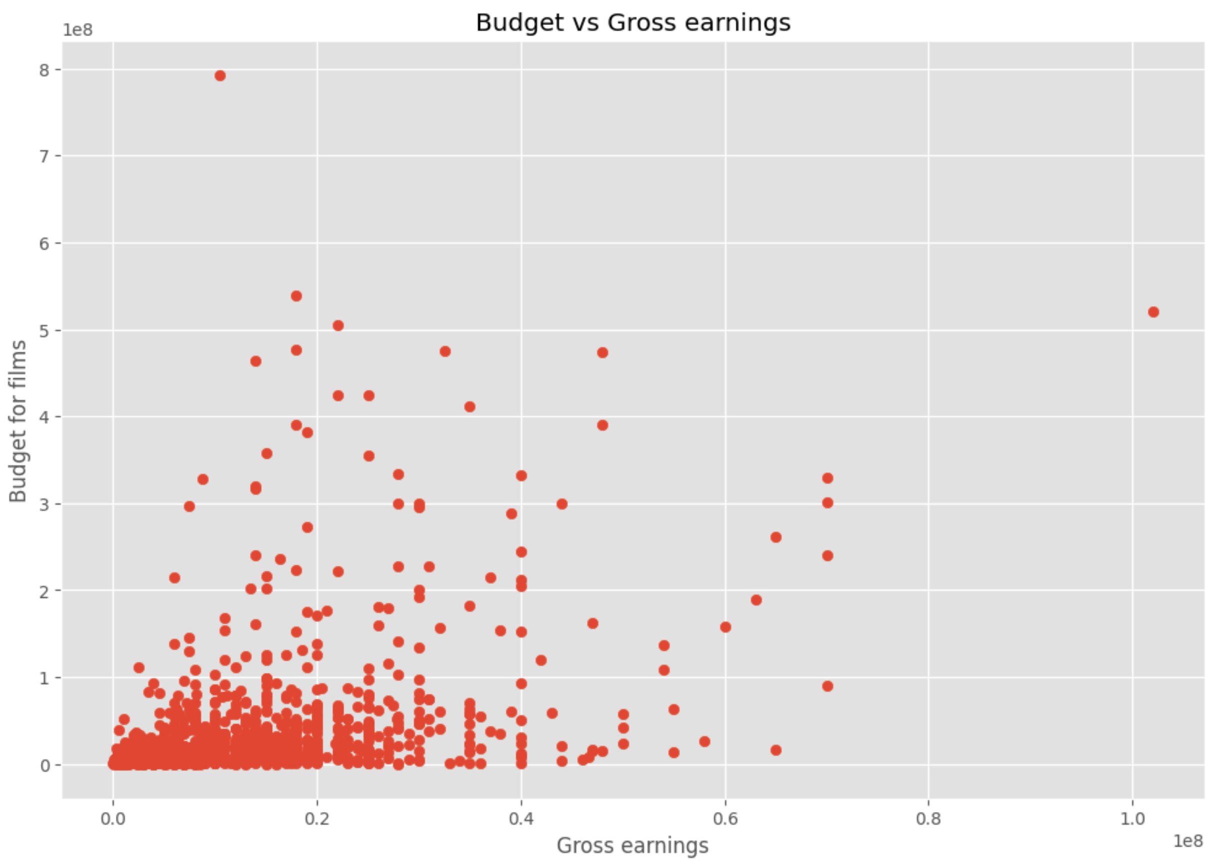
Out[12]:

	name	rating	genre	year	released	score	votes	
0	The Shining	R	Drama	1980	1980-06-13	8.4	927000	
1	The Blue Lagoon	R	Adventure	1980	1980-07-02	5.8	65000	Ran
2	Star Wars: Episode V - The Empire Strikes Back	PG	Action	1980	1980-06-20	8.7	1200000	Irvi
3	Airplane!	PG	Comedy	1980	1980-07-02	7.7	221000	Jim
4	Caddyshack	R	Comedy	1980	1980-07-25	7.3	108000	Ha
5	Friday the 13th	R	Horror	1980	1980-05-09	6.4	123000	Cu
6	The Blues Brothers	R	Action	1980	1980-06-20	7.9	188000	Jo
7	Raging Bull	R	Biography	1980	1980-12-19	8.2	330000	
8	Superman II	PG	Action	1980	1981-06-19	6.8	101000	Rich
9	The Long Riders	R	Biography	1980	1980-05-16	7.0	10000	
10	Any Which Way You Can	PG	Action	1980	1980-12-17	6.1	18000	I
11	The Gods Must Be Crazy	PG	Adventure	1980	1984-10-26	7.3	54000	
12	Popeye	PG	Adventure	1980	1980-12-12	5.3	30000	Rob
13	Ordinary People	R	Drama	1980	1980-09-19	7.7	49000	
14	Dressed to Kill	R	Crime	1980	1980-07-25	7.1	37000	
15	Somewhere in Time	PG	Drama	1980	1980-10-03	7.2	27000	
16	9 to 5	PG	Comedy	1980	1980-12-19	6.9	29000	Co
17	The Fog	R	Horror	1980	1980-02-08	6.8	66000	
18	Cruising	R	Crime	1980	1980-02-15	6.5	20000	
19	Heaven's Gate	R	Adventure	1980	1981-04-24	6.8	14000	

	name	rating	genre	year	released	score	votes	
1095	Mortal Thoughts	R	Mystery	1991	1991-04-19	5.7	8400	Ala
1096	Another You	R	Comedy	1991	1991-07-26	5.4	3700	
1097	For the Boys	R	Comedy	1991	1991-11-27	6.4	5300	M
1098	Beastmaster 2: Through the Portal of Time	PG-13	Action	1991	1991-08-30	4.1	3000	Sy
1099	Eve of Destruction	R	Action	1991	1991-01-18	4.9	2100	
1100	A Rage in Harlem	R	Comedy	1991	1991-05-03	5.9	2000	
1101	The Super	R	Comedy	1991	1991-10-04	5.6	5400	I
1102	At Play in the Fields of the Lord	R	Drama	1991	1991-12-06	6.8	3100	
1103	True Colors	R	Drama	1991	1991-03-15	6.3	5000	He
1104	A Kiss Before Dying	R	Crime	1991	1991-04-26	5.7	5100	
1105	Bingo	PG	Adventure	1991	1991-08-09	4.9	2700	
1106	Body Parts	R	Horror	1991	1991-08-02	5.6	4100	
1107	One Good Cop	R	Action	1991	1991-05-03	5.9	3900	

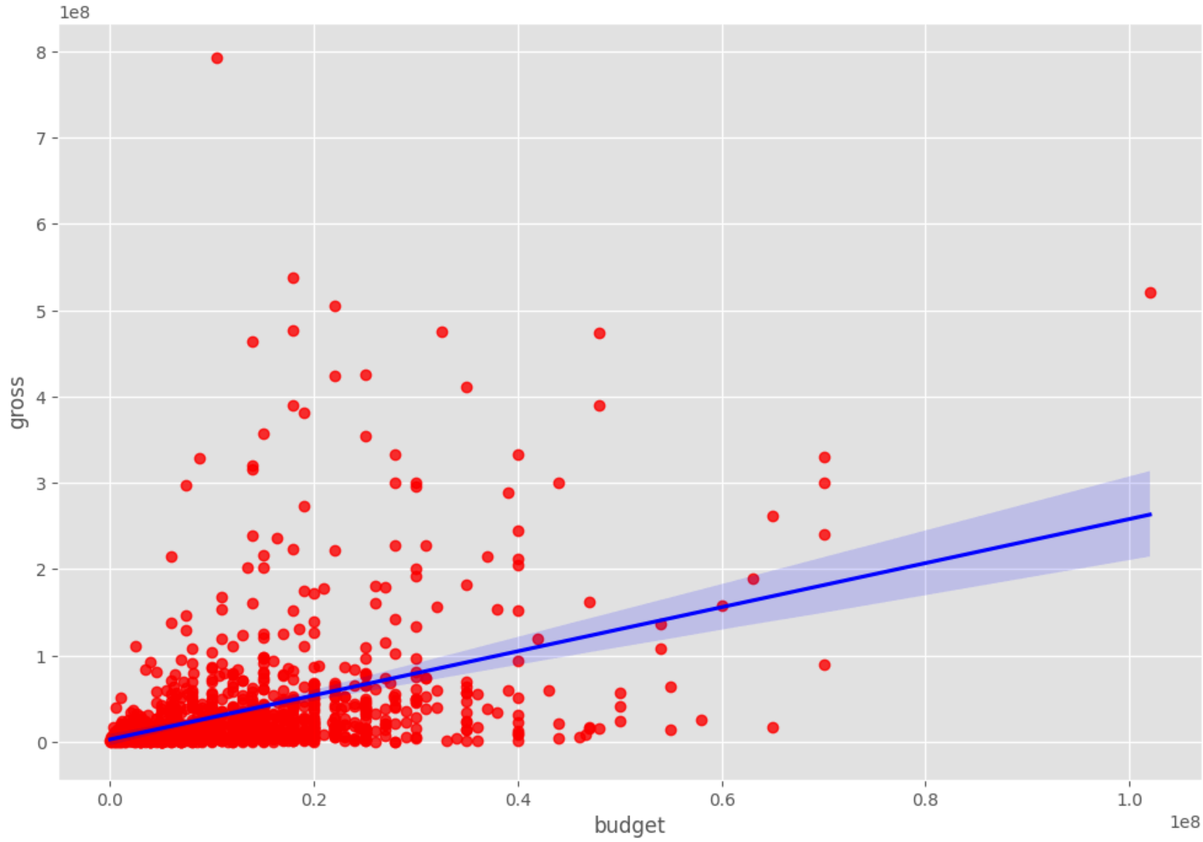
In [13]:

```
# Scattre plot with budget vs gross

plt.scatter(x=df['budget'], y=df['gross'])

plt.title('Budget vs Gross earnings')

plt.xlabel('Gross earnings')

plt.ylabel('Budget for films')

plt.show()
```



In [14]:
```
# Plot budget vs gross using seaborn

sns.regplot(x='budget', y='gross', data=df, scatter_kws={"color":"red"}, lir
```

Out[14]: <Axes: xlabel='budget', ylabel='gross'>



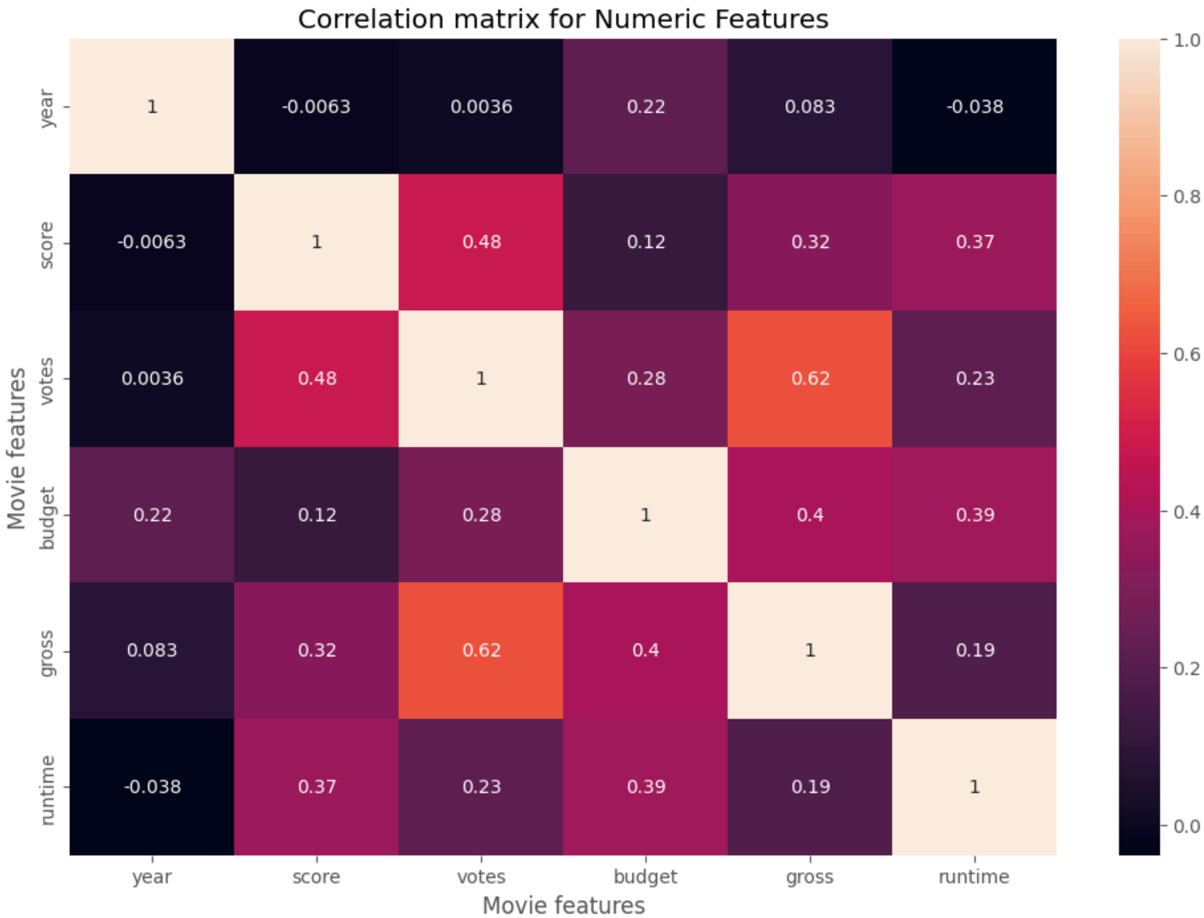
In [16]:
```
# Looking at correlation
df.select_dtypes(include='number').corr(method='pearson')
```

Out[16]:

	year	score	votes	budget	gross	runtime
year	1.000000	-0.006319	0.003583	0.220788	0.083008	-0.038301
score	-0.006319	1.000000	0.480645	0.117596	0.323590	0.369157
votes	0.003583	0.480645	1.000000	0.279210	0.622154	0.228393
budget	0.220788	0.117596	0.279210	1.000000	0.395682	0.386728
gross	0.083008	0.323590	0.622154	0.395682	1.000000	0.192768
runtime	-0.038301	0.369157	0.228393	0.386728	0.192768	1.000000

In [19]:
```
# High correlation between gross and votes
```

In [19]:
```
correlation_matrix = df.select_dtypes(include='number').corr(method='pearsor

sns.heatmap(correlation_matrix, annot= True)

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

plt.xlabel('Movie features')

plt.ylabel('Movie features')

plt.show()
```



In [20]:

```
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
```

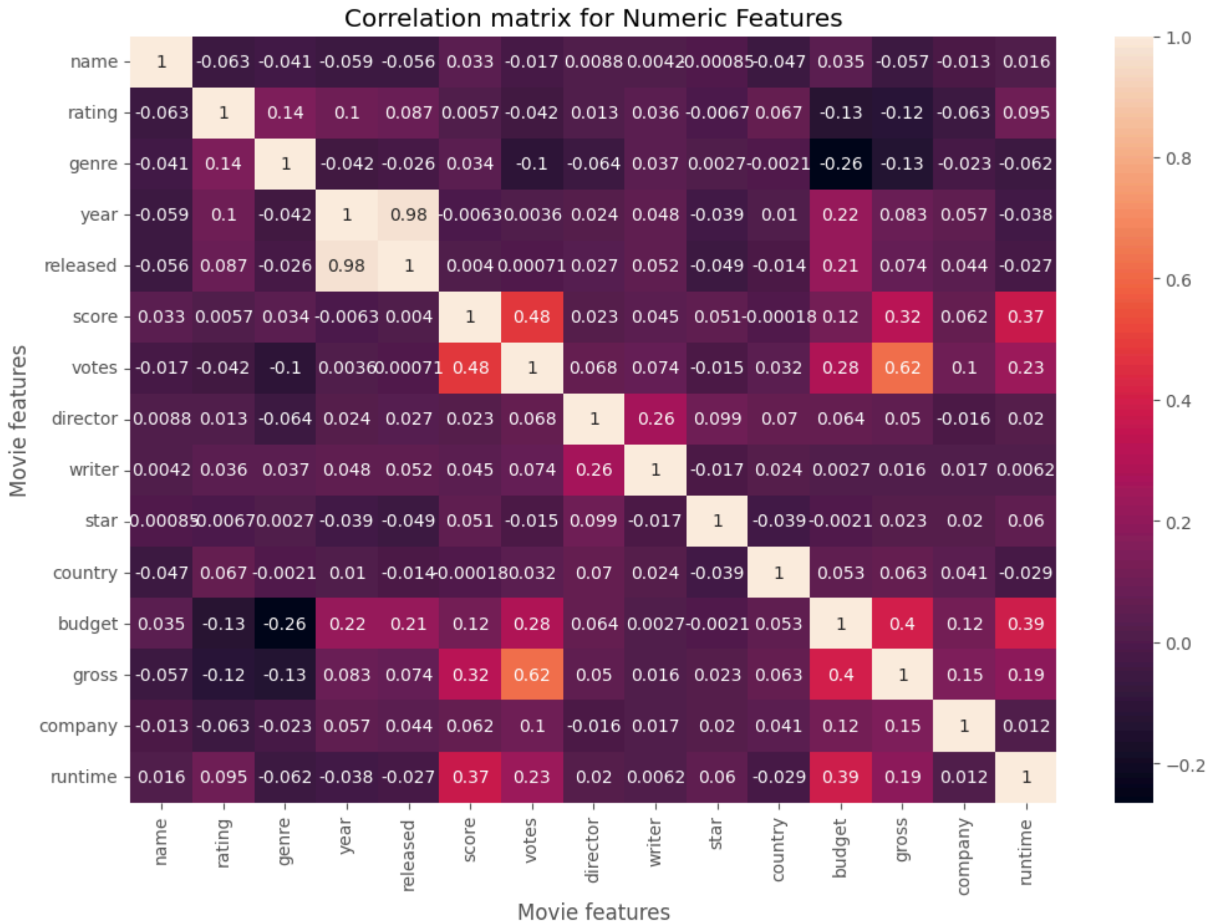
Out[20]:

	name	rating	genre	year	released	score	votes	director	writer
0	985	7	6	1980	8	8.4	927000	492	719
1	848	7	1	1980	12	5.8	65000	421	305
2	782	5	0	1980	9	8.7	1200000	186	470
3	37	5	4	1980	12	7.7	221000	229	369
4	148	7	4	1980	14	7.3	108000	168	81
5	322	7	9	1980	4	6.4	123000	480	781
6	849	7	0	1980	9	7.9	188000	266	141
7	661	7	3	1980	26	8.2	330000	332	327
8	804	5	0	1980	39	6.8	101000	432	367
9	929	7	3	1980	5	7.0	10000	543	63
10	62	5	0	1980	25	6.1	18000	53	711
11	889	5	1	1980	182	7.3	54000	212	345
12	629	5	1	1980	24	5.3	30000	444	434
13	590	7	6	1980	19	7.7	49000	452	432
14	255	7	5	1980	14	7.1	37000	45	80
15	762	5	6	1980	21	7.2	27000	218	638
16	6	5	4	1980	26	6.9	29000	72	577
17	883	7	9	1980	1	6.8	66000	251	392
18	205	7	5	1980	2	6.5	20000	552	802
19	381	7	1	1980	33	6.8	14000	344	519
20	878	5	0	1980	15	6.7	22000	107	759
21	1097	5	8	1980	16	5.3	12000	448	632
22	142	7	5	1980	9	7.2	17000	510	784
23	49	7	5	1980	0	6.2	22000	391	590
24	158	5	6	1980	33	6.1	604	306	164
25	643	7	4	1980	22	6.2	24000	181	549
26	898	7	4	1980	6	6.3	4300	125	250
27	534	7	4	1980	23	6.0	11000	302	659
28	990	7	0	1980	11	7.1	9000	437	465
29	908	7	0	1980	8	5.3	3900	356	598
30	950	5	0	1980	4	5.1	3100	70	511
31	785	5	4	1980	20	7.3	22000	559	819
32	507	7	4	1980	19	6.8	4500	279	70

	name	rating	genre	year	released	score	votes	director	writer
1089	957	7	0	1991	524	6.0	3800	321	159
1090	1058	7	5	1991	573	6.5	3800	485	709
1091	964	2	2	1991	582	6.8	2600	290	671
1092	664	7	6	1991	554	6.6	4700	327	90
1093	654	5	4	1991	548	5.8	6700	368	251
1094	286	6	0	1991	531	5.9	9000	431	667
1095	531	7	10	1991	528	5.7	8400	10	812
1096	61	7	4	1991	545	5.4	3700	337	821
1097	315	7	4	1991	564	6.4	5300	325	554
1098	90	6	0	1991	552	4.1	3000	514	25
1099	275	7	0	1991	517	4.9	2100	114	216
1100	21	7	4	1991	530	5.9	2000	33	114
1101	991	7	4	1991	556	5.6	5400	459	700
1102	71	7	6	1991	565	6.8	3100	171	602
1103	1043	7	6	1991	524	6.3	5000	174	453
1104	13	7	5	1991	529	5.7	5100	199	314
1105	107	5	1	1991	548	4.9	2700	336	376
1106	125	7	9	1991	547	5.6	4100	121	615
1107	588	7	0	1991	530	5.9	3900	175	306

In [21]:

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

sns.heatmap(correlation_matrix, annot= True)

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

plt.xlabel('Movie features')

plt.ylabel('Movie features')

plt.show()
```



In [22]:
```
correlation_mat = df_numerized.corr()
corr_pairs = correlation_mat.unstack()
corr_pairs
```

```
Out[22]: name      name        1.000000
                   rating     -0.062527
                   genre      -0.040568
                   year       -0.059270
                   released   -0.056298
                   score       0.032979
                   votes      -0.017372
                   director    0.008806
                   writer      0.004160
                   star       -0.000852
                   country    -0.047315
                   budget      0.035302
                   gross      -0.056939
                   company    -0.013220
                   runtime     0.016358
         rating    name       -0.062527
                   rating      1.000000
                   genre       0.141235
                   year        0.102554
                   released    0.087159
                   score       0.005737
                   votes      -0.042434
                   director    0.013426
                   writer      0.035815
                   star       -0.006734
                   country     0.067382
                   budget     -0.125042
                   gross      -0.116418
                   company    -0.063346
                   runtime     0.095424
         genre     name       -0.040568
                   rating      0.141235
                   genre       1.000000
                   year       -0.042129
                   released   -0.026211
                   score       0.034008
                   votes      -0.101856
                   director   -0.064215
                   writer      0.037447
                   star        0.002719
                   country    -0.002066
                   budget     -0.264239
                   gross      -0.131967
                   company    -0.022781
                   runtime    -0.061529
         year      name       -0.059270
                   rating      0.102554
                   genre      -0.042129
                   year        1.000000
                   released    0.978338
                   score      -0.006319
                   votes       0.003583
                   director    0.024055
                   writer      0.047829
                   star       -0.038926
                   country     0.010144
```

	budget	0.220788
	gross	0.083008
	company	0.056962
	runtime	-0.038301
released	name	-0.056298
	rating	0.087159
	genre	-0.026211
	year	0.978338
	released	1.000000
	score	0.004043
	votes	0.000708
	director	0.027373
	writer	0.052320
	star	-0.048765
	country	-0.013692
	budget	0.210905
	gross	0.074048
	company	0.043520
	runtime	-0.026950
score	name	0.032979
	rating	0.005737
	genre	0.034008
	year	-0.006319
	released	0.004043
	score	1.000000
	votes	0.480645
	director	0.023295
	writer	0.044893
	star	0.051408
	country	-0.000182
	budget	0.117596
	gross	0.323590
	company	0.061902
	runtime	0.369157
votes	name	-0.017372
	rating	-0.042434
	genre	-0.101856
	year	0.003583
	released	0.000708
	score	0.480645
	votes	1.000000
	director	0.067950
	writer	0.074264
	star	-0.015312
	country	0.032470
	budget	0.279210
	gross	0.622154
	company	0.103676
	runtime	0.228393
director	name	0.008806
	rating	0.013426
	genre	-0.064215
	year	0.024055
	released	0.027373
	score	0.023295
	votes	0.067950

	director	1.000000
	writer	0.260299
	star	0.099348
	country	0.069799
	budget	0.063775
	gross	0.049715
	company	-0.016137
	runtime	0.020369
writer	name	0.004160
	rating	0.035815
	genre	0.037447
	year	0.047829
	released	0.052320
	score	0.044893
	votes	0.074264
	director	0.260299
	writer	1.000000
	star	-0.016638
	country	0.024477
	budget	0.002747
	gross	0.015603
	company	0.016878
	runtime	0.006188
star	name	-0.000852
	rating	-0.006734
	genre	0.002719
	year	-0.038926
	released	-0.048765
	score	0.051408
	votes	-0.015312
	director	0.099348
	writer	-0.016638
	star	1.000000
	country	-0.038505
	budget	-0.002095
	gross	0.023099
	company	0.020439
	runtime	0.059706
country	name	-0.047315
	rating	0.067382
	genre	-0.002066
	year	0.010144
	released	-0.013692
	score	-0.000182
	votes	0.032470
	director	0.069799
	writer	0.024477
	star	-0.038505
	country	1.000000
	budget	0.052961
	gross	0.063399
	company	0.040723
	runtime	-0.029232
budget	name	0.035302
	rating	-0.125042
	genre	-0.264239

```
          year        0.220788
          released    0.210905
          score       0.117596
          votes       0.279210
          director    0.063775
          writer      0.002747
          star       -0.002095
          country     0.052961
          budget      1.000000
          gross       0.395682
          company     0.117075
          runtime     0.386728
gross     name       -0.056939
          rating     -0.116418
          genre      -0.131967
          year        0.083008
          released    0.074048
          score       0.323590
          votes       0.622154
          director    0.049715
          writer      0.015603
          star        0.023099
          country     0.063399
          budget      0.395682
          gross       1.000000
          company     0.147032
          runtime     0.192768
company   name       -0.013220
          rating     -0.063346
          genre      -0.022781
          year        0.056962
          released    0.043520
          score       0.061902
          votes       0.103676
          director   -0.016137
          writer      0.016878
          star        0.020439
          country     0.040723
          budget      0.117075
          gross       0.147032
          company     1.000000
          runtime     0.011635
runtime   name        0.016358
          rating      0.095424
          genre      -0.061529
          year       -0.038301
          released   -0.026950
          score       0.369157
          votes       0.228393
          director    0.020369
          writer      0.006188
          star        0.059706
          country    -0.029232
          budget      0.386728
          gross       0.192768
          company     0.011635
```

```
            runtime     1.000000
dtype: float64
```

In [23]:
```
sorted_pairs = corr_pairs.sort_values()
sorted_pairs
```

```
Out[23]: genre     budget     -0.264239
         budget    genre      -0.264239
         gross     genre      -0.131967
         genre     gross      -0.131967
         rating    budget     -0.125042
         budget    rating     -0.125042
         gross     rating     -0.116418
         rating    gross      -0.116418
         genre     votes      -0.101856
         votes     genre      -0.101856
         director  genre      -0.064215
         genre     director   -0.064215
         rating    company    -0.063346
         company   rating     -0.063346
         rating    name       -0.062527
         name      rating     -0.062527
         runtime   genre      -0.061529
         genre     runtime    -0.061529
         year      name       -0.059270
         name      year       -0.059270
                   gross      -0.056939
         gross     name       -0.056939
         name      released   -0.056298
         released  name       -0.056298
         star      released   -0.048765
         released  star       -0.048765
         name      country    -0.047315
         country   name       -0.047315
         votes     rating     -0.042434
         rating    votes      -0.042434
         year      genre      -0.042129
         genre     year       -0.042129
                   name       -0.040568
         name      genre      -0.040568
         star      year       -0.038926
         year      star       -0.038926
         star      country    -0.038505
         country   star       -0.038505
         year      runtime    -0.038301
         runtime   year       -0.038301
                   country    -0.029232
         country   runtime    -0.029232
         runtime   released   -0.026950
         released  runtime    -0.026950
                   genre      -0.026211
         genre     released   -0.026211
                   company    -0.022781
         company   genre      -0.022781
         votes     name       -0.017372
         name      votes      -0.017372
         writer    star       -0.016638
         star      writer     -0.016638
         company   director   -0.016137
         director  company    -0.016137
         votes     star       -0.015312
         star      votes      -0.015312
```

country	released	-0.013692
released	country	-0.013692
name	company	-0.013220
company	name	-0.013220
star	rating	-0.006734
rating	star	-0.006734
score	year	-0.006319
year	score	-0.006319
budget	star	-0.002095
star	budget	-0.002095
genre	country	-0.002066
country	genre	-0.002066
star	name	-0.000852
name	star	-0.000852
score	country	-0.000182
country	score	-0.000182
votes	released	0.000708
released	votes	0.000708
star	genre	0.002719
genre	star	0.002719
writer	budget	0.002747
budget	writer	0.002747
votes	year	0.003583
year	votes	0.003583
score	released	0.004043
released	score	0.004043
name	writer	0.004160
writer	name	0.004160
score	rating	0.005737
rating	score	0.005737
runtime	writer	0.006188
writer	runtime	0.006188
name	director	0.008806
director	name	0.008806
year	country	0.010144
country	year	0.010144
runtime	company	0.011635
company	runtime	0.011635
rating	director	0.013426
director	rating	0.013426
writer	gross	0.015603
gross	writer	0.015603
runtime	name	0.016358
name	runtime	0.016358
company	writer	0.016878
writer	company	0.016878
director	runtime	0.020369
runtime	director	0.020369
company	star	0.020439
star	company	0.020439
gross	star	0.023099
star	gross	0.023099
score	director	0.023295
director	score	0.023295
	year	0.024055
year	director	0.024055

writer	country	0.024477
country	writer	0.024477
released	director	0.027373
director	released	0.027373
country	votes	0.032470
votes	country	0.032470
name	score	0.032979
score	name	0.032979
	genre	0.034008
genre	score	0.034008
name	budget	0.035302
budget	name	0.035302
rating	writer	0.035815
writer	rating	0.035815
	genre	0.037447
genre	writer	0.037447
country	company	0.040723
company	country	0.040723
released	company	0.043520
company	released	0.043520
writer	score	0.044893
score	writer	0.044893
year	writer	0.047829
writer	year	0.047829
gross	director	0.049715
director	gross	0.049715
score	star	0.051408
star	score	0.051408
writer	released	0.052320
released	writer	0.052320
budget	country	0.052961
country	budget	0.052961
company	year	0.056962
year	company	0.056962
runtime	star	0.059706
star	runtime	0.059706
score	company	0.061902
company	score	0.061902
gross	country	0.063399
country	gross	0.063399
budget	director	0.063775
director	budget	0.063775
rating	country	0.067382
country	rating	0.067382
votes	director	0.067950
director	votes	0.067950
country	director	0.069799
director	country	0.069799
released	gross	0.074048
gross	released	0.074048
votes	writer	0.074264
writer	votes	0.074264
gross	year	0.083008
year	gross	0.083008
rating	released	0.087159
released	rating	0.087159

```
rating    runtime     0.095424
runtime   rating      0.095424
star      director    0.099348
director  star        0.099348
rating    year        0.102554
year      rating      0.102554
votes     company     0.103676
company   votes       0.103676
          budget      0.117075
budget    company     0.117075
          score       0.117596
score     budget      0.117596
genre     rating      0.141235
rating    genre       0.141235
company   gross       0.147032
gross     company     0.147032
runtime   gross       0.192768
gross     runtime     0.192768
released  budget      0.210905
budget    released    0.210905
          year        0.220788
year      budget      0.220788
votes     runtime     0.228393
runtime   votes       0.228393
director  writer      0.260299
writer    director    0.260299
budget    votes       0.279210
votes     budget      0.279210
score     gross       0.323590
gross     score       0.323590
runtime   score       0.369157
score     runtime     0.369157
budget    runtime     0.386728
runtime   budget      0.386728
budget    gross       0.395682
gross     budget      0.395682
votes     score       0.480645
score     votes       0.480645
votes     gross       0.622154
gross     votes       0.622154
released  year        0.978338
year      released    0.978338
released  released    1.000000
score     score       1.000000
rating    rating      1.000000
name      name        1.000000
genre     genre       1.000000
year      year        1.000000
director  director    1.000000
votes     votes       1.000000
writer    writer      1.000000
star      star        1.000000
budget    budget      1.000000
country   country     1.000000
gross     gross       1.000000
company   company     1.000000
```

```
runtime   runtime     1.000000
dtype: float64
```

In [24]:
```python
high_corr = sorted_pairs[(sorted_pairs)>0.5]
high_corr
```

Out[24]:
```
votes     gross       0.622154
gross     votes       0.622154
released  year        0.978338
year      released    0.978338
released  released    1.000000
score     score       1.000000
rating    rating      1.000000
name      name        1.000000
genre     genre       1.000000
year      year        1.000000
director  director    1.000000
votes     votes       1.000000
writer    writer      1.000000
star      star        1.000000
budget    budget      1.000000
country   country     1.000000
gross     gross       1.000000
company   company     1.000000
runtime   runtime     1.000000
dtype: float64
```

In [ ]: