```
import pandas as pd
from google.colab import files

files.upload()
```

```
Choose Files top_movies_by_title
• top_movies_by_title.csv(text/csv) - 11026 bytes, last modified: 9/27/2021 - 100% done
Saving top_movies_by_title.csv to top_movies_by_title (1).csv
{'top_movies_by_title.csv': b'Title,Studio,Gross,Gross (Adjusted),Year\n101 Dalmatians,Disney,144880014,869280100,1961\n2001: A
```

```
##### Problem 1 #####

table1 = pd.read_csv('top_movies_by_title.csv')
```

```
##### Problem 2 #####

table1[['Title','Year']]
```

	Title	Year
0	101 Dalmatians	1961
1	2001: A Space Odyssey	1968
2	9 to 5	1980
3	A Star Is Born (1976)	1976
4	Air Force One	1997
...	...	...
195	Up	2009
196	West Side Story	1961
197	What's Up, Doc?	1972
198	Who Framed Roger Rabbit	1988
199	Young Frankenstein	1974

200 rows × 2 columns

```
##### Problem 3 #####

table1.loc[table1['Gross'] > 100000000]
```

	Title	Studio	Gross	Gross (Adjusted)	Year	Ratio
0	101 Dalmatians	Disney	144880014	869280100	1961	6.000000
2	9 to 5	Fox	103290500	334062200	1980	3.234201
4	Air Force One	Sony	172956409	327752300	1997	1.894999
5	Airport	Universal	100489151	575168200	1970	5.723685
6	Aladdin	Buena Vista (Disney)	217350219	456248400	1992	2.099139
...	...	...	...	...	...	...
192	Transformers: Dark of the Moon	Paramount/Dreamworks	352390543	385069700	2011	1.092736
193	Transformers: Revenge of the Fallen	Paramount/Dreamworks	402111870	468938100	2009	1.166188
194	Twister	Warner Bros.	241721524	475786700	1996	1.968326
195	Up	Buena Vista (Disney)	293004164	341681900	2009	1.166133
198	Who Framed Roger Rabbit	Buena Vista (Disney)	156452370	331176600	1988	2.116789

```
##### Problem 4 #####

table1.loc[(table1['Gross'] > 100000000) & (table1['Gross'] < 200000000)].sort_values(['Gross'], ascending=False ).head(10)
```

	Title	Studio	Gross	Gross (Adjusted)	Year	Ratio
51	Gone with the Wind	MGM	198676459	1757788200	1939	8.847491
69	Indiana Jones and the Last Crusade	Paramount	197171806	429923500	1989	2.180451
188	Toy Story	Buena Vista (Disney)	191796233	381654400	1995	1.989895
52	Grease	Paramount	188755690	669632000	1978	3.547612
129	Snow White and the Seven Dwarfs	Disney	184925486	948300000	1937	5.128011
35	Dances with Wolves	Orion	184208848	380183900	1990	2.063874

```
##### Problem 5 #####

table1.loc[ table1 ['Title'].str.contains('the') ].head(10)
```

	Title	Studio	Gross	Gross (Adjusted)	Year	Ratio
13	Around the World in 80 Days	UA	42000000	562153800	1956	13.384614
17	Back to the Future	Universal	210609762	513740700	1985	2.439301
22	Beauty and the Beast	Buena Vista (Disney)	218967620	394664300	1991	1.802387
28	Butch Cassidy and the Sundance Kid	Fox	102308889	613853300	1969	6.000000
33	Close Encounters of the Third Kind	Columbia	132088635	483507200	1977	3.660475
38	Duel in the Sun	Selz.	20408163	443877500	1946	21.749998
51	Gone with the Wind	MGM	198676459	1757788200	1939	8.847491
54	Guardians of the Galaxy	Buena Vista (Disney)	333176600	358244800	2014	1.075240

```
##### Problem 6 #####

a = table1.sort_values(['Gross'], ascending = False).head(3)
print(a)
b = table1.sort_values(['Gross (Adjusted)'], ascending = False).head(3)
print(b)
```

```
                          Title               Studio  ...  Year     Ratio
137  Star Wars: The Force Awakens  Buena Vista (Disney)  ...  2015  1.000000
15                         Avatar                  Fox  ...  2009  1.112574
185                       Titanic            Paramount  ...  1997  1.789400

[3 rows x 6 columns]
                  Title Studio      Gross  Gross (Adjusted)  Year     Ratio
51   Gone with the Wind    MGM  198676459        1757788200  1939  8.847491
133           Star Wars    Fox  460998007        1549640500  1977  3.361491
176  The Sound of Music    Fox  158671368        1239013800  1965  7.808679
```

```
##### Problem 7 #####

c = table1.groupby('Studio').count()['Title'].sort_values(ascending=False)
print(c)
print('\nMovies made by Columbia: {}'.format(c['Columbia']))
print('Movies made by Fox: {}'.format(c['Fox']))
```

```
Studio
Buena Vista (Disney)       29
Warner Bros.               29
Fox                        26
Paramount                  25
Universal                  22
Disney                     11
Columbia                   10
MGM                         7
UA                          6
Sony                        6
New Line                    5
Paramount/Dreamworks        4
Dreamworks                  3
RKO                         3
Lionsgate                   3
Sum.                        2
```

```
TriS                        2
Warner Bros. (New Line)     1
IFC                         1
MPC                         1
NM                          1
Orion                       1
Selz.                       1
AVCO                        1
Name: Title, dtype: int64

Movies made by Columbia: 10
Movies made by Fox: 26
```

```
##### Problem 8 #####

table1.loc[table1['Studio'] == 'Columbia'].sort_values('Gross', ascending=False).head(1)
```

	Title	Studio	Gross	Gross (Adjusted)	Year
49	Ghostbusters	Columbia	242212467	619211400	1984

```
##### Problem 9 #####

table1['Ratio'] = table1['Gross (Adjusted)'] / table1['Gross']
table1.sort_values('Ratio', ascending=False).head(10)
```

	Title	Studio	Gross	Gross (Adjusted)	Year	Ratio
153	The Four Horsemen of the Apocalypse	MPC	9183673	399489800	1921	43.500003
143	The Bells of St. Mary's	RKO	21333333	545882400	1945	25.588238
121	Sergeant York	Warner Bros.	16361885	418671800	1941	25.588238
38	Duel in the Sun	Selz.	20408163	443877500	1946	21.749998
144	The Best Years of Our Lives	RKO	23650000	478500000	1946	20.232558
65	House of Wax	Warner Bros.	23750000	439627700	1953	18.510640
146	The Caine Mutiny	Columbia	21750000	386173500	1954	17.755103
178	The Ten Commandments	Paramount	65500000	1139700000	1956	17.400000

```
##### Problem 10 #####

f = table1.groupby('Year').count().sort_values('Title',ascending=False)
print(f)
#print(f.groupby('Title').groups)

for i in f.groupby('Title').groups:
   if i == max(f.groupby('Title').groups):
     print('\nYear(s) with maximum amount of movies are:')
     for j in f.groupby('Title').groups[i]:
       print(' ', j, '  =  ', i)
```

```
      Title  Studio  Gross  Gross (Adjusted)  Ratio
Year
2015      6       6      6                 6      6
2009      6       6      6                 6      6
2004      6       6      6                 6      6
2002      6       6      6                 6      6
1997      5       5      5                 5      5
...     ...     ...    ...               ...    ...
1957      1       1      1                 1      1
1969      1       1      1                 1      1
1983      1       1      1                 1      1
1937      1       1      1                 1      1
1921      1       1      1                 1      1

[70 rows x 5 columns]

Years with maximum amount of movies are:
  2015   =   6
  2009   =   6
  2004   =   6
  2002   =   6
```