```
import matplotlib.pyplot as plt

import pandas as pd

df=pd.read_csv ('movies.csv')
```

```
print(df.columns)
print(df.shape)
```

```
Index(['millones', 'budget', 'company', 'country', 'director', 'genre',
       'gross', 'name', 'rating', 'released', 'runtime', 'score', 'star',
       'votes', 'writer', 'year'],
      dtype='object')
(6820, 16)
```

```
print(df.head(8))
print(df.tail)
```

```
3  1986
4  1986
5  1986
6  1986
7  1986
<bound method NDFrame.tail of        millones    budget                              company country  \
0          8.0   8000000          Columbia Pictures Corporation     USA
1          6.0   6000000                     Paramount Pictures     USA
2         15.0  15000000                     Paramount Pictures     USA
3         18.5  18500000  Twentieth Century Fox Film Corporation     USA
4          9.0   9000000                    Walt Disney Pictures     USA
...        ...       ...                                    ...     ...
6815       0.0         0               Fox Searchlight Pictures      UK
6816       0.0         0                Siempre Viva Productions     USA
6817       3.5   3500000                 Warner Bros. Animation     USA
6818       0.0         0                    Borderline Presents     USA
6819       0.0         0             Les Productions du Tr����  France

             director      genre      gross                             name  \
0          Rob Reiner  Adventure   52287414                      Stand by Me
1         John Hughes     Comedy   70136369          Ferris Bueller's Day Off
2          Tony Scott     Action  179800601                          Top Gun
3       James Cameron     Action   85160248                           Aliens
4      Randal Kleiser  Adventure   18564613          Flight of the Navigator
...               ...        ...        ...                              ...
6815  Mandie Fletcher     Comedy    4750497  Absolutely Fabulous: The Movie
6816   Paul Duddridge      Drama      28368             Mothers and Daughters
6817          Sam Liu  Animation    3775000          Batman: The Killing Joke
6818    Nicolas Pesce      Drama      25981             The Eyes of My Mother
6819     Nicole Garcia      Drama      37757        From the Land of the Moon

       rating    released  runtime  score             star   votes  \
0          R  22/08/1986       89    8.1      Wil Wheaton  299174
1       PG-13  11/06/1986      103    7.8  Matthew Broderick  264740
2          PG  16/05/1986      110    6.9        Tom Cruise  236909
3          R  18/07/1986      137    8.4  Sigourney Weaver  540152
4          PG  01/08/1986       90    6.9       Joey Cramer   36636
...       ...         ...      ...    ...              ...     ...
6815       R  22/07/2016       91    5.4  Jennifer Saunders    9161
6816    PG-13  06/05/2016       90    4.9        Selma Blair    1959
6817       R  25/07/2016       76    6.5      Kevin Conroy   36333
6818       R  02/12/2016       76    6.2    Kika Magalh���    6947
6819       R  28/07/2017      120    6.7   Marion Cotillard    2411

                writer  year
0        Stephen King  1986
1         John Hughes  1986
2            Jim Cash  1986
3       James Cameron  1986
4       Mark H. Baker  1986
...                ...   ...
6815  Jennifer Saunders  2016
6816     Paige Cameron  2016
6817    Brian Azzarello  2016
6818      Nicolas Pesce  2016
6819       Milena Agus  2016

[6820 rows x 16 columns]>
```

```
print(df.info())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6820 entries, 0 to 6819
Data columns (total 16 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   millones  6820 non-null   float64
 1   budget    6820 non-null   int64
 2   company   6820 non-null   object
 3   country   6820 non-null   object
 4   director  6820 non-null   object
 5   genre     6820 non-null   object
 6   gross     6820 non-null   int64
 7   name      6820 non-null   object
 8   rating    6820 non-null   object
 9   released  6820 non-null   object
 10  runtime   6820 non-null   int64
 11  score     6820 non-null   float64
 12  star      6820 non-null   object
 13  votes     6820 non-null   int64
 14  writer    6820 non-null   object
 15  year      6820 non-null   int64
dtypes: float64(2), int64(5), object(9)
memory usage: 852.6+ KB
None
```

```
#seleccion de 8 columnas

# Selección  de columnas
df2 = df[['budget', 'score', 'released', 'genre', 'company', 'country', 'director']]
print(df.shape)
print(df2)
print(df2)
```

```
(6820, 16)
        budget  score     released      genre  \
0      8000000    8.1   22/08/1986  Adventure
1      6000000    7.8   11/06/1986     Comedy
2     15000000    6.9   16/05/1986     Action
3     18500000    8.4   18/07/1986     Action
4      9000000    6.9   01/08/1986  Adventure
...        ...    ...          ...        ...
6815         0    5.4   22/07/2016     Comedy
6816         0    4.9   06/05/2016      Drama
6817   3500000    6.5   25/07/2016  Animation
6818         0    6.2   02/12/2016      Drama
6819         0    6.7   28/07/2017      Drama

                                 company country        director
0           Columbia Pictures Corporation     USA       Rob Reiner
1                     Paramount Pictures     USA      John Hughes
2                     Paramount Pictures     USA       Tony Scott
3     Twentieth Century Fox Film Corporation  USA   James Cameron
4                   Walt Disney Pictures     USA   Randal Kleiser
...                                    ...     ...             ...
6815               Fox Searchlight Pictures      UK  Mandie Fletcher
6816             Siempre Viva Productions     USA   Paul Duddridge
6817               Warner Bros. Animation     USA          Sam Liu
6818                  Borderline Presents     USA   Nicolas Pesce
6819           Les Productions du Tr����  France    Nicole Garcia

[6820 rows x 7 columns]
        budget  score     released      genre  \
0      8000000    8.1   22/08/1986  Adventure
1      6000000    7.8   11/06/1986     Comedy
2     15000000    6.9   16/05/1986     Action
3     18500000    8.4   18/07/1986     Action
4      9000000    6.9   01/08/1986  Adventure
...        ...    ...          ...        ...
6815         0    5.4   22/07/2016     Comedy
6816         0    4.9   06/05/2016      Drama
6817   3500000    6.5   25/07/2016  Animation
6818         0    6.2   02/12/2016      Drama
6819         0    6.7   28/07/2017      Drama

                                 company country        director
0           Columbia Pictures Corporation     USA       Rob Reiner
1                     Paramount Pictures     USA      John Hughes
2                     Paramount Pictures     USA       Tony Scott
3     Twentieth Century Fox Film Corporation  USA   James Cameron
```

```
       4                     Walt Disney Pictures    USA    Randal Kleiser
    ...                                        ...    ...              ...
    6815            Fox Searchlight Pictures     UK  Mandie Fletcher
    6816            Siempre Viva Productions    USA    Paul Duddridge
    6817            Warner Bros. Animation      USA          Sam Liu
    6818                 Borderline Presents    USA    Nicolas Pesce
    6819           Les Productions du Tr����  France   Nicole Garcia

    [6820 rows x 7 columns]
```

```python
print(df['year'].describe())
```

```
count    6820.000000
mean     2001.000293
std         8.944501
min      1986.000000
25%      1993.000000
50%      2001.000000
75%      2009.000000
max      2016.000000
Name: year, dtype: float64
```

```python
print(df['budget'].describe())
```

```
count    6.820000e+03
mean     2.458113e+07
std      3.702254e+07
min      0.000000e+00
25%      0.000000e+00
50%      1.100000e+07
75%      3.200000e+07
max      3.000000e+08
Name: budget, dtype: float64
```

```python
print(df['released'].describe())
```

```
count           6820
unique          2403
top       04/10/1991
freq              10
Name: released, dtype: object
```

```python
print(df['genre'].value_counts())
```

```
genre
Comedy       2080
Drama        1444
Action       1331
Crime         522
Adventure     392
Biography     359
Horror        277
Animation     277
Mystery        38
Fantasy        32
Thriller       18
Romance        15
Family         14
Sci-Fi         13
Musical         4
War             2
Western         2
Name: count, dtype: int64
```

```python
print(df['company'].value_counts().head(10))
```

```
company
Universal Pictures                      302
Warner Bros.                            294
Paramount Pictures                      259
Twentieth Century Fox Film Corporation  205
New Line Cinema                         172
Columbia Pictures Corporation           166
Touchstone Pictures                     131
Columbia Pictures                       108
Walt Disney Pictures                    102
Metro-Goldwyn-Mayer (MGM)               101
Name: count, dtype: int64
```

```python
import matplotlib.pyplot as plt
import pandas as pd

# 1. Histograma de presupuestos
plt.subplot(221)
plt.hist(df2['budget'], color='blue', bins=20)
plt.title("Distribución de Presupuestos")
plt.xlabel("Presupuesto")
plt.ylabel("Frecuencia")
plt.subplots_adjust(left=0.1,
                    bottom=0.1,
                    right=0.9,
                    top=0.9,
                    wspace=0.4,
                    hspace=0.5)
# 2. Línea de tendencia de ganancias totales a través del tiempo
plt.subplot(222)
plt.plot(yearly["released"], yearly["budget"], color="green")
plt.title("Tendencia de Ganancias Totales por Año")
plt.xlabel("Año de Lanzamiento")
plt.ylabel("Ganancias Totales")

# 3. Barras con top 10 compañías
plt.subplot(223)
plt.bar(top10_companies.index, top10_companies.values, color="magenta")
plt.title("Top 10 Compañías por Ganancias")


plt.tight_layout()
plt.show()
```