In [ ]:
```
#########################################################
# DataFrames part III
# Author : Rodrique KAFANDO
# Destination : Master FD&IA - UV-BF
# Online : 27.07.2021
#########################################################
```

This chapter is focused on data extraction

- filter, select, reindex, etc.
- comment selectionner plusieurs colonnes
- comment indexer plusieurs cellules
- comment modifier le contenu d'une cellule
- etc.

In [1]:
```
import pandas as pd
```

In [136…
```
df = pd.read_csv('./pandas/jamesbond.csv')
df.head()
```

Out[136…

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
0	Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
1	From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
2	Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
3	Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
4	Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

The .set_index() and the .reset_index() Methods

In [5]:
```
# The set_index() permet de specifier une colonne existante comme index du DataFrame
# 1) elle peut se faire directrment lors de la lecture du fichier CSV
df = pd.read_csv('./pandas/jamesbond.csv', index_col = 'Film')
df
```

Out[5]:

	Year	Actor	Director	Box Office	Budget	Bond Actor Salary

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Film						
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Never Say Never Again	1983	Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
Casino Royale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

In [30]:
```
# 2) mais lorsque nous avons déjà les données sous forme de df,
nous devons utiliser .set_index()
```

```
df = pd.read_csv('./pandas/jamesbond.csv')
df
```

Out[30]:

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
0	Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
1	From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
2	Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
3	Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
4	Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
5	You Only Live Twice	1967	Sean Connery	Lewis Gilbert	514.2	59.9	4.4
6	On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
7	Diamonds Are Forever	1971	Sean Connery	Guy Hamilton	442.5	34.7	5.8
8	Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
9	The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
10	The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
11	Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
12	For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
13	Never Say Never Again	1983	Sean Connery	Irvin Kershner	380.0	86.0	NaN
14	Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
15	A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
16	The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
17	Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
18	GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
19	Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
20	The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
21	Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
22	Casino Royale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
23	Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
24	Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
25	Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

In [31]:
```
# specifier inplace = True pour rendre la nouvelle valeur permanente
df.set_index('Film')
```

Out[31]:

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Never Say Never Again	1983	Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
Casino Royale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

In [32]:
```
df.set_index(keys = 'Film', inplace = True)
df.head(3)
# Maintenant, comment fair epour utiliser 'Year' comme index, au lieu de 'Film'?
```

Out[32]:

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2

In [34]:
```
# comment revenir à la specification de l'index initiale? en utilisant reset_index()
df.reset_index().head()
```

Out[34]:

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
0	Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
1	From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
2	Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
3	Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
4	Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

Retrieve rows by index Label with *.loc[]* accessor

In [39]:
```
# nous revenons sur les mêmes données, et nous utilisaons 'Film'
```

```
comme colonne index
df = pd.read_csv('./pandas/jamesbond.csv')
```

In [40]:
```
df.set_index(keys = 'Film', inplace = True)
```

In [41]:
```
df.head()
```

Out[41]:

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [45]:
```
# par exemple, si nous voulons les informations sur le film dont le nom est 'Casino Royale' or Goldfinger
df.loc[['Casino Royale', 'Goldfinger']]
```

Out[45]:

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
Casino Royale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2

In [43]:
```
df.loc['Goldfinger']
```

Out[43]:
```
Year                    1964
Actor             Sean Connery
Director           Guy Hamilton
Box Office                820.4
Budget                     18.6
Bond Actor Salary           3.2
Name: Goldfinger, dtype: object
```

In [51]:
```
# comme les listes, nous pouvons aussi definir une plage
df.loc['Dr. No':'Thunderball']
```

Out[51]:

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7

Retrieve rows by index Position with *.iloc[]* accessor

In [53]:
```
# la syntax est similaire à celle de .loc[]
df.head(3)
```

Out[53]:

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2

In [55]:
```
df.reset_index(inplace=True )
```

In [56]:
```
df.head(3)
```

Out[56]:

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
0	Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
1	From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
2	Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2

In [57]:
```
# acceder la ligne à la position 5
df.iloc[5]
```

Out[57]:
```
Film                You Only Live Twice
Year                               1967
Actor                       Sean Connery
Director                   Lewis Gilbert
Box Office                         514.2
Budget                              59.9
```

```
Bond Actor Salary                4.4
Name: 5, dtype: object
```

In [58]:
```
df.iloc[0:10:2]
```

Out[58]:

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
0	Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
2	Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
4	Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
6	On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
8	Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN

In [62]:
```
df.iloc[[2,5,10,8]]
```

Out[62]:

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
2	Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
5	You Only Live Twice	1967	Sean Connery	Lewis Gilbert	514.2	59.9	4.4
10	The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
8	Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN

In [ ]:
```
# Remarques, .iloc[] est utilisable dans le cas où l'on peut utiliser .loc[]
```

In [ ]:
```

```

Utiliser des seconds arguments avecles accesseurs .loc et .iloc pour plus de flexibilité

In [63]:
```
df.head()
```

Out[63]:

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
0	Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
1	From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
2	Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
3	Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
4	Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [67]:
```
# set 'Film' as index column
df.set_index('Film', inplace  = True)
```

In [72]:
```
# nous pouvons specifier 1) le label de l'index, et la valeur de la colonne correspondante
# ex : qui est le directeur du film dont le titre est 'Thunderball'?
# df.loc['lable', 'nom colonne' ou [noms colonnes]]
df.loc['Thunderball', ['Actor','Budget']]
#df.loc[['Goldfinger','Thunderball'], ['Actor','Budget']]
```

Out[72]:
```
Actor     Sean Connery
Budget            41.9
Name: Thunderball, dtype: object
```

In [73]:
```
# autrement nous pouvons utiliser un intervalle de valeurs
df.loc[['Goldfinger'], 'Actor':'Budget']
```

Out[73]:

Film	Actor	Director	Box Office	Budget
Goldfinger	Sean Connery	Guy Hamilton	820.4	18.6

In [76]:
```
#
df.loc['Goldfinger':'The Spy Who Loved Me', 'Actor':'Budget']
```

Out[76]:

Film	Actor	Director	Box Office	Budget
Goldfinger	Sean Connery	Guy Hamilton	820.4	18.6
Thunderball	Sean Connery	Terence Young	848.1	41.9
Casino Royale	David Niven	Ken Hughes	315.0	85.0
You Only Live Twice	Sean Connery	Lewis Gilbert	514.2	59.9
On Her Majesty's Secret Service	George Lazenby	Peter R. Hunt	291.5	37.3
Diamonds Are Forever	Sean Connery	Guy Hamilton	442.5	34.7
Live and Let Die	Roger Moore	Guy Hamilton	460.3	30.8
The Man with the Golden Gun	Roger Moore	Guy Hamilton	334.0	27.7
The Spy Who Loved Me	Roger Moore	Lewis Gilbert	533.0	45.1

In [79]:
```
#.ILOC, je vous laisse vous exercer avec .iloc[]
```

```
df.iloc[3]
```

Out[79]:
```
Year                          1965
Actor                Sean Connery
Director             Terence Young
Box Office                   848.1
Budget                        41.9
Bond Actor Salary              4.7
Name: Thunderball, dtype: object
```

set the value for a specific cell

In [87]:
```
# acteur du filmn dont le nom est ?
df.loc['Thunderball', 'Actor']
```

Out[87]:
```
'Sean Connery'
```

In [88]:
```
df.loc['Thunderball', 'Actor'] = 'Mr. Sean Connery'
```

In [89]:
```
df.loc['Thunderball', 'Actor']
```

Out[89]:
```
'Mr. Sean Connery'
```

In [90]:
```
# now if i want to modify a multiple values
df.loc['Thunderball', ['Actor', 'Box Office', 'Budget']]
```

Out[90]:
```
Actor         Mr. Sean Connery
Box Office               848.1
Budget                    41.9
Name: Thunderball, dtype: object
```

In [91]:
```
# modifions ces valeurs
df.loc['Thunderball', ['Actor', 'Box Office', 'Budget']] = ['Sean
Connery',700, 50]
```

In [92]:
```
df.loc['Thunderball', ['Actor', 'Box Office', 'Budget']]
```

Out[92]:
```
Actor         Sean Connery
Box Office             700
Budget                  50
Name: Thunderball, dtype: object
```

Set multiple values in a DF

In [94]:
```
df.head()
```

Out[94]:

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	700.0	50.0	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [101…
```
# here's the wrong way to overwrite multiple values
val = df["Actor"] == 'Sean Connery'
df[val]["Actor"]
```

Out[101…
```
Film
Dr. No                    Sean Connery
From Russia with Love     Sean Connery
Goldfinger                Sean Connery
Thunderball               Sean Connery
You Only Live Twice       Sean Connery
Diamonds Are Forever      Sean Connery
Never Say Never Again     Sean Connery
Name: Actor, dtype: object
```

In [102…
```
# en affectant de cette maniere, c'est comme si nous creons une copie du DF,
# cela voudrait dire que le DF d'origine ne changera pas, verifions cela plus bas
# le message d'error nous indique les bonnes manières
df[val]["Actor"] = 'Mr. Sean Connery'
```

```
/home/rodrique/anaconda3/lib/python3.7/site-packages/ipykernel_launcher.py:1:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy
  """Entry point for launching an IPython kernel.
```

In [103…
```
df.head()
```

Out[103…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	700.0	50.0	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [107…
```
# la solution idéale serait:
val = df["Actor"] == 'Sean Connery'
# .loc utilise directement la reference, et non une copie
df.loc[val, 'Actor'] = 'Mr Sean Connery'
```

In [108…
```
df.head()
```

Out[108…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Mr Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Mr Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Mr Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Mr Sean Connery	Terence Young	700.0	50.0	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

Rename Index Labels or Columns in DataFrame

Rename() methode

In [110…
```
# use mapper() with axis => dictionnary
df.rename(mapper = {'Goldfinger':'Gold  finger',
                    'Casino Royale':'CasinoRoyale' })
```

Out[110…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Mr Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Mr Sean Connery	Terence Young	543.8	12.6	1.6

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Gold finger	1964	Mr Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Mr Sean Connery	Terence Young	700.0	50.0	4.7
CasinoRoyale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Mr Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Mr Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Never Say Never Again	1983	Mr Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
CasinoRoyale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

In [111…

```
# use mapper() with axis => dictionnary
df.rename(mapper = {'Goldfinger':'Gold  finger',
                    'Casino Royale':'CasinoRoyale'}, axis =0)
```

Out[111…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Film						
Dr. No	1962	Mr Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Mr Sean Connery	Terence Young	543.8	12.6	1.6
Gold finger	1964	Mr Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Mr Sean Connery	Terence Young	700.0	50.0	4.7
CasinoRoyale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Mr Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Mr Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Never Say Never Again	1983	Mr Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
CasinoRoyale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

In [117…

```
# use mapper() with axis => dictionnary
```

```
df.rename(mapper = {'Goldfinger':'Gold  finger',
                    'Casino Royale':'CasinoRoyale' }, axis =
'index', inplace = True)
```

In [ ]:
```
# use mapper() with axis => dictionnary
df.rename(mapper = {'Goldfinger':'Gold  finger',
                    'Casino Royale':'CasinoRoyale'}, axis = 'rows'
)
```

In [ ]:
```

```

In [ ]:
```
# Use index => same style as mapper, but don't take axis parameter
df.rename(index = {'Goldfinger':'Gold  finger',
                    'Casino Royale':'CasinoRoyale'}, inplace =
True )
```

In [118…
```
df.head()
```

Out[118…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Mr Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Mr Sean Connery	Terence Young	543.8	12.6	1.6
Gold finger	1964	Mr Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Mr Sean Connery	Terence Young	700.0	50.0	4.7
CasinoRoyale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

Rename columns name

In [119…
```
# rename with mapper(), Why nothong happens ? Because we need to
specify the axis level
df.rename(mapper = {'Actor':'Author', 'Director':'supervisor',
'Box Office':'Address' } )
```

Out[119…

Film	Year	Author	supervisor	Address	Budget	Bond Actor Salary
Dr. No	1962	Mr Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Mr Sean Connery	Terence Young	543.8	12.6	1.6

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Gold finger	1964	Mr Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Mr Sean Connery	Terence Young	700.0	50.0	4.7
CasinoRoyale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Mr Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Mr Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Never Say Never Again	1983	Mr Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
CasinoRoyale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

In [120…

```
# so we have
df.rename(mapper = {'Actor':'Author', 'Director':'supervisor',
'Box Office':'Address' }, axis = 1)
```

Out[120…

	Year	Author	supervisor	Address	Budget	Bond Actor Salary

Film	Year	Author	supervisor	Address	Budget	Bond Actor Salary
Film						
Dr. No	1962	Mr Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Mr Sean Connery	Terence Young	543.8	12.6	1.6
Gold finger	1964	Mr Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Mr Sean Connery	Terence Young	700.0	50.0	4.7
CasinoRoyale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Mr Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Mr Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Never Say Never Again	1983	Mr Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
CasinoRoyale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

In [122…

```
df.rename(mapper = {'Actor':'Author', 'Director':'supervisor',
'Box Office':'Address' }, axis = 'columns')
```

Out[122…

Film	Year	Author	supervisor	Address	Budget	Bond Actor Salary
Dr. No	1962	Mr Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Mr Sean Connery	Terence Young	543.8	12.6	1.6
Gold finger	1964	Mr Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Mr Sean Connery	Terence Young	700.0	50.0	4.7
CasinoRoyale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Mr Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Mr Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Never Say Never Again	1983	Mr Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
CasinoRoyale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

In [123…

```
# or finally use the columns parameter instead the mapper
df.rename(columns = {'Actor':'Author', 'Director':'supervisor',
```

```
'Box Office':'Address' })
```

Out[123…

Film	Year	Author	supervisor	Address	Budget	Bond Actor Salary
Dr. No	1962	Mr Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Mr Sean Connery	Terence Young	543.8	12.6	1.6
Gold finger	1964	Mr Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Mr Sean Connery	Terence Young	700.0	50.0	4.7
CasinoRoyale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Mr Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Mr Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Never Say Never Again	1983	Mr Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5	76.9	5.1
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2	133.9	10.0
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
CasinoRoyale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2	181.4	8.1
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN

rename by using the columns attrbute

In [124…
```
# be careful, the list provided must have the same len() as the number of columns
df.columns = ['Year of','Author Name','Director','Box Office','Budget TTC','Bond Actor Salary']
```

In [139…
```
df.head()
```

Out[139…

	Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
0	Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
1	From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
2	Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
3	Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
4	Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

Delete Rows or columns from DF

In [140…
```
# drop() Method, can call ddirectely on DF => rows or index, or axis  = 0
# need to specify the index value that we want to delete, ex : 'From Russia with Love', or more than one
df.drop('From Russia with Love')
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
<ipython-input-140-97ac04f4f790> in <module>
      1 # drop() Method, can call ddirectely on DF => rows or index, or axis = 0
      2 # need to specify the index value that we want to delete, ex : 'From Russia with Love', or more than one
----> 3 df.drop('From Russia with Love')

~/anaconda3/lib/python3.7/site-packages/pandas/core/frame.py in drop(self, labels, axis, index, columns, level, inplace, errors)
   4168                 level=level,
   4169                 inplace=inplace,
-> 4170                 errors=errors,
   4171             )
   4172

~/anaconda3/lib/python3.7/site-packages/pandas/core/generic.py in drop(self, labels, axis, index, columns, level, inplace, errors)
   3885             for axis, labels in axes.items():
   3886                 if labels is not None:
-> 3887                     obj = obj._drop_axis(labels, axis, level=level, errors=errors)
   3888
   3889             if inplace:

~/anaconda3/lib/python3.7/site-packages/pandas/core/generic.py in _drop_axis(self, labels, axis, level, errors)
```

```
   3919                 new_axis = axis.drop(labels, level=level, errors=erro
rs)
   3920             else:
-> 3921                 new_axis = axis.drop(labels, errors=errors)
   3922             result = self.reindex(**{axis_name: new_axis})
   3923

~/anaconda3/lib/python3.7/site-packages/pandas/core/indexes/base.py in drop(s
elf, labels, errors)
   5280         if mask.any():
   5281             if errors != "ignore":
-> 5282                 raise KeyError(f"{labels[mask]} not found in axis")
   5283             indexer = indexer[~mask]
   5284         return self.delete(indexer)

KeyError: "['From Russia with Love'] not found in axis"
```

In [132…

```
df.drop(['Author Name', 'Director'], axis = 1) # or axis =
'columns'
```

Out[132…

Film	Year of	Box Office	Budget TTC	Bond Actor Salary
Dr. No	1962	448.8	7.0	0.6
From Russia with Love	1963	543.8	12.6	1.6
Gold finger	1964	820.4	18.6	3.2
Thunderball	1965	700.0	50.0	4.7
CasinoRoyale	1967	315.0	85.0	NaN
You Only Live Twice	1967	514.2	59.9	4.4
On Her Majesty's Secret Service	1969	291.5	37.3	0.6
Diamonds Are Forever	1971	442.5	34.7	5.8
Live and Let Die	1973	460.3	30.8	NaN
The Man with the Golden Gun	1974	334.0	27.7	NaN
The Spy Who Loved Me	1977	533.0	45.1	NaN
Moonraker	1979	535.0	91.5	NaN
For Your Eyes Only	1981	449.4	60.2	NaN
Never Say Never Again	1983	380.0	86.0	NaN
Octopussy	1983	373.8	53.9	7.8
A View to a Kill	1985	275.2	54.5	9.1
The Living Daylights	1987	313.5	68.8	5.2
Licence to Kill	1989	250.9	56.7	7.9
GoldenEye	1995	518.5	76.9	5.1
Tomorrow Never Dies	1997	463.2	133.9	10.0
The World Is Not Enough	1999	439.5	158.3	13.5
Die Another Day	2002	465.4	154.2	17.9
CasinoRoyale	2006	581.5	145.3	3.3
Quantum of Solace	2008	514.2	181.4	8.1
Skyfall	2012	943.5	170.2	14.5

Film	Year of	Box Office	Budget TTC	Bond Actor Salary
Spectre	2015	726.7	206.3	NaN

In [137…

```
# pop() method, which is permanent method, and returns the poped
series. It's takes a single series
actor = df.pop('Author Name')
```

```
---------------------------------------------------------------------------
KeyError                                  Traceback (most recent call last)
~/anaconda3/lib/python3.7/site-packages/pandas/core/indexes/base.py in get_lo
c(self, key, method, tolerance)
   2894             try:
-> 2895                 return self._engine.get_loc(casted_key)
   2896             except KeyError as err:

pandas/_libs/index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas/_libs/index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas/_libs/hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHas
hTable.get_item()

pandas/_libs/hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHas
hTable.get_item()

KeyError: 'Author Name'

The above exception was the direct cause of the following exception:

KeyError                                  Traceback (most recent call last)
<ipython-input-137-79cf67175810> in <module>
      1 # pop() method, which is permanent method, and returns the poped valu
e. It's takes a single
----> 2 actor = df.pop('Author Name')

~/anaconda3/lib/python3.7/site-packages/pandas/core/frame.py in pop(self, ite
m)
   4365         3  monkey        NaN
   4366         """
-> 4367         return super().pop(item=item)
   4368
   4369     @doc(NDFrame.replace, **_shared_doc_kwargs)

~/anaconda3/lib/python3.7/site-packages/pandas/core/generic.py in pop(self, i
tem)
    659
    660     def pop(self, item: Label) -> Union["Series", Any]:
--> 661         result = self[item]
    662         del self[item]
    663         if self.ndim == 2:

~/anaconda3/lib/python3.7/site-packages/pandas/core/frame.py in __getitem__(s
elf, key)
   2900             if self.columns.nlevels > 1:
   2901                 return self._getitem_multilevel(key)
-> 2902             indexer = self.columns.get_loc(key)
   2903             if is_integer(indexer):
   2904                 indexer = [indexer]

~/anaconda3/lib/python3.7/site-packages/pandas/core/indexes/base.py in get_lo
c(self, key, method, tolerance)
   2895                 return self._engine.get_loc(casted_key)
   2896             except KeyError as err:
```

```
-> 2897                 raise KeyError(key) from err
   2898
   2899         if tolerance is not None:

KeyError: 'Author Name'
```

In [144…

```
# we also have the del keyword, ex :
del df['Actor']
```

Create random sample with sample Method

In [146…

```
df = pd.read_csv('./pandas/jamesbond.csv', index_col = 'Film')
df.head()
```

Out[146…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [155…

```
# the sample() method
df.sampldf.sample()e() # return a given sample
```

Out[155…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN

In [156…

```
df.sample(n = 4) # sepecify N, the number of rows to be extracted randomly
```

Out[156…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
You Only Live Twice	1967	Sean Connery	Lewis Gilbert	514.2	59.9	4.4
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Never Say Never Again	1983	Sean Connery	Irvin Kershner	380.0	86.0	NaN

In [162…
```
df.sample(frac = .50) # will return 25% of the rows, randomly
```

Out[162…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4	154.2	17.9
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
Diamonds Are Forever	1971	Sean Connery	Guy Hamilton	442.5	34.7	5.8
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5	158.3	13.5
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Casino Royale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3

In [163…
```
# specify axis value to select columns randomly
df.sample(frac = .50, axis  = 1)
```

Out[163…

Film	Budget	Director	Bond Actor Salary
Dr. No	7.0	Terence Young	0.6
From Russia with Love	12.6	Terence Young	1.6
Goldfinger	18.6	Guy Hamilton	3.2
Thunderball	41.9	Terence Young	4.7
Casino Royale	85.0	Ken Hughes	NaN
You Only Live Twice	59.9	Lewis Gilbert	4.4
On Her Majesty's Secret Service	37.3	Peter R. Hunt	0.6
Diamonds Are Forever	34.7	Guy Hamilton	5.8
Live and Let Die	30.8	Guy Hamilton	NaN
The Man with the Golden Gun	27.7	Guy Hamilton	NaN
The Spy Who Loved Me	45.1	Lewis Gilbert	NaN
Moonraker	91.5	Lewis Gilbert	NaN
For Your Eyes Only	60.2	John Glen	NaN

Film	Budget	Director	Bond Actor Salary
Never Say Never Again	86.0	Irvin Kershner	NaN
Octopussy	53.9	John Glen	7.8
A View to a Kill	54.5	John Glen	9.1
The Living Daylights	68.8	John Glen	5.2
Licence to Kill	56.7	John Glen	7.9
GoldenEye	76.9	Martin Campbell	5.1
Tomorrow Never Dies	133.9	Roger Spottiswoode	10.0
The World Is Not Enough	158.3	Michael Apted	13.5
Die Another Day	154.2	Lee Tamahori	17.9
Casino Royale	145.3	Martin Campbell	3.3
Quantum of Solace	181.4	Marc Forster	8.1
Skyfall	170.2	Sam Mendes	14.5
Spectre	206.3	Sam Mendes	NaN

The .nsmallest() and .nlargest() columns or rows in a DF

In [170…

```
# elle ppeut etre souvent bien efficcient par rapport à la method sort_values()
# ex:
df.sort_values('Box Office', ascending = False).head()
```

Out[170…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN
Casino Royale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3

In [174…

```
#  .nlargest() nous permet d'extraire les top N plus grandes valeurs
df.nlargest(n = 5, columns = ['Box Office','Budget'])
```

Out[174…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN
Casino Royale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3

In [172…
```
df.nsmallest(5,'Box Office')
```

Out[172…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Licence to Kill	1989	Timothy Dalton	John Glen	250.9	56.7	7.9
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5	37.3	0.6
The Living Daylights	1987	Timothy Dalton	John Glen	313.5	68.8	5.2
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [175…
```
# can also be called directly on a series
df["Budget"].nlargest(4)
```

Out[175…
```
Film
Spectre                     206.3
Quantum of Solace           181.4
Skyfall                     170.2
The World Is Not Enough     158.3
Name: Budget, dtype: float64
```

In [ ]:
```
df.nlargest()
```

In [178…
```
df.nlargest( 5, ['Box Office','Budget'])
```

Out[178…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Skyfall	2012	Daniel Craig	Sam Mendes	943.5	170.2	14.5
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Spectre	2015	Daniel Craig	Sam Mendes	726.7	206.3	NaN
Casino Royale	2006	Daniel Craig	Martin Campbell	581.5	145.3	3.3

Filtering DF with where() method

In [179…

```
df.head()
```

Out[179…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [183…

```
# Where return un DF avec des valeurs reelles pour les lignes qui s'accordent à la requete,
# le reste sera replacé par des Nan
#########################"
# commençons par un filtre simple
v_ = df['Actor'] == 'Sean Connery'
df[v_]
```

Out[183…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
You Only Live Twice	1967	Sean Connery	Lewis Gilbert	514.2	59.9	4.4
Diamonds Are Forever	1971	Sean Connery	Guy Hamilton	442.5	34.7	5.8
Never Say Never Again	1983	Sean Connery	Irvin Kershner	380.0	86.0	NaN

In [187…

```
# avec where
df.where(v_)
```

Out[187…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962.0	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963.0	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964.0	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965.0	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	NaN	NaN	NaN	NaN	NaN	NaN
You Only Live Twice	1967.0	Sean Connery	Lewis Gilbert	514.2	59.9	4.4
On Her Majesty's Secret Service	NaN	NaN	NaN	NaN	NaN	NaN
Diamonds Are Forever	1971.0	Sean Connery	Guy Hamilton	442.5	34.7	5.8
Live and Let Die	NaN	NaN	NaN	NaN	NaN	NaN
The Man with the Golden Gun	NaN	NaN	NaN	NaN	NaN	NaN
The Spy Who Loved Me	NaN	NaN	NaN	NaN	NaN	NaN
Moonraker	NaN	NaN	NaN	NaN	NaN	NaN
For Your Eyes Only	NaN	NaN	NaN	NaN	NaN	NaN
Never Say Never Again	1983.0	Sean Connery	Irvin Kershner	380.0	86.0	NaN
Octopussy	NaN	NaN	NaN	NaN	NaN	NaN
A View to a Kill	NaN	NaN	NaN	NaN	NaN	NaN
The Living Daylights	NaN	NaN	NaN	NaN	NaN	NaN
Licence to Kill	NaN	NaN	NaN	NaN	NaN	NaN
GoldenEye	NaN	NaN	NaN	NaN	NaN	NaN
Tomorrow Never Dies	NaN	NaN	NaN	NaN	NaN	NaN
The World Is Not Enough	NaN	NaN	NaN	NaN	NaN	NaN
Die Another Day	NaN	NaN	NaN	NaN	NaN	NaN
Casino Royale	NaN	NaN	NaN	NaN	NaN	NaN
Quantum of Solace	NaN	NaN	NaN	NaN	NaN	NaN
Skyfall	NaN	NaN	NaN	NaN	NaN	NaN
Spectre	NaN	NaN	NaN	NaN	NaN	NaN

In [ ]:
```
# nous pouvons specifier une valeur à al place de Nan
df.where(v_, other = 'forget')
```

Filter a DF with the .query() method

In [ ]:
```
# .query() method works with columns names that do not have spaces
# we first need to replace spaces with '_' par example
```

In [194…
```
# the fastest way's
df.columns
```

Out[194…
```
Index(['Year', 'Actor', 'Director', 'Box_Office', 'Budget',
       'Bond_Actor_Salary'],
      dtype='object')
```

In [196…
```
df.columns = [col.replace(' ', '_') for col in df.columns]
# or
# df.columns = ['Year','Actor', 'Director',
'Box_Office','Budget','Bond_Actor_Salary']
```

In [199…
```
df.head()
```

Out[199…

Film	Year	Actor	Director	Box_Office	Budget	Bond_Actor_Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [203…
```
# let's use .query() method
df.query('Actor == "Roger Moore"')
```

Out[203…

Film	Year	Actor	Director	Box_Office	Budget	Bond_Actor_Salary
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3	30.8	NaN
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0	27.7	NaN
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0	45.1	NaN
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0	91.5	NaN
For Your Eyes Only	1981	Roger Moore	John Glen	449.4	60.2	NaN
Octopussy	1983	Roger Moore	John Glen	373.8	53.9	7.8
A View to a Kill	1985	Roger Moore	John Glen	275.2	54.5	9.1

In [205…
```
# or
df.query('Actor != "Roger Moore"').head()
```

Out[205…

Film	Year	Actor	Director	Box_Office	Budget	Bond_Actor_Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [212…
```
df.query('Actor == "Sean Connery" and Director == "Terence Young"
')
```

Out[212…

Film	Year	Actor	Director	Box_Office	Budget	Bond_Actor_Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7

In [225…
```
df.query('Actor in ["Sean Connery", "David Niven"]')
```

Out[225…

Film	Year	Actor	Director	Box_Office	Budget	Bond_Actor_Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN
You Only Live Twice	1967	Sean Connery	Lewis Gilbert	514.2	59.9	4.4
Diamonds Are Forever	1971	Sean Connery	Guy Hamilton	442.5	34.7	5.8
Never Say Never Again	1983	Sean Connery	Irvin Kershner	380.0	86.0	NaN

Quick review on .apply() method

In [226…
```
df.head()
```

Out[226…

Film	Year	Actor	Director	Box_Office	Budget	Bond_Actor_Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

.apply() on a single series

In [220…
```
# ex
def convert_to_string_and_add_million(number):
    return str(number)+ 'MILLIONS!'
```

In [223…
```
df['Box_Office'].apply(convert_to_string_and_add_million)
```

Out[223…
```
Film
Dr. No                              448.8MILLIONS!
From Russia with Love               543.8MILLIONS!
Goldfinger                          820.4MILLIONS!
Thunderball                         848.1MILLIONS!
Casino Royale                       315.0MILLIONS!
You Only Live Twice                 514.2MILLIONS!
On Her Majesty's Secret Service     291.5MILLIONS!
Diamonds Are Forever                442.5MILLIONS!
Live and Let Die                    460.3MILLIONS!
The Man with the Golden Gun         334.0MILLIONS!
The Spy Who Loved Me                533.0MILLIONS!
Moonraker                           535.0MILLIONS!
For Your Eyes Only                  449.4MILLIONS!
Never Say Never Again               380.0MILLIONS!
Octopussy                           373.8MILLIONS!
A View to a Kill                    275.2MILLIONS!
The Living Daylights                313.5MILLIONS!
Licence to Kill                     250.9MILLIONS!
GoldenEye                           518.5MILLIONS!
Tomorrow Never Dies                 463.2MILLIONS!
The World Is Not Enough             439.5MILLIONS!
Die Another Day                     465.4MILLIONS!
Casino Royale                       581.5MILLIONS!
Quantum of Solace                   514.2MILLIONS!
Skyfall                             943.5MILLIONS!
Spectre                             726.7MILLIONS!
Name: Box_Office, dtype: object
```

In [224…
```
# identically , we can do it with the Budget column
```

```
df['Budget'].apply(convert_to_string_and_add_million)
```

Out[224…

```
Film
Dr. No                                    7.0MILLIONS!
From Russia with Love                    12.6MILLIONS!
Goldfinger                               18.6MILLIONS!
Thunderball                              41.9MILLIONS!
Casino Royale                            85.0MILLIONS!
You Only Live Twice                      59.9MILLIONS!
On Her Majesty's Secret Service          37.3MILLIONS!
Diamonds Are Forever                     34.7MILLIONS!
Live and Let Die                         30.8MILLIONS!
The Man with the Golden Gun              27.7MILLIONS!
The Spy Who Loved Me                     45.1MILLIONS!
Moonraker                                91.5MILLIONS!
For Your Eyes Only                       60.2MILLIONS!
Never Say Never Again                    86.0MILLIONS!
Octopussy                                53.9MILLIONS!
A View to a Kill                         54.5MILLIONS!
The Living Daylights                     68.8MILLIONS!
Licence to Kill                          56.7MILLIONS!
GoldenEye                                76.9MILLIONS!
Tomorrow Never Dies                     133.9MILLIONS!
The World Is Not Enough                 158.3MILLIONS!
Die Another Day                         154.2MILLIONS!
Casino Royale                           145.3MILLIONS!
Quantum of Solace                       181.4MILLIONS!
Skyfall                                 170.2MILLIONS!
Spectre                                 206.3MILLIONS!
Name: Budget, dtype: object
```

In [230…

```
# now, let's use this function more nicely
# instead of doing this operation column by column, we can use a simple loop as:
columns = ['Box_Office', 'Budget','Bond_Actor_Salary' ]
for tmp in columns:
    df[tmp] = df[tmp].apply(convert_to_string_and_add_million)
```

In [232…

```
df.head(30)
```

Out[232…

Film	Year	Actor	Director	Box_Office	Budget	Bond_Actor_Salary
Dr. No	1962	Sean Connery	Terence Young	448.8MILLIONS!	7.0MILLIONS!	0.6MILLIONS!
From Russia with Love	1963	Sean Connery	Terence Young	543.8MILLIONS!	12.6MILLIONS!	1.6MILLIONS!
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4MILLIONS!	18.6MILLIONS!	3.2MILLIONS!
Thunderball	1965	Sean Connery	Terence Young	848.1MILLIONS!	41.9MILLIONS!	4.7MILLIONS!
Casino Royale	1967	David Niven	Ken Hughes	315.0MILLIONS!	85.0MILLIONS!	nanMILLIONS!
You Only Live Twice	1967	Sean Connery	Lewis Gilbert	514.2MILLIONS!	59.9MILLIONS!	4.4MILLIONS!

Film	Year	Actor	Director	Box_Office	Budget	Bond_Actor_Salary
On Her Majesty's Secret Service	1969	George Lazenby	Peter R. Hunt	291.5MILLIONS!	37.3MILLIONS!	0.6MILLIONS!
Diamonds Are Forever	1971	Sean Connery	Guy Hamilton	442.5MILLIONS!	34.7MILLIONS!	5.8MILLIONS!
Live and Let Die	1973	Roger Moore	Guy Hamilton	460.3MILLIONS!	30.8MILLIONS!	nanMILLIONS!
The Man with the Golden Gun	1974	Roger Moore	Guy Hamilton	334.0MILLIONS!	27.7MILLIONS!	nanMILLIONS!
The Spy Who Loved Me	1977	Roger Moore	Lewis Gilbert	533.0MILLIONS!	45.1MILLIONS!	nanMILLIONS!
Moonraker	1979	Roger Moore	Lewis Gilbert	535.0MILLIONS!	91.5MILLIONS!	nanMILLIONS!
For Your Eyes Only	1981	Roger Moore	John Glen	449.4MILLIONS!	60.2MILLIONS!	nanMILLIONS!
Never Say Never Again	1983	Sean Connery	Irvin Kershner	380.0MILLIONS!	86.0MILLIONS!	nanMILLIONS!
Octopussy	1983	Roger Moore	John Glen	373.8MILLIONS!	53.9MILLIONS!	7.8MILLIONS!
A View to a Kill	1985	Roger Moore	John Glen	275.2MILLIONS!	54.5MILLIONS!	9.1MILLIONS!
The Living Daylights	1987	Timothy Dalton	John Glen	313.5MILLIONS!	68.8MILLIONS!	5.2MILLIONS!
Licence to Kill	1989	Timothy Dalton	John Glen	250.9MILLIONS!	56.7MILLIONS!	7.9MILLIONS!
GoldenEye	1995	Pierce Brosnan	Martin Campbell	518.5MILLIONS!	76.9MILLIONS!	5.1MILLIONS!
Tomorrow Never Dies	1997	Pierce Brosnan	Roger Spottiswoode	463.2MILLIONS!	133.9MILLIONS!	10.0MILLIONS!
The World Is Not Enough	1999	Pierce Brosnan	Michael Apted	439.5MILLIONS!	158.3MILLIONS!	13.5MILLIONS!
Die Another Day	2002	Pierce Brosnan	Lee Tamahori	465.4MILLIONS!	154.2MILLIONS!	17.9MILLIONS!
Casino Royale	2006	Daniel Craig	Martin Campbell	581.5MILLIONS!	145.3MILLIONS!	3.3MILLIONS!
Quantum of Solace	2008	Daniel Craig	Marc Forster	514.2MILLIONS!	181.4MILLIONS!	8.1MILLIONS!
Skyfall	2012	Daniel Craig	Sam Mendes	943.5MILLIONS!	170.2MILLIONS!	14.5MILLIONS!
Spectre	2015	Daniel Craig	Sam Mendes	726.7MILLIONS!	206.3MILLIONS!	nanMILLIONS!

In [ ]:

The .apply() method with rows values

In [233…
```
df = pd.read_csv('./pandas/jamesbond.csv', index_col = 'Film')
df.head()
```

Out[233…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [ ]:
```

```

In [235…
```
# cette fois ci nous utilisons apply() sur les rows du DF
# ex : si nous voulons attribuer automatiquement un review sur chacun des films, en function du nom de l'acteur
# et le budget de realisation
def movie_review(row):
    actor, budget = row[1], row[4]
    if actor  == "Sean Connery":
        return 'excellent'
    elif actor == 'Roger Moore' and budget >= 40:
        return 'moyen'
    else:
        return 'aucune idea'
```

In [239…
```
df.apply(movie_review, axis = 1) # or axis = 'columns'
```

In [ ]:
```
# we can store this data in new col
df['review'] = df.apply(movie_review, axis = 1)
```

In [240…
```
df.head()
```

Out[240…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary	review
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6	excellent

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary	review
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6	excellent
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2	excellent
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7	excellent
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN	aucune idea

Create a DF copy with .copy() Method

- cree une copie du DF et le sauvegarde en memo

In [241…
```
df = pd.read_csv('./pandas/jamesbond.csv', index_col = 'Film')
df.head()
```

Out[241…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [242…
```
# verifions une premiere operation
df['Director']
```

Out[242…
```
Film
Dr. No                                 Terence Young
From Russia with Love                  Terence Young
Goldfinger                              Guy Hamilton
Thunderball                            Terence Young
Casino Royale                             Ken Hughes
You Only Live Twice                    Lewis Gilbert
On Her Majesty's Secret Service        Peter R. Hunt
Diamonds Are Forever                    Guy Hamilton
Live and Let Die                        Guy Hamilton
The Man with the Golden Gun             Guy Hamilton
The Spy Who Loved Me                   Lewis Gilbert
Moonraker                              Lewis Gilbert
For Your Eyes Only                         John Glen
Never Say Never Again                 Irvin Kershner
Octopussy                                  John Glen
A View to a Kill                           John Glen
The Living Daylights                       John Glen
```

```
Licence to Kill                      John Glen
GoldenEye                         Martin Campbell
Tomorrow Never Dies              Roger Spottiswoode
The World Is Not Enough               Michael Apted
Die Another Day                        Lee Tamahori
Casino Royale                     Martin Campbell
Quantum of Solace                    Marc Forster
Skyfall                                Sam Mendes
Spectre                                Sam Mendes
Name: Director, dtype: object
```

In [243…

```
# sauvegardons le dans une variable
direct = df['Director']
direct.head()
```

Out[243…

```
Film
Dr. No                    Terence Young
From Russia with Love     Terence Young
Goldfinger                 Guy Hamilton
Thunderball               Terence Young
Casino Royale                Ken Hughes
Name: Director, dtype: object
```

In [245…

```
# essayons de modifier le nom d'un directeur
direct['From Russia with Love'] = 'Mr Terence Young'
```

```
/home/rodrique/anaconda3/lib/python3.7/site-packages/ipykernel_launcher.py:2:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/s
table/user_guide/indexing.html#returning-a-view-versus-a-copy
```

In [246…

```
# le problm est que nous essayons de remplacer une valeur du df
d'origine, et nous avons donc cet avertissement
# nous pouvons verifier que le director di=u film en question a
ete effectivement modifié dans la base de donnée
df.head()
```

Out[246…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Mr Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [247…

```
# alors pour eviter cela, nous devos creer une copy() de la valeur
en cours que nous souhaitons modifier
```

```
#
df = pd.read_csv('./pandas/jamesbond.csv', index_col = 'Film')
df.head()
```

Out[247…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [248…

```
direct = df['Director'].copy()
direct.head()
```

Out[248…

```
Film
Dr. No                    Terence Young
From Russia with Love     Terence Young
Goldfinger                 Guy Hamilton
Thunderball               Terence Young
Casino Royale                Ken Hughes
Name: Director, dtype: object
```

In [249…

```
# nous pouvons maintenant modifier cette valeur tranquillement
direct['From Russia with Love'] = 'Mr Terence Young'
```

In [251…

```
direct.head()
```

Out[251…

```
Film
Dr. No                       Terence Young
From Russia with Love    Mr Terence Young
Goldfinger                    Guy Hamilton
Thunderball                  Terence Young
Casino Royale                   Ken Hughes
Name: Director, dtype: object
```

In [253…

```
# la valeur dans le DF initial ne change pas
df.head()
```

Out[253…

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Dr. No	1962	Sean Connery	Terence Young	448.8	7.0	0.6
From Russia with Love	1963	Sean Connery	Terence Young	543.8	12.6	1.6

Film	Year	Actor	Director	Box Office	Budget	Bond Actor Salary
Goldfinger	1964	Sean Connery	Guy Hamilton	820.4	18.6	3.2
Thunderball	1965	Sean Connery	Terence Young	848.1	41.9	4.7
Casino Royale	1967	David Niven	Ken Hughes	315.0	85.0	NaN

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

Bonus

- https://www.datasciencecentral.com/profiles/blogs/python-pandas-an-in-depth-tool-for-data-analytics#ToobaAhmedAlvi44319

In [ ]: