Manip 3: pratique pandas

In [4]:
```
#Pandas
import pandas as pd

#1- créer le dataframe

data = {'prenom': ['Alain', 'Abdel', 'Annie', 'Toto', 'Momo'],
        'nom': ['Flouflou', 'Clairclair', "?", 'Icano', 'Jolie'],
        'age': [42, 52, 36, 24, 73],
        'noteintra': [48, 84, 91, "?", "?"],
        'notefinal': ["92,50", "94,00", 77, 63, 75]}
```

In [5]:
```
df = pd.DataFrame(data, columns = ['prenom', 'nom', 'age', 'noteintra', 'notefina

#2. save dataframe as csv
df.to_csv('etudiants.csv')
```

In [6]:
```
#3. Charger csv
df = pd.read_csv('etudiants.csv')
df
```

Out[6]:

	Unnamed: 0	prenom	nom	age	noteintra	notefinal
0	0	Alain	Flouflou	42	48	92,50
1	1	Abdel	Clairclair	52	84	94,00
2	2	Annie	?	36	91	77
3	3	Toto	Icano	24	?	63
4	4	Momo	Jolie	73	?	75

In [7]:
```
#4. Charger csv  sans les headers

df = pd.read_csv('etudiants.csv', header=None)
df
```

Out[7]:

	0	1	2	3	4	5
0	NaN	prenom	nom	age	noteintra	notefinal
1	0.0	Alain	Flouflou	42	48	92,50
2	1.0	Abdel	Clairclair	52	84	94,00
3	2.0	Annie	?	36	91	77
4	3.0	Toto	Icano	24	?	63
5	4.0	Momo	Jolie	73	?	75

In [8]:
```
#5. Charger csv avec specification de noms de colonnes

df = pd.read_csv('etudiants.csv', names=['UID', 'Prenom', 'Nom', 'Age', 'Note int
df
```

Out[8]:

	UID	Prenom	Nom	Age	Note intra	Note final
0	NaN	prenom	nom	age	noteintra	notefinal
1	0.0	Alain	Flouflou	42	48	92,50
2	1.0	Abdel	Clairclair	52	84	94,00
3	2.0	Annie	?	36	91	77
4	3.0	Toto	Icano	24	?	63
5	4.0	Momo	Jolie	73	?	75

In [10]:
```
#6. Charger csv en specifiant "?" comme valeurs manquantes

df = pd.read_csv('etudiants.csv', na_values=['?'])
pd.isnull(df)
df
```

Out[10]:

	Unnamed: 0	prenom	nom	age	noteintra	notefinal
0	0	Alain	Flouflou	42	48.0	92,50
1	1	Abdel	Clairclair	52	84.0	94,00
2	2	Annie	NaN	36	91.0	77
3	3	Toto	Icano	24	NaN	63
4	4	Momo	Jolie	73	NaN	75

In [12]:
```
#7. Charger csv avec skip des 3 premieres rows

df = pd.read_csv('etudiants.csv', na_values=['?'], skiprows=3)
df
```

Out[12]:

	2	Annie	?	36	91	77
0	3	Toto	Icano	24	NaN	63
1	4	Momo	Jolie	73	NaN	75

In [16]:
```
#8.Ajouter une colonnes

df['devoir'] =pd.Series([82.5, 92.5, 92.0, 99.0, 90.5])
df
```

Out[16]:

	2	Annie	?	36	91	77	devoir
0	3	Toto	Icano	24	NaN	63	82.5
1	4	Momo	Jolie	73	NaN	75	92.5

In [ ]: