Manipulation des données avec Pandas

1. Importer les packages Numpy et Pandas

```
import pandas as pd
import numpy as np
```

2. Création d'un DataFrame ayant une seule ligne de données

```
props=['age',"nb d'enfants"] # labels de colonnes
noms=['Ali'] # labels de lignes
X=np.array([[30,2]]) # matrice des valeurs
df=pd.DataFrame(X,columns=props,index=noms) # dataframe
df

age nb d'enfants
Ali 30 2
```

3. Création d'un DataFrame ayant plusieurs lignes de données

```
props=['age',"nb d'enfants"]
noms=['Ali','Sonia','Salma']
X=np.array([[30,1],[35,2],[40,3]])
df=pd.DataFrame(X,columns=props,index=noms)
df

age nb d'enfants
Ali 30 1
Sonia 35 2
Salma 40 3
```

4. Récupération d'une colonne

```
df['age']
Ali
         30
Sonia
         35
Salma
         40
Name: age, dtype: int32
df.age
Ali
         30
Sonia
         35
Salma
         40
Name: age, dtype: int32
```

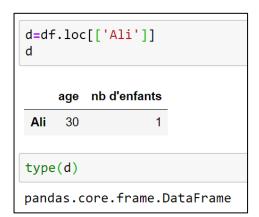
5. Récupération d'une ligne sous forme d'une Série

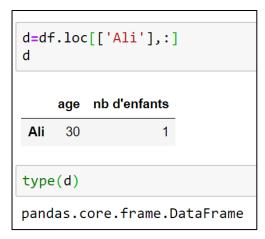
```
d=df.loc['Ali']
d

age          30
nb d'enfants    1
Name: Ali, dtype: int32

type(d)
pandas.core.series.Series
```

6. Récupération d'une ligne sous forme d'une DataFrame





7. Concaténation de 2 DataFrames

```
# concatenation de 2 dataframes
props=['age',"nb d'enfants"]
noms1=['Ali','Sonia','Salma']
X1=np.array([[30,1],[35,2],[40,3]])
df1=pd.DataFrame(X1,columns=props,index=noms1)
noms2=['Wael','Fadwa']
X2=np.array([[45,2],[55,4]])
df2=pd.DataFrame(X2,columns=props,index=noms2)
df=pd.concat([df1,df2])
       age nb d'enfants
   Ali
       30
                    2
        35
 Sonia
        40
                    3
 Salma
       45
                    2
 Wael
                    4
Fadwa
       55
```

8. Slicing de DataFrame:

```
# slicing dataframe
df.loc[['Ali','Salma'],:]

age nb d'enfants

Ali 30 1

Salma 40 3
```

9. Modification de la matrice de DataFrame :

```
# modifier une partie de la matrice de dataframe
df.loc[['Ali', 'Salma'],:]=np.array([[60,5],[65,5]])
       age nb d'enfants
   Ali
        60
                    5
                    2
 Sonia
        35
 Salma
        65
                    5
 Wael
        45
                    2
Fadwa
        55
                     4
```

10. Opérations mathématiques (Exemple produit scalaire)

```
# produit scalaire
dfa=df.loc['Ali',:]
dfs=df.loc['Salma',:]
np.dot(dfa,dfs)

3925

x1=np.array([60,5])
x2=np.array([65,5])
np.dot(x1,x2)

3925
```