

# UTILISATION KAGGLE

## Importation et creation de donnees

### Creation de donnees

Les dataframes :

```
donnees = pd.DataFrame({"column" : ["donnee1","donnee2"],"column2" :  
["donnee2 - 1","donnee2 - 2"],...}
```

Les series :

```
donneest = pd.Series(["donnee1","donnee2","donnee1"],index = ["index1","index2","index3"],name =  
"nomSerie")s
```

### Importation de donnees

CSV:

```
GET : pd.read_csv("../input/fichier/nom.csv",index_col = 0)
```

```
PUT : pd.DataFrame.to_csv(path_or_buf = "../nom.csv",self = nom)
```

SQL:

## Manipulation des structures de donnees

Code	Description	Exemple
<code>.rename(columns={ 'nom_column' : 'nv_nom', ...})</code>	changer le nom des colonnes	...
<code>.dropna()</code>	remove all the rows that contain a missing value	...
<code>.arange(n)</code>	...	...
<code>.resize((n,m))</code>	...	...
<code>.copy()</code>	...	copy original
<code>.reset_index()</code>	...	copy original

## Information sur les structures de donnees

Code	Description	Exemple
<code>.shape</code>	nombre de ligne et de colonnes	...
<code>.columns</code>	...	...
<code>.info()</code>	information generale sur la dataframe. Type de donne : bool, int64, float64 and object	...
<code>.describe()</code>	multiple info : nombre de valeurs non manquantes, moyenne, ecart type, etendue, mdiane, quartiles et 0,25	...
<code>.describe(include= ['object', 'bool'])</code>	statistiques sur les caractéristique non numerique ( indiquer explicitement les types de donnees d'intrt dans le parametre include )	...
...	...	...

# Selection et manipulation des donnees

## Selection des donnees

Code	Description	Exemple
donnees.nom_column	Data frame de la colonne en question	...
.nom_column.iloc[n]	n ieme valeur de la colonne nom_column	...
donnees.iloc[i,j]	valeur colonne i et ligne j(exclu)	...
(n).iloc[i,j]	[ :n] (tout jusqu'a n) et iloc	...
.groupby()	grouper les donnees selon un critere nom	...
.sample(donnees)	...	...

Condition:

==, <, >, <=, >= comparaison de valeur avec donnee selectionnee

Reference logique

~ pour le not; | pour le or; & pour le and

## Informations sur les donnees

Donnee seul :

Code	Description	Exemple
.dtype	...	...
.astype('type')	conversion de la donee	...
.isnull()	...	...
.fillna('nom')	...	...
...	...	...

Ensemble de Donnee :

Code	Description	Exemple
.count()	...	...
.sum()	...	...
.max()	donnee max	...
.min()	donnee min	...
.mean()	moyenne	...
.std()	...	...
.argmax()	index donnee max	...
.argmin()	index donnee min	...
stat.mods	...	...
...	...	...
...	...	...

## Manipulation des donnees

Code	Description	Exemple
<code>data.unique()</code>	supprime les doublons	...

## Autre

Code	Description	Exemple
<code>data.</code>	...	...

# Graphiques

## Simple

## Complexes

## Autres