Python et panda

Ressources :

[data (free.fr)](http://maths.bro.free.fr/data/data.html) (bug sur firefox)

login : livre

mot de passe :Python\_bro\_remy

## Jupyter Notebook

Format **Notebook :** ipynb permet d’avoir code et sa sortie avec visualisations, équations, dans un seul document

***jupyter notebook*** est un **Environnement de calcul interactif.**

Fourni avec anaconda . qui contaient toutes les bib populaires et permet de travailler en local

* Jupyter notebook peut s’ouvrir avec anaconda navigator -> localhost
* Le dossier d’écriture est dans « utilisateurs » ou « justo »

Rappel Python

fichier .py importé dans un autre script.

**Pylab :** tracer une liste en fonction d’une autre

**Random**: loi uniforme bernoulli, binomiale

**Numpy :**  tableaux, matrices

**Scipy :** calcul probabilités selon loi

Modules complémentaire.

Gestionnaire de modules (packages) : **Conda**

Terminal (base) : python accessible depuis le powershell de anaconda browser

conda install -c anaconda seaborn

pour ouvrir jupyter notebook on peut le selectionner depuis anaconda navigator> home > jupyter notebook ou depuis cmd base (« environnement python »)

pour installer panda depuis cmd si py déjà installé -> pip install panda

T = pandas.csv\_read(‘matable.csv’) charge le fichier matable.csv dans la variable T (qui sera de type DataFrame)

La fonction Head(i) retourne un dataframe avec un i lignes

Un [dataframe](https://www.w3schools.com/python/pandas/pandas_dataframes.asp) est une structure de données à deux dimension (matrice)

Quelques méthodes de dataframe :

* T[‘macolonne1’] pour voir les index et valeurs de macolonne1
* T.loc[4, ‘macol1’] renvoie la valeur de la 5eme ligne de macol1
* T.query(‘macol1 == “valeur” and E.T

[**Tutoriel pandas linogaliana**](https://linogaliana-teaching.netlify.app/pandastp/)

le package [pynsee](https://github.com/InseeFrLab/Py-Insee-Data) qui simplifie l’accès aux données de l’Insee disponibles sur le site web [insee.fr](https://www.insee.fr/fr/accueil) ou via des API

Fonctions utilies :

dataF.sort\_values(by= ‘’colone2”)