

population_recuperation_donnees

July 2, 2023

1 Récupération des données

Ce notebook donne quelques exemples de codes qui permettent de récupérer les données utilisées par d'autres notebooks. Le module [actuariat_python](#) est implémenté avec Python 3. Pour les utilisateurs de Python 2.7, il suffira de recopier le code chaque fonction dans le notebook (suivre les liens insérés dans le notebook).

```
[1]: %matplotlib inline
import matplotlib.pyplot as plt
plt.style.use('ggplot')
# le code qui suit n'est pas indispensable, il génère automatiquement un menu
# dans le notebook
from jyquickhelper import add_notebook_menu
add_notebook_menu()
```

Populating the interactive namespace from numpy and matplotlib

```
[1]: <IPython.core.display.HTML object>
```

1.0.1 Population française janvier 2017

Les données sont disponibles sur le site de l'INSEE [Pyramide des âges au 1er janvier](#). Elles sont disponibles au format Excel. Le format n'est pas le plus simple et il a le don d'être parfois illisible avec pandas. Le plus simple est de le convertir au format texte avec Excel.

```
[2]: url = "https://www.insee.fr/fr/statistiques/fichier/1892086/pop-totale-france.xls"
url = "pop-totale-france.txt"
import pandas
df=pandas.read_csv(url, sep="\t", encoding="latin-1")
df.head(n=5)
```

```
[2]: Population totale par sexe et âge au 1er janvier 2017, France Unnamed: 1 \
0 Mis à jour : janvier 2017 NaN
1 NaN NaN
2 Année de naissance Âge révolu
3 2016 0
4 2015 1

      Unnamed: 2      Unnamed: 3 Unnamed: 4
0      NaN      NaN      NaN
1      NaN      NaN      NaN
2 Nombre d'hommes  Nombre de femmes  Ensemble
```

3	382 585	364 392	746 977
4	390 810	373 164	763 974

```
[3]: df=pandas.read_csv(url, sep="\t", encoding="latin-1", skiprows=3)
df.head(n=5)
```

```
[3]:   Année de naissance  Âge révolu  Nombre d'hommes  Nombre de femmes  Ensemble
0             2016           0         382 585         364 392    746 977
1             2015           1         390 810         373 164    763 974
2             2014           2         402 728         386 467    789 195
3             2013           3         405 636         387 097    792 733
4             2012           4         416 074         396 365    812 439
```

```
[4]: df.tail(n=5)
```

```
[4]:
           Année de naissance  Âge révolu \
100             1916 ou avant  100 ou plus
101                      NaN      Total
102                      NaN      NaN
103             Champ : France inclus Mayotte.      NaN
104  Source : Insee, estimations de population (rés...      NaN

           Nombre d'hommes  Nombre de femmes  Ensemble
100             3 015             15 617      18 632
101          32 455 859          34 534 967  66 990 826
102             NaN             NaN      NaN
103             NaN             NaN      NaN
104             NaN             NaN      NaN
```

La récupération de ces données est implémentée dans la fonction `population_france_year` :

```
[5]: from actuariat_python.data import population_france_year
df = population_france_year()
```

```
[6]: df.head(n=3)
```

```
[6]:   naissance  age  hommes  femmes  ensemble
0       2016    0  382585  364392    746977
1       2015    1  390810  373164    763974
2       2014    2  402728  386467    789195
```

```
[7]: df.tail(n=3)
```

```
[7]:   naissance  age  hommes  femmes  ensemble
98       1918   98   2474   10089    12563
99       1917   99   1362    6765     8127
100      1916  100   3015   15617    18632
```

D'après cette table, il y a plus de personnes âgées de 110 ans que de 109 ans. C'est dû au fait que la dernière ligne agrège toutes les personnes âgées de plus de 110 ans.

1.0.2 Table de mortalité 2000-2002 (France)

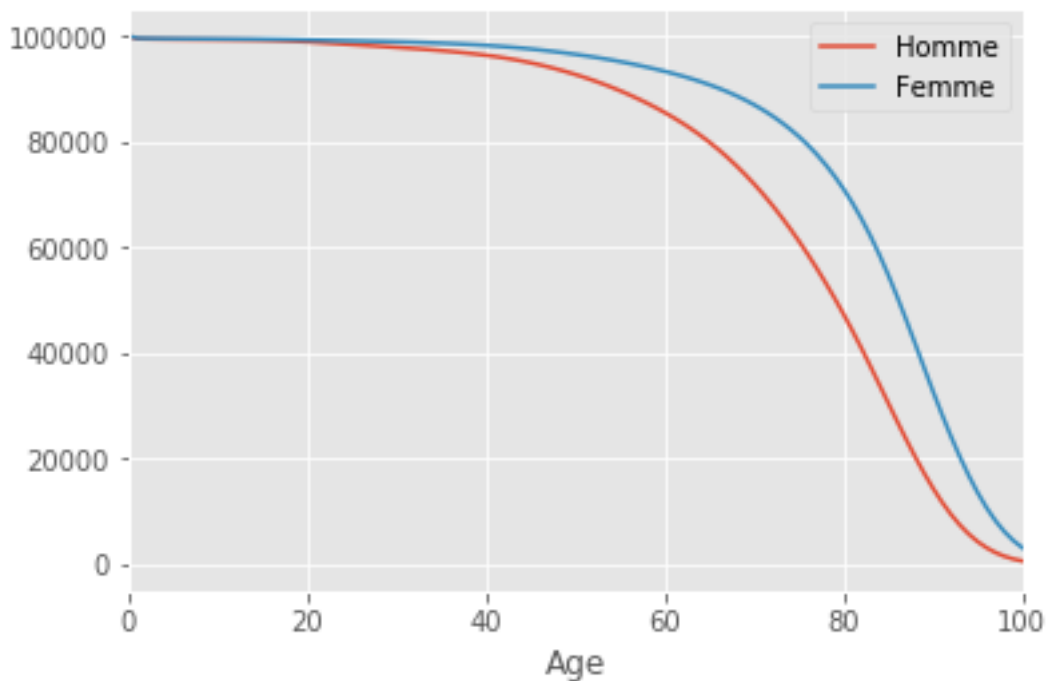
On utilise quelques raccourcis afin d'éviter d'y passer trop de temps. Les données sont fournis au format Excel à l'adresse : <http://www.institutdesactuaires.com/gene/main.php?base=314>. La fonction `table_mortalite_france_00_02` permet de les récupérer.

```
[8]: from actuariat_python.data import table_mortalite_france_00_02
df=table_mortalite_france_00_02()
df.head()
```

```
[8]:   Age      Homme      Femme
0  0.0  100000.0  100000.0
1  1.0   99511.0   99616.0
2  2.0   99473.0   99583.0
3  3.0   99446.0   99562.0
4  4.0   99424.0   99545.0
```

```
[9]: df.plot(x="Age",y=["Homme", "Femme"],xlim=[0,100])
```

```
[9]: <matplotlib.axes._subplots.AxesSubplot at 0x193a55bee48>
```



1.0.3 Taux de fécondité (France)

On procède de même pour cette table avec la fonction `fecondite_france`. Source : [INSEE : Fécondité selon l'âge détaillé de la mère](#).

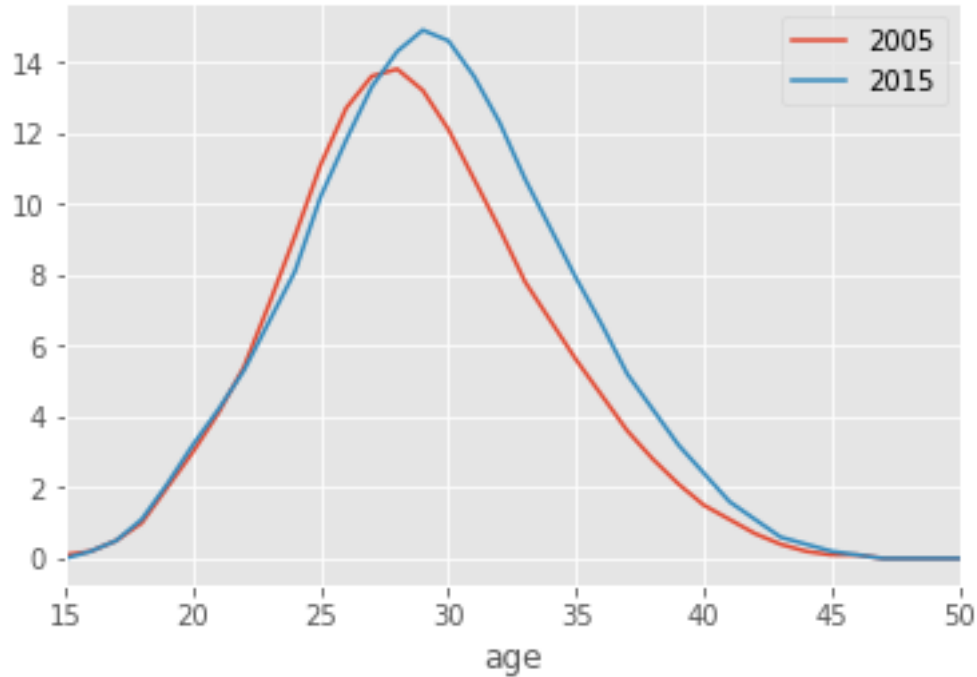
```
[10]: from actuariat_python.data import fecondite_france
df=fecondite_france()
df.head()
```

```
[10]:   age  2005  2015
3  15.0    0.1    0.0
4  16.0    0.2    0.2
5  17.0    0.5    0.5
```

```
6 18.0 1.0 1.1
7 19.0 2.0 2.1
```

```
[11]: df.plot(x="age", y=["2005", "2015"])
```

```
[11]: <matplotlib.axes._subplots.AxesSubplot at 0x193a583fba8>
```



1.0.4 Table de mortalité étendue 1960-2010

table de mortalité de 1960 à 2010 qu'on récupère à l'aide de la fonction `table_mortalite_euro_stat`.

```
[12]: from actuariat_python.data import table_mortalite_euro_stat
      table_mortalite_euro_stat()
```

```
[12]: 'mortalite.txt'
```

```
[13]: import os
      os.stat("mortalite.txt")
```

```
[13]: os.stat_result(st_mode=33206, st_ino=9007199255282328, st_dev=311511962,
      st_nlink=1, st_uid=0, st_gid=0, st_size=105741052, st_atime=1480548419,
      st_mtime=1480548432, st_ctime=1480548419)
```

```
[14]: import pandas
      df = pandas.read_csv("mortalite.txt", sep="\t", encoding="utf8", low_memory=False)
      df.head()
```

```
[14]:
```

	annee	valeur	age	age_num	indicateur	genre	pays
0	2012	0.00000	Y01	1.0	DEATHRATE	F	AD
1	2014	0.00042	Y01	1.0	DEATHRATE	F	AL
2	2009	0.00080	Y01	1.0	DEATHRATE	F	AM
3	2008	0.00067	Y01	1.0	DEATHRATE	F	AM
4	2007	0.00052	Y01	1.0	DEATHRATE	F	AM

```
[15]: df [ ((df.age=="Y60") | (df.age=="Y61")) & (df.annee == 2000) & (df.pays=="FR") & (df.
genre=="F")]
```

```
[15]:
```

	annee	valeur	age	age_num	indicateur	genre	pays
87406	2000	5.020000e-03	Y60	60.0	DEATHRATE	F	FR
88955	2000	4.860000e-03	Y61	61.0	DEATHRATE	F	FR
485730	2000	2.580000e+01	Y60	60.0	LIFEXP	F	FR
487268	2000	2.490000e+01	Y61	61.0	LIFEXP	F	FR
882568	2000	5.010000e-03	Y60	60.0	PROBDEATH	F	FR
884117	2000	4.850000e-03	Y61	61.0	PROBDEATH	F	FR
1281364	2000	9.949900e-01	Y60	60.0	PROBSURV	F	FR
1282913	2000	9.951500e-01	Y61	61.0	PROBSURV	F	FR
1679610	2000	9.307600e+04	Y60	60.0	PYLIVED	F	FR
1681148	2000	9.261800e+04	Y61	61.0	PYLIVED	F	FR
2075898	2000	9.331000e+04	Y60	60.0	SURVIVORS	F	FR
2077436	2000	9.284300e+04	Y61	61.0	SURVIVORS	F	FR
2472219	2000	2.405594e+06	Y60	60.0	TOTPYLIVED	F	FR
2473757	2000	2.312517e+06	Y61	61.0	TOTPYLIVED	F	FR

```
[16]:
```