titanic

February 26, 2018

1 A. Etude de l'âge des passagers

In [1]:
```
import pandas as pa
```

In [2]:
```
T=pa.read_csv('/Users/fredericbro/Documents/Lycee(henrimoissan)/Prof (henri-moissan)/Ann
```

In [3]:
```
T.head()
```

Out[3]:

	pclass	survived	name	sex	age	sibsp	parch	ticket	fare	cabin	embarked	boat	body	home.dest
0	1	1	Allen, Miss. Elisabeth Walton	female	29.0000	0	0	24160	211.3375	B5	S	2	NaN	St Louis, MO
1	1	1	Allison, Master. Hudson Trevor	male	0.9167	1	2	113781	151.5500	C22 C26	S	11	NaN	Montreal, PQ / Chesterville, ON
2	1	0	Allison, Miss. Helen Loraine	female	2.0000	1	2	113781	151.5500	C22 C26	S	NaN	NaN	Montreal, PQ / Chesterville, ON
3	1	0	Allison, Mr. Hudson Joshua Creighton	male	30.0000	1	2	113781	151.5500	C22 C26	S	NaN	135.0	Montreal, PQ / Chesterville, ON
4	1	0	Allison, Mrs. Hudson J C (Bessie Waldo Daniels)	female	25.0000	1	2	113781	151.5500	C22 C26	S	NaN	NaN	Montreal, PQ / Chesterville, ON

In [4]:
```
T.describe()
```

Out[4]:

	pclass	survived	age	sibsp	parch
count	1309.000000	1309.000000	1046.000000	1309.000000	1309.000000
mean	2.294882	0.381971	29.881135	0.498854	0.385027
std	0.837836	0.486055	14.413500	1.041658	0.865560
min	1.000000	0.000000	0.166700	0.000000	0.000000
25%	2.000000	0.000000	21.000000	0.000000	0.000000
50%	3.000000	0.000000	28.000000	0.000000	0.000000

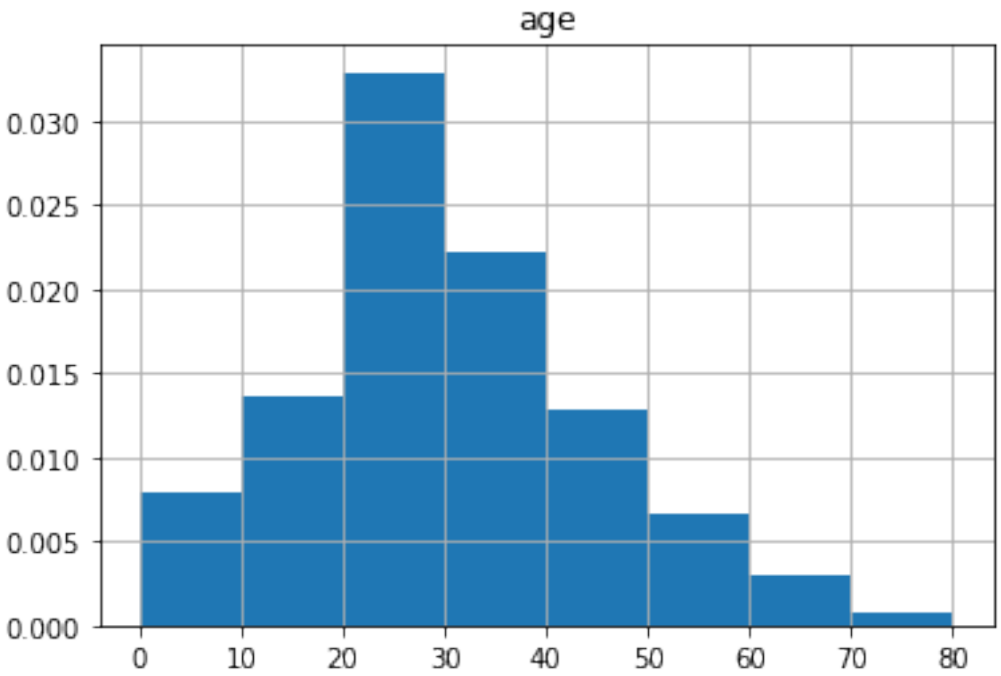
```
75%       3.000000     1.000000    39.000000     1.000000     0.000000
max       3.000000     1.000000    80.000000     8.000000     9.000000

              fare        body
count  1308.000000  121.000000
mean     33.295479  160.809917
std      51.758668   97.696922
min       0.000000    1.000000
25%       7.895800   72.000000
50%      14.454200  155.000000
75%      31.275000  256.000000
max     512.329200  328.000000
```

In [5]:
```
import pylab as pl
```

In [6]:
```
T.hist(column='age',bins=range(0,90,10),normed=True)
```

Out[6]:
```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x10efdf358>]], dtype=object)
```

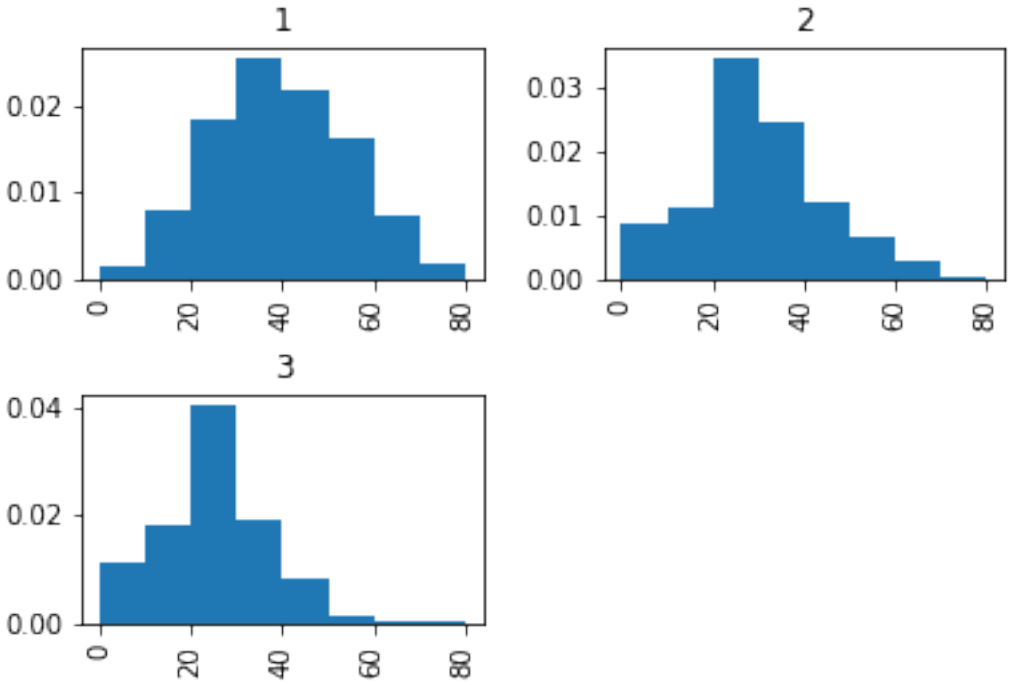
In [7]:
```
pl.show()
```



In [8]:
```
T.hist(column='age',bins=range(0,90,10),normed=True,by='pclass')
```

Out[8]:
```
array([[<matplotlib.axes._subplots.AxesSubplot object at 0x1131fc0f0>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x1132394a8>],
       [<matplotlib.axes._subplots.AxesSubplot object at 0x1132714a8>,
        <matplotlib.axes._subplots.AxesSubplot object at 0x1132ac518>]], dtype=object)
```

```
In [9]: pl.show()
```



2 B. Etude relative aux survivants

2.1 1. Proportion de survivants

```
In [10]: T1=T[T['survived']==1]

In [11]: len(T1)/len(T)

Out[11]: 0.3819709702062643
```

2.2 2. Proportion de survivants par classe

```
In [12]: a=T1['pclass'].value_counts(sort=False)

In [13]: a

Out[13]: 1    200
         2    119
         3    181
         Name: pclass, dtype: int64
```

La variable **a** est la série statistique qui à chaque classe associe le nombre des survivants.

```
In [14]: b=T['pclass'].value_counts(sort=False)
```

```
In [15]: b

Out[15]: 1    323
         2    277
         3    709
         Name: pclass, dtype: int64
```

La variable **b** est la série statistique qui à chaque classe associe le nombre des passagers.

```
In [16]: a/b

Out[16]: 1    0.619195
         2    0.429603
         3    0.255289
         Name: pclass, dtype: float64
```

2.3 3. Proportion de survivants par sexe

```
In [17]: c=T1['sex'].value_counts()

In [18]: d=T['sex'].value_counts()

In [19]: c/d

Out[19]: female    0.727468
         male      0.190985
         Name: sex, dtype: float64
```

2.4 4. La règle "les femmes et les enfants d'abord" fut-elle respectée ?

a) Tableau **T2** concernant **survivants** femmes **ou** les enfants de moins de 18 ans

```
In [20]: T2=T1[(T1['sex']=='female') | ((T1['sex']=='male')&(T1['age']<18))]
```

b) Tableau **T3** concernant les femmes **ou** les enfants de moins de 18 ans

```
In [21]: T3=T[(T['sex']=='female') | ((T['sex']=='male')&(T['age']<18))]
```

c) Conclusion

```
In [22]: len(T2)/len(T3)

Out[22]: 0.6751824817518248

In [23]: len(T1)/len(T)

Out[23]: 0.3819709702062643
```