

# Notebook 5

## Analyse cartographique

Avec l'outil Geopandas, nous pouvons visualiser nos données sous formes de cartes. Pour cela, il faut récupérer sur internet un fichier csv associé à la carte de notre choix (ces fichiers comporte une colonne géométrie afin de tracer les pays/régions). De plus, les cartes peuvent permettre de dégager certaines dynamiques géographiques et de donner un autre point de vue sur un jeu de données.

Tout d'abord, nous allons tout simplement afficher le taux de suicide par pays d'après les données que nous avons récupéré dans les notebooks précédents afin de voir si l'on peut en dégager certaines tendances.

In [2]:

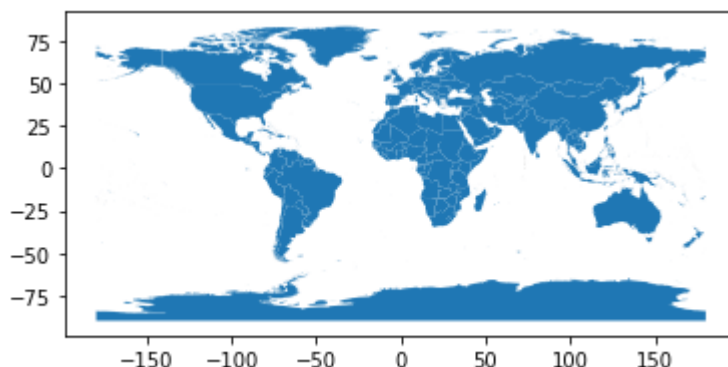
```
import pandas as pd
import matplotlib.pyplot as plt
import geopandas as gpd
from pandas import DataFrame
```

In [3]:

```
fp = "./Data/TM_WORLD_BORDERS-0.3.shp"
map_df = gpd.read_file(fp) #On importe une carte du monde pour y transposer nos données
```

In [4]:

```
map_df.plot();
```



In [5]:

```
df = pd.read_csv("./Data/age-standardized.csv")
df.columns=['Country', 'Suicide Rate']
merged = map_df.merge(df, left_on = 'NAME', right_on = 'Country', how = 'left')
```

In [6]:

```
variable = 'Suicide Rate'
```

```

vmin, vmax = 0, 30.200000
fig, ax = plt.subplots(1, figsize=(30, 15));
merged.plot(column=variable, cmap='YlOrBr', linewidth=0.8, ax=ax, edgecolor='0.8', legend = True,
            legend_kwds={'orientation': "horizontal"});
ax.set_title("Taux de suicide par pays dans le monde (pour 100k habitant s)", fontdict={'fontsize': '25', 'fontweight' : '3'});
ax.axis('off');

```

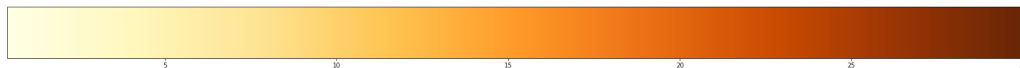
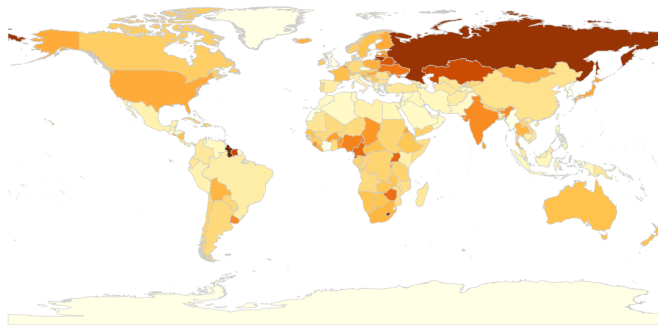
C:\Users\Utilisateur\Anaconda3\lib\site-packages\matplotlib\colors.py:527: RuntimeWarning: invalid value encountered in le

```

ss
    xa[xa < 0] = -1

```

Taux de suicide par pays dans le monde (pour 100k habitants)



Au premier abord, il semblerait que les pays à religion musulmane possèdent un taux de suicide bien moins élevé que dans le reste du globe. Nous allons donc vérifier ça :

In [7]:

```

Religion = pd.read_csv("../Data/religion.csv")
# On va s'intéresser uniquement aux statistiques pour l'année 2010 (année
  la plus récente proposée dans le dataset)
Religion = Religion[Religion.year == 2010]
Religion = Religion[['state', 'islam_percent']]
Religion.columns = ['name', 'islam_percent']
carte = map_df.merge(Religion, left_on = 'NAME', right_on = 'name', how='left')

```

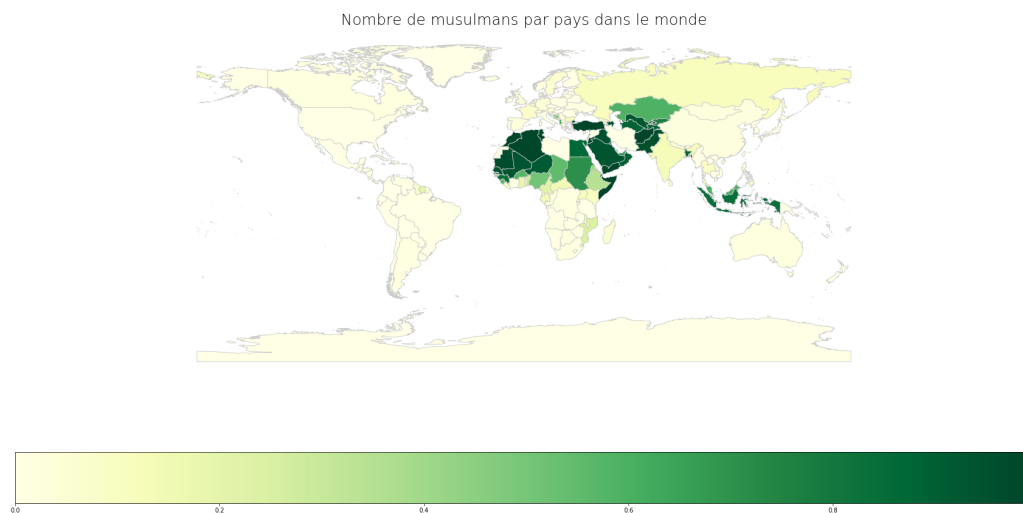
In [8]:

```

var = 'islam_percent'
vmin, vmax = 0, 1
fig, ax = plt.subplots(1, figsize=(30, 15));
carte.plot(column=var, cmap='YlGn', linewidth=0.8, ax=ax, edgecolor='0.8',
            legend = True,
            legend_kwds={'orientation': "horizontal"});
ax.set_title('Nombre de musulmans par pays dans le monde', fontdict={'font
size': '25', 'fontweight' : '3'});
ax.axis('off');

```

```
C:\Users\Utilisateur\Anaconda3\lib\site-packages\matplotlib\colors.py:527: RuntimeWarning: invalid value encountered in less
    xa[xa < 0] = -1
```

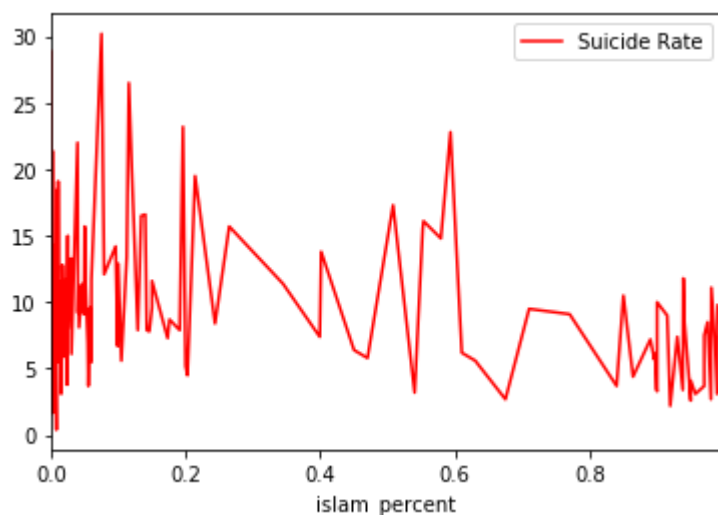


On observe que les pays d'Afrique du Nord où la religion majoritaire est l'Islam présentent un taux de suicide plus bas que la moyenne.

Comparant les taux de suicide en fonction du pourcentage de musulmans dans le pays:

In [9]:

```
SuiciReligion = df.merge(Religion, left_on='Country', right_on='name')
SuiciReligion = SuiciReligion.sort_values(by = 'islam_percent')
SuiciReligion.plot(kind='line', x='islam_percent', y='Suicide Rate', color = 'Red');
```



On observe une légère baisse du taux de suicide dans les pays avec plus de 25% de Musulmans. Néanmoins on n'observe pas de nette figure de corrélation.

Pourtant, en Amérique du Sud notamment, on observe de faibles taux de suicide. Etant donné que le suicide est un acte condamné par toutes les religions majeures dans le monde, on se demande si le taux de suicide pourrait être corrélé avec le taux de pratiquants dans le pays. Pour cela, nous allons nous intéresser au taux d'athées par pays.

In [10]:

```
atheism = pd.read_csv("./Data/atheism.csv")
atheism.columns=['country', 'AtheistRate', 'pop2019']
atheism = atheism.iloc[:, [0,1]]
carte2 = map_df.merge(atheism, left_on = 'NAME', right_on = 'country', how=
'left')
```

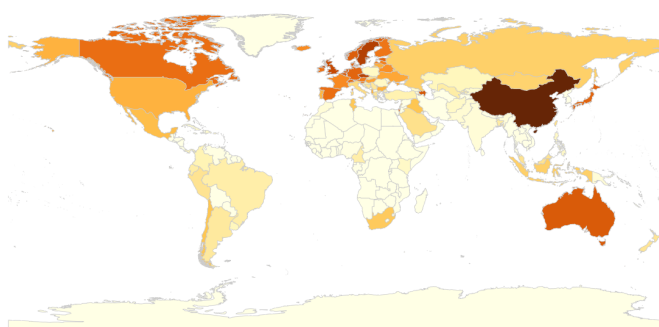
In [11]:

```
var = 'AtheistRate'
vmin, vmax = 0, 100
fig, ax = plt.subplots(1, figsize=(30, 15))
carte2.plot(column=var, cmap='YlOrBr', linewidth=0.8, ax=ax, edgecolor='0.
8', legend = True,
             legend_kwds={'orientation': "horizontal"});
ax.set_title("Pourcentage d'athés par pays dans le monde", fontdict={'font
size': '25', 'fontweight' : '3'})
ax.axis('off');
```

```
C:\Users\Utilisateur\Anaconda3\lib\site-packages\matplotlib\c
olors.py:527: RuntimeWarning: invalid value encountered in le
ss
```

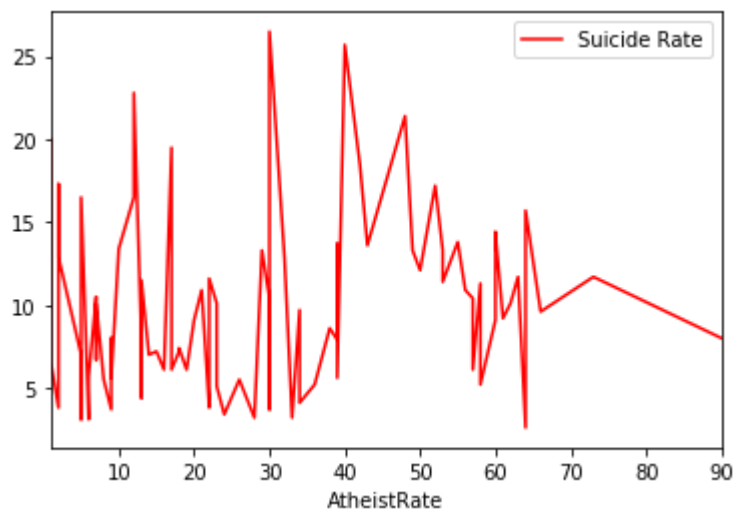
```
xa[xa < 0] = -1
```

Pourcentage d'athés par pays dans le monde



In [12]:

```
SuiciAtheism = df.merge(atheism, left_on='Country', right_on='country')
SuiciAtheism = SuiciAtheism.sort_values(by = 'AtheistRate')
SuiciAtheism.plot(kind='line', x='AtheistRate', y='Suicide Rate', color='R
ed');
```



Finalement on n'observe (encore une fois) aucune corrélation entre le taux de suicide et le nombre d'athés.

## Conclusion

**En conclusion, même si les données géographiques ne nous ont pas permis d'établir de corrélation par rapport aux facteurs que nous avons dégagés au préalable, elle nous permettent tout de même d'observer les pays/régions où le taux de suicide est le plus élevé pour potentiellement y consacrer une étude plus tard afin d'expliquer ces taux plus élevés.**