

## Importer les packages

In [1]:

```
import pyforest
from Utils import replace_columns
from utils import convert_dtype
from UTILS import CRW
# Hide warnings
import warnings
warnings.filterwarnings('ignore')
```

## Importer les données

In [2]:

```
Data1 = replace_columns(pd.read_csv('data/forest_data/Dakar-Forest (2000-2020).csv'))
Data2 = replace_columns(pd.read_csv('data/forest_data/Diourbel-Forest (2000-2020).csv'))
Data3 = replace_columns(pd.read_csv('data/forest_data/Fatick-Forest (2000-2020).csv'))
Data4 = replace_columns(pd.read_csv('data/forest_data/Kaffrine-Forest (2000-2020).csv'))
Data5 = replace_columns(pd.read_csv('data/forest_data/Kaolack-Forest (2000-2020).csv'))
Data6 = replace_columns(pd.read_csv('data/forest_data/Kedougou-Forest (2000-2020).csv'))
Data7 = replace_columns(pd.read_csv('data/forest_data/Kolda-Forest (2000-2020).csv'))
Data8 = replace_columns(pd.read_csv('data/forest_data/Louga-Forest (2000-2020).csv'))
Data9 = replace_columns(pd.read_csv('data/forest_data/Matam-Forest (2000-2020).csv'))
Data10 = replace_columns(pd.read_csv('data/forest_data/Saint louis-Forest (2000-2020).csv'))
Data11 = replace_columns(pd.read_csv('data/forest_data/Sedhiou-Forest (2000-2020).csv'))
Data12 = replace_columns(pd.read_csv('data/forest_data/Tambacounda-Forest (2000-2020).csv'))
Data13 = replace_columns(pd.read_csv('data/forest_data/Thies-Forest (2000-2020).csv'))
Data14 = replace_columns(pd.read_csv('data/forest_data/Ziguinchor-Forest (2000-2020).csv'))
```

## Exploration des données

In [3]:

```
Data1.dtypes
```

Out[3]:

```
Date          object
Zone Perdue    object
dtype: object
```

### Convertir les variables

In [4]:

```
Data1 = convert_dtype(Data1)
Data2 = convert_dtype(Data2)
Data3 = convert_dtype(Data3)
Data4 = convert_dtype(Data4)
Data5 = convert_dtype(Data5)
Data6 = convert_dtype(Data6)
Data7 = convert_dtype(Data7)
Data8 = convert_dtype(Data8)
Data9 = convert_dtype(Data9)
Data10 =convert_dtype(Data10)
Data11 =convert_dtype(Data11)
Data12 = convert_dtype(Data12)
Data13 = convert_dtype(Data13)
Data14 = convert_dtype(Data14)
```

**La superficie totale des zones couvertes d'arbres (ha) dans chaque région**

In [5]:

```
Tree_covered_area = pd.read_csv('data/forest_data/Tree_covered_area.csv').drop(['Unnamed: 0'], axis = 1)
```

In [6]:

```
Tree_covered_area
```

Out[6]:

	region	Tree_covered_area
0	Dakar	350.18
1	Diourbel	16.98
2	Fatick	69919.79
3	Kaffrine	1684.56
4	Kaolack	13747.08
5	Kedougou	1387613.15
6	Kolda	1117358.42
7	Louga	19.29
8	Matam	59.95
9	Saint Louis	617.87
10	Sedhiou	619421.31
11	Tambacounda	1118735.78
12	Thies	9060.77
13	Ziguinchor	460975.69

#### Visualisation de la perte de superficie des zones couvertes d'arbres

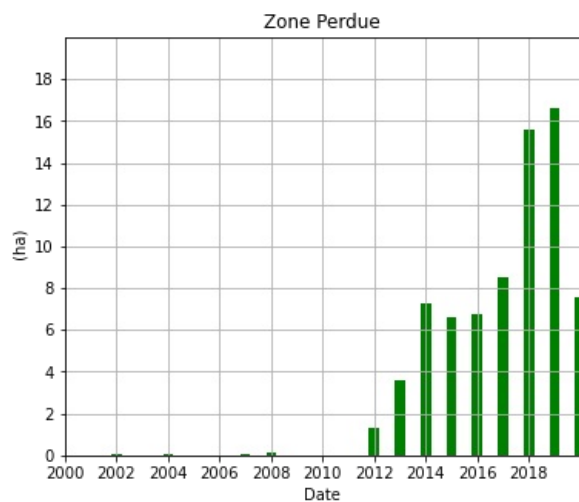
In [7]:

```
def visualize_lost(data, color, val1, val2):  
    fig = plt.figure(figsize = (6, 5))  
  
    # creating the bar plot  
    plt.bar(data['Date'], data['Zone Perdue'], color =color,  
            width = 0.4)  
  
    plt.xlabel('Date')  
    plt.ylabel(" (ha)")  
    plt.title("Zone Perdue")  
    plt.xlim([2000, 2020])  
    plt.ylim([0, 20])  
    plt.yticks(np.arange(0, val1, val2))  
    plt.xticks(np.arange(2000, 2020, 2))  
    plt.grid()  
    plt.show()
```

#### Région de Dakar

In [8]:

```
visualize_lost(Data1, 'green', 20, 2)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Dakar**

In [9]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data1['Zone Perdue'])* 100) / Tree_covered_area  
['Tree_covered_area'][0]))
```

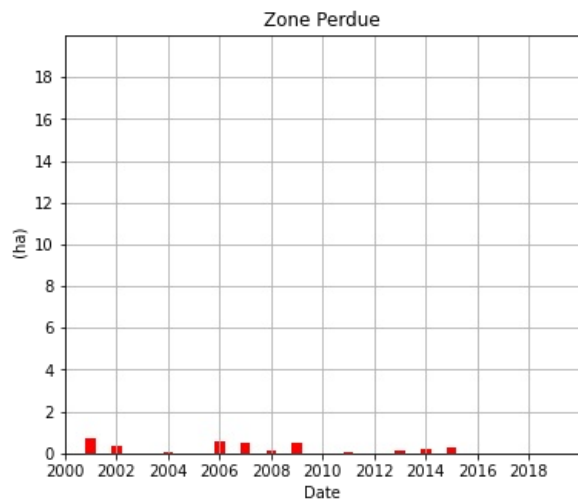
La superficie de la zone perdue est de : 21.19807485238061 %

- A travers la figure ci-dessus, on constate que la région de Dakar a enregistré une faible perte de superficie des zones couvertes d'arbres de l'année 2000 à 2011. Par contre de l'année 2012 à 2020, on observe une augmentation des zones couvertes d'arbres perdues. Ainsi on note une perte de 21.19 % enregistrée de 2000 jusqu'à 2020.
- Cette augmentation de la perte de superficie des zones couvertes d'arbres ces dernières années fait de la région de Dakar une région fortement menacée de déforestation.

**Région de Diourbel**

In [10]:

```
visualize_lost(Data2, 'red', 20, 2)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Diourbel**

In [11]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data2['Zone Perdue'])* 100) / Tree_covered_area  
['Tree_covered_area'][1]))
```

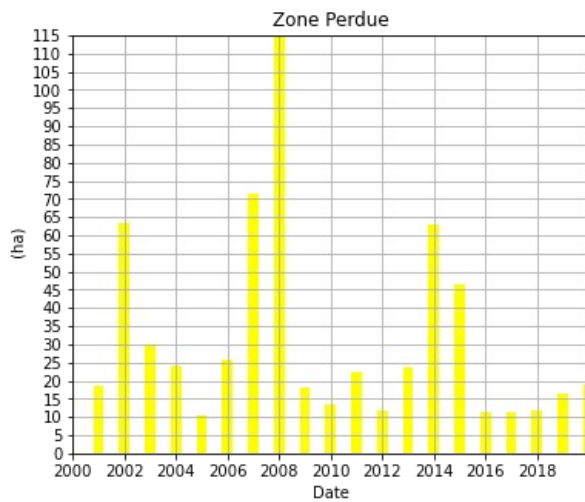
La superficie de la zone perdue est de : 21.922399765751525 %

- Dans la région de Diourbel, on observe une faible perte de zones couvertes d'arbres des années 2000 à 2020. Comparée à la superficie totale de zones couvertes d'arbres, on voit que la superficie perdue est estimée à 21.92 %.
- Vu les faibles pertes de superficie de zones couvertes d'arbres lors des différentes années, on peut considérer que la région de Diourbel n'est pas trop menacée par la déforestation.

**Région de Fatick**

In [12]:

```
visualize_lost(Data3, 'yellow', 120, 5)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Fatick**

In [13]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data3['Zone Perdue'])* 100) / Tree_covered_area  
['Tree_covered_area'][2]))
```

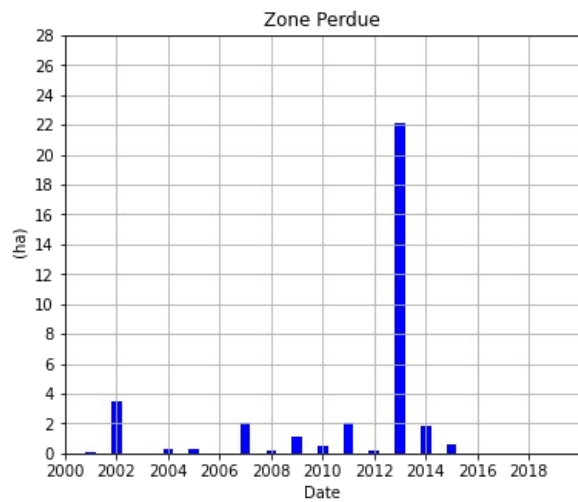
La superficie de la zone perdue est de : 0.8957581077324069 %

- A travers la figure ci-dessus on observe tantôt une augmentation et diminution de la perte de superficie des zones couvertes d'arbres avec plus 115 ha perdue en 2008. La perte de superficie est estimé à moins de 1 % de 2000 à 2020.
- Malgré une augmentation de la perte de superficie lors de certaines années, le pourcentage de zones perdues reste faible par rapport la zone totale couverte d'arbres.

**Région de Kaffrine**

In [14]:

```
visualize_lost(Data4, 'blue', 30, 2)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Kaffrine**

In [15]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data4['Zone Perdue'])* 100) / Tree_covered_area  
['Tree_covered_area'][3]))
```

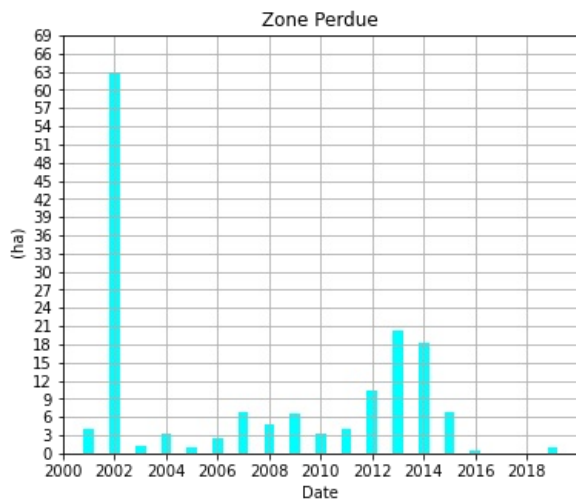
La superficie de la zone perdue est de : 2.065246160651859 %

- On observe à travers la figure ci-dessus que la région Kaffrine enregistre de faibles perte de zones arborées durant ces dernières années avec une superficie d'environ 2 % perdue par rapport à la superficie totale de zones couvertes d'arbres.
- Donc on peut qualifier que la région Kaffrine est faiblement menacé par déforestation.

**Région de Kaolack**

In [16]:

```
visualize_lost(Data5, 'cyan', 70, 3)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Kaolack**

In [17]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data5['Zone Perdue'])* 100) / Tree_covered_area  
['Tree_covered_area'][4]))
```

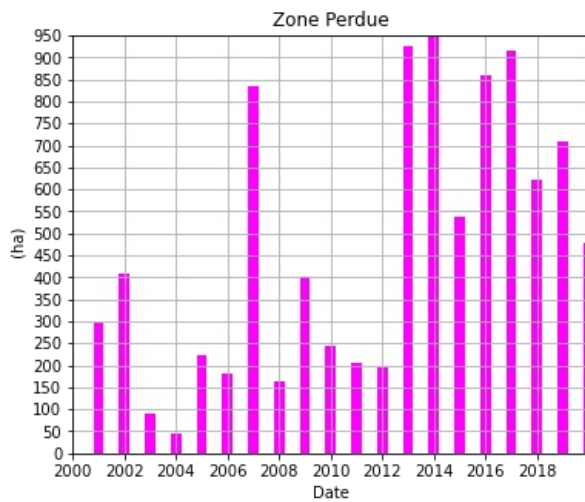
La superficie de la zone perdue est de : 1.1431121878490882 %

- De la figure ci-dessus, on observe de faibles pertes de superficie durant les années 2003 à 2006 et les années 2016 et 2020 dans la région de Kaolack. Par contre de fortes pertes sont notées durant l'année 2002 et les années 2012 à 2014.
- En estimant le taux de superficie perdue durant les années 2000 à 2020, on obtient une valeur d'environ 1 %. Donc on peut conclure que la région de Kaolack est faiblement menacée par la déforestation.

**Région de Kédougou**

In [18]:

```
visualize_lost(Data6, 'magenta', 1000, 50)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Kédougou**

In [19]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data6['Zone Perdue'])* 100) / Tree_covered_area  
['Tree_covered_area'][5]))
```

La superficie de la zone perdue est de : 0.6774087468616645 %

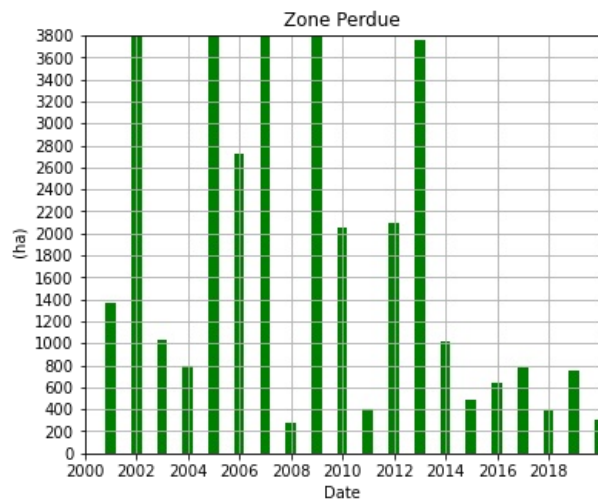
- Malgré que le taux de superficie perdue soit faible (0.677 %), la région de Kédougou est l'une des régions les plus menacée par la déforestation. On voit à travers la figure ci-dessus une forte augmentation des zones couvertes d'arbres perdues durant les années 2013 à 2020 avec plus 1000 ha perdus en 2014.

**Région de Kolda**



In [20]:

```
visualize_lost(Data7, 'green', 4000, 200)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Kolda**

In [21]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data7['Zone Perdue'])* 100) / Tree_covered_area  
['Tree_covered_area'][6]))
```

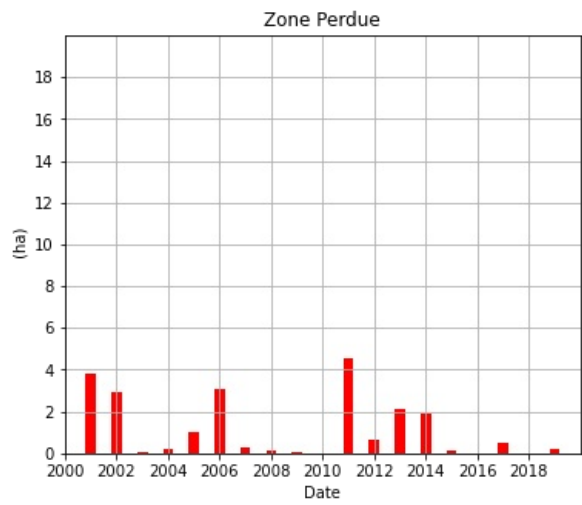
La superficie de la zone perdue est de : 4.309916410319083 %

- La région Kolda est caractérisée par de fortes pertes de superficie de zones couvertes d'arbres avec des pertes estimées à plus de 4000 ha en 2002, 2005, 2007 et 2009. Malgré un taux de superficie perdue faible (4.30 %), le région de Kolda est l'une des régions fortement menacée par la déforestation.

**Région de Louga**

In [22]:

```
visualize_lost(Data8, 'red', 20, 2)
```



*Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Louga*

In [23]:

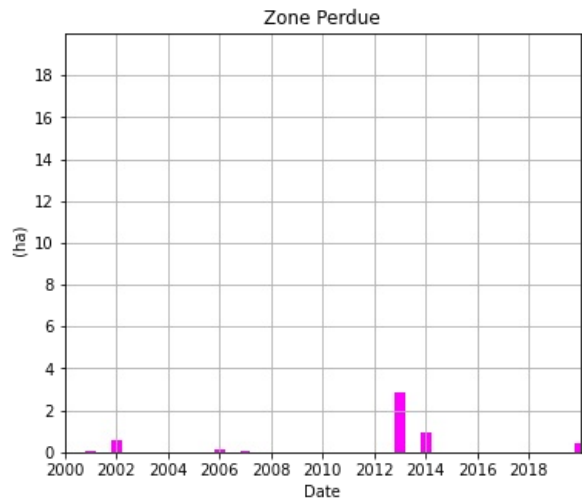
```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data8['Zone Perdue'])* 100) / Tree_covered_area ['Tree_covered_area'][7]))
```

La superficie de la zone perdue est de : 112.705252750727 %

**Région de Matam**

In [24]:

```
visualize_lost(Data9, 'magenta' ,20, 2)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Matam**

In [25]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data9['Zone Perdue'])* 100) / Tree_covered_area['Tree_covered_area'][8]))
```

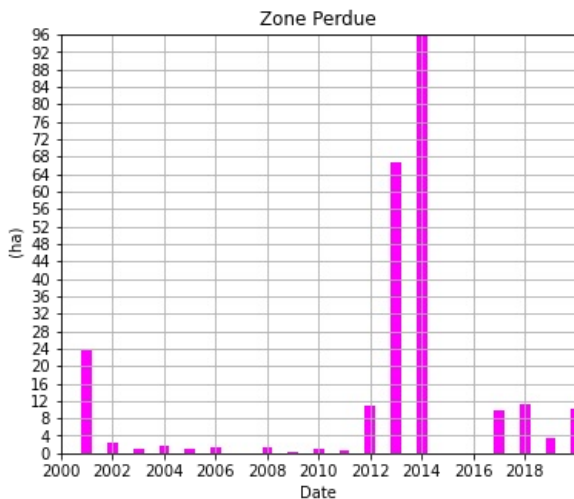
La superficie de la zone perdue est de : 8.550422338846824 %

- A travers la figure ci-dessus représentant la perte de superficie de zones couvertes d'arbres dans la région de Matam, on observe de très faibles pertes de superficie durant presque toutes les années de 2000 à 2020.
- On conclut que la région de Matam est très faiblement menacée par la déforestation.

**Région de Saint Louis**

In [26]:

```
visualize_lost(Data10, 'magenta' ,100, 4)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Saint Louis**

In [27]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data10['Zone Perdue'])* 100) / Tree_covered_area['Tree_covered_area'][9]))
```

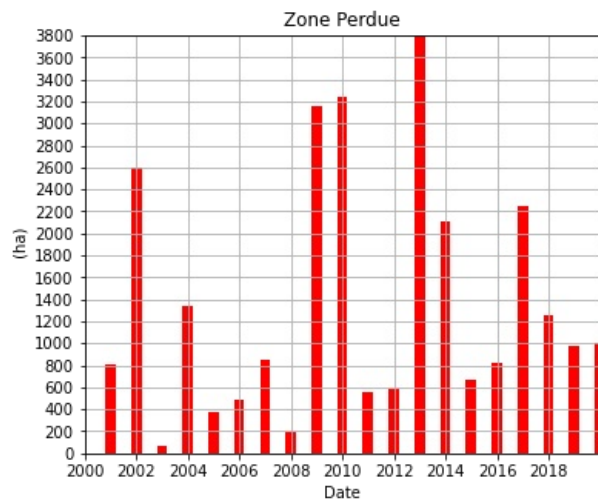
La superficie de la zone perdue est de : 39.94109561602256 %

- En partant de la figure ci-dessus, on constate que la région de Saint Louis a enregistré de faibles pertes de superficie de zones couvertes d'arbres durant les années 2002 à 2011 contrairement aux années 2001, 2012, 2013 et 2014 où la région a enregistré de fortes pertes de superficie avec plus de 100 ha perdus lors de l'année 2014.
- Durant ces 20 dernières années la région de Saint Louis a perdu environ 40 % de sa superficie de zones couvertes d'arbres entraînant une forte menace de déforestation.

**Région de Sedhiou**

In [28]:

```
visualize_lost(Data11, 'red', 4000, 200)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Sédhiou**

In [29]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data11['Zone Perdue'])* 100) / Tree_covered_area['Tree_covered_area'][10]))
```

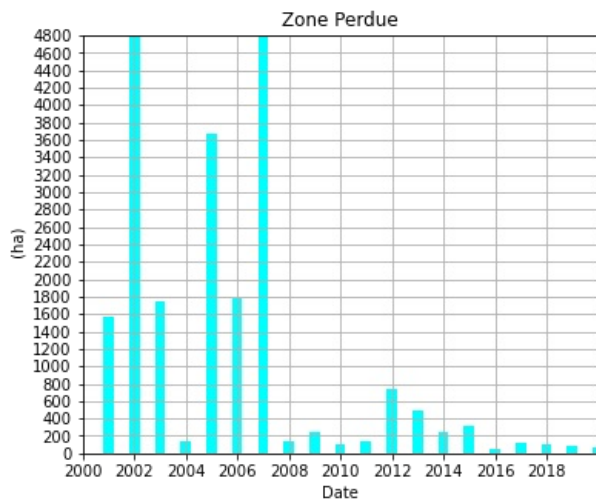
La superficie de la zone perdue est de : 4.726915696193469 %

- Comme les régions Kolda et Kédougou, la région de Sédhiou a aussi enregistré de fortes pertes de superficie durant ces 20 dernières années avec plus de 3000 ha perdus lors des années 2009, 2010 et 2013.
- Malgré une faible superficie perdue de 4.73 %, la région est l'une des régions fortement menacée par la déforestation.

**Région de Tambacounda**

In [30]:

```
visualize_lost(Data12, 'cyan', 5000, 200)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Tambacounda**

In [31]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data12['Zone Perdue'])* 100) / Tree_covered_area['Tree_covered_area'][11]))
```

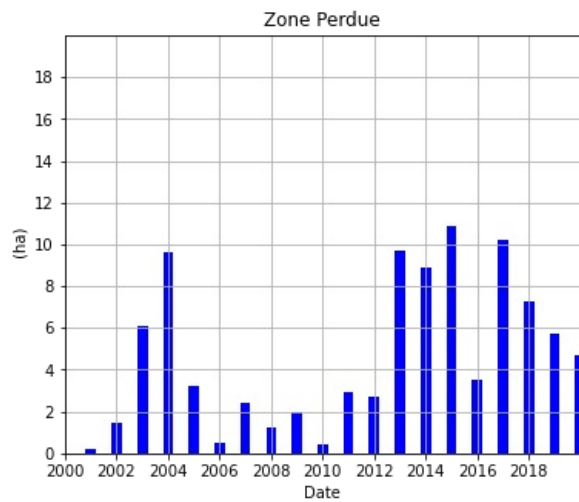
La superficie de la zone perdue est de : 2.089150501833774 %

- A travers le diagramme ci-dessus, on voit que durant les années 2000 à 2007, la région de Tambacounda a enregistré de grandes pertes de superficie avec plus de 5000 ha perdues lors des 2002 et 2007. Par contre depuis 2008 on constate une baisse très forte de superficie perdues avec moins de 100 ha de pertes 2020.
- Avec un taux d'environ 2.1 % de superficie de zones couvertes d'arbres perdus, la région de Tambacounda est très faiblement menacée par la déforestation.

**Région de Thiés**

In [32]:

```
visualize_lost(Data13, 'blue', 20, 2)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Thiès**

In [33]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data13['Zone Perdue'])* 100) / Tree_covered_area['Tree_covered_area'][12]))
```

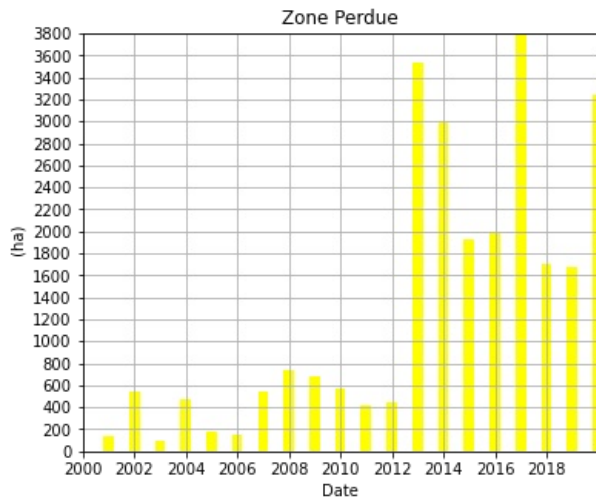
La superficie de la zone perdue est de : 1.0328927302806623 %

- Malgré une hausse des zones couvertes d'arbres perdues lors des années 2013 à 2020, la région de Thiès a enregistré un taux de perte de superficie de 1 % durant les 20 ans dernières. Donc on peut considérer que la region de Thiès est faiblement par la déforestation.

**Région de Ziguinchor**

In [34]:

```
visualize_lost(Data14, 'Yellow', 4000, 200)
```



**Superficie des zones couvertes d'arbres perdues en (%) par rapport à la superficie totale des zones couvertes d'arbres dans la région de Ziguinchor**

In [35]:

```
print('La superficie de la zone perdue est de : {} %'.format((sum(Data14['Zone Perdue'])* 100) / Tree_covered_area['Tree_covered_area'][13]))
```

La superficie de la zone perdue est de : 5.597783570721179 %

- Comme les régions de Kolda, Kédougou et Sedhiou, la région de Ziguinchor a enregistré de fortes pertes de superficie de zones couvertes d'arbres avec plus de 3500 ha perdues durant les années 2013 et 2017.
- Malgré un taux de perte de superficie 5.59 % pas tellement élevé par rapport à la superficie totale, la région de Ziguinchor est l'une des régions fortement menacée par la déforestation.

## Labelisation des données

In [36]:

```
Data1['risque_déforestation'] = [1]*20
Data2['risque_déforestation'] = [0]*20
Data3['risque_déforestation'] = [0]*20
Data4['risque_déforestation'] = [0]*20
Data5['risque_déforestation'] = [0]*20
Data6['risque_déforestation'] = [1]*20
Data7['risque_déforestation'] = [1]*20
Data8['risque_déforestation'] = [0]*20

Data9['risque_déforestation'] = [0]*20
Data10['risque_déforestation'] = [1]*20
Data11['risque_déforestation'] = [1]*20
Data12['risque_déforestation'] = [0]*20
Data13['risque_déforestation'] = [0]*20
Data14['risque_déforestation'] = [1]*20
```

In [37]:

```
Data1.head()
```

Out[37]:

	Date	Zone Perdue	risque_déforestation
1	2001	0.000000	1
2	2002	0.074471	1
3	2003	0.000000	1
4	2004	0.074460	1
5	2005	0.000000	1

In [38]:

```
Data2.head()
```

Out[38]:

	Date	Zone Perdue	risque_déforestation
1	2001	0.744093	0
2	2002	0.372197	0
3	2003	0.000000	0
4	2004	0.074518	0
5	2005	0.000000	0

- 1: signifie que la région est fortement menace par la déforestation
- 0: signifie que la région est faiblement menace par la déforestation

**Le jeu de données issu de la concaténation des données sur l'urbanisation et celles forestières**

In [43]:

```
Data_Urbanization_Forest = pd.read_csv('data/concat_data/Data_Urbanization_Forest.csv')
```

In [44]:

```
Data_Urbanization_Forest.head()
```

Out[44]:

	region	Date	Effectif de la population	Population rurale	Population urbaine	Taux d'urbanisation	Nombre de ménages ruraux	Nombre de ménages urbains	Taille moyenne des ménages	Taille moyenne des ménages ruraux	Taille moyenn de ménage urbain
0	Dakar	1970	724461.692160	27614.737826	696846.954334	96.188240	2785.348196	97871.763249	7.55	9.914286	7.12000
1	Dakar	1971	759203.252607	28805.871927	730397.380680	96.205776	2928.701576	101632.288592	7.52	9.835714	7.18666
2	Dakar	1972	795610.844586	30048.384407	765562.460179	96.223231	3079.629441	105546.295062	7.48	9.757143	7.25333
3	Dakar	1973	833764.362638	31344.491420	802419.871218	96.240606	3238.545239	109620.200986	7.45	9.678571	7.32000
4	Dakar	1974	873747.532649	32696.504713	841051.027936	96.257900	3405.885908	114397.582690	7.42	9.600000	7.35200

- Les valeurs de la variable 'Superficie Perdue' des années 1970 à 2001 ont été remplacées par la moyenne de la variable dans chaque région.