TP1 SERIES TEMPS

Réaliser par MEKA MOISE CHRISTIAN JUNIOR 21T2561

Il s'agit de l'implémentation du TP1 sur les taux de positivité du sars-cov2 au Cameroun pour la période allant de 2020 jusqu'à 2023

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

In [2]:
```
FILEPATH="SARS-Cov2-PCRdata.csv"
extension=FILEPATH.split(".")[-1]
if extension=="csv":
    df=pd.read_csv(FILEPATH, sep=";")
else:
    df=pd.read_excel(FILEPATH, index_col=0)
```

In [3]:
```
extension
```

Out[3]:
```
'csv'
```

In [4]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 146 entries, 0 to 145
Data columns (total 6 columns):
 #   Column       Non-Null Count  Dtype  
---  ------       --------------  -----  
 0   Year         146 non-null    int64  
 1   Period       146 non-null    object 
 2   Full Period  146 non-null    object 
 3   Ech Analyse  146 non-null    int64  
 4   Ech Positif  146 non-null    int64  
 5   Tx           146 non-null    float64
dtypes: float64(1), int64(3), object(2)
memory usage: 7.0+ KB
```

In [5]:
```
df.head()
```

Out[5]:

	Year	Period	Full Period	Ech Analyse	Ech Positif	Tx
0	2020	S42	20-S42	380	24	6.315789
1	2020	S43	20-S43	4050	189	4.666667
2	2020	S44	20-S44	4682	235	5.019223
3	2020	S45	20-S45	3715	332	8.936743
4	2020	S46	20-S46	4167	342	8.207343

In [6]:
```
metrique="Tx"
period="Full Period"
```

In [7]:
```
df[metrique].describe()
```

Out[7]:
```
count    146.000000
mean       5.876844
std        6.334876
min        0.000000
25%        1.070534
50%        3.840060
75%        8.687123
max       27.196943
Name: Tx, dtype: float64
```

In [8]:
```
df=df.dropna()
```

Question 1

In [9]:
```
#moyenne
np.mean(df[metrique])
```

Out[9]:
```
5.8768439308500735
```

In [10]:
```
#varaince
np.std(df[metrique])**2
```

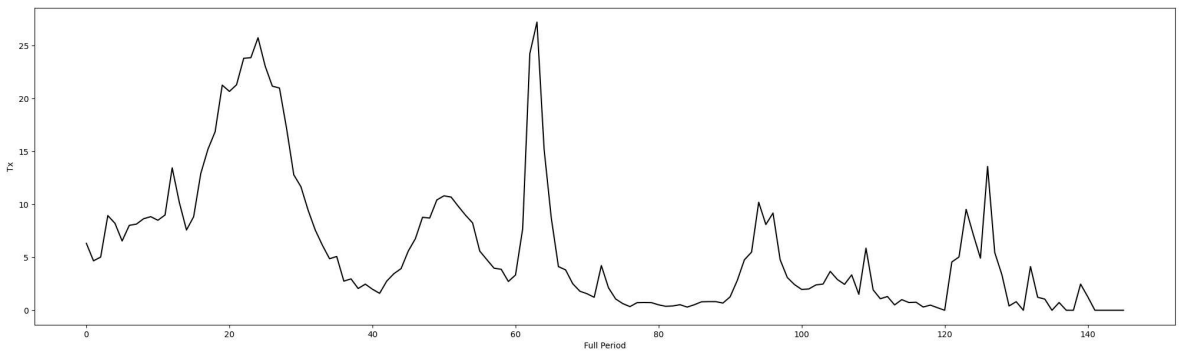
Out[10]:
```
39.85578662438383
```

In [11]:
```
#ecart-type
np.std(df[metrique])
```

Out[11]:
```
6.313143957204194
```

Question 2

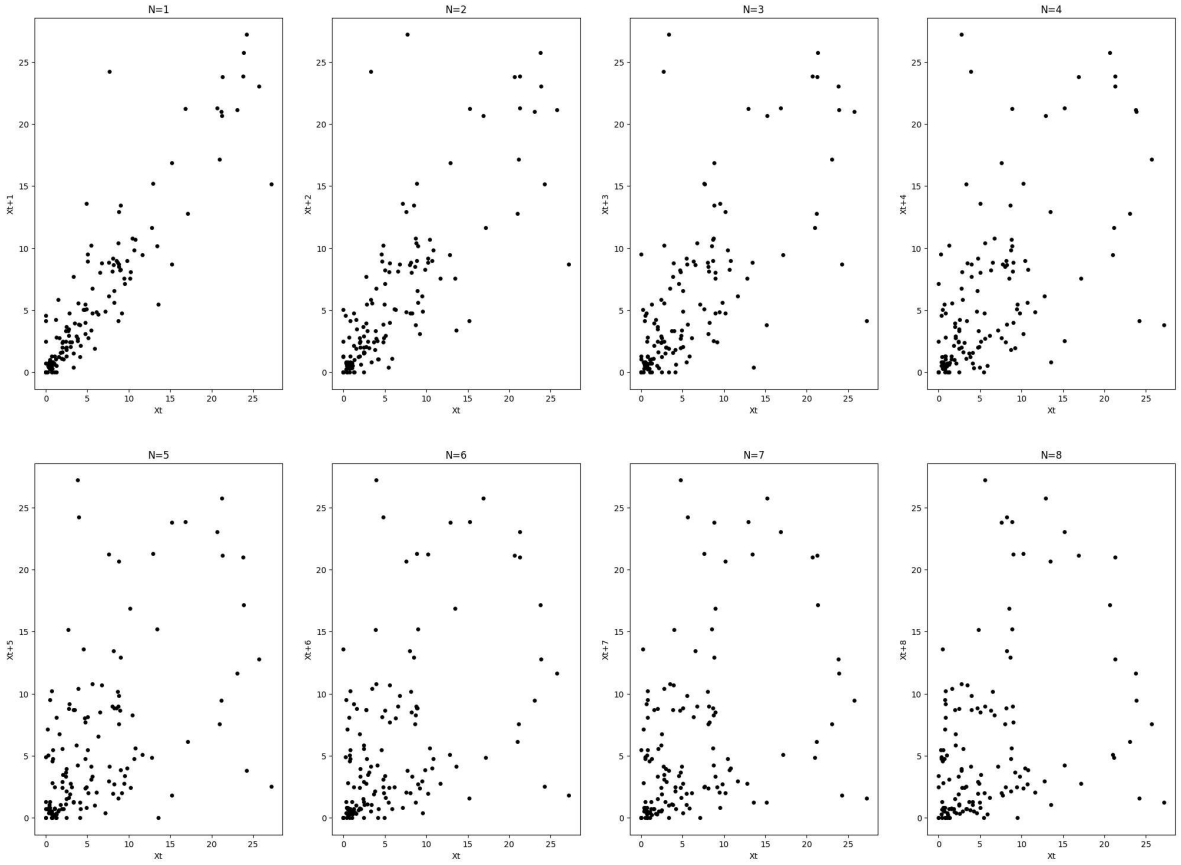
In [12]:
```
plt.figure(figsize=[25,7])
plt.plot(df[metrique], color='black')
plt.xlabel(period)
plt.ylabel(metrique)
plt.show()
```



Question 3

In [13]:
```
###NX
plt.figure(figsize=[25,18])
```

```
for N in range(1,9):
    plt.subplot(2,4,N)
    debut=0
    fin=df.shape[0]
    plt.scatter(df[metrique][debut:fin-N], df[metrique][debut+N:fin], color='bla
    plt.xlabel("Xt")
    plt.ylabel(f"Xt+{N}")
    plt.title(f"N={N}")
plt.show()
```



Question 4

In [14]:
```
#Calcul l'auto-Covariance empirique
def auto_cov(data, K, moy):
    debut=0
    fin=len(data)
    Xt=data[debut:fin-K]
    Xt_k=data[debut+K:fin]

    cov=0
    for i in range(fin-K):
        cov+=(Xt[i]-moy)*(Xt_k[i]-moy)
    return cov/(fin-K)
```

In [15]:
```
def auto_cor(data, K):
    moy=np.mean(data)
    cov_0=auto_cov(data, 0, moy)
    cov_K=auto_cov(data, K, moy)
    return cov_K/cov_0
```

In [16]:
```
from tqdm import tqdm
auto_cor_all=list()
```

```
data=list(df[metrique])
for i in tqdm(range(1, 51)):
    auto_cor_all.append(auto_cor(data, i))
```

```
100%|██████████| 50/50 [00:00<00:00, 808.61it/s]
```

In [17]:
```
indexes=[i for i in range(1,51)]
plt.bar(indexes,auto_cor_all, align='edge', width= 0.25, color='black')
plt.xlabel("Ordre")
plt.ylabel("Auto-Correlation")
```

Out[17]:
```
Text(0, 0.5, 'Auto-Correlation')
```