

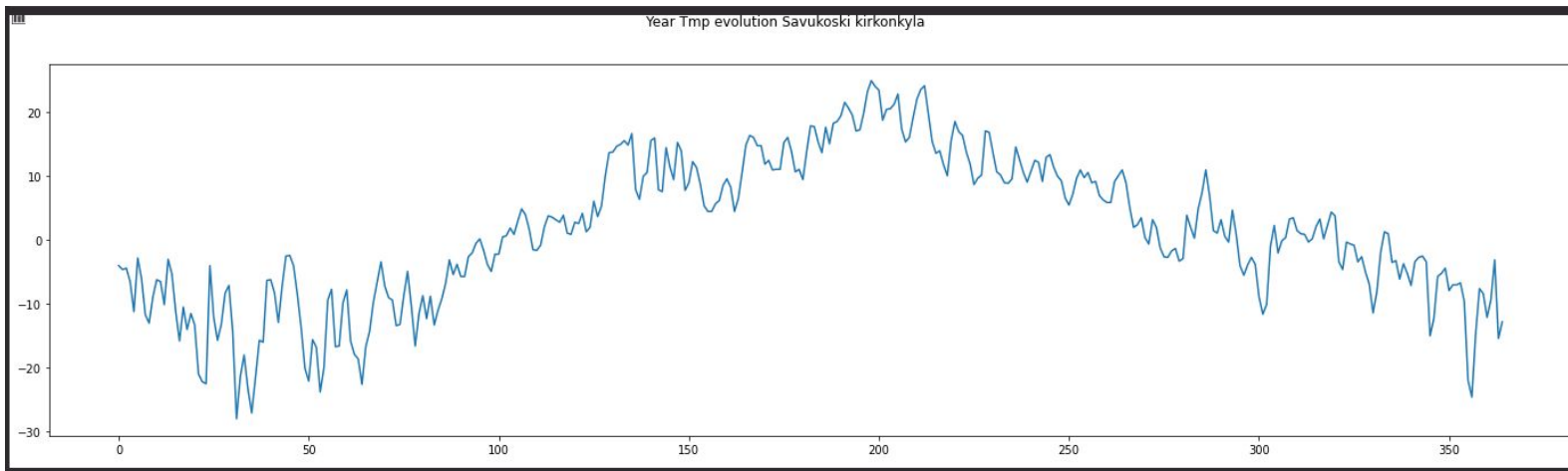
README du groupe 2

Auvray Baptiste, Théophile Choynet, Thomas Le Floch, Yoann Mazza, Sacha Monderer

Contexte :

L'objectif de ce TP est de comparer des jeux de données de climat pour déterminer la capitale européenne d'où viennent les données du fichier Climat-SI.xlsx. Le fichier Savukoskikirkonkyla.xlsx servira de référence.

Ici la température de l'air ambiant est relevée par jour.



Pour réaliser cet objectif, nous aurons besoin de mettre en œuvre un environnement de traitement graphique, ici Google Colab Research, pour pouvoir ensuite corriger le jeu de données traité et proposer un candidat potentiel.

Réalisation:

Comme énoncé plus tôt, pour notre environnement graphique nous avons utilisé Google Colab Research afin de modéliser les données grâce aux bibliothèques Matplotlib et Panda notamment.

D'abord, nous avons réalisé la moyenne et l'écart type par mois du fichier Climat-SI.xlsx :

```
round(data.mean()) #moyenne par mois
```

janvier	-9.0
février	-5.0
mars	1.0
avril	8.0
mai	13.0
juin	17.0
juillet	18.0
août	18.0
septembre	10.0
octobre	4.0
novembre	-0.0
décembre	-7.0
dtype:	float64

Puis le min/max par mois et par année :

```
▶ data.min()
```

```
janvier      -23.0  
février      -12.0  
mars         -8.0  
avril        2.0  
mai          5.0  
juin         11.0  
juillet      13.0  
août         14.0  
septembre    6.0  
octobre      1.0  
novembre     -8.0  
décembre     -11.0  
dtype: float64
```

```
[ ] data.min().min()
```

```
-23.0
```

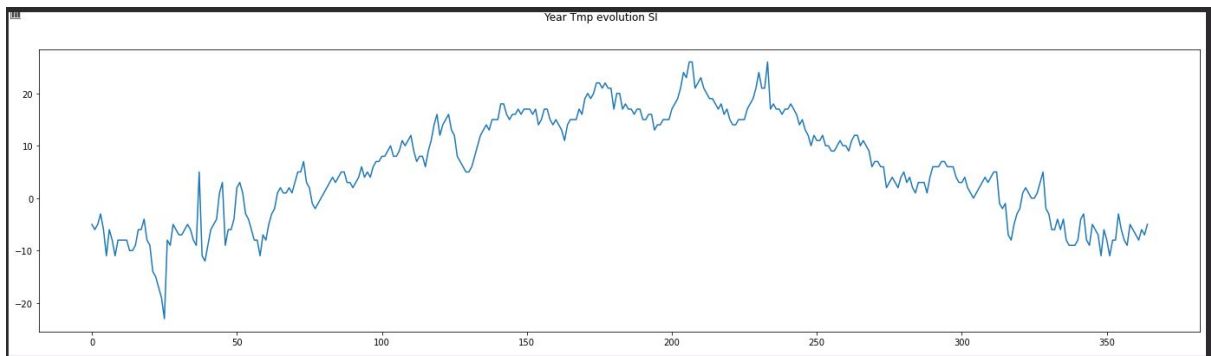
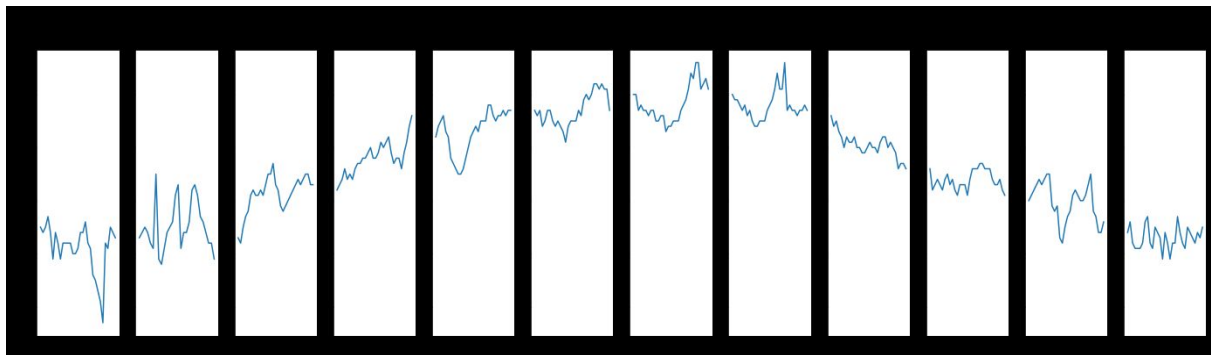
```
[ ] data.max()
```

```
janvier      -3.0  
février      5.0  
mars         7.0  
avril        16.0  
mai          18.0  
juin         22.0  
juillet      26.0  
août         26.0  
septembre    16.0  
octobre      7.0  
novembre     5.0  
décembre     -3.0  
dtype: float64
```

```
[ ] data.max().max()
```

```
26.0
```

Après nous traçons la courbe des températures en vue mensuel et en vue annuel :



Une fois toutes ces opérations effectuées nous recommençons avec le jeu de données Climat-SI-erreur.xlsx qu'il va falloir corriger dans un premier temps :

Pour la correction des données, nous avons choisi, pour les données corrompues ou erronées (différence supérieur ou égal à 15° C), de faire la moyenne des températures du jour d'avant et d'après. Nous effectuons ensuite les mêmes traitements que pour le fichier Climat-SI-erreur.xlsx :

Données corrigées :

```
print(totalyearerr)
```

```
[-5, -6, -5, -3, -6, -11, -6, -8, -11, -8, -8, -8, -8, -10, -10, -9, -6, -6, -4, -8, -9, -14, -15, -17, -19, -23, -8, -9, -5, -6, -7, -7.0, -6.0, -5.0, -6.0, -8.0, -9.0, 5.0, -3.5, -12.0, -9.0, -6.0, -5.0, -4.0, 1.0, 3.0, -9.0, -6.0, -6.0, -4.0, 2.0, 3.0, 1.0, -3.0, -4.0, -6.0, -8.0, -8.0, -11.0, -11.0, -11.0, -11.0, -7.0, -8.0, -5.0, -3.0, -2.0, 1.0, 2.0, 1.0, 1.0, 2.0, 1.0, 3.0, 5.0, 5.0, 7.0, 4.5, 2.0, -1.0, -2.0, -1.0, 0.0, 1.0, 2.0, 3.0, 4.0, 3.0, 4.0, 5.0, 5.0, 3.0, 3.0, 2.0, 3.0, 4.0, 6.0, 4.0, 5.0, 4.0, 6.0, 7.0, 7.0, 8.0, 8.0, 9.0, 10.0, 8.0, 8.0, 9.0, 11.0, 10.0, 11.0, 12.0, 9.0, 7.0, 8.0, 8.0, 6.0, 9.0, 11.0, 14.0, 16.0, 16.0, 12, 14, 15, 16, 13, 12, 8, 7, 6, 5, 5, 6, 8, 10, 12, 13, 14, 13, 15, 15, 15, 18, 18, 16, 15, 16, 16, 17, 16, 17, 17, 17.0, 16.0, 17.0, 14.0, 15.0, 17.0, 17.0, 15.0, 14.0, 14.0, 14.0, 13.0, 11.0, 14.0, 15.0, 15.0, 15.0, 17.0, 16.0, 19.0, 20.0, 19.0, 20.0, 22.0, 22.0, 21.0, 22.0, 21.0, 21.0, 17.0, 17.0, 20.0, 20.0, 17.0, 18.0, 17.0, 17.0, 16.0, 17.0, 17.0, 15.0, 15.0, 16.0, 16.0, 13.0, 14.0, 14.5, 15.0, 15.0, 15.0, 17.0, 18.0, 19.0, 21.5, 24.0, 23.0, 26.0, 26.0, 21.0, 22.0, 23.0, 21.0, 20, 19, 19, 18, 17, 18, 16, 17, 15, 14, 14, 15, 15, 15, 16.5, 18, 19, 21, 24, 21, 21, 26, 17, 18, 17, 17, 16, 17, 17, 18, 17, 16.0, 14.0, 15.0, 13.0, 12.0, 10.0, 12.0, 11.0, 11.0, 12.0, 10.0, 10.0, 9.0, 9.0, 10.0, 11.0, 10.0, 10.0, 9.0, 11.0, 12.0, 12.0, 10.0, 11.0, 10.0, 9.0, 6.0, 7.0, 7.0, 6.0, 6.0, 6.0, 2.0, 3.0, 4.0, 3.0, 2.0, 4.0, 3.5, 3.0, 4.0, 2.0, 1.0, 3.0, 3.0, 3.0, 1.0, 4.0, 6.0, 6.0, 6.0, 7.0, 7.0, 6.0, 6.0, 6.0, 4.0, 3.0, 3.0, 4.0, 2.0, 1.0, 0.0, 1.0, 2.0, 3.0, 4.0, 3.0, 4.0, 5.0, 5.0, -1.0, -2.0, -1.0, -7.0, -8.0, -5.0, -3.0, -2.0, 1.0, 2.0, 1.0, 0.0, 0.0, 1.0, 3.0, 5.0, -2.0, -3.0, -6.0, -6.0, -4.0, -4.0, -6, -4, -8, -9, -9, -9, -8, -4, -3, -8, -9, -5, -6, -7, -11, -6, -8, -11, -8, -8, -7, -6, -8, -9, -5, -6, -7, -8, -6, -7, -5]
```

Moyenne de l'année :

```
[77] statistics.mean(totalyearerr)

5.661290322580645
```

Min/Max de l'année:

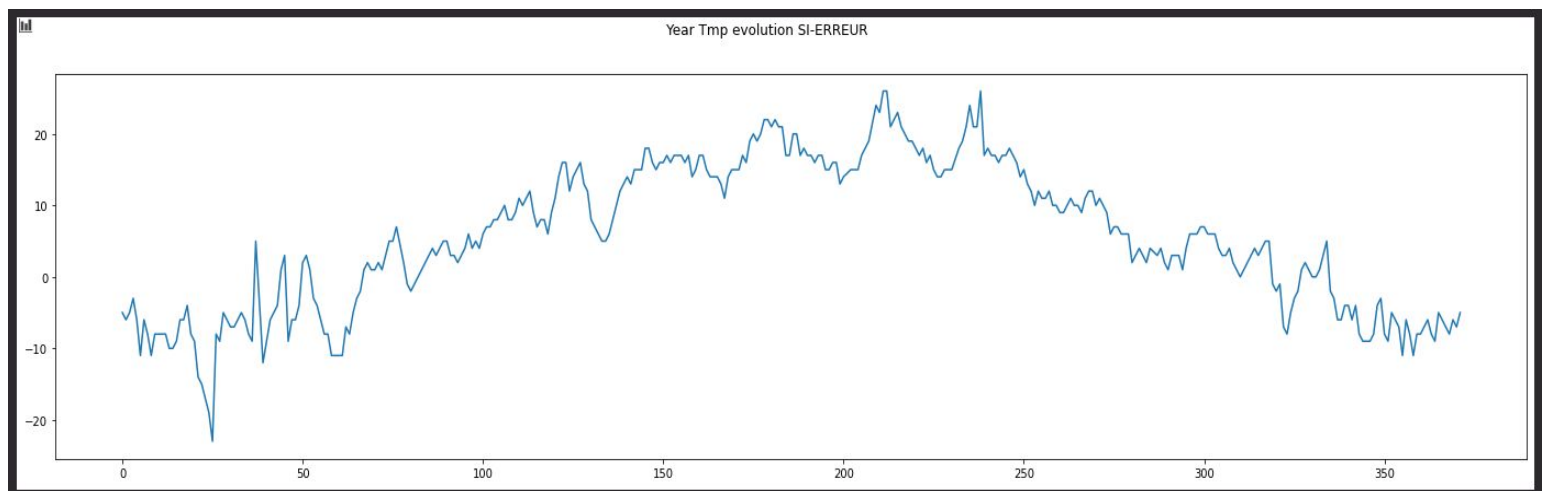
```
[79] min(totalyearerr)

-23

[80] max(totalyearerr)

26.0
```

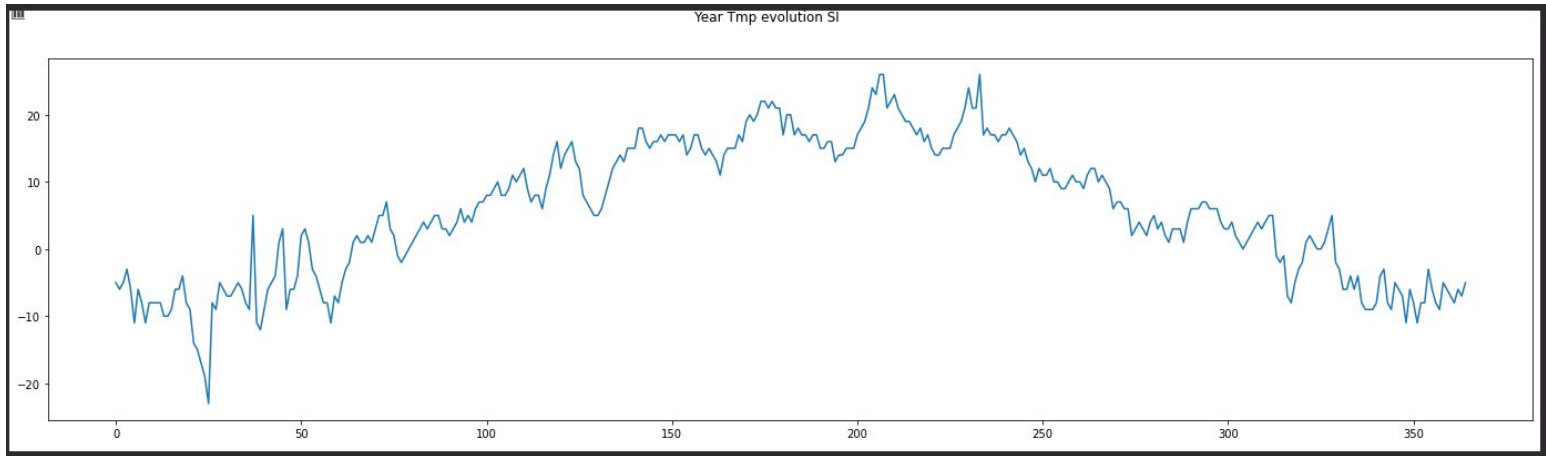
Courbe des températures :



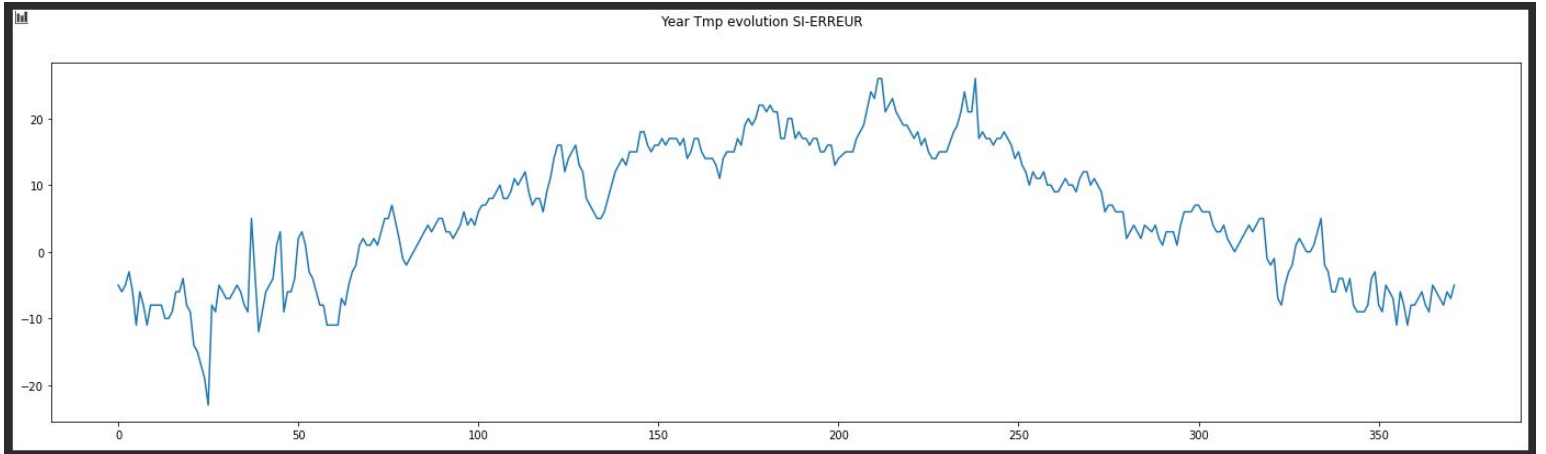
Conclusion :

En comparants les graphiques des 2 jeux de données (Climat-SI et Climat-SI-erreur) :

Climat-SI :



Climat-SI-erreur :



On remarque que les données corrigées concordent bien avec les données réelles.

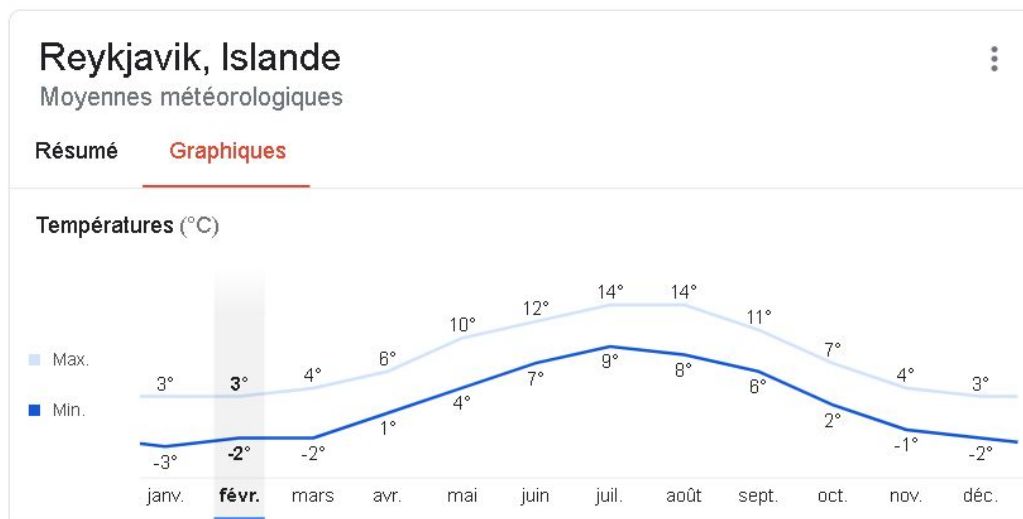
Température moyenne Helsinki :



Température moyenne Stockholm :



Température moyenne Reykjavik :



Notre jeux de données correspond davantage au graphique des températures de la ville d'Helsinki. Nous en déduisons donc que Helsinki est la capitale du jeu de données Climat-Si et Climat-Si-erreur.