

TP3 - Statistique température

Le but de ce travail est de générer des statistiques à partir des fichiers de données brutes générées par un ensemble de capteurs.

Pour simuler les données générées par les capteurs, nous allons créer un premier job qui crée les données brutes. Les fichiers générés doivent tous avoir la structure du nom suivante :

`"capture_'+capture_id+'.csv"`(Image 1).



Image 1

les fichiers sont des fichiers csv avec la structure suivante (image 2):

```
temperature;datetime
94;10-01-2022 17:07:33
90;10-01-2022 17:07:33
55;10-01-2022 17:07:33
52;10-01-2022 17:07:33
53;10-01-2022 17:07:33
```

Image 2

le second job que nous allons créer, lira tous les fichiers générés et générera deux fichiers, le premier est un fichier csv nommé output.csv (image 3) qui contiendra les statistiques de tous les capteurs (nom, max, min et moyenne), le second est nommé output_error.csv qui contiendra toutes les valeurs erronées (<-10 et >70) à la fin du job, afficher une fenêtre avec le message "job terminé" (image 5)

```
cap;avg;max;min  
capteur_1;31;69;-9  
capteur_10;25;69;-9  
capteur_100;26;69;-9  
capteur_11;27;69;-9
```

Image 3

```
cap;temperature  
capteur_1;-14  
capteur_1;81  
capteur_1;79  
capteur_1;92
```

Image 4

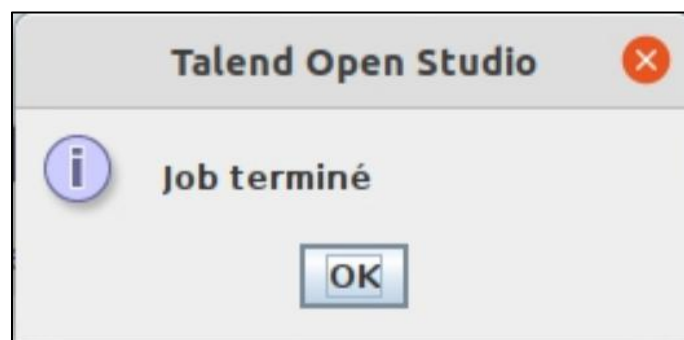


Image 5

changer le deuxième job pour générer un nouveau fichier output_error_stat.csv qui contient le nombre de valeurs aberrantes pour chaque capteur.