

Données et métadonnées : mise en forme et utilisation des données téléchargées sur le site de l'observatoire Edumed

Mai 2020

Fabrice Mourau ^(1,2)

1 : Collège Pierre de Coubertin, Education Nationale

2 : Observatoire Edumed, IDEX-UCA, OCA, UMR-Géoazur

Téléchargement et format des données EDUMED

Le Data center de l'observatoire Edumed propose la visualisation graphique des données issues des instruments du réseau mais aussi le téléchargement de séries temporelles dans l'onglet « Données ». L'utilisateur du site a ainsi la possibilité de directement travailler avec ses élèves sur les valeurs chiffrées ou de créer une étude de cas qui sera utilisée en cours (Figure 1).

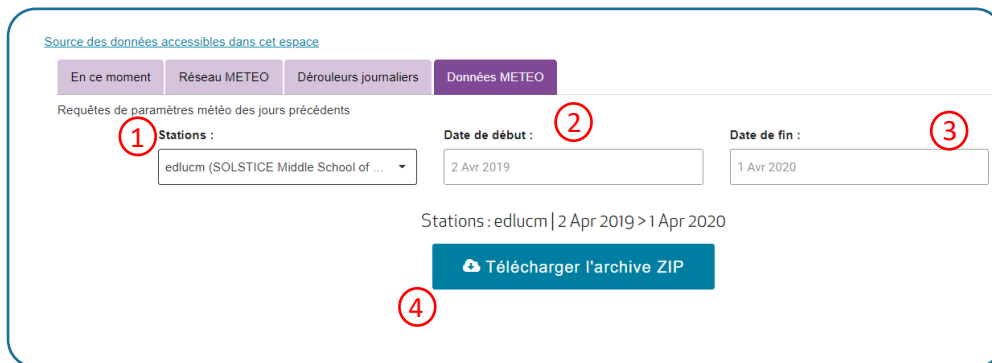


Figure 1 : L'interface de l'onglet « Données » du site EDUMED permet de télécharger une série temporelle sous forme d'archive compressée. On sélectionne la (ou les) station(s) météo (1), puis la date de début (2) et la date de fin (3) de série temporelle. Enfin, on clique sur l'icône Télécharger l'archive ZIP (4).

Les données téléchargées sont écrites dans en CSV (Comma Separated Value), un format standard qui peut être lu par n'importe quel logiciel tableur comme Microsoft Excel, LibreOffice Calc ou Apple Numbers. Dans ce format, le tableau de données est écrit sous la forme d'un fichier texte. Une ligne de texte correspond à une ligne du tableau et chaque colonne est séparée de la suivante par un séparateur. Dans le cas des fichiers Edumed, le séparateur est un **point-virgule** (figure 2).

```
01;04;2020
02;04;2020
03;04;2020
```

Figure 2 : Exemple d'un fichier CSV tel qu'il est écrit (en haut) et comment il est lu par un tableur (en bas)

01	04	2020
02	04	2020
03	04	2020

Données et métadonnées

A titre d'exemple, on a téléchargé les données de la station météo du collège de Simiane pour la journée du 1^{er} avril 2020 (Figure 3). On voit que les données sont accompagnées d'informations leur qualité, la station elle-même et sur la façon dont elles sont écrites : ce sont les métadonnées.

date_sample	tempout	humout	patmo	rainday	windvel	winddir	solarrad	windex	windchill	heatindex	pbare	rainratehou	windvelavg10	et	
01/04/2020 00:01	282.45	79	None	0.0000	1.79	81	None	None	281.45	282.55	279.25	101890.00	0.0000	1.79	None
01/04/2020 00:03	282.45	79	None	0.0000	2.24	64	None	None	281.45	282.55	279.25	101890.00	0.0000	1.79	None
01/04/2020 00:05	282.45	80	None	0.0000	0.89	97	None	None	282.05	282.55	279.25	101890.00	0.0000	1.79	None
01/04/2020 00:07	282.45	80	None	0.0000	1.34	80	None	None	282.05	282.55	279.25	101900.00	0.0000	1.34	None
01/04/2020 00:09	282.45	80	None	0.0000	1.34	47	None	None	282.05	282.55	279.25	101890.00	0.0000	1.34	None
01/04/2020 00:11	282.45	80	None	0.0000	0.89	84	None	None	282.05	282.55	279.25	101890.00	0.0000	1.34	None
01/04/2020 00:13	282.35	80	None	0.0000	0.89	82	None	None	282.05	282.55	279.25	101900.00	0.0000	1.34	None
01/04/2020 00:15	282.35	79	None	0.0000	1.34	118	None	None	282.05	282.55	278.75	101900.00	0.0000	1.34	None
01/04/2020 00:17	282.35	80	None	0.0000	3.13	90	None	None	280.95	282.55	279.25	101890.00	0.0000	1.34	None

Métadonnée

Comment, par qui et avec quelle précision

Les valeurs mesurées par la station

Données

Figure 3 : Les données météo demandées par l'utilisateur sont accompagnées de métadonnées qui nous informe sur les données elles-mêmes.

Les métadonnées sont très importantes, elles seules vont permettre une utilisation rigoureuse des mesures réalisées par la sonde. Nous conseillons de créer une nouvelle feuille de calcul dans le fichier nommée *Metadata*, de couper et de coller les métadonnées à l'intérieur (figure 4). A ce niveau, vous allez devoir enregistrer votre document sous un autre format que le .csv qui n'accepte pas plusieurs feuilles dans un même fichier : Fichier → Enregistrer sous. Utilisez le format .xlsx pour un document Microsoft Excel ou .ods pour un LibreOffice Calc. (Figure 4).

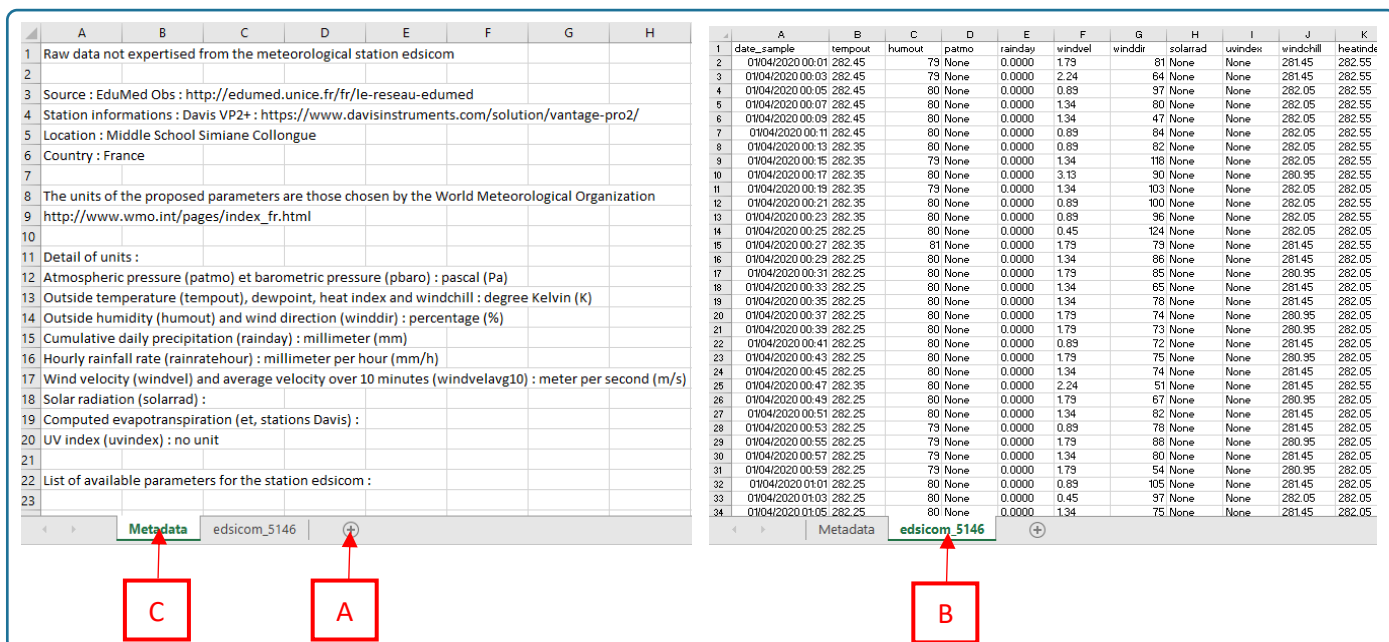


Figure 4 : Pour créer une feuille de calcul, cliquer sur le symbole + (A), puis double cliquer sur l'onglet pour le renommer. Les données (B) et métadonnées (C) sont ensuite réparties sur 2 feuilles de calcul dans le fichier enregistré aux formats xlsx ou ods.

Mise en forme du tableau de valeur

En observant la feuille du tableau de données, on voit que les nombres entiers sont écrits à droite de leur cellule alors que les nombres décimaux sont écrits à gauche. C'est une conséquence de l'utilisation des fichiers csv : la virgule et le point-virgule sont utilisés comme séparateur de colonne alors que les décimales sont notées avec un point : c'est le format anglo-saxon. Pour mettre en forme correctement le fichier, dans Calc il suffit d'utiliser la commande Rechercher et remplacer du menu Edition. On recherche les « . » et on les remplace par des « , ». Dans Excel 2019, la commande est située dans la loupe à l'extrême droite du bandeau « Accueil » (figure 5).

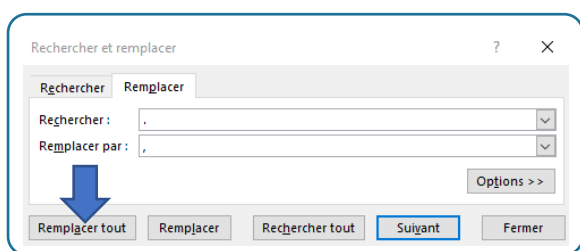


Figure 5 : Le menu rechercher et remplacer permet de remplacer tous les marqueurs décimaux « . » par des virgules en cliquant sur « Remplacer tout ».

Exemple d'une activité pour le cycle 4 (Cinquième)

Dans l'exemple ci-après, on veut demander à des élèves de cycle 4 de tracer un diagramme qui représente les variations de la hauteur d'eau de la rivière Issole, suite à un épisode méditerranéen. En fonction de l'âge des élèves, des acquis dans l'usage de l'outil tableur et des compétences que l'on veut évaluer, les consignes données aux élèves pourront être très graduées. Ici, les données « brutes » issues du site ont déjà été mises en forme et l'enseignant va directement utiliser une étude de cas.

Place dans le programme :

L'exemple qui suit s'insère dans la leçon sur le risque inondation, il s'adresse à des élèves de cinquième qui découvrent l'interface EDUMED. L'apprentissage de l'outil s'intègre à une approche spiralaire, car il sera réutilisé en quatrième (risque sismique et risque volcanique), en troisième (paléoclimat, paléogéographie, évolution), au lycée puis à l'université. Nous avons organisé identiquement chaque item du Data center afin que les élèves puissent utiliser l'interface web Edumed-obs du collège à la L3 en réinvestissant les acquis des années précédentes. C'est un gain de temps et d'efficacité énorme pour l'élève et l'enseignant qui illustre parfaitement l'intérêt de centraliser les données sur une même plateforme.

Planète Terre 4A : Variations de la hauteur d'eau de l'Issole, suite à un épisode méditerranéen

Objectif : tracer grâce à LibreOffice Calc un diagramme en colonne et lignes qui représente les précipitations et les variations de hauteur d'eau de l'Issole pendant un épisode méditerranéen.

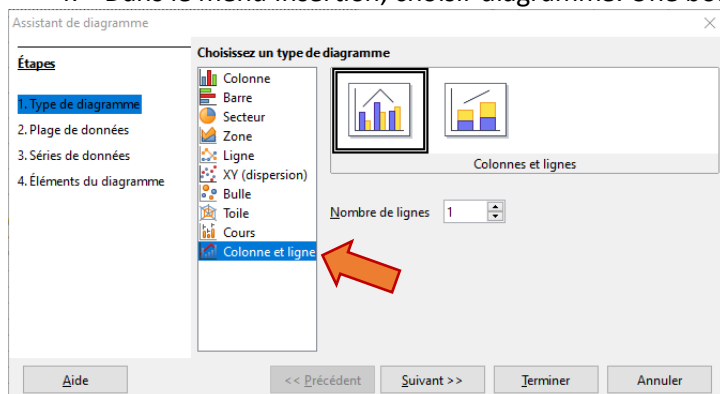
- ➔ A partir du navigateur internet Chrome, ouvrir la page web de l'observatoire EDUMED : edumed.unice.fr
- ➔ Ouvrir le Data center HYDRO, puis cliquer sur l'onglet « Etudes de cas HYDRO »
- ➔ Télécharger le fichier « Variations de la hauteur d'eau de l'Issole, suite à un épisode méditerranéen » et ouvrir le fichier DATA_ISSOLE2019.ods à l'aide de LibreOffice CALC.

A partir des informations de l'onglet « métadonnées », trouve les informations suivantes et note-les sur ta fiche :

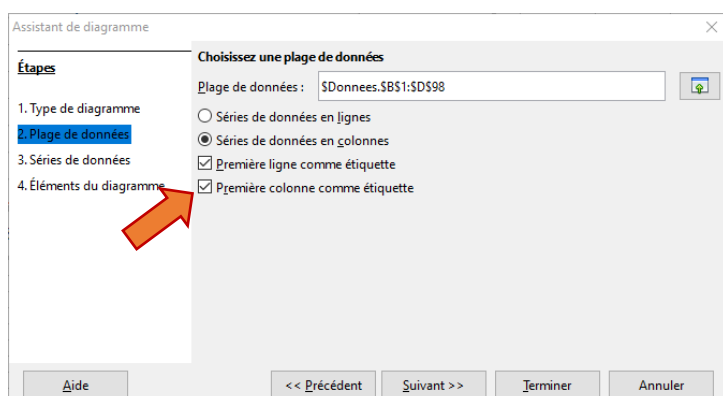
1. Quel est la source des données météo ?
2. Quel est le nom de la station météo ?
3. Dans quelle commune est-elle située ?
4. Quelle unité est utilisée pour mesurer les précipitations ?
5. Quelle est la source des données hydrologiques ?
6. Quel est le nom de la rivière instrumentée ?
7. Quelle unité est utilisée pour mesurer la hauteur d'eau ?

Dans l'onglet « données » :

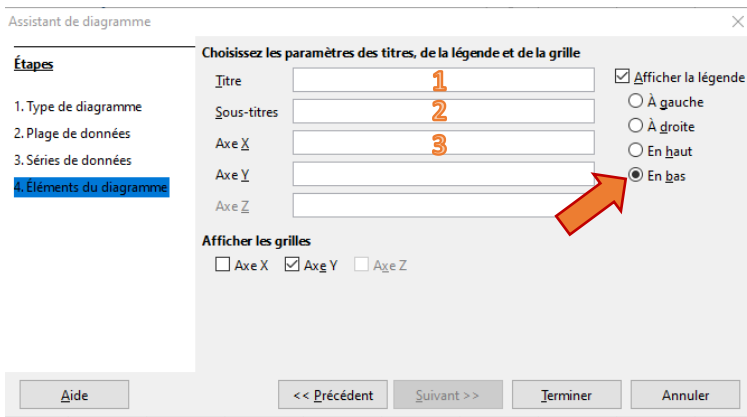
1. Noter le jour de début des mesures :
2. Noter le jour de fin des mesures :
3. Sélectionner les colonnes B, C et D
4. Dans le menu Insertion, choisir diagramme. Une boîte de dialogue doit s'ouvrir :



Choisir un diagramme en colonne et lignes, puis cliquer sur suivant.



Cliquer sur la case « Première colonne comme étiquette », afin que les données de la première colonne sélectionnée (les heures) soient affichées sur l'axe des abscisses (l'axe horizontal que le logiciel appellera axe X), puis cliquer sur suivant



Choisir d'afficher la légende En bas, puis écrire le Titre (1), les sous-titres (2) et la légende de l'axe X (l'axe horizontal) (3). Puis, choisir terminer.

Aide méthodologique 1

(A fournir séparément, en fonction des acquis des élèves et dans un but de différenciation pédagogique. L'évaluation des acquis tiendra compte de l'utilisation ou non de cette aide)

1 : Ecrire le titre.

Quand on étudie des séries temporelles, le titre commence par « Evolution », puis on cite les variables* représentées sur le diagramme et le lieu de mesure de chacune (Rappelle-toi, les lieux de mesure sont écrits dans les métadonnées).

2 : Ecrire les sous-titres :

Dans le sous-titre, on précise les dates de début et de fin de la série temporelle et la source des données entre parenthèses. Ex (sce :)

3 : Ecrire la légende d'un l'axe :

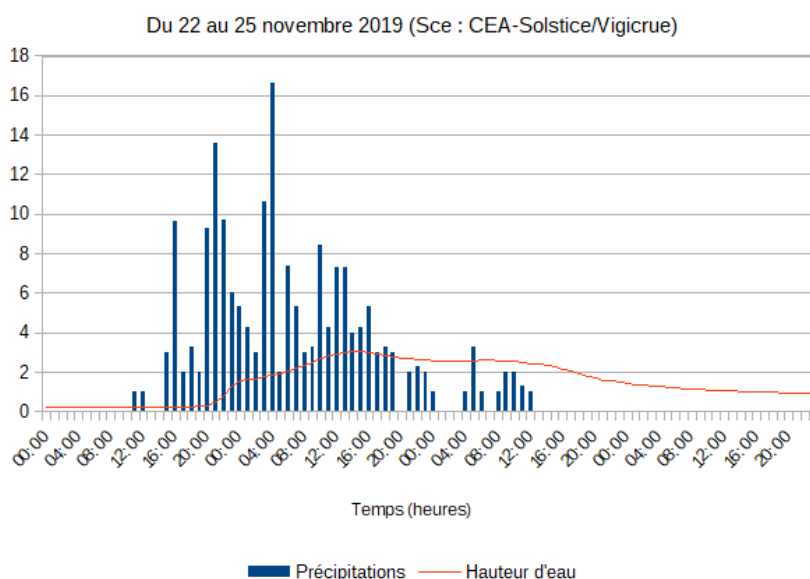
La légende de l'axe indique le nom de la variable mesurée et l'unité choisie pour la mesurer. Ici, le Temps est mesuré en heures. Il faut donc écrire Temps (Heures).

*Variable : c'est une grandeur mesurée et dont la valeur varie au cours du temps. On trouve le nom des variables dans les entêtes de colonne, ici ce sont les cumuls de précipitations et la hauteur d'eau.

Aide méthodologique 2

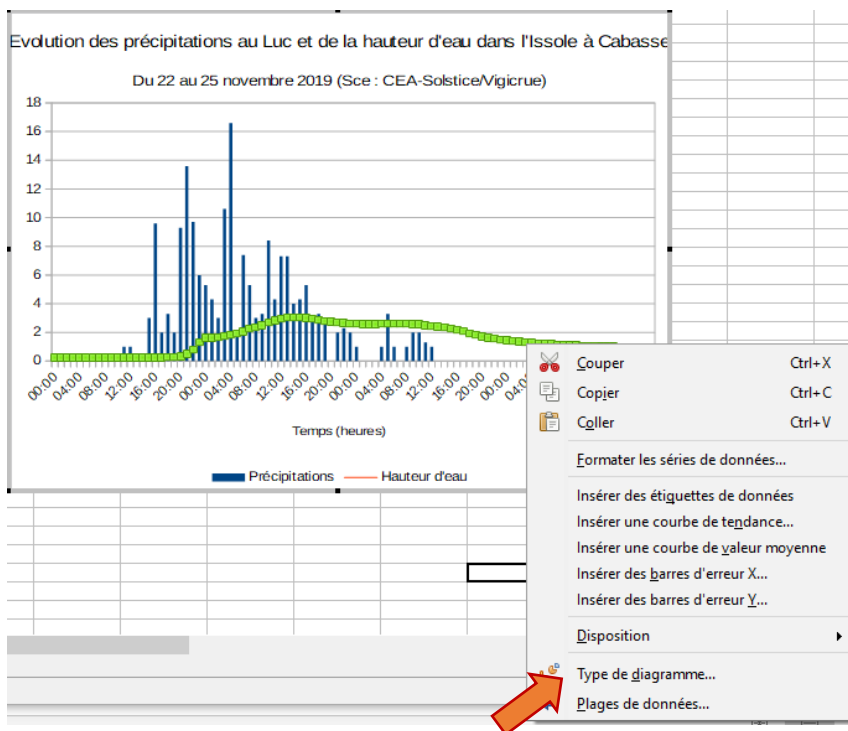
(A fournir séparément : chaque élève doit réussir à produire un document, on peut donc fournir aux élèves les plus en difficulté un document à recopier. Il lui suffit alors d'identifier les bons items sur le logiciel et de recopier l'exemple fourni par l'enseignant. Bien entendu, l'évaluation tiendra compte de l'utilisation de cette aide).

Evolution des précipitations au Luc et de la hauteur d'eau dans l'Issole à Cabasse



Mettre en forme du diagramme et finaliser les légendes des axes

On voit que le diagramme est incomplet, il va falloir ajouter un axe secondaire et finir de légender les axes.



1. Double cliquer sur la zone de diagramme et sélectionner la courbe en cliquant directement sur elle (tous les points de construction de la courbe deviennent visibles en vert). Une fois la courbe sélectionnée, faire un clic droit. Une boîte de dialogue s'ouvre et choisir « Formater les séries de données ». Dans l'onglet « Option », choisir axe Y secondaire et valider.

2. Il suffit maintenant de sélectionner le bouton « Titres » et de compléter la légende des axes Y primaires et secondaires.

La correction est fournie aux élèves, elle leur permettra dans la leçon suivante de lire les hauteurs d'eau maximales atteintes et de calculer le temps mis par la rivière pour se mettre en charge après les premières précipitations. Un travail à partir de la carte topographique, permettra de définir les zones de la communes mises en danger par l'élévation du niveau de l'Issole. C'est une introduction à la carte de l'aléa...

Evolution des précipitations au Luc et de la hauteur d'eau dans l'Issole à Cabasse

