

Données structurées : Fichiers CSV

Travaux Pratiques, Seconde SNT

Définition 1 — Fichier CSV

Un fichier d'extension `.csv` (*Comma-Separated Values*) est un fichier texte permettant de stocker des données sous forme de tableau. Chaque ligne correspond à un enregistrement et les différentes valeurs sont séparées par un caractère spécial, souvent un point-virgule " ; " ou une virgule " , ".

Ces fichiers peuvent être ouverts avec un simple éditeur de texte, mais aussi avec un tableur comme **LibreOffice Calc**.

Consignes générales

Ce TP est à réaliser en binôme. Penser à **sauvegarder régulièrement** son travail dans le dossier TP_CSV. Les réponses écrites sont à rédiger directement sur cette feuille.

NOM : Prénom :

NOM : Prénom :

N° de groupe :

Partie I : Découverte du format CSV

1) Lire la définition et la consigne au-dessus, puis inscrire vos noms.

Dans cette partie, on va manipuler un fichier `.csv` à l'aide d'un éditeur de texte.

2) Créer un dossier TP_CSV (peu importe où) sur votre ordinateur, **télécharger** le fichier `pokemon.csv` à l'adresse suivante : <https://nsi.mdehais.fr/SNT/donnees/> et **déplacer** ce fichier dans le dossier TP_CSV.

3) Ouvrir Visual Studio Code, puis **aller** dans *Fichier > Ouvrir un dossier* et **sélectionner** le dossier TP_CSV. Vous voyez désormais le fichier `pokemon.csv` dans l'explorer VsCode.

4) Observer le fichier `pokemon.csv`. La **toute première ligne** contient les **descripteurs** (aussi appelés en-têtes) : ce sont les noms de chaque colonne. **Recopier** ci-dessous ces descripteurs :

--

5) Chaque ligne suivante (dans le fichier) correspond à un Pokémon. En s'aidant des descripteurs, **compléter** le tableau ci-après pour les Pokémon attribués selon le numéro de groupe. **La répartition des pokemons attribuée à votre groupe est disponible à l'adresse suivante** <https://nsi.mdehais.fr/SNT/donnees/>.

💡 Astuce

Pour trouver vos pokemons dans le fichier, il vous faudra utiliser le raccourci **Ctrl+F** dans VsCode, qui permet de chercher un mot dans un document.

Nom	Type1	Type2	PV	Attaque	Defense	Att. Spé	Def. Spé	Vitesse	Légitimaire

6) Nous savons que le pokemon **Carapuce** porte le numéro 7, qu'il est uniquement de *type Eau*, que sa **vitesse** vaut 43, que sa **défense spéciale** vaut 64, qu'il a 44 **points de vie**, que son **attaque** vaut 48, qu'il n'est **pas légendaire**, que son **attaque spéciale** vaut 50 et que sa **défense** vaut 65. **Modifier** le fichier `pokemon.csv` pour rajouter **Carapuce** au bon endroit. Penser à bien **sauvegarder**!

7) Dans un fichier CSV, les valeurs de chaque ligne sont délimitées par un **caractère séparateur** qui permet de distinguer une colonne de la suivante. En observant le fichier, **identifier** le séparateur utilisé ici. **Quel problème** cela pourrait-il poser si les noms des Pokémon contenaient ce même caractère ?

Partie II : Analyse des données

La puissance d'un Pokémon peut s'évaluer grâce à ses statistiques : PV, Attaque, Défense, Attaque Spéciale, Défense Spéciale et Vitesse.

1) Compléter le tableau suivant en calculant à la main la somme et la moyenne des six statistiques de chacun des Pokémon de votre groupe.

Nom du Pokémon	Somme des statistiques	Moyenne des statistiques

2) Parmi les Pokémon de votre groupe, lequel est le **plus puissant** selon la somme de ses statistiques ?

.....

.....

3) Parmi les Pokémon de votre groupe, y a-t-il au moins un Pokémon **Léendaire** ? Si oui, lequel ?

.....

.....

4) Pour chacun des Pokémons de votre groupe, **identifier** sa **meilleure statistique** (la plus élevée).

Nom du Pokémon	Meilleure statistique	Valeur de la stat

5) Trouvez-vous pratique d'analyser ces données dans un éditeur de texte ? Vous semble-t-il envisageable de calculer la moyenne des HP des 700 pokemons du fichier?

Partie III : En route vers le tableur !

Définition 2 — Tableur : LibreOffice Calc

Un **tableur** est un logiciel permettant de lire, afficher et traiter des données structurées comme les fichiers CSV. Les données sont organisées dans des **cellules**, identifiées par une lettre (colonne) et un nombre (ligne). Par exemple, la cellule **B3** se trouve à la colonne B, ligne 3.

Une cellule peut contenir une valeur ou une **formule**. Toute formule commence par le signe =.

1) Lancer l'application **LibreOffice Calc**. Dans le menu *Fichier*, **cliquer** sur *Ouvrir* puis **sélectionner** le fichier `pokemon.csv`.

Fenêtre d'import CSV

LibreOffice affiche une fenêtre avant d'ouvrir le fichier. Dans la section *Options de séparateur*, il faut cocher le bon séparateur. Laisser les autres options par défaut et cliquer sur **OK**.

2) Comparer l'affichage avec ce que l'on avait dans VS Code. Vérifier que **Carapuce** est bien présent. Que peut-on dire de l'aspect visuel des données ?

3) Voilà la ligne correspondant au pokemon **Salamèche** qui aurait du se trouver dans le fichier `pokemon.csv` : "4,Salamèche,Feu,,39,52,43,60,50,65,False". **Ajouter** la ligne correspondante grâce à Libre Office Calc en faisant un clic droit à la bonne ligne puis *insérer* pour créer une ligne vide puis rentrer les bonnes informations au bon endroit.

4) Compléter le tableau suivant avec les **noms de cellules** correspondants (exemple de format : B3), les informations portent sur les pokemons de votre groupe.

Information recherchée	Cellule
PV du premier Pokémon	
Nom du quatrième Pokémon	
Vitesse du dernier Pokémon	
Type1 du cinquième Pokémon	

5) Dans la cellule vide juste après la colonne **Légendaire** (sur la ligne du premier Pokémon), saisir la formule `=SOMME(` puis sélectionner les six statistiques du premier pokemon avec la souris. Valider avec Entrée.

Quelle formule exacte a été écrite **si vous appuyez une seule fois sur la case où vous venez d'écrire**.

6) En faisant un **cliquer-glisser sur le petit carré bleu** en bas à droite de la cellule, puis faites défiler jusqu'en bas. À quoi sert cette manipulation ?

Astuce

Vous pouvez faire de grandes sélections grâce à un **Shift Clic**.

7) Choisir un nom de colonne adapté pour cette nouvelle colonne et le saisir dans la cellule d'en-tête.

8) Utiliser la fonction `=MAX()` dans la ligne vide en dessous des données pour obtenir la **valeur maximale** de chacune des six statistiques (une formule par colonne).

9) Utiliser la fonction `=MOYENNE()` pour calculer la moyenne des statistiques de chaque Pokémon dans une nouvelle colonne.

10) Sélectionner l'ensemble des données (de la cellule A1 jusqu'à la dernière cellule remplie). Dans le menu **Données**, cliquer sur **Trier**.

11) Trier les données de façon à ce que le Pokémon ayant la **somme de statistiques la plus élevée** apparaisse en premier.

12) En réutilisant les tris, compléter les tableaux suivants, dont les informations portent sur l'intégralité des Pokemons

Meilleurs PV	Meilleure Attaque	Meilleure Défense

Meilleure Attaque speciale	Meilleure Défense speciale	Meilleure Vitesse

13) Le tableur permet également de **filtrer** les données. Dans le menu **Données**, sélectionner **AutoFiltre**. Des flèches apparaissent sur chaque colonne. Utiliser le filtre sur la colonne Type1 pour n'afficher que les Pokémon de type **Fée**.

Combien de Pokémon de type Feu contient ce fichier ?

14) En comparant l'expérience du tableur avec celle de VS Code, quels avantages trouvez-vous à l'utilisation d'un tableur pour travailler avec des fichiers CSV ?

.....

.....

.....

.....