

Thème 4 : Les données structurées

TP3 : Exploiter une base de données

Une base de données connue : Le Pokédex



Le **Pokédex** est une encyclopédie recensant tous les **Pokémon** existants et permettant d'enregistrer leurs informations. Inventé par le professeur Chen, il est remis en début d'aventure à tout dresseur afin qu'il puisse récolter un maximum de données sur ces créatures.

Cependant, avec l'apparition du *Big Data* et des objets connectés, le professeur n'a plus besoin d'envoyer des enfants de 13 ans parcourir le monde dans le but de récolter des données... Tous les Pokédex sont maintenant connectés et la base de données des Pokémon est complète (mais en anglais, vous allez devoir faire avec !)



Chaque espèce de Pokémon possède des **statistiques de base** fixes suivantes

	Type principal	Le type principal du Pokémon
	Type secondaire	Le type secondaire du Pokémon
	Point de vie (HP)	un Pokémon avec beaucoup de Points de Vie peut recevoir plus d'attaques avant d'être KO
	L'Attaque (Attack)	un Pokémon avec beaucoup d'Attaque fait plus de dégâts lorsqu'il lance des attaques "physiques"
	La Défense (Défense)	un Pokémon avec beaucoup de Défense perd moins de vie lorsqu'il reçoit des attaques "physiques"
	L'Attaque spéciale (Special Att)	un Pokémon avec beaucoup d'Attaque spéciale fait plus de dégâts lorsqu'il lance des attaques "spéciales"
	La Défense spéciale (Special Def)	un Pokémon avec beaucoup de Défense spéciale perd moins de vie lorsqu'il reçoit des attaques "spéciales"
	La Vitesse (Speed)	un Pokémon avec plus de Vitesse que son adversaire lance son attaque en premier

→ **Activité 1** : Ouvrir le fichier pokedex.csv avec un éditeur de texte type notepad++ (ou TextEdit sur Mac) et observer la façon dont sont présentées les données

1. Que représente la première ligne ?

2. Combien de données a-t-on sur chaque ligne ?

3. Comment sont séparées les différentes données sur chaque ligne ?

→ **Activité 2** : Ouvrir le fichier pokedex.csv avec un logiciel de tableur-grapheur (Microsoft Excel, Apple Numbers, Google Sheets, Libre Office Calc...)

4. Où se trouvent les données présentes sur une ligne ?

L'objectif va être d'obtenir une donnée par cellule. Pour cela nous allons convertir les données disponibles.

EXCEL → DONNEES → CONVERTIR

LIBRE OFFICE → DONNEES → TEXTE EN COLONNE



Choisir comme séparateur « VIRGULE » et décocher le reste. On voit un aperçu de ce à quoi va ressembler le fichier une fois converti. Vérifier que chaque donnée sera dans une colonne différente.

Chaque colonne est maintenant utilisée pour un descripteur.

→ **Sauvegarder le fichier (au format CSV ou XLS ou Numbers ou ODS). Vous pouvez utiliser Excel pour trier vos données. Tout sélectionner puis utiliser le bouton « Trier et filtrer » et répondre aux questions ci-dessous :**

3.1 Questions

5. Quel Pokémon est de type « Psychic » et « Fire » ?

6. Quel Pokémon est de type « Dragon » et « Electric » ?

7. Quel Pokémon est de type « Rock » avec plus de 100HP ?

8. Quel est le top 5 des Pokémon hors légendaires ayant le meilleur total des caractéristiques ?

9. Quel est le top 5 des Pokémon ayant le moins bon total des caractéristiques ?

10. Quels sont les Pokémon de type « Grass » avec plus de 100 en Attaque spéciale et Défense spéciale ?

11. Parmi les Pokémon légendaires de la seconde génération, lequel a le plus de Vitesse ?

12. Quel Pokémon électrique non légendaire à toutes ses caractéristiques au-dessus de la moyenne ?

13. Quel est le top 5 des Pokémon légendaires en prenant en compte la somme de leur attaque spéciale et défense spéciale ?

14. Quelle est la moyenne des HP des Pokémon de la 4^e génération (3 chiffres après la virgule) ?

15. Quelle est la somme de toutes les caractéristiques des Pokémon ?

16. Parmi le starter de la 1^{er} génération (Bulbasaur, Squirtle, Charmander), qui est le meilleur ?

3.2 La Datavisualisation



La *datavisualisation* est « l'art » de représenter des données de façon visuelle. Cela peut se concrétiser par différents types de graphiques (diagrammes, courbes, ...) ou des cartographies, La présentation sous une forme illustrée rend les données plus lisibles et compréhensibles.



Dans cette activité, il s'agit d'apprendre à utiliser un outil de datavisualisation qui s'appelle « Tableau public ». Vous en trouverez des exemples sur ce site : <https://public.tableau.com/fr-fr/s/>

Regardez quelques vidéos de tutoriels : <https://public.tableau.com/app/learn/how-to-videos>



→ **Activité 3** (à faire à la maison) : À partir du fichier de données pokedex.xlsx que vous avez sauvegardé, vous devez créer une présentation visuelle (infographie) sur les Pokémon afin de déterminer quel est le type de Pokémon le plus puissant selon les critères de votre choix. **Pour cela vous pouvez utiliser le site Tableau public ou le faire directement dans Excel.**

Vous illustrerez votre infographie avec **trois types de diagrammes différents**. Il n'est pas nécessaire d'y inclure tous les Pokémon, vous pouvez vous contenter d'un « top 5 » dans chacune des trois catégories. Vous m'enverrez votre travail par l'ENT pour que je puisse le vérifier **AVANT** le cours.