

Traitement des données : Pokédex



I. Structure de données pokédex

Le Pokédex est une encyclopédie recensant tous les Pokémon existant et permettant d'enregistrer leurs informations. Inventé par le professeur Chen, il est remis en début d'aventure à tous les dresseurs afin qu'ils puissent récolter un maximum de données sur ces créatures. Cependant, l'apparition du Big Data et des objets connectés a changé la donne, et le professeur n'a plus besoin d'envoyer des enfants de 13 ans parcourir le monde dans le but de collecter des données. Tous les Pokédex sont maintenant connectés à la base de données des Pokémon, et celle-ci est désormais complète !



Chaque espèce de Pokémon possède des **statistiques** de base (fixes) qui sont décrites comme telle :

- Point de vie (**HP**) : un Pokémon avec beaucoup de points de vie pourra recevoir plus de dégâts avant d'être KO.
- Attaque (**Att**) : un Pokémon avec beaucoup d'attaque fera plus de dégâts lorsqu'il lance des attaques physiques.
- Défense (**Def**) : un Pokémon avec beaucoup de défense perdra moins de vie lorsqu'il reçoit une attaque physique.
- Attaque Spéciale (**Att spe**) : un Pokémon avec beaucoup d'attaque spéciale fera plus de dégâts lorsqu'il lance des attaques spéciales.
- Défense Spéciale (**Def spe**) : un Pokémon avec beaucoup de défense spéciale perdra moins de vie lorsqu'il reçoit une attaque spéciale.
- Vitesse (**Vit**) : un Pokémon avec plus de vitesse que son adversaire lance son attaque en premier. Nous avons mis la main sur les données du Pokédex (jusqu'à la huitième génération en tout cas) sous la forme d'un fichier CSV, et nous allons pouvoir à présent exploiter ces données !

Télécharger le fichier [pokedex.csv](#) puis l'ouvrir avec LibreOffice.

La première fenêtre qui s'ouvre permet d'importer le fichier CSV (voir copie d'écran plus bas).

Vérifier que seul « Point-virgule » est coché (et surtout pas Virgule).

Cliquer ensuite sur « OK ».

Import de texte - [Pokédex.csv]

Importer

Jeu de caractères : Unicode (UTF-8)

Locale : Par défaut - Français (France)

À partir de la ligne : 1

Options de séparateur

Largeur fixe Séparé par

Tabulation Virgule Point-virgule Espace Autre

Fusionner les séparateurs Espaces superflus

Séparateur de chaîne de caractères : "

Autres options

Formater les champs entre guillemets comme texte Détecter les nombres spéciaux

Évaluer les formules Détecter la notation scientifique

Champs

Type de colonne :

	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Standard	Standard
1	Num	Nom	Gen	Status	Type 1	Type 2	Taille (m)	Poids (kg)	HP	Att	Def	Att spe	Def s	
2	1	Bulbizarre	1	Commun	Plante	Poison	0,7	6,9	45	49	49	65	65	
3	2	Herbizarre	1	Commun	Plante	Poison	1	13	60	62	63	80	80	
4	3	Florizarre	1	Commun	Plante	Poison	2	100	80	82	83	100	100	
5	3	Florizarre (Mega)	1	Mega	Plante	Poison	2,4	155,5	80	100	123	122	120	

Pour que la première ligne reste toujours visible, aller dans le menu :

Affichage > Fixer des cellules > Fixer la première ligne

Pour pouvoir effectuer des recherches et trier les données dans le tableau, aller dans le menu :

Données > AutoFiltre

II. Tri simple

Maintenant que nous disposons d'une base de données conséquente, il faut pouvoir l'exploiter correctement. L'une des premières actions que l'on peut effectuer sur une base de données est de trier cette base en fonction d'un descripteur. Dans le cas du Pokédex, les éléments sont triés en fonction de leur numéro, mais il peut être plus avantageux de les trier par rapport à un autre descripteur ! Imaginons par exemple que nous recherchions un Pokémon de type « Feu » et « Eau ». Il est tout à fait possible de chercher directement dans notre base un Pokémon ayant ces deux caractéristiques, mais si nous n'avons pas de chance, le Pokémon que nous recherchons est en fin de base (c'est le cas ici), et donc cela va nous demander un temps fou. Nous allons donc opter pour une stratégie plus efficace en triant différemment les éléments de notre base !

À savoir :

A) Pour trier ou filtrer une base en fonction d'un descripteur donné, il suffit d'utiliser le filtre mis en place précédemment. Il suffit de cliquer sur le bouton en tête de la colonne selon laquelle on souhaite trier ou filtrer. Cela va nous permettre de ranger les éléments de notre base (par ordre croissant ou décroissant au choix) en fonction d'un descripteur, ou de ne choisir que les éléments qui nous intéressent.

B) Il est également possible de trier par rapport à deux ou plusieurs colonnes. Un premier tri sera alors effectué par rapport à la première colonne choisie, et en cas d'égalité entre deux éléments ou plus, on effectuera un deuxième tri par rapport à la deuxième colonne sélectionnée. Pour cela, aller dans :

Données > Trier...

C) On pourra également aller dans :

Données > Plus de filtres > Filtre standard...

On peut aller sélectionner les données qui nous intéressent selon des critères variés.

Informations importantes concernant les réponses à apporter :

- Si le Pokémon que vous voulez rentrer en réponse possède une forme spéciale, il faudra la préciser. Par exemple écrire "Dardargnan (Mega)";

- Lorsque plusieurs Pokémon répondent à la question posée, les écrire à la suite dans l'ordre que vous voulez en les séparant d'un espace.
- Dans certains cas, on vous pose des questions sur des Pokémon qui possèdent deux types, il faut alors choisir ces deux types dans chacune des colonnes "Type 1" et "Type 2" pour éviter de rater certains Pokémon.

Répondre aux questions suivantes en utilisant les outils évoqués précédemment :

1. Quels sont les Pokémon les plus lourds ?
2. Quel est le Pokémon de type « Plante » le plus grand ?
3. Quels sont les Pokémon de type « Feu » et « Psy » ?
4. Quels sont les Pokémon de type « Dragon » et « Electrik » ?
5. Quels sont les Pokémon de type « Roche » ayant une base de point de vie strictement supérieure à 110 ?
6. Quel est le Pokémon ayant la meilleure défense (si deux Pokémon ont la même défense, on les différenciera avec la défense spéciale) ?
7. Quel est le Pokémon légendaire de deuxième génération le plus rapide ?

III. Formule simple

1. Formule

Maintenant que l'on sait parfaitement trier nos données afin de les exploiter au mieux, nous allons pouvoir passer à une autre action fondamentale dans le traitement de données : les formules ! Une formule dans un tableur commence toujours par un signe égal « = » (cela permet de faire comprendre au programme qu'il s'agit d'une formule à calculer), suivi de valeurs numériques et d'opérations. Il est important de retenir que toutes les règles mathématiques s'appliquent ici (et notamment les parenthèses). Il est également possible de remplacer une valeur numérique par la valeur contenue dans une case en spécifiant les "coordonnées" de la case en question.

Exemple Imaginons que l'on souhaite par exemple calculer l'attaque moyenne de notre premier Pokémon. On peut alors sélectionner une case vide (par exemple la case O2), et entrer la formule suivante $=\text{(J2+L2)}/2$, les cases J2 et L2 correspondant à l'attaque et l'attaque spéciale du Pokémon.

2. Étendre une formule

Il est possible d'étendre une formule placée dans une case (c'est-à-dire reproduire cette formule) sur plusieurs lignes (ou colonnes). Pour cela, il suffit de placer le curseur de la souris dans le coin inférieur droit de la case que vous voulez étendre (une croix noire apparaît alors) et de l'étirer sur les différentes lignes (ou colonnes) souhaitées en restant appuyé sur le bouton gauche de la souris. Si la case que vous étirez ainsi comporte des cases spécifiques dans la formule, ces dernières seront automatiquement remplacées en respectant la logique et la structure de votre formule.

Exemple Si on étend la case O2 sur les 10 lignes du dessous (de O3 à O12), la formule va automatiquement s'ajuster, calculant ainsi l'attaque moyenne des 10 prochains Pokémon.

À savoir : Si on souhaite étirer une formule sur toute une colonne, le plus simple est de double-cliquer sur le coin inférieur droit de la cellule.

Si on effectue un tri après avoir écrit des formules, celles-ci vont automatiquement s'ajuster par rapport à notre tri.

3. Utiliser des fonctions

Il est possible d'automatiser certains calculs à l'aide de fonctions particulières. Les plus utilisées étant les fonctions SOMME et MOYENNE qui permettent de calculer respectivement la somme et la moyenne d'un grand nombre de données. Une autre fonction très utile (surtout lorsque l'on calcule des moyennes) est la fonction ARRONDI qui permet d'arrondir un nombre à n chiffres après la virgule, où n est une valeur à préciser.

Exemple La formule =MOYENNE(G2:G16) permet de calculer la moyenne des valeurs contenues dans les cases G2 à G16 incluses. Et si on veut arrondir ce résultat avec deux chiffres après la virgule, on peut modifier la formule par =ARRONDI(MOYENNE(G2:G16);2).

Répondre aux questions suivantes :

1. Notre base de données est encore incomplète car il manque le total des statistiques pour l'ensemble des Pokémon.

Quelle formule faut-il rentrer dans la case **O2** afin d'obtenir le total des statistiques du Pokémon ?

Rappel : les statistiques sont HP, Att, Def, Arr spe, Def spe et Vit.

Pour la suite, compléter la base de données en étirant cette formule sur toute la colonne.

2. Quel Pokémon commun possède le total de statistiques le plus élevé ?
3. Taper dans la cellule **I1030** (située à quelques lignes d'écart du tableau, en bas) la formule donnant la moyenne des points de vie (**HP**) de l'ensemble des Pokémon, arrondie à 2 chiffres après la virgule. Étirer ensuite cette formule vers la droite jusqu'à la colonne **N** de la vitesse (**Vit**).

Aide : On pourra utiliser l'outil disponible dans **Données > Plus de filtres > Filtre standard...**

4. Quel est la moyenne, arrondie à 3 chiffres après la virgule, des points de vie des Pokémon de la septième génération ?

Aide : Utiliser l'outil disponible dans le menu **Données > Sous-totaux...**

Pour obtenir le nombre de chiffres après la virgule souhaité, faire un clic droit dans la cellule à configurer et choisir « Formater des cellules ».

IV. Instruction conditionnelle

Nous avons vu jusqu'à maintenant comment trier une base de données et utiliser quelques fonctions simples afin d'exploiter au mieux ces données, mais il est facile de voir que nous sommes encore assez limités dans nos actions. Prenons un exemple simple : on veut déterminer le nombre de Pokémon de type « Plante ». A priori, rien de bien difficile, mais ici, on ne peut pas utiliser la fonction SOMME (car on ne peut sommer que des valeurs numériques) et une deuxième difficulté apparaît, le Pokémon peut-être de Type 1 « Plante » ou de Type 2 « Plante ». Dans ce cas de figure, nous n'avons d'autres choix que de compter à la main (et cela va rapidement être pénible car il y en a 95 tout de même).

Afin d'éviter ce travail fastidieux, on va introduire une nouvelle fonction : l'instruction conditionnelle SI. L'instruction SI permet d'établir des comparaisons logiques (ou tests) entre une valeur et le résultat attendu, et va renvoyer une réponse en fonction du résultat de cette comparaison. Si la comparaison est valide, la fonction renverra la première réponse possible, sinon elle renverra la deuxième réponse.

Exemple Si on écrit dans une case vide (on peut reprendre la case O2) la formule :

=SI(E2="Plante";1;0)

le programme va comparer le contenu de la cellule E2 avec le mot « Plante ». Si la cellule E2 contient le mot « Plante », le programme renverra en réponse 1, et sinon le programme renverra 0.

Pour la comparaison, il est important de mettre le mot Plante entre guillemets afin de faire comprendre au programme qu'il s'agit d'un mot.

Il est également possible de tester si une case est vide en écrivant rien entre les guillemets. Par exemple, le test E2="" va vérifier que la case E2 ne contient rien.

Remarque En étendant notre formule à l'ensemble des lignes, on obtiendra donc une colonne de 0 et de 1 en fonction du type du Pokémon. Ainsi en sommant tous ces résultats, on obtiendra le nombre de Pokémon ayant pour Type 1 « Plante ».

C'est un bon début, mais cela ne répond pas tout à fait à la question initiale car nous n'avons toujours pas pris en compte le Type 2 du Pokémon. Pour régler ce problème, nous allons devoir rajouter une comparaison, et pour cela, nous allons utiliser les fonctions logiques :

- La fonction logique OU permet d'effectuer plusieurs tests. Si au moins l'un des tests est valide, l'ensemble des tests sera validé.
- La fonction logique ET permet d'effectuer plusieurs tests. Si au moins l'un des tests n'est pas valide, l'ensemble des tests ne sera pas validé.

Dans notre exemple, nous devons tester si le Pokémon est de Type 1 « Plante » ou de Type 2 « Plante ». Si on revient sur le premier Pokémon, on a vu que le test afin de savoir s'il est de Type 1 « Plante » est **E2="Plante"**. De même, si on veut tester s'il est de Type 2 « Plante », il suffit d'effectuer le test **F2="Plante"**. Notre problème ici est que l'on aimerait faire les deux tests, et si l'un est valide on veut renvoyer la valeur 1, et 0 sinon.

Pour cela on va donc utiliser la fonction logique OU. Cela nous donnera la formule suivante :

=SI(OU(E2="Plante";F2="Plante");1;0).

Cette formule se lit : "si E2 contient le mot « Plante » ou F2 contient le mot « Plante » renvoie la valeur 1, sinon renvoie la valeur 0".

Il ne reste plus qu'à étendre cette formule à l'ensemble des lignes, et à sommer tous les résultats pour obtenir le nombre de Pokémon de type « Plante ».

Répondre aux questions suivantes :

1. Quel test affiche 1 si le Pokémon ne possède pas un double type et 0 sinon ? (Mettre juste la formule que vous rentreriez pour Bulbizarre.)
2. Quel est le nombre de Pokémon n'ayant pas de type 2 ?
3. Combien de Pokémon possèdent le type "Dragon" ?
4. Quel est le nombre de Pokémon ayant strictement plus de 100 de base en attaque (Att) et en attaque spéciale (Att spe) ?
5. Quel est le nombre de Pokémon de type « Spectre » et « Acier » ?
6. Quel est le nombre de Pokémon de type « Ténèbres » ayant strictement plus de 100 de base en attaque spéciale (Att spe) et en défense spéciale (Def spe) ?
7. **Réalisez un tableau détaillant le nombre de Pokémon de chaque type.**
Quel est le type le plus représenté parmi l'ensemble des Pokémon ?

Pour les plus rapides, chercher à obtenir les réponses précédentes en utilisant l'outil :

Données > Table dynamique