



**ACADÉMIE
DE NANTES**

*Liberté
Égalité
Fraternité*

Damien RIVIERE

Professeur de mathématiques – Collège Pierre Dubois – Laval (53)

Académie de Nantes – driviere@ac-nantes.fr

Expérimentation pédagogique sur le thème

Big Data et Intelligence Artificielle

**Des activités en collège et lycée
ayant pour thème les notions de Big Data et d'IA**

« Vérifier la loi de Benford »

Cycle 4 – 5e

Testée dans 2 classes de 5e au collège Pierre Dubois.

1. INTRODUCTION

Le but de l'activité est d'utiliser le tableur pour vérifier que la loi de Benford, que l'on a pu observer sur 2 exemples avec un nombre de données raisonnable, est valide sur un exemple comprenant un grand nombre de données. Dans un premier temps, les élèves de 5^{ème} ont découvert la loi de Benford à partir des 2 premières activités (sur les prix et la population des départements français) de l'activité déjà présente sur le site de l'académie de Nantes : <https://www.pedagogie.ac-nantes.fr/mathematiques/enseignement/groupes-de-recherche/traam/2019-2021/verifier-la-loi-de-benford-1362472.kjsp?RH=1160079471359>. Dans un 2^{ème} temps, ils ont découvert des fonctions du tableur en travaillant sur un fichier concernant les personnages d'Harry Potter et enfin ils ont vérifié la loi de Benford sur la superficie de tous les pays du monde.

2. DECOUVERTE DE FONCTIONS SOUS TABLEUR.

Pour que les élèves soient plus à l'aise dans l'étude des superficies des pays, une première activité a été proposée pour qu'ils puissent bien comprendre des fonctions du tableur qui leur seront utiles. Je me suis appuyé sur les personnages d'Harry Potter. Voici l'énoncé de l'activité et la fiche d'aide qui ont été donnés aux élèves.

Vous pourrez trouver un exemple de réalisation d'élèves dans les documents joints.

5 ^{ème}	S'ENTRAINER Le tableur	2022-2023
------------------	---	-----------

Les données : vous disposez d'un fichier contenant les prénoms, noms et dates de naissance de certains des personnages principaux d'Harry Potter.

Prénoms	NOMS	dates de naissance
Lily	Evans	30.01.1960
Arthur	Weasley	06.02.1950
Luna	Lovegood	13.02.1981
Ron	Weasley	01.03.1980
Remus	Lupin	10.03.1960
James	Potter	27.03.1960
Fred	Weasley	01.04.1978
George	Weasley	01.04.1978
Drago	Malefoy	05.06.1980
Neville	Londubat	30.07.1980
Harry	Potter	31.07.1980
Ginny	Weasley	11.08.1981
Hermione	Granger	19.09.1979
Sirius	Black	03.11.1959
Bill	Weasley	29.11.1971
Severus	Rogue	09.01.1960

Votre 1ère mission

Vous devez créer un code pour chaque personnage à l'aide des fonctions GAUCHE et DROITE du tableur.

Le code est composé des 3 premières lettres du prénom, de la première lettre du nom et de l'année de naissance.

Exemple :

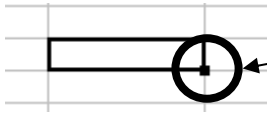
Lily	Evans	30.01.1960
LilE1960		

Votre 2^{ème} mission

Vous devez compter le nombre de personnages dont le nom commence par W, B ... et le nombre de personnages qui sont nés en 1959, 1960 ... à l'aide de la fonction NB.SI

commençant par ...	nombre de noms	année de naissance	nombre de personnages
E		1950	
W	6	1959	
L		1960	
P		1971	
M		1978	
G		1979	
B		1980	4
R		1981	

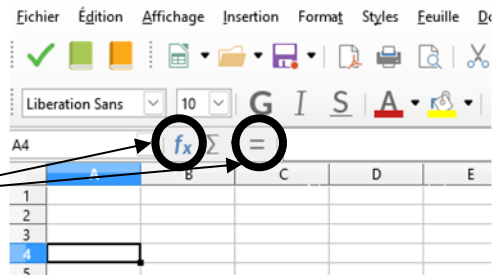
RAPPEL : Etirer une formule



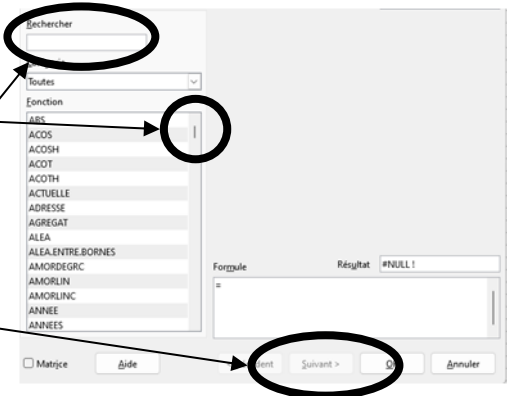
Pour que la formule se fasse sur plusieurs cellules il vous suffit de cliquer sur le petit carré de la première cellule -et d'aller jusqu'à la dernière cellule.

Pour utiliser une fonction

Cliquez d'abord sur « = » puis sur fx



Recherchez la fonction Soit en faisant une recherche, Soit en faisant défiler toutes les fonctions.



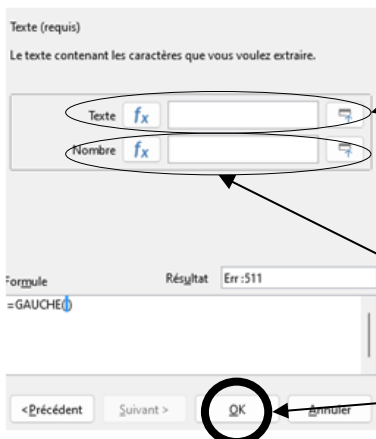
Cliquez sur « suivant » Puis suivez les instructions.

La fonction GAUCHE et la Fonction DROITE

GAUCHE permet de prendre les premières lettres ou nombres d'une cellule

DROITE permet de prendre les dernières lettres ou nombres d'une cellule

GAUCHE(Texte; Nombre)

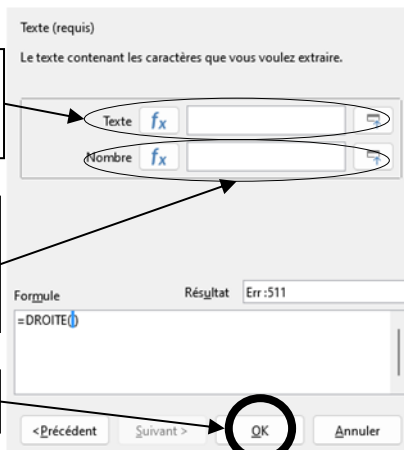


Cliquez sur la cellule du texte d'origine.

Choisissez le nombre de lettre que vous voulez garder

Cliquez sur « OK »

DROITE(Texte; Nombre)



La fonction NB.SI

NB.SI permet de compter le nombre de cellules qui vérifie une condition.

NB.SI(Plage; Critère)

The screenshot shows the 'NB.SI' dialog box in Excel. It has a title bar 'Plage (requis)' and a subtitle 'La plage de cellules à évaluer par les critères donnés.' Below this are two input fields: 'Plage' and 'Critère', both with an 'fx' icon and a selection button. Below these is a 'Formule' field containing '=NB.SI(' and a 'Résultat' field showing 'Err:511'. At the bottom are buttons for '<Précédent', 'Suivant >', 'OK', and 'Annuler'. The 'OK' button is circled in red. Three callout boxes with arrows point to the dialog:

- Top right: Cliquez sur les cellules dans lesquelles vous voulez compter.
- Middle right: Mettre entre «" "» ce que vous voulez compter.
- Bottom right: Cliquez sur « OK »

La fonction CONCATENER

CONCATENER permet de regrouper des textes de plusieurs cellules pour former un texte dans une seule cellule.

CONCATENER(Texte1; Texte2; ...)

The screenshot shows the 'CONCATENER' dialog box in Excel. It has a title bar 'Texte1 (requis)' and a subtitle 'Élément de texte pour la concaténation.' Below this are four input fields: 'Texte1', 'Texte2', 'Texte3', and 'Texte4', each with an 'fx' icon and a selection button. Below these is a 'Formule' field containing '=CONCATENER(' and a 'Résultat' field showing '""'. At the bottom are buttons for '<Précédent', 'Suivant >', 'OK', and 'Annuler'. The 'OK' button is circled in red. Three callout boxes with arrows point to the dialog:

- Top right: Cliquez sur la cellule du premier texte que vous voulez regrouper.
- Middle right: Cliquez sur la cellule du deuxième texte que vous voulez regrouper...
- Bottom right: Cliquez sur « OK »

Voici un exemple de réalisation d'élèves

	A	B	C	D	E	F	G
1							
2							
3							
4	Prénom	NOM	date de naissance	3 premières lettres du Prénom	1ère lettre du NOM	année de naissance	code créé
5							
6	Lily	Evans	30.01.1960	Lil	E	1960	LilE1960
7	Arthur	Weasley	06.02.1950	Art	W	1950	ArtW1950
8	Luna	Lovegood	13.02.1981	Lun	L	1981	LunL1981
9	Ron	Weasley	01.03.1980	Ron	W	1980	RonW1980
10	Remus	Lupin	10.03.1960	Rem	L	1960	RemL1960
11	James	Potter	27.03.1960	Jam	P	1960	JamP1960
12	Fred	Weasley	01.04.1978	Fre	W	1978	FreW1978
13	George	Weasley	01.04.1978	Geo	W	1978	GeoW1978
14	Drago	Malefoy	05.06.1980	Dra	M	1980	DraM1980
15	Neville	Londubat	30.07.1980	Nev	L	1980	NevL1980
16	Harry	Potter	31.07.1980	Har	P	1980	HarP1980
17	Ginny	Weasley	11.08.1981	Gin	W	1981	GinW1981
18	Hermione	Granger	19.09.1979	Her	G	1979	HerG1979
19	Sirius	Black	03.11.1959	Sir	B	1959	SirB1959
20	Bill	Weasley	29.11.1971	Bill	W	1971	BillW1971
21	Severus	Rogue	09.01.1960	Sev	R	1960	SevR1960
22							
23							
24				commençant par ...	nombre de noms	année de naissance	nombre de personnages
25				E	1	1950	
26				W	6	1959	
27				L	3	1960	
28				P	2	1971	
29				M	1	1978	
30				G	1	1979	
31				B	1	1980	
32				R	1	1981	
33							

3. ETUDE DES SUPERFICIES DES PAYS.

Dans un 2^{ème} temps, un fichier contenant les superficies de tous les pays du monde a été fourni aux élèves. Leur mission était de vérifier que la loi de Benford était bien respectée.

Rapidement les élèves ont compris qu'il fallait trouver le 1^{er} nombre de chaque superficie et ont utilisé la fonction GAUCHE pour remplir une première colonne :

	A	B	C
1	Pays ou territoire	Superficie (km ²)	
2	Russie	17 125 191	1
3	Canada	9 984 670	9
4	États-Unis	9 629 091	9
5	République populaire de Chine	9 600 000	9
6	Hong Kong	1 104	1
7	Macao	30	3
8	Brésil	8 515 767	8
9	Australie	7 692 060	7
10	Inde	3 287 263	3
11	Argentine	2 780 400	2

Une difficulté est apparue pour les élèves : elle concerne la superficie du Vatican (0,44 km²). La fonction GAUCHE donne « 0 » alors que la loi de Benford étudie le 1^{er} chiffre différent de 0. Après discussion avec les élèves sur ce problème, 2 solutions ont été retenues :

- On considère qu'au vu du nombre de pays étudié, si on ne tient pas compte d'un pays cela ne changera pas la globalité des résultats.
- Vu qu'il n'y a qu'un seul pays qui est concerné, on peut faire la modification à la main.

J'ai expliqué aux élèves que l'on aurait pu faire une programmation sous le tableur en utilisant par exemple la fonction SI, mais que ce n'était pas l'objet de l'activité.

Ils ont ensuite compté le nombre de 1, de 2 et ... qu'il y avait dans la colonne qu'ils venaient de créer, en utilisant la fonction NB.SI :

	A	B	C	D	E
1	Pays ou territoire	Superficie (km ²)			
2	Russie	17 125 191	1		63
3	Canada	9 984 670	9		43
4	États-Unis	9 629 091	9		31
5	République populaire de Chine	9 600 000	9		24
6	Hong Kong	1 104	1		16
7	Macao	30	3		16
8	Bésil	8 515 767	8		16
9	Australie	7 692 060	7		10
10	Inde	3 287 263	3		13

La plupart des élèves ont saisi dans chaque cellule la formule. J'ai expliqué le principe du \$ avec seulement 3 groupes d'élèves.

Ils ont enfin réalisé un graphique représentant les résultats obtenus :

