



MINISTÈRE
DE L'ÉDUCATION
NATIONALE
ET DE LA JEUNESSE

*Liberté
Égalité
Fraternité*

VOIE PROFESSIONNELLE

CAP

2^{DE}

1^{RE}

T^{LE}

Mathématiques

Accompagnement
renforcé

Décoder un texte par analyse fréquentielle

Domaine

Organisation et gestion des données, Fonctions (cycle 4).

Statistique et probabilité (seconde professionnelle)

Capacités

- Organiser des données statistiques en choisissant un mode de représentation adapté à l'aide des fonctions statistiques d'une calculatrice, d'un tableur
- Extraire des informations d'une représentation graphique d'une série statistique (seconde professionnelle)

Connaissances

- Effectifs, fréquences (cycle 4)
- Représentation d'une série statistique par un diagramme en bâtons (seconde professionnelle)

Compétences du cycle 4 associées

Lire et interpréter des données sous forme brute, sous forme de tableau, de diagramme

Objectifs

- Utiliser les TIC dans le cadre de la résolution d'un problème.
- Décomposer une tâche complexe en plusieurs tâches simples.
- Calculer des fréquences et les représenter graphiquement à l'aide d'un diagramme en bâtons.
- Analyser et exploiter des données extraites d'un dossier documentaire.

Modalités

Durée de la séance : une séance de 55 minutes

- Un travail en îlots est préconisé durant la séance.
- La constitution de groupes de besoin est à prioriser. Ces groupes sont à établir grâce à l'exploitation des tests de positionnement de début de seconde et/ou des tests de mi-parcours.
- La séance proposée est à destination d'un groupe d'élèves éventuellement issus de différentes classes et ayant été identifiés comme présentant des difficultés dans le domaine : « Organisation des données » lors des tests de positionnement d'entrée en seconde et/ou des tests de mi-parcours.

Énoncé de l'activité d'accompagnement renforcé

Une célèbre citation scientifique a été codée à l'aide du chiffre de César, dont la longueur de la clé est inconnue :

« HYUD DU IU FUHT HYUD DU IU SHUU JEKJ IU JHQDIVEHCU »

À l'aide des documents suivants, trouver l'auteur de cette citation.

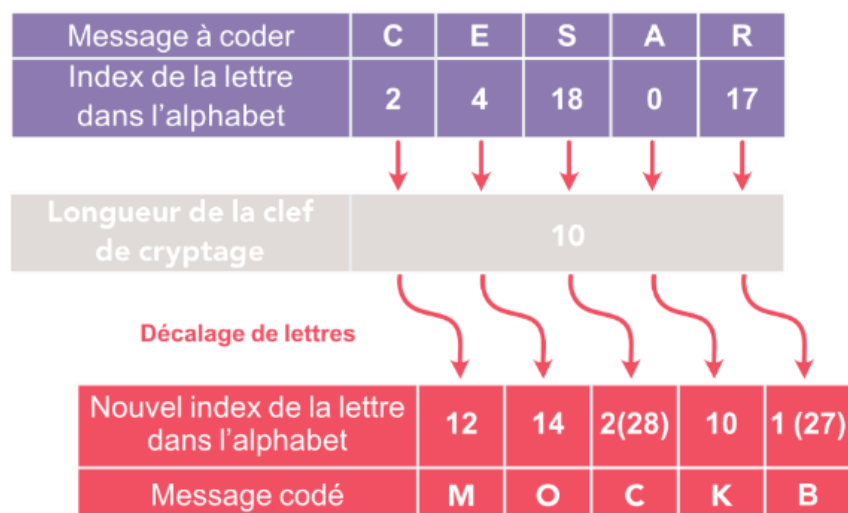
Document 1 : Fréquences d'apparition des lettres dans un texte en Français.

Lettre	A	B	C	D	E	F	G	H	I	J	K	L	M
Fréquence (%)	8,15	0,9 7	3,15	3,7 3	17,39	1,12	0,9 7	0,8 5	7,31	0,4 5	0,0 2	5,6 9	2,8 7

Lettre	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Fréquence (%)	7,12	5,2 8	2,8 0	1,21	6,64	8,14	7,2 2	6,3 8	1,64	0,0 3	0,41	0,2 8	0,15

SOURCE : WWW.APPRENDRE-EN-LIGNE.NET

Document 2 : Principe de codage par le chiffre de César pour une clé de longueur 10.



Document 3 : Occurrences des lettres dans un texte codé avec la même clé de cryptage que celle utilisée pour coder la citation

Lettre	Nombre d'occurrences
A	496
B	14 973
C	8 074
D	20 423
E	15 775
F	8 564
G	3 547
H	19 558
I	23 317
J	20 935
K	17 407
L	4 441
M	32

Lettre	Nombre d'occurrences
N	1 359
O	926
P	233
Q	21 255
R	2 394
S	9 834
T	12 030
U	41 947
V	3 150
W	2 681
X	2 445
Y	21 805
Lettre	Nombre d'occurrences

Commentaires de l'activité

Cette activité est à traiter au cours d'une séance d'accompagnement renforcé, et plus particulièrement dans le cadre d'un travail autour de la consolidation des acquis.

Indicateur de réussite

- Repérer que la lettre la plus fréquente en français est le **E** et que la moins fréquente est le **W**.
- Repérer que la lettre la plus fréquente dans le texte codé est le **U** et que la moins fréquente est le **M**.

- Vérifier que l'écart entre les lettres **E** et **W** est le même qu'entre les lettres **U** et **M**.
- Dérouler les correspondances lettre à lettre puis remplacer les lettres dans le texte pour trouver la phrase suivante : « Rien ne se perd, rien ne se crée, tout se transforme »

Déroulé

Une phase de 55 minutes

Modalité de travail	Objectifs
Phase individuelle (5 min)	Présentation de l'activité et lecture silencieuse par les élèves Phase d'appropriation et de recherche Élaboration des premières représentations
Phase collective (5 min)	Vérification de la compréhension de l'activité et de la consigne Reformulation de la situation par les élèves Phase de questionnement afin de s'assurer de la compréhension de l'activité et explicitation
Phase de travail en îlots (25 min)	Accompagnement des élèves selon leurs besoins – Questionnement sur les démarches envisagées Résolution de l'activité - Différenciation
Phase collective (10 min)	Explicitation par les élèves, à l'oral, de leurs diverses propositions de démarche
Phase collective (10 min)	Décontextualisation - Institutionnalisation Trace écrite

Analyse

Procédures utilisées par les élèves pour résoudre la tâche.

Toute démarche proposée par l'élève doit être prise en compte dans le scénario pédagogique, qu'elle soit erronée, incomplète ou aboutie.

Obstacles pouvant être rencontrés par les élèves

- Obstacle 1 : l'élève ne parvient pas à s'approprier la situation ;
- Obstacle 2 : l'élève ne parvient pas à mettre en place une stratégie de résolution ;
- Obstacle 3 : l'élève n'est pas en mesure de déterminer la clé de cryptage ;
- Obstacle 4 : l'élève ne maîtrise pas les contenus disciplinaires nécessaires au traitement de la situation ;
- Obstacle 5 : l'élève ne parvient pas à rendre compte de la démarche élaborée ;

Obstacle 6 : l'élève ne parvient pas à comparer les deux diagrammes en bâtons¹ ;
 Obstacle 7 : l'élève n'est pas en mesure d'utiliser la clé de cryptage pour décoder la citation.

Pistes de différenciation

	Obstacles 1 et 6	Obstacle 2	Obstacles 3 et 7	Obstacle 4	Obstacle 5
Différenciation des contenus	Proposer un questionnement détaillé	Proposer des supports avec décomposition de la démarche	Proposer un tableau avec les lettres de l'alphabet et pour chacune, la lettre codée correspondante ou un autre exemple de codage avec la clé de cryptage utilisée pour coder la citation	Proposer un exemple de calcul de fréquence	
Différenciation des processus	Proposer des supports de différents types (pour favoriser l'appropriation d'informations) : aide-mémoire, liste de vérification corrective, grille d'attendus, jokers, coups de pouce, Qrcode, capsules vidéo, notice technique d'utilisation des TIC, etc.				
Différenciation des productions					Laisser le libre choix à l'élève du mode de restitution
Différenciation des outils	Proposer un document mettant en parallèle les deux	Proposer un protocole de résolution		Permettre l'accès à différentes ressources	Proposer un outil de présentation (carte mentale,

¹ Dans le cas où l'élève aurait choisi de représenter graphiquement les fréquences ou effectifs des lettres à l'aide d'un diagramme en bâtons

	diagrammes en bâtons				diaporama, etc.)
--	-------------------------	--	--	--	---------------------

Protocole de résolution à proposer

Les différentes étapes de ce protocole peuvent également être utilisées indépendamment, tout ou partie, et ainsi constituer des aides ponctuelles sous forme de jokers, QR code, etc.

- Déterminer la lettre la plus fréquente dans la langue française ;
- Représenter graphiquement d'une part la répartition des fréquences d'apparition théoriques et d'autre part celle du texte codé ;
- Déterminer la clé du cryptage ;
- Décoder la citation² proposée dans le contexte à l'aide de la clé de cryptage déterminée ;
- Retrouver l'auteur de la citation une fois décodée.

Pistes d'exploitation en amont

Automatismes

En plus de permettre à l'enseignant de diagnostiquer les difficultés persistantes en vue d'une remédiation ultérieure, l'instauration de rituels mathématiques en début de séance, ou de manière asynchrone présente un intérêt multiple :

- Répondre à la demande institutionnelle concernant le module « Automatismes »³ ;
- Réactiver les prérequis du cycle 4 nécessaires à la séance ;
- Consolider les acquis antérieurs ;
- Rendre disponibles des réflexes en situation de résolution de problème ;
- Exploiter les erreurs rencontrées.

Les éléments présentés ci-dessous ont pour objectif de traiter les automatismes calculatoires sous deux formes différentes : des questions flash et des activités de niveau intermédiaire. Ces propositions pourront être adaptées par l'enseignant pour créer des situations à prise d'initiative. Elles ont pour objectif de nourrir la réflexion des enseignants sur les capacités et connaissances relatives au domaine de la statistique et des probabilités. L'enseignant pourra en utiliser tout ou partie en prenant soin de les intégrer au moment de la séquence qu'il jugera le plus propice, en fonction des objectifs visés et de l'évolution du niveau de compréhension des apprenants.

² Si l'élève présente des difficultés d'appropriation du processus de chiffrement, un étayage supplémentaire pourra lui être proposé.

³ <https://eduscol.education.fr/document/25972/download>

Propositions de questions flash

① Quelle est la fréquence de boules vertes contenues dans l'urne ci-contre ?

- ☐ 0,25 ☐ 0,50
☐ 5 ☐ 0,12

Couleur des boules	Nombre de boules
Bleu	4
Rouge	9
Vert	5
Jaune	2

② Le réservoir d'un véhicule contient 45 litres. **En considérant qu'il consomme 6 litres pour 100 kilomètres, quelle distance le conducteur pourra-t-il parcourir ?**

- ☐ 51 kilomètres ☐ 4 500 kilomètres ☐ 39 kilomètres ☐ 750 kilomètres

③ Quelle masse de glucides contient un bol de 50 g de céréales ?

- ☐ 82 g ☐ 141 g
☐ 0,82 g ☐ 41 g

Nutriments	Teneur pour 100 g
Protides	6
Glucides	82
Lipides	12

Question flash	Descriptif	Analyse des distracteurs selon les propositions de réponse
1	Calculer une fréquence	0,12 : l'élève effectue un calcul erroné par manque de stratégie calculatoire ou choisit au hasard la réponse. 0,50 ou 5 : l'élève utilise la valeur de l'effectif sans calculer l'effectif total.
2	Faire un calcul de proportionnalité	51 : l'élève calcule la somme des valeurs extraites du contexte (45 et 6). 4500 : l'élève calcule le produit de la valeur de la capacité du réservoir par 100. 39 : l'élève calcule la différence des valeurs extraites du contexte (45 et 6).

3	Faire un calcul de proportionnalité	82 g : l'élève ne prend pas en compte la masse de céréales contenue dans le bol. 0,82 g : l'élève calcule la valeur relative au pourcentage de glucides. 141 g : l'élève calcule la masse de glucides dans le bol, puis la somme de cette valeur avec la masse de référence (100 g).
---	-------------------------------------	--

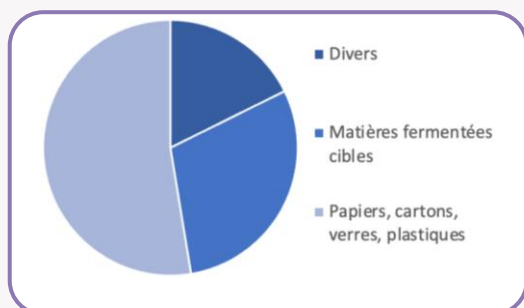
Propositions de situations de niveau intermédiaire

- ④ On réalise une étude sur la composition moyenne des poubelles de deux familles. Les tableaux suivants résument la composition moyenne des poubelles de deux familles.

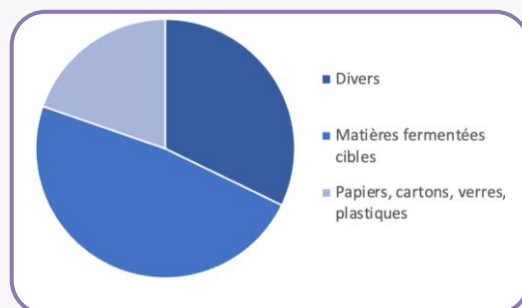
Composition moyenne de la poubelle de la famille A	
Papiers, cartons, verres, plastiques	32,2%
Matières fermentées cibles	19,7%
Divers	48,1%

Composition moyenne de la poubelle de la famille B	
Papiers, cartons, verres, plastiques	29,7%
Matières fermentées cibles	17,7%
Divers	52,6%

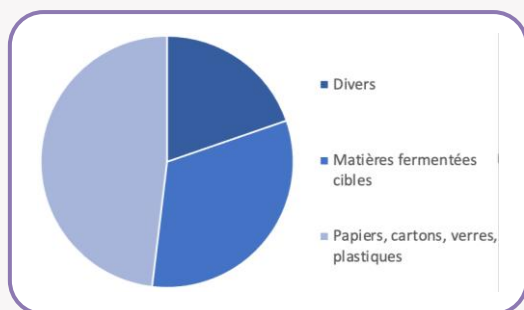
Parmi les quatre diagrammes circulaires ci-dessous, déterminer celui représentant la composition moyenne de la poubelle de la famille A.



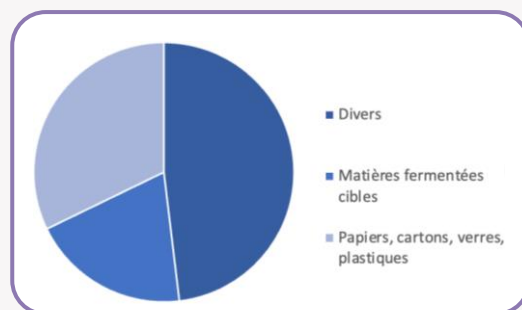
☐ Graphique A



☐ Graphique B



☐ Graphique C



☐ Graphique D

⑤ Le tableau suivant présente la répartition des plateformes les plus regardées pour un panel donné :

Plateforme utilisée	A	B	C	D	E
Effectif n_i	40	80	160	200	140

Quelle représentation graphique choisir pour représenter cette série statistique ?

- ☐ un diagramme à lignes brisées ☐ un histogramme
☐ un diagramme circulaire ☐ un diagramme en boîte à moustaches

⑥ Les salaires des employés d'une PME ont été recensés dans le tableau suivant.

Montant du salaire (€)	1900	1760	2200	3400	1200	2050
------------------------	------	------	------	------	------	------

Déterminer le salaire médian de cette entreprise.

- ☐ 2 085 € ☐ 1 975 € ☐ 2 800 € ☐ 10 801,67 €

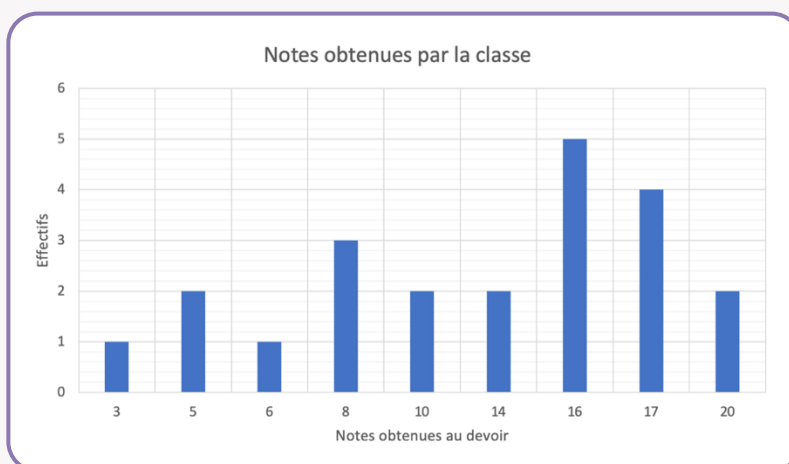
⑦ Le temps consacré chaque semaine par les élèves d'un lycée à regarder des séries est donné dans le tableau suivant.

Durée en heures	[0 ; 4 [	[4 ; 8 [	[8 ; 12 [	[12 ; 20 [	[20 ; 28 [
Effectif n_i	40	80	160	200	140

Déterminer une valeur arrondie au centième de la fréquence des élèves regardant des séries plus de 12h par semaine.

- ☐ 0,32 ☐ 0,45 ☐ 340 ☐ 0,55

8 Le graphique ci-dessous représente les résultats des élèves d'une classe suite à une évaluation.



Déterminer une valeur arrondie à l'unité de la note moyenne obtenue par cette classe.

Situation	Descriptif	Analyse des distracteurs selon les propositions de réponse
4	Lire et interpréter des données brutes : mettre en relation un tableau de valeurs avec un diagramme circulaire	Graphique A : l'élève confond la représentation de la composition de la poubelle B avec celle de la poubelle A. Graphique B : l'élève réalise une interversion entre les trois composants de la poubelle de la famille A. Graphique C : l'élève réalise une interversion entre les trois composants de la poubelle de la famille A.
5	Organiser des données statistiques en choisissant un mode de représentation adapté	Un diagramme à lignes brisées : l'élève ne prend pas en compte la nature du caractère étudié. Un histogramme : l'élève ne prend pas en compte la nature du caractère étudié. Un diagramme en boîte à moustaches : l'élève effectue son choix au hasard (le diagramme en boîte à moustaches n'a, a priori, pas été étudié au cycle 4).
6	Calculer un indicateur de position à partir de données brutes	2 085 € : l'élève réalise le calcul de la moyenne pour déterminer le salaire médian de la série statistique. 2 800 € : l'élève calcule le salaire médian à l'aide des valeurs non triées par ordre croissant. 10 801,67 € : l'élève réalise le calcul de la moyenne pour déterminer le salaire médian de la série statistique et effectue une erreur de calcul (oubli de la parenthèse lors de la division).

7	Calculer des fréquences dans une série statistique	0,32 : l'élève divise l'effectif de la classe [12 ; 20 [par l'effectif total et ne prend pas en compte la classe [20 ; 28[dans le calcul. 340 : l'élève additionne les effectifs des deux classes relatifs au terme : « plus de 12h ». 0,45 : l'élève confond le terme « plus de » et « au plus », il effectue donc le quotient de la somme des effectifs des classes correspondantes par l'effectif total.
8	Calculer un indicateur de position à partir d'une représentation graphique donnée	17 : l'élève calcule l'étendue de la série statistique. 15 : l'élève calcule la note médiane de la série statistique. 16 : l'élève confond la moyenne de la série statistique avec le mode de celle-ci.

Pistes d'exploitation en aval

Activité de réinvestissement

En complément des notions relatives aux programmes de cycle 4 (interpréter, représenter et traiter des données) et de seconde professionnelle (statistique et probabilités), on peut envisager une activité réinvestissant les capacités, connaissances et contenus relatifs de la notion de fonction affine.

Le cryptage affine est une méthode de chiffrement qui consiste à coder chaque lettre de l'alphabet en utilisant les coefficients d'une fonction affine donnée.

À titre d'exemple, on considère la fonction affine $f: x \mapsto 3x + 7$ pour définir la méthode de cryptage ci-dessous dont la clé est donc (3 ; 7) :

Message clair	L	A	C	E	T
Index de la lettre dans l'alphabet	11	0	2	4	19
↓ ↓ ↓ ↓ ↓					
3 × Index + 7					
↓ ↓ ↓ ↓ ↓					
Nouvel index de la lettre dans l'alphabet	14(40)	7	13	19	12 (64)
Message codé	O	H	N	T	M

Coder la citation que vous avez trouvée dans l'activité précédente à l'aide d'un codage affine de clé (5 ; 8)

Remarque : il convient de choisir le coefficient directeur **a** de la fonction utilisée dans le cryptage de manière à ce que les nombres **a** et 26 soient premiers entre eux.

Activité de prolongement

La cryptographie correspond à l'art de coder un message qui se voudrait compréhensible uniquement de l'émetteur et du destinataire dudit message. En prolongement du contexte proposé, une activité traitant le codage d'un message par le chiffre de César pourrait être proposée dans le cadre d'un travail autour du module transversal 'Algorithmique et programmation'⁴.

La fonction Python **code** ci-contre permet de coder un mot écrit en majuscules non accentuées à l'aide du chiffre de César.

```
alpha = "ABCDEFGH IJKLMNOPQRSTUVWXYZ"

def code(texte,clef):
    resultat = ""
    for lettre in texte:
        rangL = alpha.find(lettre)
        indice = (rangL + clef) % 26
        resultat += alpha[indice]
    return resultat
```

1 - Expliquer les principales étapes mobilisées dans cette fonction.

2 - Modifier la fonction code afin qu'elle permette de chiffrer une phrase écrite en majuscules non accentuées.

Éléments de correction

1 - L'étude et l'utilisation des listes Python n'étant pas au programme de seconde professionnelle, il convient dans cet exemple d'utiliser un outil alternatif : les chaînes de caractères.

```
alpha = "ABCDEFGH IJKLMNOPQRSTUVWXYZ"
```

```
def code(texte,clef):
    resultat = ""
    for lettre in texte:
        rangL = alpha.find(lettre)
        indice = (rangL + clef) % 26
        resultat += alpha[indice]
    return resultat
```

Création d'une chaîne de caractère regroupant l'ensemble des lettres de l'alphabet

Pour chaque lettre du texte

On recherche le rang alphabétique de cette lettre

On détermine le rang alphabétique de la lettre codée à l'aide de la clef de cryptage

On ajoute à la variable **resultat** la lettre codée

On renvoie la valeur de la variable **resultat** qui correspond au message codé

⁴ <https://eduscol.education.fr/document/25978/download>

2 - La prise en compte des espaces implique d'utiliser une instruction conditionnelle comme dans l'algorithme ci-dessous et la fonction Python correspondante :

Fonction **code2** (texte, clef) :

```
resultat ← Chaîne de caractères vide
Pour chaque lettre du texte à coder
  Si la lettre est un espace alors
    Ajouter un espace à la variable resultat
  Sinon
    rangL ← rang alphabétique de la lettre
    indice ← rang alphabétique de la lettre
    codée
    Ajouter la lettre codée à la variable resultat
  Fin Si
Fin Pour
Renvoyer la valeur de la variable resultat
Fin
```

```
def code2(texte,clef):
    resultat = ""
    for lettre in texte:
        if lettre == ' ':
            resultat += ' '
        else:
            rangL = alpha.find(lettre)
            indice = (rangL+clef)%26
            resultat += alpha[indice]
    return resultat
```

Pistes de prolongement complémentaire

- Il serait intéressant de sensibiliser les élèves à la partie déchiffrement d'un message codé à l'aide du code pour une clé donnée. Pour cela, on utilise le même principe que la fonction **code2** proposée ci-dessus mais en modifiant le sens du décalage des lettres. Dans cet exemple, cela se traduit par une modification de la variable indice, correspondant à l'index de la lettre codée dans l'alphabet c'est-à-dire son rang alphabétique :

indice = (rangL+clef)%26



indice = (rangL-clef)%26

- Dans un contexte différent, on pourrait envisager l'étude de chiffrement et/ou déchiffrement pour un message non nettoyé⁵, c'est à dire qui contient les formes de ponctuation, des caractères accentués ou encore un mélange de minuscules et de majuscules.

⁵ <https://eduscol.education.fr/document/25978/download> (page 37)