STATISTIQUES

Exemple

Voici la liste des âges en mois d'élèves de troisième :

180	176	179	176	182	178	184	181	179	173	182
187	175	181	174	183	178	180	178	184	173	175
173	180	195	179	174	182	172	174	195	183	192
190	181	173	177	180	183	180	183	186	180	

Définitions

On appelle effectif total le nombre de valeurs de la série.

Ici, l'effectif total est 43.

On appelle effectif de A le nombre de fois où A apparait dans la série.

L'effectif de **178** est 3 car il y a 3 personnes ayant **178** mois.

On appelle fréquence de A le quotient de l'effectif de A par l'effectif total.

La fréquence de **178** est $\frac{3}{43}$, ce qui signifie que 3 élèves sur les 43 du groupe ont **178** mois.

$$\text{Fréquence de A} = \frac{\text{Effectif de A}}{\text{Effectif total}}$$

Les fréquences sont (souvent) exprimées en pourcentages.

La fréquence de **178** est $\frac{3}{43} \approx 7\%$.

$$\text{Fréquence de A en \%} = \text{Fréquence de A} \times 100 = \frac{\text{Effectif de A}}{\text{Effectif total}} \times 100$$

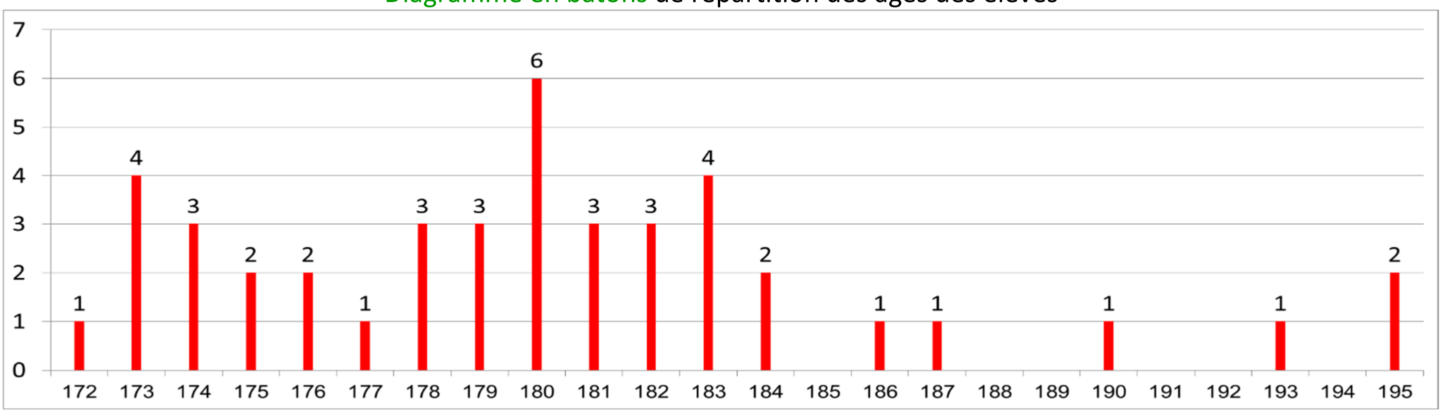
Exemple des âges

Age en mois	172	173	174	175	176	177	178	179	180	181	182	183	184	186	187	190	192	195	Total
Effectif	1	4	3	2	2	1	3	3	6	3	3	4	2	1	1	1	1	2	43
Fréquence	1/43	4/43	3/43	2/43	2/43	1/43	3/43	3/43	6/43	3/43	3/43	4/43	2/43	1/43	1/43	1/43	1/43	2/43	1
Fréquence en %	2 %	9 %	7 %	5 %	5 %	2 %	7 %	7 %	14 %	7 %	7 %	9 %	5 %	2 %	2 %	2 %	2 %	5 %	100 %

÷ 43

× 100

Diagramme en bâtons de répartition des âges des élèves



Définitions

On appelle *minimum* la plus petite valeur de la série.

Le minimum est 172.

On appelle *maximum* la plus grande valeur de la série.

Le maximum est 195.

On appelle *étendue* l'écart entre le minimum et le maximum.

L'étendue est 195 - 172 = 23 mois.

On appelle *médiane* une valeur qui partage la série en deux sous-parties de même effectif tel que dans une sous-partie sont regroupées toutes les valeurs inférieures ou égales à la médiane et dans l'autre sont regroupées toutes les valeurs supérieures ou égales à la médiane.

On appelle *premier quartile* la plus petite valeur de la série telle que 25% au moins des valeurs lui soit inférieure ou égale.

On appelle *troisième quartile* la plus petite valeur de la série telle que 75% au moins des valeurs lui soit inférieure ou égale.

Remarques

Une médiane peut être une valeur de la série (ou non).
Les quartiles sont obligatoirement des valeurs de la série.

Comment déterminer une médiane ?

On ordonne (on trie) la série (par ordre croissant).
On détermine l'effectif total N.
Si N est impair, la médiane est une des valeurs de la série ; c'est la $\frac{N+1}{2}$ ème.
Si N est pair, une médiane est entre deux valeurs de la série ; entre la $\frac{N}{2}$ ème et la $\frac{N}{2}+1$ ème.

Astuce pour déterminer la position de la médiane

On calcule $\frac{N+1}{2}$
Si on trouve un nombre entier, c'est la position de la médiane dans la série
Si ce n'est pas un nombre entier, la position de la médiane est donnée par les deux entiers consécutifs encadrant $\frac{N+1}{2}$.

Exemple 1

Dans la série 1 ; 2 ; 5 ; 7 ; 8 ; 5 ; 6 ; 9 ; 10, l'effectif total est 9, donc la médiane est la 5ème valeur de la série ordonnée.

La série ordonnée est 1 ; 2 ; 5 ; 5 ; **6** ; 7 ; 8 ; 9 ; 10 donc la médiane est **6**.

On prend les 2 entiers consécutifs encadrant $\frac{10+1}{2}=5{,}5$ On calcule $\frac{9+1}{2}=5$

Exemple 2

Dans la série 1 ; 2 ; 6 ; 7 ; **11 ; 15** ; 15 ; 17 ; 17 ; 17, l'effectif total est 10, donc une médiane est entre la 5ème valeur et la 6ème valeur de la série ordonnée.

La série est déjà ordonnée donc la médiane est **entre 11 et 15**. On peut prendre n'importe quelle valeur, mais on choisit le « milieu » de l'intervalle. Ici, une médiane est 13.

Exemple des âges

Dans l'exemple des âges, l'effectif total est 43. Une médiane est la 22-ème valeur de la série ordonnée.

La série ordonnée commence par : 172, 173, 173, 173, 173, 174, 174, 174, 175, 175, 176, 176, 177, 178, 178, 178, 179, 179, 179, 180, 180, **180**, 180, 180, 180, 181, 181, 181 …

Inutile de trier toute la série car il nous faut juste la 22ème valeur.

La médiane est **180** mois.

Comment déterminer les quartiles ?

On ordonne la série par ordre croissant.
On détermine l'effectif total : N.
On calcule $\frac{N}{4}$. Si ce n'est pas un entier, on arrondi à l'entier supérieur. Ce nombre est le rang du premier quartile.
On calcule $\frac{3}{4}\times \mathrm{N}=\frac{3\mathrm{N}}{4}$. Si ce n'est pas un entier, on arrondi à l'entier supérieur. Ce nombre est le rang du troisième quartile.

Exemple 1

On cherche les premier et troisième quartiles de la série :

5 13 27 38 4 16 32 49

La série ordonnée est : 4 ; 5 ; 13 ; 16 ; 27 ; 32 ; 38 et 49.

L'effectif total est 8.

On calcule $\frac{8}{4}$ = 2 et $\frac{3}{4} \times 8$ = 6.

Le premier quartile est la 2ème valeur de la série ordonnée ; ici c'est 5.
Le troisième quartile est la 6ème valeur de la série ordonnée ; ici c'est 32.

Exemple 2

On cherche les premier et troisième quartiles de la série :

7 9 **15** 18 19 25 **29** 32 37

La série est déjà ordonnée.

L'effectif total est 9.

*On calcule $\frac{9}{4}$ = 2,25 ; on arrondit à **3**.*

*On calcule $\frac{3}{4} \times 9$= 6,75 ; on arrondit à **7**.*

Le premier quartile est la **3**ème valeur de la série ordonnée ; ici c'est 15.
Le troisième quartile est la **7**ème valeur de la série ordonnée ; ici c'est 29.

Exemple des âges

La série ordonnée commence par : 172, 173, 173, 173, 173, 174, 174, 174, 175, 175, 176, 176, 177, 178, 178, 178, 179, 179, 179, 180, 180, 180, 180, 180, 180, 181, 181, 181, 182, 182, 182, 183, 183, 183, 183, 184, 184, 186, …

L'effectif total est 43.

On calcule $\frac{43}{4}$ = 10,75 ; on arrondit à 11.

On calcule $\frac{3}{4} \times 43$ = 32,25 ; on arrondi à 33.

Le premier quartile est la 11ème valeur de la série ordonnée ; ici c'est 176.
Le troisième quartile est la 33ème valeur de la série ordonnée ; ici c'est 183.

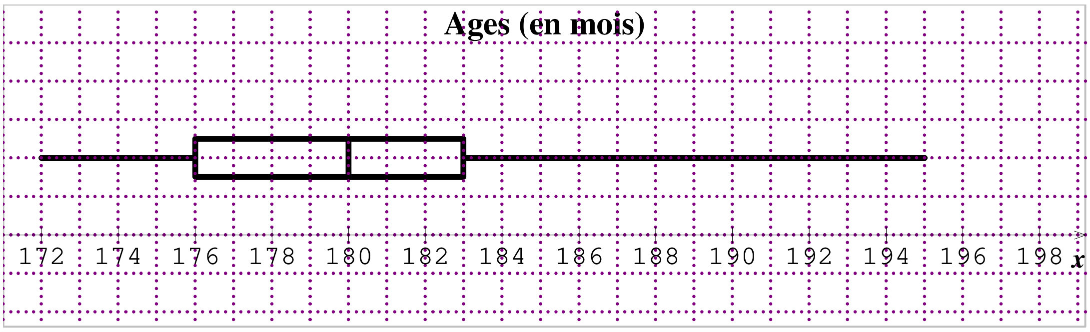
Utilisation d'un tableur

Pour calculer la médiane d'une série de valeur comprise entre la cellule A3 et la cellule F5, je tape la formule =mediane(A3:F5)

Pour calculer le premier quartile d'une série de valeur comprise entre la cellule A3 et la cellule F5, je tape la formule =quartile(A3:F5;1)

Pour calculer le troisième quartile d'une série de valeur comprise entre la cellule A3 et la cellule F5, je tape la formule =quartile(A3:F5;3)

Pour les quartiles, le tableur donne une valeur qui n'est pas obligatoirement une valeur de la série …
Certaines fois, on présente ses valeurs sur une « boite à moustaches » (ou box plot).
En voici une (simplifiée) pour la série des âges :



Comment calculer la moyenne

Méthode 1

1. On additionne toutes les valeurs.
2. On divise par l'effectif total.

Exemple des âges

Soit M la moyenne de cette série.

M = (180 + 176 + 179 + 176 + 182 + 178 + 184 + 181 + 179 + 173 + 182 + 187 + 175 + 181 + 174 + 183 + 178 + 180 + 178 + 184 + 173 + 175 + 173 + 180 + 195 + 179 + 174 + 182 + 172 + 174 + 195 + 183 + 192 + 190 + 181 + 173 + 177 + 180 + 183 + 180 + 183 + 186 + 180) ÷ 43

= 7750 ÷ 43 ≈ 180,2 mois.

L'âge moyen est de 7750 ÷ 43 ≈ 180,2 mois.

Comment calculer la « moyenne pondérée »

Méthode 2

1. On calcule la somme de toutes les valeurs en utilisant le tableau d'effectifs.
2. On divise par l'effectif total.

Exemple des âges

Age en mois	172	173	174	175	176	177	178	179	180	181	182	183	184	186	187	190	193	195
Effectif	1	4	3	2	2	1	3	3	6	3	3	4	2	1	1	1	1	2

Soit M la moyenne de cette série.

M = (172 + 4×173 + 3×174 + 2×175 + 2×176 + 177 + 3×178 + 3×179 + 6×180 + 3×181 + 3×182 + 4×183 + 2×184 + 186 + 187 + 190 + 193 + 2×195) ÷ 43

= 7750 ÷ 43 ≈ 180,2 mois.

L'âge moyen est de 7750 ÷ 43 ≈ 180,2 mois.

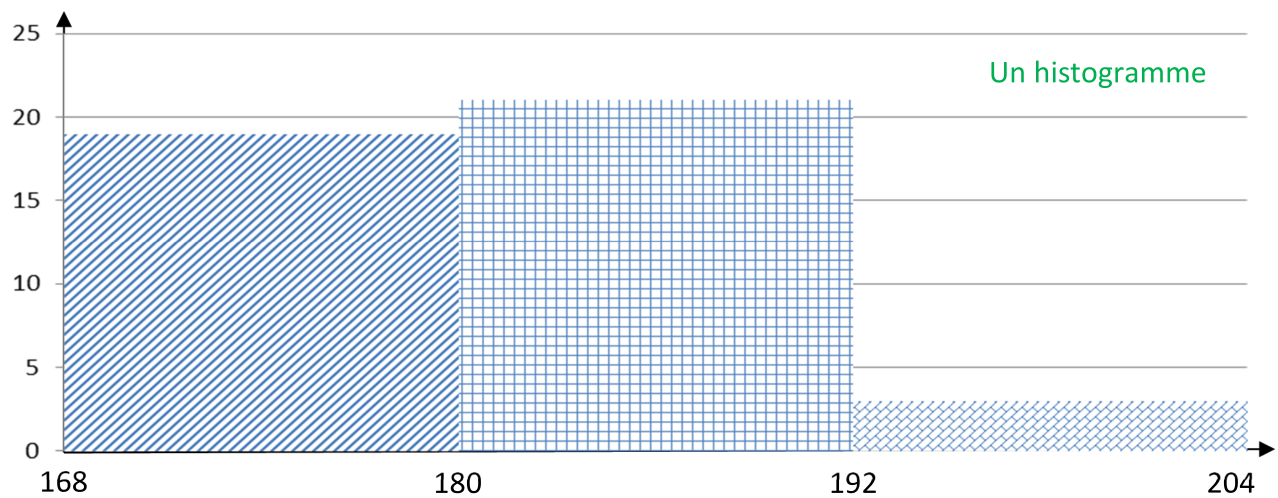
Partage en classes de valeurs

On regroupe les valeurs de la série en classes de valeurs.

Par exemple, on peut regrouper les personnes qui ont le même nombre d'années.

Age en mois	172	173	174	175	176	177	178	179	180	181	182	183	184	186	187	190	193	195
Effectif	1	4	3	2	2	1	3	3	6	3	3	4	2	1	1	1	1	2

Age en mois	Effectif	Fréquence	Fréquence en %
[168 ; 180[	19	19/43	44 %
[180 ; 192[	21	21/43	49 %
[192 ; 204[	3	3/43	7 %
Total	43	1	100 %



Définition

On appelle *centre de la classe*, le milieu de l'intervalle définissant la classe.

Exemple des âges.

Le milieu de l'intervalle [168 ; 180[est 174. Pour le calculer on effectue $\frac{168+180}{2}$.

Age en mois	Effectif	Fréquence	Fréquence en %	Centre de la classe
[168 ; 180[	19	19/43	44 %	174
[180 ; 192[	21	21/43	49 %	186
[192 ; 204[	3	3/43	7 %	198
Total	43	1	100 %	

<u>Calcul</u> d'une valeur approchée de la moyenne.
On suppose que les valeurs sont regroupées aux centres des classes.
Soit M une valeur approchée de la moyenne de cette série.
M = (19×174 + 21×186 + 3×198) ÷ 43 = 7806 ÷ 43 ≈ 181,5 mois.
Une valeur approchée de l'âge moyen est $\boxed{7806 \div 43 \approx 181{,}5 \text{ mois}}$.

Remarques

Le résultat est très bon.
La précision de la mesure est le mois et on trouve 1,3 mois d'écart avec la valeur exacte.
Cette méthode (malgré sa forte approximation) donne souvent de très bons résultats.

Exemple des âges

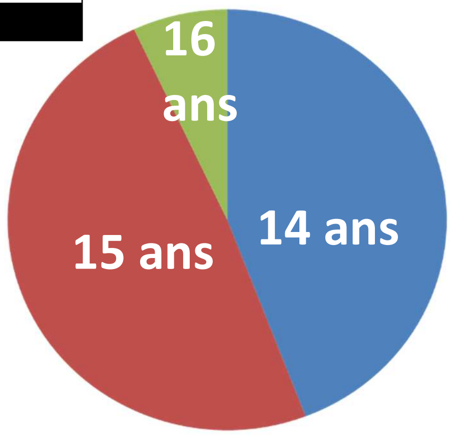
Age en mois	Effectif	Fréquence	Fréquence en %	Angle sur le diagramme circulaire en degré	Centre de la classe
[168 ; 180[	19	19/43	44 %	159°	174
[180 ; 192[	21	21/43	49 %	176°	186
[192 ; 204[	3	3/43	7 %	25°	198
Total	43	1	100 %	360°	

×3,6

×360

$\times \frac{360}{43}$

Un diagramme circulaire



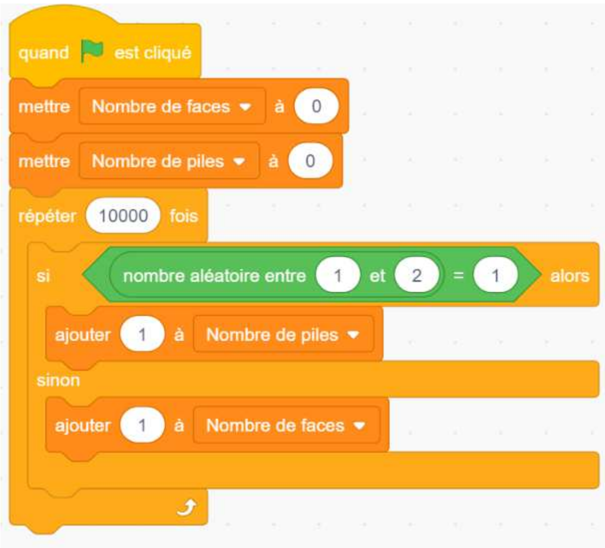
https://www.lesmathsdherve.net/statistiques-parcours/

https://www.lesmathsdherve.net/statistiques-videos-2/

https://www.lesmathsdherve.net/statistiques-aide/

Exemple des pièces

On lance une pièce de monnaie.
On recommence l'expérience.
Compter le nombre de piles et le nombre de faces.



	A	B	C	D	E
1	**Nombre de piles**	=NB.SI(A3:J100;1)		**Nombre de faces**	=NB.SI(A3:J100;2)
2	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)
3	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)
4	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)
5	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)	=ALEA.ENTRE.BORNES(1;2)

Exemple des dés

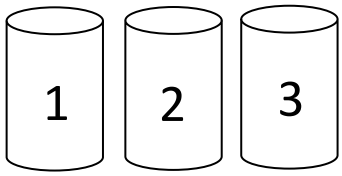
On lance deux dés numérotés de 1 à 6.
On additionne les faces visibles.
On recommence l'expérience.
On compte le nombre d'apparitions de 2, de 3, de 4, … de 12.

	A	B	C	D	E	F	G	H	I	J	K	L
1	1	2	3	4	5	6	7	8	9	10	11	12
2	=NB.SI($A$4:$L$86;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)	=NB.SI($A$4:$L$28;A$1)
3												
4	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)
5	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)	=ALEA.ENTRE.BORNES(1;6) +ALEA.ENTRE.BORNES(1;6)

Exemple des boites

Trois boites, numérotées de 1 à 3, sont posées sur une table. Une récompense est cachée sous l'une des boites.

Dans un premier temps, je choisis une boite.



Le meneur de jeu, qui sait où se trouve la récompense, me montre ensuite une boite dans laquelle il n'y a pas la récompense et me demande si je confirme mon choix ou si je veux en changer.

Le meneur de jeu me donne ce qu'il y a dans la boite que j'ai alors choisi.

Y a-t-il une stratégie plus intéressante que l'autre ?

La première idée qui vient à l'esprit est qu'il ne reste que 2 boites et donc qu'on a autant de chance de gagner en changeant qu'en conservant son choix initial.

Nous allons simuler cette expérience avec un tableur.

	A	B	C	D	E
1	Position de la récompense	Choix du candidat	Le meneur de jeu montre la boite	Si je ne change pas j'ai …	Si je change, j'ai …
2	1	3	2	perdu	gagné

En A2, je veux un nombre choisi aléatoirement entre 1, 2 et 3. Je tape ***=ALEA.ENTRE.BORNES(1;3)***

En B2, je veux un nombre choisi aléatoirement entre 1, 2 et 3. Je tape ***=ALEA.ENTRE.BORNES(1;3)***

En C2 je dois déterminer une boite non choisie par le candidat et perdante.
Si le joueur n'a pas choisi la bonne boite, je remarque que la somme des numéros des boites est 6 donc je n'ai qu'à taper ***= 6 – A2 – B2***
Si le joueur a choisi la bonne boite, il faut choisir aléatoirement une des deux autres :

- si le joueur a choisi 1, il faut que je trouve 2 ou 3 ; je tape alors ***=ALEA.ENTRE.BORNES(2;3)***
- si le joueur a choisi 2, il faut que je trouve 1 ou 3 ; je tape alors ***=1+ALEA.ENTRE.BORNES(0;1)*2***
- si le joueur a choisi 3, il faut que je trouve 1 ou 2 ; je tape alors ***=ALEA.ENTRE.BORNES(1;2)***

En C2, je tape **=si(A2<>B2 ; 6-A2-B2 ; si(A2=1; ALEA.ENTRE.BORNES(2;3);0)**
+ si(A2=2; 1+ALEA.ENTRE.BORNES(0;1)*2;0) + si(A2=3; ALEA.ENTRE.BORNES(1;2);0))

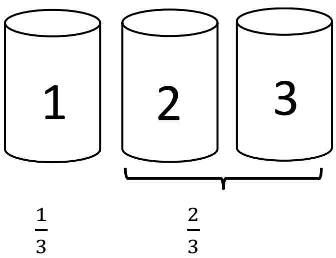
En D2 je tape =si(B2=A2 ; "gagné" ; "perdu")

En E2, je tape =si(D2="gagné" ; "perdu" ; "gagné")

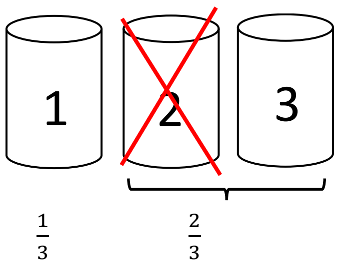
On conjecture qu'il est préférable de changer plutôt que de garder le choix initial.

Prouvons-le.

Dans l'exemple ci-dessous, on choisit la boîte n°1.



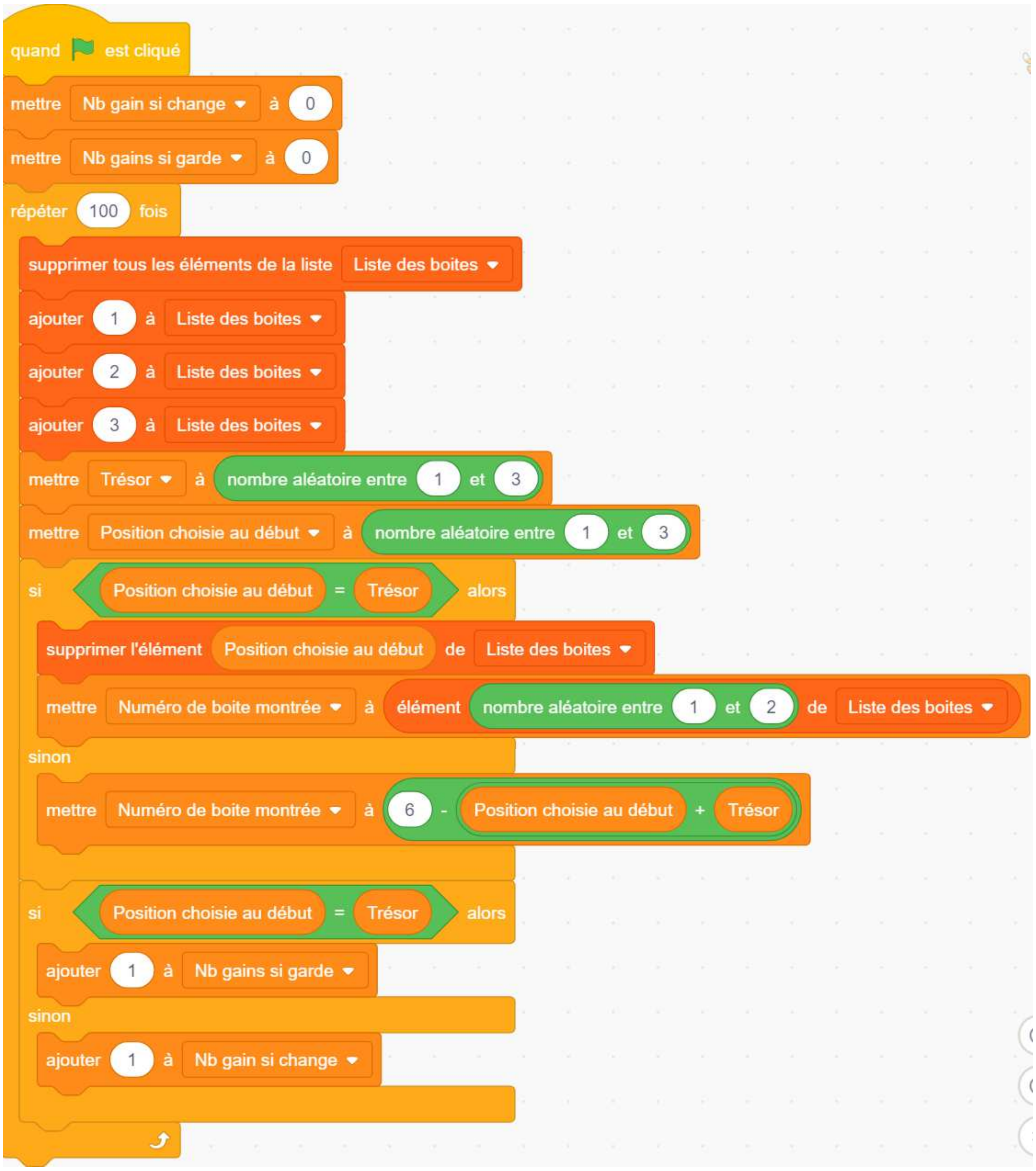
Au début, il y a 1 chance sur 3 que la récompense soit dans une boite et 2 chances sur 3 qu'elle soit dans l'autre une des deux autres boites.



Si on enlève une boite, il reste 2 chances sur 3 qu'il ne soit pas dans cette boite.
Si on sait qu'il n'est pas dans une de ces deux boites, il y a toujours 2 chances sur 3 qu'elle soit dans la boite non choisie.

Il y a donc 2 fois plus de chance de gagner en changeant de boite.

La stratégie gagnante est donc de changer de boite à chaque fois.



Si on regarde bien le programme, cette dernière condition montre que la probabilité est $\frac{2}{3}$ sans avoir besoin de le tester.

Exemple des anniversaires

On cherche à trouver la probabilité que dans un groupe d'élèves, il y ait au moins deux élèves qui aient leur anniversaire le même jour.