

# Science des données II : module 5



Approche pédagogique du cours de science des données

Philippe Grosjean & Guyliann Engels

Université de Mons, Belgique  
Laboratoire d'Écologie numérique



<https://wp.sciviews.org>  
[sdd@sciviews.org](mailto:sdd@sciviews.org)

## De quoi allons-nous discuter ?

Analyse des **avis pédagogiques** de l'an dernier => approche en classe inversée pas toujours bien comprise.



- Explication de notre approche pédagogique
- Debriefing de vos résultats Q1
- Retour et propositions de votre part

C'est quoi la classe inversée ?

Vous avez une minute ?

Pour comprendre  
La classe inversée

0:03 / 1:18



(lien vers la vidéo)

UMONS

Notre approche : **pédagogie active en classe inversée** (vous apprenez *d'abord* à la maison, nous appliquons *ensuite* en présentiel -quand on n'est pas confinés-).

*I hear and I forget.*

*I see and I remember.*

*I do and I understand.*

— Confucius

## C'est quoi la pédagogie active ?

Les  
pédagogies  
actives  
pourquoi ne  
pas essayer?



0:00 / 3:05



(lien vers la vidéo)

UMONS

# Et moi, je fais quoi dans tout cela ?

*Lisez ceci... et réagissez (question Wooclap juste après) !*



- Vous êtes **acteur de votre apprentissage**, les enseignants sont des **facilitateurs** (plus en retrait par rapport à l'approche classique).
- Plus de séparation entre **cours théorique** et **exercices** ; vos échanges avec le professeur et le ou les assistants sont similaires.
- Les **élèves-assistants** sont coachés tout autant que vous pour vous faciliter l'apprentissage de manière active.
- **Vous posez les questions**, et vos enseignants vous répondent **individuellement**.

# ECTS

**European Credits Transfer System**, créé en 1988 de manière standardisée par la Commission Européenne comme correspondant à une **charge de travail totale** pour l'étudiant de **25 à 30 heures**.

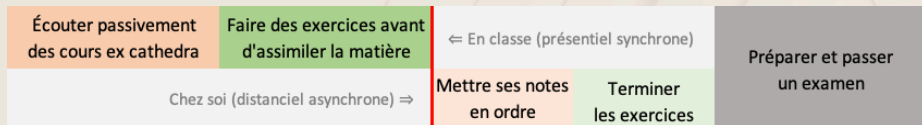


**Comment voulez-vous passer vos 25-30h/ECTS ?**

Note : 8 modules pour 3 ECTS dans notre cours, donc environ 11h par module.



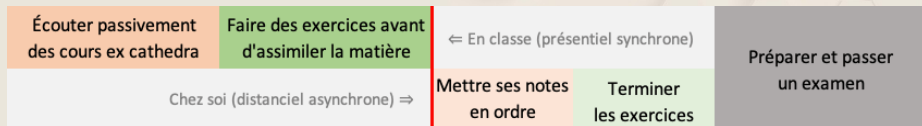
## Cours classique *ex cathedra* + séances d'exercices



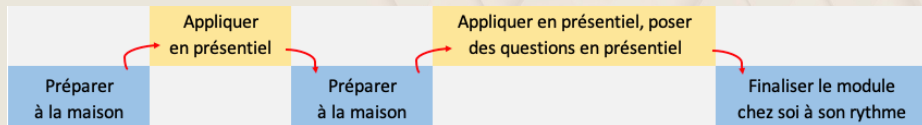
- Le réel apprentissage se déroule **après** les séances de cours et d'exercices
- Un examen est nécessaire pour vérifier vos acquis

# Optimisation du temps de travail... comparé à la classe inversée

## Cours classique *ex cathedra* + séances d'exercices



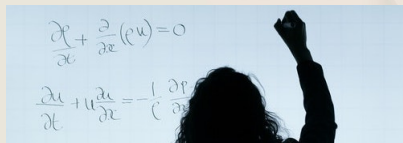
## Approche en classe inversée



- Aucune séance en présentiel sans préparation
- Chaque heure de travail pleinement consacrée à l'apprentissage
- Vous êtes actifs **tout le temps** et vous gérez à **votre rythme**
- **Pas besoin d'un examen à la fin** : travail évalué dans sa globalité

# Le professeur est un coach et un facilitateur

*Mais que font les enseignants alors ? Lisez et réagissez (question Wooclap après).*



- Le professeur (et l'assistant) ne mettent **pas** leur savoir en avant. C'est vous qui construisez votre *propre* savoir.
- Ils **ne répondent pas directement** à vos questions : ils vous mettent sur une piste et vous font réfléchir pour trouver la réponse *par vous-même*.
- Ils se mettent en retrait, mais sont **disponibles pour vous aider** (Discord, mail, etc.)

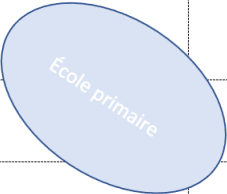
**C'est déroutant car le professeur n'apparaît plus comme l'omniscient qui transmet de manière unilatérale son savoir aux étudiants !**

## Objectifs pour l'acquisition de compétences

Processus cognitif	Types de connaissances			
	A. factuelles	B. conceptuelles	C. procédurales	D. métacognitives
1. Mémoriser				
2. Comprendre				
3. Appliquer				
4. Analyser				
5. Évaluer				
6. Créer				

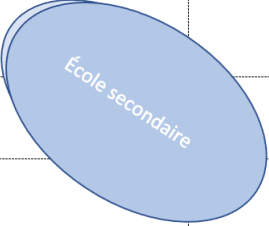
Taxonomie de Bloom modifiée par Krathwohl & Anderson.

## Objectifs pour l'acquisition de compétences

Processus cognitif	Types de connaissances			
	A. factuelles	B. conceptuelles	C. procédurales	D. métacognitives
1. Mémoriser				
2. Comprendre				
3. Appliquer				
4. Analyser				
5. Évaluer				
6. Créer				

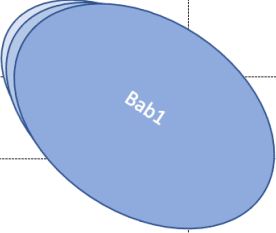
Taxonomie de Bloom modifiée par Krathwohl & Anderson.

## Objectifs pour l'acquisition de compétences

Processus cognitif	Types de connaissances			
	A. factuelles	B. conceptuelles	C. procédurales	D. métacognitives
1. Mémoriser				
2. Comprendre				
3. Appliquer				
4. Analyser				
5. Évaluer				
6. Créer				

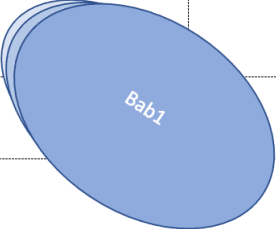
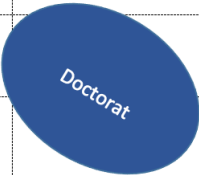
Taxonomie de Bloom modifiée par Krathwohl & Anderson.

# Objectifs pour l'acquisition de compétences

Processus cognitif	Types de connaissances			
	A. factuelles	B. conceptuelles	C. procédurales	D. métacognitives
1. Mémoriser				
2. Comprendre				
3. Appliquer				
4. Analyser				
5. Évaluer				
6. Créer				

Taxonomie de Bloom modifiée par Krathwohl & Anderson.

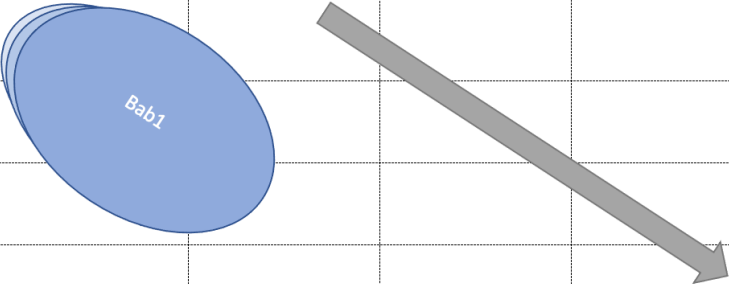
# Objectifs pour l'acquisition de compétences

Processus cognitif	Types de connaissances			
	A. factuelles	B. conceptuelles	C. procédurales	D. métacognitives
1. Mémoriser				
2. Comprendre				
3. Appliquer				
4. Analyser				
5. Évaluer				
6. Créer				

Taxonomie de Bloom modifiée par Krathwohl & Anderson.

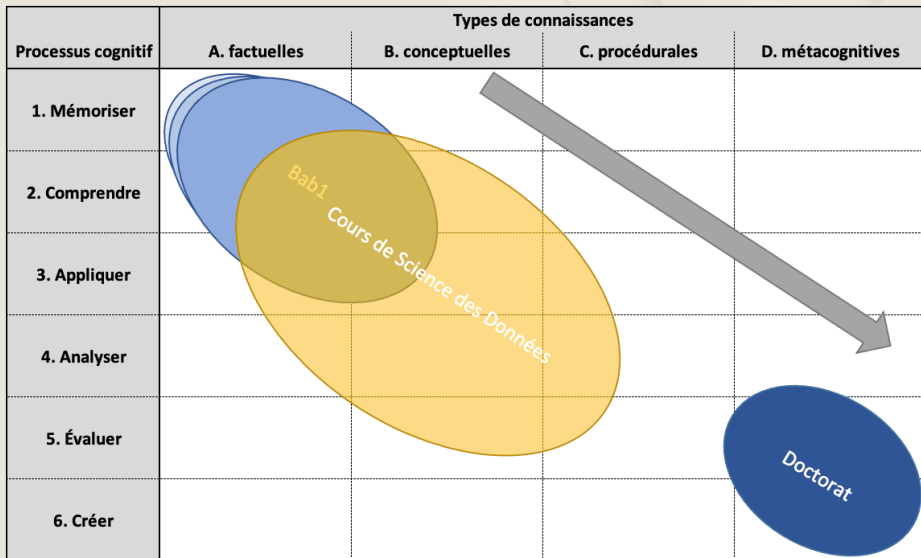


# Objectifs pour l'acquisition de compétences

Processus cognitif	Types de connaissances			
	A. factuelles	B. conceptuelles	C. procédurales	D. métacognitives
1. Mémoriser				
2. Comprendre				
3. Appliquer				
4. Analyser				
5. Évaluer				
6. Créer				

Taxonomie de Bloom modifiée par Krathwohl & Anderson.

# Objectifs pour l'acquisition de compétences



Taxonomie de Bloom modifiée par Krathwohl & Anderson.

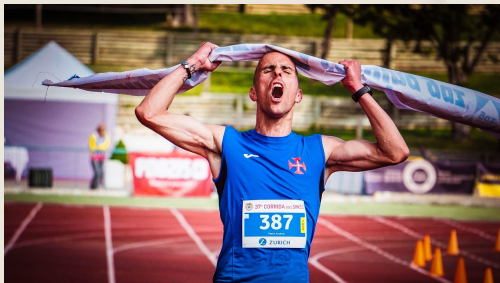
# Vous avez la parole !



- Répondez aux questions (pages Wooclap suivantes), levez la main pour prendre la parole. La discussion est ouverte.

## Debriefing des notes du Q1

**Bravo, vous avez bien travaillé !**



- 9 étudiants sur 10 ont fait (quasiment) tous les exercices
- Ceux qui ont fait tous les exercices ont une note bonne à excellente
- Travail important pour y arriver... on en est conscients !

## Erreurs récurrentes : entre-aide et citation des sources

Entre-aidez-vous, nous l'encourageons, mais **citez vos sources**.



### Cas 1 :

- **Josua** découvre une analyse intéressante.
- Il l'explique à **Roxane**, la même analyse se retrouve dans deux rapports.
- **Roxane** cite ses sources [**Josua, comm. pers.**].
- Josua a des points pour avoir aidé Roxane.
- Roxane a des points pour avoir correctement cité ses sources.

## Erreurs récurrentes : entre-aide et citation des sources

Entre-aidez-vous, nous l'encourageons, mais **citez vos sources**.

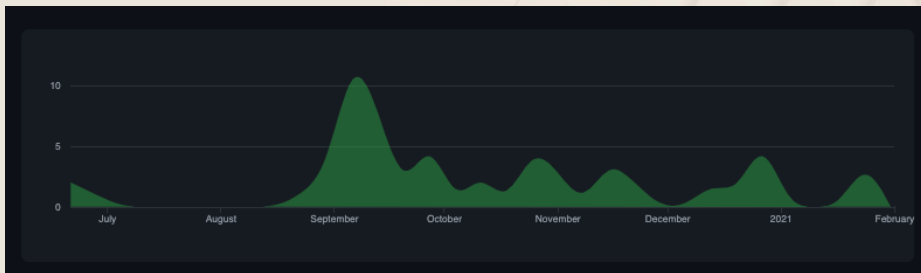


### Cas 2 :

- **Peter** partage son code R avec **Déborah**, la même analyse se retrouve dans deux rapports.
- Personne ne cite ses sources.
- Cela apparaît comme du **plagiat** (ne sera plus pardonné au Q2 !)

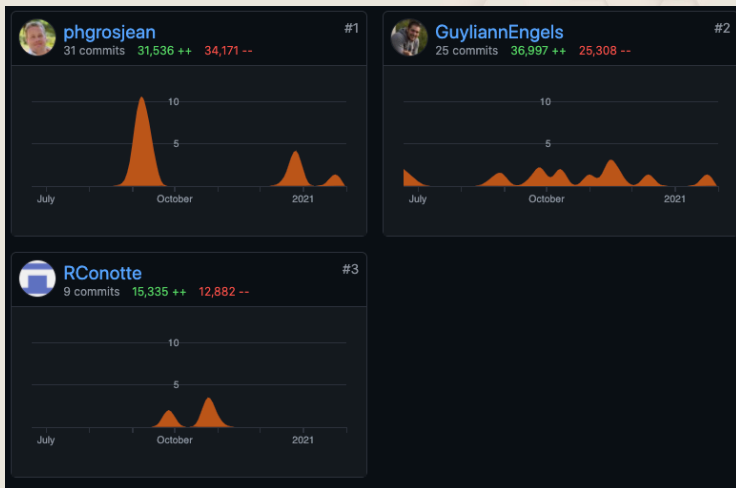
## Erreurs récurrentes : répartitions du travail

GitHub propose des outils pour visualiser l'activité d'un dépôt (Insight -> Contributors).



L'activité globale peut être **divisée par collaborateur**.

L'activité de chaque membre est **disponible**.



*Vérifiez dans vos projets que votre activité est bien prise en compte.*



## Erreurs récurrentes : ne pas vérifier ce qu'on fait

Partant des données suivantes :

```
pop2020 <- tibble(  
  age = factor(c("Moins de 18 ans", "18 à 64 ans", "65 ans et plus"),  
               levels = c("Moins de 18 ans", "18 à 64 ans", "65 ans et plus"),  
  population = c(2312040, 6976123, 2204478))  
knitr::kable(pop2020, col.names = c("Groupe d'age", "Nombre de personnes"),  
              caption = "Population au premier septembre 2020 en Belgique")
```

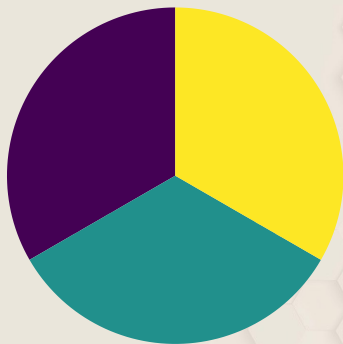
**Table 1:** Population au premier septembre 2020 en Belgique

Groupe d'age	Nombre de personnes
Moins de 18 ans	2312040
18 à 64 ans	6976123
65 ans et plus	2204478

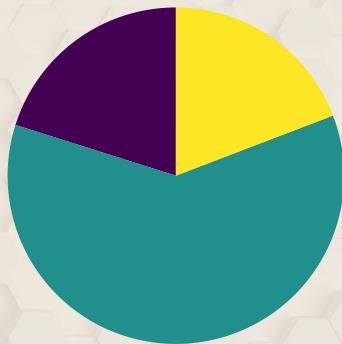
Quel est le graphique correct ? Pourquoi l'autre est-il faux ?

age     Moins de 18 ans     18 à 64 ans     65 ans et plus

**A**

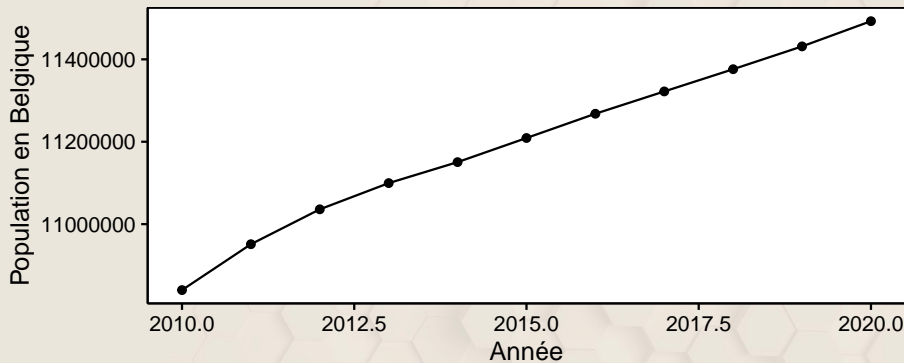


**B**

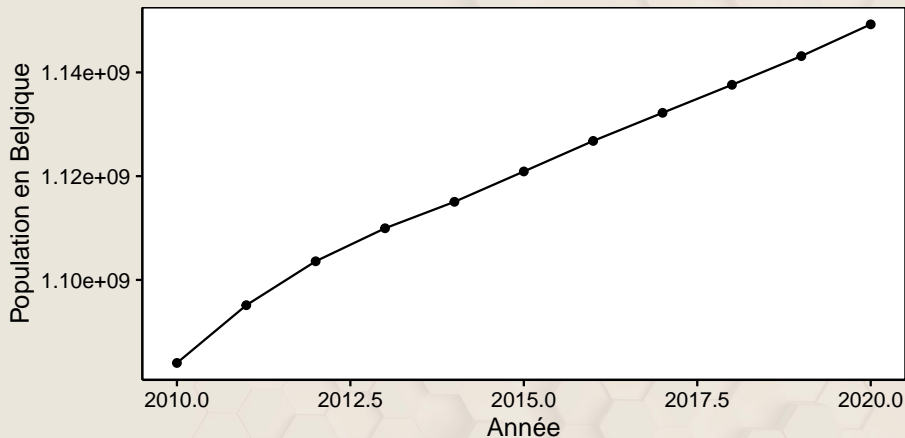


## Erreurs récurrentes : axes des graphiques

```
chart(pop, density ~ year) +  
  geom_point() +  
  geom_line() +  
  labs(y = "Population en Belgique", x= "Année")
```

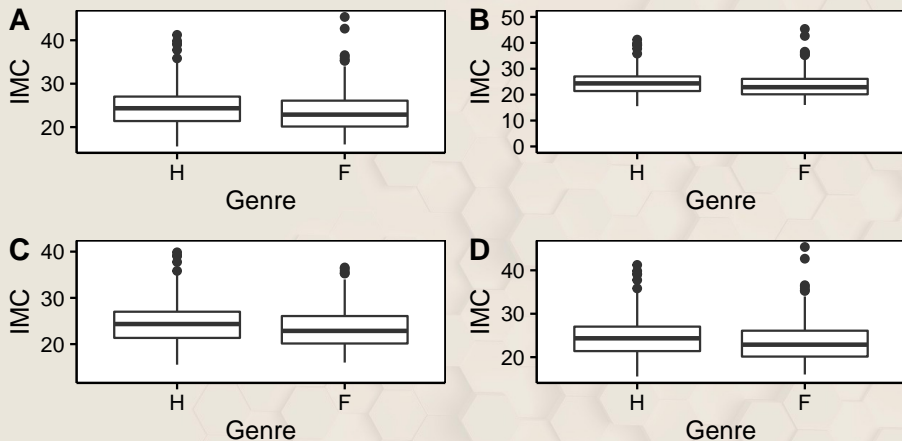


Repérez l'erreur sur le graphique ci-dessous.



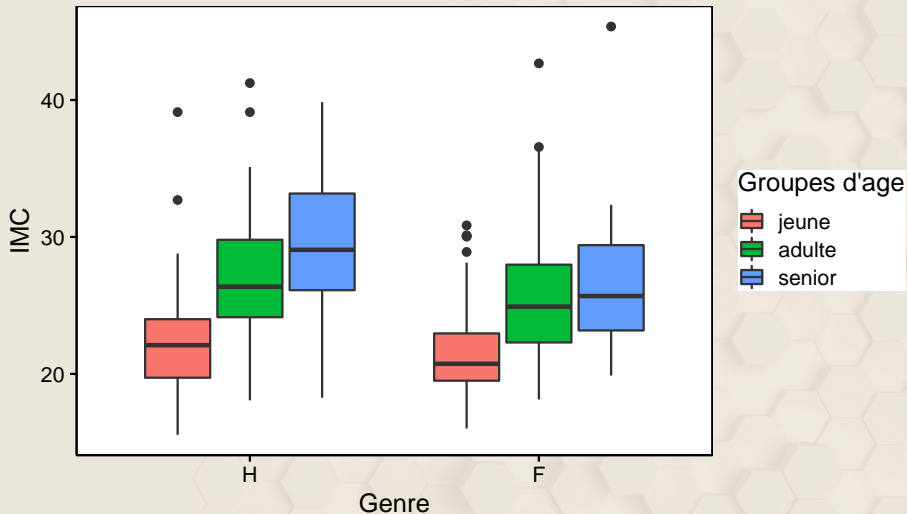
## Erreurs récurrentes : graphiques multiples inopportuns

Variation de l'indice de masse corporelle en fonction du genre et du groupe d'âges.



*Ce graphique multiple n'est pas interprétable.*

Ce graphique présente la même information que le graphique multiple précédent.



*Un graphique multiple ne doit pas être utilisé sur des données indissociables.*

## Erreurs récurrentes : valeurs manquantes

```
drop <- tibble(  
  x = runif(10),  
  y = runif(10),  
  z = c(5, 3, rep(NA, 6), 3, 2)  
)  
knitr::kable(drop[1:3,])
```

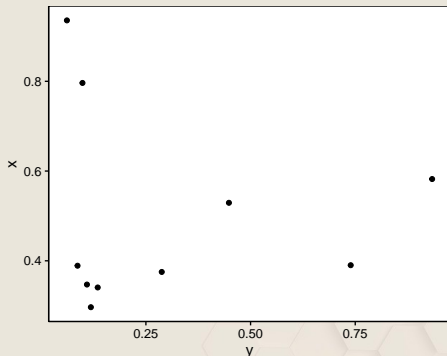
x	y	z
0.5822085	0.9336553	5
0.2965706	0.1185422	3
0.7963689	0.0985016	NA

Élimination naïve des données manquantes.

```
drop1 <- drop_na(drop)
```

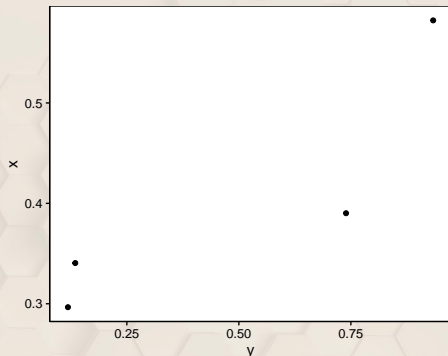
## Utilisation du tableau drop

```
chart(drop, x ~ y) +  
  geom_point()
```



## Utilisation du tableau drop1

```
chart(drop1, x ~ y) +  
  geom_point()
```

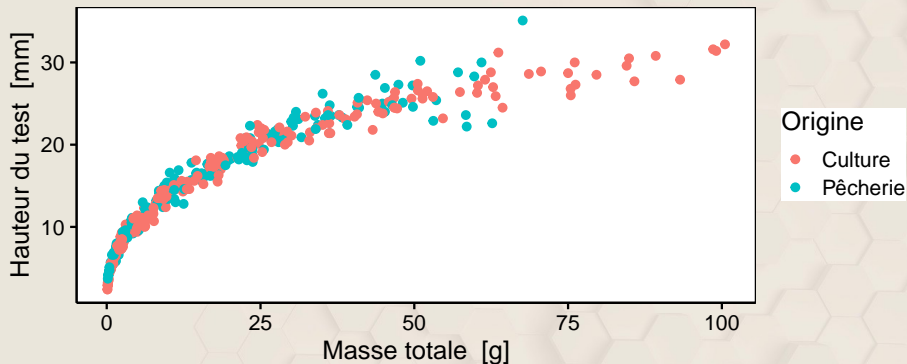


*Soyez très vigilant avec les NA. Les variables  $x$  et  $y$  ne contenaient pas de valeurs manquantes au départ ici !*



## Erreurs récurrentes : interprétation simpliste

L'interprétation d'un graphique est importante. Elle met en avant votre compréhension de la situation.



Description/interprétation simpliste : “les oursins proviennent du milieu naturel et d'élevage.”

## Attentes au Q2

- **Corriger et ne plus commettre** toutes ces erreurs.
- La **réalisation technique** d'un graphique ou d'une analyse ne sera plus suffisante (taxonomie de Bloom : **appliquer**). Vous allez maintenant vous focaliser sur votre capacité à **analyser** vos résultats.



# Vous avez la parole !



- Répondez aux questions (page Wooclap suivante), levez la main pour prendre la parole.