



# Les Ateliers Archéomatiques

<http://isa.univ-tours.fr/spip.php?rubrique245>

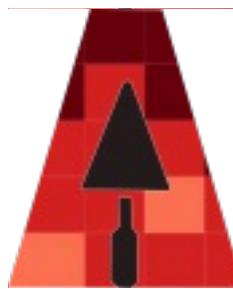
---

Lundi 27 et Mardi 28 Mars 2017 - MSH Val-de-Loire

# Atelier Archéomatique 17.1

<http://isa.univ-tours.fr/spip.php?article355>

---



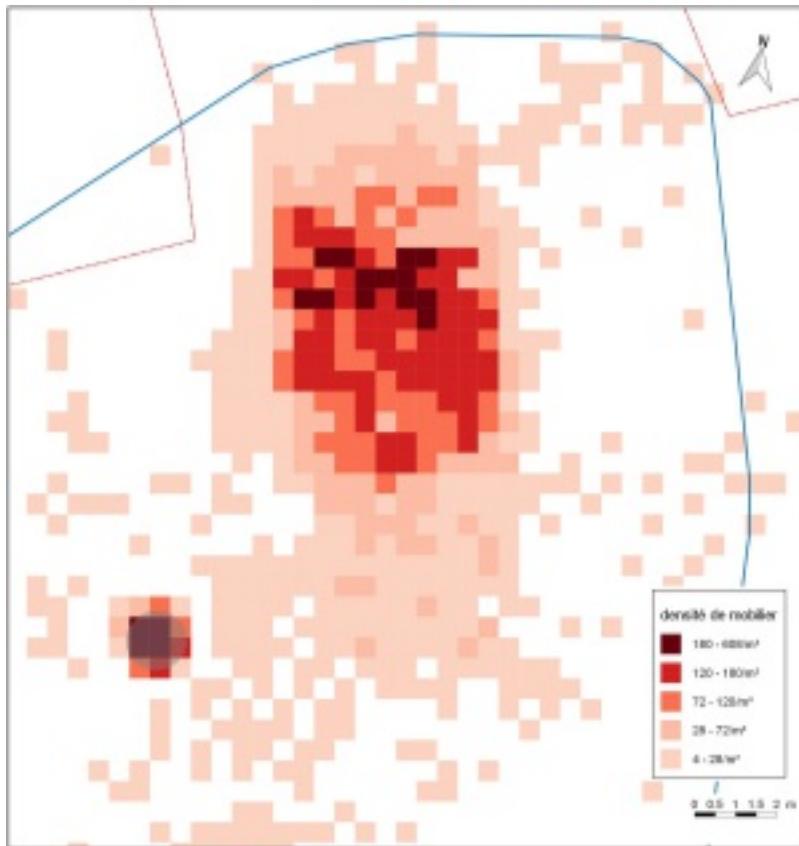
**L'analyse par maille en archéologie:  
un outil d'aide à l'analyse et de représentation  
cartographique quantitative.**

Sylvain Badey

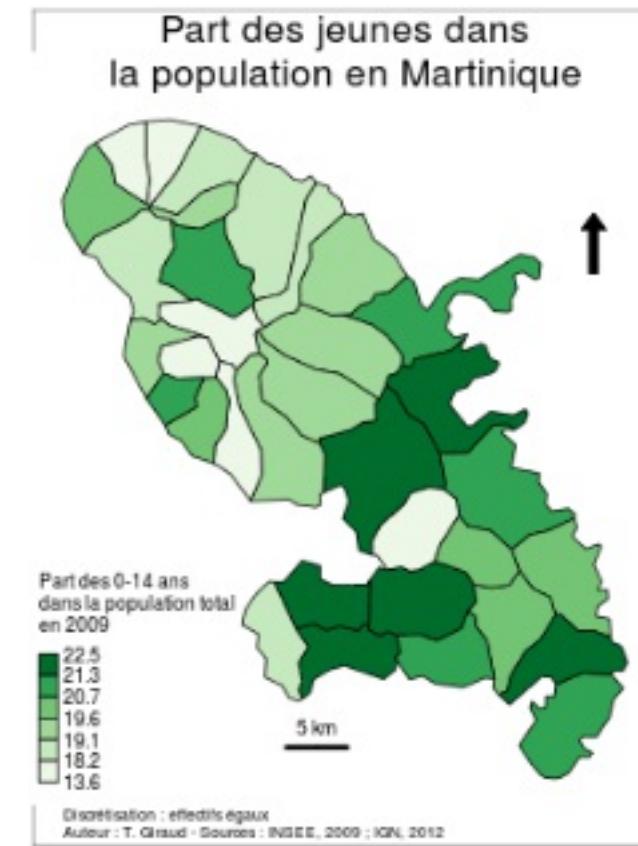


# Introduction

## Analyse par maille



## Carte choroplète

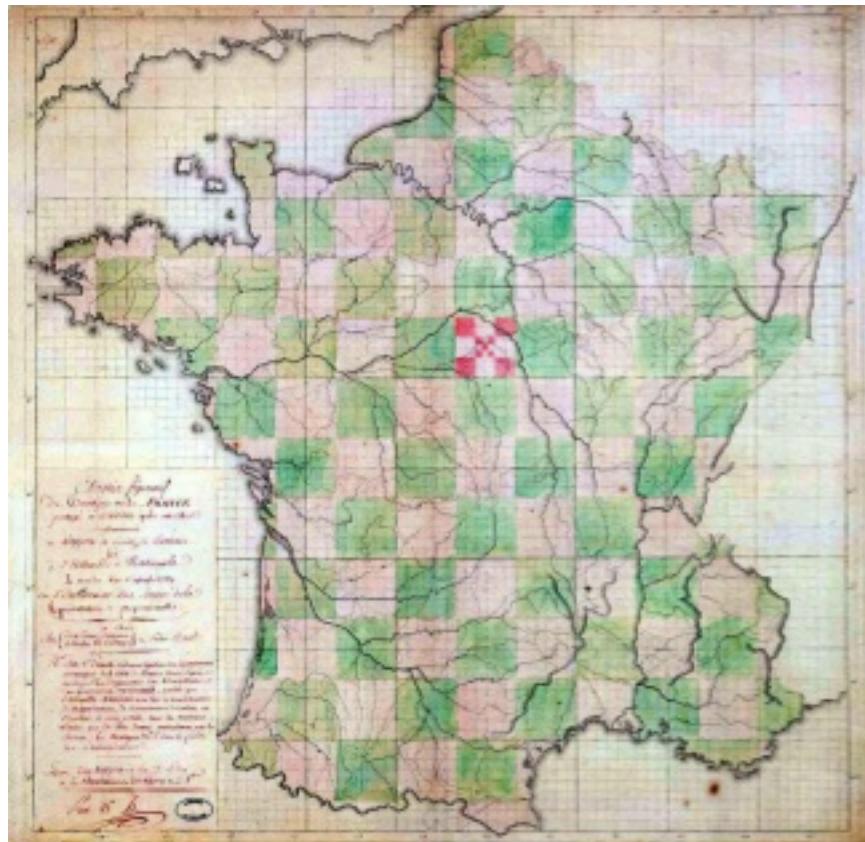


# Pourquoi faire une analyse par maille ?

---

Parce que le territoire étudié est déjà partitionné par un carroyage

*projet de découpage géométrique de la France en 89 départements formés chacuns de 9 communes, proposé par l'Abbé Seyiès à l'assemblée nationale en 1789*



# Pourquoi faire une analyse par maille ?

---

Parce que le territoire étudié est déjà partitionné par un carroyage

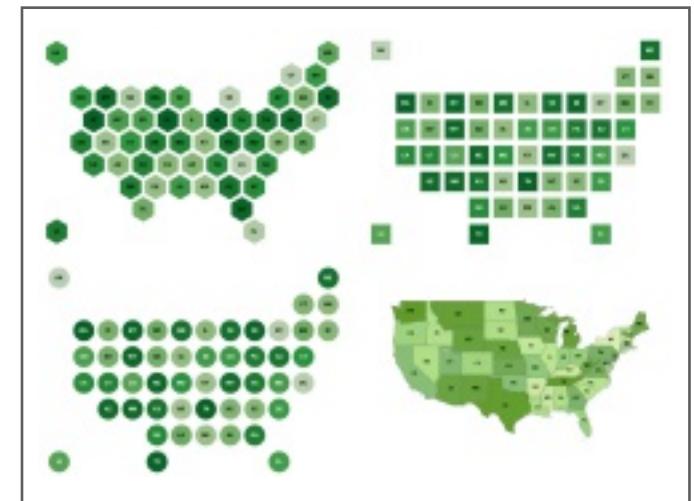
*Système de carroyage à Épieds-en-Beauce (Loiret), 2010. Un carroyage de 1 m de côté a été mis en place sur la totalité de l'enclos (400 m<sup>2</sup>) afin de fouiller la couche de rejet de bûcher funéraire.*



# Les étapes indispensables

---

- **définir la maille (taille, géométrie et position)**
- choisir l'information à représenter (brute ou calculée)
- découper les classes (méthodes de discrétisation)
- cartographier (règles de la sémiologie graphique)

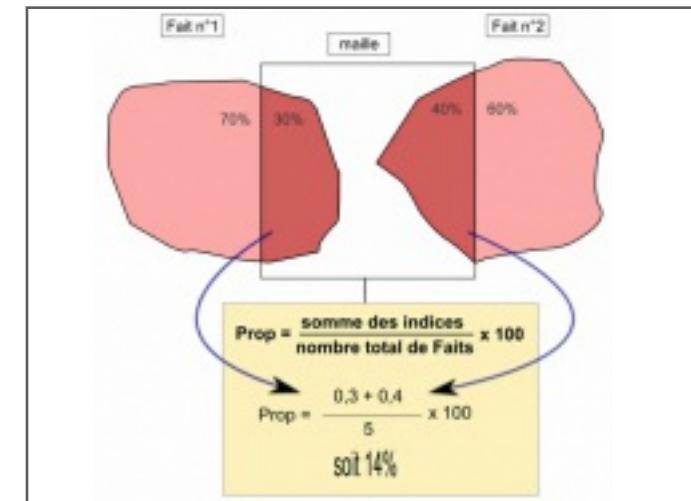


<https://clarity-us.com/wp-content/uploads/2016/02/blog-hex-maps-4.png>

# Les étapes indispensables

---

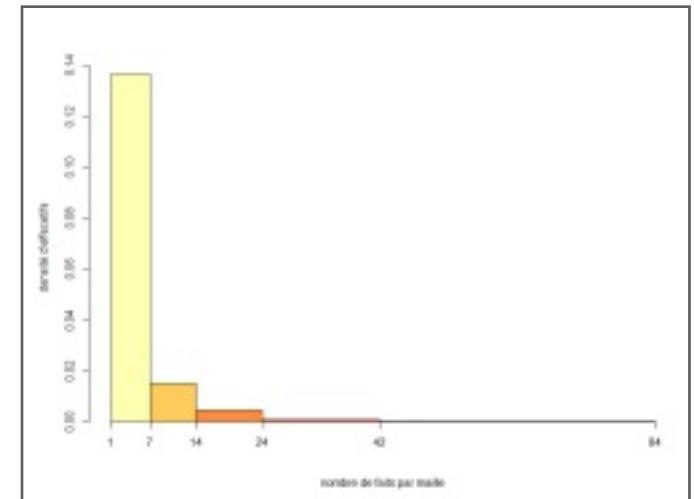
- définir la maille (taille, géométrie et position)
- choisir l'information à représenter (brute ou calculée)
- découper les classes (méthodes de discréttisation)
- cartographier (règles de la sémiologie graphique)



# Les étapes indispensables

---

- définir la maille (taille, géométrie et position)
- choisir l'information à représenter (brute ou calculée)
- découper les classes (méthodes de discrétisation)
- cartographier (règles de la sémiologie graphique)



# Les étapes indispensables

---

- définir la maille (taille, géométrie et position)
- choisir l'information à représenter (brute ou calculée)
- découper les classes (méthodes de discrétisation)
- cartographier (règles de la sémiologie graphique)



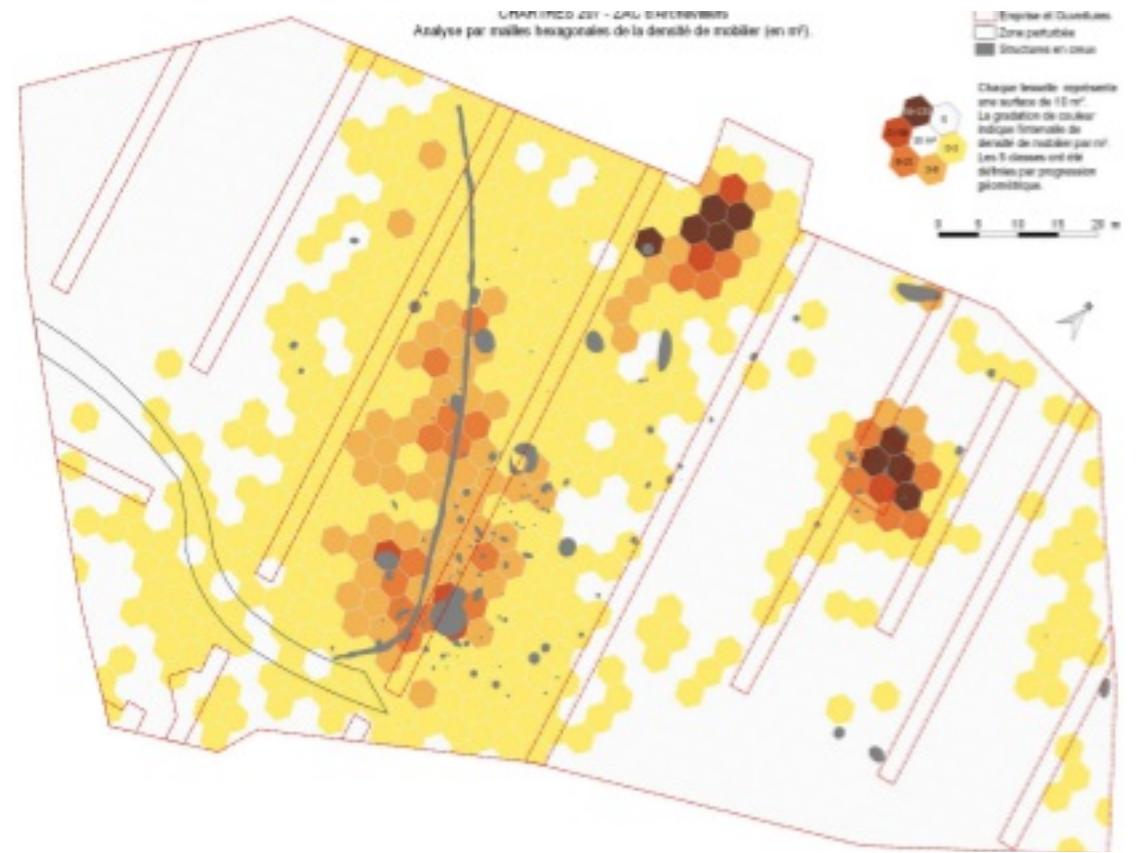
Jacques Bertin, Sémiologie graphique, 1973; Design by Jurriaan Schrofer



# Définir la maille: la forme

## La maille hexagonale

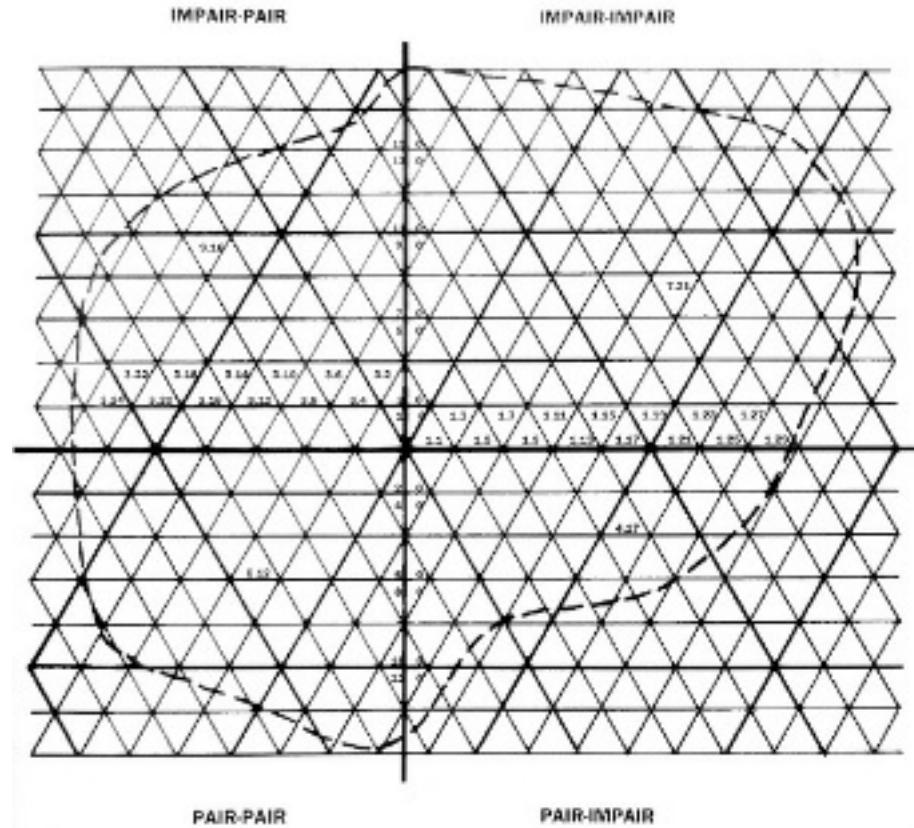
*Étude des concentrations de mobilier de la fouille Chartres 207 - ZAC d'archevilliers (RO: M.F. Creusillet) à l'aide d'une analyse par maille hexagonale dans laquelle chaque tesselle fait 10m<sup>2</sup>.*



# Définir la maille: la forme ni carré, ni hexagone ?

*Repérage triangulaire pour la fouille  
subaquatique d'installations de l'An Mil du  
Lac de Paladru . L'unité de base est un  
triangle équilatéral de 5 mètres de côté  
divisé en 25 petits triangles métriques. Le  
repérage se fait à partir d'un point 0 central  
et d'un axe nord-sud.*

Note: 6 triangles équilatéraux = 1 hexagone.



# Définir la maille: et une maille 3D, c'est possible ?

---

*Répartition des surfaces de faits mis au jour lors d'un diagnostic. La variable hauteur (basée sur la même variable quantitative utilisée pour définir les classes de valeurs) est ici utilisée à la fois pour accentuer visuellement les concentrations et pour pouvoir comparer les valeurs absolues de la variables (différence de hauteur entre deux mailles de la même couleur).*



# Le diagnostic de la ZAC d'Ozans, commune d'Etrechet (36)

---

Projet d'aménagement d'une plate-forme multimodale de 800 ha.

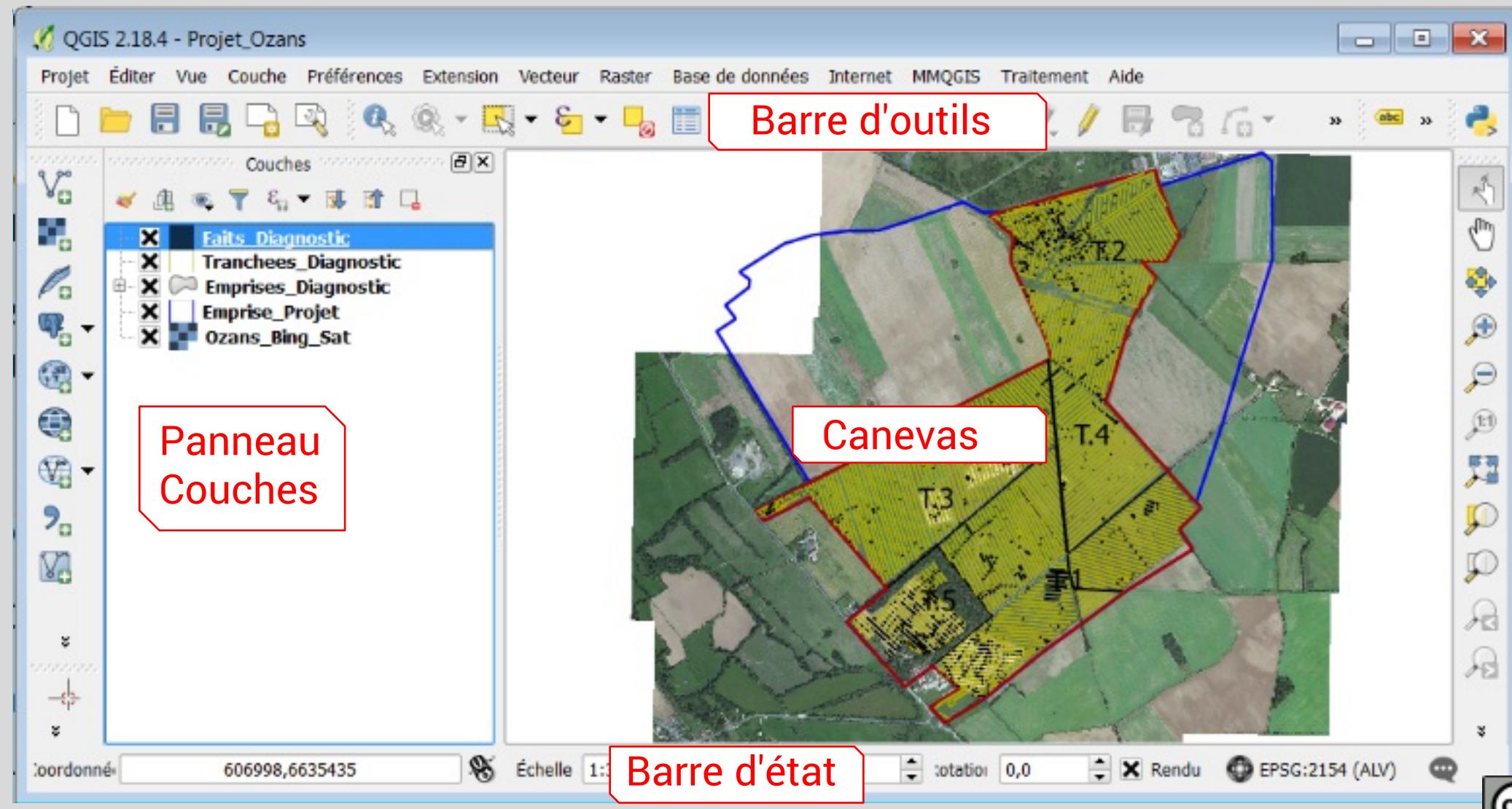
1ère phase de prescriptions de diagnostics archéologiques, divisée en 5 tranches.

412 tranchées mécaniques continues:

- 2.5m de large
- espacement de 18m
- longueur totale cumulée d'environ 124 km
- surface totale décapée de 248 032 m<sup>2</sup>
- taux global d'ouverture de 11%

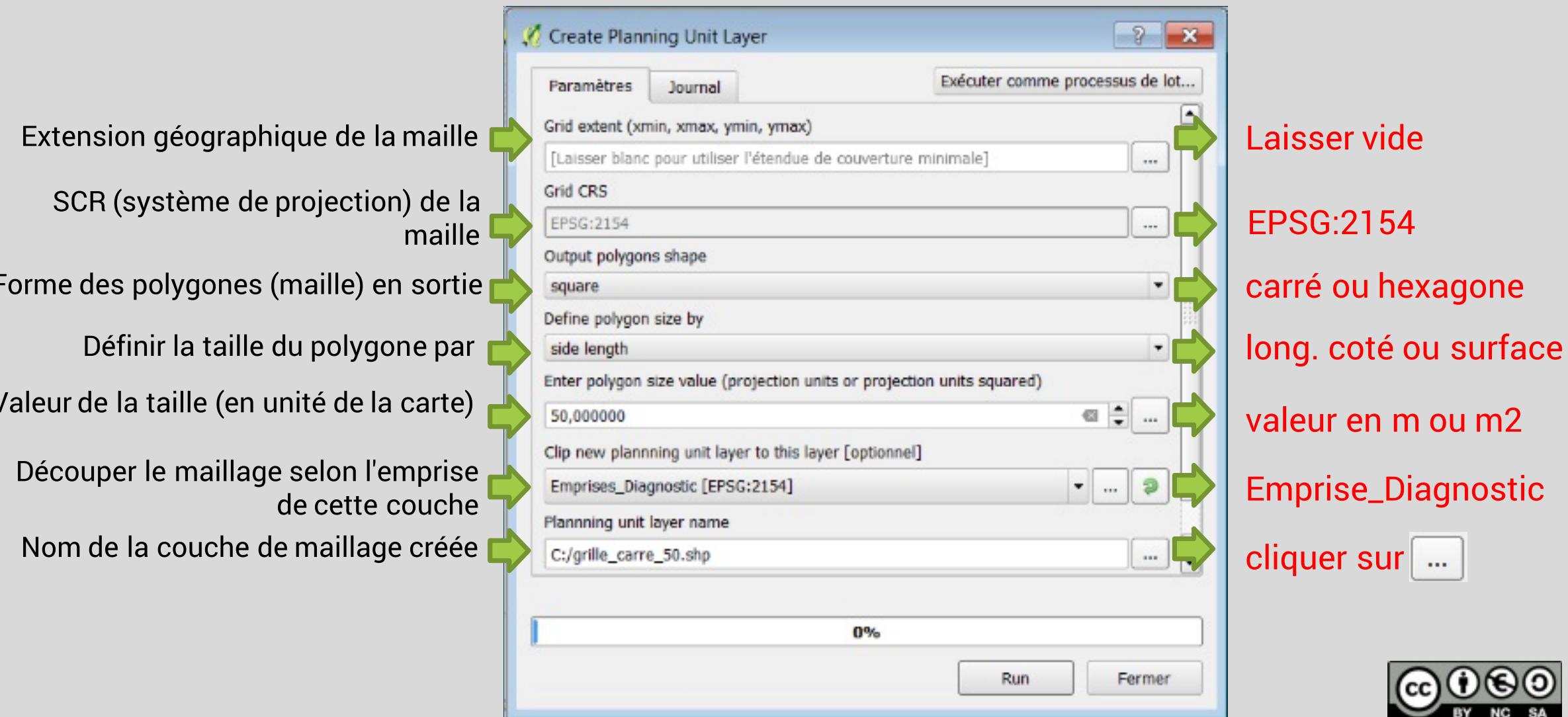


# Projet\_Ozans.qgs: L'interface générale de QGIS



# Projet\_Ozans.qgs: Crée un maillage

Traitement > Boîte à Outils > Qmarxan Tools Algorithms > Data preparation > **Create Planning Unit Layer**



Projet\_Ozans.qgs: transformation des polygones en centroïdes puis comptage des points dans la maille.

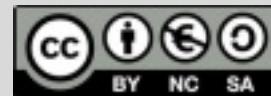
---

**Créer une couche de points d'après le centroïdes de polygones:**

Menu Vecteur ➔ Outils de géométrie ➔ Centroïdes de polygones

**Compter le nombre de points dans une maille:**

Menu Vecteur ➔ Outils d'analyse ➔ Compter les points dans les polygones



# Projet Ozans.qgs: Jointure par localisation

(récupérer un résumé quantitatif des faits dans chaque maille)

Vecteur > Outils de gestion de données > Jointure par localisation

couche dans laquelle on veut récupérer la donnée (maille)

couche contenant les données (faits)

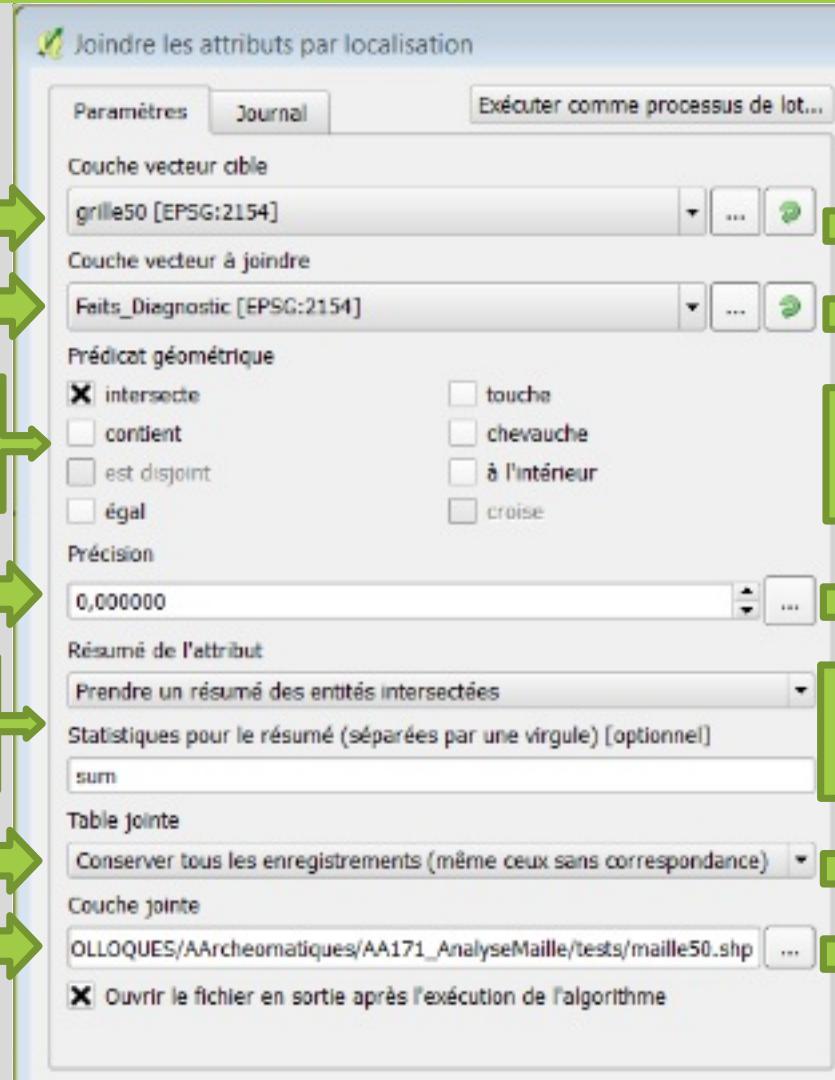
Relation topologique entre la 1ère couche et la 2ème

Précision ??

Calculs (résumés statistiques) que l'on veut récupérer

Conserver (ou non) les mailles n'ayant pas de données

Nom de la couche de maillage créée



couche de maillage

couche de faits

intersecte = contient et chevauche

rien

sum (somme des attributs des entités intersectées)

conserver tous les enregistrements

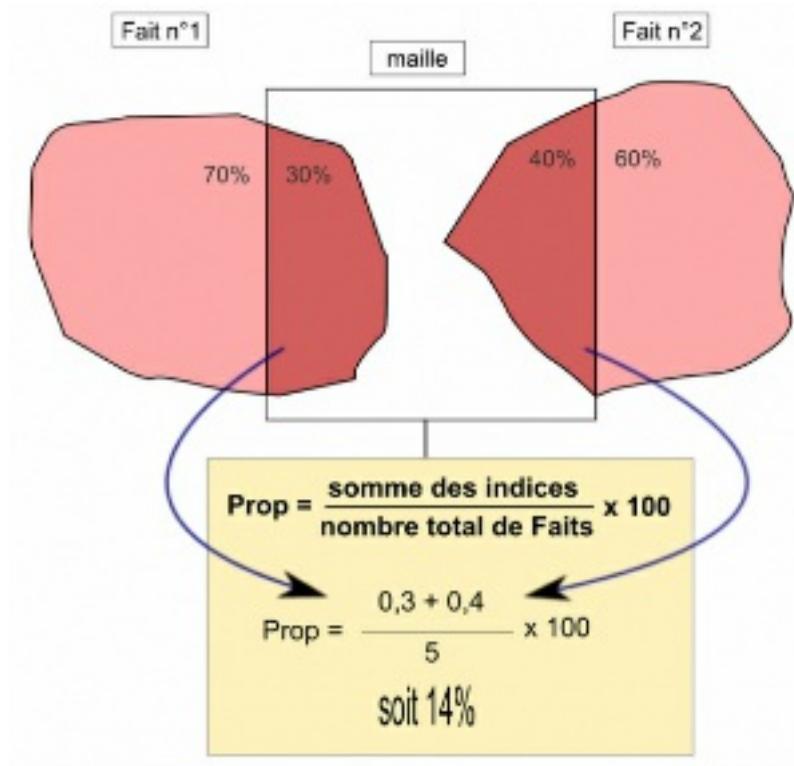
cliquer sur ...

# Valeur de la maille: carroyage ventilé

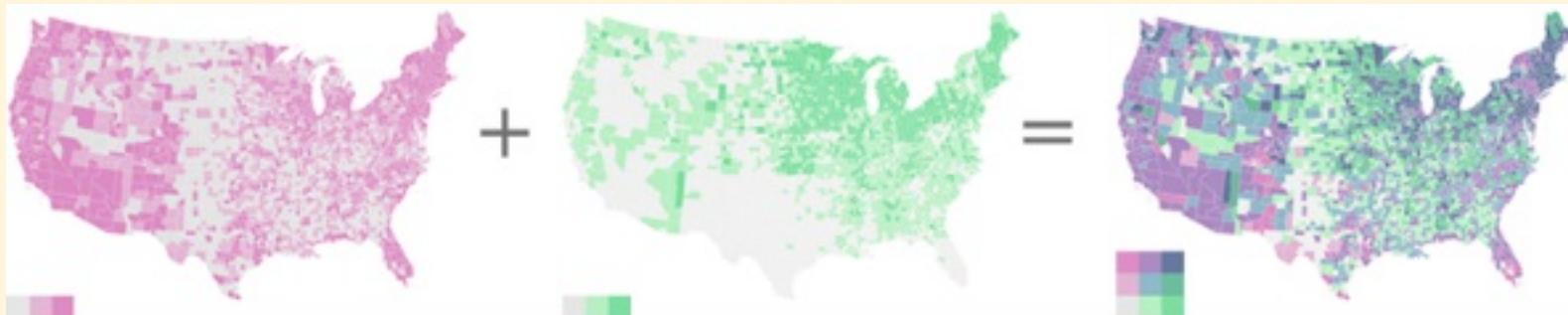
---

*Analyse par maille du diagnostic d'Ozans.  
Dans chaque maille est représenté une  
quantité relative de surface de faits cad un  
pourcentage de surface de faits rapporté au  
nombre total de faits.*

*Cet indice permet de comparer des analyses  
par mailles par périodes entre elles.*



# Valeur de la maille: et si on représentait 2 valeurs ?



*Il est possible d'agréger 2 cartes choroplèthes univariées avec 3 classes chacunes en une carte choroplète bivariée à 9 classes. mais il faut:*

- Vérifier la pertinence de la compilation des deux variables.
- Limiter le nombre de classes pour chaque variable (3 classes par variables idéalement).
- Bien choisir le double camaïeu de couleur



# Projet\_Ozans.qgs: faire une carte "graduée"

Clic droit sur la couche de maille > Propriétés > Onglet Style

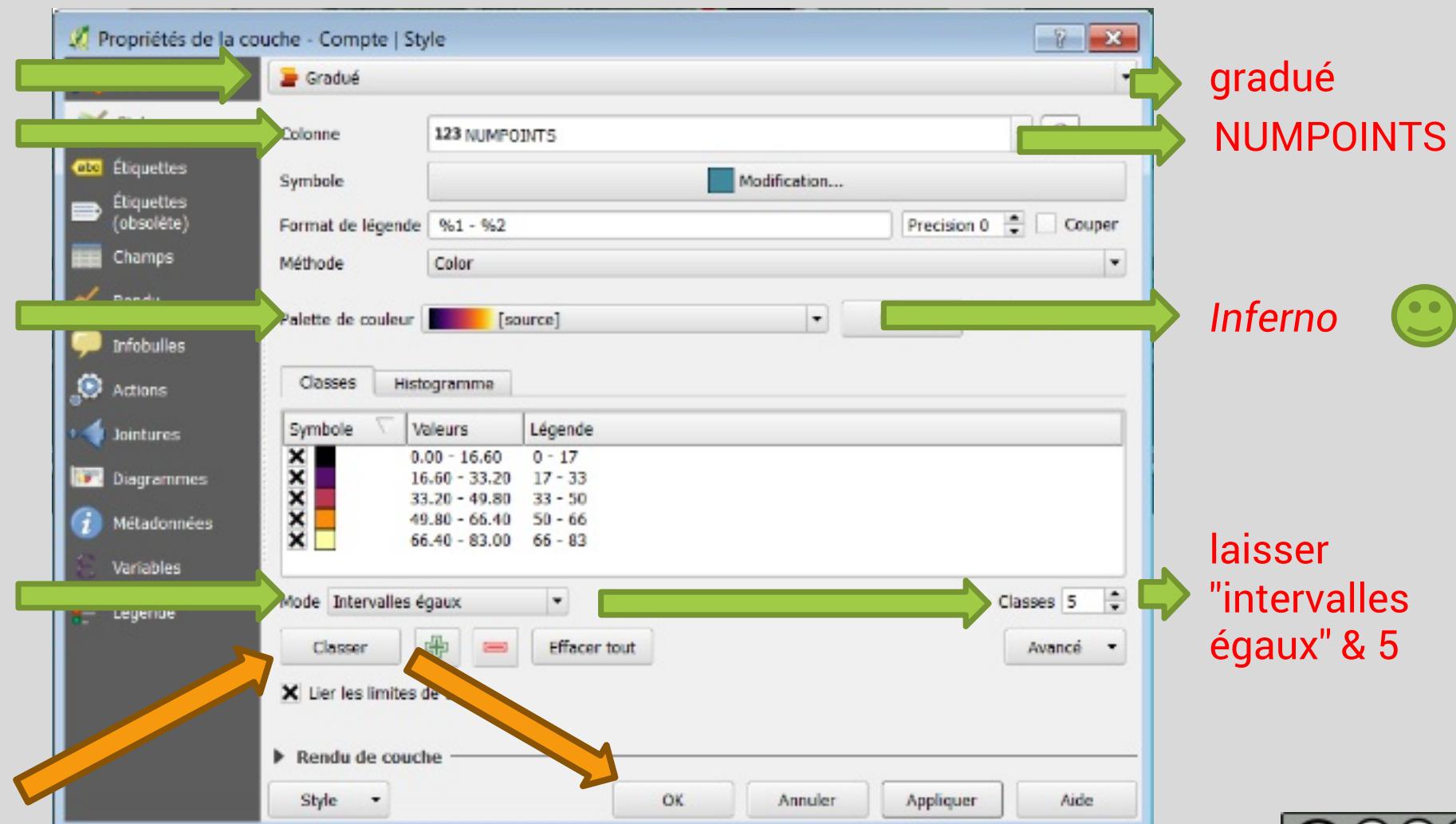
Style (simple, catégorisé, gradué, etc.)

Champs contenant la donnée quantitative

Palette de couleur

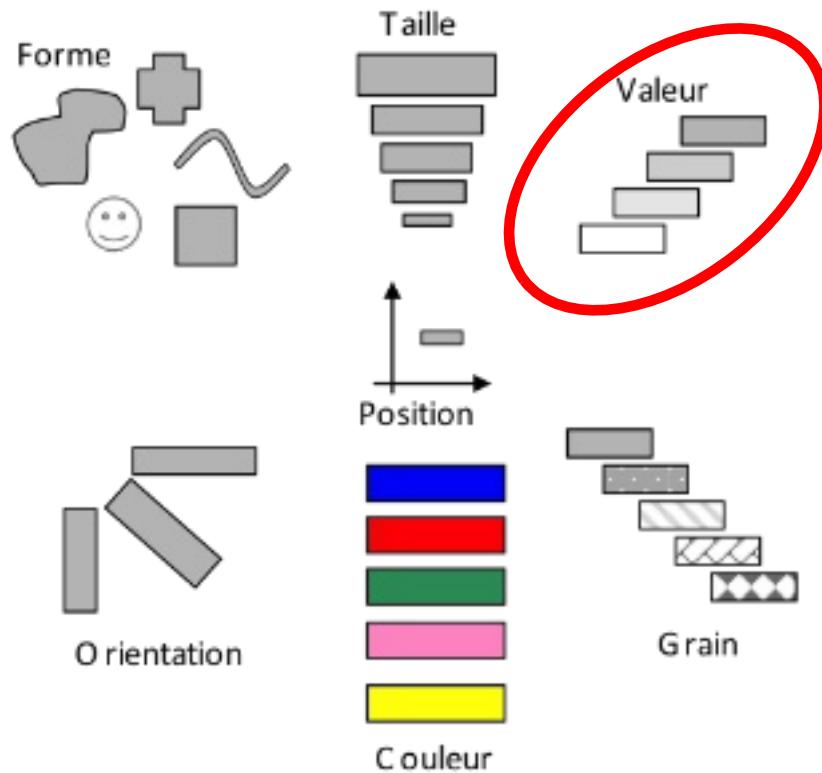
Mode de partition des classes & nombre de classes

Effectuer la graduation



[Classer] puis [OK]

# sémiologie graphique: la variable rétinienne valeur



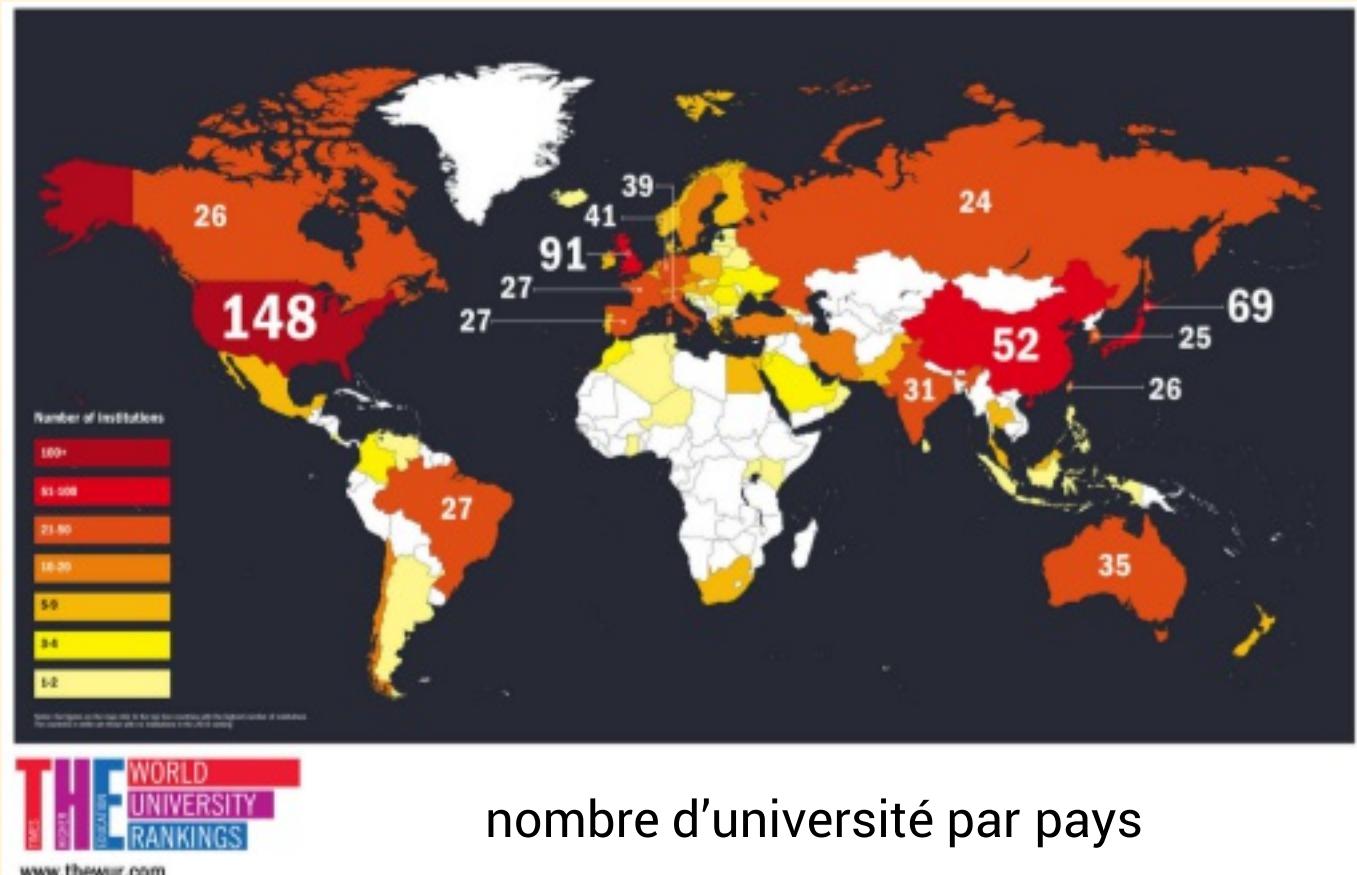
point	line	area		associative	selective	ordered	quantitative
			size				
			value				
			texture				
			colour				
			orientation				
			shape				

# sémiologie graphique : valeur = quantité relative

#mapfail

*Cette carte représente une variable quantitative absolue (le nombre d'universités par pays) avec la variable rétinienne valeur.*

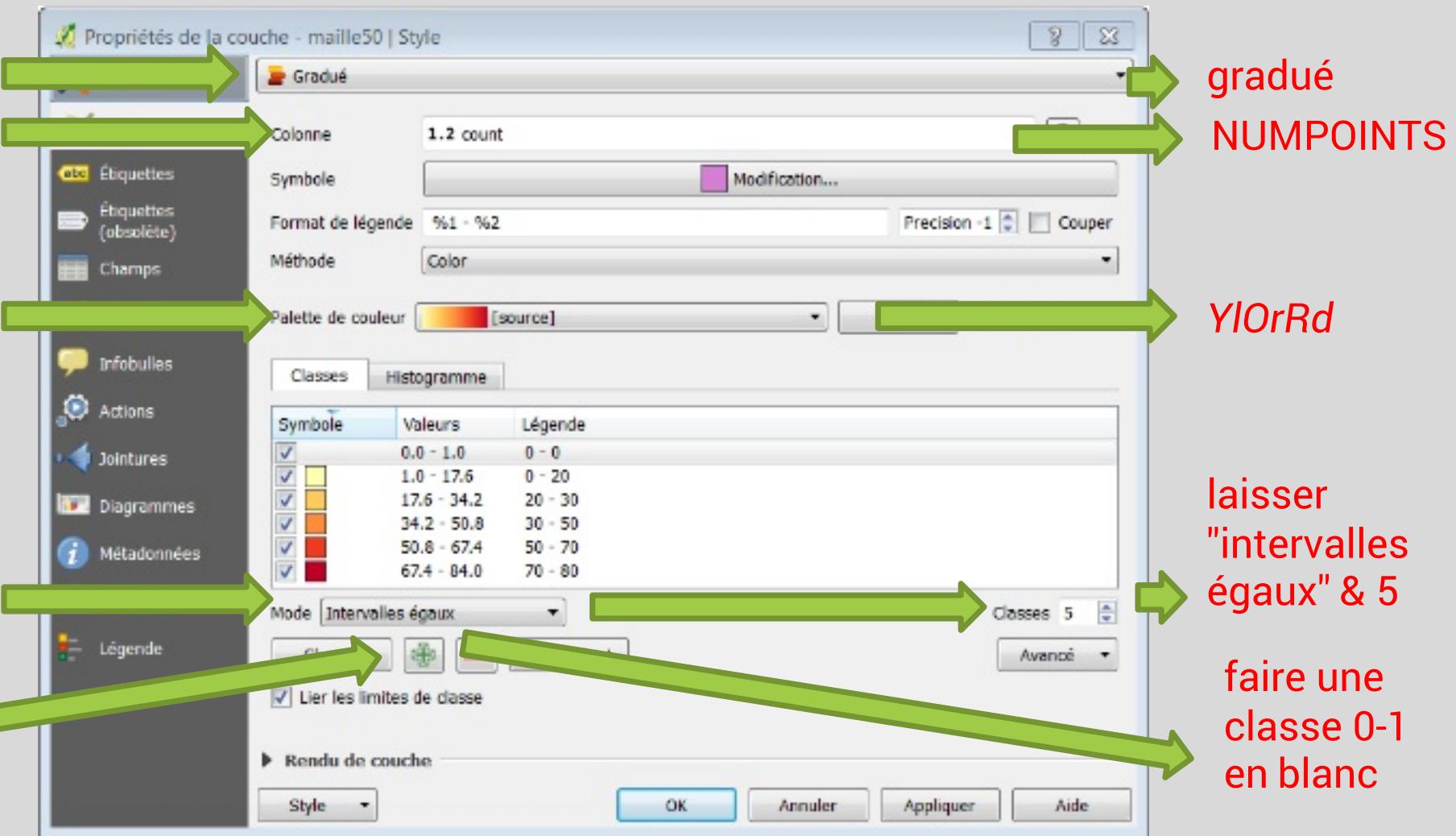
*La France et la Russie ont la même couleur alors que si on rapportait ce nombre par habitants ou à la superficie la hiérarchisation visuelle ne serait pas du tout la même !*



# Projet\_Ozans.qgs: faire une carte "graduée"

Clic droit sur la couche de maille > Propriétés > Onglet Style

Style (simple, catégorisé, gradué, etc.)



Méthode & nombre de classes

Ajouter une classe

laisser  
"intervalles  
égaux" & 5

faire une  
classe 0-1  
en blanc

# Sémiologie graphique: Double gradation positive/négative

---

*le logo 5-C conçu par le professeur Serge Hercberg (Inserm)*

*Ici on n'utilise pas la variable rétinienne valeur mais une double gradation de la variable couleur ordonnée selon le spectre visible de la lumière en partant du jaune (centre du spectre) d'un côté vers l'ultraviolet (les verts) et l'infrarouge (orange et rouge) pour exprimer une opposition.*

*En outre le créateur joue sur la dimension symbolique et culturelle des couleurs: vert = c'est bien, rouge = pas bien*



# Résumer la série statistique: Les limites de la moyenne

---



Albert

@amisdalbert



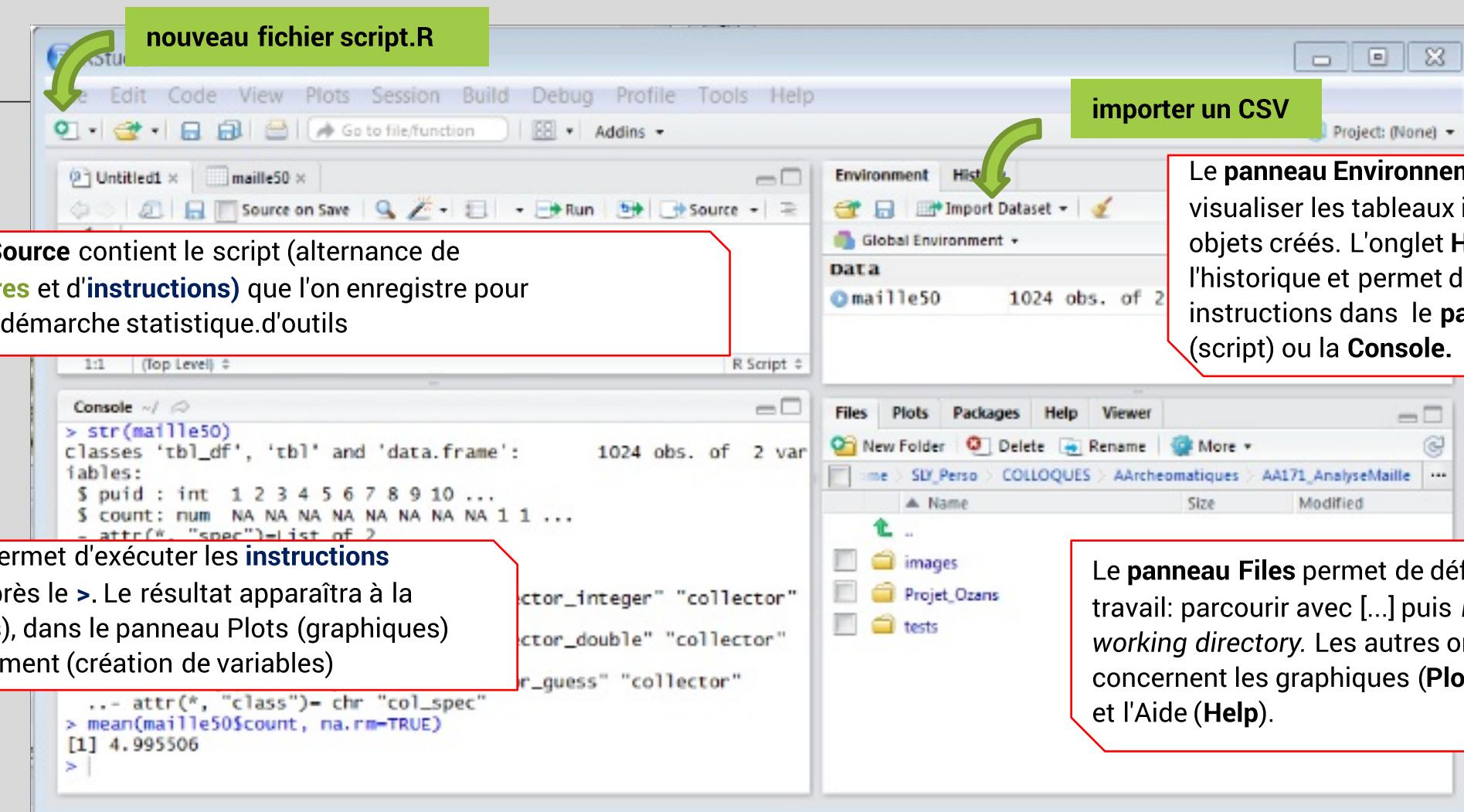
Suivre



Les limites d'une moyenne : si Bill Gates entre dans un bar, en moyenne, tout le monde y est millionnaire.

*teasing pour une vidéo de la chaîne youtube "les statistiques expliquées à mon chat" intitulée pourquoi gagnez vous moins que le salaire moyen. <https://t.co/AJOG2djDf5>*

# L'interface générale de Rstudio



# R & Rstudio: principes de fonctionnement

---

```
# le logiciel permet d'exécuter des commandes à écrire, dans le panneau console, derrière #
le signe > et a valider avec [Entrée]
# le résultat s'affichera soit dans le panneau console (calcul), soit dans une fenêtre plot
# (graphique), soit dans le panneau environnement (table ou variable)
# mettre un commentaire avec un # il ne sera pas exécuté
# une fonction est une commande que va exécuter le logiciel
# la syntaxe est fonction(argument1, argument2), par exemple
mean(maille$count)
# une variable est un objet que l'on pourra appeler ultérieurement dans une fonction
# par exemple
dens <- maille$count
mean(dens)
# pour installer un package (extension) qui amène des nouvelles fonctions on utilise
install.packages("nom_dupackage")
# pour activer un package on utilise
library(nom_du_package)
```



# Projet Ozans: Résumer la série statistique



**QGIS:** exporter la table maille50 au format CSV (clic droit > Enregistrer sous..)



**RStudio:**

**Importer le fichier csv à l'aide du bouton [Import DataSet] > à partir d'un CSV**

Dans le panneau Console, tapez les lignes suivantes une à une suivies de la touche [Entrée]

# afficher la table

maille50

# regarder la structure de la table

str(maille50)

# appeler **dens** la variable correspondant à la variable (champs) count de la table **maille50**

dens <- maille50\$count

mean(dens) # calcule la moyenne

median(dens) # calcule la médiane

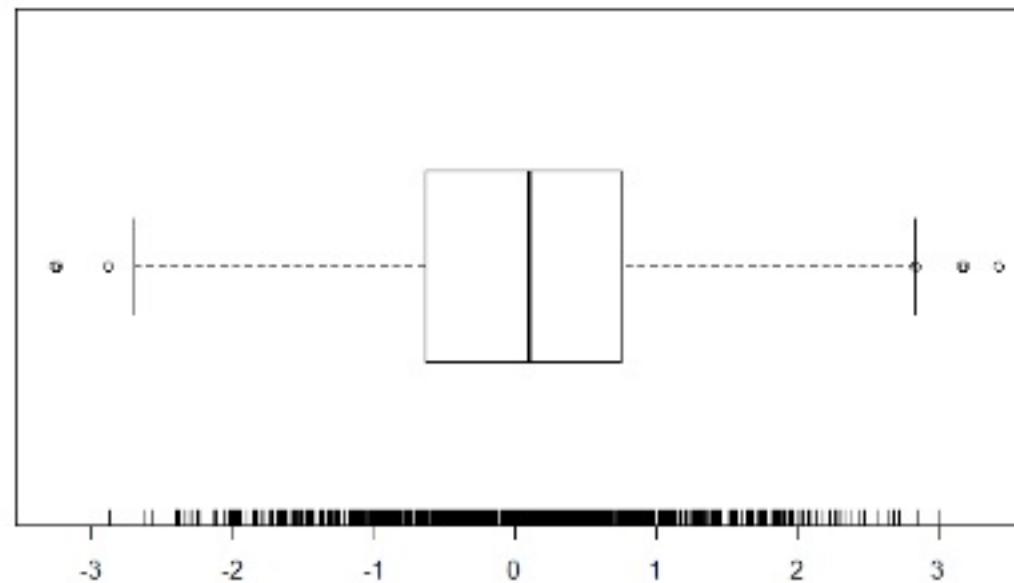
sd(dens) # calcule l'écart-type

max(dens)-min(dens) # calcule l'étendue

summary(dens) # donne un résumé statistique

# Analyser la distribution: la boîte à moustache

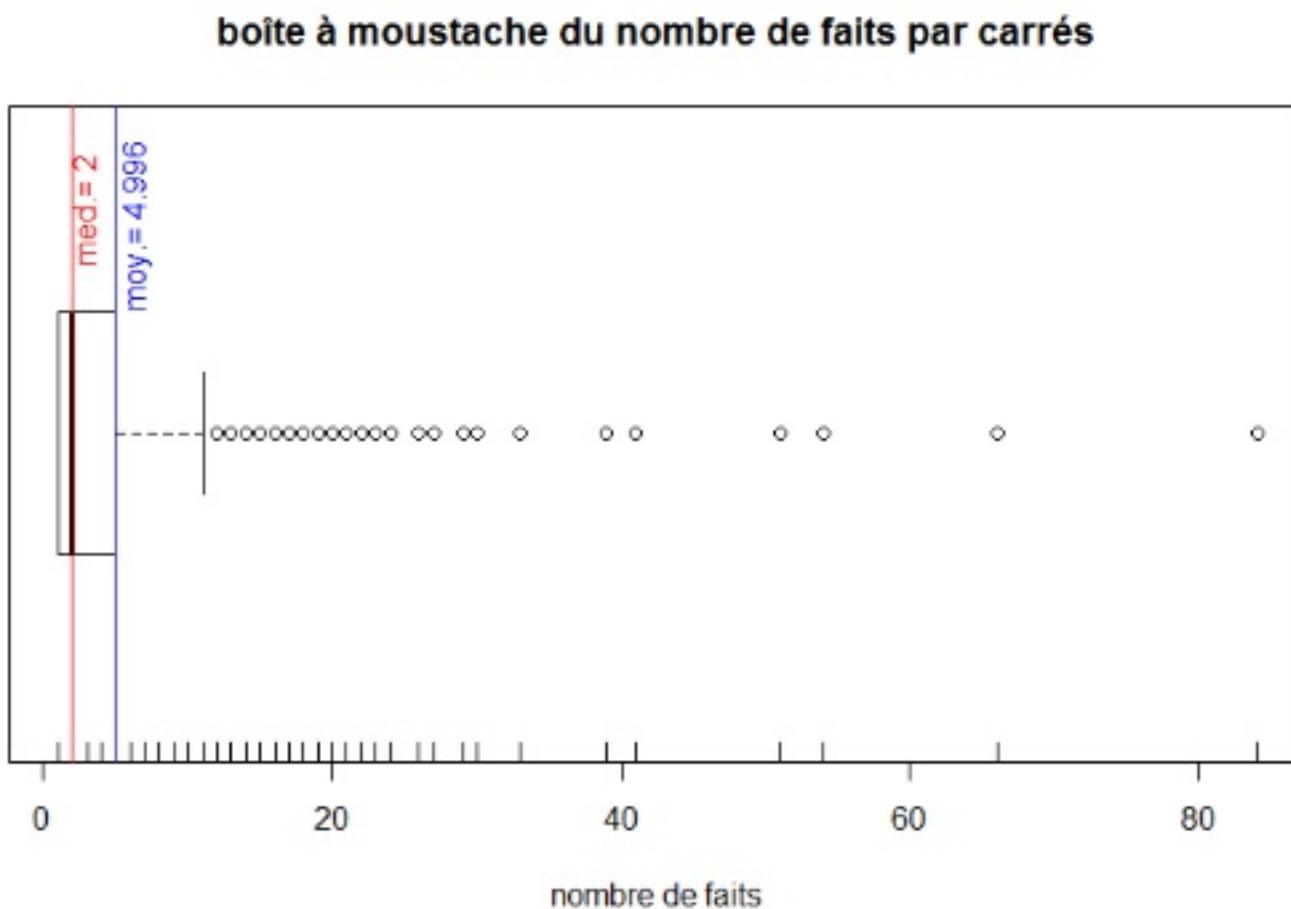
---



distribution d'une série statistique suivant une loi normale,  $rnorm(1000)$

# Analyser la distribution: la boîte à moustache

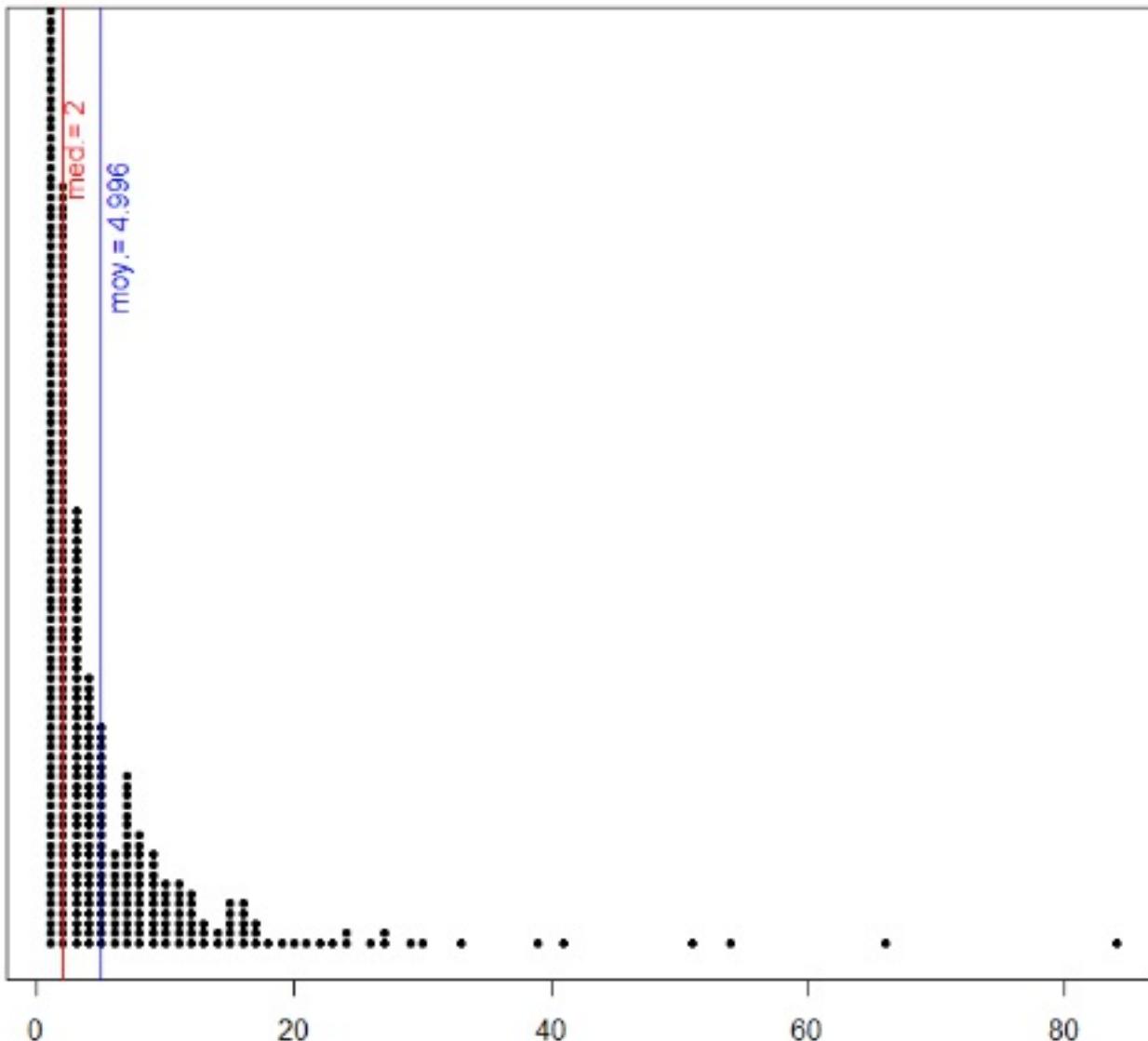
---



R

```
# faire la boxplot  
> boxplot(dens, horizontal = TRUE, main  
= "boîte à moustache", xlab="nombres de  
faits")  
  
# ajouter les individus en abscisse  
> rug(dens)  
  
# tracer la moyenne  
> abline (v=mean(dens), col="red")  
  
# ajouter l'étiquette  
> moy <- mean(dens)  
> text(moy,1.4, paste  
('moy.=',round(moy, digits=3)), pos=1,  
srt=90, col = "red")
```

# Analyser la distribution: le scalogramme



R

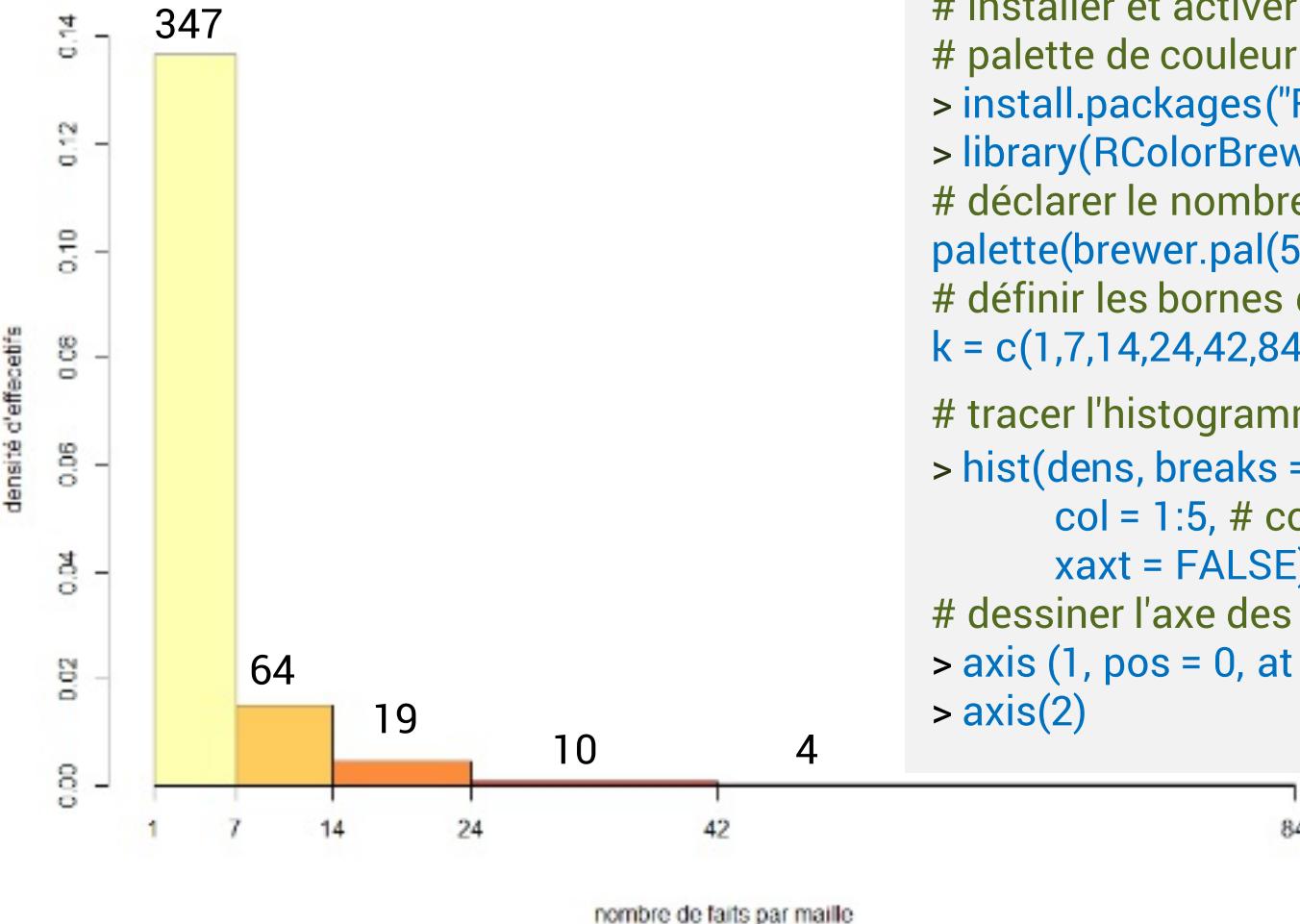
```
# tracer la série statistique avec des  
points  
> stripchart(dens)
```

```
# les "agiter" pour mieux les voir  
> stripchart(dens, method = "jitter")
```

```
# faire un scalogramme, un vrai !  
> stripchart(dens,  
# placer les points à la verticale  
method = "stack",  
# choisir la forme du point  
pch = 20,  
# mettre l'origine à 0  
at = 0)
```



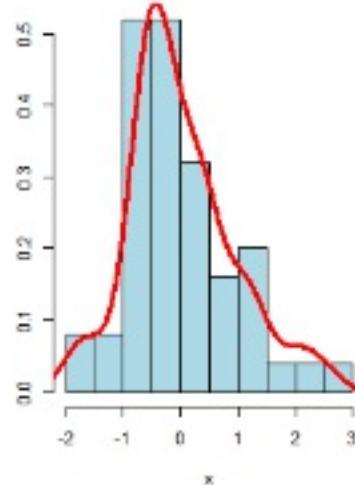
# Analyser la distribution: l'histogramme



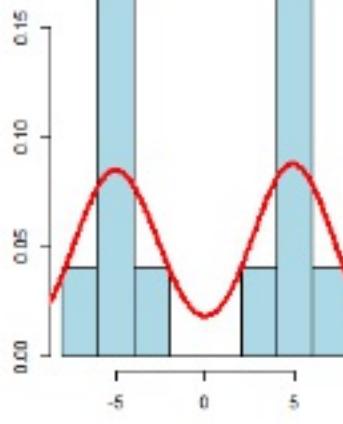
```
# tracer un histogramme  
> hist(dens)  
# installer et activer le package RColor Brewer pour utiliser une  
# palette de couleur "juste"  
> install.packages("RColorBrewer")  
> library(RColorBrewer)  
# déclarer le nombre de couleurs et la palette  
palette(brewer.pal(5,"YlOrRd"))  
# définir les bornes de classes à la main  
k = c(1,7,14,24,42,84)  
# tracer l'histogramme  
> hist(dens, breaks = k, # breaks = bornes de classes  
       col = 1:5, # couleurs à utiliser dans la palette active  
       xaxt = FALSE) # ne pas dessiner l'axe des abscisses  
# dessiner l'axe des abscices avec les bornes de classes  
> axis (1, pos = 0, at = k)  
> axis(2)
```



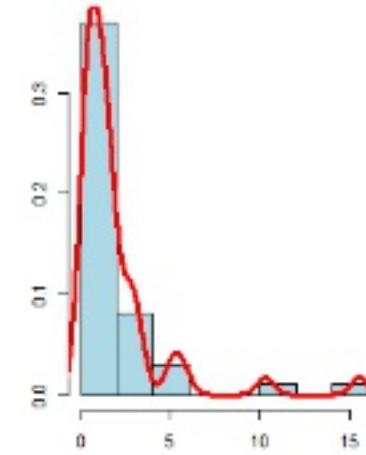
# Analyser la distribution: la forme



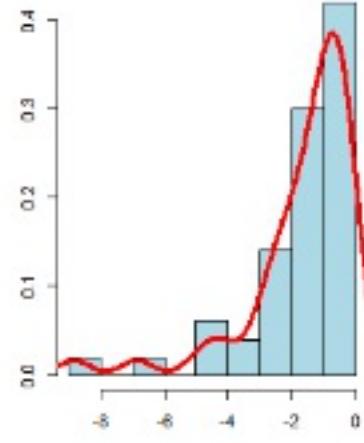
unimodale



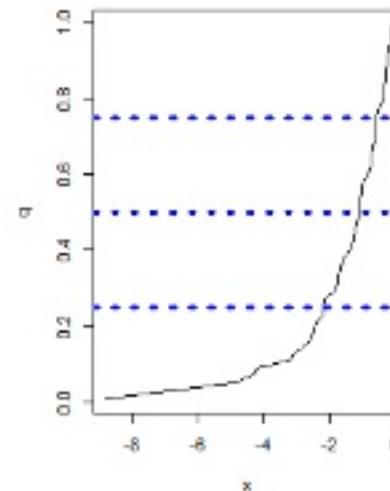
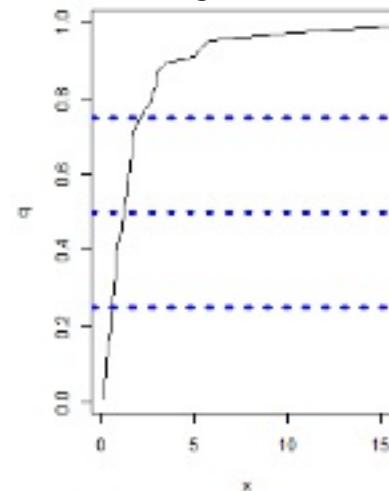
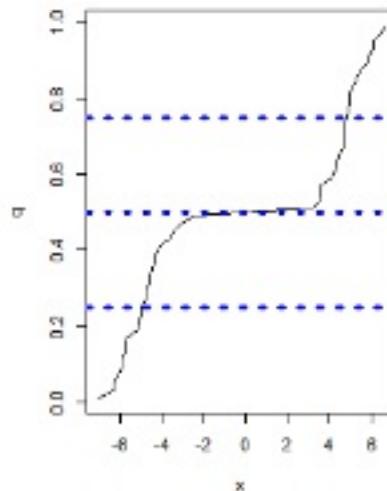
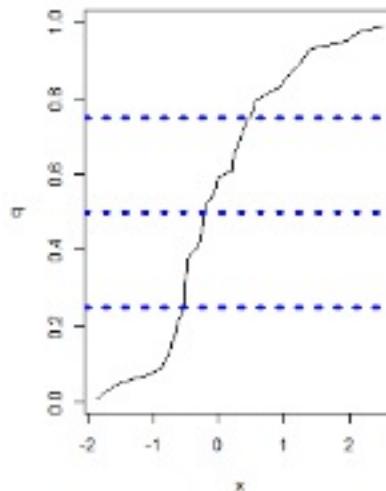
bimodale



dissymétrique  
à gauche



dissymétrique  
à droite



# Discrétiser: Les règles

---

- Les classes doivent couvrir l'ensemble de la distribution et elles doivent être contiguës (jointives).
- Une valeur ne doit appartenir qu'à une classe et une seule.
- Les classes ne doivent pas être vides.
- Les valeurs limites doivent être précises et rapidement appréhendables.
- Éviter de placer dans deux classes distinctes des valeurs non significativement différentes.
- Ne pas définir des seuils avec un nombre de décimales supérieur à celui de la précision des données.



# Projet Ozans: Discréter avec et/ou

---

## QGIS:

- Clic droit sur la couche maille 50 > propriétés > Style > Gradué
- Choisir la méthode de discréttisation (mode) puis bouton [Classer]
- Passer à l'onglet histogramme pour visualiser un **diagramme de distribution**
- Bouton **[Charger les valeurs]**, barres d'histogramme = 100

Note: Qgis fait des classes de type (borne\_inf - borne\_sup)

## Rstudio:

- Ouvrir le fichier Projet\_Ozans/AA\_maille\_scriptR.R (File > Open File)
- Exécuter les lignes de commandes en les sélectionnant dans le **panneau Source** puis **[Ctrl]+[Entrée]**

Note: classIntervals fait (par défaut) des classes de type [borne\_inf - borne\_sup)

Pour chaque méthode de discréttisation:

Observez les bornes de classes sur le diagramme de distribution de QGIS sur votre poste et regardez le résultat cartographique vidéoprojeté!



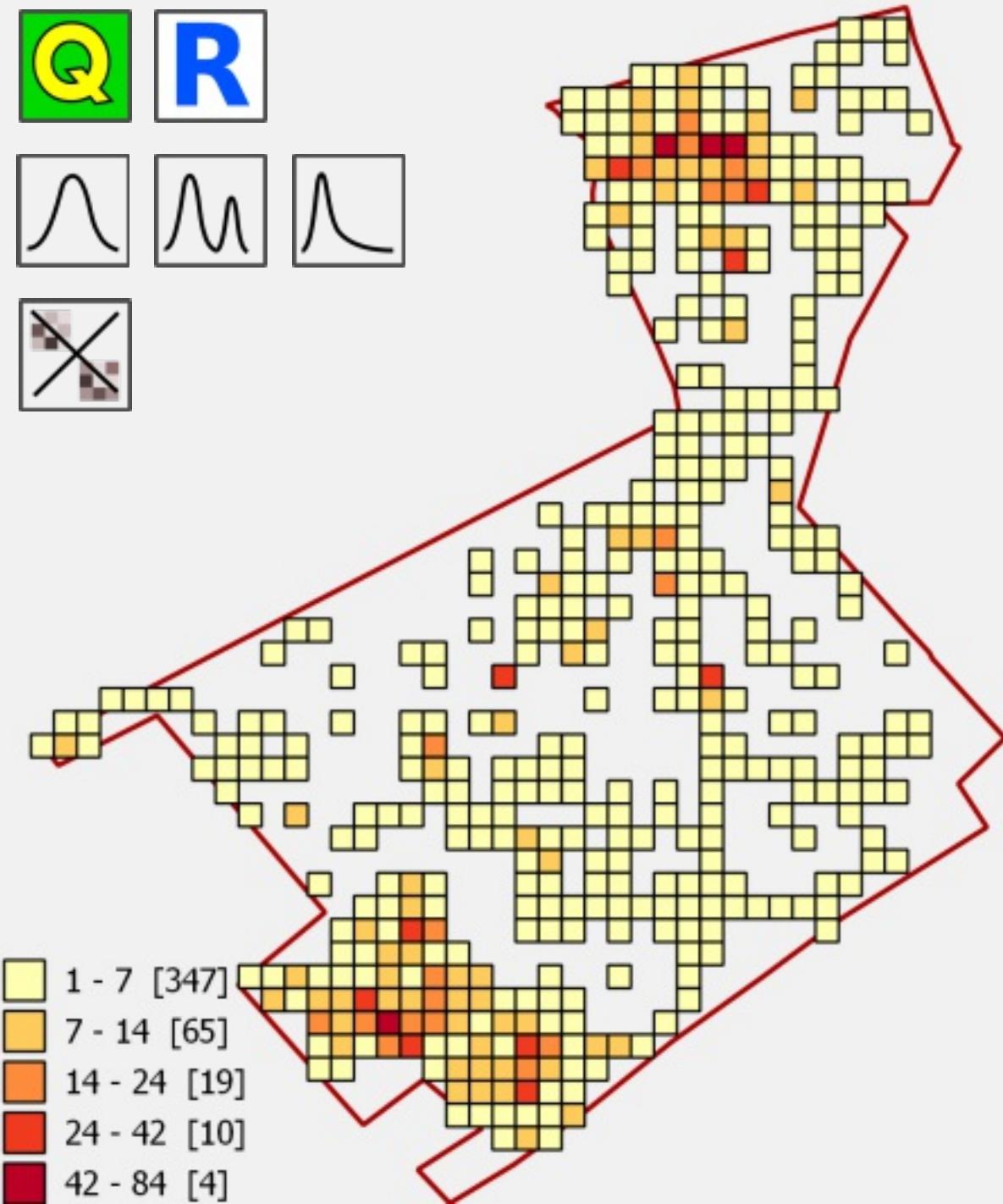
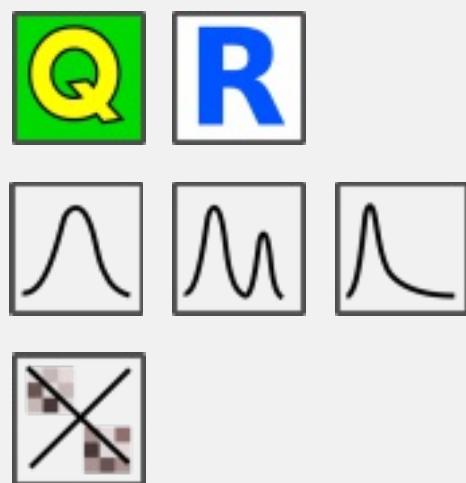
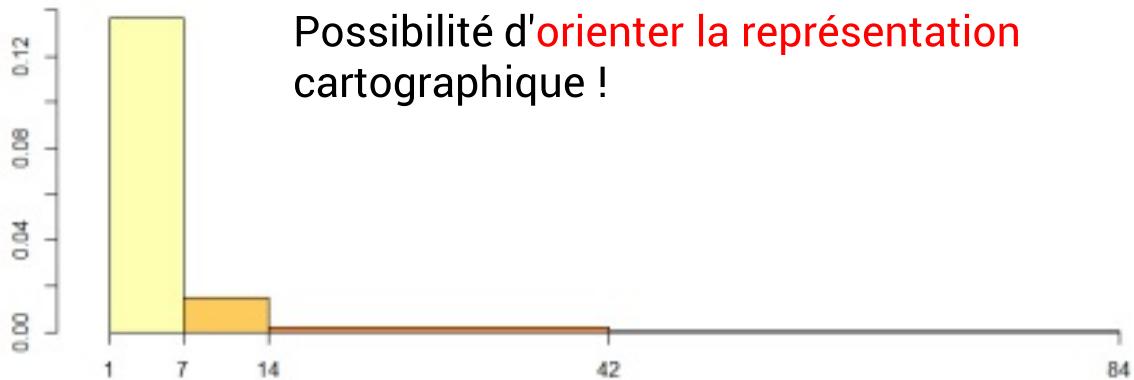
# Discrétiser: Les méthodes Seuils observés

On choisit comme bornes des classes les trous de la distribution (histogramme)

⊕ Simple à mettre en œuvre (si des regroupements apparaissent clairement dans la distribution)

mise en évidence des valeurs extrêmes ou des regroupements spécifiques

⊖ Risque de produire des classes surchargées .



# Discrétiser: Les méthodes Amplitudes égales

**amplitude des classes = effectif/nombre de classes**



Facile à utiliser et rapidement comprise par le lecteur

Valorise les valeurs extrêmes

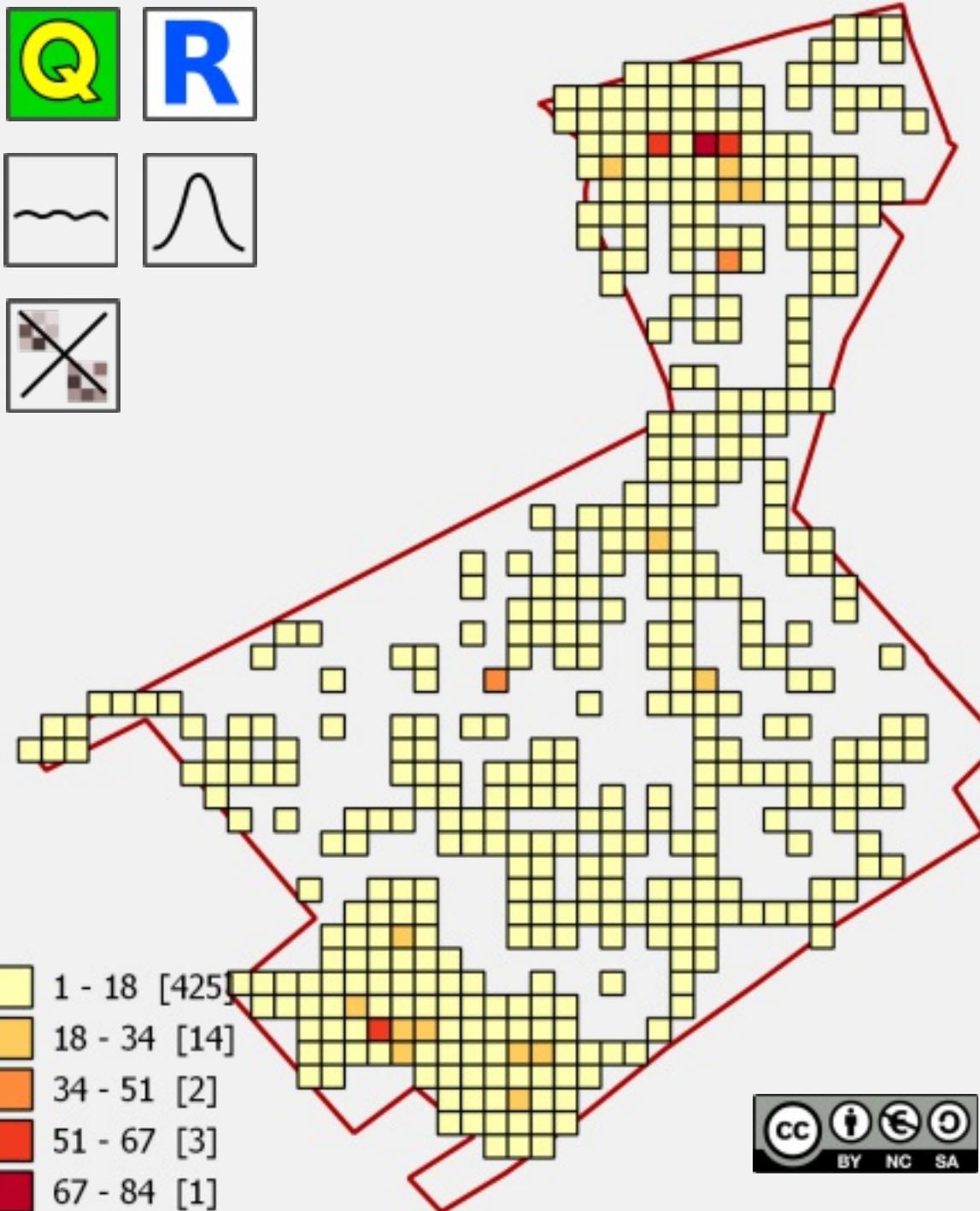
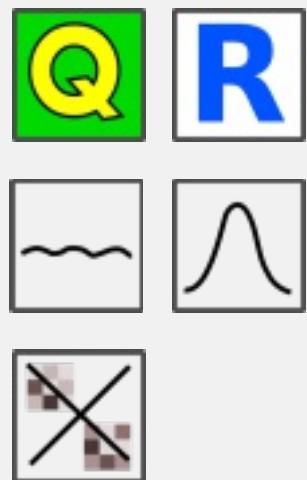
Efficace pour les distributions symétriques et normales



Inadaptée aux distributions dissymétriques

Les discontinuités de la distribution peuvent faire apparaître des classes vides

Des classes sont souvent surchargées (image cartographique ternes)



# Discrétiser: Les méthodes Quantiles (Effectifs égaux)

nombre de mailles par classes / nombre de classes



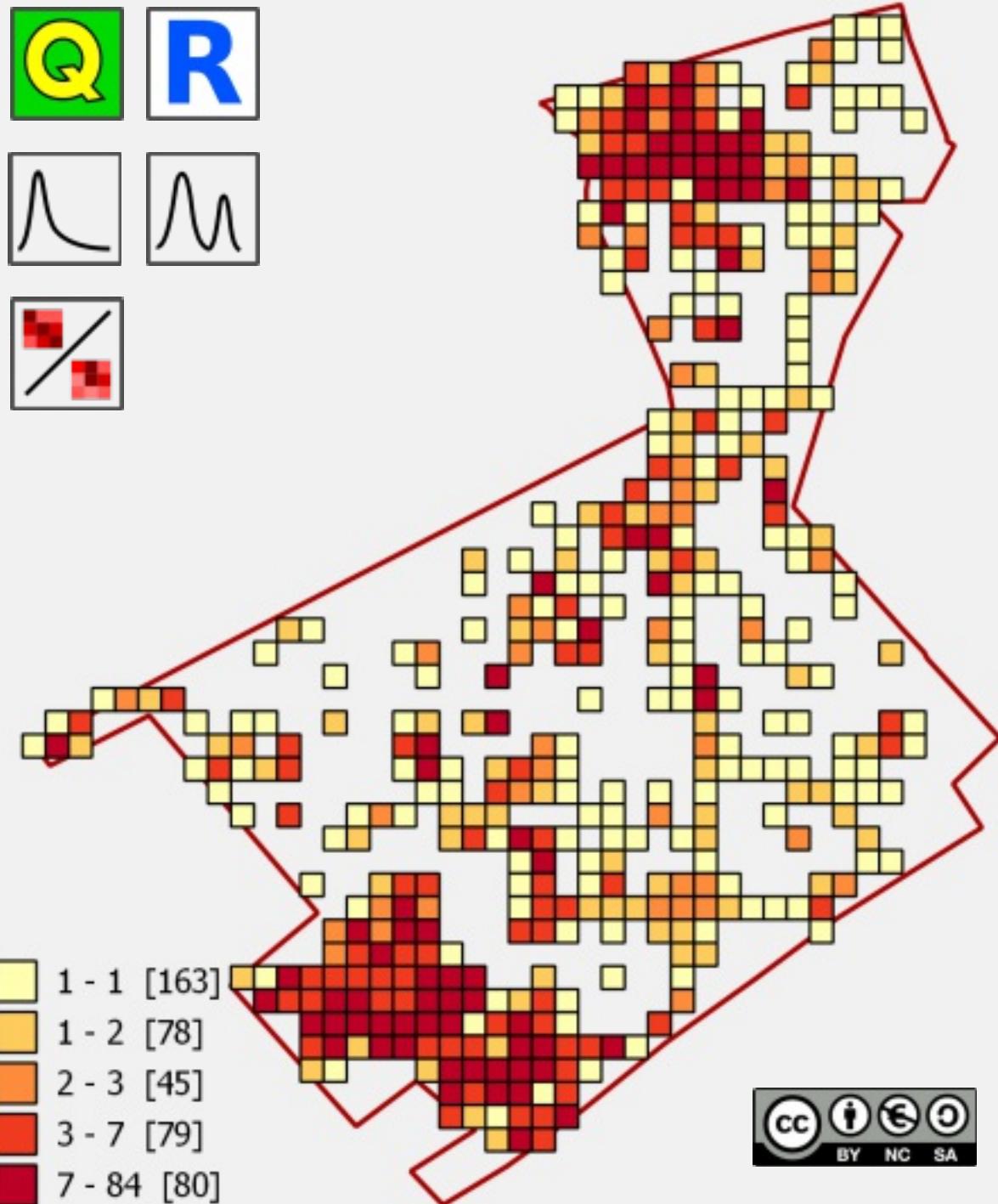
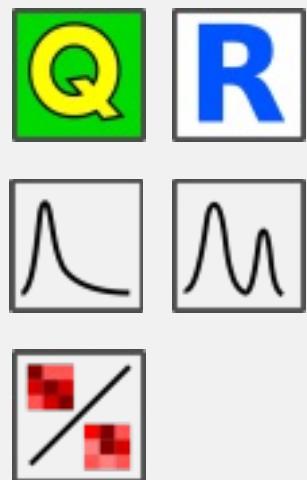
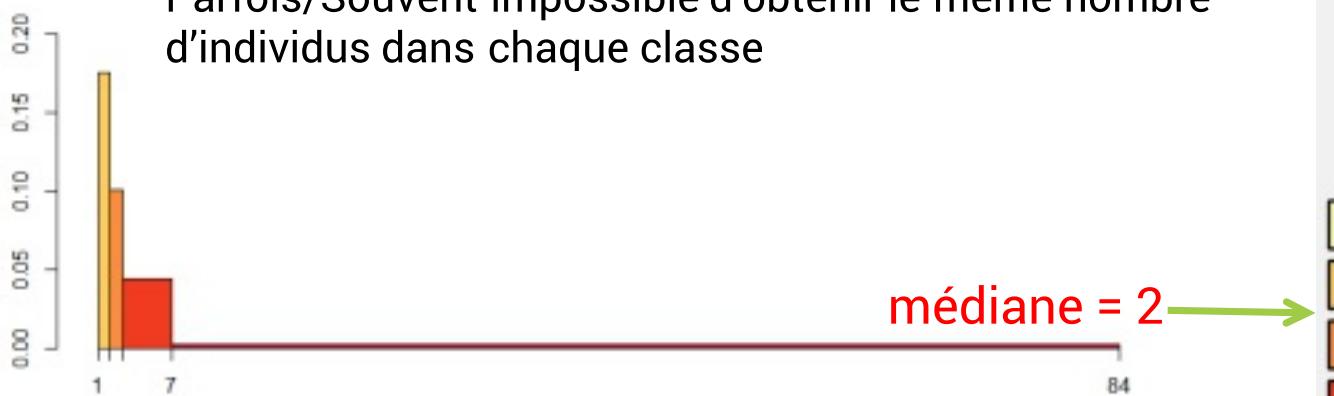
Efficace pour toutes les distributions

Utile pour comparer plusieurs séries statistiques (mais comparaison d'ordre de grandeurs et non de valeurs)



Ne tient pas compte de la distribution et des valeurs exceptionnelles

Parfois/Souvent impossible d'obtenir le même nombre d'individus dans chaque classe



# Discrétiser: Les méthodes Standardisée (écart-types)

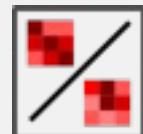
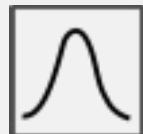
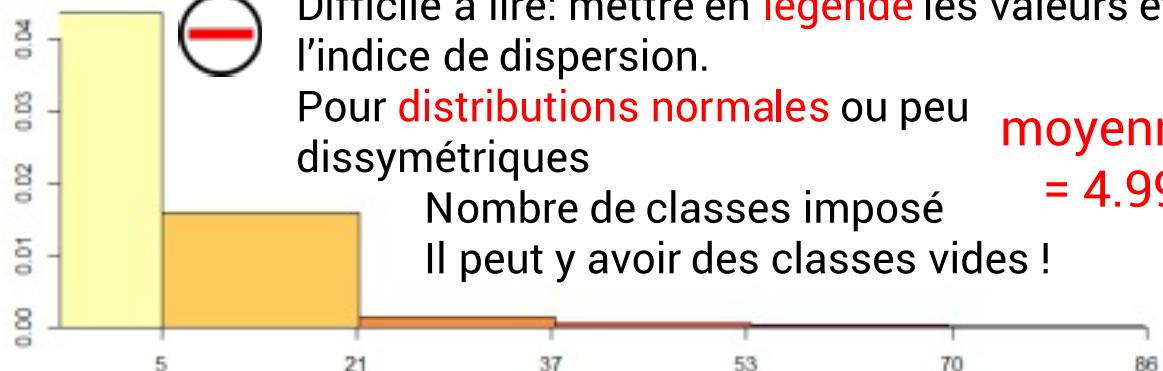
Origine = moyenne & amplitude de classe = écart-type(s)



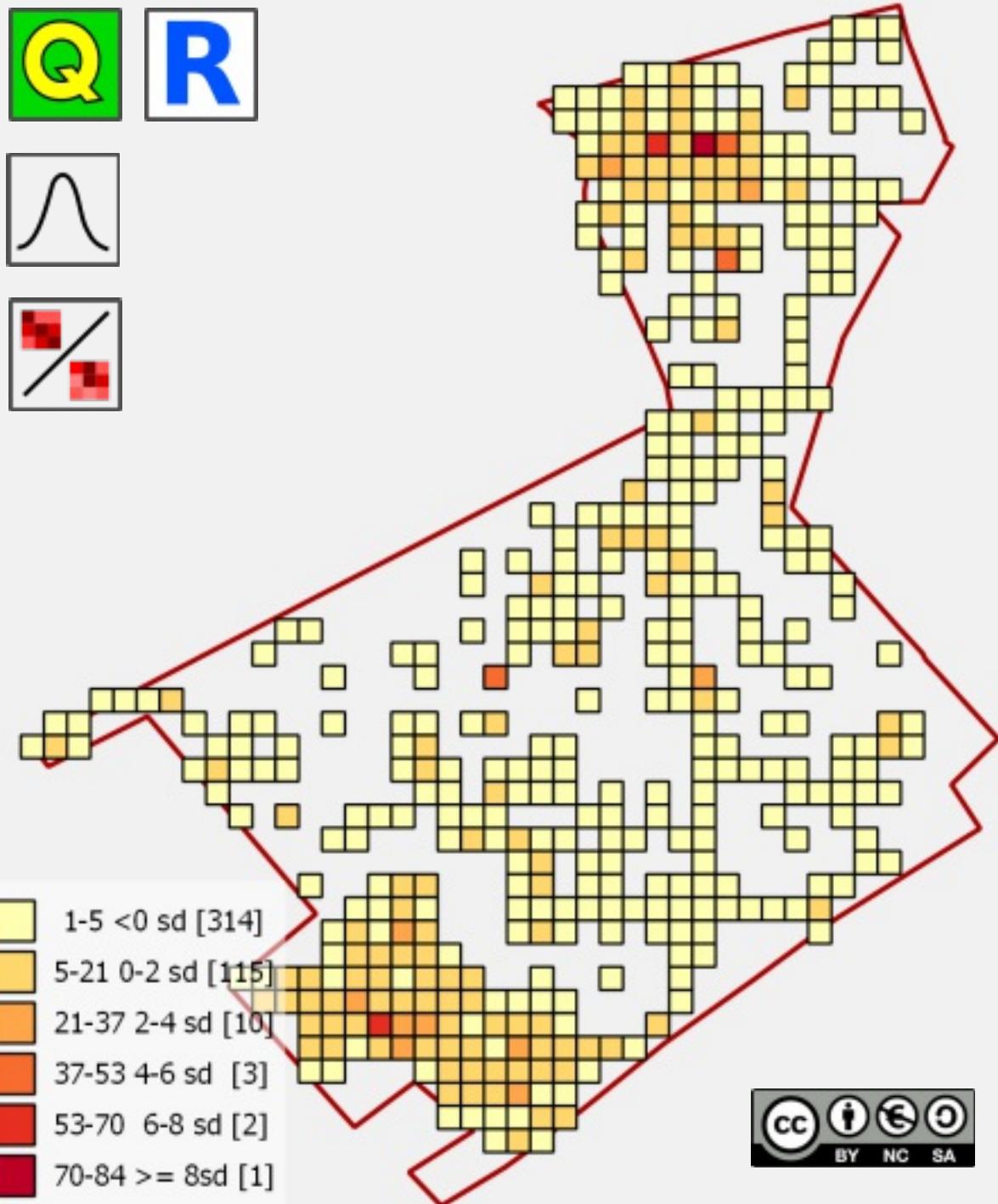
Permet des **comparaisons** indépendamment des problèmes liés à la taille de la variable

Toutes les classes ont une même étendue égale à l'écart type, sauf éventuellement les extrêmes

Très utile pour mettre en avant des **concentrations** et/ou à l'inverse des « **vides** »



■	1-5 <0 sd [314]
■	5-21 0-2 sd [115]
■	21-37 2-4 sd [10]
■	37-53 4-6 sd [3]
■	53-70 6-8 sd [2]
■	70-84 >= 8sd [1]



# Discrétiser: Les méthodes progression géométrique

**l'amplitude des classes augmente par progression géométrique**

On calcule la raison,  $r = 10^{[(\text{LOG}(\text{MAX}) - \text{LOG}(\text{MIN})) / k]}$   
puis chaque borne = borne précédente multipliée par r

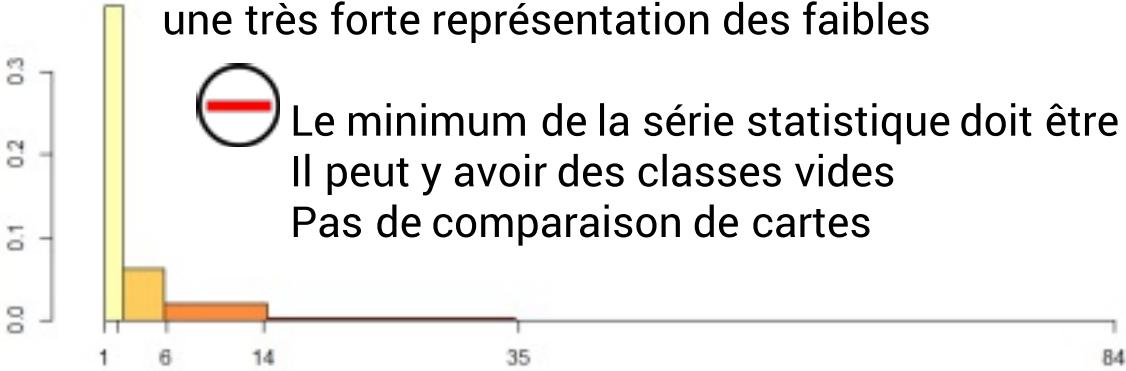


Les **faibles valeurs sont plus finement représentées**, et les fortes valeurs sont regroupées dans la dernière classe.

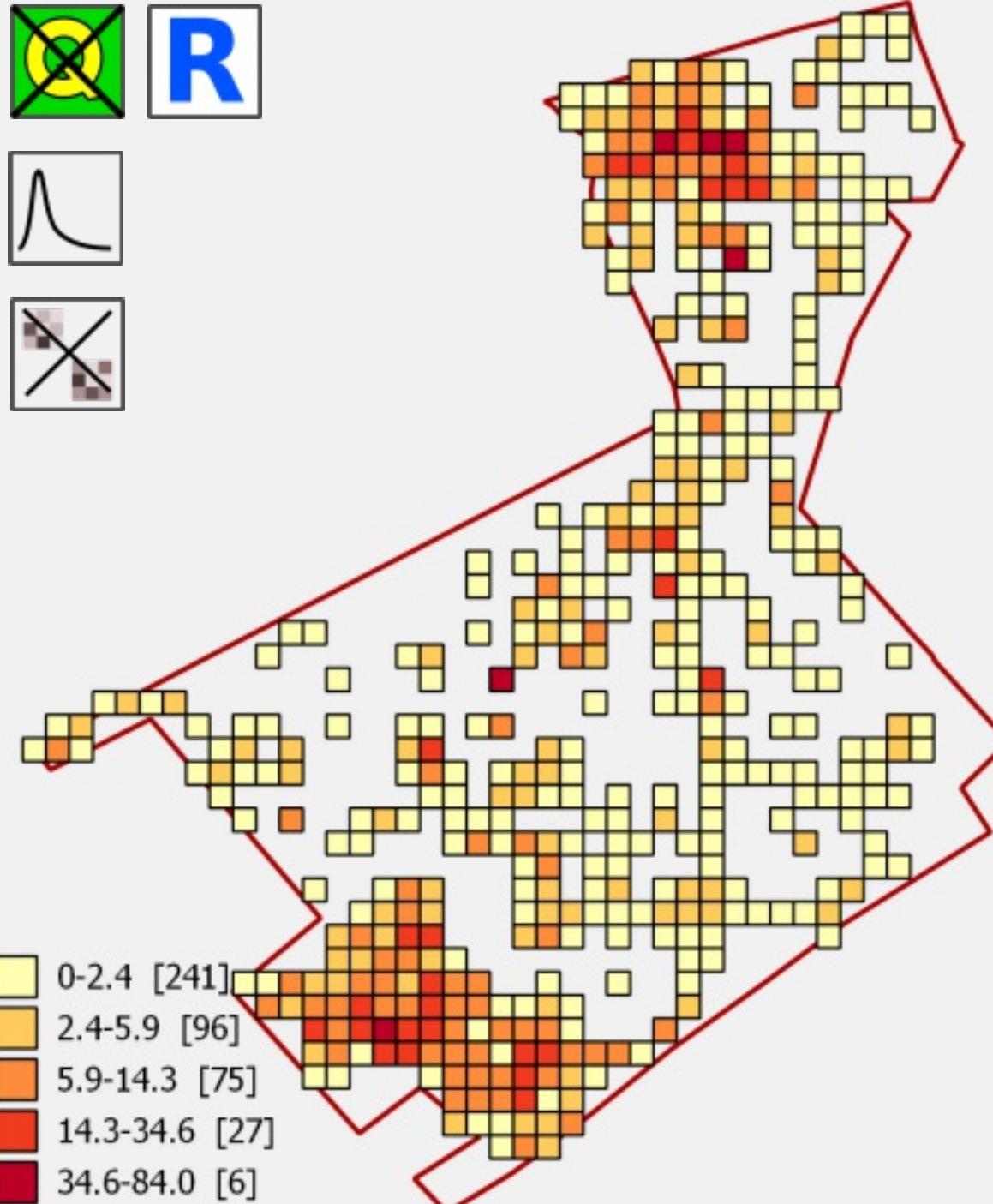
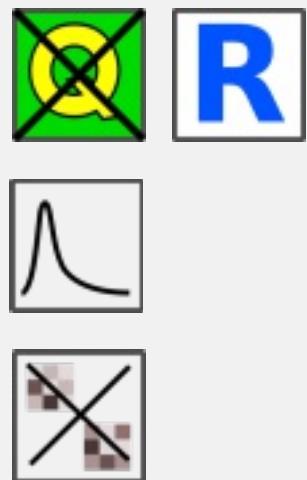
Utile pour les **distributions asymétriques** qui comportent une très forte représentation des faibles



Le minimum de la série statistique doit être > 0  
Il peut y avoir des classes vides  
Pas de comparaison de cartes



84



# Discrétiser: Les méthodes Ruptures naturelles (Jenks)

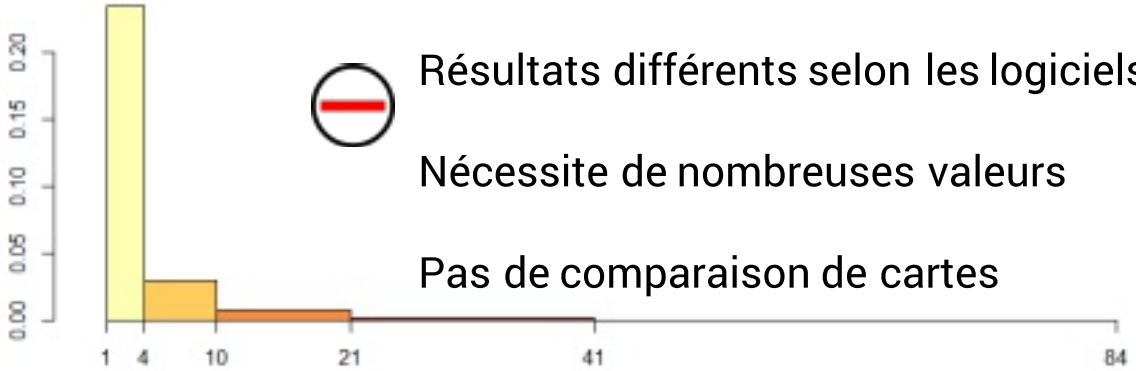
algorithme qui minimise la variance dans les classes et la maximise entre les classes



Propose un découpage où les individus d'une même classe sont les plus proches possibles et isole les différents groupes

Respecte bien la **structure des données**

Résultat **cartographique harmonieux** (mais on est pas là pour faire de l'archéojolie!)



Résultats différents selon les logiciels

Nécessite de nombreuses valeurs

Pas de comparaison de cartes

Q R

Wavy bell-like double bell-like

