



# Cartographie thématique

Mise à niveau L3 Géographie et aménagement

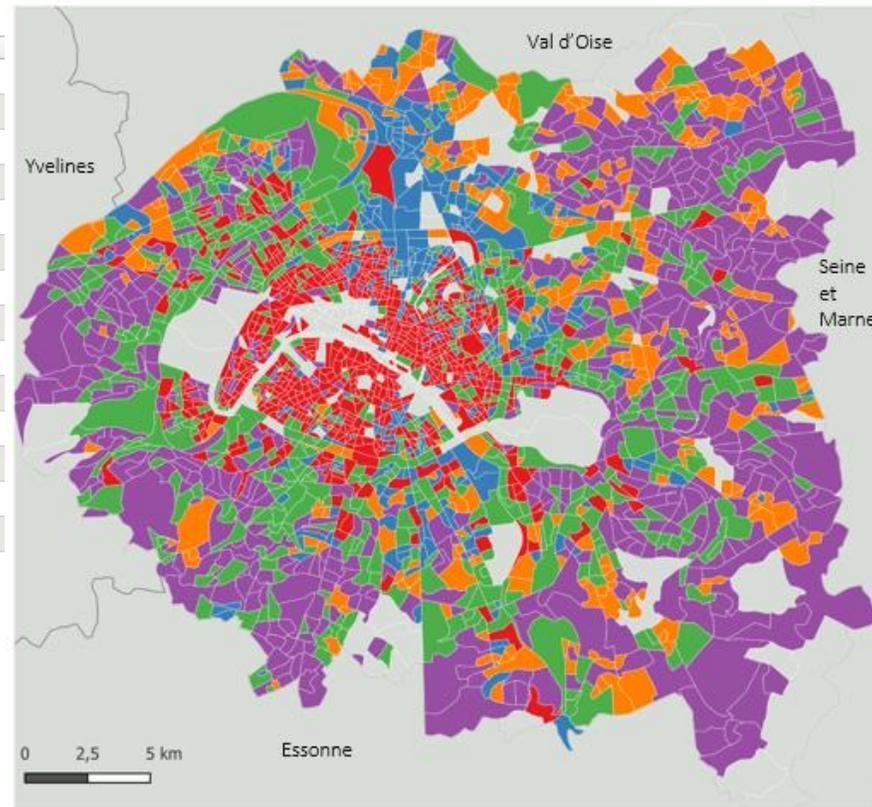
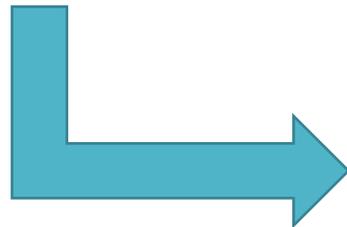
Christina Aschan-Leygonie ([christina.aschan@univ-lyon2.fr](mailto:christina.aschan@univ-lyon2.fr))

# Introduction

La cartographie **thématique** (= statistique) est le passage d'une information sous forme de données géographiques statistiques à une image graphique

La cartographie utilise les propriétés de l'image visuelle pour faire apparaître les relations d'ordre et de ressemblance entre les données

CODE_IRIS	NOM_IRIS	TYP_IRIS	MPSEUL	MSFAM	MCOSENF	MCOAENF	MFAMONO
751114113	Folie Méricourt 13	H	468,0573407607...	32,62903014675...	189,9764203666...	75,91660560804...	37,40490591342...
920440109	Wilson-Rivay 1	H	780,7223134913...	39,61502494494...	250,2870777707...	285,1588420233...	117,5745582968...
930270105	Anatole France	H	523,3170118390...	104,7575439416...	288,1755065631...	566,7465066472...	291,1296728440...
751197316	Villette 16	H	1010,900638288...	106,4436662812...	284,4285434956...	352,3505514135...	155,2873737968...
751176716	Batignolles 16	H	553,4129378033...	25,80855267243...	176,0117192139...	170,2522701605...	72,89553498592...
940210107	Sorbiers	H	326,2436444835...	32,85117918717...	107,0690277557...	273,5715089688...	272,2122385873...
920090106	Hugo-Bel-Air	H	439,1428709425...	15,00871739145...	273,0296738369...	319,3794288850...	53,90113931360...
920630104	Plaine-Gare 4	H	433,6588113382...	17,16356727559...	242,3312122047...	267,8094971821...	51,45399212036...
751103703	Saint-Vincent de Paul 3	H	588,4116546173...	75,56141689550...	297,7607763028...	180,1120565173...	88,84477395348...
930050802	Ambourget 2	H	158,0153063749...	40,41176501828...	63,06036649218...	393,5374130607...	175,0065886612...
930500112	Iris 12	H	329,0687050126...	18,85139752123...	222,4449696819...	445,6920916184...	163,4770404303...
920440121	Eiffel 4	H	451,5933510137...	32,22944370613...	146,0313467785...	265,8077845671...	260,4304991671...
940680204	Le Vieux Saint-Maur 4	H	377,3901246237...	27,83033556852...	231,4299348693...	219,9071058340...	98,97060420664...
920400505	Colombiers 2	H	648,7896950613...	32,27629084510...	442,3746198952...	709,4564944154...	272,1046479855...



Types d'IRIS de Paris et la première couronne en fonction de la composition familiale en 2017

(typologie issu d'une classification multivariée à partir de 5 variables)

## Caractéristiques des IRIS

- Très forte sur-représentation de ménages avec des personnes vivant seuls et une très forte sous-représentation de familles avec enfants
- Sur-représentation de ménages sans familles et des familles monoparentales et sous-représentation de couples (avec ou sans enfants)
- Caractéristiques de ménages proche de la moyenne de Paris et de la première couronne
- Couples (avec ou sans enfants) très sur-représentés  
Personnes qui ne vivent pas en couple sous-représentées
- Très forte sur-représentation de ménages avec enfants et une très forte sous-représentation de ménages sans enfants
- Iris non résidentiels
- Départements voisins

Sources : INSEE, IGN  
Réalisation : C.Aschan, 2022

# I. Les données statistiques

# I. Les données statistiques à cartographier

Un peu de **vocabulaire statistique** et le vocabulaire SIG correspondante

**Tableau élémentaire** = table attributaire

**Individu** (ou unité géographique) = entité

**Variable statistique** (ou caractère) = champ (de la table attributaire)

**Modalité** = attribut (information alphanumérique (texte ou chiffres) dans la table attributaire)

# Vocabulaire statistique

## Exemple

Les 10 premières aires urbaines dans le Monde d'après l'ONU en 2005

**INDIVIDUS**

**VARIABLES**

Aires Urbaines	Population	Superficie en km <sup>2</sup>	Pays	rang
Tokyo	37 730 064	7 019	Japon	1
Mexico	23 239 783	2 072	Mexique	2
Séoul	22 692 652	8 680	Etats-Unis	3
New-York	22 214 083	1 150	Corée du Sud	4
Bombay (Mumbai)	21 900 967	1 968	Brésil	5
Karachi	21 142 625	2 590	Indonésie	6
Sao Paulo	20 853 705	2 564	Japon	7
Manille	19 888 419	1 295	Inde	8
Bangkok	18 916 890	4 360	Inde	9
Los Angeles	16 895 945	4 320	Etats-Unis	10

**POPULATION**

**MODALITES**

# Identifiez ces items sur une carte

Individus ?

Les IRIS de Paris et première couronne : (Îlots Regroupés pour l'Information Statistique) sont les plus petites unités spatiales statistiques, inférieures à l'échelon communal

Variable statistique ?

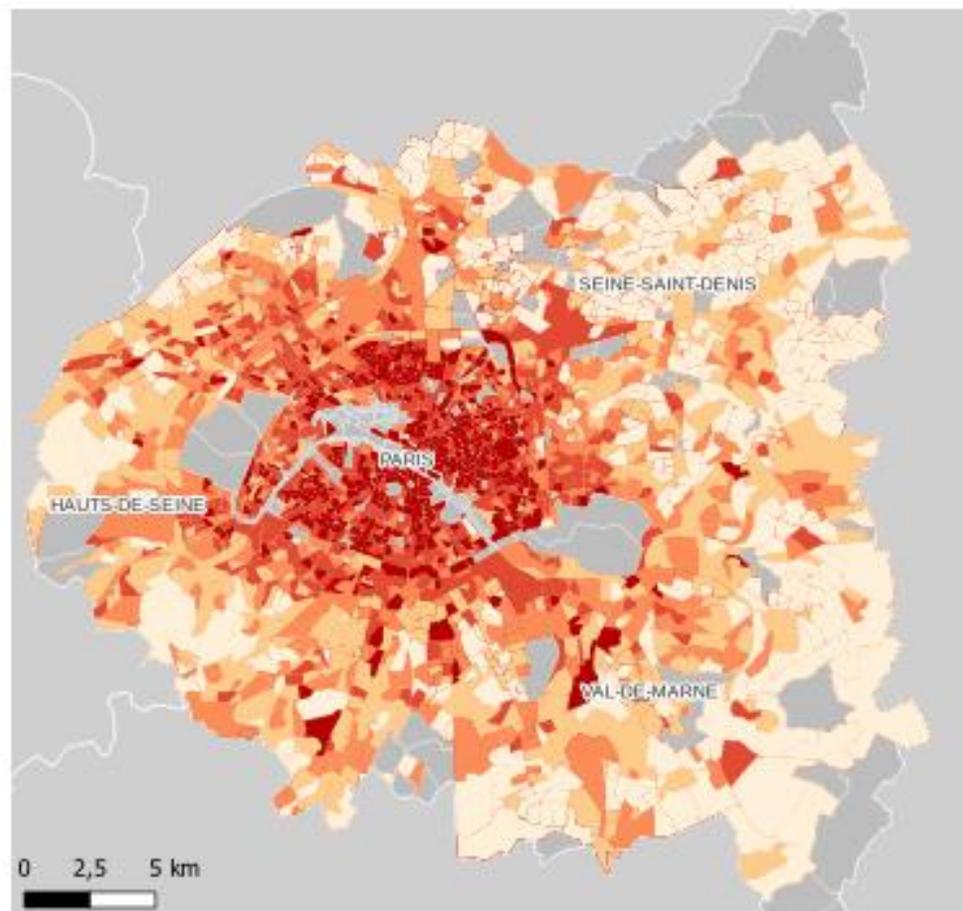
Part des ménages à une personne

Modalité ?

Valeur en %

Tableau élémentaire

CODE_IRIS	NOM_IRIS	TYP_IRIS	MPSEUL	MSFAM	MCOSENF	MCOAENF	MFAMONO
751114113	Folie Méricourt 13	H	468,0573407607...	32,62903014675...	189,9764203666...	75,91660560804...	37,40490591342...
920440109	Wilson-Rivay 1	H	780,7223134913...	39,61502494494...	250,2870777707...	285,1588420233...	117,5745582968...
930270105	Anatole France	H	523,3170118390...	104,7575439416...	288,1755065631...	566,7465066472...	291,1296728440...
751197316	Villette 16	H	1010,900638288...	106,4436662812...	284,4285434956...	352,3505514135...	155,2873737968...
751176716	Batignolles 16	H	553,4129378033...	25,80855267243...	176,0117192139...	170,2522701605...	72,89553498592...
940210107	Sorbiers	H	326,2436444835...	32,85117918717...	107,0690277557...	273,5715089688...	272,2122385873...
920090106	Hugo-Bel-Air	H	439,1428709425...	15,00871739145...	273,0296738369...	319,3794288850...	53,90113931360...
920630104	Plaine-Gare 4	H	433,6588113382...	17,16356727559...	242,3312122047...	267,8094971821...	51,45399212036...
751103703	Saint-Vincent de Paul 3	H	588,4116546173...	75,56141689550...	297,7607763028...	180,1120565173...	88,84477395348...
930050802	Ambourget 2	H	158,0153063749...	40,41176501828...	63,06036649218...	393,5374130607...	175,0065886612...
930500112	Iris 12	H	329,0687050126...	18,85139752123...	222,4449696819...	445,6920916184...	163,4770404303...
920440121	Eiffel 4	H	451,5933510137...	32,22944370613...	146,0313467785...	265,8077845671...	260,4304991671...
940680204	Le Vieux Saint-Maur 4	H	377,3901246237...	27,83033556852...	231,4299348693...	219,9071058340...	98,97060420664...
920400505	Colombiers 2	H	648,7896950613...	32,27629084510...	442,3746198952...	709,4564944154...	272,1046479855...



Part des ménages avec personnes vivant seules dans les IRIS de Paris et la première couronne en 2021

Part des ménages à une personne (en %)

- 8,6 - 30,4
- 30,4 - 37,2
- 37,2 - 44
- 44 - 51,7
- 51,7 - 73,9
- IRIS sans habitants

Sources : INSEE, IGN

## II. Les types de variables

## II. Les types de variables

**Fondamental** : connaître le type de variable avant de faire une carte

### Selon le type de variable statistique

- Les calculs statistiques vont être différents
- Le choix de la représentation cartographique va être différent

### Deux grands familles de variables

- Qualitatives (QL)
- Quantitatives (QT)

# Les types de variables

## Qualitatives

Modalités (« valeurs »):  
caractéristiques, catégories,  
codes

### Nominale

Modalités : ne sont pas quantifiables et ne peuvent pas être ordonnées

*Ex. Type de logement (maison individuelle, appartement...); Type d'industrie (chimie, textile...)*

### Ordinale

Modalités : ne sont pas quantifiables, mais peuvent être ordonnées

*Ex. Type de route (autoroute, nationale...); Fréquence d'utilisation du vélo (jamais, une ou deux fois par semaine, ..., tous les jours)*

## Quantitatives

Modalités : quantifiables et correspondent à des nombres (sauf codes)

### Stock

Modalités : effectifs, quantités ; la somme a un sens ; zéro signifie absence de phénomène  
*Ex. Nombre d'habitants, PNB*

### Intensité

#### Rapport

Modalités : taux (%), indices... ; la somme n'a pas de sens

*Ex. PNB/habitant, densité*

#### Intervalle

Modalités : valeurs sur une échelle ; zéro ne signifie pas absence de phénomène

*Ex : Température, altitude*

## Exemples:

## Types de variables

Aires Urbaines	Quantitatif de stock		Pays	Quantitatif d'intensité	
	Population	Superficie en km <sup>2</sup>		Rang pour la population	Densité (hab/km <sup>2</sup> )
Tokyo	33 413 000	7 019	Japon	1	4 760
Mexico	21 701 925	2 072	Mexique	2	10 474
New-York	21 199 865	8 680	Etats-Unis	3	2 442
Séoul	20 156 800	1 150	Corée du Sud	4	17 528
Sao Paulo	19 194 100	1 968	Brésil	5	9 753
Djakarta	18 206 700	2 590	Indonésie	6	7 030
Osaka-Kyoto-Kobé	17 646 900	2 564	Japon	7	6 883
Delhi	17 367 300	1 295	Inde	8	13 411
Bombay (Mumbai)	17 340 900	4 360	Inde	9	3 977
Los Angeles	16 895 945	4 320	Etats-Unis	10	3 911

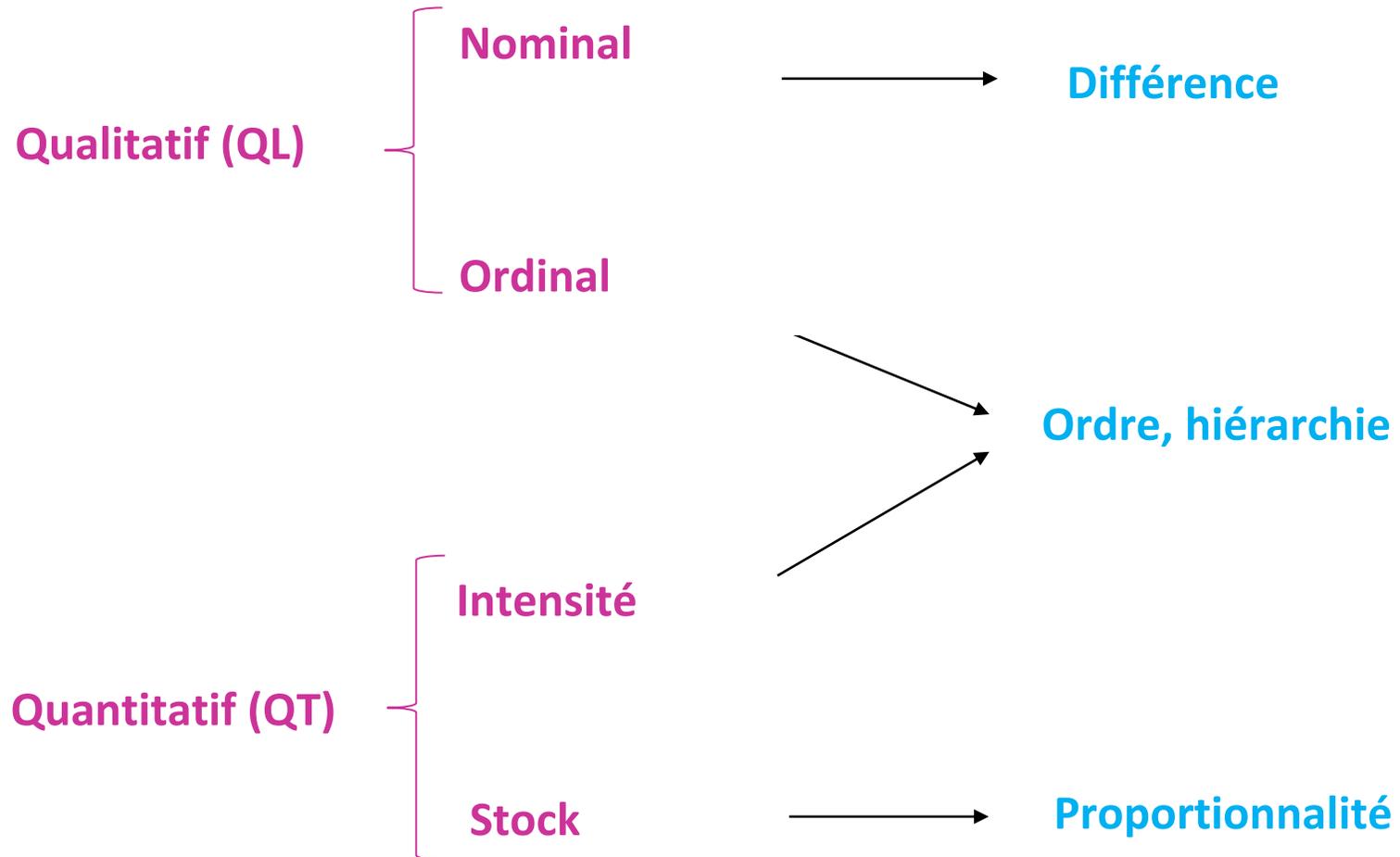
Qualitatif nominal

Qualitatif ordinal

# III. Les relations entre les modalités

## Type de variables / caractères

## Relation entre modalités



Exemples:  
Relations  
entre les  
modalités  
des variables

Aires Urbaines	Population	Superficie en km <sup>2</sup>	Pays	Rang pour la population	Densité (hab/km <sup>2</sup> )
Tokyo	33 413 000	7 019	Japon	1	4 760
Mexico	21 701 925	2 072	Mexique	2	10 474
New-York	21 199 865	8 680	Etats-Unis	3	2 442
Séoul	20 156 800	1 150	Corée du Sud	4	17 528
Sao Paulo	19 194 100	1 968	Brésil	5	9 753
Djakarta	18 206 700	2 590	Indonésie	6	7 030
Osaka-Kyoto-Kobé	17 646 900	2 564	Japon	7	6 883
Delhi	17 367 300	1 295	Inde	8	13 411
Bombay (Mumbai)	17 340 900	4 360	Inde	9	3 977
Los Angeles	16 895 945	4 320	Etats-Unis	10	3 911

Proportionnalité

Ordre/hiérarchie

Différence

Ordre/hiérarchie

# IV. Les variables visuelles

# Des variables statistiques... aux variables visuelles

Les **variables statistiques** conditionnent le choix des variables visuelles...

...en fonction de leur capacité à traduire visuellement **les rapports pouvant exister entre modalités** :

- La différence
- L'ordre / la hiérarchie
- La proportionnalité

# Sémiologie graphique et variables visuelles

## Sémiologie graphique

C'est l'ensemble des règles permettant l'utilisation d'un système graphique de signes pour la transmission d'une information correcte et accessible à un lecteur

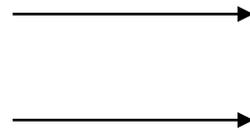
# Sémiologie graphique et variables visuelles

## Les variables visuelles

- Découlent des principes de la sémiologie graphique
- Constituent le « vocabulaire » du langage cartographique
- Permettent de différencier, comparer, ordonner
- Facilitent la mémorisation des informations transcrites sur la carte

**Les variables visuelles sont choisies en fonction de leur capacité à traduire le rapport entre les modalités :**

**1** **Différence**



**Forme**

**Couleur**

**(teintes différentes)**

**Ordre / hiérarchie**

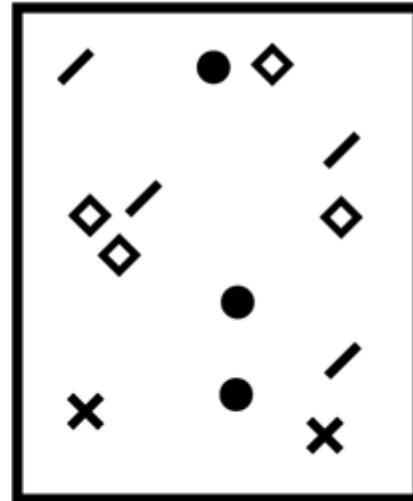
**Proportionnalité**

Trois implantations possibles du figuré cartographique : ponctuelle, linéaire, zonale

RELATION : différence

FORME

ponctuelle



Villes industrielles :

- Sucrerie (cross symbol)
- Fonderie (cross symbol)
- Textile (diamond symbol)
- Chimie (circle symbol)

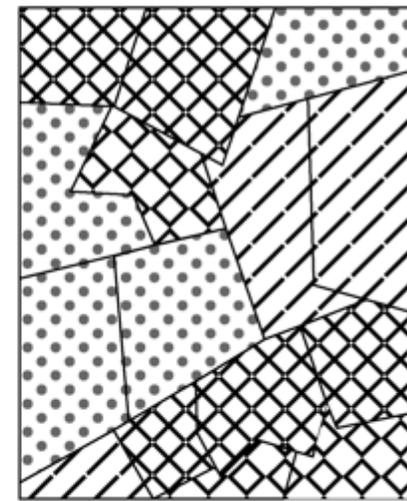
IMPLANTATION linéaire



Voies de communications :

- Route (solid line)
- Canal (dotted line)
- Chemin de fer (dashed line)

zonale



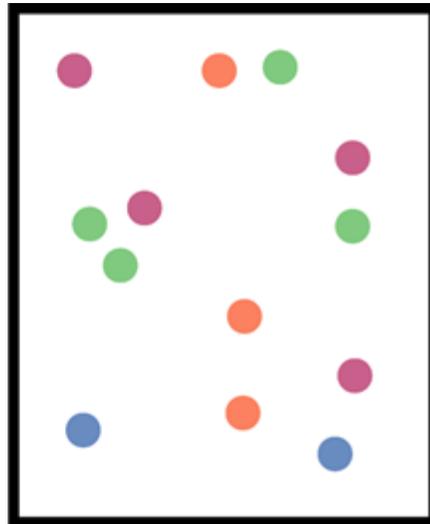
Forêts :

- Pins (cross-hatched pattern)
- Hêtres (diagonal hatched pattern)
- Bouleaux (cross-hatched pattern)
- Chênes (dotted pattern)

# RELATION : différence

COULEUR  
(teintes différentes  
d'intensité égale)

ponctuelle



Villes industrielles :

- Sucrierie
- Textile
- Fonderie
- Chimie

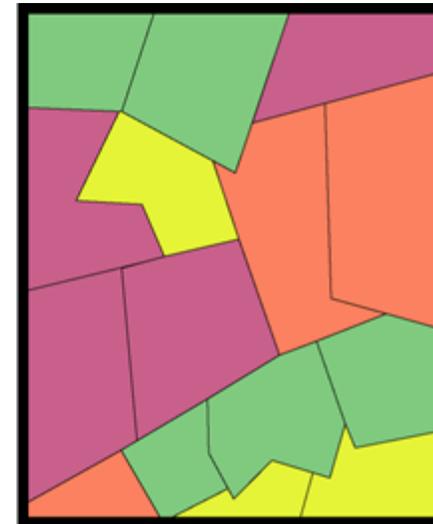
IMPLANTATION  
linéaire



Voies de communications :

- Route
- Chemin de fer
- Canal

zonale



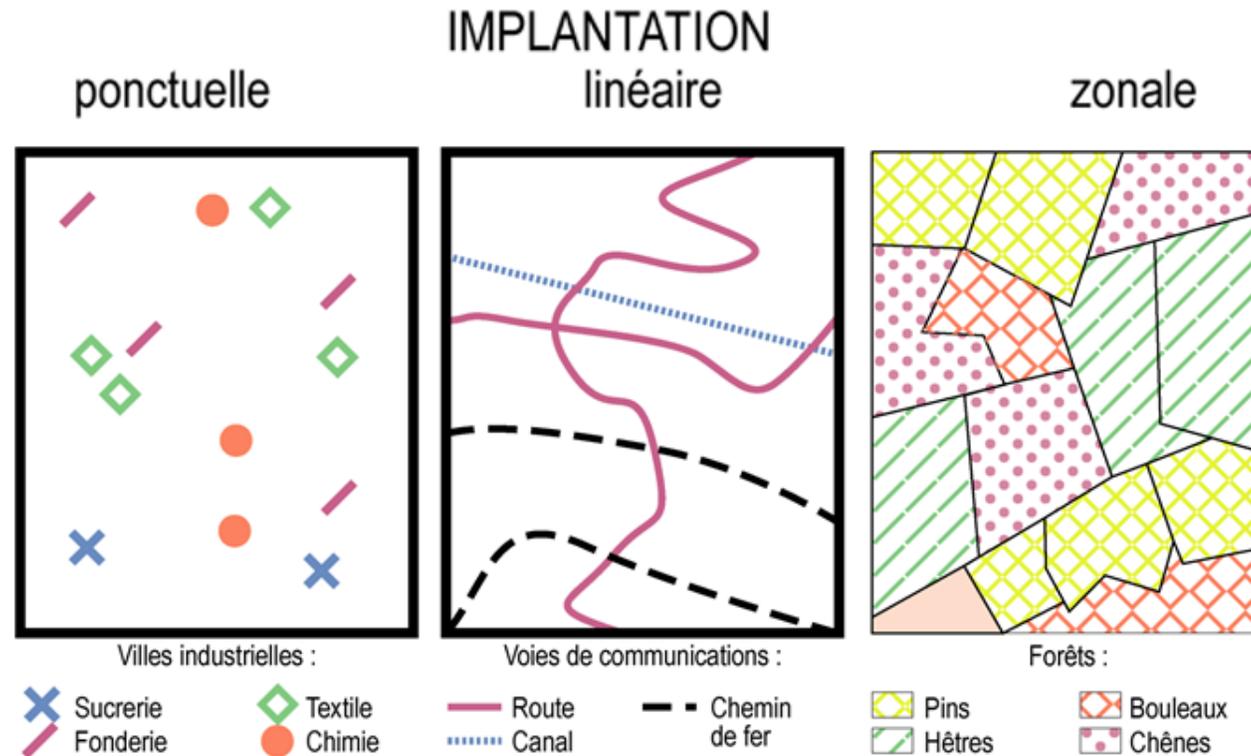
Forêts :

- Pins
- Bouleaux
- Hêtres
- Chênes

# Combiner les variables visuelles exprimant la différence (forme et couleur) renforce leur message

## RELATION : différence

FORME  
et COULEUR



**Les variables visuelles sont choisies en fonction de leur capacité à traduire le rapport entre les modalités :**

**Différence**



**Forme**



**Couleur**

**(teintes différentes)**

**2** **Ordre / hiérarchie**



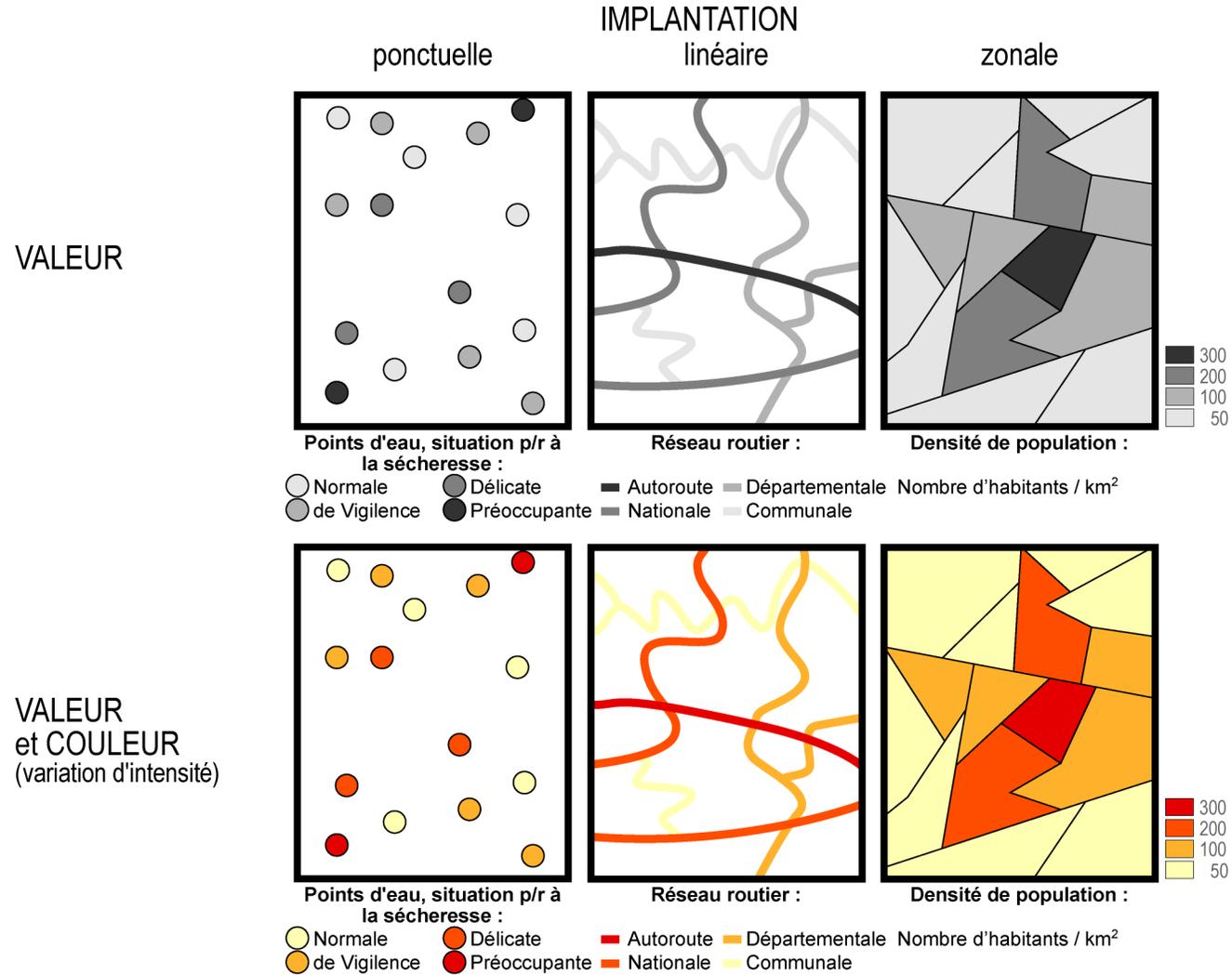
**Valeur :**

**- % Noir p/r au Blanc**

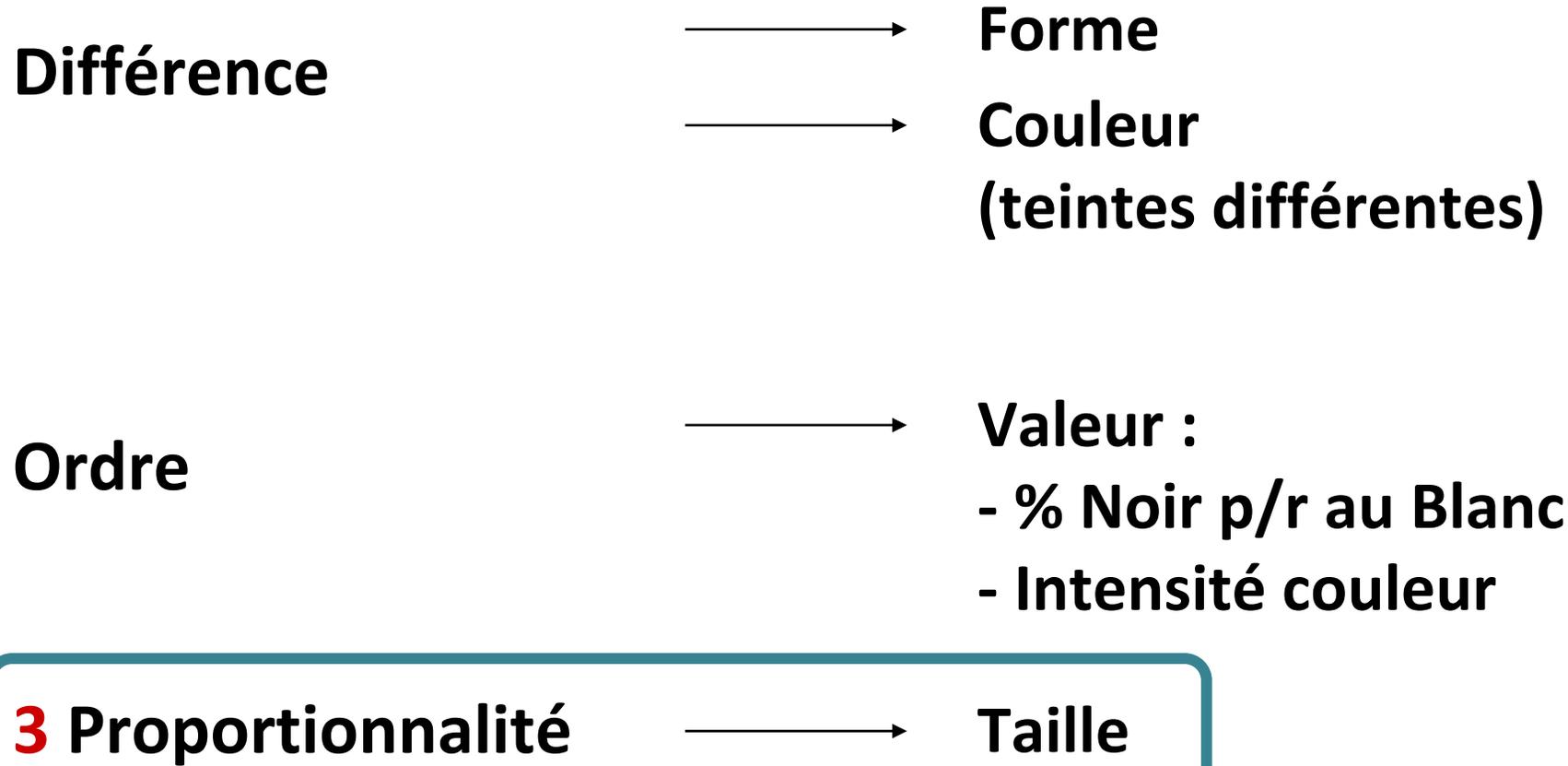
**- Intensité couleur**

**Proportionnalité**

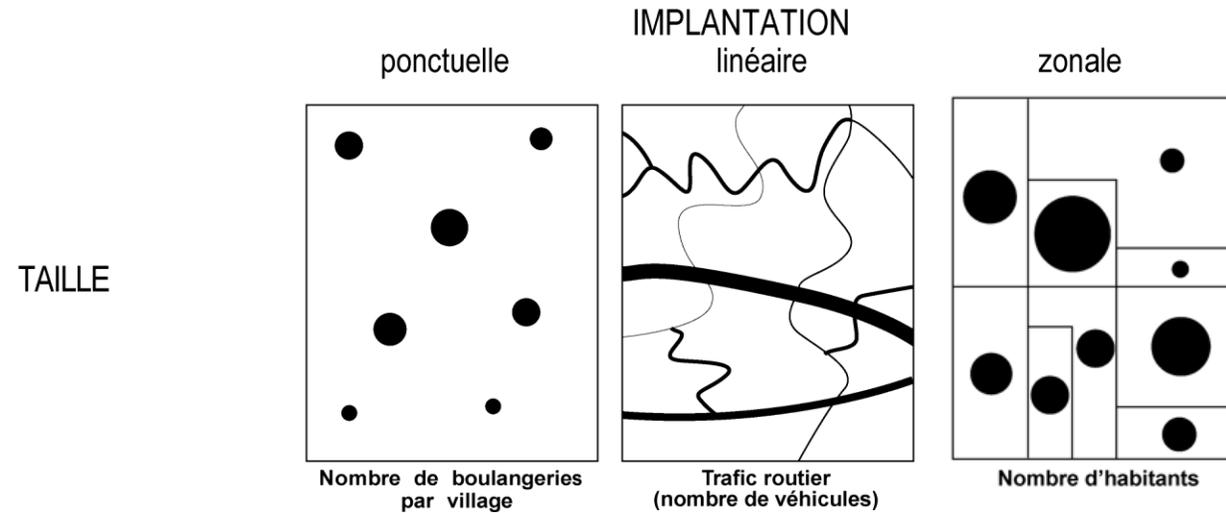
RELATION : ordre



**Les variables visuelles sont choisies en fonction de leur capacité à traduire le rapport entre les modalités :**



## RELATION : proportionnalité

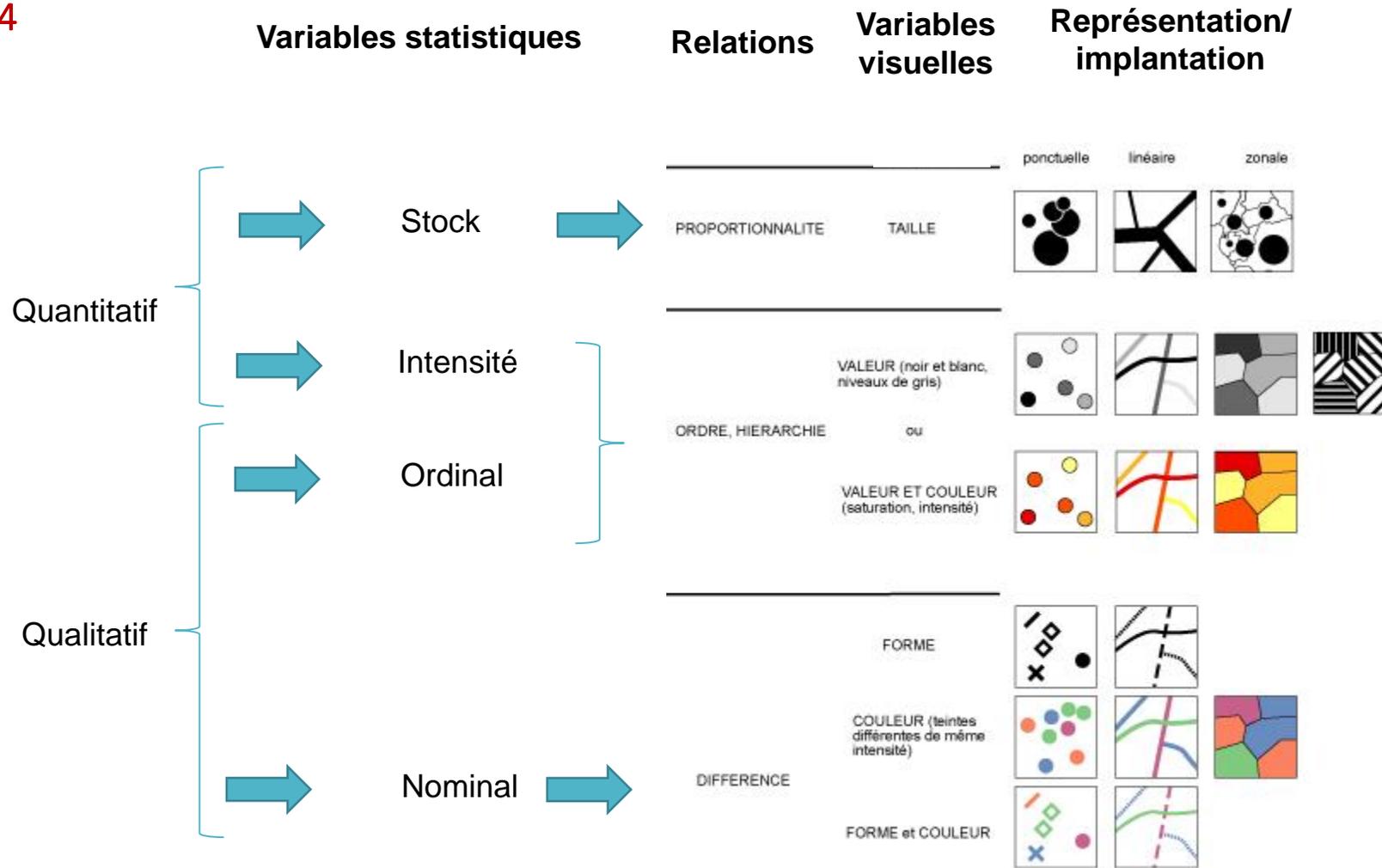


Une variation de longueur/largeur, hauteur, surface d'un figuré unique (ponctuel ou linéaire) traduit directement une variation de quantité, et donc une relation de proportionnalité.

En implantation zonale, on se contente souvent de représenter la variation d'une quantité à partir du figuré ponctuel situé au centre de l'unité géographique concernée

# Ce récapitulatif est à connaître « par cœur »

Fiche outil n° 4

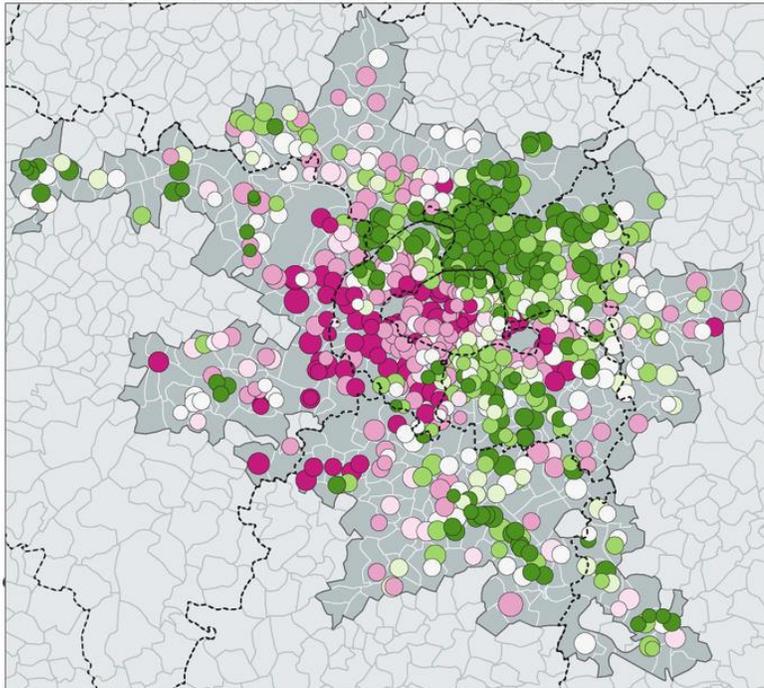


Variables statistiques	Description	Exemples	Relations entre modalités	Variables visuelles	Représentation selon l'implantation des unités géographiques (individus statistiques)	
Quantitative de stock (somme possible)	Série statistique qui exprime un <b>dénombrément</b> ou une <b>quantité mesurée à un moment ou sur une période donnés</b> . Les modalités ont une unité de mesure et peuvent être additionnées (le total fait sens pour la population étudiée). Le zéro signifie "absence de".	<i>Nombre d'habitants par région française</i> <i>Nombre de véhicules en circulation par section de route</i> <i>Débit sur un cours d'eau</i>	PROPORTIONNALITE	TAILLE	ponctuelle	
					linéaire	
Quantitative d'intensité (Somme aberrante)	Série statistique exprimant : • des modalités dans une unité dont le <b>zéro est arbitraire</b> • des modalités résultant d'un <b>rapport</b> de deux données quantitatives de stocks (unité de population (%), de surface ou de volume (km², m³...)) Série statistique chiffrée dont seuls les intervalles entre modalités ont un sens, pas la somme.	<i>Température, altitude</i>  <i>Espérance vie, densité de pop°</i> <i>taux de natalité, nombre de cas pour 1000 habitants</i>	ORDRE, HIERARCHIE	VALEUR  % de noir par rapport au blanc (aplats ou trames)  ou  variation d'intensité de la couleur	zonale	
Qualitative ordinale	Série statistique de modalités non chiffrées, ordonnées (la signification des mots est ordonnée). Pas d'opération mathématique possible en dehors du tri des modalités par ordre croissant ou décroissant (rang).	<i>Petit, moyen, grand</i> <i>1er, 2e, 3e rang</i> <i>autoroute, RN, RD, Chemin maternelle, primaire, collège, lycée</i> <i>étage montagnard, alpin, nival</i>				
Qualitative nominale	Série statistique de modalités non chiffrées et équivalentes, non ordonnées. Aucune opération mathématique n'est possible. Le seul classement envisageable est alphabétique (arbitraire).	<i>Secteurs d'activité (agriculture, commerces...)</i> <i>types de roches (calcaire, granite...), espaces boisés (pins, hêtres, bambous...) etc...</i>	DIFFERENCE	FORME  COULEUR (TEINTES différentes, de même intensité)		

Pour mieux mémoriser, ajoutez vos propres exemples !

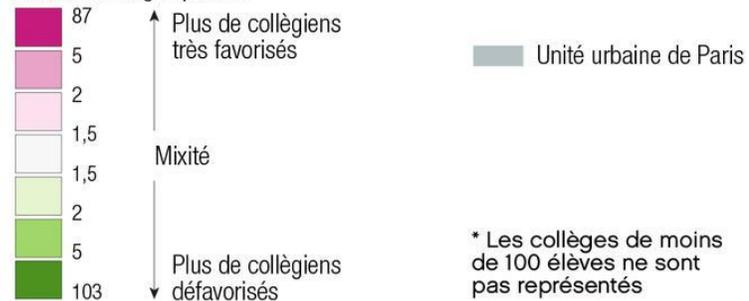
# Des exemples pour retenir le vocabulaire

## Dans l'agglomération parisienne, la ségrégation scolaire se lit à l'échelle des établissements



### Rapport entre les plus favorisés et les moins favorisés

Dans les collèges publics



POUR MIEUX COMPRENDRE

Exemple de lecture : « Dans la majorité des collèges publics de **Seine-Saint-Denis**, il y a au moins 5 fois plus de collégiens défavorisés que de collégiens très favorisés. »

Source : MEN-DEPP, Base Scolarité, rentrée 2015  
Réalisation : CGET, 2018

Quels individus ?

Collèges

Quel type de variable statistique ?

QT d'intensité (rapport entre le nombre d'élèves défavorisés et le nombre favorisé)

Quelle implantation ?

Ponctuelle

Quelle relation entre les modalités ?

Ordre

Quelle variable visuelle ?

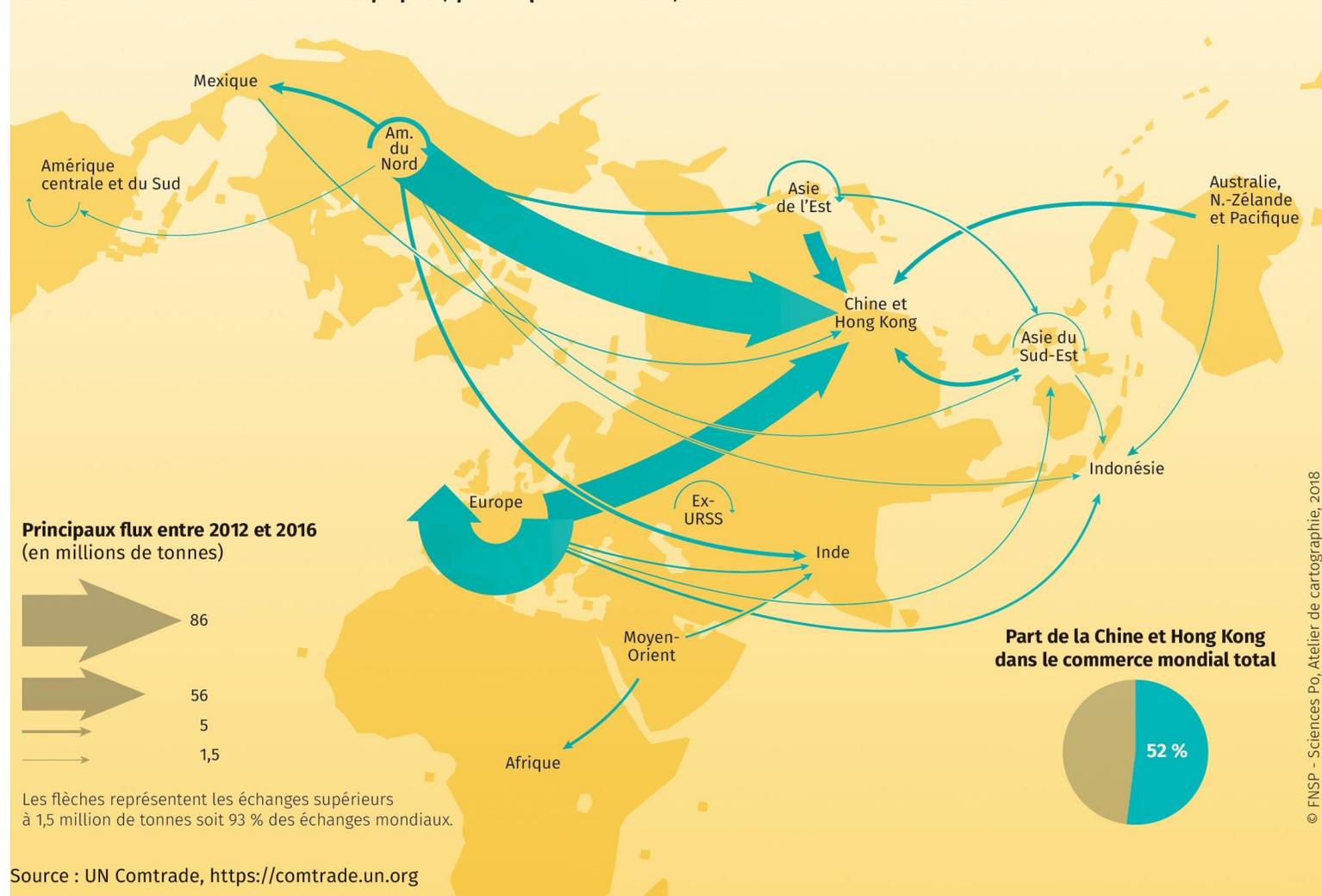
Valeur



POUR MIEUX COMPRENDRE

Exemple de lecture : « Dans la majorité des collèges publics de **Seine-Saint-Denis**, il y a au moins 5 fois plus de collégiens défavorisés que de collégiens très favorisés. »

## Commerce mondial de déchets papier, plastique et textile, 2012-2016



Quels individus ?

Continents, régions, pays

Quel type de variable statistique ?

QT stock (échanges commerciales déchets papier, plastiques, textile)

Quelle implantation ?

Linéaire

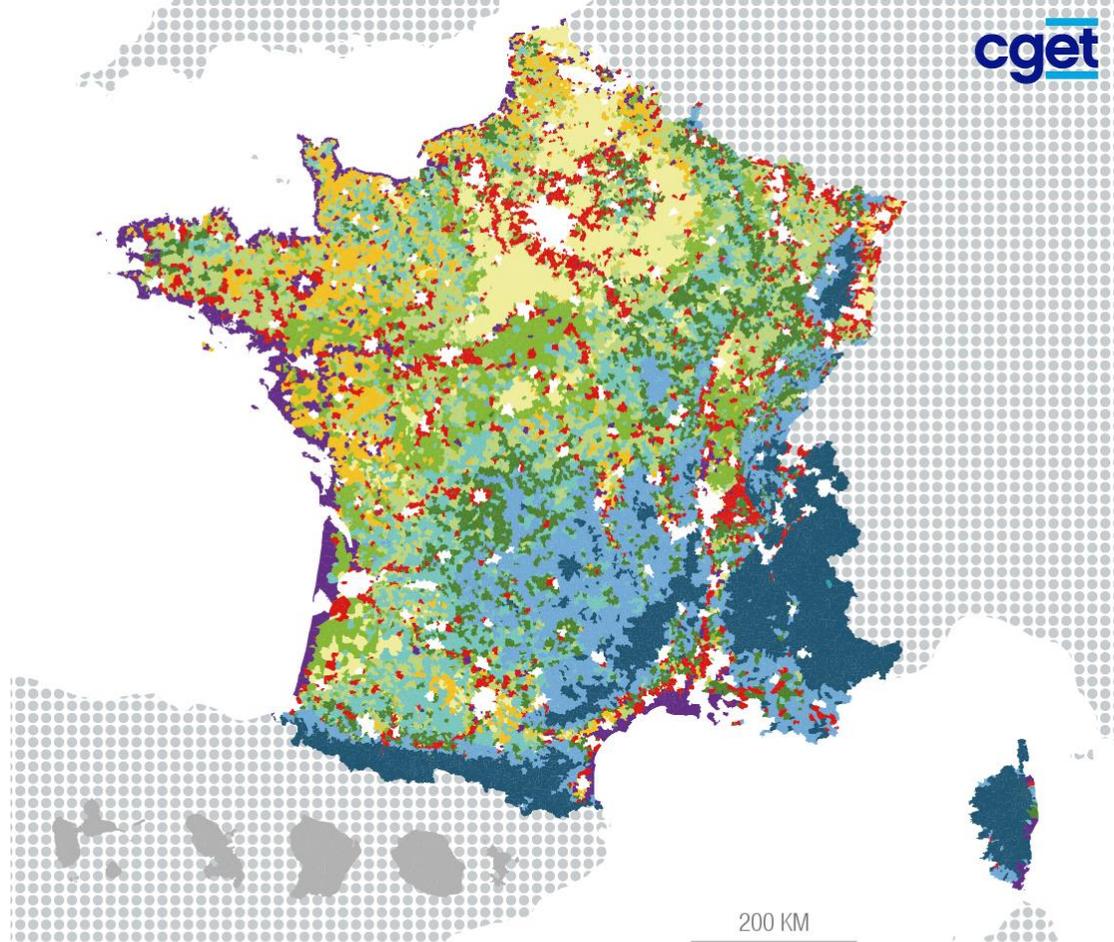
Quelle relation entre les modalités ?

Proportionnalité

Quelle variable visuelle ?

Taille

<https://bibnum.sciencespo.fr/s/catalogue/ark:/46513/sc16xfxd>



Typologie des campagnes en fonction des paysages

 Espace rural artificialisé : présence forte du bâti en paysage diversifié	 Collines et vallées, paysages diversifiés
 Paysage artificialisé en contextes semi-naturel ou humide, notamment littoraux	 Collines et vallées en prairies, peu artificialisé
 Paysage de cultures avec forte marque du bâti	 Relief marqué, paysages diversifiés : forêts, prairies, milieux semi-naturels
 Plaines et plateaux, grands openfields	 Montagne, très forts reliefs et végétation semi-naturelle
 Plaines et plateaux, campagne diversifiée avec bâti marqué	 Unités urbaines de plus de 10 000 emplois
 Plaines et plateaux, prairies et forêts	 Hors champs de la typologie

Quels individus ?

Unités homogènes paysagères

Quel type de variable statistique ?

QL nominale (types de campagnes en fonction des paysages)

Quelle implantation ?

Zonale

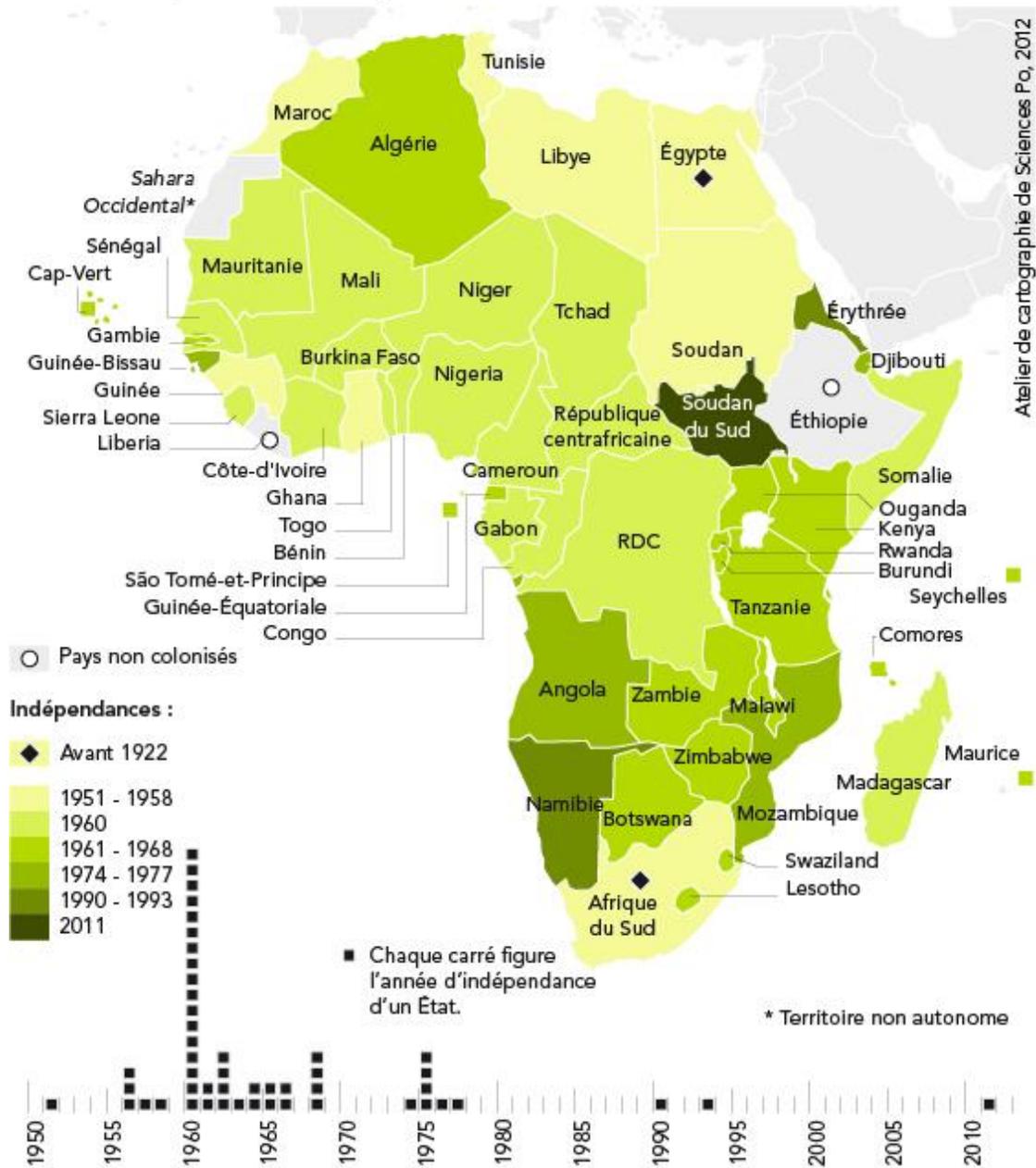
Quelle relation entre les modalités ?

Différence

Quelle variable visuelle ?

Couleur

# Chronologie des indépendances africaines



Source : Rédaction Afrique Contemporaine, Agence Française de Développement, De Boeck  
<https://bibnum.sciencespo.fr/s/catalogue/ark:/46513/sc16d5wz>

Quels individus ?

Etats

Quel type de variable statistique ?

QL ordinaire (année d'indépendance)

Quelle implantation ?

Zonale

Quelle relation entre les modalités ?

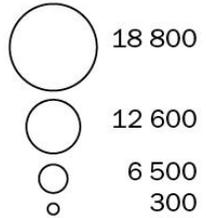
Ordre

Quelle variable visuelle ?

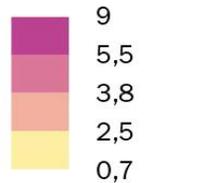
Valeur

**FIG. 56 Villes africaines, 2015**

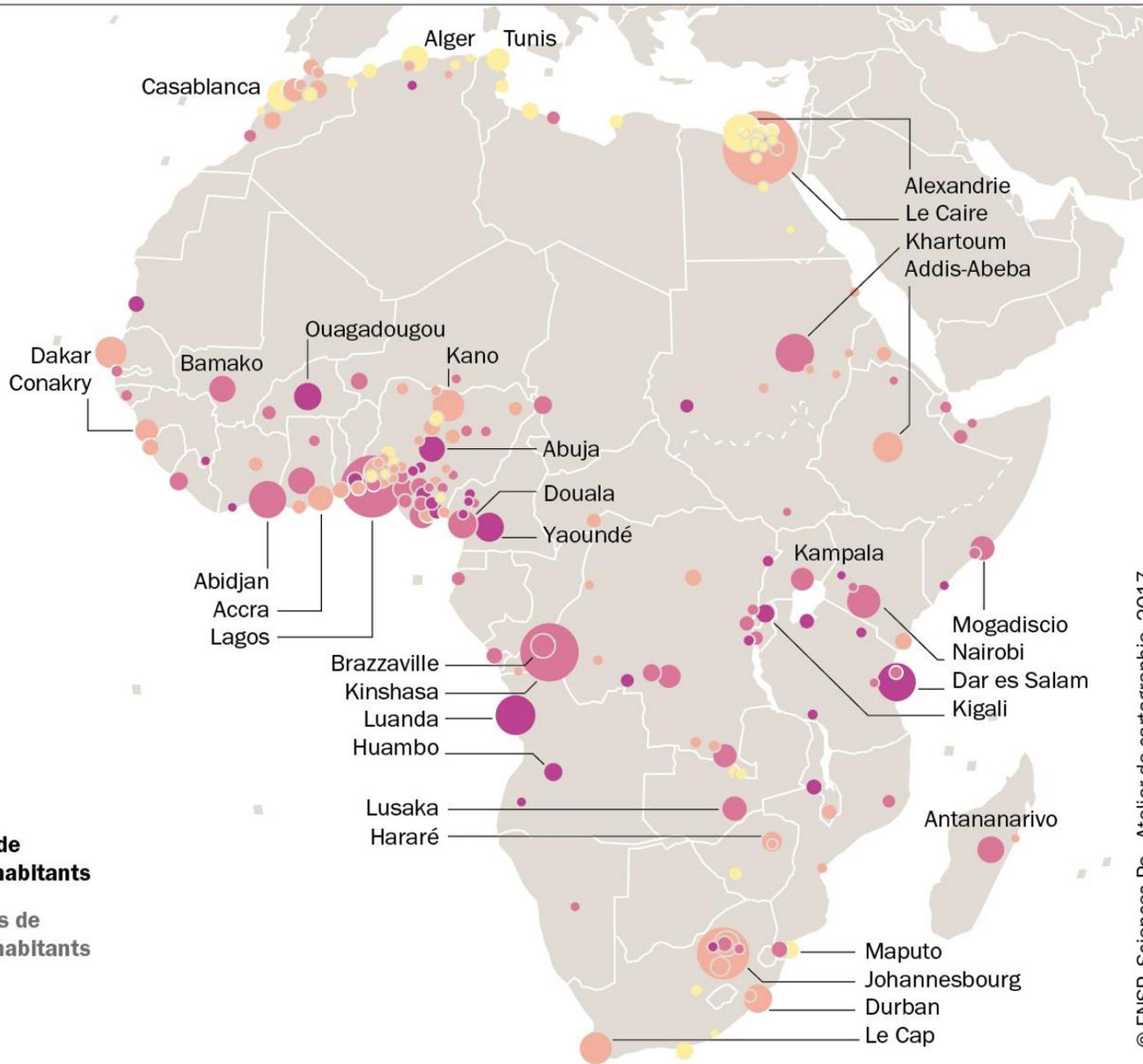
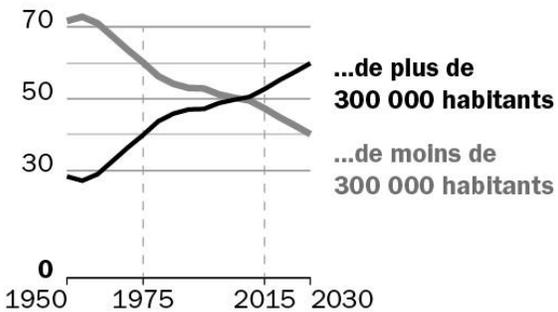
**Villes de plus de 300 000 habitants, 2015**  
(en milliers)



**Taux de croissance annuel entre 1975 et 2015 (en %)**



**Part de la population urbaine vivant dans les villes... (en %)**



© FNSP. Sciences Po - Atelier de cartographie, 2017

Quels individus ?

Villes africaines de plus de 300 000 hab

Quel type de variable statistique ?

- 1) QT d'intensité (taux de croissance)
- 2) QT de stock (taille des villes)

Quelle implantation ?

ponctuelle

Quelle relation entre les modalités ?

- 1) ordre/hiérarchie
- 2) proportionnalité

Quelle variable visuelle ?

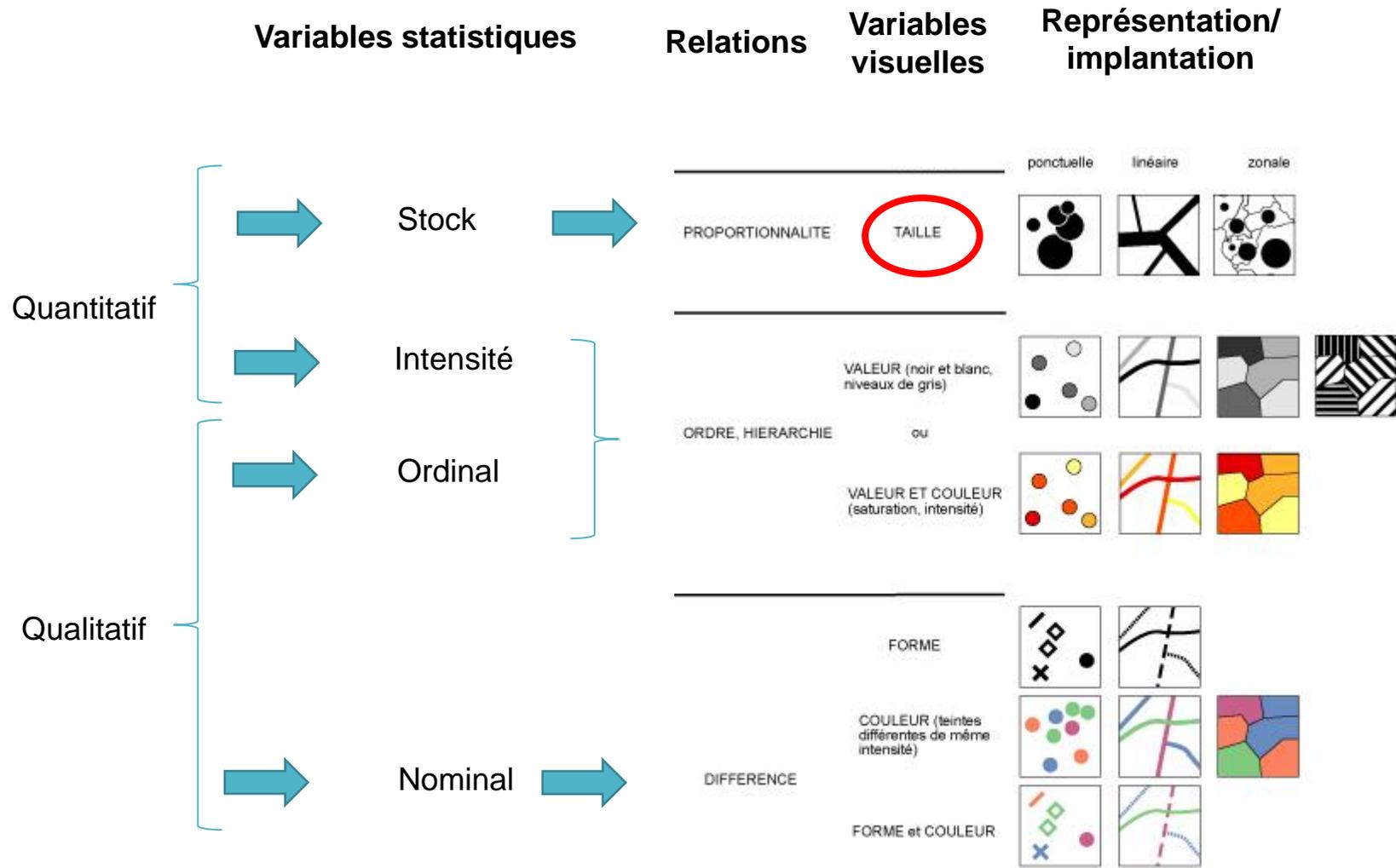
- 1) valeur
- 2) taille

# Précisions sur 2 variables visuelles

# La variable visuelle TAILLE

Traduit la relation de proportionnalité  
entre les modalités

# RAPPEL



# Comment choisir la taille minimum et maximum des cercles proportionnels ?

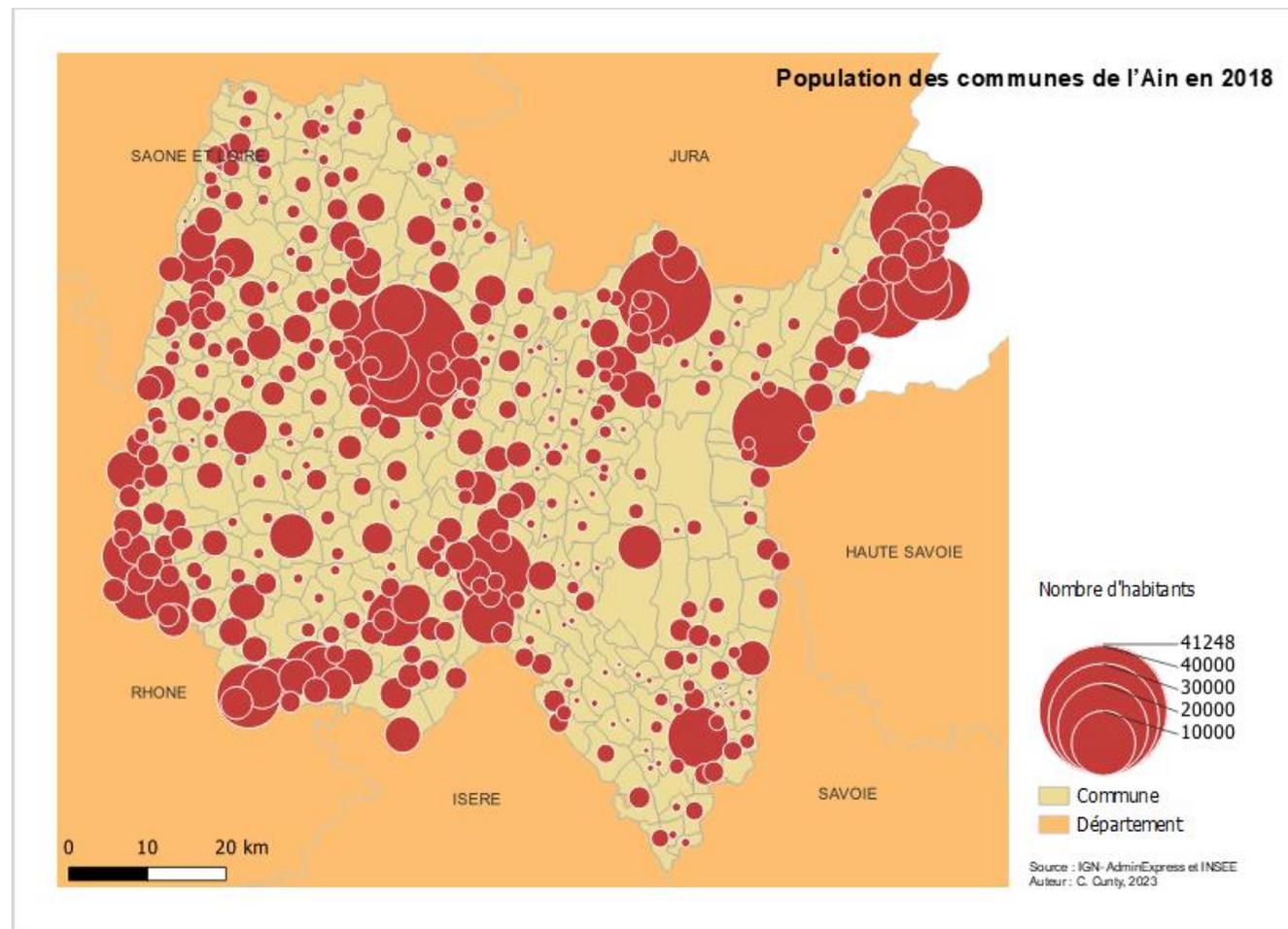
1. Calcul du **rapport de proportion** entre la valeur la plus élevée et la moins élevée :

$$\Rightarrow 41\,248 / 27 = 1\,527,7 \approx 1\,500$$

La commune de Armix est 1 500 fois moins peuplée que la commune de Bourg-en-Bresse

2. **Rendre visible ce rapport de proportion** sur la carte :

$\Rightarrow$  Ici, un rapport proportionnel cartographique de 1/1 500 soit les cercles de 0.02 à 30 (30/1 500 = 0.02) si on décide que la taille maximal d'un cercle = 30

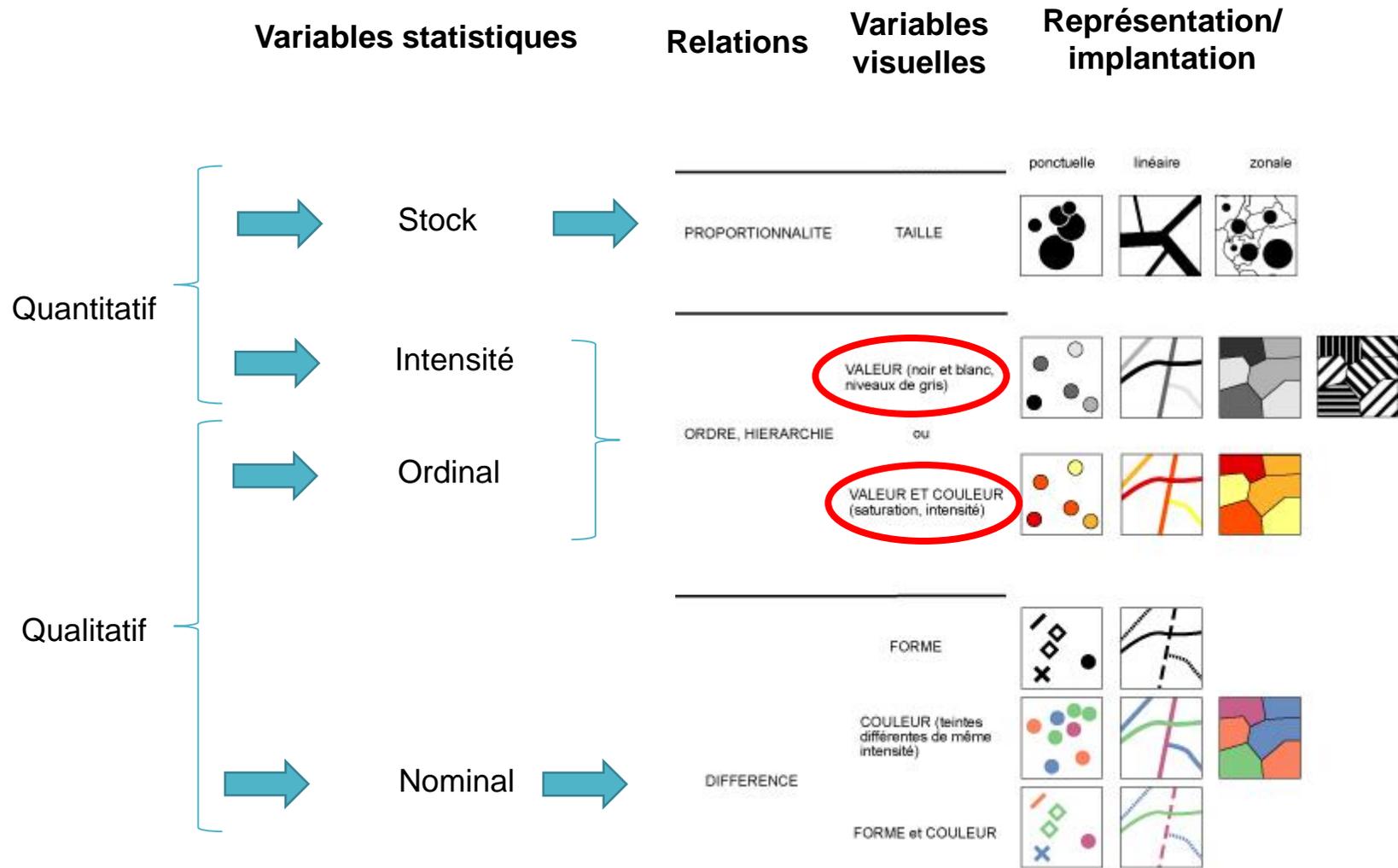


# La Variable Visuelle VALEUR

Traduit l'ordre et la hiérarchie

Comment choisir des couleurs qui traduisent visuellement l'ordre, la hiérarchie ?

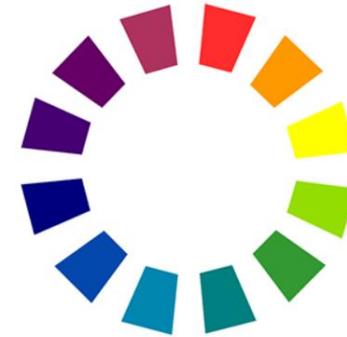
# RAPPEL



# Paramètre de la couleur

La « couleur » au sens commun est définie par 3 paramètres :

- **La teinte** : nature de la couleur (ex: bleu, jaune ...)
- **La saturation** : le *degré* de pureté/dilution (vive ou terne)
- **L'intensité** : du *plus sombre* au *plus clair* (dépend des deux autres)



# Teinte, saturation et intensité

Attention, ces trois paramètres ont des effets visuels différents

Site conseillé  
<http://colorbrewer2.org>

Relations

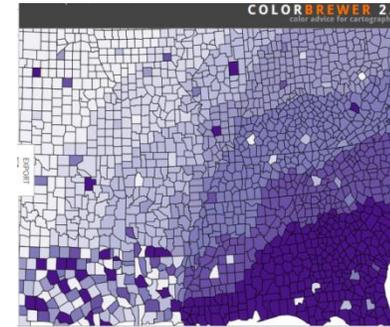
Paramètres  
couleurs

Exemples

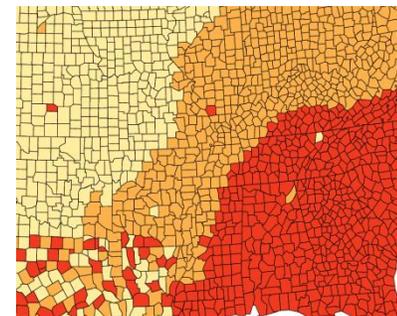
Ordre /  
Hiérarchie

Saturation

Intensité



*Camaïeu  
de violet*



*Gradation  
harmonique  
chaude (rouge /  
jaune*

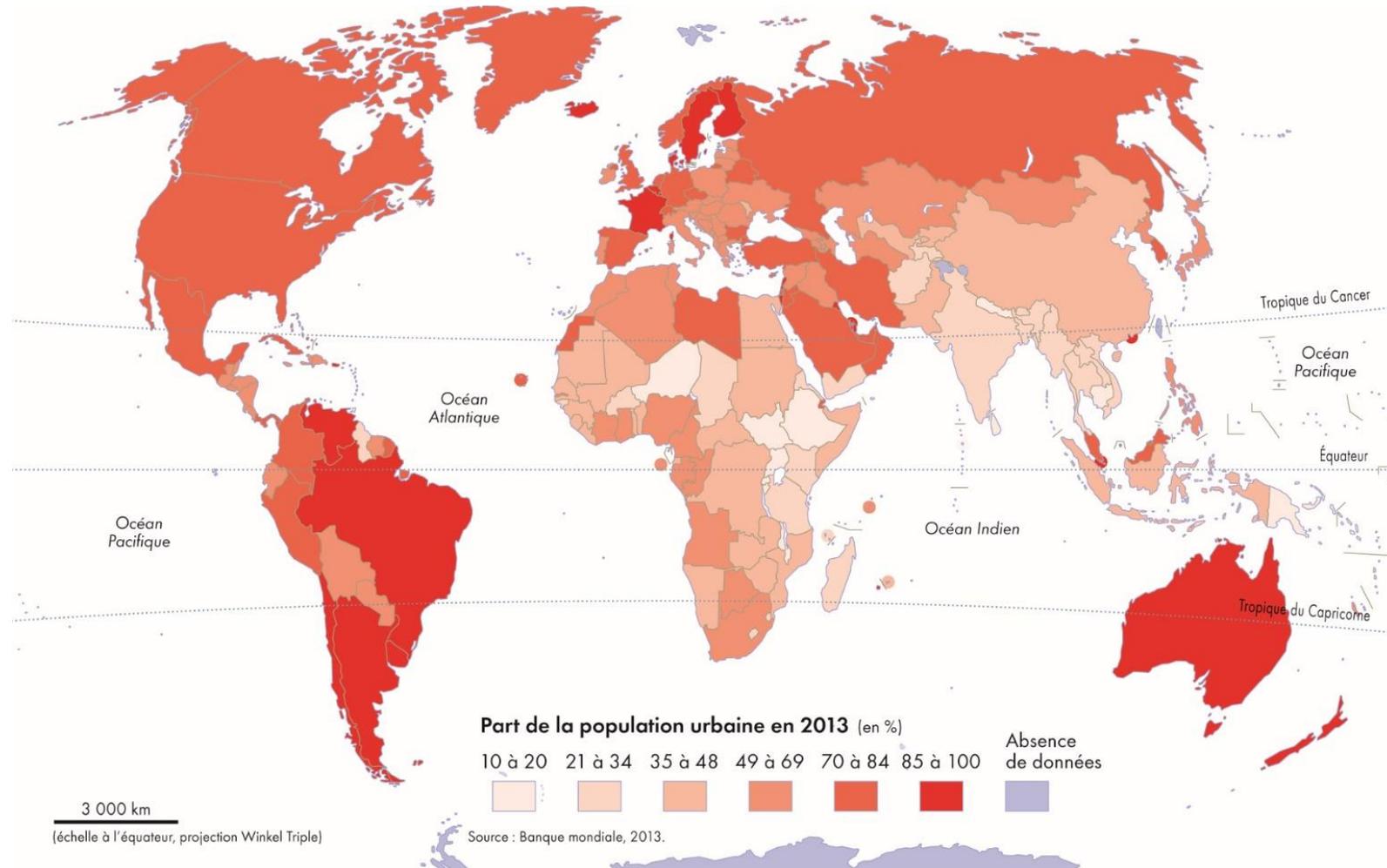
**Variables  
visuelles  
Valeurs**

Source : <http://colorbrewer2.org>

# Exemple de variation de saturation d'une teinte – le camaïeu

La teinte rouge est déclinée par saturation

La population de citadins

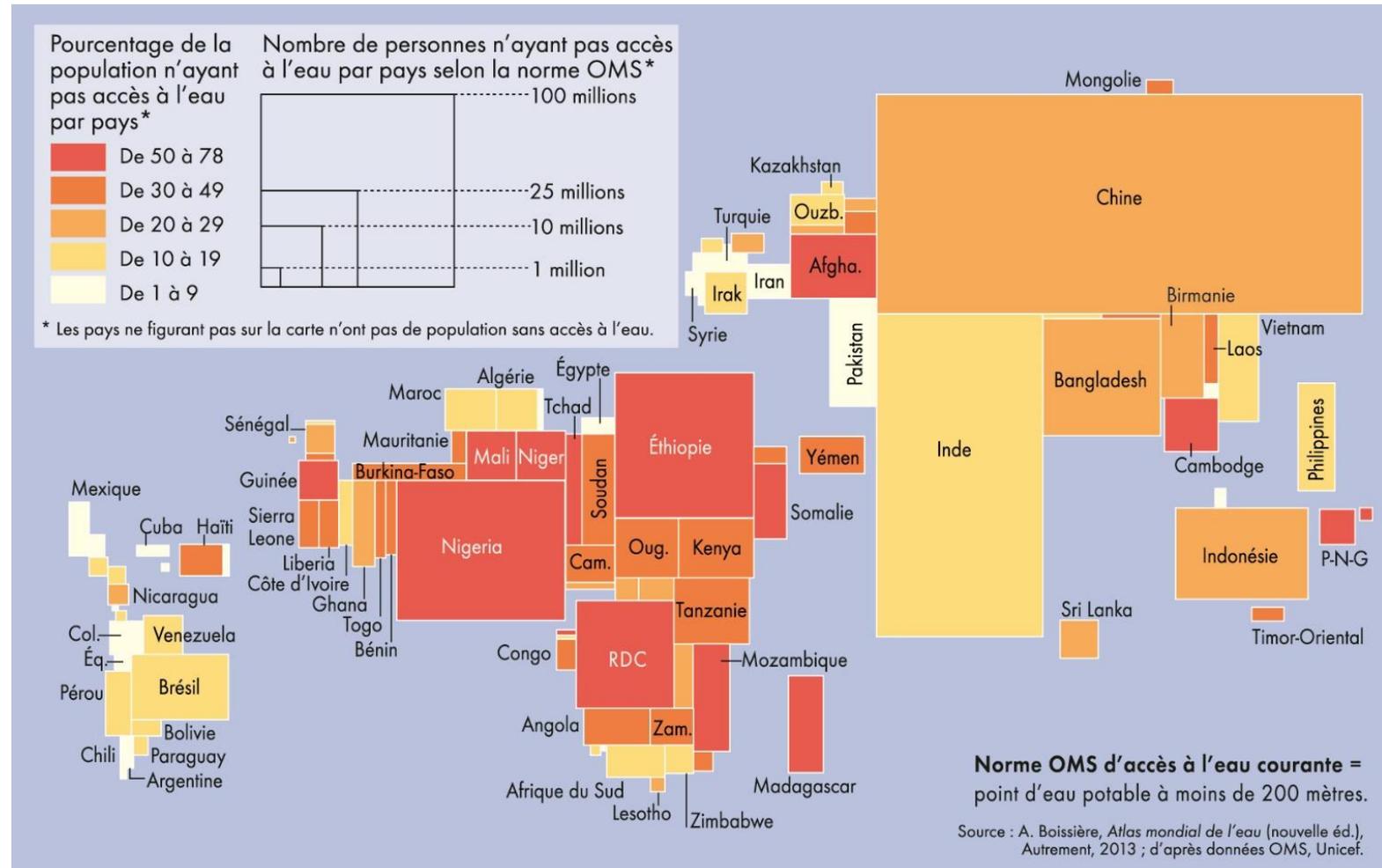


Source : Tétart F. (dir), 2014. *Grand Atlas 2015*. Courrier International, Autrement, p. 65

# Exemple de gradation harmonique

Variation de l'intensité en utilisant des teintes ordonnées

## Le non accès à l'eau

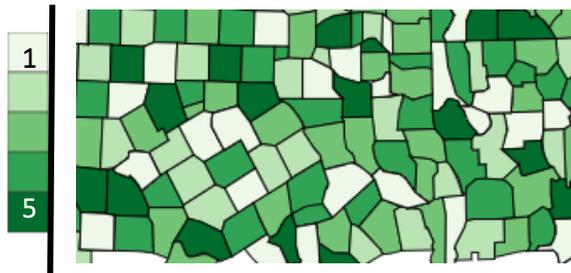


Source : Tétart F. (dir), 2014. *Grand Atlas 2015*. Courrier International, Autrement, p. 70

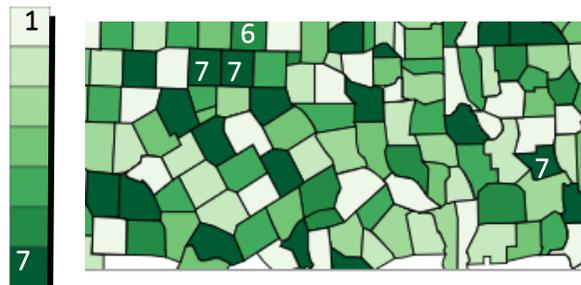
# Les capacités de l'œil sont limitantes

D'après des tests de perception, l'œil humain n'est pas capable de distinguer plus de 7 paliers de valeurs ordonnées (en couleur comme en Noir et Blanc).

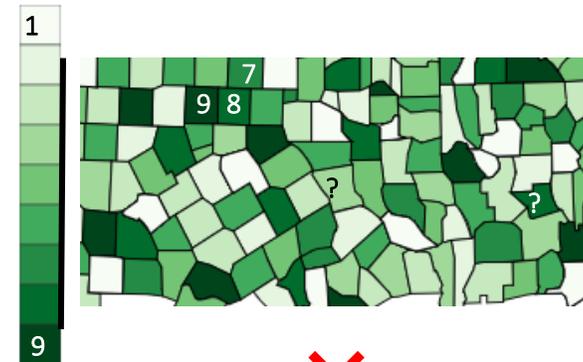
5 paliers : lecture aisée



7 paliers : lecture possible



9 paliers : lecture non fiable



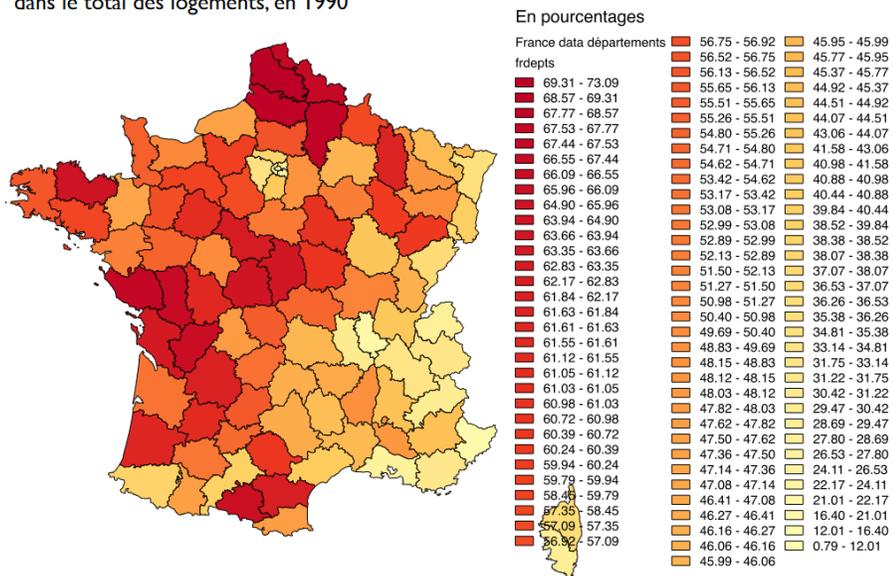
# La discrétisation de la série statistique d'une variable QT d'intensité

Faire des classes

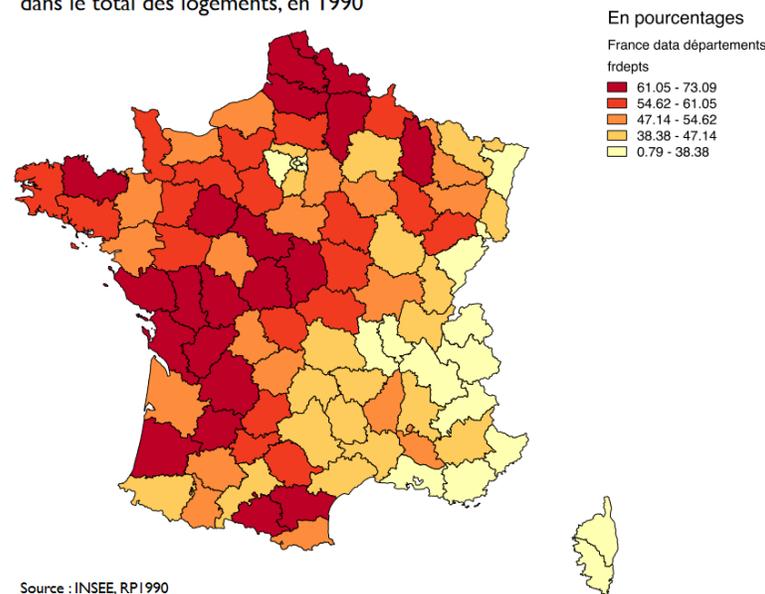
# Pourquoi discrétiser des données avant de faire une carte ?

Limites physiologiques de perception (maximum 7 paliers de valeurs ordonnées)

Part des maisons individuelles dans le total des logements, en 1990



Part des maisons individuelles dans le total des logements, en 1990



Cartographie d'une variable quantitatives d'intensité sans et avec discrétisation  
Auteur : L. JÉGOU

## Quelles données discrétiser ?

Variables statistiques quantitatives d'intensité -> pour représenter l'ordre par la variable visuelle « valeur »

ATTENTION : les variables quantitatives de stock, représentées par la VV taille ne doivent pas être discrétisées (sinon, il y a perte d'information)

# Définition de la discrétisation

Simplification de l'information pour qu'elle soit lisible sur la carte

Représentation non neutre de la réalité (**biais**)

Opération qui permet de **découper en classes** (sous ensembles) une série statistique continue

Regroupement des objets géographiques

- aux **caractéristiques voisines** dans une **même classe** (homogénéité interne)
- et ceux aux **caractéristiques éloignées** dans des **classes distinctes** (différentiation externe)

Cette simplification doit s'appuyer sur :

- l'effectif de la population, c'ad le nombre d'individus => nombre de classes
- les caractéristiques de distribution statistique => position des limites de classes

# Comment choisir le nombre de classes?

## Grand compromis entre significativité statistique et lisibilité cartographique

- Résumer au mieux la distribution (vérité)  
=> *Optimum statistique* : le plus grand nombre de classes
- Construire une carte efficace (clarté)  
=> *Optimum cartographique* : un faible nombre de classes (7 maximum)

### Choix du nombre de classes

- Nombre de classes dépend de l'effectif total (nombre d'individus)
  - 20 individus → 5 classes
  - 100 individus → 6 classes
  - Au-delà de 100 individus → 7 classes (maximum pour un dégradé sur une teinte)



# L'influence du nombre de classes et les bornes des classes

## Hors cours regardez

Carte interactive sur la Part des nuitées de résidents étrangers en 2015

Auteur : M@ppemonde , <http://mappemonde.mgm.fr/geovisu/119/>

Le **nombre de classes** influence le résultat cartographique :

Sur l'application faire varier le nombre de classes, qu'observez-vous ?

Le choix des **bornes de classes** peut aussi **changer** le **résultat visuel** et donc le **message** de la carte, à nombre de classes constant !

Sur l'application faire varier le type de discrétisation (quantile, Jenks...), qu'observez-vous (légende, résultat cartographique...) ?

# Comment choisir les limites des classes ?

*Règle n°1* : Le choix passe toujours par une analyse préalable de la **distribution statistique** de la variable

*Règle n°2* : du « contexte » : la **carte unique** ou la carte est **comparée à d'autres**

*Règle n°3* : le choix peut être influencé par **l'objectif de la carte (message, public)**

Par exemple : Choix de classes régulières (ex : de 20% en 20%) dans la presse, quelque soit la distribution

On privilégie des bornes précises à une décimale dans un contexte académique ou on arrondit à des valeurs faciles à mémoriser dans la presse.

# RÈGLE N° 1 : COMMENT CONNAÎTRE LA DISTRIBUTION?

Tableau élémentaire

Etudiants	Notes
A	2
B	8
C	6
D	3
E	15
F	7
G	2
H	8
I	10
J	7
K	8

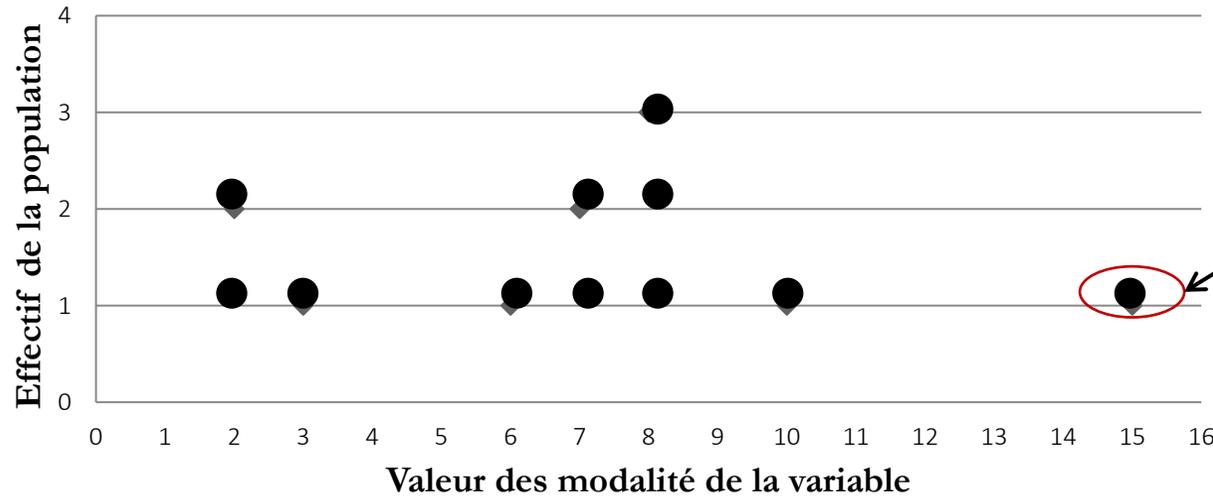


Modalité de la Variable (Notes)	Effectifs = nb d'individus (étudiants)
2	2
3	1
6	1
7	2
8	3
10	1
15	1
Total d'individus	11

Tableau Dénombrement



Observer le diagramme de distribution



# RÈGLE N° 1 : COMMENT CONNAÎTRE LA DISTRIBUTION ?

Tableau élémentaire

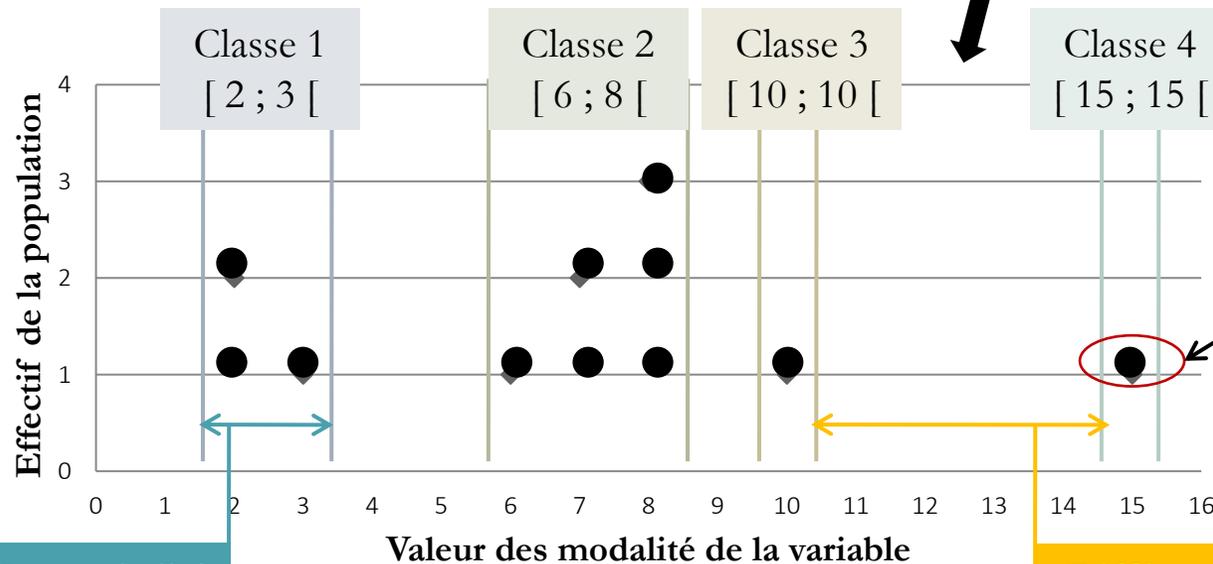
Etudiants	Notes
A	2
B	8
C	6
D	3
E	15
F	7
G	2
H	8
I	10
J	7
K	8



Modalité de la Variable	Effectifs de la population = Nb d'individus
2	2
3	1
6	1
7	2
8	3
10	1
15	1
Total d'individus	11

Tableau Dénombrement

Observer le diagramme de distribution



Homogénéité interne

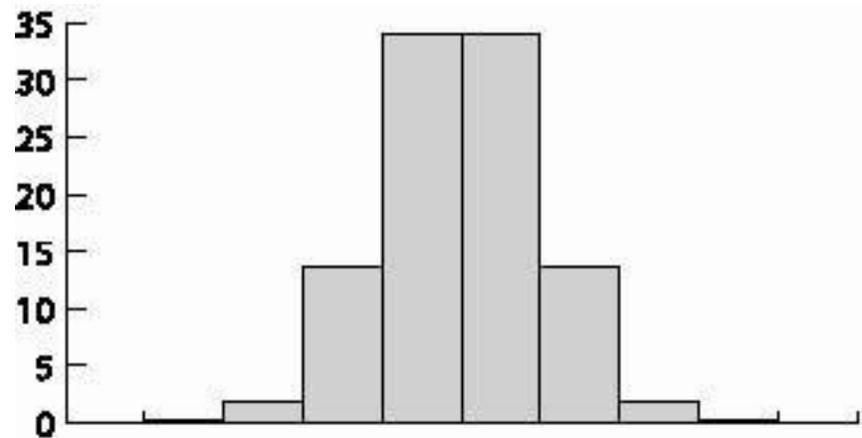
Différenciation externe

Individu E

# RÈGLE N° 1 : COMMENT CONNAÎTRE LA DISTRIBUTION ?

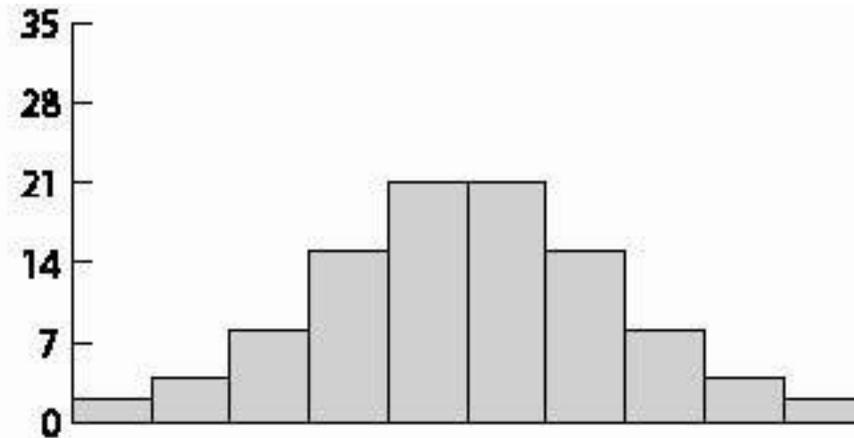
Des formes de distribution caractéristiques

Distribution symétrique



Diagrammes de distribution :

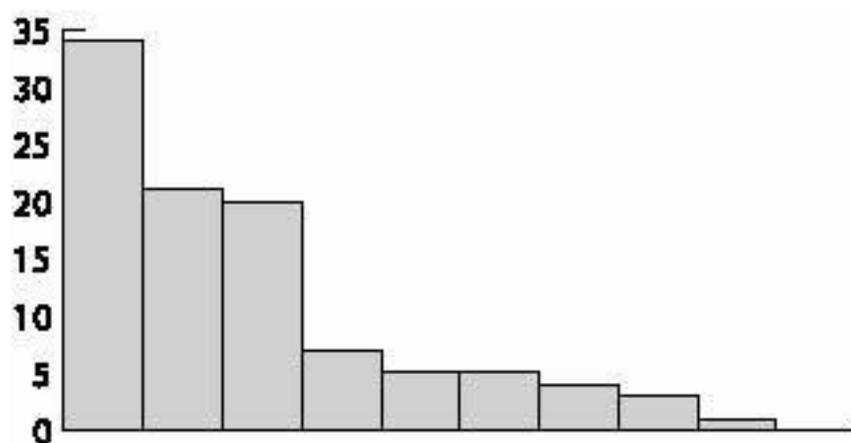
- En ordonné (axe Y) : Nombre d'individus ou %
- En abscisse (axe X) Valeurs de la variable (ex. densité de la population)



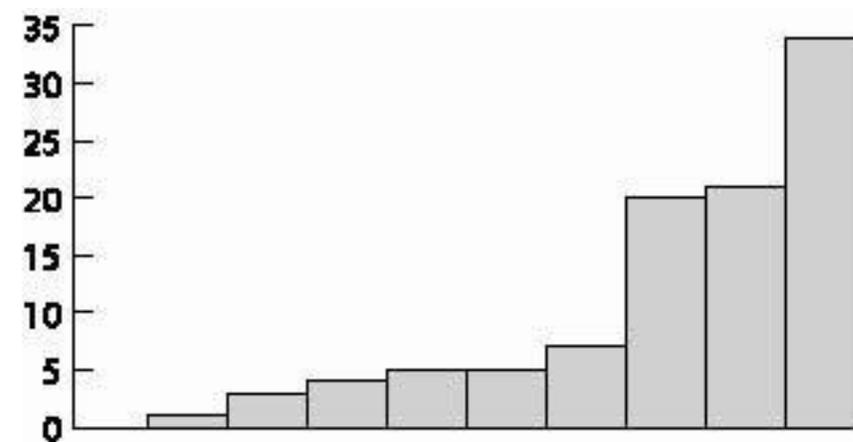
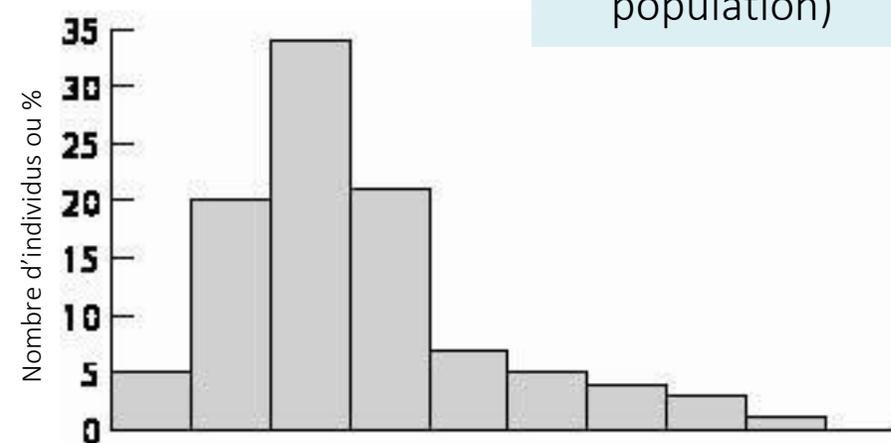
# RÈGLE N° 1 : COMMENT CONNAÎTRE LA DISTRIBUTION ?

Des formes de distribution caractéristiques

Distribution dissymétrique à gauche



Distribution dissymétrique à droite



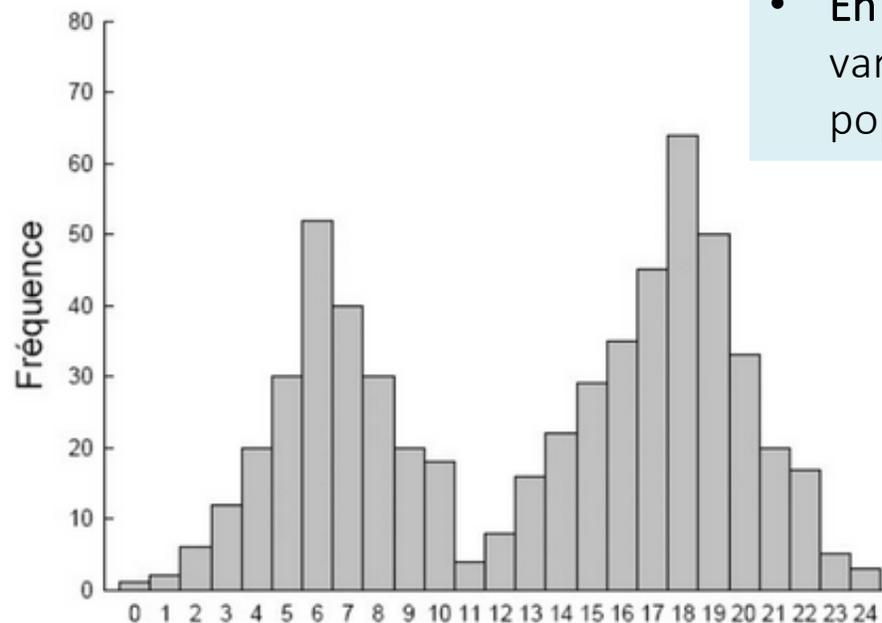
Diagrammes de distribution :

- En ordonné (axe Y) : Nombre d'individus ou %
- En abscisse (axe X) Valeurs de la variable (ex. densité de la population)

# RÈGLE N° 1 : COMMENT CONNAÎTRE LA DISTRIBUTION?

Des formes de distribution caractéristiques

Plurimodale

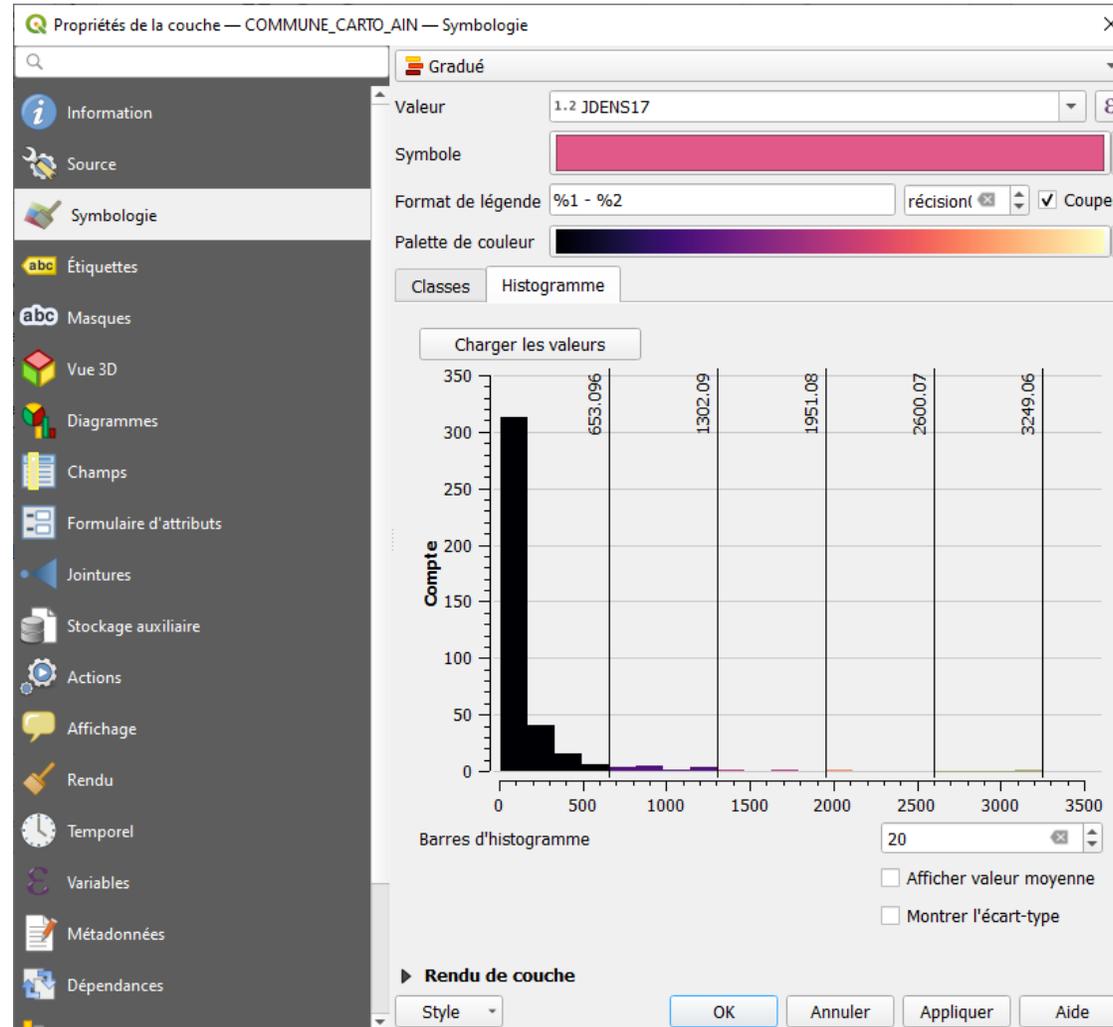


Diagrammes de distribution :

- En ordonné (axe Y) : Nombre d'individus ou %
- En abscisse (axe X) Valeurs de la variable (ex. densité de la population)

# RÈGLE N° 1 : COMMENT CONNAÎTRE LA DISTRIBUTION?

Avec QGIS => dans Symbologie Histogramme



# Comment répartir les individus dans les classes ?

Il existe plusieurs méthodes faciles à mettre en œuvre et couramment utilisées en géographie

Méthode des seuls naturels – JENKS

Classes d'**effectifs** égaux (quantiles)

Classes basées sur la **moyenne et l'écart-type** (standardisation)

- Ne permet pas la comparaison
- Permet la comparaison

# Discrétisation basée sur les seuils naturels - JENKS

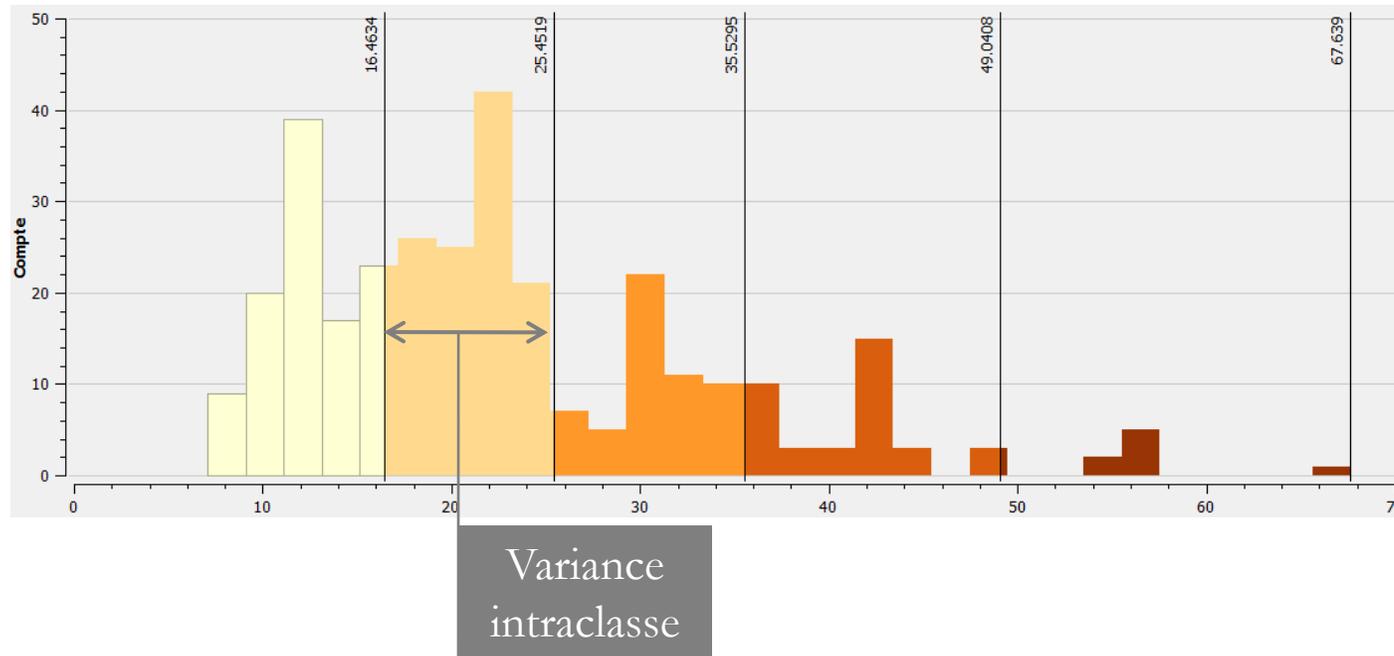
## Méthode de Jenks

Procédure :

Mode  Ruptures naturelles (Jenks) ▼

Méthode :

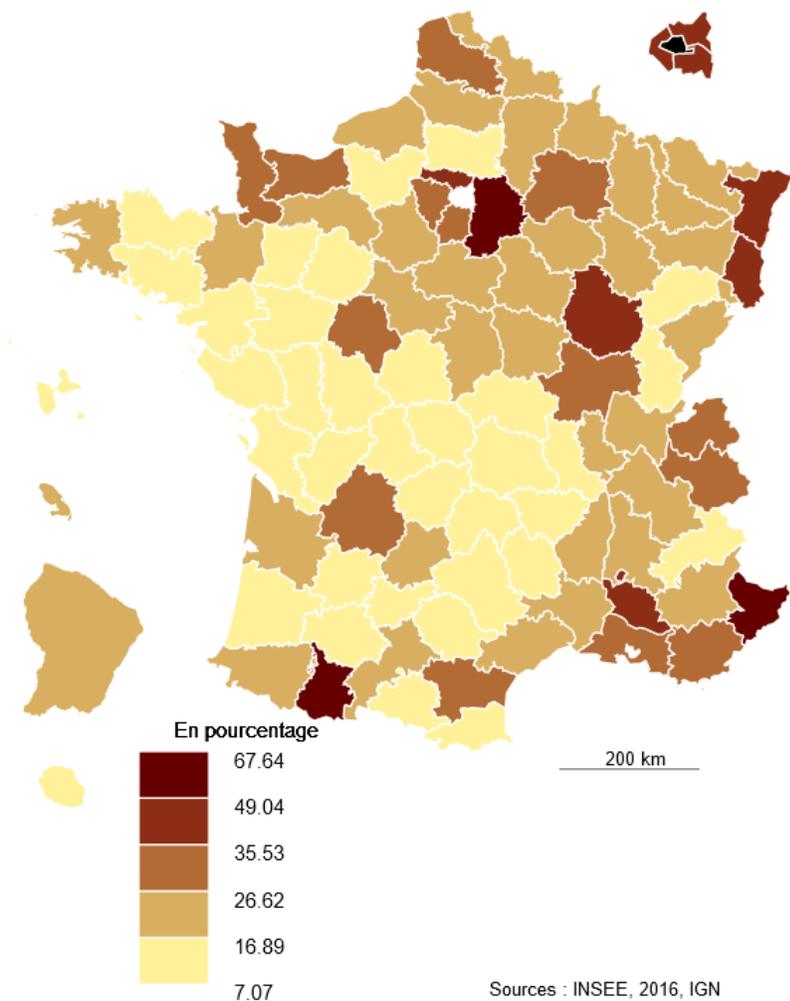
- Calcul informatique
- Recherche à réduire la variance intraclasse et à maximiser la différence entre les moyennes des classes



# DISCRÉTISATION BASÉE SUR LES SEUILS NATURELS - JENKS

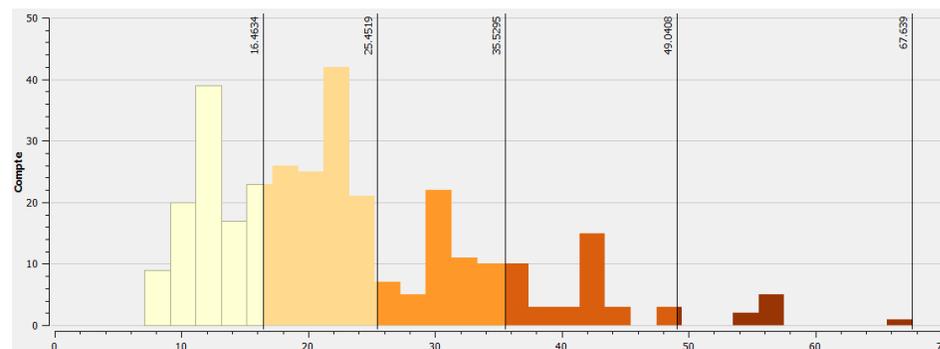
## Résultat cartographique pour Jenks

Part des nuitées de résidents étrangers en 2015

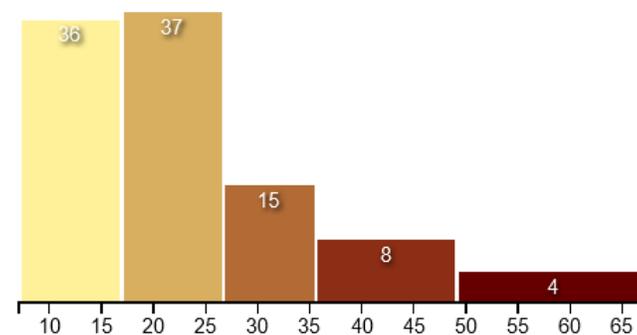


Sources : INSEE, 2016, IGN  
Réalisation : M@ppemonde - 12/2016

Tableau de dénombrement



Discrétisation : Jenks



Quantiles   Équivalences   Prog. arithmétique   Prog. géométrique  
Moyennes   Standard   Jenks

Nbre. de classes

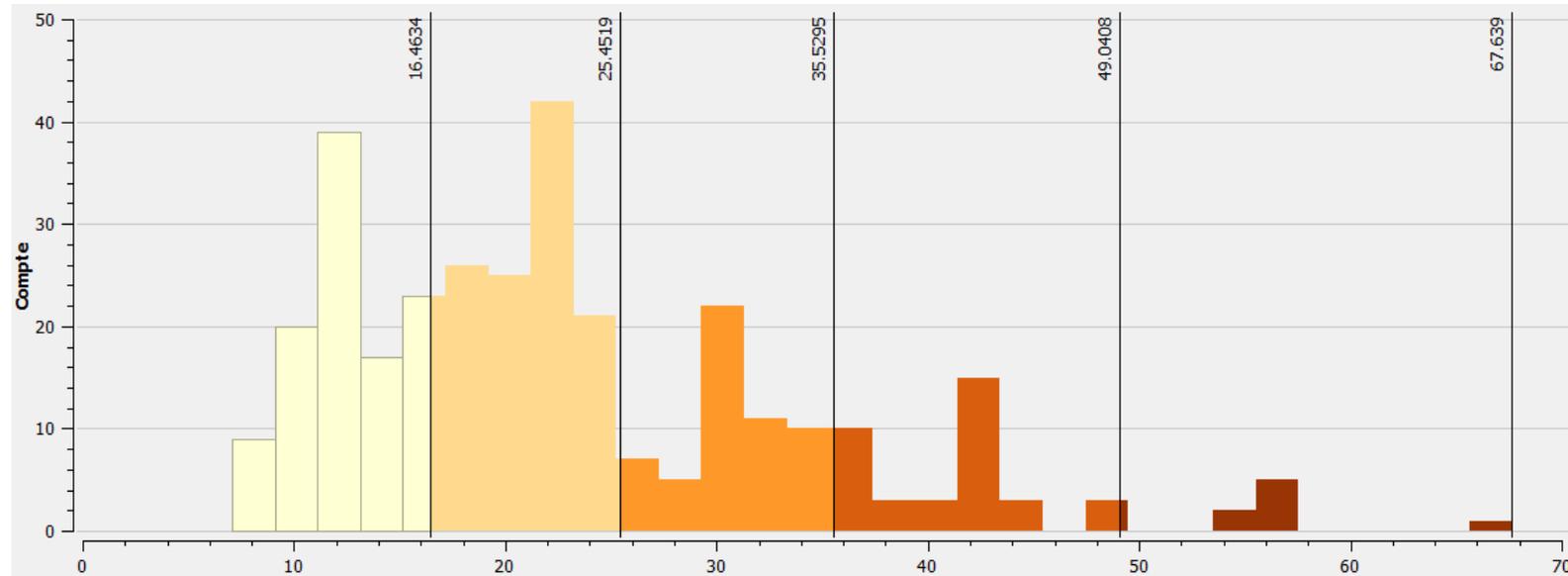
Nbre. de barres de fréquences

*Nous n'avons pas assez de temps pour regarder les différents modes de discrétisation*

Voir les diapos suivantes pour d'autres méthodes de discrétisation et l'ouvrage de Zanin et Lambert (référence donnée au début du cours)

# DISCRÉTISATION BASÉE SUR LES SEUILS NATURELS - JENKS

À retenir



+ Fait référence aux caractéristiques de la distribution (ruptures visibles avec un pas fin)

MAIS

- Convient mal aux distributions homogènes (rares)
- Ne convient pas aux comparaisons

# DISCRÉTISATION BASÉE SUR LES AMPLITUDES ÉGALES

Procédure :

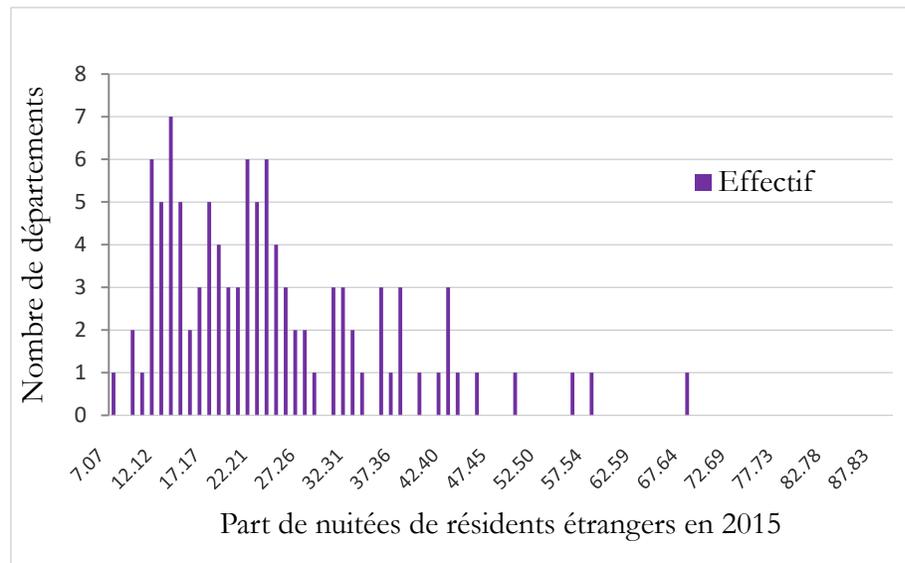
Mode

Méthode :

Premier calcul : l'amplitude chaque classe

Amplitude ou pas = (étendue) / (nb de classes)

Étendu = Max - Min



Min	7,07
Max	67,64
Taille pop.	100
<b>Calcul de l'amplitude</b>	
Étendu	
Nb de classe	
Amplitude	

# DISCRÉTISATION BASÉE SUR LES AMPLITUDES ÉGALES

*Deuxième calcul:* les limites de classes

- On détermine les limites des 5 classes en ajoutant la valeur du pas (1,48) au minimum (26), puis à la valeur ainsi trouvée (27,48), jusqu'à atteindre le maximum (33,4).

Étape 1 : Calcul de l'amplitude

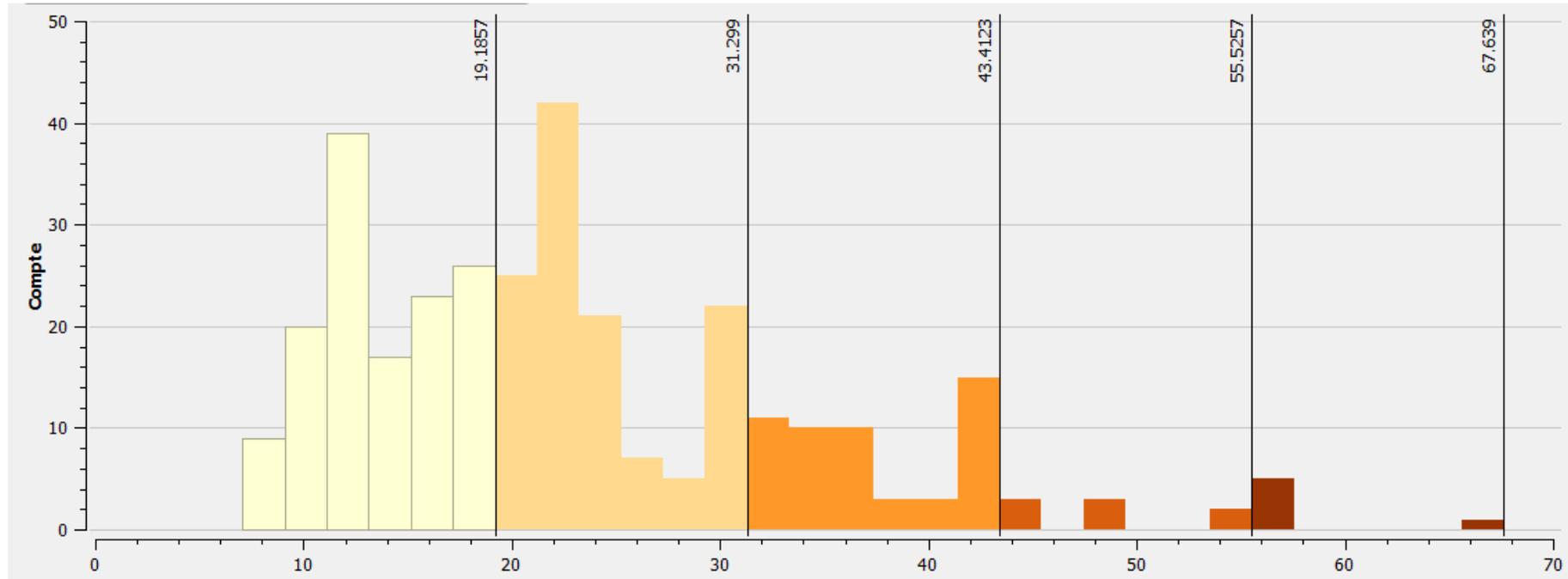
Min	7,07
Max	67,64
<b>Calcul de l'amplitude</b>	
Étendu	$67,64 - 7,07 = 60,57$
Nb de classe	5
Amplitude	$60,57 / 5 = 12,114$

Étape 2 : Calcul des limites des classe

Min	7,07	7,07
Min + Amplitude	$7,07 + 12,114$	19,184
Min + 2 Amplitude	$7,07 + (2 * 12,114)$	31,298
Min + 3 Amplitude	$7,07 + (3 * 12,114)$	43,412
Min + 4 Amplitude	$7,07 + (4 * 12,114)$	55,526

# DISCRÉTISATION BASÉE SUR LES AMPLITUDES ÉGALES

Résultat de la discrétisation sur le tableau de dénombrement



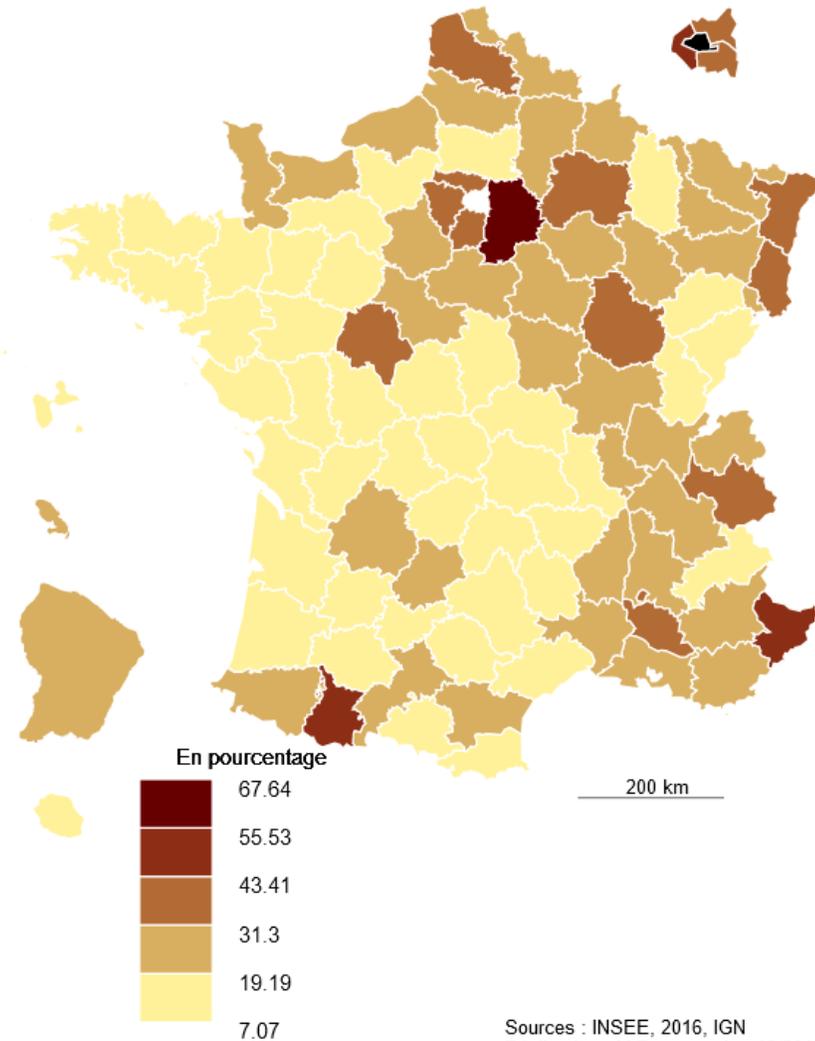
Classes calculées

Classe 1	[7,07 ; 19,184[
Classe 2	[19,184 ; 31,298[
Classe 3	[31,298 ; 43,412 [
Classe 4	[43,412 ; 55,526[
Classe 5	[55,526 ; 67,64[

# DISCRÉTISATION BASÉE SUR LES AMPLITUDES ÉGALES

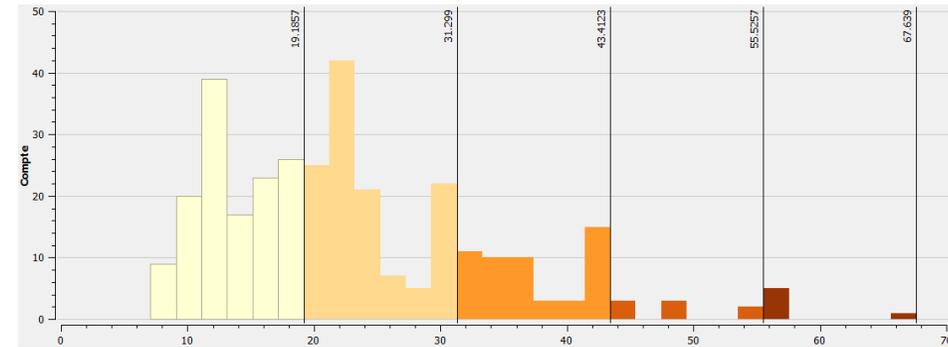
## Résultat cartographique

Part des nuitées de résidents étrangers en 2015

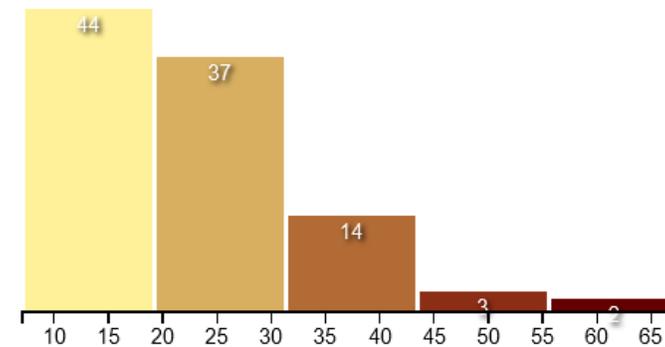


Sources : INSEE, 2016, IGN  
Réalisation : M@ppemonde - 12/2016

Tableau de dénombrement



Discrétisation : equivalences



Quantiles Équivalences Prog. arithmétique Prog. géométrique

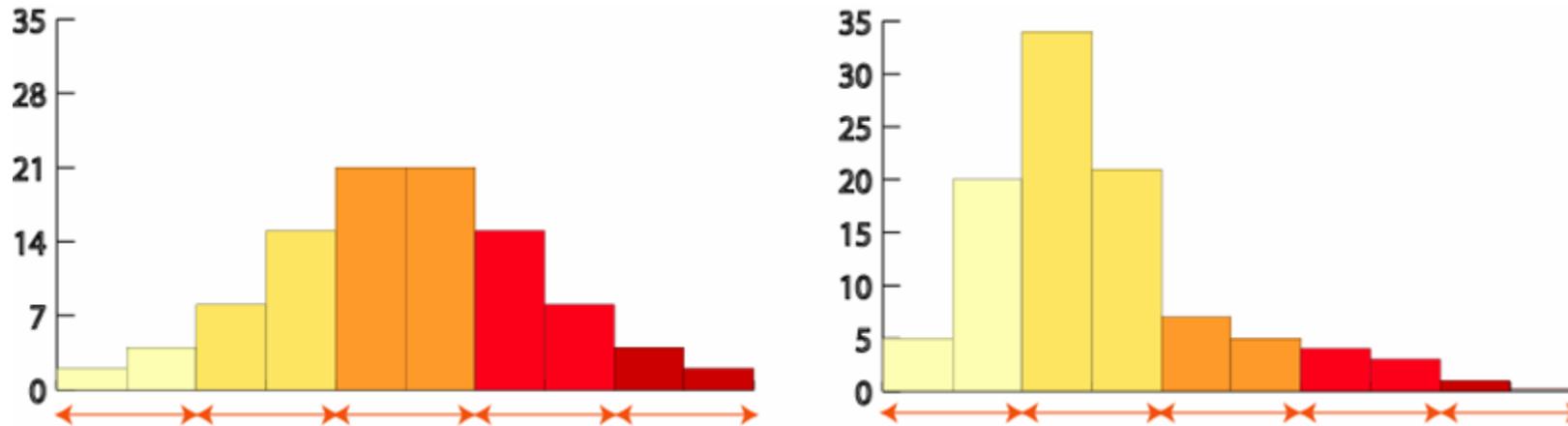
Moyennes Standard Jenks

Nbre. de classes

Nbre. de barres de fréquences

# DISCRÉTISATION BASÉE SUR LES AMPLITUDES ÉGALES

À retenir



+ Facile à construire, facile à lire, conserve la forme de la distribution

## MAIS

- Est rarement l'optimum cartographique (seulement si la distribution est uniforme)
- Ne convient pas aux distributions dissymétriques
- Peut contenir des classes vides

# DISCRÉTISATION BASÉE SUR « JOLIE RUPTURES »

Procédure :

Mode  Jolies ruptures

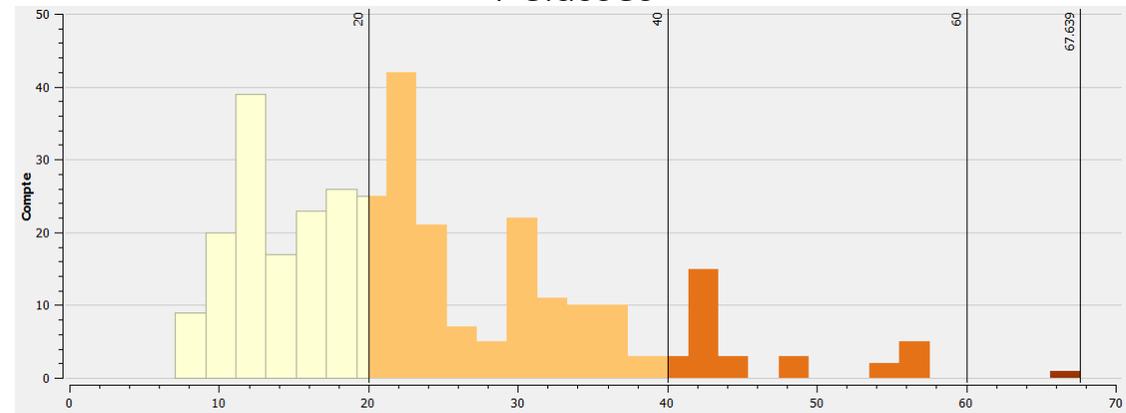
Méthode :

On choisit une discrétisation d'amplitudes égales pour faciliter la lecture de la carte

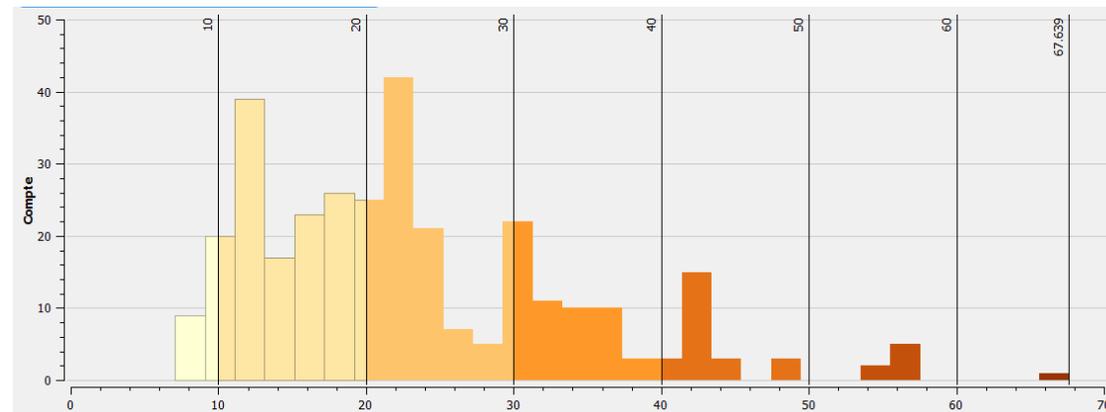
Ici, deux discrétisation possible :

- soit 4 classes de 10 en 10
- Soit 7 classe de 20 en 20

4 Classes



7 Classes

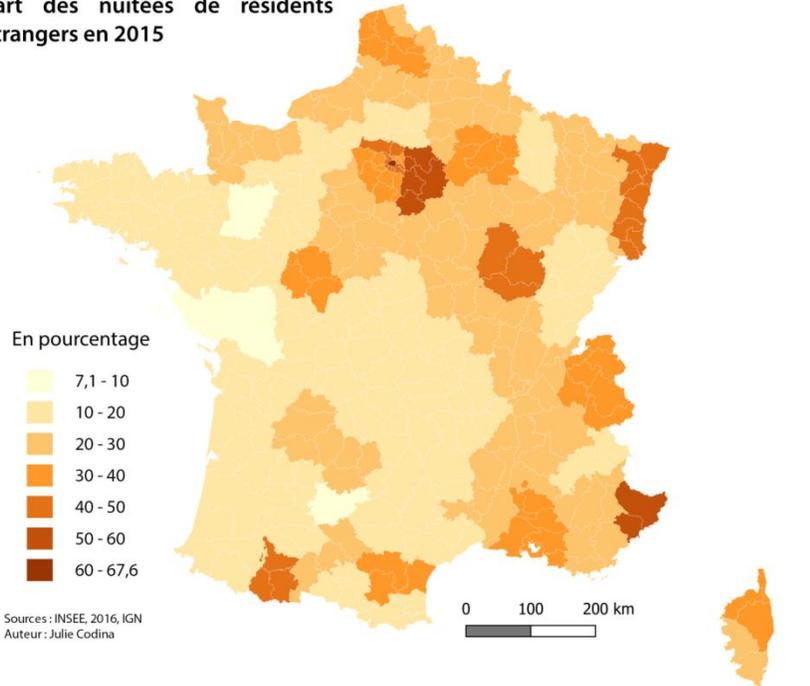


# DISCRÉTISATION BASÉE SUR « JOLIE RUPTURES »

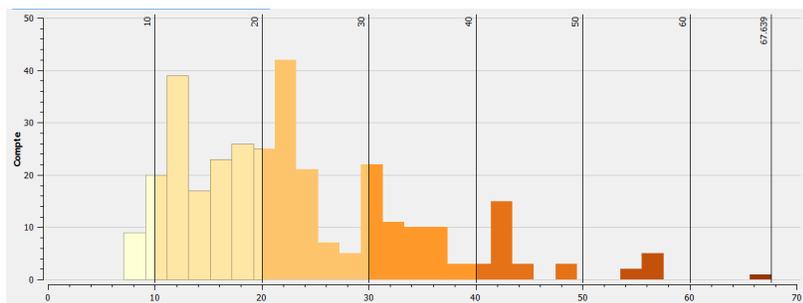
## Résultat cartographique

### 7 Classes

Part des nuitées de résidents étrangers en 2015

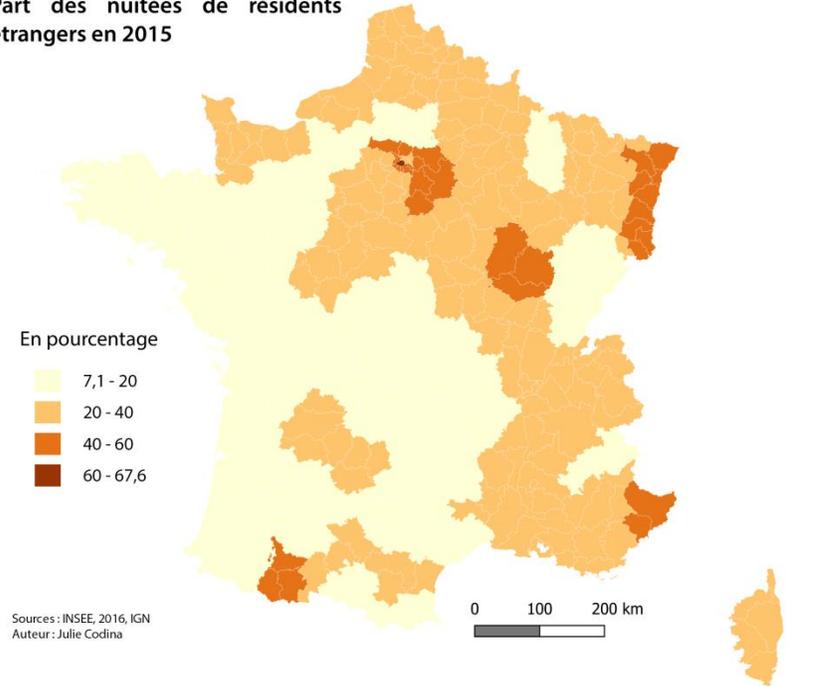


### Tableau de dénombrement

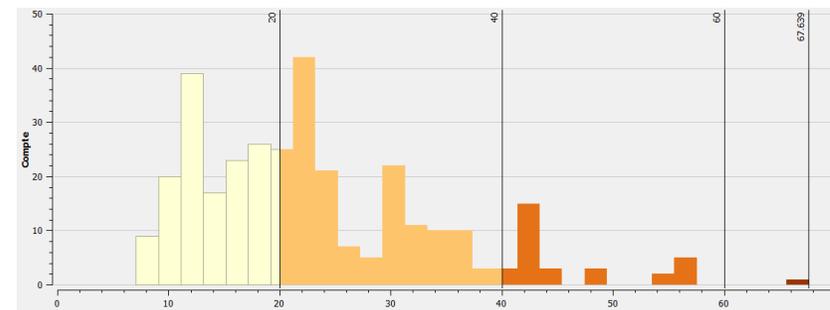


### 4 Classes

Part des nuitées de résidents étrangers en 2015

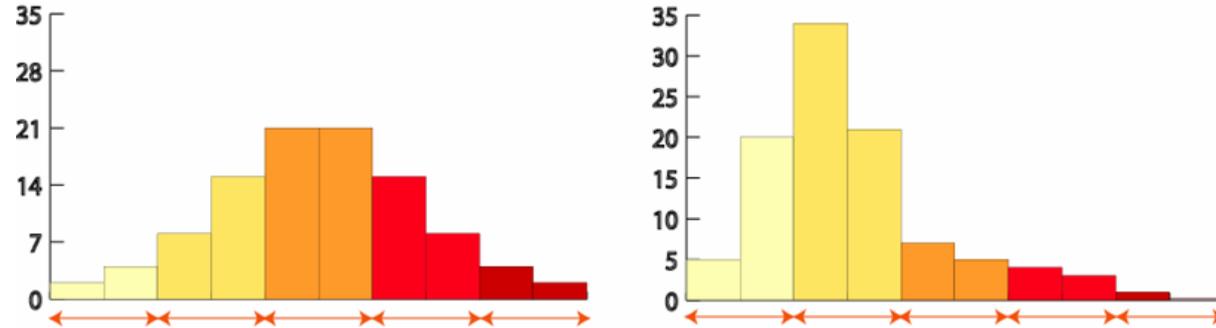


### Tableau de dénombrement



# DISCRÉTISATION BASÉE SUR « Jolie Ruptures »

## Résultat cartographique



+ Facile à construire, facile à lire (grand public, presse), conserve la forme de la distribution

## MAIS

- Est rarement l'optimum cartographique (seulement si la distribution est uniforme)
- Ne convient pas aux distributions dissymétriques
- Peut contenir des classes vides
- Ne convient pas dans une démarche scientifique car simplifie trop les données

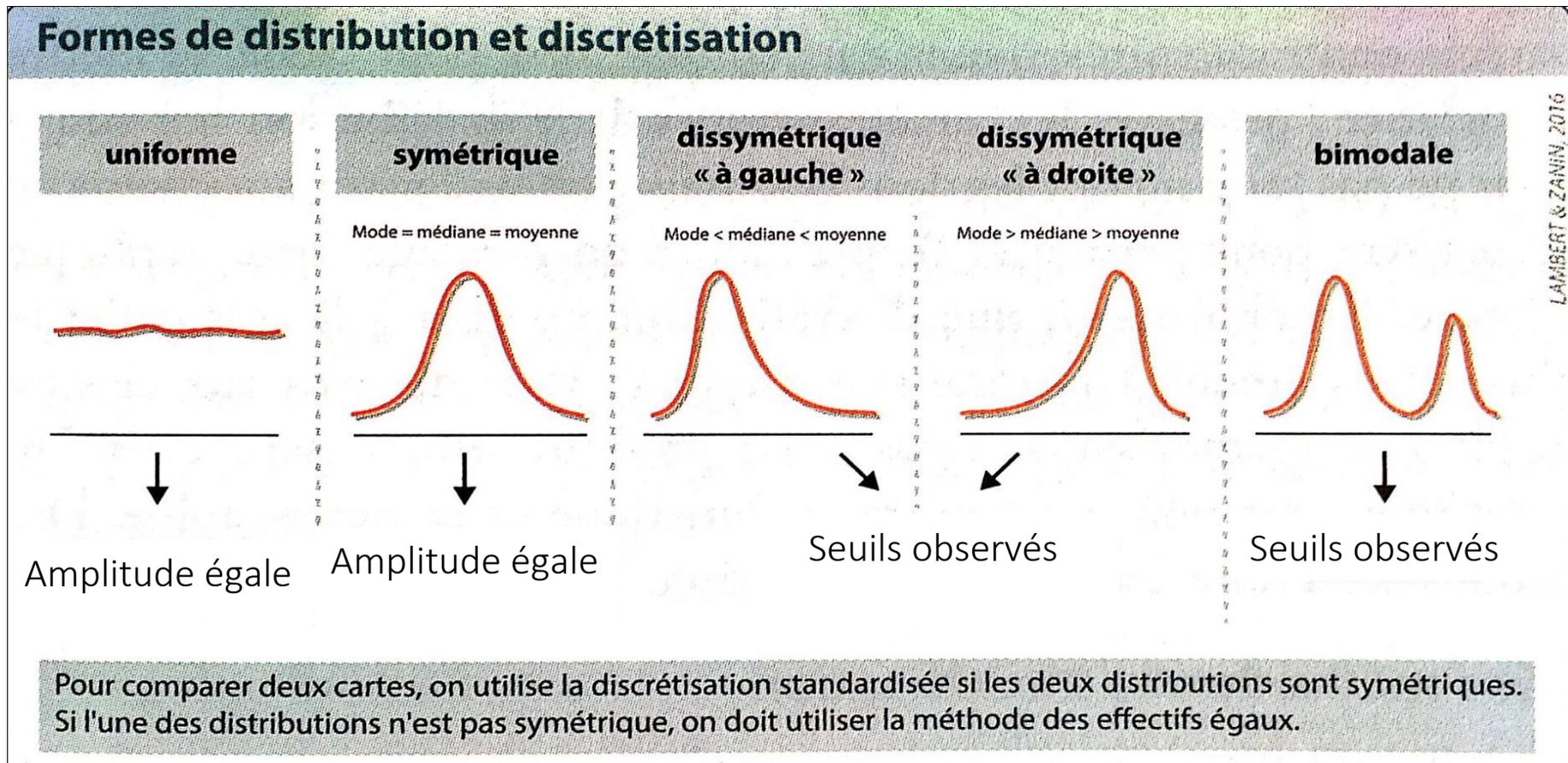
# CONCLUSION

À retenir  
Exemple de cartes

# COMMENT CHOISIR UNE DISCRÉTISATION?

La discrétisation dépend principalement :

- De la forme de la distribution
- De la situation de la carte (carte unique ou ensemble de cartes à comparer)
- .... Sans oublier le public visé



## À RETENIR :

Il n'y a pas de discrétisation optimale unique.

Il y a une démarche *cohérente par rapport* à une (ou plusieurs) *distribution statistique* donnée et un *objectif* donné.

Il faut donc sans cesse **faire et justifier des choix entre deux impératifs apparemment contradictoires** :

- un impératif de **vérité scientifique** (dont l'optimum est un **grand nombre de classes**, et une **discrétisation fidèle** à la forme de la distribution)
- Un impératif de **lisibilité cartographique** en fonction d'un **public** (dont l'optimum est un faible nombre de classes avec nombre d'UG par classe proche et des **amplitudes égales**)