

Systemes d'Information Géographique

<https://go.epfl.ch/sig>

Cartographie thématique et sémiologie graphique

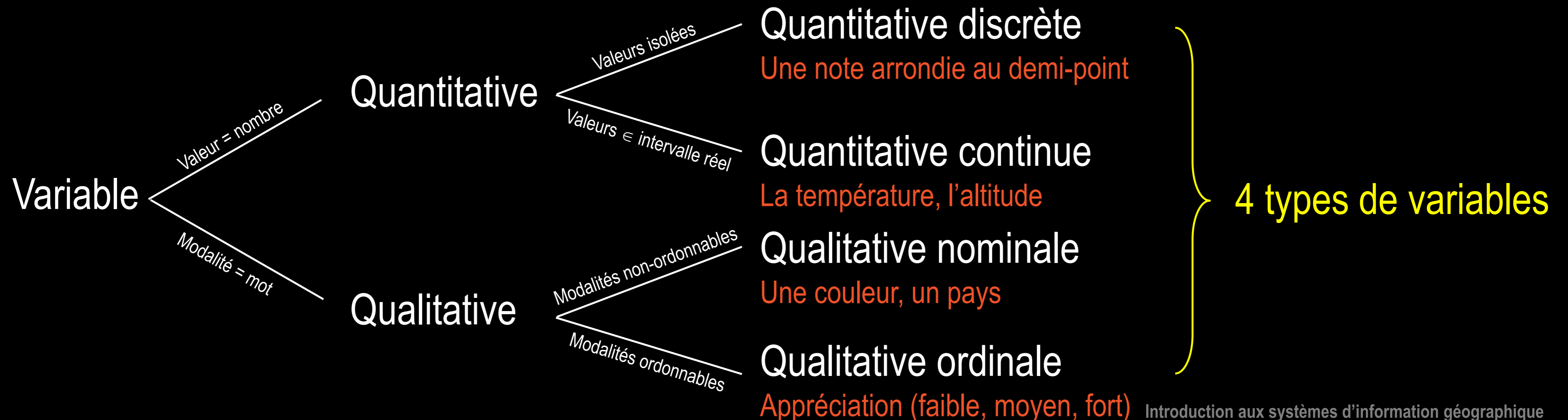
Attributs thématiques et mise en classes

Stéphane Joost, Gabriel Kathari (GEOME-LGB)

Types de données

- Variables **quantitatives**
- Grandeurs numériques
- Ou **discrètes**: nombre fini de valeurs
- Ou **continues**: nombre infini de valeurs

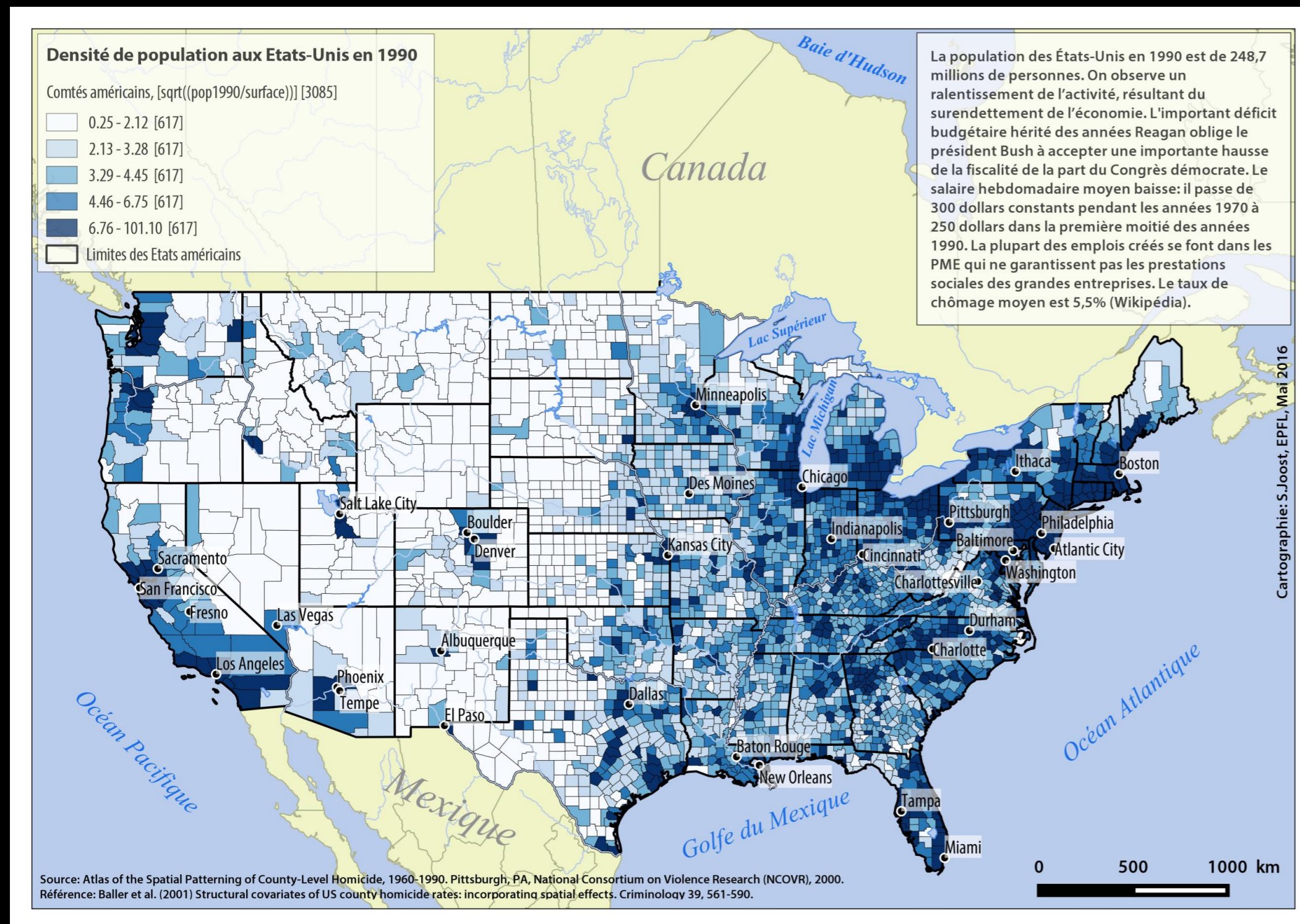
- Variables **qualitatives**
- Ses modalités sont des mots
- **Nominales**: non-ordonnables
- **Ordinales**: ordonnables





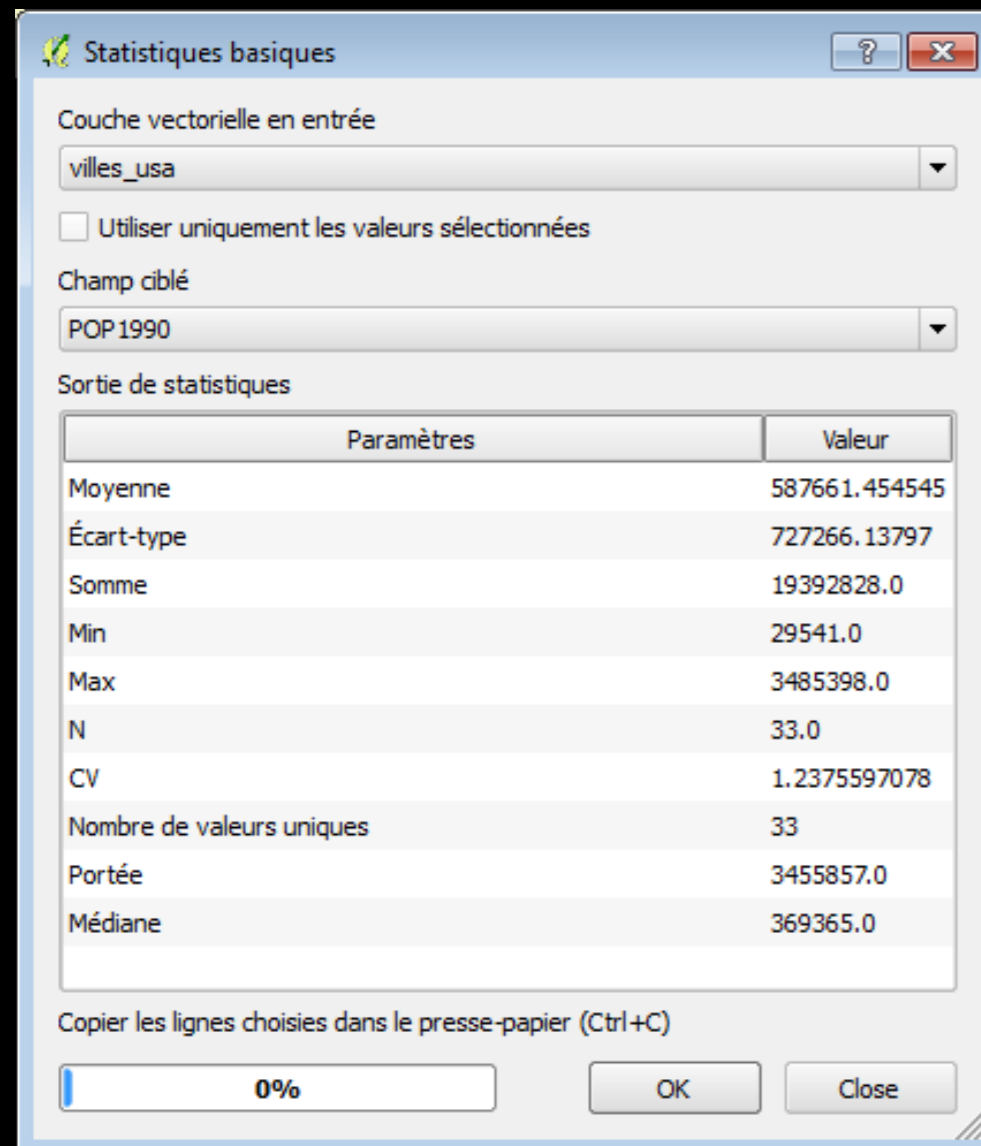
Comment les cartographier ?

Données quantitatives brutes et relatives – règle

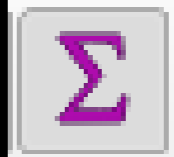


- Variable quantitative **absolue**, brute
- Variable visuelle = taille
- Exemple: nombre d'habitants
- Symboles (cercles) proportionnels
- Variable thématique **relative**
- Valable en tout point des surfaces représentées
- Variables visuelles = valeur et/ou couleur
- Exemple: densité de population

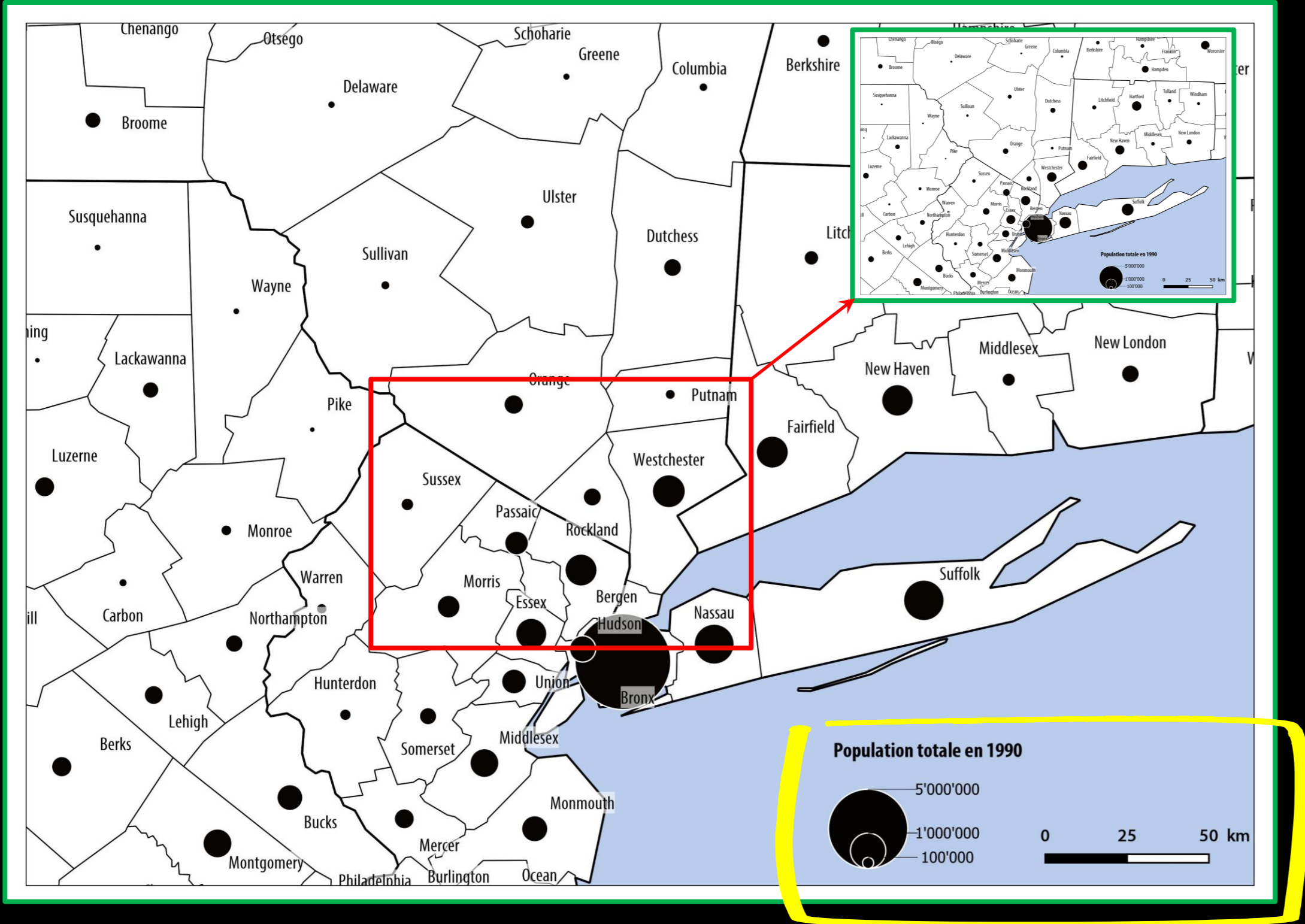
Cartes en symboles proportionnels



Statistic	Value
Count	81
Sum	4,11769e+06
Mean	50 835,7
Median	49 278
St dev (pop)	10 675,1
St dev (sample)	10 741,6
Minimum	31 306
Maximum	98 586
Range	67 280
Minority	31 306
Majority	31 306
Variety	81
Q1	43 464
Q3	55 686
IQR	12 222
Missing (null) values	0

- Relation mathématique entre valeurs et surface des symboles
- Connaître la distribution statistique
- Maximum, minimum, **étendue**
- Dans QGIS, «Statistiques basiques» ou «Statistiques» 
- **Dispersion** faible, cercles de même taille 😞
- Dispersion forte, très petits versus très gros cercles 😞
- Cas extrêmes: transformer le ratio

Symboles proportionnels – contrainte spatiale



Symboles proportionnels – le ratio valeur-surface

- Variable peu dispersée

id	var1	var2	var3	var4	var5	Var6	var7
1	23	bleu	5.5	11.5	1	ouv	pet
2	21	vert	9.7	12.2	2	ferm	grd
3	16	vert	2.3	11.0	4	ouv	moy
4	24	gris	5.9	10.8	2	ferm	grd
5	18	rose	6.2	9.9	1	ouv	moy
6	22	gris	2.3	9.8	3	ferm	grd
7	17	vert	4.4	11.5	3	ferm	grd
8	24	bleu	4.8	10.4	4	ouv	grd
9	23	gris	5.2	11.3	1	ouv	moy
10	20	rose	5.6	10.8	2	ouv	grd
11	19	vert	4.9	11.2	4	ferm	grd
12	24	bleu	1.5	10.4	3	ouv	grd

- Identifier la valeur maximum = 24
- Un rayon de 8mm = ~200 pixels²
- $S = \pi \cdot r^2$ et $r = \sqrt{\frac{S}{\pi}}$
- Ratio de 24 à 200 (S) et règle de 3 pour calculer le rayon, soit r=8;
- 23 → S=190 et r=7.7;
- 21 → S=174 et r=7.4;
- 20 → S=166 et r=7.2, etc.

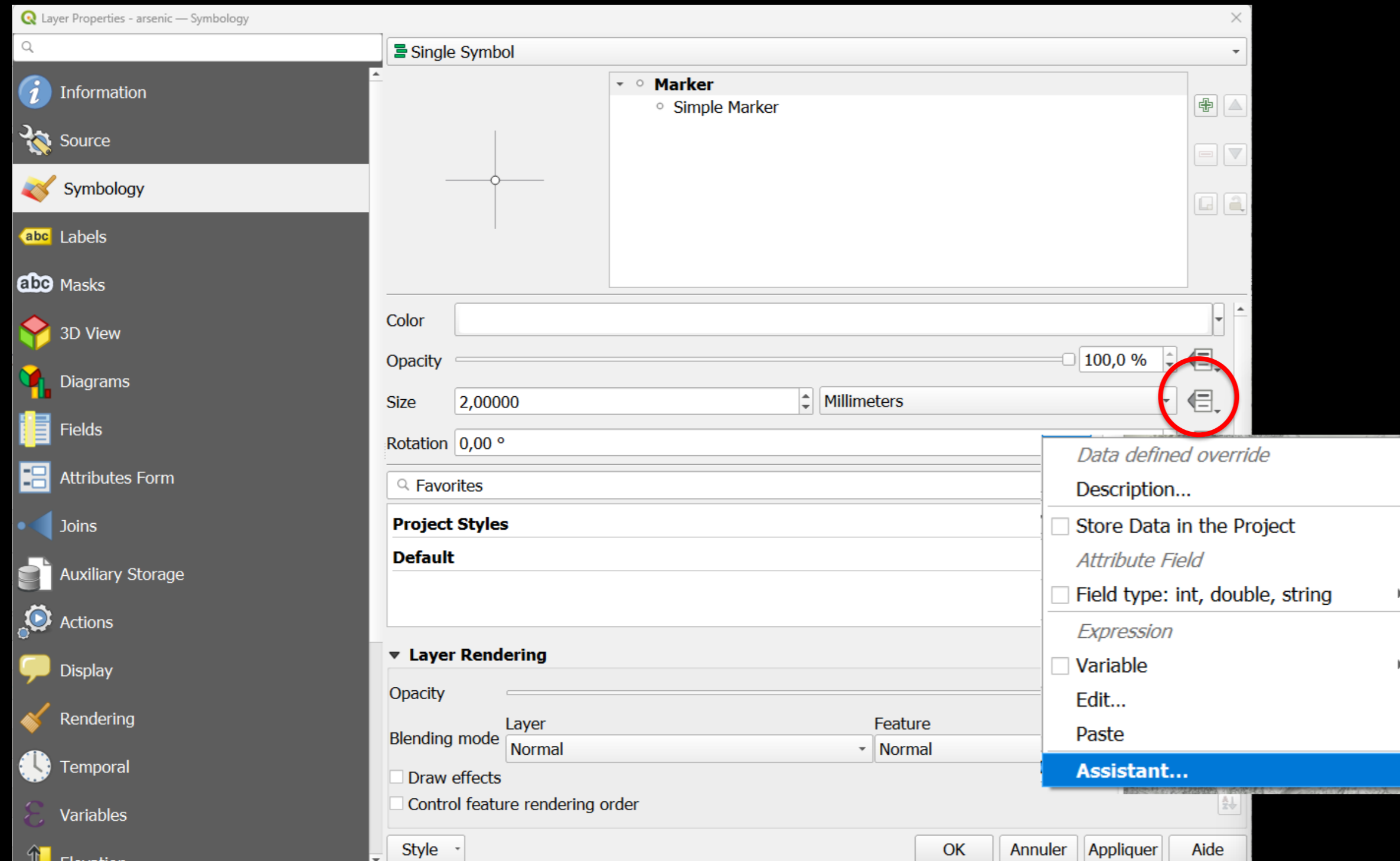
Symboles proportionnels – le ratio valeur-surface (2)

- Variable très dispersée

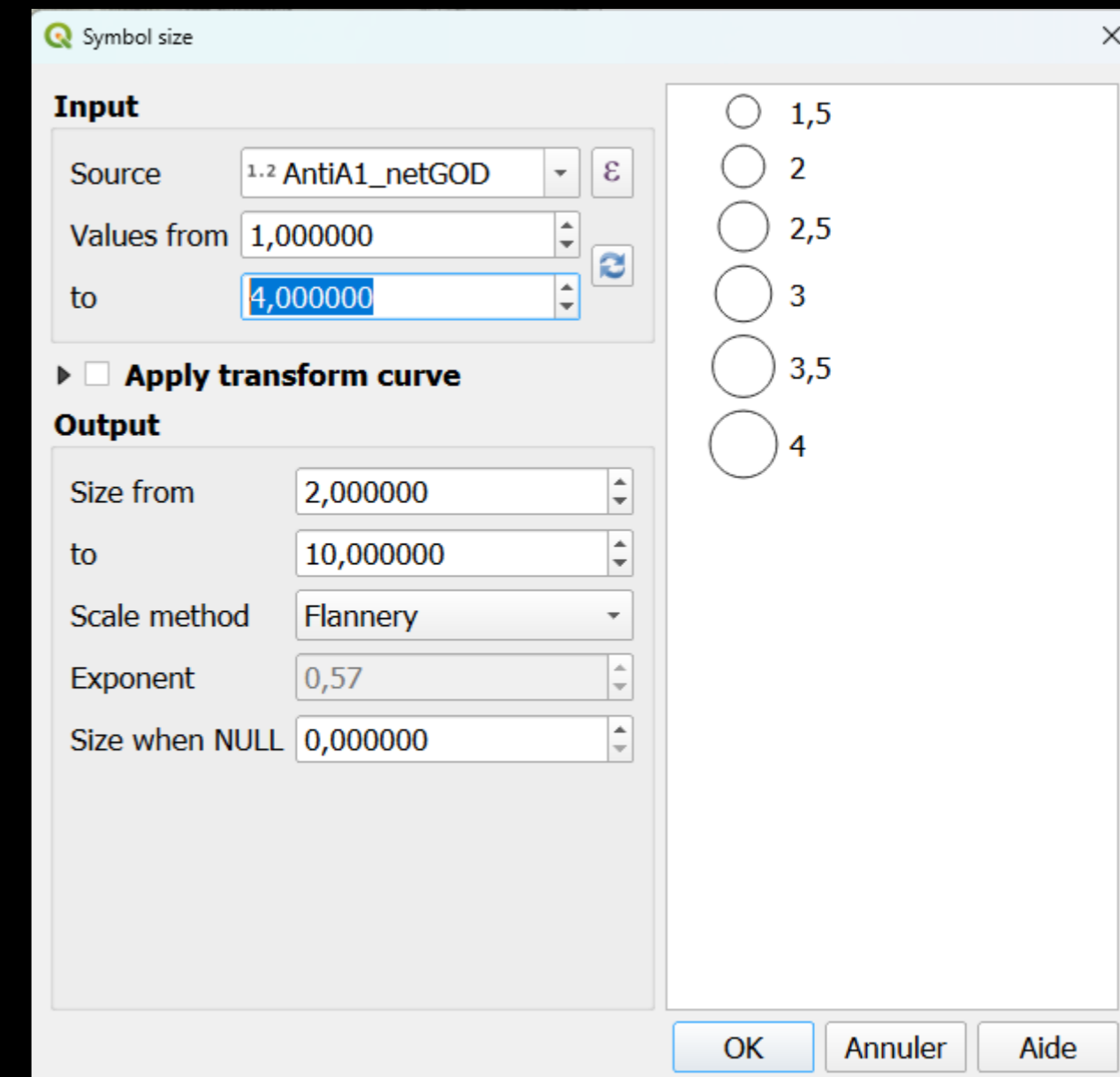
id	var1	var2	var3	var4	var5	Var6	var7
1	23	bleu	5.5	11.5	1	ouv	pet
2	21	vert	9.7	12.2	2	ferm	grd
3	16	vert	2.3	11.0	4	ouv	moy
4	24	gris	5.9	10.8	2	ferm	grd
5	18	rose	6.2	9.9	1	ouv	moy
6	22	gris	2.3	9.8	3	ferm	grd
7	17	vert	4.4	11.5	3	ferm	grd
8	24	bleu	4.8	10.4	4	ouv	grd
9	23	gris	5.2	11.3	1	ouv	moy
10	20	rose	5.6	10.8	2	ouv	grd
11	19	vert	4.9	11.2	4	ferm	grd
12	24	bleu	1.5	10.4	3	ouv	grd

- Ratio basé sur valeur la plus forte inadapté
- Transformer le ratio
- Logarithme ou racine
- Perte du lien simple entre les valeurs et les surfaces des symboles
- Mentionner le type de transformation dans la légende

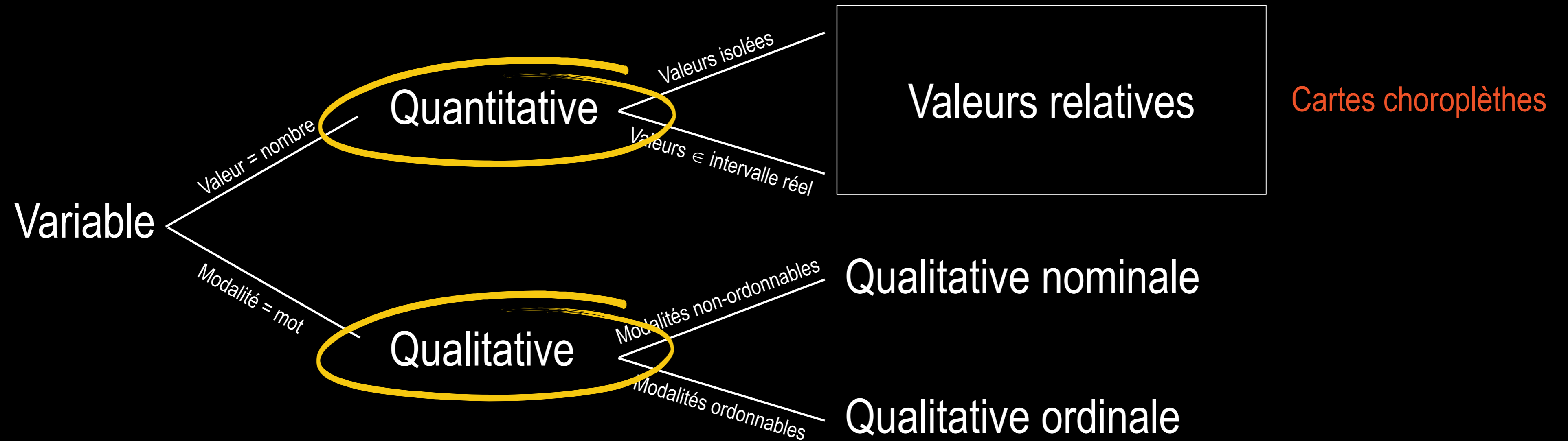
Symboles proportionnels dans QGIS



- Style
- Symbole unique
- Assistant taille



Cartes thématiques en plages de couleurs

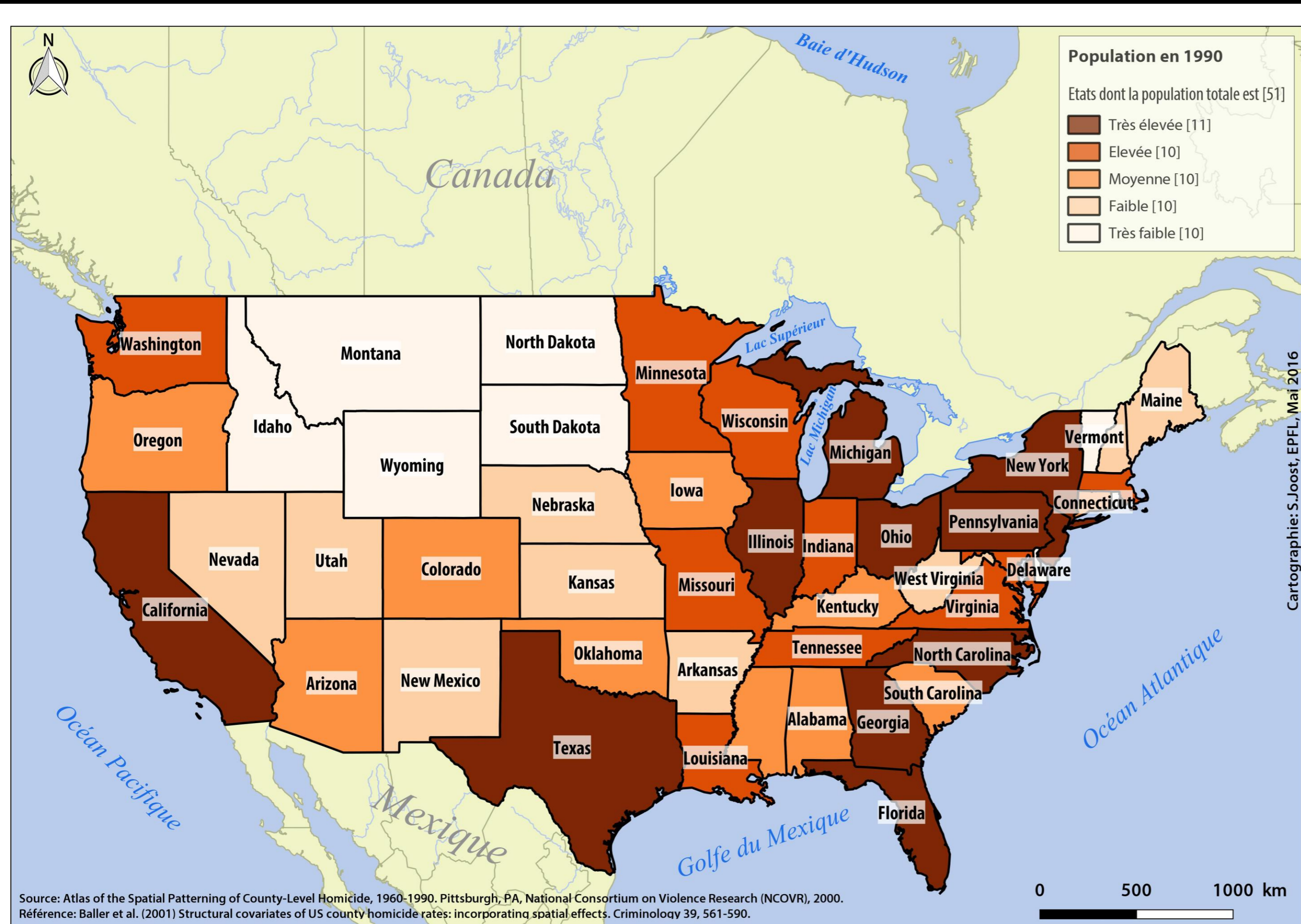


- Fabrication d'une carte en plages de couleurs
- Etape numéro 1: **la discrétisation**

Discrétisation ou mise en classes

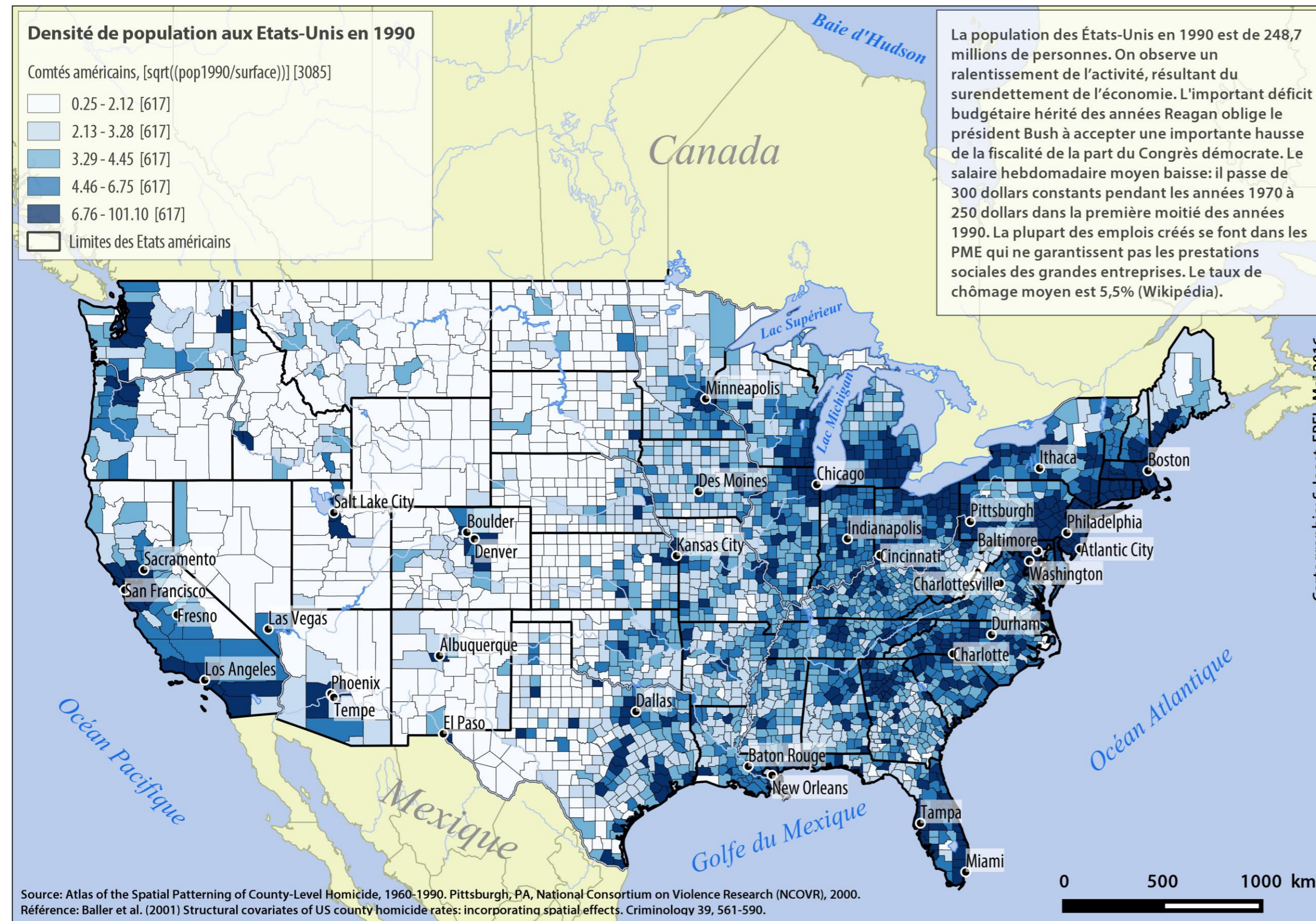
- Découper en classes
 - Simplifier en regroupant
 - Classes homogènes et distinctes
-
- Satisfaire les règles de la sémiologie graphique et de la statistique
 - Perdre le moins d'information possible
 - Applicable aux variables quantitatives (relatives), qualitatives nominales et ordinales

Discrétisation de variables qualitatives



- Nominale: généralisation de l'information en fonction de la hiérarchie ou d'un autre critère
- Maximiser la discrimination entre unités spatiales
- Ordinales: utilisation du rang

Cartes choroplèthes

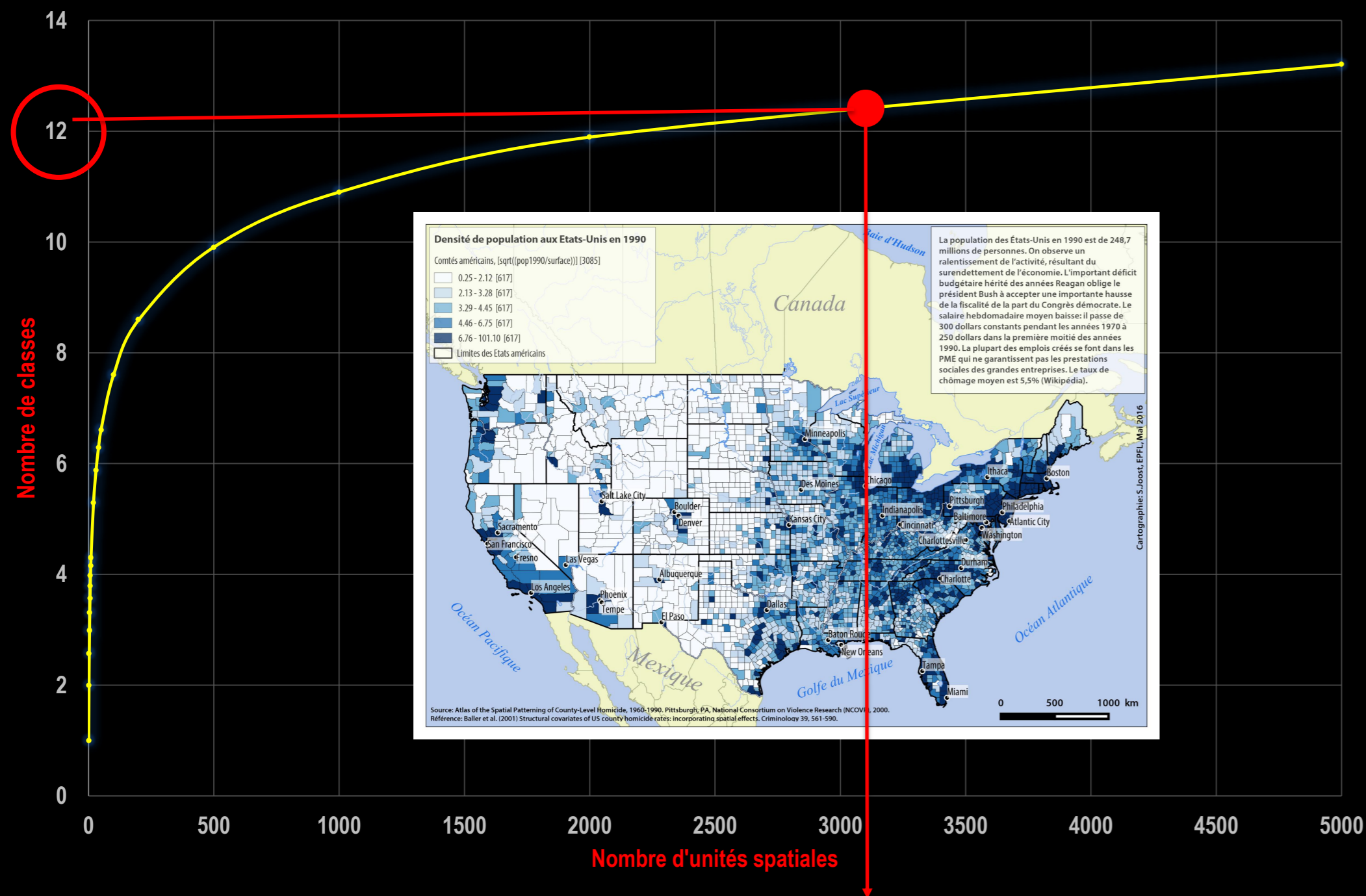


- **Plethos** = quantité
- **Khorê** = ~aire géographique
- Comptages, mesures
- Variables quantitatives relatives (ou transformées)

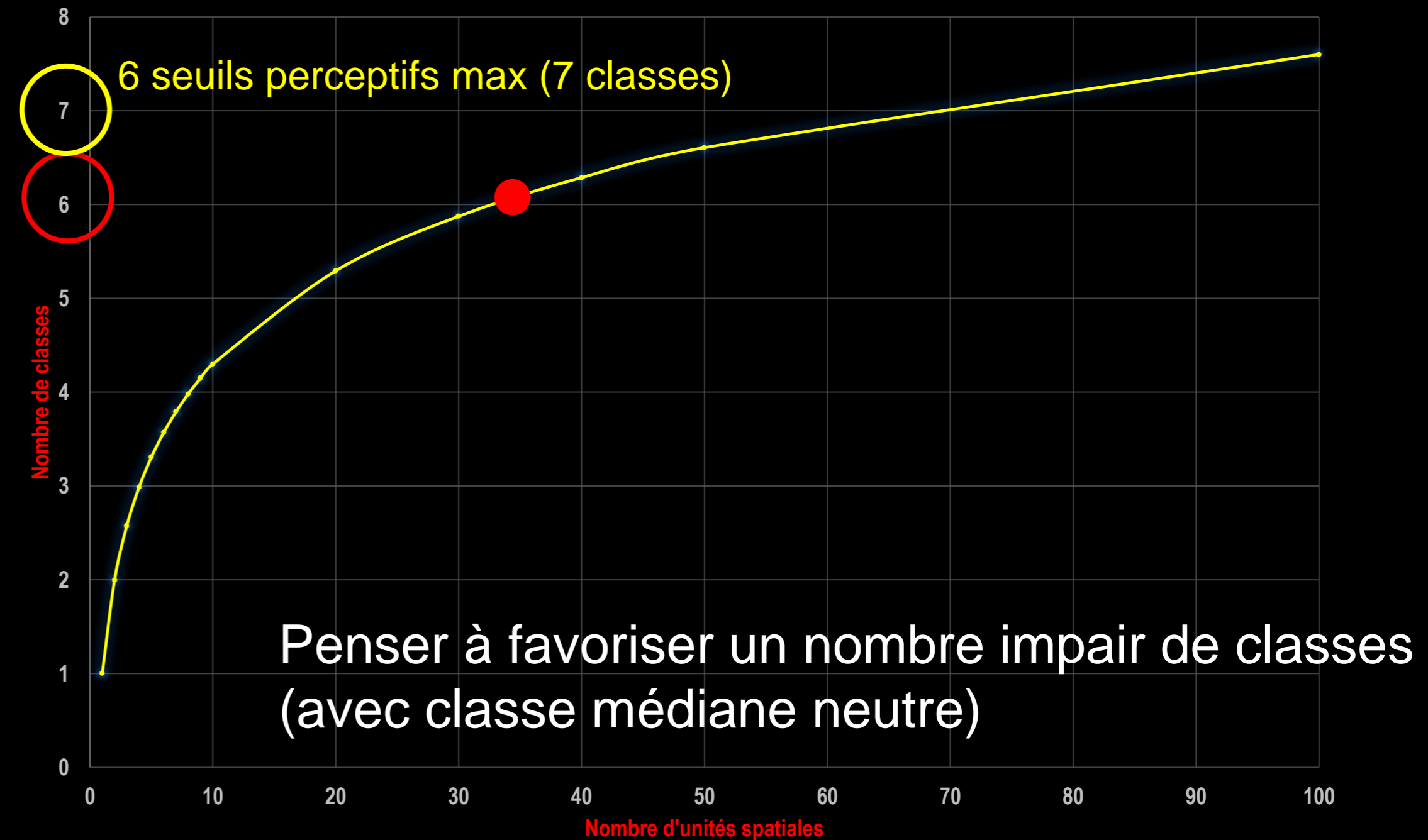
Discrétisation de variables quantitatives relatives

Comment trouver le nombre idéal de classes ? 5 ou 7 classes dans la plupart des cas

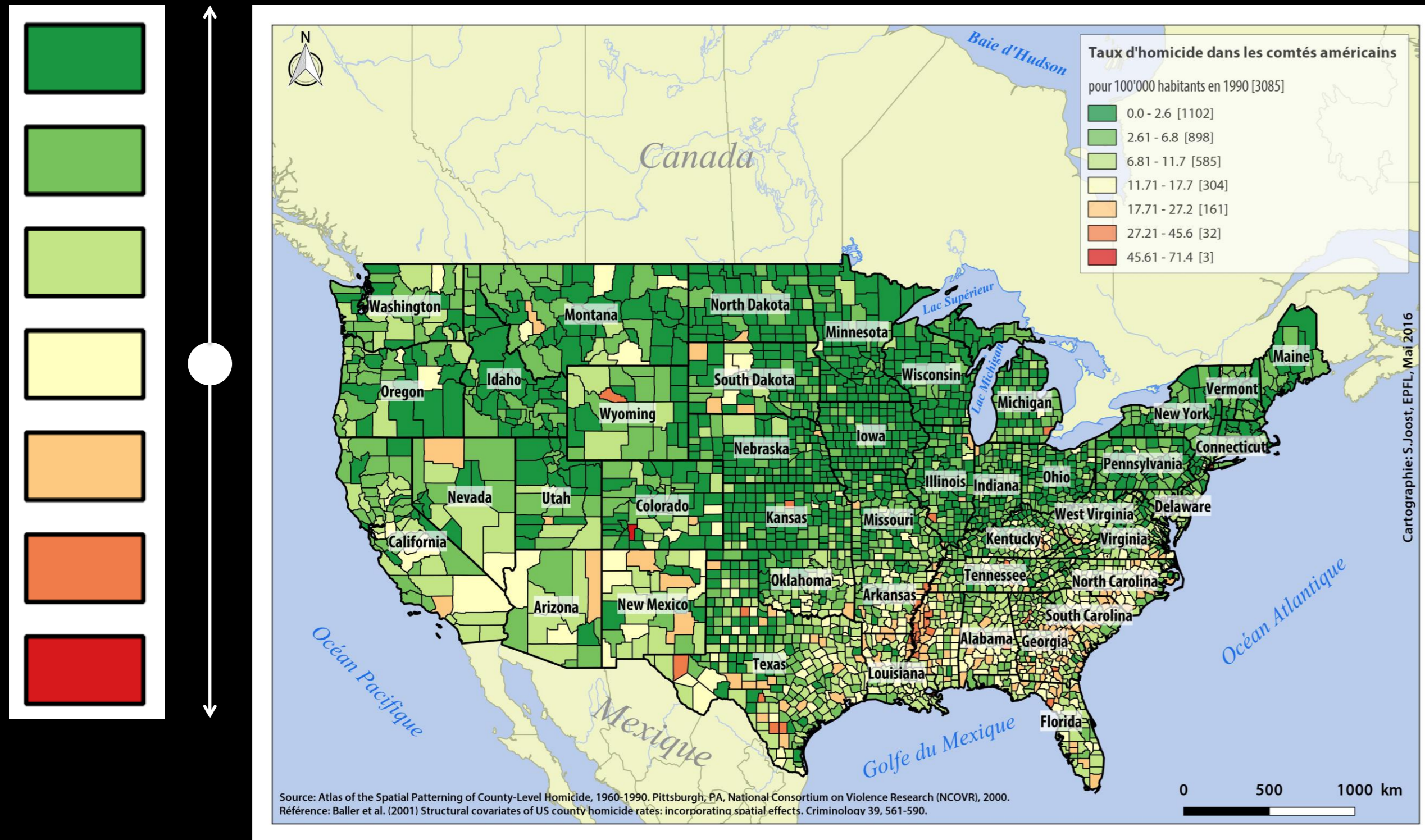
$$N_{cl} = 1 + 3,3 \log_{10}(N_{obs}) \quad (\text{Indice de Huntsberger})$$



12 classes ! Problème de perception

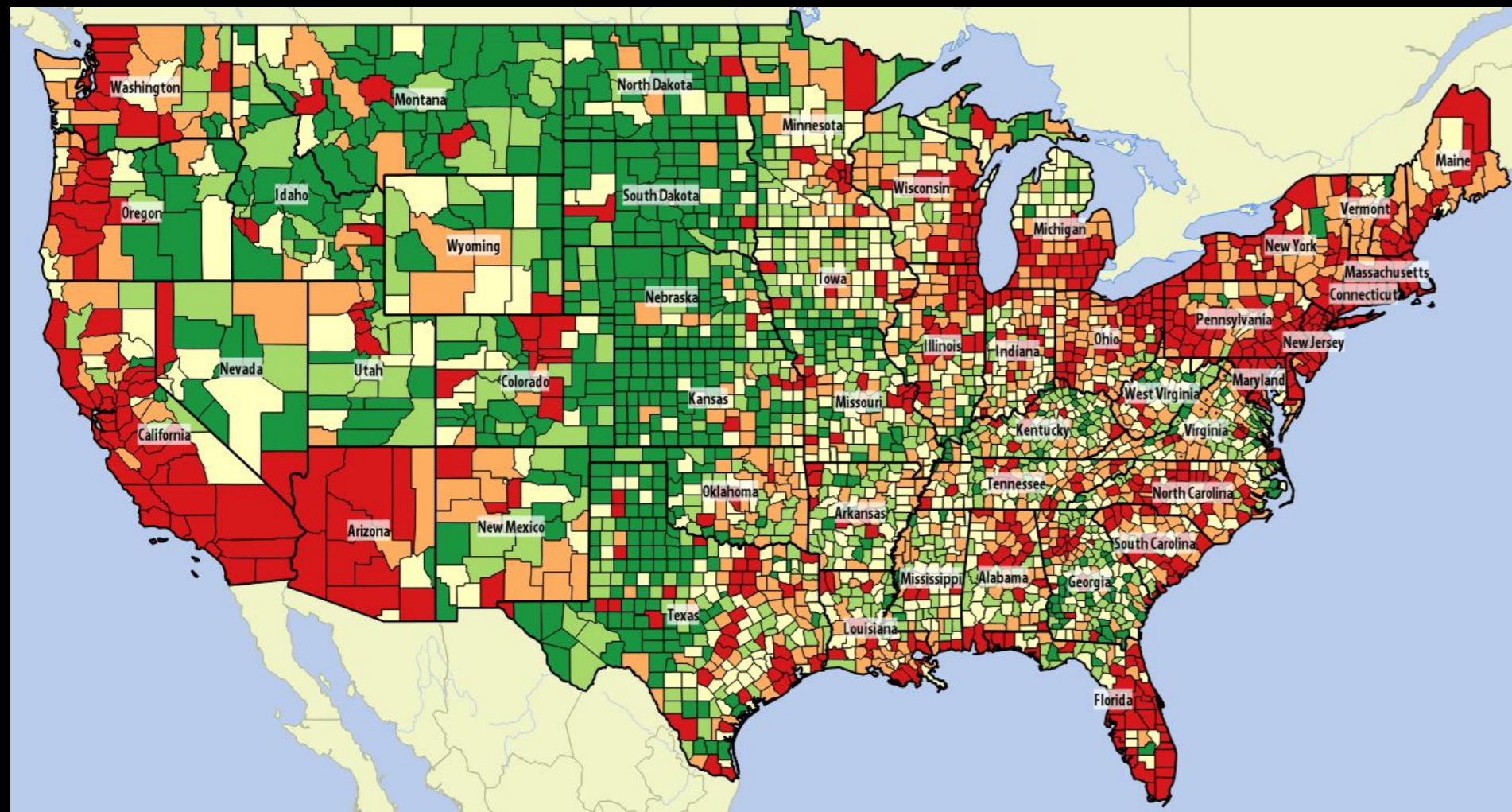
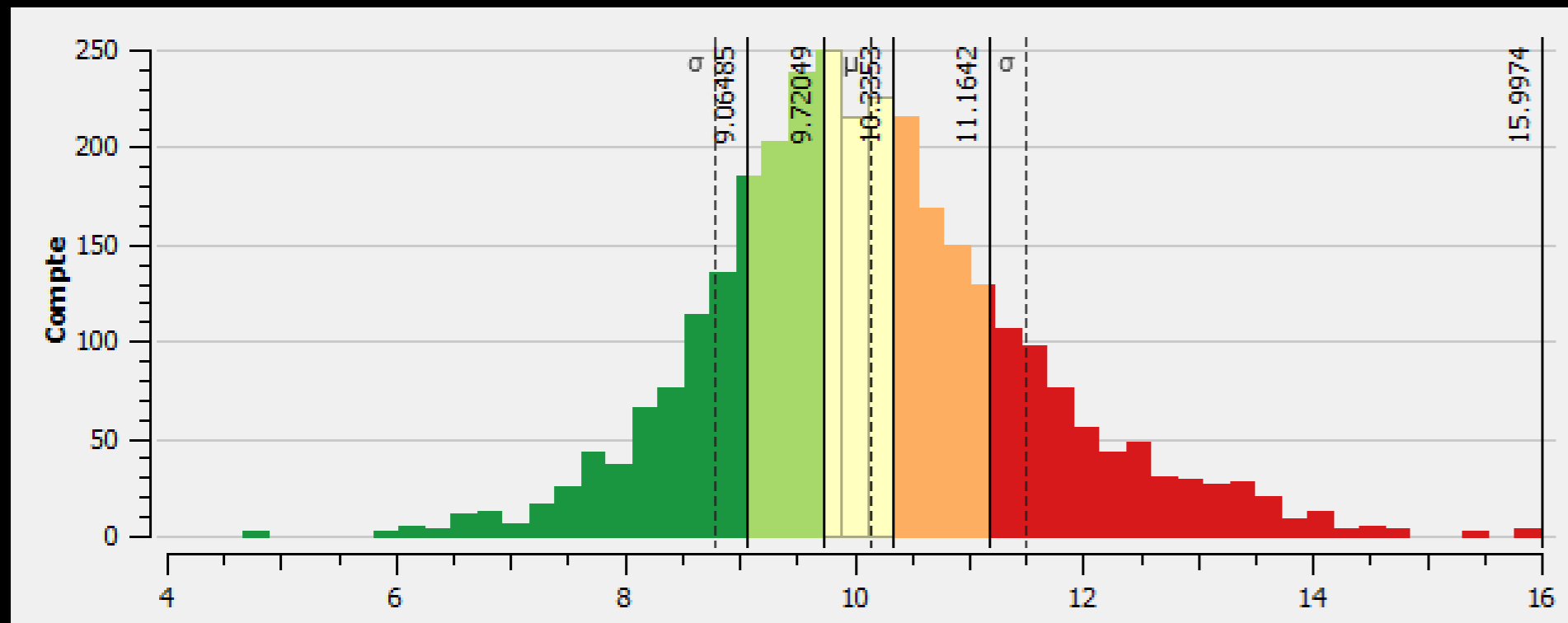


Nombre de classes – grand nombre d'unités spatiales



- Augmenter le nombre de seuils perceptifs
- Deux teintes
- Une classe intermédiaire
- Nombre impaire de classes
- Maximum théorique: 12-1 classes
- Choix des couleurs cohérent avec le thème traité
- Explications nécessaires

Choix d'une méthode de discrétisation



- Méthode pour répartir les données dans les classes?
- Variable qualitative nominale: pas de méthode spécifique, définir un critère commun de regroupement
- Variable qualitative ordinale: conserver la hiérarchie de la distribution
- Variables quantitatives, 3 types d'information à considérer:
 - Ordre de grandeur, valeurs centrales
 - Dispersion: écart-type, intervalle inter-quantiles
 - Forme de la distribution: normale, asymétrique, exponentielle, unimodale, bimodale, plurimodale, etc.

Méthodes de discrétisation (1)

Méthode	Définition	Calcul	Remarques	Type de distribution statistique
Ecart à la moyenne	Toutes les classes ont une même étendue égale à l'écart type, sauf les classes extrêmes	Moyenne et écart type	Si le nombre de classes est impair, la classe centrale est à cheval sur la moyenne. Intérêt: se repérer par rapport à la moyenne, mettre en évidence les extrêmes, comparer plusieurs cartes	Normale (Gauss) avec concentration des données autour de la moyenne, éventuellement peu asymétrique
Égale amplitude	Les intervalles de classes sont égaux (intervalles constants)	(Valeur maximale - valeur minimale) / nombre de classes	Méthode simple, facile à interpréter mais peu utilisée car ne convient pas si la distribution des valeurs est asymétrique : les classes peuvent être très inégales et certaines vides. Pas de comparaison possible	Distribution uniforme (données réparties uniformément sur toute l'amplitude de la série), normale (Gauss) avec concentration des données autour de la moyenne

Source: d'après «Rappels sur les discrétisations», F. Demoraes, M. Souris, T. Serrano, Laboratoire de Cartographie Appliquée, Paris X. Nanterre, & E. Ployon, Université de Savoie, Formation SIG-Santé 2010

Méthodes de discrétisation (2)

Méthode	Définition	Calcul	Remarques	Type de distribution statistique
Seuils naturels	Seuils observés	Observation de l'histogramme de fréquences cumulées triées dans l'ordre croissant	Prend en compte les discontinuités de la série Cartes non comparables	Toute distribution qui présente des discontinuités Distribution plurimodale
Quantiles	Chaque classe a le même nombre d'individus	Effectif total / nombre de classes Limites de classes= nb d'individus défini par le calcul ci-dessus, statistique d'ordre	Ne tient pas compte des valeurs exceptionnelles. Des limites de classes peuvent êtres discutables (valeurs très proches réparties dans des classes différentes) Permet les comparaisons	Distribution uniforme ou autre distribution sans discontinuités A éviter si présence de valeurs extrêmes ou si trop grand nombre de valeurs proches
Box map	Classes basées sur l'analyse en boxplot ou boîte à moustaches	6 classes: quartile 1 (Q1), Q2, Q3, Q4, outliers inférieurs ($Q1 - [1.5 * (Q3 - Q1)]$), outliers sups. ($Q1 + [1.5 * (Q3 - Q1)]$)	Statistique d'ordre, permet les comparaisons, implémentée dans Geoda	Toute distribution, adéquat pour traiter des valeurs exceptionnelles

Source: d'après «Rappels sur les discrétisations», F. Demoraes, M. Souris, T. Serrano, Laboratoire de Cartographie Appliquée, Paris X. Nanterre, & E. Ployon, Université de Savoie, Formation SIG-Santé 2010

Règles de base

- Les classes doivent couvrir l'ensemble de la distribution et doivent être contiguës
- Une valeur ne doit appartenir qu'à une seule classe
- Les classes ne peuvent pas être vides
- Le choix des valeurs-seuils doit reposer sur des critères robustes
- Des valeurs non-significativement différentes ne doivent pas être placées dans des classes distinctes



Habillage et production de la carte