



Méthodes de discrétisation

Hélène Mathian

*Ecole d'été « Statistiques, Cartographies
et Analyse spatiale »*

Yaoundé 2006



De l'information à la représentation cartographique

La collecte des données et leurs traitements constituent des phases très importantes.

Une carte est **une** représentation possible des données . Elle constitue **un** point de vue.

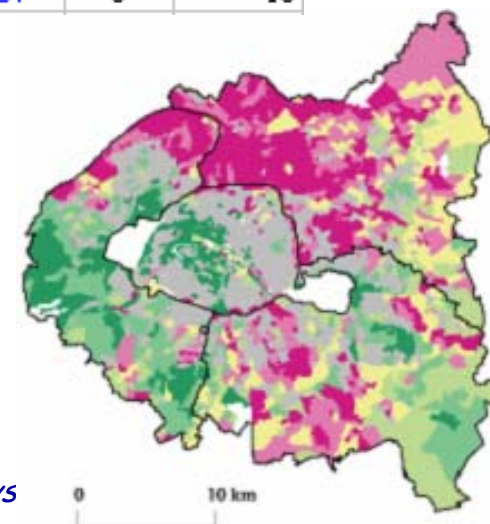
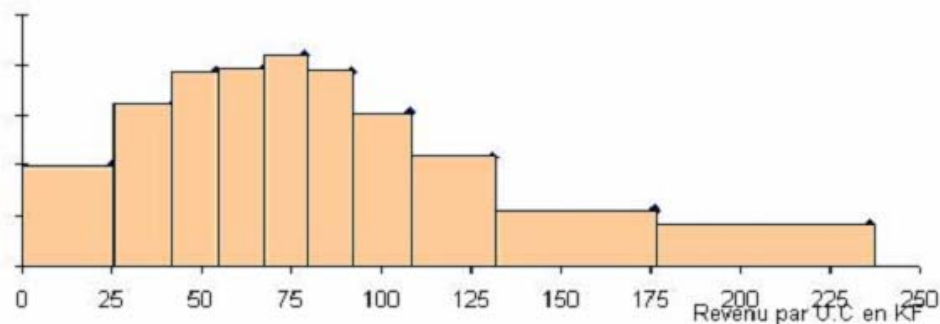
On peut avoir des données fausses déguisées en jolie carte... et des jolies données mal cartographiées.

- Pourquoi ? *Quel est le but ?*
- Quoi ? *Quelle description ? Quelles mesures ?*
- Comment et où ? *Quand et qui ?*

Un jeu ... de données

Distribution statistique et distribution spatiale

CTOT	NOM	PSDC99	densité	tx82 99	emp/act	%chô meurs	Revenu	Muni 2001	Type Com.	nb crèches
75101	Paris 1er Arr.	16888	7573	-0.54	6.12	10.52	167906	RPR	2	3
75102	Paris 2e Arr.	19585	20191	-0.47	4.83	13.34	109396	GP	5	2
75103	Paris 3e Arr.	34248	23458	-0.31	1.51	11.67	134496	GP	2	4
92002	ANTONY	59855	6139	0.54	0.68	7.92	111026	RPR	2	6
92004	ASNIERES-SUR-SEINE	75837	13739	0.38	0.54	11.62	96344	RPR	5	13
92007	BAGNEUX	37252	9130	-0.47	0.76	13.44	75792	PC	6	6
92009	BOIS-COLOMBES	23885	12571	0.03	0.40	9.93	103931	RPF	3	6
93015	COUBRON	4612	1083	0.42	0.26	7.40	103528	DVD	3	3
93027	LA COURNEUVE	35310	4209	0.30	0.89	24.34	54706	PC	7	7
93029	DRANCY	62263	8280	0.20	0.44	17.20	68974	UDF	6	9
93030	DUGNY	8641	1959	0.13	0.37	17.00	58843	RPR	7	5
93031	EPINAY-SUR-SEINE	46409	10971	-0.47	0.41	18.89	66413	UDF	6	10



Discrétiser pour cartographier

- Cartographier une série quantitative nécessite de discrétiser les valeurs

Discrétiser = découper une série statistiques en classes

- **Grand compromis statistique - cartographie**

- **Résumer** au mieux la distribution (conserver l'hétérogénéité)
 - Optimum: le plus grand nombre de classes
- **Construire une carte efficace** (synthétiser)
 - Optimum: un faible nombre de classes et d'effectifs égaux.
- Choisir des combinaison de **variables visuelles** restituant au mieux les spécificités de la distribution statistique.
 - Règles de sémiologie graphique

Les règles de la réduction statistique de l'information en vue d'une cartographie.

- ◆ Conserver l'ordre de grandeur
- ◆ Conserver la forme de la distribution
- ◆ Conserver la dispersion
 - Maximiser l'hétérogénéité entre les classes
 - Maximiser l'homogénéité à l'intérieur des classes
- ◆ Décider de conserver ou non les singularités

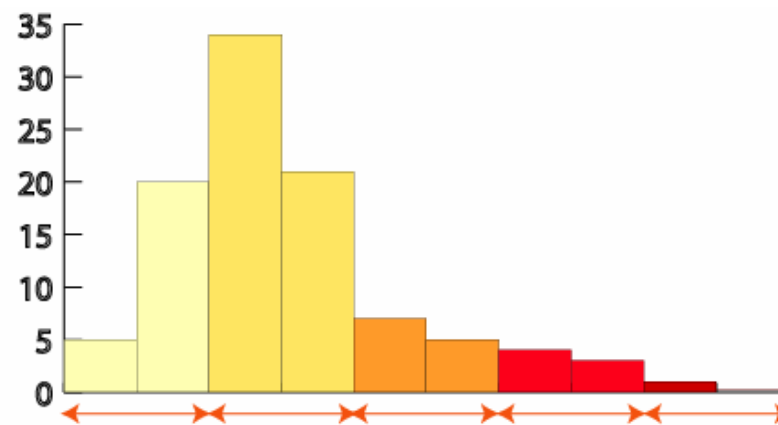
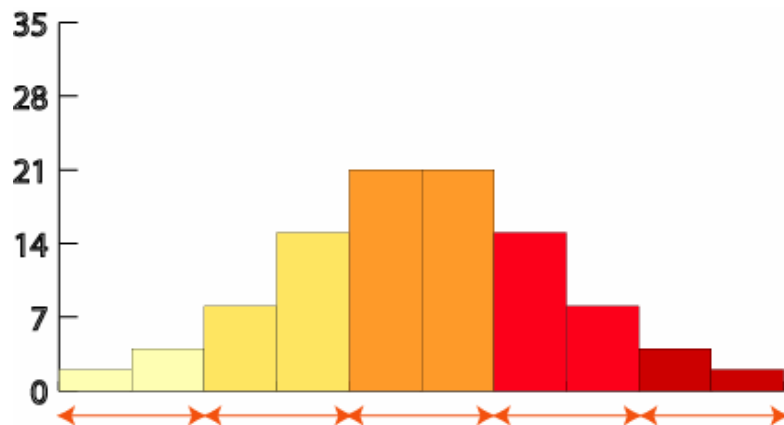
Il n'y a pas une discrétisation optimale.

Il y a une démarche *cohérente* par rapport à un *objectif* donné et une *distribution* donnée.

Les méthodes de discrétisation

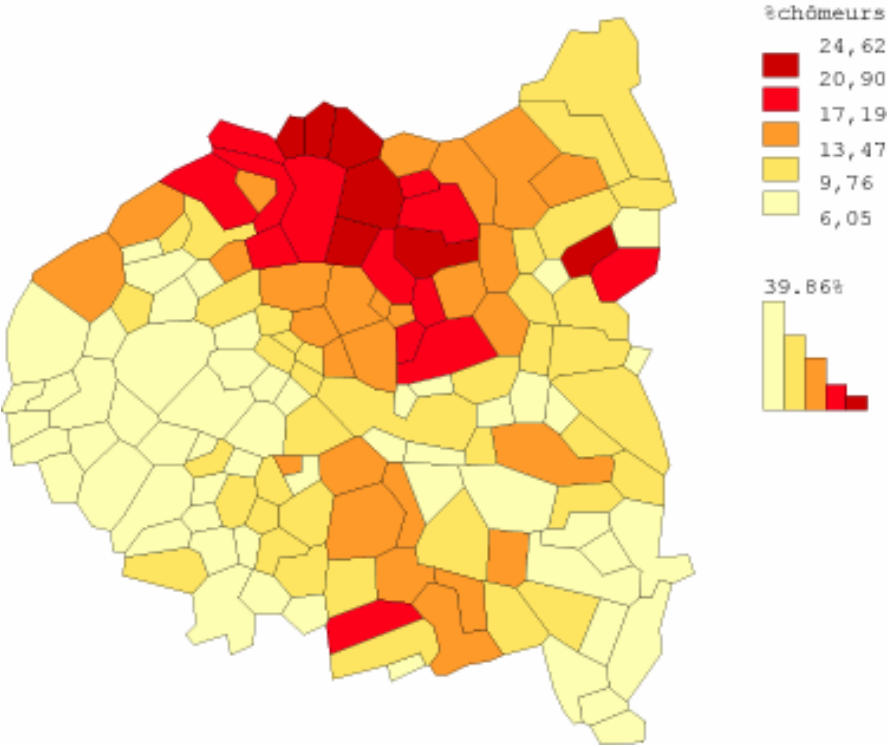
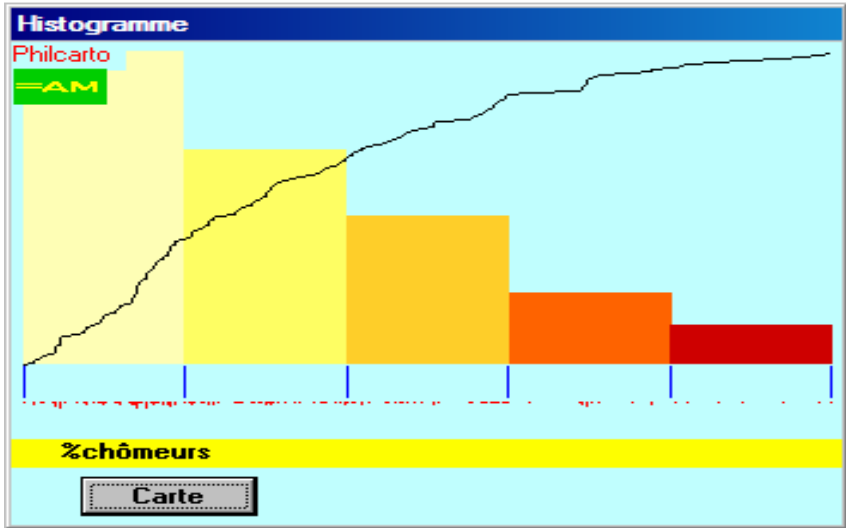
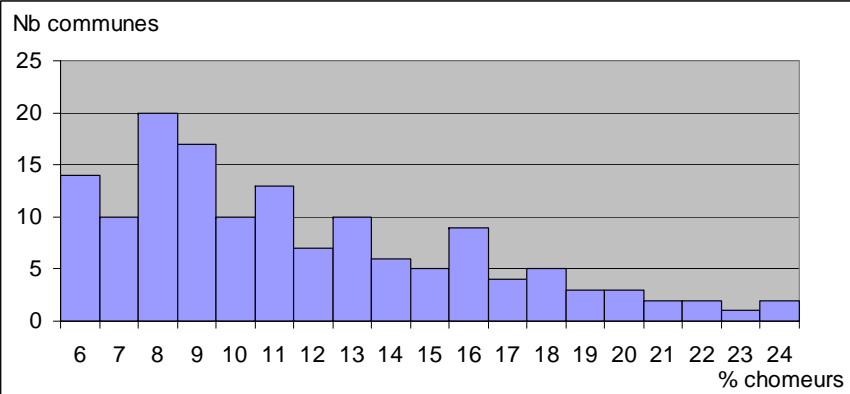
- Classes d'amplitudes égales
- Classes basées sur la moyenne et l'écart-type
- Classes d'effectifs égaux (ou quantiles)
- Classes en progression géométrique
- Méthode des seuils naturels
- Méthode de Jenks

Discrétisation en classes d'amplitudes égales

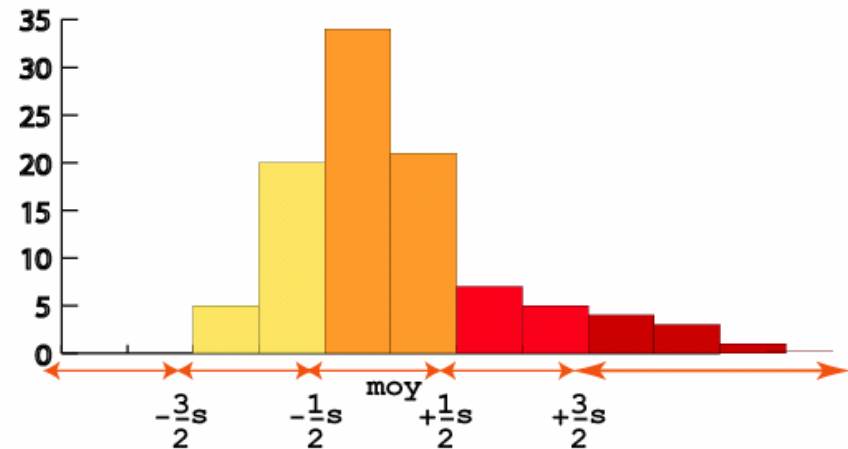
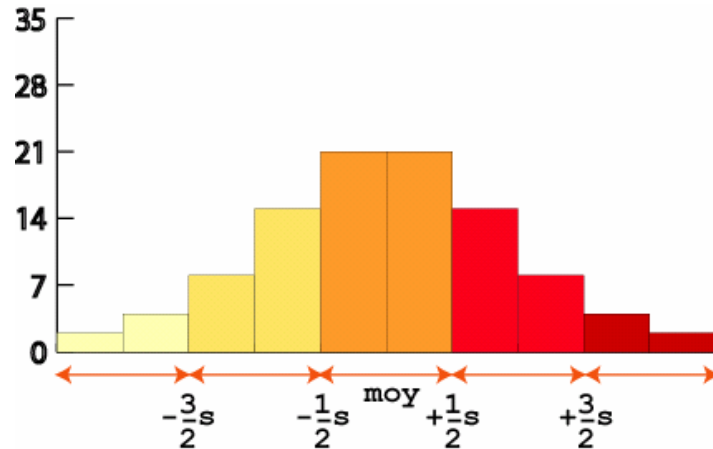


- + Facile à construire, facile à lire, conserve la forme de la distribution
- Optimum carto pour distribution uniforme (rare) -- pour les distributions asymétriques

Classes d'amplitudes égales

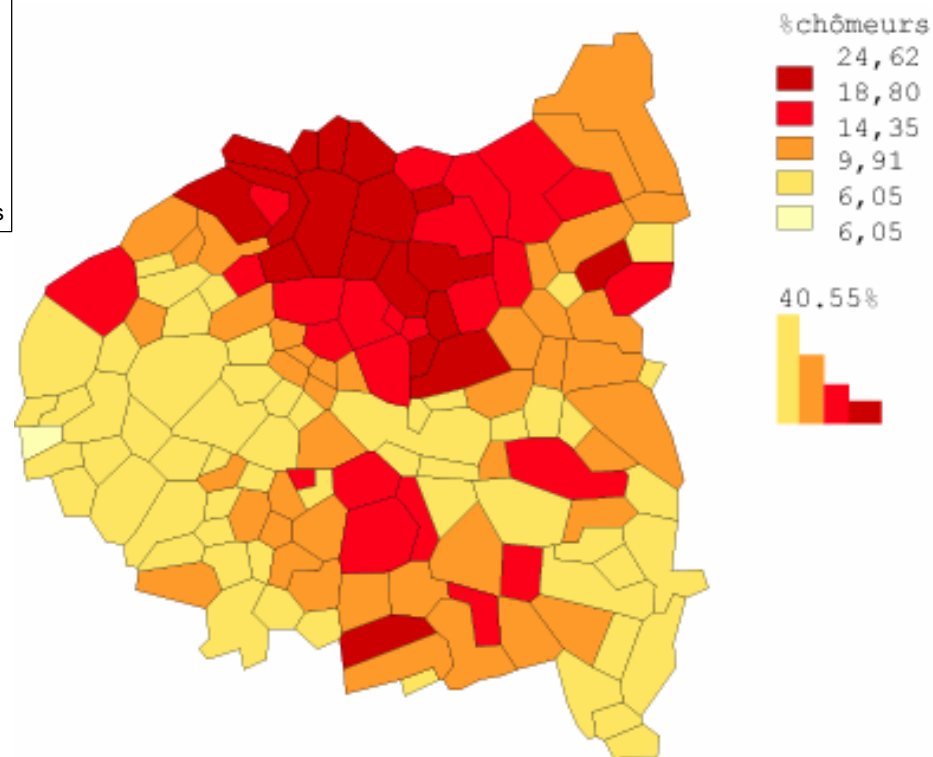
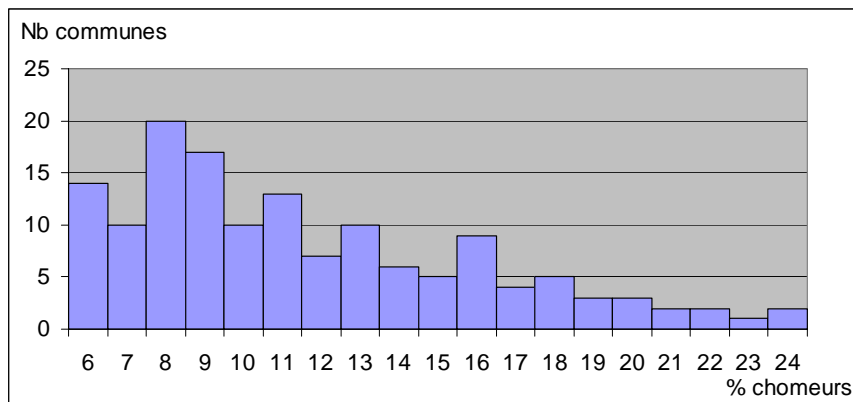


Discrétisation basée sur la moyenne et l'écart-type

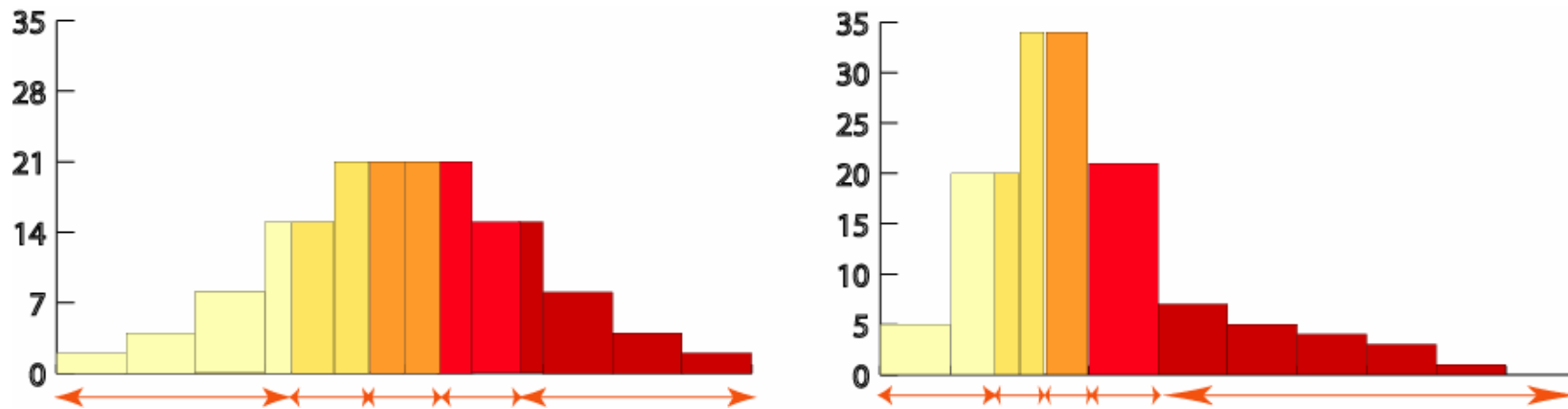


- + Fait référence aux caractéristiques de la distribution (moy, s)
- + permet la comparaison dans une unité commune
- Convient mal aux distributions asymétriques

Classes basées sur la moyenne et l'écart-type

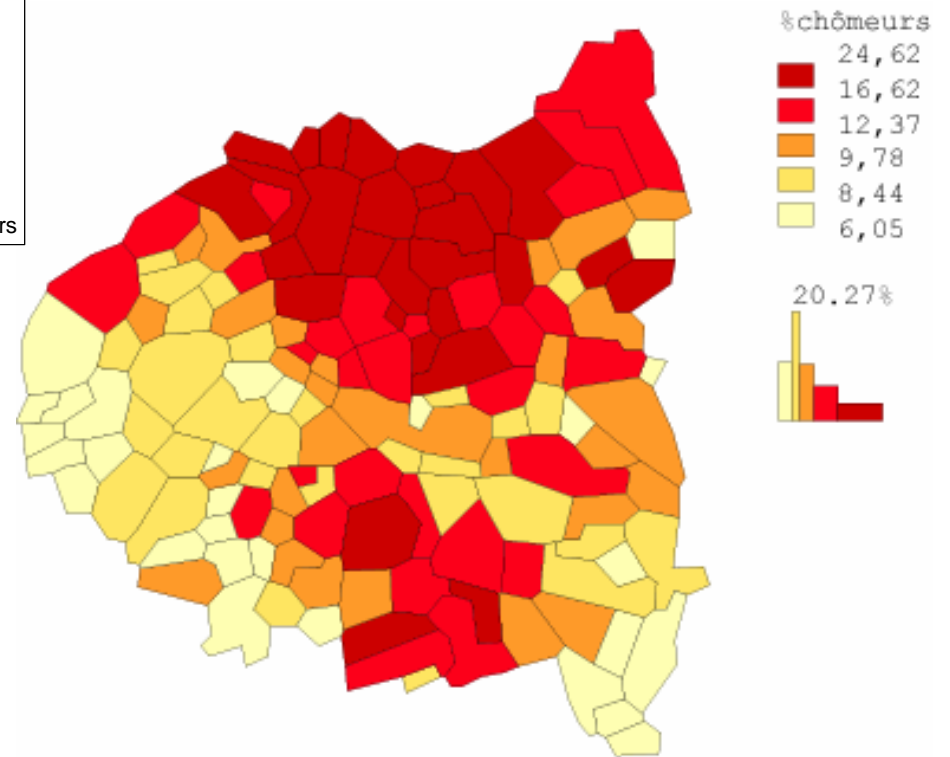
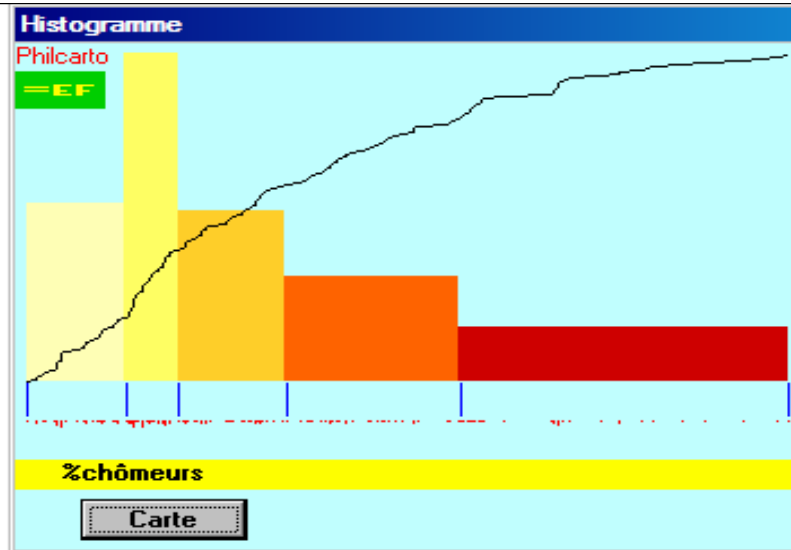
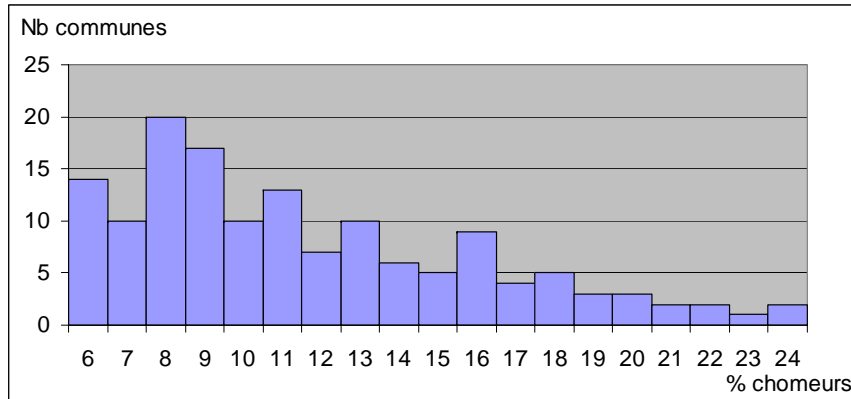


Discrétisation en classes d'effectifs égaux (méthodes des quantiles)

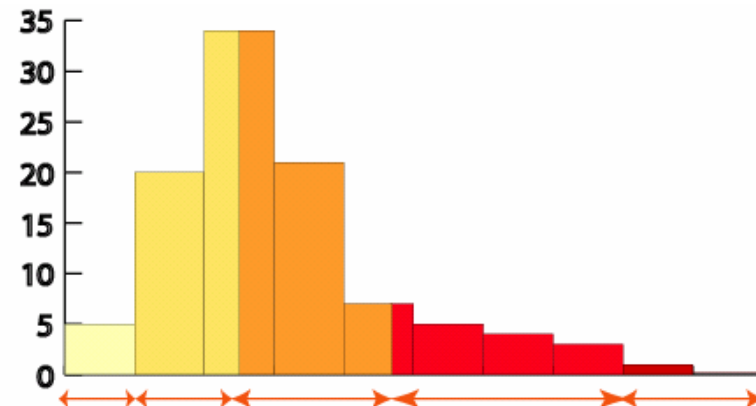
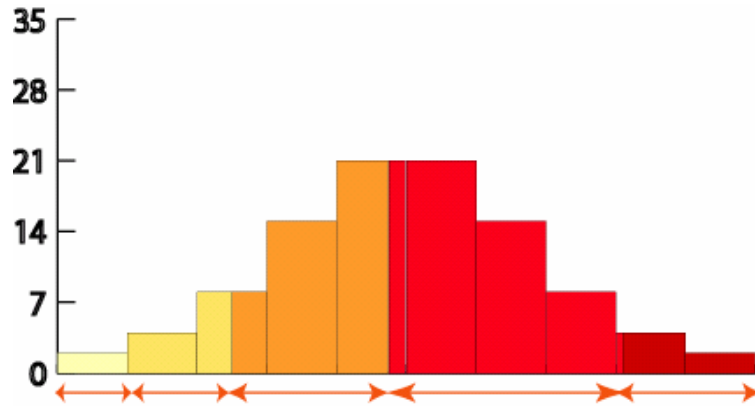


- + Transmet une information maximale
- + Privilégie les rangs plus que les valeurs
- Transforme complètement la distribution

Classes d'effectifs égaux

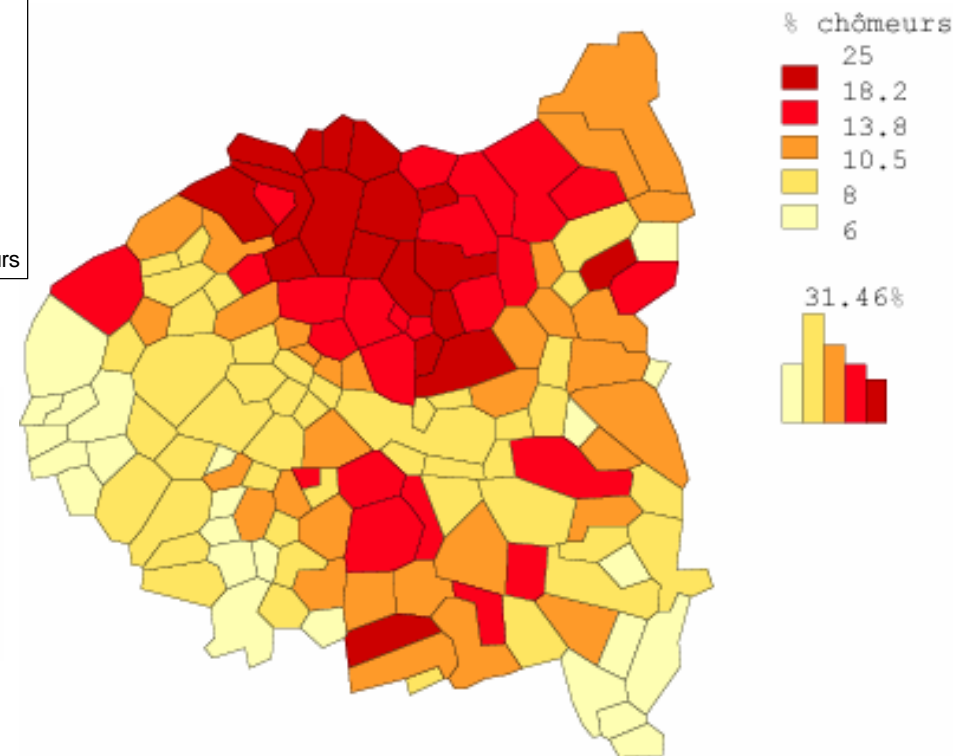
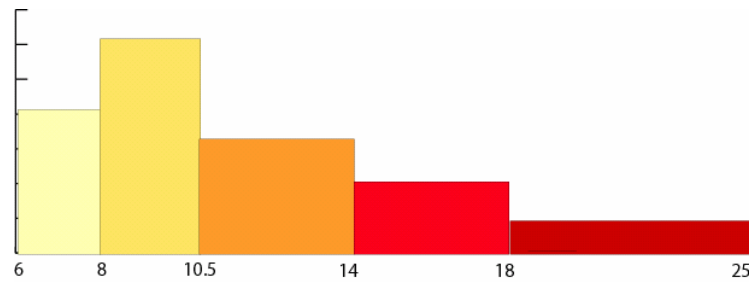
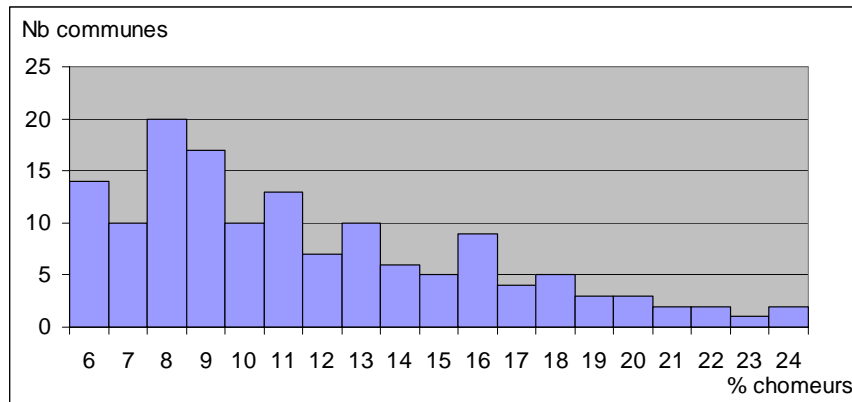


Discrétisation en classes en progression géométrique



+ Convient aux distributions asymétriques produites par exemple par des processus multiplicatifs

Classes en progression géométrique

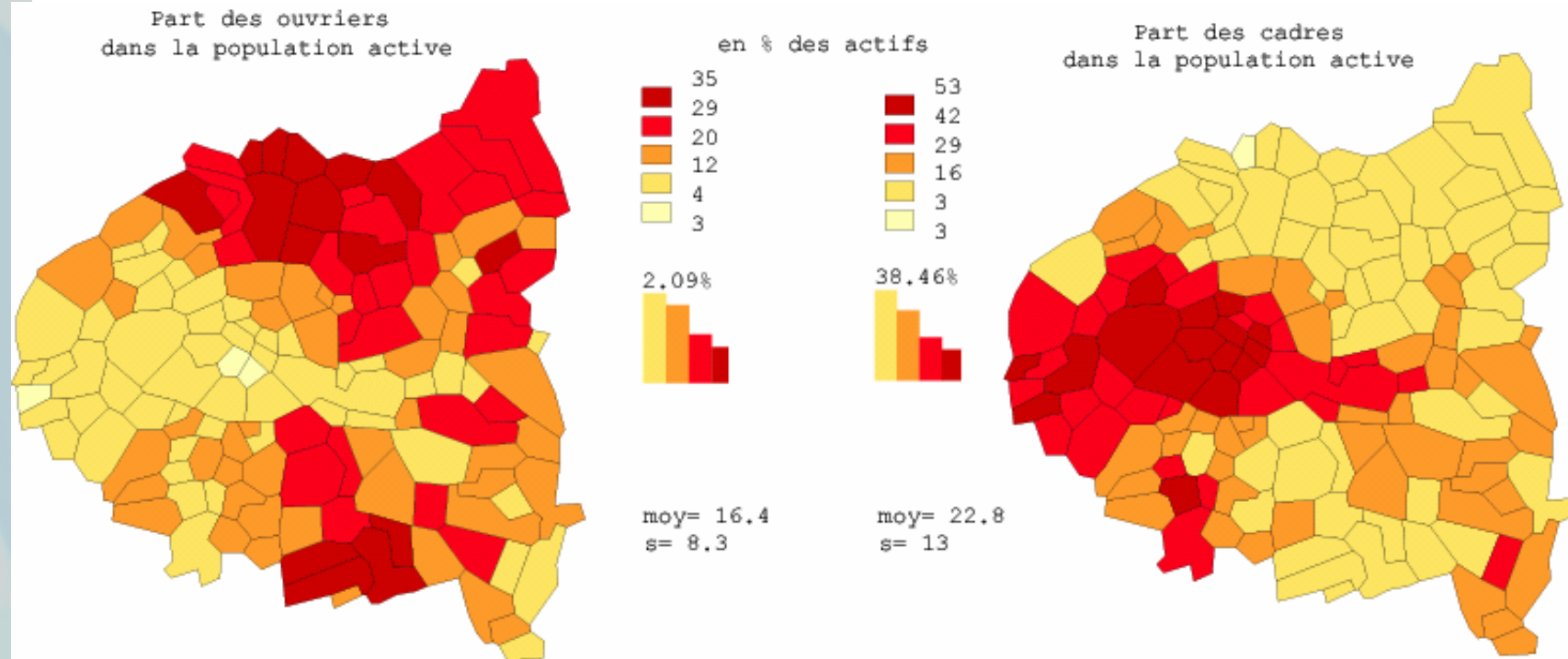


Discrétiser pour comparer

- La comparaison est un objectif spécifique de la cartographie
- Comparaison des positions relatives ou des positions absolues ?
- Comparaison
 - Des distributions de 2 phénomènes sur un même espace
 - Dans le temps (cartographie d'une évolution)
 - Des distributions d'un même phénomène sur 2 espaces différents ou à 2 échelons différents.

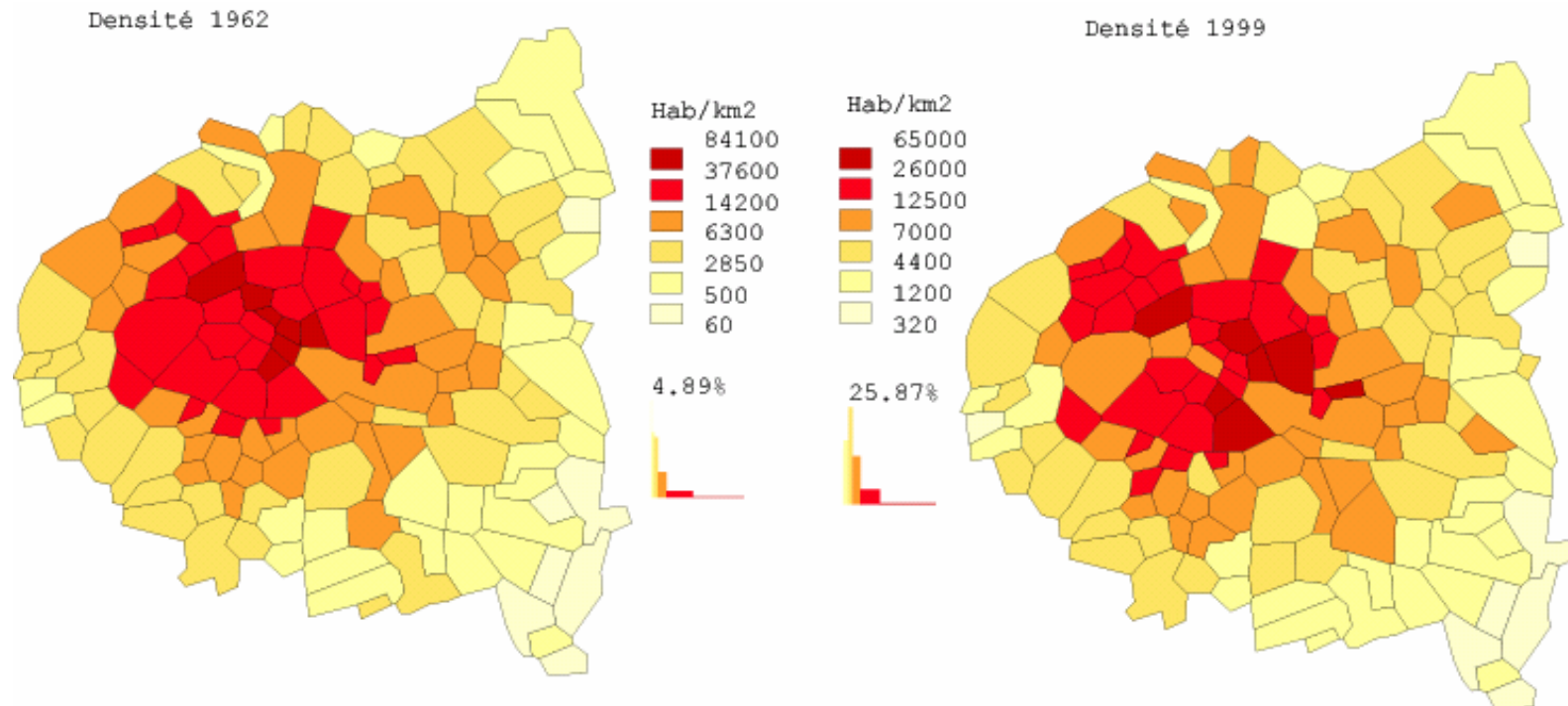
Comparaison des positions relatives

- 2 phénomènes sur un même espace:
La part des cadres et des ouvriers.



Comparaison des positions relatives

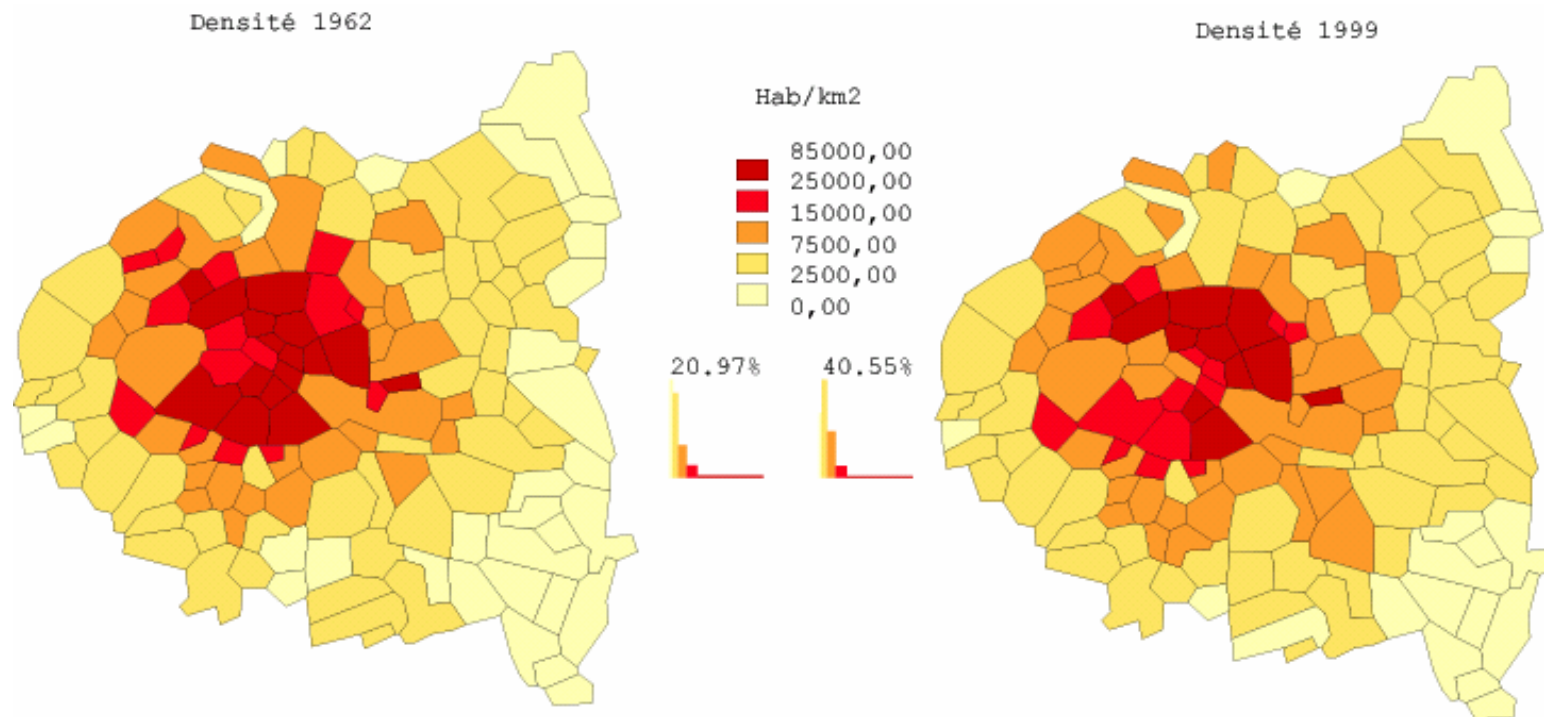
- Évolution d'un phénomène:
les densités communales entre 1962 et 1999



Méthode: min,p5,Q1,Q2,Q3,p95,max

Comparaison des positions absolues

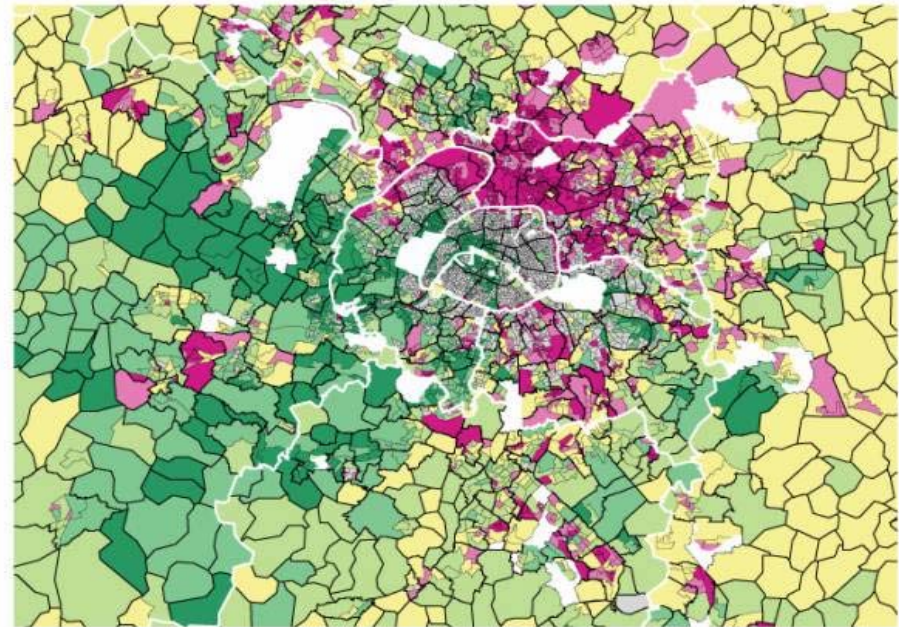
- Évolution d'un phénomène: les densités communales entre 1962 et 1999



Les problèmes que posent le territoire aux statistiques et à la cartographie

◆ Problèmes de comparabilité:

- Effet de contour
- Effet de bordure
- Effet de définition



L'exemple du découpage en Iris:

- Homogénéité de définition sémantique
- Hétérogénéité de définition spatiale