



INITIATION A LA CARTOGRAPHIE AUTOMATIQUE

Avant propos

Ce cours vise à présenter les méthodes et techniques de la cartographie statistique par ordinateur, dite cartographie automatique. Il se base sur les acquis du cours de cartographie générale, pour présenter les limites et avantages de l'automatisation grâce à l'outil micro-informatique.

L'essentiel des méthodologies présentées requiert l'utilisation d'un logiciel de cartographie automatique performant, accompagné si possible d'un logiciel tableur. La bonne compréhension de ce cours nécessite donc une importante phase de travaux personnels, de pratique des méthodologies, c'est pourquoi chaque leçon est complétée par la manipulation d'un tel logiciel en situation réelle sous la forme d'exercices guidés.

Le logiciel utilisé est SCALP, un logiciel « open source » et gratuit disponible à l'adresse Internet suivante : <http://w3.univ-tlse2.fr/geoprdc/scap/java/>

Ce logiciel constitue un complément idéal au présent cours, car il a été conçu dans ce but et vous donne un accès simplifié aux techniques présentées.

Bibliographie :

- BEGUIN Michèle, PUMAIN Denise - 1994 - *La représentation des données géographiques*-
Ed. Armand Colin – Paris.
- BERTIN Jacques - 1977 - *La Graphique et le Traitement Graphique de l'Information*.
Ed. Flammarion – Paris.
- BRUNET Roger - 1987 - *La Carte Mode d'Emploi*. - Ed. Fayard/RECLUS – Paris
CAUVIN Colette, REYMOND Henri, SERRADJ Abdelaziz - 1987 " Discrétisation et représentation cartographique." - Ed. RECLUS, Montpellier.
- ZANIN Christine, TRÉMÉLO Marie-Laure - *Savoir faire une carte, aide à la conception d'une carte thématique univariée*, Ed. Belin, Coll. Sup Géographie, Paris 2003.

Webographie :

Annuaire de ressources pour la cartographie (logiciels, fonds de carte...) :
<http://w3.univ-tlse2.fr/geoprdc/bazarcarto/index.php>

SCAP-Java, logiciel open source et gratuit de cartographie automatique :
<http://w3.univ-tlse2.fr/geoprdc/scap/java/>

Portails et forums :

<http://www.forumsig.org> et <http://www.portailsig.org>
<http://www.georezo.net>

Tutoriaux en cartographie :

<http://www.usherbrooke.ca/geotel/geomatique/tutoriels.html>
http://www.ensg.ign.fr/FAD/Supports_de_Cours.html

Dictionnaire de la cartographie :

<http://www.microplan.fr/Dictionnaire.htm>

I- Positionnement et pré-requis de la cartographie automatique

Le présent cours développe les techniques de la cartographie générale dans un domaine particulier : la conception cartographique assistée par micro-ordinateur (CCAO), dans le domaine de la cartographie statistique.

La CCAO va aider le cartographe dans les différentes phases de la réalisation d'une représentation cartographique, de l'analyse/traitement de la variable à l'édition du document final, mais, à cause des limites des logiciels, *elle ne s'applique que dans certains cas et qu'à certains types de carte.*

Ce cours va donc reprendre une grande partie du vocabulaire et des acquis du [cours de cartographie générale](#), qu'il est donc préférable d'avoir parcouru auparavant. On reprendra en effet tous les concepts concernant les [variables graphiques](#), [l'imposition](#) et la recherche des [composantes](#), tout en maintenant les mêmes objectifs généraux de lisibilité, d'[objectivité](#) et d'[efficacité](#).

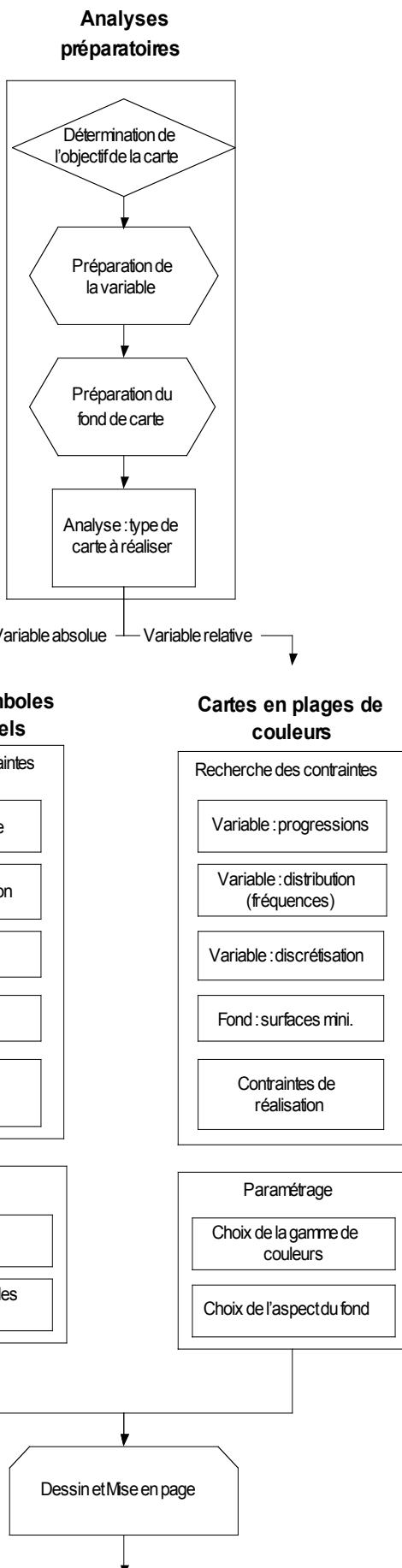
On va ici développer une petite partie des techniques de cartographie générale pour présenter les méthodes et les outils qui permettent de les automatiser. Grâce à ces techniques, la réalisation complète d'une carte thématique est automatisée dans de nombreux aspects, notamment dans ses opérations les plus longues et fastidieuses (discrétisation, dessin), pour permettre au cartographe d'aboutir rapidement à une représentation graphique de données statistique sur un fond de carte.

Mais cette rapidité potentielle d'exécution de la carte peut être perçue comme une invitation à laisser le logiciel décider à votre place. Ce serait utiliser un outil sans en guider les mouvements, ce qui aboutit souvent à des accidents... C'est pourquoi il faut suivre une méthodologie précise, basée sur la compréhension et la résolution des contraintes. Le schéma de la page suivante résume cette méthodologie.

Ce cours est structuré par les principes suivants :

- La carte à réaliser doit être la représentation la plus claire et la plus fidèle (objective) de l'information donnée.
- La carte à réaliser sera thématique et non synthétique.
- La variable à représenter est quantitative, donc ordonnée et proportionnelle.

Méthodologie de la Cartographie Automatique



II - Détermination de l'objectif de la représentation cartographique

a - Toute représentation est une simplification.

Une représentation graphique est la transcription sous forme de dessin de valeurs numériques. Un histogramme, une courbe, un diagramme en secteurs, une carte, sont des transcriptions de valeurs numériques sous une forme simplifiée, visuelle. Une variable quantitative devient alors une série de formes dessinées, selon des principes, des règles, des usages conventionnels.

Lors de cette transformation, on passe de valeurs très précises, numériques, à des formes dessinées, c'est donc une simplification assez grande de l'information. C'est cette simplification qui permet à l'information d'être plus facilement et rapidement compréhensible : il est plus intéressant de lire un graphique ou une carte que de lire le tableau de données qui en est la source.

Mais, naturellement, la simplification de l'information doit être réalisée de manière particulièrement soigneuse, au risque de passer à côté des caractéristiques importantes de la variable à représenter et donc d'en donner une image faussée au travers de la représentation graphique.

b - Pour simplifier utilement, il faut analyser l'information et fixer un objectif à la représentation.

En dehors des principes généraux qui guident la représentation (être clair et objectif, thématique et quantitatif), il faut fixer un objectif à la carte que l'on va réaliser, quant à l'information qu'elle doit *transmettre* et de sa destination. A quoi doit servir la carte ? Quel est le *public* visé ?

Cet objectif servira de guide lors de la sélection des caractéristiques de l'information à garder lors de la simplification et surtout lors du choix du type précis de représentation à effectuer. En fait, le principe d'objectivité reste le guide principal, mais le nombre de techniques de représentation disponibles laisse de la marge à l'expression de l'information.

Par exemple, lors de la réalisation d'une carte concernant une variable relative comme la densité de population, on utilise des plages de couleurs, et l'on va donc simplifier l'information de départ en regroupant en classes les valeurs d'origine, ces classes devenant les valeurs, les couleurs, de la carte.

Cette opération de regroupement en classes (qui sera détaillée plus loin) fait perdre énormément de l'information de départ, dans l'absolu : on passe, dans le cas d'une carte des départements français, de 99 valeurs à six ou sept seulement. C'est donc une simplification de l'ordre de 90% qui est effectuée !

Comment, lors d'une telle simplification, préserver l'information importante ? Le sens même de la variable numérique ?

Il existe plusieurs méthodes de découpage en classe (les discrétilisations), qui seront examinées dans le détail plus loin, mais comment faire le choix entre tel ou tel découpage ?

Il faut fixer un objectif à la représentation, et vérifier au long de la procédure que le résultat prévu reste conforme à cet objectif.

Pour reprendre l'exemple de la cartographie des densités dans les départements français, on peut fixer plusieurs objectifs à la carte, suivant l'utilité que cette dernière doit avoir (de manière caricaturale, bien sur) :

Auteur de la carte	Utilité de la carte	Objectif 1	Objectif 2
Chercheur Géographe	Etude, exploration	Avoir un aperçu global, déterminer des groupes spatiaux	Repérage des atypicités, des originalités
Gouvernement	Démonstration	Montrer l'homogénéité, l'harmonie	Désactiver les revendications potentielles
Elu local	Démonstration	Montrer l'hétérogénéité	Valoriser/dévaloriser son département

Ainsi, la simplification nécessaire pour une bonne compréhension de la représentation graphique amène parfois à un choix difficile, que seule une réflexion préalable sur l'utilité de la carte peut aider.

Nous nous placerons, dans ce cours, dans la position neutre, visant à minimiser la perte d'information, et privilégieront l'objectif de la réalisation de cartes d'aide à l'étude, d'exploration.

III - Préparation de la variable à représenter

a - Adaptation du niveau d'agrégation spatiale à l'objectif de la carte

Avant de se lancer dans la partie automatisée de la représentation cartographique, il faut fournir au logiciel que l'on utilise plusieurs données :

- Un fond de carte numérique adapté à la représentation;
- Une ou plusieurs séries de données numériques contenant l'information à représenter.

La question du fond de carte sera traitée dans [une leçon suivante](#), nous nous intéresserons toutefois ici au niveau d'agrégation du fond de carte. Il s'agit du type des unités géographiques qui composent le fond de carte, souvent des entités administratives : communes, départements... En effet, une fois l'information récoltée (cf. les liens Internet donnés), l'objectif de la carte fixée, il faut se poser la question suivante :

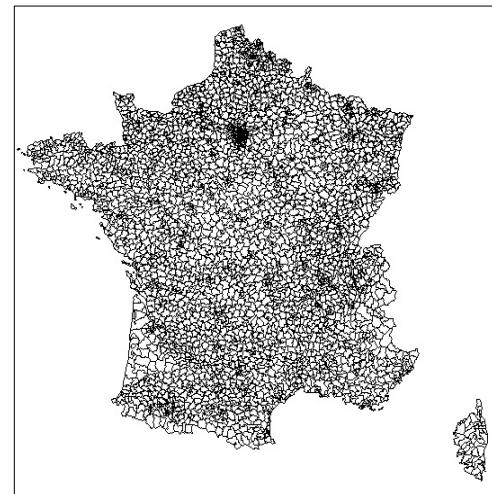
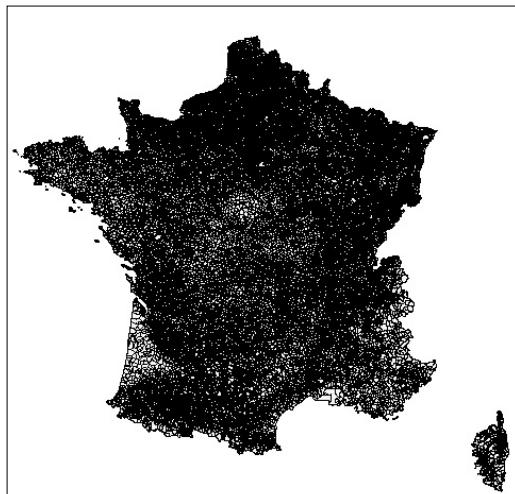
Le niveau de précision géographique des informations récoltées est-il adapté à l'objectif fixé à la carte ?

Souvent (et c'est tout à fait sain) la récolte de l'information numérique se fait au niveau le plus fin possible, au niveau auquel l'information a été produite par l'organisme de statistiques ou l'enquêteur sur le terrain. Or, on l'a vu, la représentation cartographique demande une simplification de l'information pour la rendre plus facilement compréhensible, et il faut donc se poser la question de la simplification du niveau d'agrégation de l'information de base pour faciliter sa lecture.

Par exemple, pour reprendre l'idée de la carte des densités en France, l'information est produite par l'INSEE au niveau communal, voire au niveau des quartiers dans les agglomérations.

Mais une carte dans un format courant ne peut présenter clairement les 36.600 communes de France, et ce niveau d'information, cette précision, n'est pas compatible avec un objectif de repérage rapide de la répartition des densités de population sur le territoire national.

Il faut donc envisager de regrouper les informations à un niveau d'agrégation spatiale supérieure : les cantons, les arrondissements, les départements.



b - Adaptation du niveau d'agrégation au fond de carte

L'acquisition de variables statistiques et leur transformation sont des opérations qui sont devenues progressivement moins difficiles avec la micro-informatique et la diffusion sur Internet. Par ailleurs, les outils de traitement numérique et statistique sont de plus en plus courants et accessibles (outils libres notamment : OpenOffice, R....).

Par contre, la disponibilité d'un fond de carte numérique adapté précisément aux données numériques disponibles n'est pas encore très courante, à cause de la spécificité de ce type de données et de la complexité des formats de fichier

En effet, la création d'un fond de carte numérique est encore une opération assez complexe, qui revient souvent à dessiner à la souris le fond de carte nécessaire grâce à un logiciel spécifique. Alors que les formats de fichiers de données sont assez standardisés (txt, dbf...), les formats de fonds de cartes sont souvent liés à des logiciels propriétaires.

Donc, avant de vous lancer dans la récolte d'informations tous azimuts sur un thème, veillez à recenser les fonds de carte qui vous seront accessibles ou à vérifier si la création de nouveaux fonds est envisageable.

Sinon, vous serez obligés de modifier vos données statistiques pour les adapter au fond de carte disponible.

c - Adaptation de l'unité de la variable, de l'indicateur, à l'objectif de la carte.

Les opérations de cartographie étant encore relativement longues, on trouve souvent des cartes qui sont mal adaptées à l'objectif de leur auteur, qui n'a pas voulu réaliser d'autres cartes que ses cartes d'étude. On commence en général par une carte (ou une série de cartes) qui vont représenter la répartition géographique des phénomènes statistiques qui nous intéressent. Ces cartes sont faites en essayant de perdre le moins d'information possible, pour essayer de mettre en avant des rapprochements géographiques. Ensuite seulement on peut synthétiser les idées issues de ces premières cartes en réalisant des analyses croisées, des cartes d'indicateurs plus complexes.

Vous l'aurez noté, les objectifs et les publics de ces différentes cartes sont très différents. Les cartes ne doivent donc pas être réalisées selon les mêmes principes directeurs.

Lorsque vous cherchez à présenter une vision synthétique sur une zone d'étude, essayez de réaliser des cartes qui elles aussi seront synthétiques et apporteront une information croisée. Les cartes et graphiques peuvent très bien se comparer aux paragraphes d'une argumentation : elles peuvent se compléter et se faire référence l'une l'autre. Une longue série de cartes simples sera moins intéressante pour votre public.

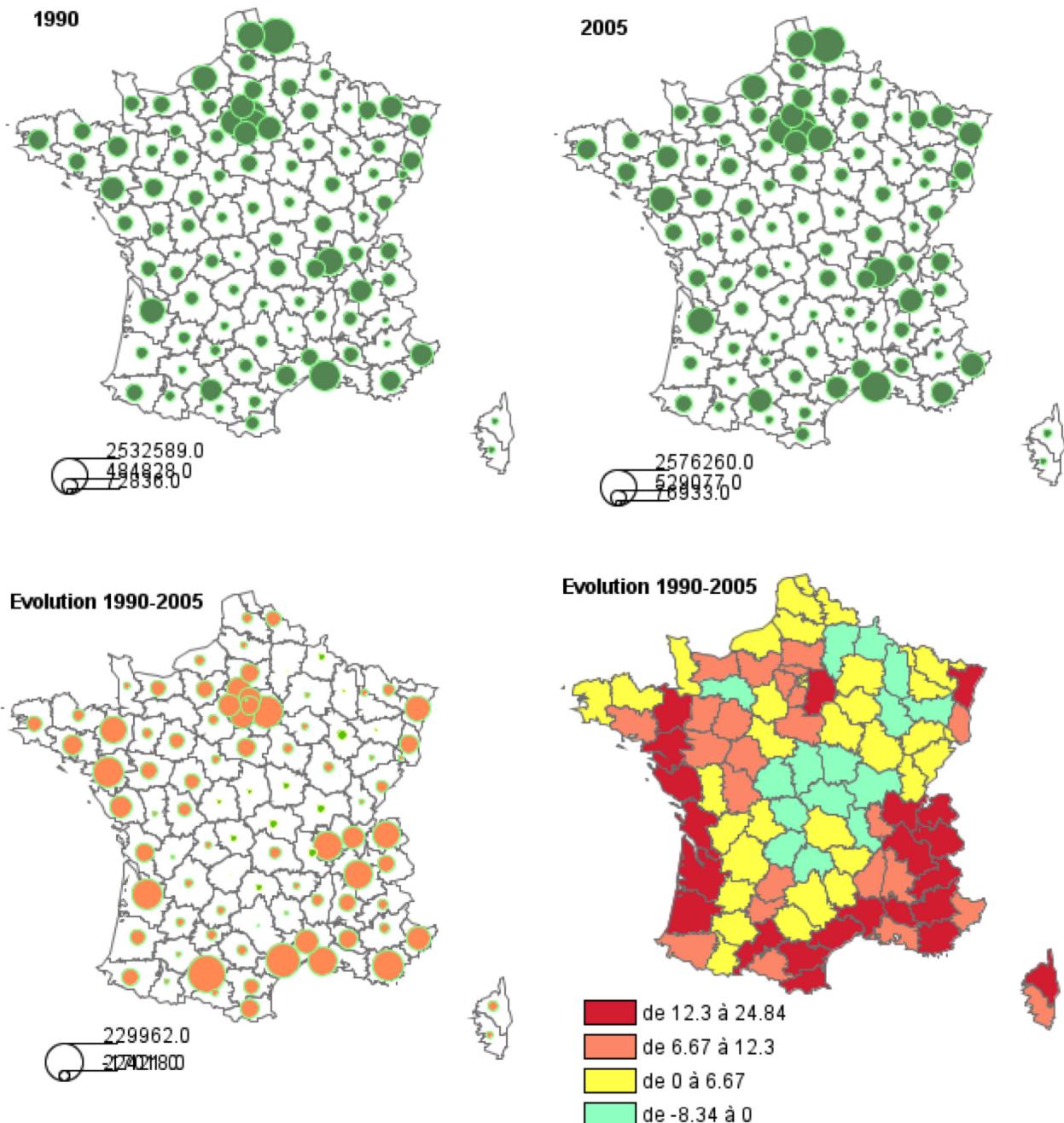
Par exemple, au lieu de multiplier les cartes en valeurs absolues sur toutes les variables dont vous disposez (population, répartition par âges, par sexe, par activités... sur plusieurs années), essayez de croiser ces variables pour produire moins de cartes mais des cartes plus synthétiques (densités, taux et rapports), qui viendront compléter une ou deux cartes en valeurs absolues. L'aboutissement de cette pratique serait alors les cartes de typologie, chaque type représentant une série de valeurs sur plusieurs variables.

On peut aussi envisager de réaliser une carte statistique double, présentant une variable absolue et une variable relative en même temps (la [leçon suivante](#) présentera les modalités du choix du type de carte à réaliser selon la variable, et une [leçon](#) est consacrée aux types de cartes complexes).

Il faut donc se poser la question suivante avant de se lancer dans la partie cartographique proprement dite :

Quelle est l'unité, l'indicateur, la façon de présenter mathématiquement la variable qui sera la plus intéressante, par rapport aux autres cartes du document en cours de réalisation ?

Par exemple, est-il intéressant de représenter la population en 1990 et la population en 2005 ou la population en 2005 et le taux de variation entre ces deux recensements ?



Une fois l'objectif de la carte fixé, la variable adaptée, il faut se poser la question du type de carte à réaliser.

IV - ANALYSE DE L'INFORMATION A REPRÉSENTER

Méthodologie du choix du type de carte.

Souvenez-vous des principes énoncés en introduction :

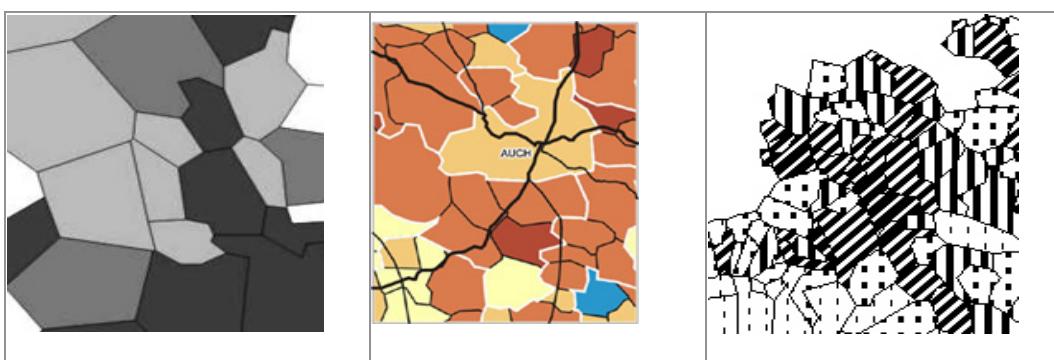
- La carte à réaliser doit être la représentation la plus claire et la plus fidèle (objective) de l'information donnée.
- La carte à réaliser sera thématique et non synthétique.
- La variable à représenter est quantitative, donc ordonnée et proportionnelle.

A partir de ces principes et selon le type de variable spécifié, le [cours de cartographie générale](#) nous a appris que seuls certains types de cartes pouvaient être réalisés. La première opération consiste à repérer quel est le type d'*objet cartographique* qui est la source d'information de la variable numérique à représenter.

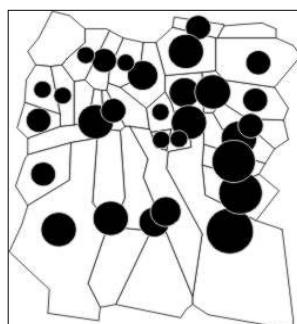
Ce repérage permet de choisir le type de fond de carte : points, lignes, ou polygones. La seconde opération concerne le choix du type de représentation entre les deux grandes familles de cartes statistiques : *plages de couleurs* (carte dite Choroplèthe) ou carte en *symboles proportionnels*? Le choix est plus ardu lorsque la composante géographique de l'information est un objet que l'on peut représenter comme un point ou comme un polygone, et c'est très souvent le cas (unités administratives).

La méthodologie de Jacques Bertin ([sémiologie graphique](#)) nous laisse le choix d'utiliser différentes techniques de dessin (les variables visuelles) pour représenter l'information. En effet, on peut représenter une variable statistique concernant une entité pouvant posséder une surface de deux manières :

- par une couleur, une valeur, une trame sur la surface occupée par cette entité;



- par un symbole ponctuel centré sur cette surface.



Il existe une règle, en CCAO, qui permet de sélectionner le type de carte à réaliser :

- Lorsque la variable statistique est absolue, brute, issue d'un comptage, on utilise la variable graphique de la taille qui représentera directement ses variations quantitatives par les variations de la surface d'un symbole.
- Lorsque la variable statistique est relative, transformée, issue d'un calcul, elle est valable en tout point de la surface de l'objet géographique, on ne peut la relier à un objet concret. On utilise donc la variable graphique de la valeur et/ou de la couleur, adaptées aux informations différentielles ordonnées.

Le point clé de ce choix est donc une interrogation sur le type de variable statistique à représenter : absolue ou relative ? Brute ou transformée ? *La réponse à cette question est cruciale, car elle va déterminer le type de carte à réaliser.*

Synthétiquement :

Type de Variable	Type de carte
Absolue, brute	Symboles proportionnels (taille)
Relative, transformée	Plages de couleurs (valeur/couleur)

Exemples de variables :

Absolue, brute	Relative, transformée
Population	Densité de population
Production en unités	Taux d'évolution
Comptage de véhicules	Part de salariés dans la population active.

Les comptages, dénombresments, quantités : habitants, tonnes, m², véhicules... sont des valeurs brutes, absolues. On représentera toujours, par exemple, les chiffres de la population en nombre d'habitants sous la forme de symboles proportionnels.

Les moyennes, taux, indices, pourcentages sont des valeurs relatives (elles se définissent toujours en rapport à une autre valeur, ne sont pas autonomes). On représentera toujours les densités, par exemple, en plages de couleurs car elles s'expriment en habitants au km².

V - Réalisation de cartes en symboles proportionnels

a - Repérage des contraintes

Positionnement

Cette leçon est la première du chapitre concernant les cartes en symboles proportionnels.

Ce type de carte ne se réalise que dans la situation suivante :

- La variable statistique à représenter est absolue, brute, discrète.
- L'implantation des individus géographique sur le fond de carte est ponctuelle, ou rendue ponctuelle par la sélection du centroïde d'une implantation zonale.

Méthodologie

Une fois la variable et le fond de carte prêts, le type de carte à réaliser défini, il ne reste plus qu'à passer à la réalisation concrète de la représentation cartographique.

Synthétiquement, la représentation cartographique d'une valeur absolue en symboles proportionnels consiste à utiliser une relation mathématique entre la valeur statistique et la surface des symboles proportionnels à dessiner sur le fond de carte. C'est en effet la surface des symboles qui est visualisée, et qui est donc le paramètre à faire varier proportionnellement.

Avant de rechercher cette relation mathématique, il faut déterminer quelles sont les contraintes qui vont limiter nos choix, contraintes issues de la variable, du fond, de l'objectif de la carte, des techniques de dessin et de reproduction disponibles...

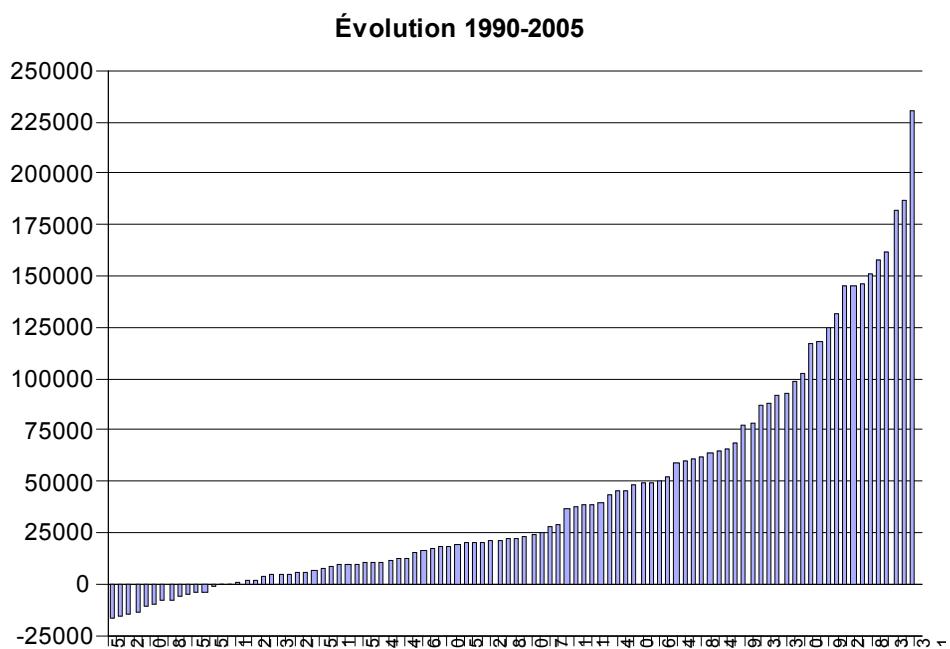
La procédure de réalisation d'une carte statistique commence toujours par une recherche des différentes contraintes qui vont s'exprimer.

En effet, ce sont ces contraintes, qui, une fois repérées, vont guider progressivement les choix précis de représentation de la variable.

Les contraintes de la variable

Le premier groupe de contraintes provient de la variable à représenter, et concerne principalement son étendue et sa dispersion :

- Étendue de la variable : différence entre la valeur la plus grande et la valeur la plus faible.



Exemple de dispersion : évolution de population dans les départements français entre 1990 et 2005 (source INSEE-RGP)

Cette étendue est une contrainte, car pour représenter la variable en utilisant la surface d'un symbole, il faut choisir un *ratio valeur de la variable / surface du symbole* qui permette de représenter correctement tous les individus. Concrètement, il faut pouvoir repérer facilement sur la carte l'individu possédant la valeur la plus faible et l'individu à l'opposé, même si ils sont très éloignés par leur valeur. Nous verrons dans la leçon suivante quelles techniques permettent de s'en sortir dans ce cas.

Si cette contrainte n'est pas correctement repérée, on risque de se retrouver à faire une carte où les plus petits symboles n'apparaîtront pas !

- Dispersion de la variable : répartition des valeurs.

La dispersion des valeurs à l'intérieur d'une variable peut jouer un rôle important sur l'aspect général de la carte en symboles.

Si la dispersion des valeurs est faible, on risque de se retrouver avec une carte où les symboles ont à peu près tous la même taille, mis à part les individus extrêmes. Au contraire, avec une dispersion trop grande, il risque de n'y avoir que des petits et des gros symboles sur la carte.

Dans ces cas extrêmes, ces problèmes de dispersion doivent être pris en compte et le ratio valeur / surface doit être transformé.

Les contraintes du fond de carte

Même si la variable semble à peu près équilibrée dans son étendue et sa dispersion, on ne peut imaginer à l'avance l'aspect général de la carte que si on s'intéresse de plus près à l'autre composant d'une carte : le fond de carte.

En effet, les individus géographiques, comme les individus de la variable, peuvent présenter des contraintes importantes, dans leur répartition dans l'espace.

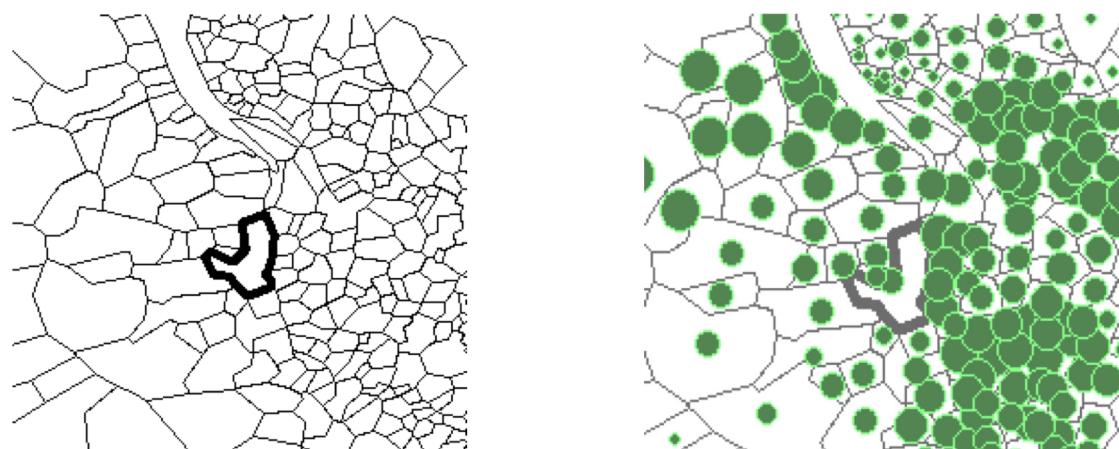
La principale contrainte concernant le fond de carte pour les analyses symboliques concerne en effet la répartition des points qui vont supporter des symboles dans le plan.

- Repérage des contraintes de répartition

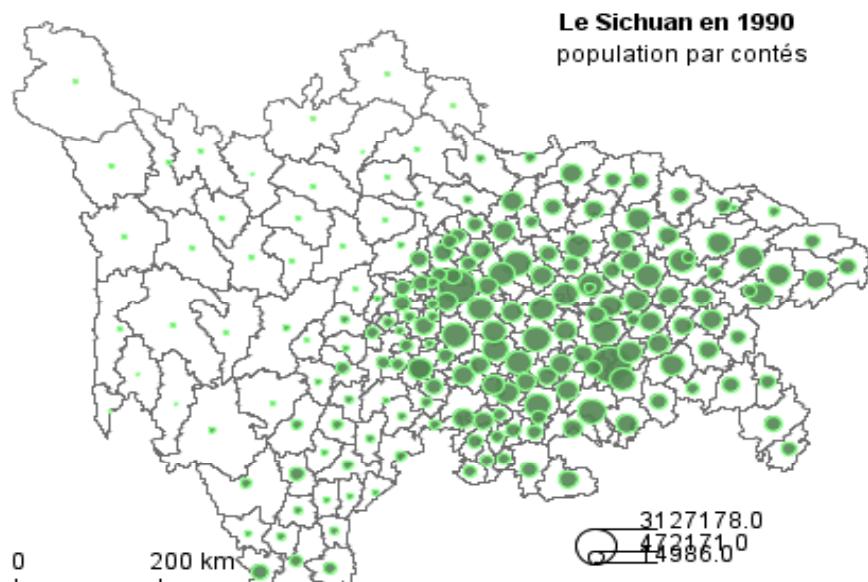
La principale contrainte à la réalisation des analyse symboliques, qui va diriger la procédure de cartographie (cf. leçon suivante), est la dimension et la position respective des symboles les plus « gros » de la carte. Le problème majeur à résoudre lorsque l'on réalise une carte par symboles proportionnels et le dimensionnement des plus gros symboles, pour qu'ils ne masquent pas le fond de carte, tout en restant assez grands pour que les plus petits symboles (qui leur restent proportionnels), soient visibles.

Ensuite on trouve les contraintes de répartition, qui sont les contraintes issues de la grande proximité des ces individus à valeurs élevée, qui risque de nuire à la lisibilité de la carte. Elles souvent liées à la présence de grandes villes dans le fond de carte. Souvent, les villes centres des agglomérations possèdent des valeurs très importantes, de loin supérieures à celles des communes qui les entourent. Le problème va être alors de représenter sur la même carte ces valeurs très différentes et très rapprochées.

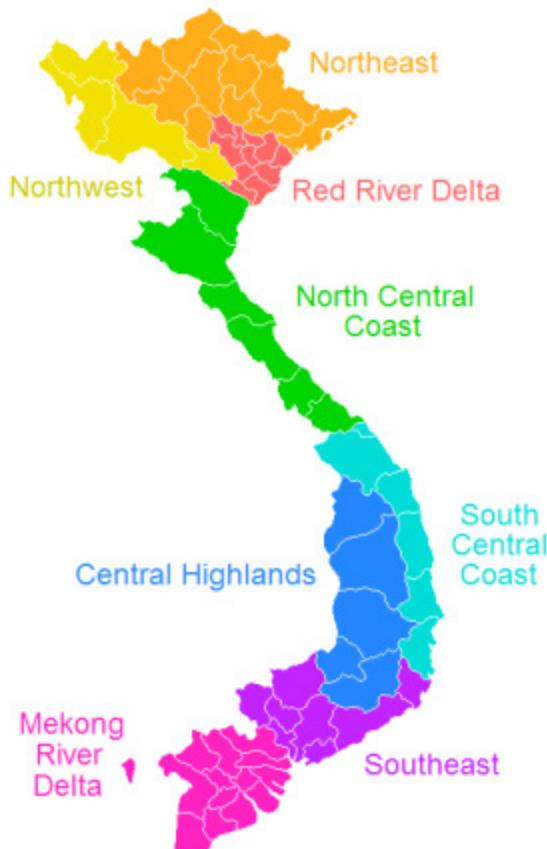
Exemple : agglomération de Bordeaux, représentation du nombre d'intercommunalités en 1999.



Exemple : la population de la province chinoise du Sichuan, concentrée autour de Chongqing et Chengdu :



A d'autres échelles, cette contrainte apparaît aussi, dans les cas de fonds de cartes possédant une forme allongée, comme dans le cas du Viêt-nam ou du Chili, par exemple.



Source : Wikipedia document



Source : CIA World Factbook atlas

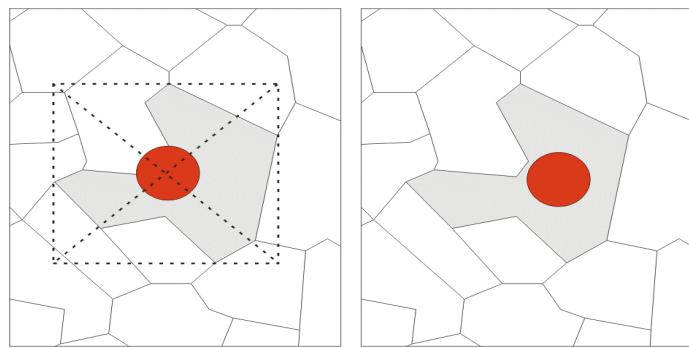
- Repérage de problèmes de positionnement des centroïdes

L'autre grande contrainte du fond de carte pour les cartes en symboles proportionnels provient de la nécessité de localiser correctement un point à l'intérieur d'une surface.

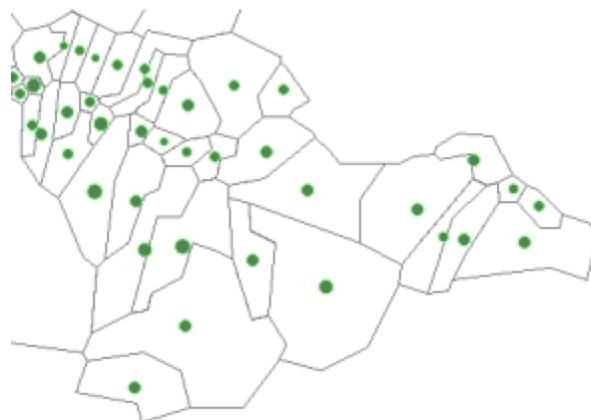
En effet, les fonds de carte informatisés sont souvent des fonds composés de polygones, de surfaces, et le logiciel va représenter les analyses symboliques en plaçant les symboles sur un point à l'intérieur des polygones, nommé centroïde.

Selon les logiciels, il est possible d'utiliser des centroïdes automatiques ou d'utiliser un fond de carte de points en superposition avec celui de polygones, pour utiliser par exemple la localisation exacte des villes préfectures pour les départements.

Mais le plus souvent le logiciel détermine seul la position de ce point, avec plus ou moins de bonheur, car souvent c'est la méthode du rectangle d'encombrement minimal qui est utilisée :



Exemple de centroïdes mal placés sur cet extrait du fond de carte des communes de l'Ariège :



Dans cette zone de montagne, les villages sont au creux des vallées, or les centroïdes sont placés au milieu géométrique du polygone représentant le territoire communal, sur un point où l'on ne trouve généralement pas de population dans la réalité.

Les contraintes de réalisation et de diffusion

Les contraintes de réalisation sont celles qui vont limiter les possibilités de représentation à cause de manques ou d'inadéquation des outils, ainsi que les contraintes qui apparaissent lorsque l'on réalise une série de cartes.

Ces contraintes doivent si possible être connues à l'avance pour éviter de se lancer dans des travaux qui ne pourront être concrétisés par une belle carte...

Par exemple, certains logiciels ne savent pas calculer correctement un ratio surface/symbole et préfèrent utiliser des symboles dits "cliparts" (ou des polices de symboles) à la place de formes géométriques simples. Ce choix aboutit à une carte dont les symboles ne sont pas proportionnels entre eux selon leur surface, mais selon leur hauteur (une seule dimension au lieu de deux).

Ce problème est très répandu et des logiciels aussi courants que Mapinfo© sont concernés. La volonté de mettre à la disposition de l'utilisateur une vaste bibliothèque de symboles différents a primé sur la rigueur méthodologique de la représentation.

Les logiciels de cartographie automatiques présentent des possibilités variables d'un logiciel à l'autre, selon leur orientation (marketing, comme Microsoft MapPoint©) ou scientifique (comme Cabral© ou SCAP©).

Ainsi, avant de réaliser les travaux préparatoires de la carte, il faut essayer de repérer si le logiciel dont on dispose est capable de prendre en compte tous les besoins relevés lors des étapes précédentes : détermination de l'objectif, analyse de la variable et du fond de carte, repérage des contraintes.

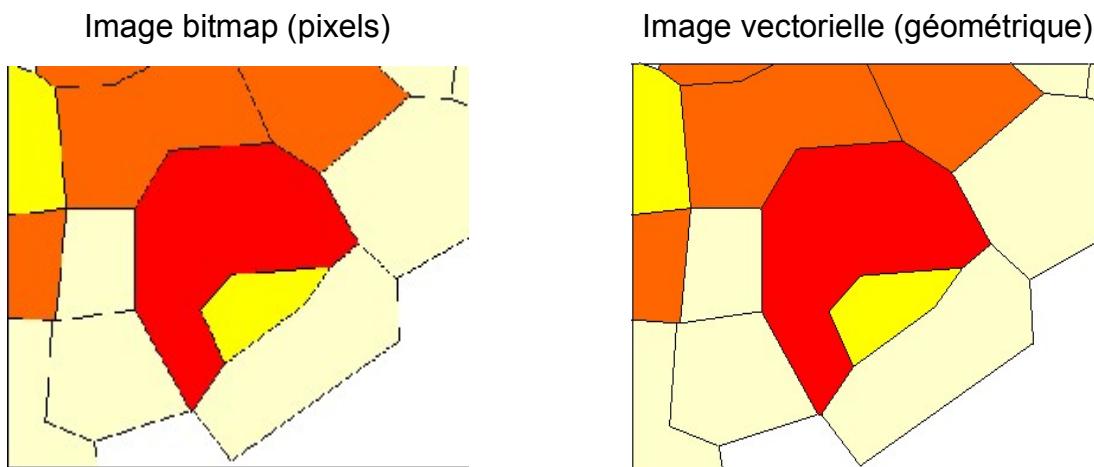
Les contraintes de série sont liées au fait que lorsque l'on réalise un document devant contenir plusieurs cartes, il faut suivre certaines règles pour essayer de produire un document harmonieux dans sa présentation :

- Utiliser toujours les mêmes styles de mise en page et de présentation pour les cartes
- Utiliser des couleurs différentes pour les symboles et si possibles liées à la variable étudiée (eau=bleu, forêt=vert, etc.)
- Dans le cas de cartes en vis à vis et/ou complémentaires, utiliser si possible le même ratio valeur/surface, pour que les symboles des légendes soient comparables.

Les contraintes de diffusion regroupent les limites à la créativité induites par le type de technique d'impression et de diffusion (reproduction) de la carte : définition de l'impression, n&b ou couleur, dimensions.

Autre problème possible, certains logiciels savent bien faire les cartes mais ont du mal à les exporter convenablement, dans des formats compatibles avec d'autres logiciels, et notamment les logiciels de mise en page et de publication assistée par ordinateur.

Les cartes à l'écran du logiciel sont en effet sous la forme de points colorés, les pixels. Ces points ne peuvent pas être utilisés pour un agrandissement de la carte pour l'impression par exemple, car on voit rapidement apparaître la forme des points qui composent l'image :



Par ailleurs, la technique de représentation des couleurs sur un écran d'ordinateur n'est pas la même que celle utilisée par une imprimante utilisant des mélanges d'encre. L'harmonie de couleurs affinée sur un écran a peu de chances de se retrouver à l'identique sur une impression papier classique.

b - Réalisation de cartes en symboles proportionnels : paramétrage et cartographie

Conseils généraux de réalisation

- La carte doit avant tout rendre lisible la variable, c'est l'information qui doit être perçue en premier lorsqu'on lit la carte. Synthétiquement on doit pouvoir mesurer l'importance de chaque information en mesurant la quantité d'encre utilisée pour la représenter, par rapport au fond blanc de la page / de l'écran.
- La carte doit avoir un objectif clair : description, exploration, mise en forme, illustration.

- La carte doit inciter le lecteur à se poser des questions à propos de la répartition spatiale de la variable, pas à propos de la méthode/technique de représentation utilisée.
- Essayez de réduire au maximum la déformation de la variable induite par la représentation graphique.
- Essayez d'utiliser le moins de place possible pour présenter le plus d'information.
- Essayez de montrer la variable à différents niveaux d'interprétation : globale, locale, structure fine.
- La carte doit être intégrée au document, liée aux textes qui l'entourent, servir une argumentation, une démonstration.

Technique du ratio valeur / surface : 1° - La taille du plus gros symbole.

Comment passer concrètement d'une série de valeurs numériques à un ensemble de symboles dessinés sur un fond de carte ?

Cette opération s'effectue automatiquement dans le logiciel de CCAO (et c'est là toute l'utilité de l'automatisation, avec le dessin) par le calcul d'un ratio, d'un coefficient, qui, une fois appliqué à une valeur de la série donne la *surface* du symbole la représentant graphiquement.

Attention, ici c'est bien la surface du symbole qui doit être proportionnelle à la valeur, car ce que l'on perçoit lorsqu'on regarde une carte c'est la présence de "taches" plus ou moins grandes qui se découpent sur le fond de la carte, et pas une mesure plus abstraite comme le rayon ou le diamètre des symboles.

Pourquoi ce rappel de techniques apprises lors des leçons du cours de Cartographie manuelle ? Tout simplement parce que les logiciels de CCAO ne savent pas forcément calculer des proportionnalités à la surface et trouver quelle sera la taille adéquate des symboles, donc la valeur du ratio.

En effet, les logiciels de CCAO attendent que vous leur fournissiez une correspondance entre une valeur de la variable et une taille de symbole (une surface). A partir de cette correspondance, par exemple 1200 habitants -> 200 pixels², ils vont en déduire le ratio et l'appliquer à toutes les autres valeurs. C'est beaucoup plus simple (parce que cela n'implique pas de calculs, juste un choix) que de demander à l'utilisateur un ratio.

- *C'est donc à vous de choisir la valeur de référence et la surface de symbole que vous allez lui affecter.*

Comment choisir ces valeurs ? Tout simplement en tenant compte des contraintes repérées précédemment.

En effet, c'est à ce moment de la méthodologie que les contraintes identifiées vont nous être utiles. Quelle est la principale contrainte de la représentation en symboles proportionnels : *la taille maximale de symbole dessinable à l'endroit de la valeur de variable la plus grande sur le fond de carte.*

Le choix de la valeur de référence du ratio valeur/surface doit donc se porter sur la valeur maximale de la variable, puisque l'on peut définir a priori quelle est la surface maximale du symbole.

Ainsi, on recherche où va se trouver la valeur maximale, et on évalue la taille maximale du symbole que l'on peut dessiner à cet endroit sans trop cacher le fond de carte, ni masquer les autres symboles.

Exemple :

Dans le cas d'un travail "à la main", on aurait procédé de cette manière, pour le dessin de cercles proportionnels :

Valeur	Surface	Rayon
987	200	7,98
457		
213		
124		
56		
10		

$$S = \pi \cdot r^2$$

$$r = \sqrt{\frac{S}{\pi}}$$

On choisit la valeur de 200 pixels² pour la taille maximale de symbole, après avoir repéré que le rayon maximal dessinable était de 8mm ($200 \approx 3,14 \times 8^2$).

On utilise ensuite ce ratio 987->200 pour calculer les surfaces des autres symboles (en proportionnalité simple, règle de trois : $457 * 200 / 987 = 92,6$), puis, à partir des surfaces, on obtient les rayons qui vont nous permettre de dessiner les symboles.

Valeur	Surface	Rayon
987	200	7,98
457	92,6	5,43
213	43,2	3,71
124	25,1	2,83
56	11,3	1,90
10	2,0	0,80

Dans le cas d'un travail en CCAO, il n'y a plus besoin de calculer à la main les valeurs de chaque rayon, ni de les dessiner. Mais il faut toujours définir quel va être le ratio qui sera utilisé pour passer des valeurs de la variable à une surface de symbole sur la carte.

Une fois le paramètre de relation valeur/surface entré dans le logiciel, la rapidité d'exécution du dessin de la carte peut être utilisée pour vérifier rapidement le résultat, sans préciser pour l'instant les autres paramètres de la représentation.

Cette vérification permettra aussi de vérifier la seconde contrainte, issue de la première, c'est à dire la taille du plus petit symbole découlant du choix du ratio effectué précédemment.

Technique du ratio valeur / surface : 2° - La taille du plus petit symbole

C'est la manifestation concrète de la contrainte de dispersion des valeurs de la variable (différence entre la plus forte et la plus faible valeur). Si la variable est trop dispersée, la surface du plus petit symbole découlant du choix du ratio à partir de la plus forte valeur risque d'être inadaptée.

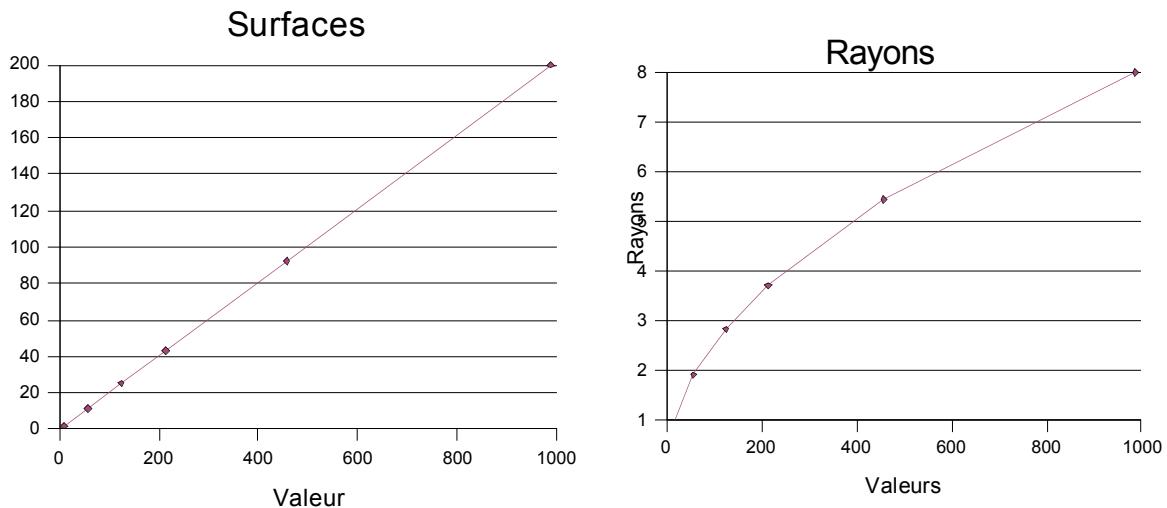
Deux solutions existent dans ce cas :

- l'utilisation d'un autre figuré pour le plus gros symbole, ce qui permet de le dessiner en utilisant une surface plus importante, donc de modifier le ratio (par exemple un figuré évidé);
- la transformation du ratio pour le faire passer d'un ratio simple, linéaire (valeur * ratio = surface), à un ratio plus complexe, logarithmique ou racine (log (valeur) * ratio = surface ou racine (valeur) * ratio = surface).

Dans tous les cas, toute utilisation d'un ratio non linéaire est à choisir en dernier recours, car on perd alors la liaison simple entre la valeur et les surfaces du symbole, donc le lecteur de la carte, qui s'attend à une proportionnalité simple, risque d'être trompé. C'est pourquoi on l'indique toujours en légende lorsque c'est le cas.

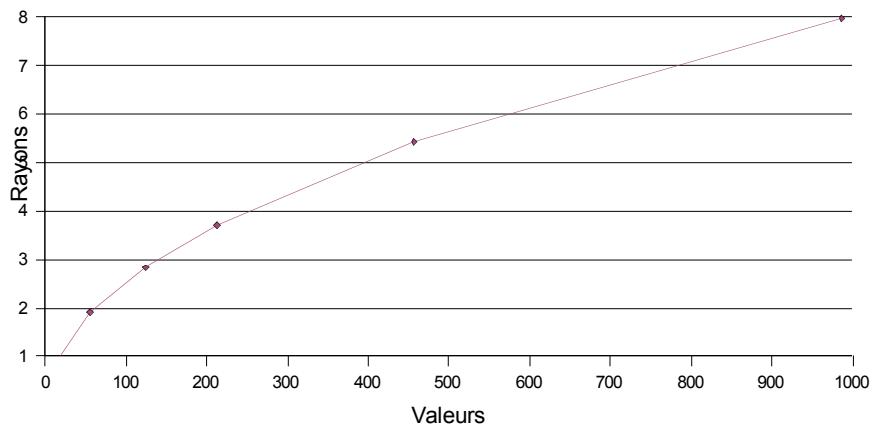
Astuce : Si le logiciel de CCAO ne vous permet pas de transformer directement le ratio pour utiliser une fonction racine ou logarithmique, il suffit de ne pas utiliser les valeurs normales de la variable, mais leurs racine ou leur logarithme népérien. Les symboles seront alors transformés, mais la légende sera fausse, il faudra la corriger.

Lorsque le ratio est simple, linéaire, voilà ce que donnent les courbes valeur/surface et valeur/rayon :

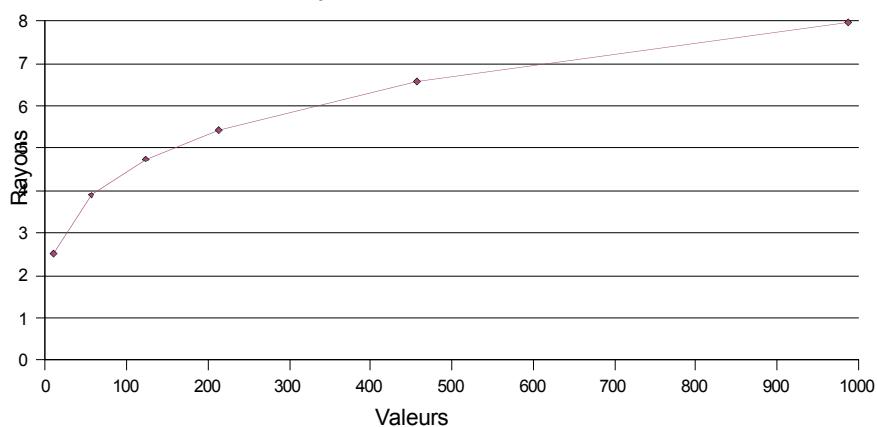


La première courbe est une droite : le ratio est linéaire. La seconde est une courbe qui ressemble à la courbe racine, c'est normal car il existe un rapport au carré entre la surface et le rayon ($S = \pi \cdot R^2$).

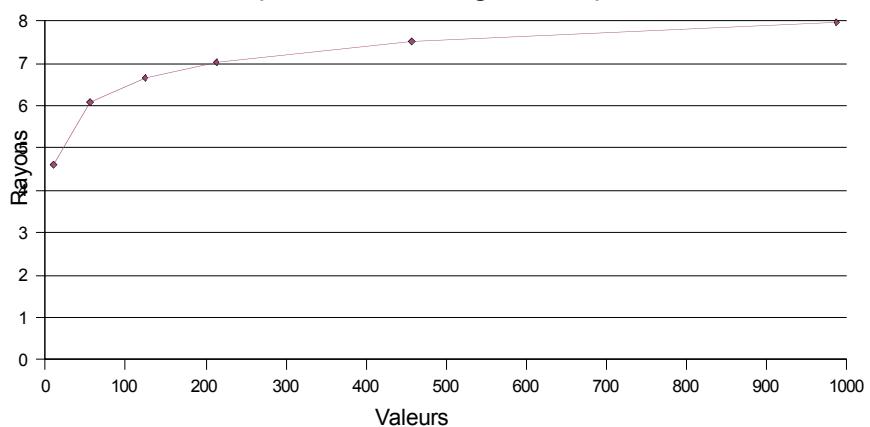
Proportionnalité linéaire



Proportionnalité racine

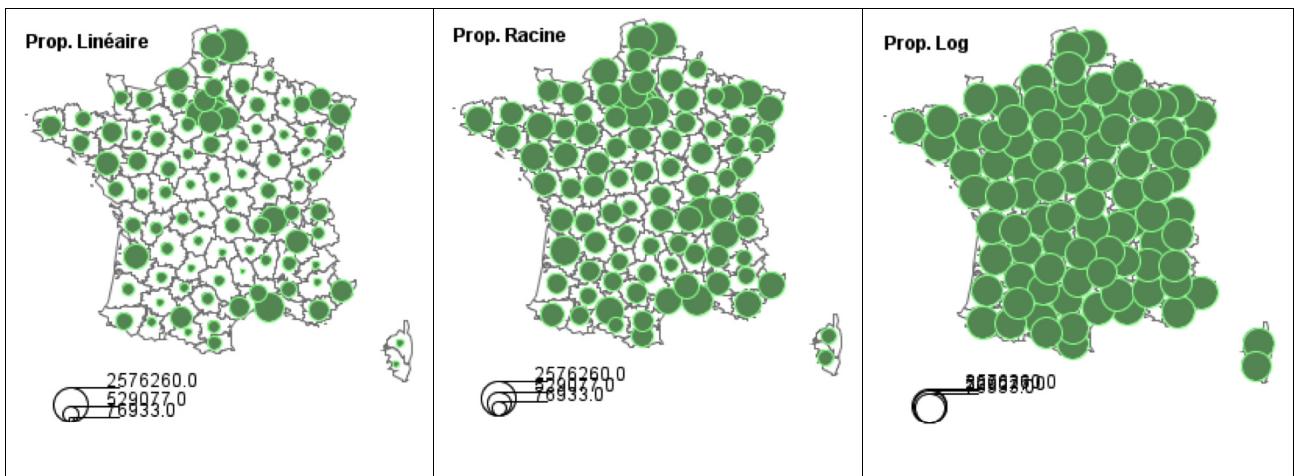


Proportionnalité Logarithmique



Lorsque le ratio est de type racine, la proportionnalité est transformée, la courbe valeur/surface n'est plus une ligne, mais une courbe racine, et la courbe valeur/rayon est un peu plus aplatie que dans l'exemple précédent. Lorsque le ratio est de type logarithmique, cette transformation est encore plus marquée.

Remarquez que le point (987,8) reste toujours fixe sur les trois séries de graphiques : c'est l'origine du ratio, la plus grande valeur/le plus gros symbole. C'est la pente de la courbe qui mène à ce point qui varie.



Une fois le calibrage des symboles défini et vérifié, reste à choisir les paramètres qui vont fixer la présentation générale de la carte.

Paramétrage de l'aspect de la représentation des symboles

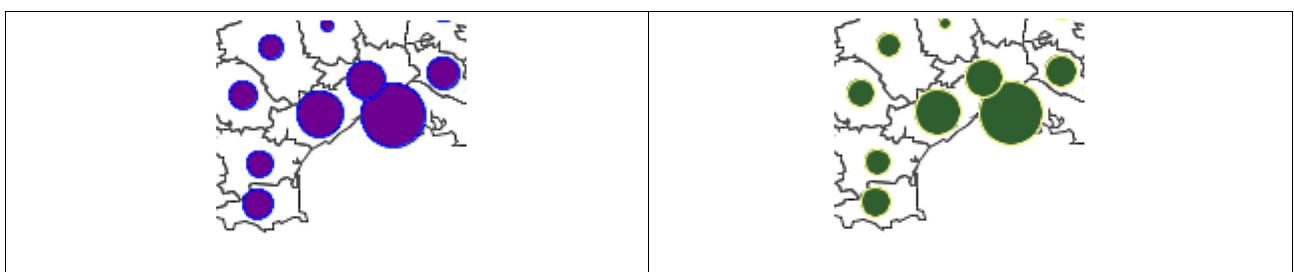
C'est à ce moment qu'interviennent les autres contraintes repérées lors de la leçon précédente : les contraintes de série et les contraintes de réalisation.

En effet, les logiciels de CCAO fonctionnant sur des écrans couleur, ils vous permettent d'utiliser les couleurs de votre choix et la finesse de trait la plus fine. Mais il va falloir tenir compte de la vie ultérieure de la carte :

- Dans quel type de document cette carte va t'elle être incluse ?
 - taille maximale de la carte
 - nombre et type de couleurs utilisables (tri ou quadrichromie)
 - épaisseur minimale des lignes/contours lisible
- Fait-elle partie d'une série de documents (carto)graphiques ?
 - Quel camaïeu de couleurs est utilisé ?
 - Doit-on faire partager les légendes à plusieurs cartes ?

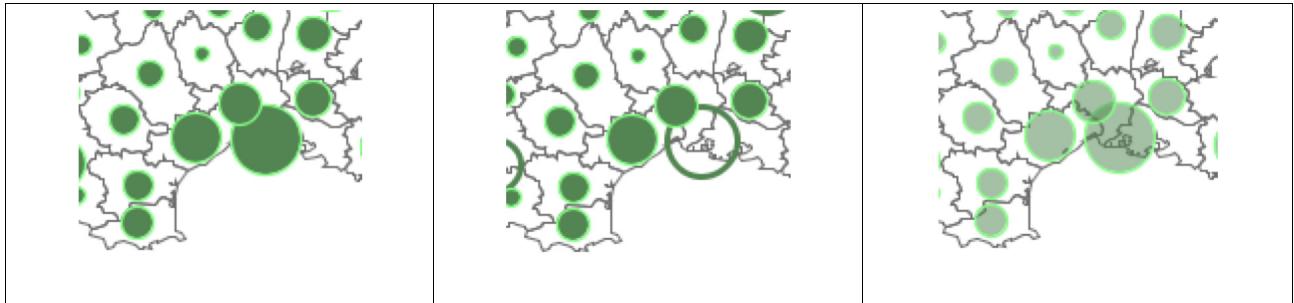
A partir de ces questions, on peut alors définir quelles couleurs utiliser pour les symboles et pour la matérialisation du fond de carte. Il existe certaines astuces de dessin pour présenter au mieux les symboles proportionnels :

- Utiliser si possible des couleurs bien contrastées pour le contour et le fond des symboles (pour bien repérer les symboles qui se superposent).



- Dans le cas de superpositions trop importantes et/ou de symboles trop grands, essayez

d'utiliser des symboles évidés, ou transparents.



- Dans le doute sur la possibilité d'utiliser la couleur, ne soyez pas timide d'utiliser le noir est blanc, qui, lorsque les nuances sont bien contrastées est très lisible.

c - Réalisation de cartes en symboles proportionnels : Présentation de la légende et mise en page

A quoi sert une légende ?

La légende est en fait la clé de lecture d'une carte.

Elle doit donner tous les renseignements nécessaires à la bonne compréhension de l'information de la carte, et ne pas induire en erreur le lecteur par des informations incomplètes ou ambiguës.

Une légende doit donc comporter :

- Un titre précisant l'unité dans laquelle est exprimée la variable.
- Une présentation graphique de la liaison variable / surface de symbole pour quelques valeurs représentatives de la variable, si possible comprenant le minimum et le maximum.

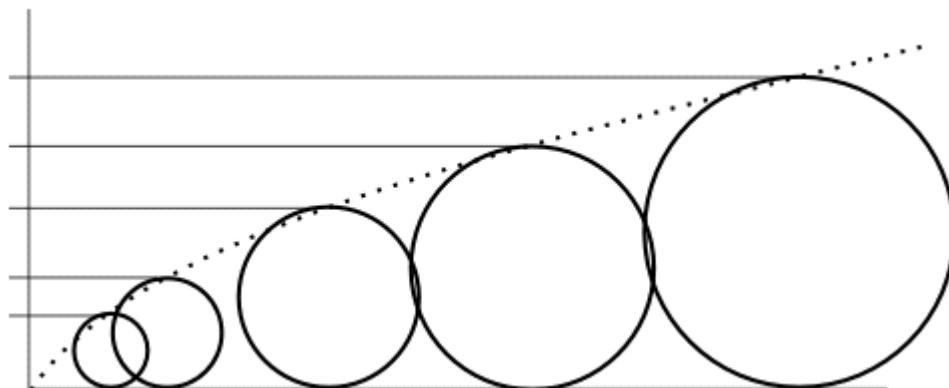
Ce n'est pas la peine de répéter le nom de la variable, le nom de la zone géographique représentée, la date de l'information en légende : c'est le rôle du titre de la carte.

En effet, une carte scientifique se doit d'être simple et claire, donnant toutes les clés de lecture dans le titre et dans la légende, sans commentaires superflus.

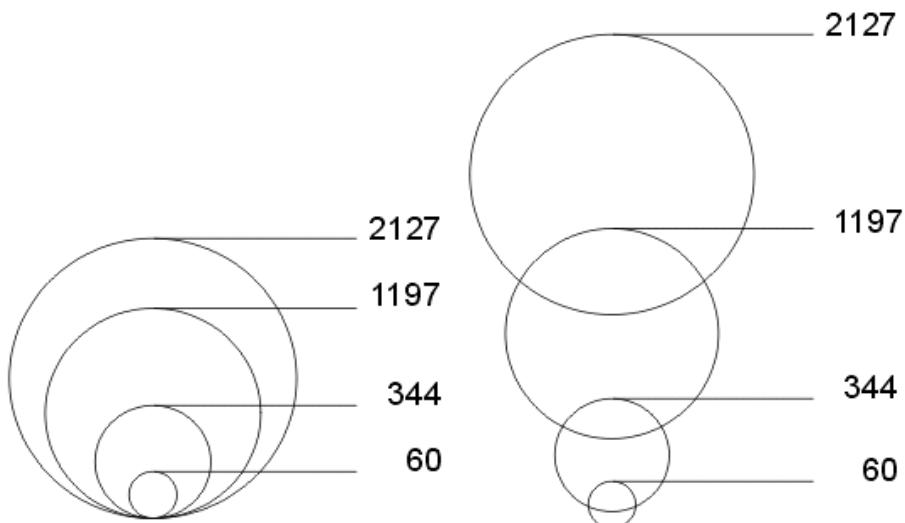
Présentation de la légende

Le problème des légendes des cartes en symboles proportionnels est qu'elles doivent présenter une sélection de valeurs (donc de symboles) représentatives de la variable. Souvent, cela se traduit par une légende prenant une place assez conséquente sur la page.

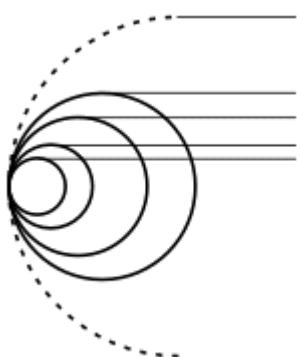
La présentation la plus claire est celle qui reprend l'idée d'un graphique, car elle matérialise la forme du ratio utilisé et informe donc le lecteur sur une éventuelle transformation du ratio :



Mais cette présentation a le défaut d'occuper beaucoup de place. On peut réduire la taille de la légende en évacuant le dessin de la courbe du ratio valeur/surface.



Parfois, le plus gros symbole est trop répété en légende, on peut alors soit le présenter en demi-symbole avec une teinte plus claire pour ne pas encombrer la carte, soit écrire en toutes lettres sa valeur en petits caractères.



Mise en page générale

La mise en page d'une carte comprend certains éléments que l'on placera autour du fond de carte. Mais retenez qu'une carte claire et agréable à lire est une carte simple, sans fioritures.

-> Une carte doit présenter, autour du fond de carte :

- Un titre qui répond aux questions : Quoi, Où, Quand ?
- Une légende qui renseigne sur l'unité utilisée et le ratio variable/surface du symbole
- Une échelle graphique (cf. exemples ci-dessous)
- Une flèche indiquant l'orientation de la carte si celle-ci n'est pas évidente (nord)
- Une indication de la source de la carte (pour informer sur la validité des données et leur date de mise au point)
- Une indication de l'auteur de la carte si celui-ci n'est pas celui du document global ou si la carte ne fait pas partie d'une série.

VI - Réalisation de cartes en plages de couleurs

a - Recherche des contraintes

Positionnement

Cette leçon est la première du chapitre concernant les cartes en plages de couleurs, dont le nom scientifique est choroplèthe.

Ce type de carte ne se réalise que dans la situation suivante :

- La variable statistique à représenter est relative, continue.
- L'implantation des individus géographique sur le fond de carte est zonale, ou possède une composante possédant une surface (exemple des cartes mixtes, cf. [la leçon sur les autres types de cartes](#)).

Méthodologie

Une fois la variable et le fond de carte prêts, le type de carte à réaliser défini, il ne reste plus qu'à passer à la réalisation concrète de la représentation cartographique.

- Celle-ci commence toujours par une recherche des différentes contraintes qui vont s'exprimer.

En effet, ce sont ces contraintes, qui, une fois repérées, vont guider les choix précis de représentation de la variable.

Une fois les contraintes identifiées, on pourra passer aux choix concernant les paramètres a priori et les paramètres spécifiques de la carte en plage de couleurs.

Les contraintes de la variable

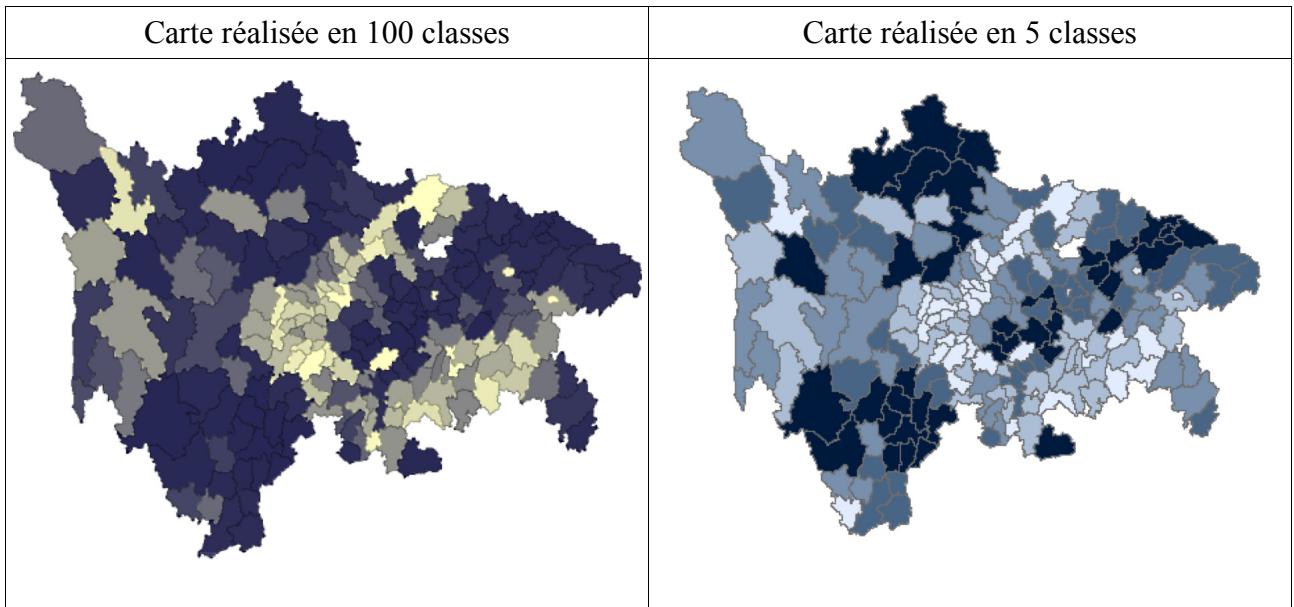
Le cas des cartes en plages de couleurs est particulier, car on a affaire à une variable quantitative continue à représenter sur une surface.

La sémiologie graphique de Jacques Bertin nous dit que les figurés zonaux à utiliser pour la meilleure « performance » de représentation sont les plages de valeur et par extension, de couleur. Ceci signifie que pour représenter au mieux une variable sur des surfaces, il est préférable d'utiliser des gradations de nuances de couleur(s).

La contrainte majeure qui se présente alors est celle de la vision humaine : en effet, si on appliquait le principe sémiologique directement, on obtiendrait alors une nuance différente par valeur de la variable, c'est à dire autant que nuances de couleurs qu'il y a d'individus sur la carte.

Or des chercheurs ont démontré que l'oeil humain ne distinguait (séparait) rapidement et de façon globale qu'au maximum de six à sept nuances différentes d'une teinte. En ergonomie, il existe même une norme (AFNOR, 1997) : « Si l'opérateur doit se rappeler la signification d'une couleur ou effectuer des recherches visuelles rapides basées sur la distinction des couleurs, il serait préférable de ne pas utiliser plus de 6 couleurs différentes ».

Comme l'objectif de la carte est de communiquer une information de manière efficace, rapide, il est donc préférable de la simplifier en tenant compte des contraintes posées par sa reproduction (impression, projection) et sa perception visuelle.



- Comment transformer une variable possédant x individus en une variable d'au maximum 7 groupes d'individus ? : en réalisant une classification, une discréétisation.

En effet, l'opération consistant à rendre discrète (avec un nombre précis de valeurs possibles) une variable continue (qui peut prendre n'importe quelle valeur dans une intervalle donné) s'appelle la discréétisation, et fera l'objet du prochain chapitre.

La variable elle-même peut ensuite poser d'autres contraintes. L'étendue et la dispersion ne sont pas ici des problèmes majeurs, ils seront à prendre en compte dans la création des classes.

Par contre, comme la transcription de l'information est basée sur des gammes de nuances, cette progression dans la couleur (clair -> foncé) implique une progression dans les valeurs des classes de la variable (faible -> élevées). Or on n'a pas toujours à faire à des variables présentant une seule progression dans ses valeurs.

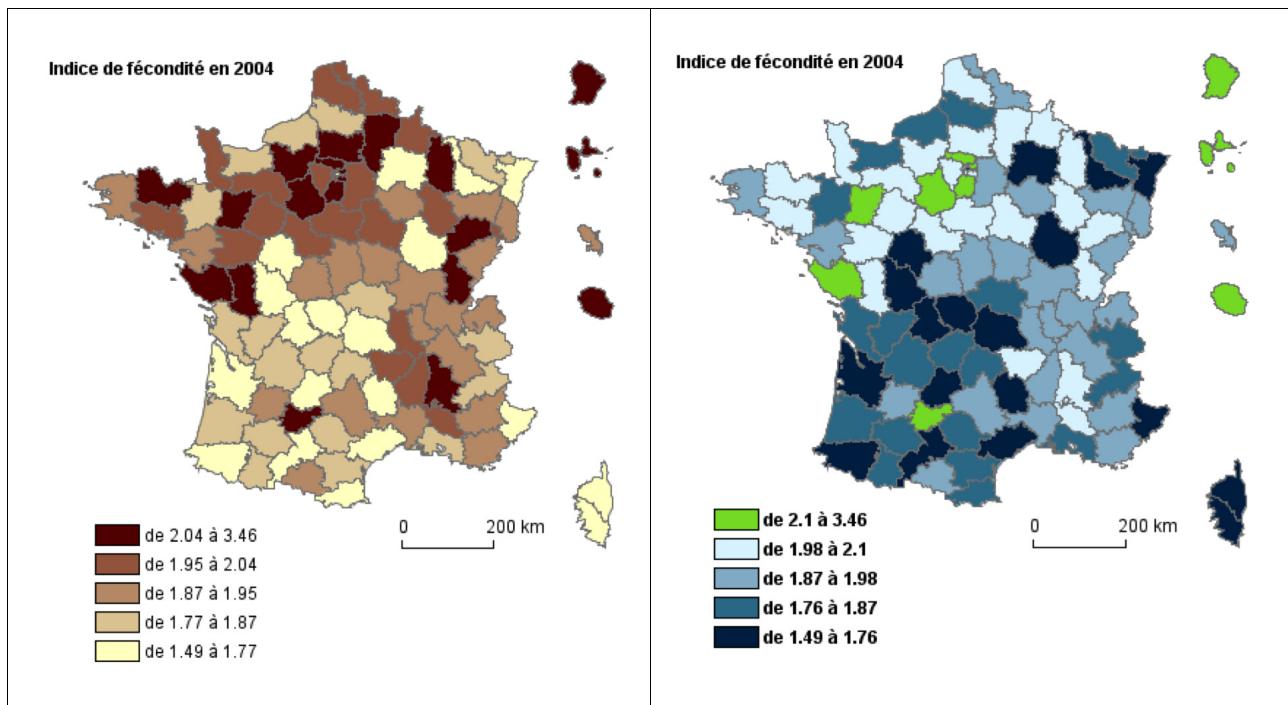
- Il faut prendre en compte la présence éventuelle de valeurs négatives, et/ou de valeurs seuils.

En effet, le guide principal de la représentation graphique étant de retranscrire le plus fidèlement possible la variable, il va falloir vérifier si cette dernière possède des divisions internes, des groupes de valeurs ayant un sens propre.

La plus classique de ces situations concerne la présence de valeurs négatives : on a alors deux progressions à partir de zéro, vers les valeur élevées et vers les valeurs faibles. N'oublions pas que les valeurs à représenter sont des valeurs relatives, des taux, on aura donc souvent affaire à des valeurs négatives.

Parfois, les variables possèdent une valeur faisant seuil, référence : l'indice 100, la moyenne, un taux plancher, etc. Ces valeurs seuils peuvent servir à découper la variable en deux sous-groupes, inférieur et supérieur. Ce qui rendra la lecture de la carte plus simple et surtout plus informative.

Par exemple lorsque l'on représente le taux de fécondité des femmes, la valeur seuil classique est celle de 2,1 enfants : celle du renouvellement de la population.



Les contraintes du fond de carte

En ce qui concerne les cartes choroplèthes, les contraintes imposées par le fond de carte sont moins importantes que pour les symboliques. En effet, le figuré étant ici directement dessiné à l'intérieur des polygones du fond, il n'y a pas d'effets de dépassement et de superposition à craindre.

Par contre, reste le problème des plus petits polygones colorés, qui doivent être visibles. Or les plus petites surfaces sont souvent celles qui vont recevoir les teintes les plus claires (et les plus grands noms, c'est très courant)

- Il faut que la teinte la plus claire de la gamme utilisée en légende soit nettement discernable de la couleur du fond de carte et facilement repérable sur la carte, même si elle se trouve dans de petites surfaces.

Ainsi, ne sélectionnez pas un découpage en classe aboutissant à une classe des valeurs les plus faibles ne possédant qu'un ou deux individus de très petite surface sur la carte.

Par extension, toujours pour la lisibilité, essayez de ne pas produire de cartes dont la répartition d'une classe n'est pas rapidement lisible (classes trop petites ou possédant des individus de trop faibles surfaces).

Les contraintes de réalisation et de diffusion

Comme la matérialisation de l'information statistique passe par une gamme de couleurs, il faut vérifier que la ou les teintes utilisées ne le soient que pour cet objectif. Séparer l'utilité de chaque couleur pour des objectifs différents permet de ne pas embrouiller la carte : une couleur neutre pour le fond de carte et l'habillage (légende, échelle, mentions), une couleur bien distincte et qui ressort bien pour l'information.

De par l'utilisation de figurés surfaciques, les contraintes de réalisation vont être importantes pour les choroplèthes. La sémiologie graphique nous indique que l'utilisation de la couleur est le meilleur moyen de retranscrire des variables différentielles, ce qui implique que l'ajout de différences de teinte à une gradation de nuance, de valeur, est intéressant.

Or la réalisation matérielle de cartes en couleur est encore assez problématique du point de vue technique, lorsque la diffusion du document se fait sur papier.

Le problème est que les logiciels de CCAO travaillant sur des écrans couleurs pouvant afficher des millions de nuances, et selon le principe des couleurs additives, ce que vous voyez à l'écran à de très fortes chances de ne pas ressembler beaucoup à ce qui va sortir de votre imprimante, qui utilise un système de couleurs très différent.

Non seulement la technique physique de représentation des couleurs change du tout au tout (couleurs soustractives), mais de plus la calibration des couleurs est un ensemble de techniques complexes difficiles à mettre en oeuvre sur du matériel grand public, même si ce dernier s'améliore progressivement.

Ainsi, votre écran peut être réglé d'une certaine manière, vos yeux percevoir les couleurs d'une manière différente et votre imprimante équilibrer les couleurs d'une troisième façon...

- Il faut donc encore une fois vous poser la question de l'avenir de votre carte : comment doit-elle être diffusée, reproduite, imprimée.

Nous verrons les astuces de dessin utilisables pour tenir compte de ces contraintes dans la leçon sur le paramétrage et la cartographie.

b- Le découpage en classes : la discréétisation

Position du problème

L'objectif d'une discréétisation est de ramener à au maximum sept classes (voire huit si l'on a deux progressions de couleur) une variable dont le nombre d'individus est changeant.

Les principes qui doivent guider la discréétisation visent à poursuivre les principes de la représentation graphique d'une variable statistique.

- Respect de l'information d'origine : la proportion des couleurs sur la carte doit être fortement liée à la répartition des individus dans la variable. L'image colorée que représente la carte doit donner une idée de l'information géographique contenue dans la variable.
- Il ne doit pas y avoir de classe vide.
- Il ne doit pas y avoir de classes non jointives (trou de valeurs entre deux classes successives).
- Les classes ne doivent pas se chevaucher (un individu ne peut appartenir qu'à une classe à la fois).

Il faut donc arriver à couvrir complètement la variable avec des classes jointives, tout en essayant de préserver la "forme" de l'information de départ.

Comment remplir ce premier objectif ? Comment savoir la forme de la répartition des individus dans la variable ? La méthode classique de la cartographie nous enjoint à observer de plus près la courbe des fréquences de la variable.

Le diagramme des fréquences

En CCAO, cette courbe des fréquences est réalisable de façon automatique, selon les logiciels (Cabral®, SCAP®, SCALP), ce qui va nous permettre de mieux visualiser la forme d'une variable. Ainsi, on aura un point de comparaison avec les discréétisation que vont nous fournir les méthodes de classification automatiques.

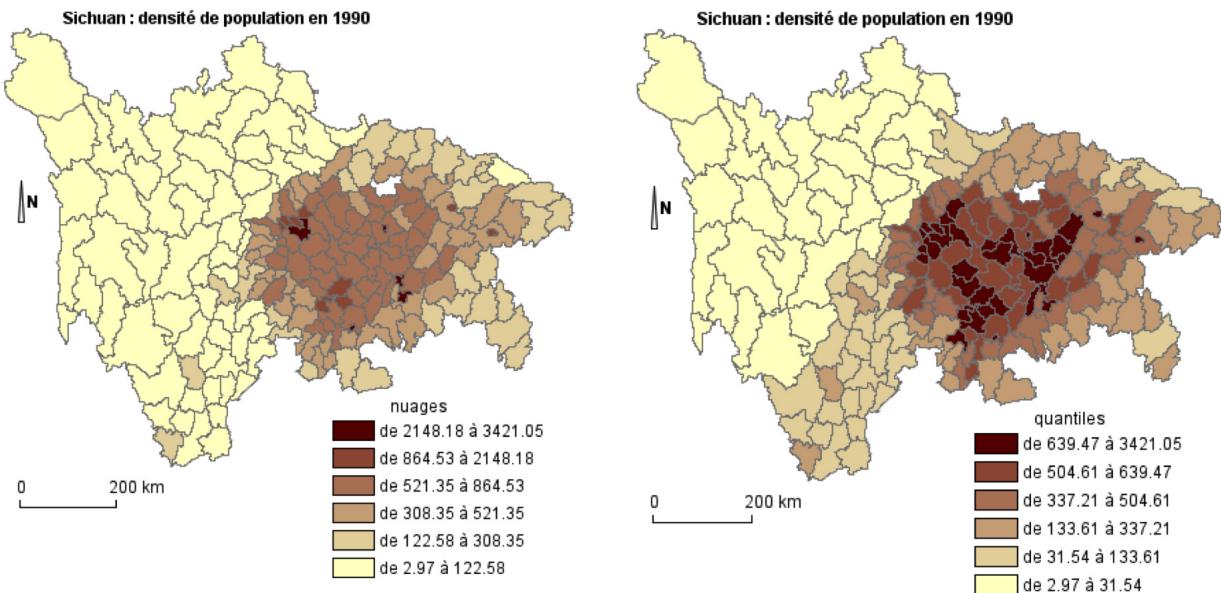
Si vous utilisez un logiciel de CCAO ne possédant pas cette fonction de dessin du diagramme des fréquences, il est envisageable de le réaliser avec un logiciel de type tableur.

La répartition des valeurs des individus dans la variable matérialisée par le graphique de fréquences va nous permettre d'en avoir une idée synthétique. C'est cette répartition qu'il va falloir essayer de reproduire avec le découpage en classes, si l'on veut préserver au maximum l'information de départ. Par exemple dans une carte de la densité de population du Sichuan en 1990, on va rapidement constater une grande hétérogénéité : beaucoup de faibles densités, peu de fortes densités qui se distinguent bien des autres valeurs. Si l'on ne choisit pas un bon découpage en classes, la carte risque d'être trompeuse.

-> Le graphique de fréquences nous donne donc une image de référence de la variable, qui va nous servir de point de comparaison avec les fréquences obtenues après discréétisation de la variable selon les différentes méthodes.

-> La méthode de discréétisation automatique dont le résultat ressemblera le plus aux fréquences de la variable d'origine sera donc celle qui sera la moins réductrice et la plus conforme.

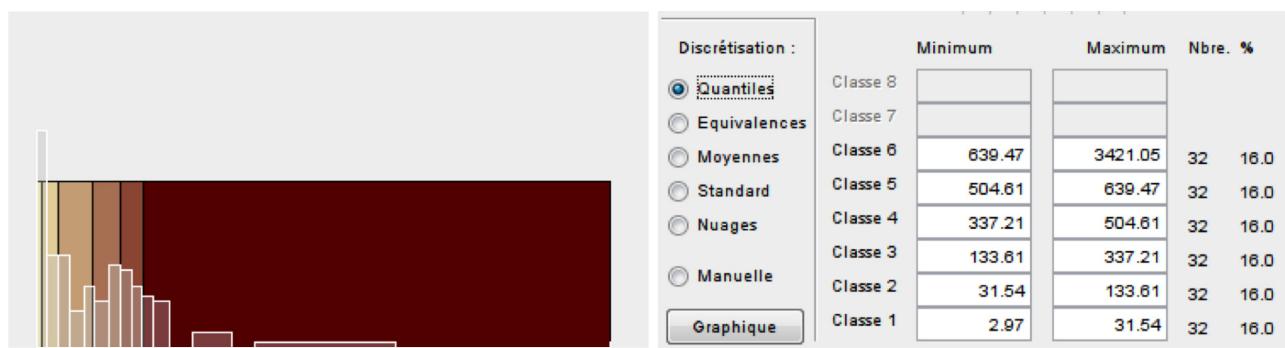
Entre ces deux cartes, seul le regroupement en classes varie, quelle est la plus juste ?



Les différentes méthodes de discréétisation automatique

Là encore, l'utilisation de l'ordinateur va nous permettre de réaliser rapidement des découpages en classes qui seraient autrement long voire impossible à réaliser. De plus, on va pouvoir comparer très rapidement le résultat de ces discréétisations avec le modèle du graphique des fréquences de la variable. La densité des contés du Sichuan en 1990 est utilisée pour les illustrations.

- Méthode des quantiles



Ce graphique présente en dégradé de rouge les classes formées par une discréétisation et en gris transparent les barres de l'histogramme des fréquences de la variable. Ce graphique est tiré du logiciel libre SCALP). En abscisses se trouve la variable, de son minima à son maxima, et en ordonnées on trouve le nombre d'individus dans les classes.

Ce sont donc deux discréétisations, l'une fine pour observer la distribution des individus, l'autre en quelques classes seulement pour réaliser une carte en plages de couleurs. Le but de la méthode revient donc à trouver la technique de regroupement et le nombre de classes appropriés pour que les deux histogrammes se ressemblent, dans leur forme.

Méthode :

La méthode des quantiles consiste tout simplement à utiliser le même nombre de valeurs par classe, ce qui produit donc des quantiles (4 classes : quartiles, 5 classes : quintiles, etc.)

Calcul :

Pour calculer le nombre de valeurs par classe, il suffit de diviser le nombre total de valeurs par le nombre de classes désiré. Ensuite on réalise les classes par comptage. On trie la variable dans l'ordre croissant et on « range » les individus dans les classes en fonction de leur rang.

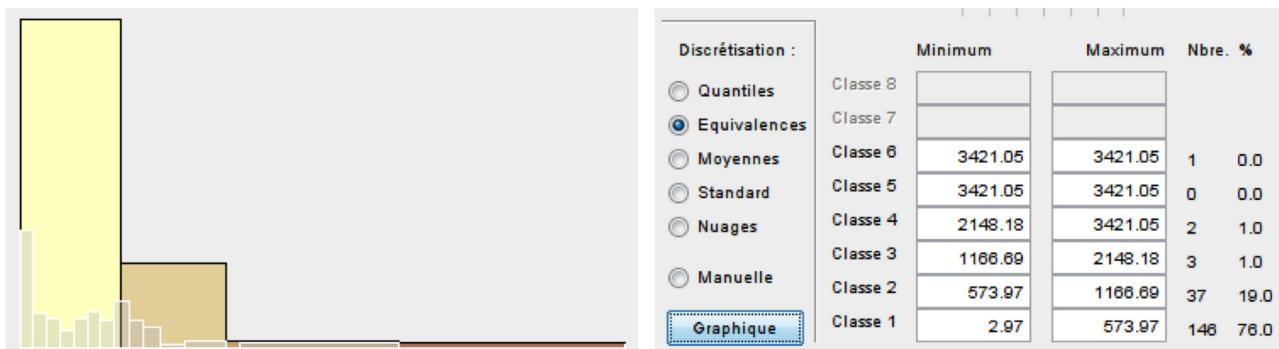
Objectif :

Cette méthode permet d'obtenir une carte dans laquelle les couleurs de la légende se répartissent à parts égales : donc une carte harmonieuse et facile à lire, si les polygones du fond de carte ne sont pas de tailles trop disparates. Dans le cas de séries de cartes sur le même fond, donc avec le même nombre de valeurs, les classes auront le même poids d'une carte à l'autre, ce qui facilite les comparaisons (chaque classe représente le même pourcentage de la variable totale).

Limites :

Mais cette méthode de discréétisation n'est pas liée à la variable cartographiée elle-même (la taille des classes n'est pas liée aux valeurs des individus mais uniquement à leur nombre), donc elle ne s'applique correctement qu'aux variables dont la forme de graphique des fréquences n'est pas trop tordue.

- **Méthode de l'équivalence**



Méthode :

Les équivalences consistent à déterminer des classes de taille égale, du point de vue de l'intervalle de valeurs. Ainsi, chaque classe couvrira une portion identique de l'étendue totale de la variable.

Calcul :

On calcule d'abord l'étendue totale de la variable (maximum-minimum), puis on divise cette étendue par le nombre de classes voulu. La valeur obtenue est l'intervalle de chacune des classes : la première classe ira du minimum à ce minimum plus l'étendue, etc. jusqu'au maximum de la variable.

Objectif :

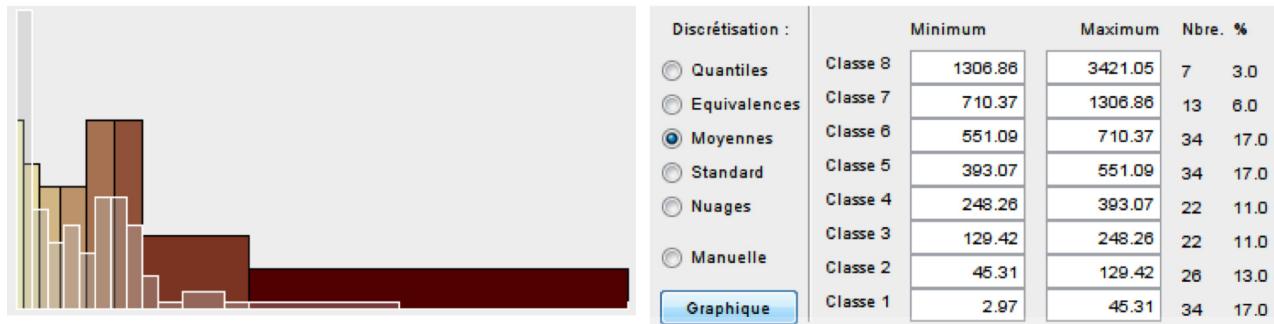
Cette méthode de discréétisation n'est pas liée à la variable, donc on peut l'utiliser pour comparer des variables distinctes, des phénomènes différents. On peut aussi y voir un moyen de mesure de la distorsion par rapport à une distribution uniforme.

Limites :

Cette méthode est peu utilisée à cause de limites importantes. La plus grave est qu'il arrive souvent qu'une classe ainsi définie ne comprenne aucune valeur, ce qui fausse la

discrétisation. C'est le cas dans l'exemple illustré ci-dessus. Ce problème se pose notamment lorsque l'étendue de la variable est tirée par des valeurs extrêmes, atypiques. Dans des cas moins graves, cette discrétisation à tendance à produire des classes de poids très inégal, et donc des cartes relativement hétérogènes.

- **Moyennes emboîtées**



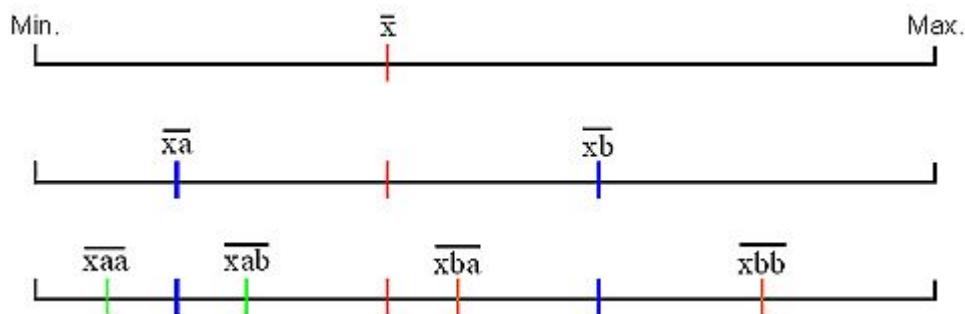
Méthode :

La discrétisation par les moyennes emboîtées consiste à découper la variable en utilisant des moyennes successives comme limites de classes.

Calcul :

Cette méthode ne peut produire que quatre ou huit classes exploitables. Voilà pourquoi, étape par étape :

- 1- Calculer la moyenne de la variable, M . Cette moyenne découpe la variable en deux classes, A et B, de part et d'autre.
- 2- Calculer la moyenne MA de la partie A de la variable. MA découpe la partie A en deux classes : A_1 et A_2 .
- 3- Calculer la moyenne MB de la partie B de la variable. MB découpe la partie B en deux classes : B_1 et B_2 .
- 4- On obtient donc quatre classes, A_1 , A_2 , B_1 et B_2 , en ayant calculé trois moyennes emboîtées.



Objectif :

Cette méthode a l'avantage d'être fortement liée à la distribution de la variable. Si vous réalisez une carte unique et que vous voulez "coller" aux mouvements de la distribution, cette méthode est généralement assez bien adaptée. Autre avantage : comme les limites des classes sont des moyennes, les cartes produites sont aisément compréhensibles.

Limites :

De par sa méthode de calcul, cette discréétisation peut produire des classes vides ou très hétérogènes, dans le cas de distributions très dissymétriques, voire plurimodales. Enfin, le nombre de classes produit peut ne pas être adapté au nombre d'individus de la variable.

- **Méthode de l'écart-type (ou standard)**



Méthode :

« Standard deviation » signifie écart-type en anglais. Cette discréétisation est basée sur la moyenne à laquelle on enlève ou on ajoute des parties ou des multiples d'écart-type.

Calcul :

On calcule d'abord la moyenne M de la variable, puis son écart-type E . Ensuite, suivant le nombre de classes voulu, on applique la procédure suivante :

3 classes:

Min, $m-0.5E$, $m+0.5E$, Max

5 classes:

Min, $m-1.5E$, $m-0.5E$, $m+0.5E$, $m+1.5E$, Max

7 classes:

Min, $m-2.5E$, $m-1.5E$, $m-0.5E$, $m+0.5E$, $m+1.5E$, $m+2.5E$, Max

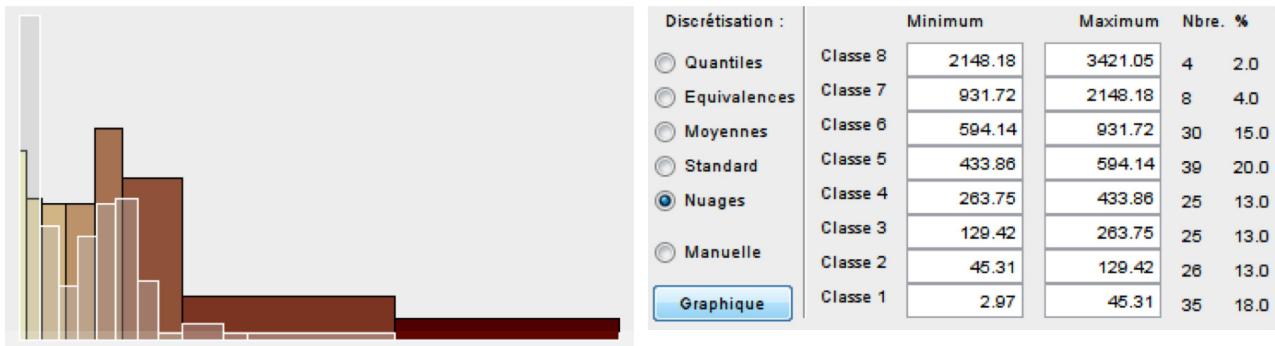
Objectif :

Cette méthode est basée sur des indicateurs statistiques de centralité et de dispersion. Elle permet d'obtenir des séries de cartes comparables entre elles, avec un découpage en classes sensible à la dispersion de la variable. Vous remarquerez que les classes centrales sont plus "courtes" que les classes externes : cette méthode s'applique correctement sur des distributions normales, en courbe de Gauss.

Limites :

C'est encore une méthode très liée à la variable et à sa dispersion. Donc à ne pas utiliser sur des dispersions trop grandes, des dissymétries trop importantes, sous peine de classes vides. De plus, comme pour les moyennes emboîtées, cette méthode s'organise autour d'une moyenne centrale. Donc elle ne fonctionnera pas très bien avec des distributions plurimodales.

- Méthode des nuages de points



Méthode :

La méthode des nuages de points est l'une des plus intéressantes pour l'analyse spatiale, car elle permet d'obtenir des classes très proches des formes de la distribution de la variable, en créant des classes de valeurs homogènes mais les classes les plus hétérogènes entre elles.

Calcul :

Le calcul de cette discréétisation est particulier : le logiciel part d'une discréétisation en quantiles pour le nombre de classes spécifié par l'utilisateur, et va faire varier les bornes de ces classes incrémentiellement pour s'approcher de l'objectif.

Objectif :

De par sa méthode de calcul, elle est mieux adaptée aux variables multimodales (plusieurs groupes), et possédant un nombre d'individus assez important pour pouvoir faire ressortir des seuils. Cette méthode est un équivalent mathématique assez intéressant de la méthode manuelle des seuils observés, ou de la méthode dite de Jenks.

Limites :

Cette méthode est gourmande en calcul et donc en puissance processeur. Elle fait partie des discréétisations statistiques liées à la variable, donc si elle est très performante sur les cartes uniques, elle n'est pas très utile pour une série de cartes.

c – Paramétrage et cartographie

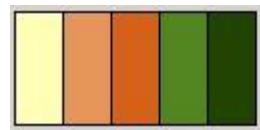
Choix des gammes de couleur

On a vu dans la leçon précédente, sur les discrétisations, comment prendre en compte une bonne partie des contraintes de la variable. On va s'intéresser maintenant aux problèmes posés par l'utilisation de couleurs dans un document graphique.

- Le principe d'utilisation de la couleur est de lier une variation d'intensité, de densité d'une teinte à la répartition des valeurs de la variable dans les classes.

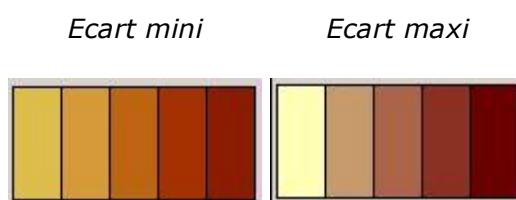
Ainsi, si on a une discrétisation en équivalence, avec le même écart entre les centres de classes, on devrait utiliser le même écart de densité de nuance entre les tons de légende.

- On utilise toujours une teinte par progression dans la variable. Une seule progression de valeur doit être transcrise par une seule progression de densité dans une seule teinte, une seule couleur.



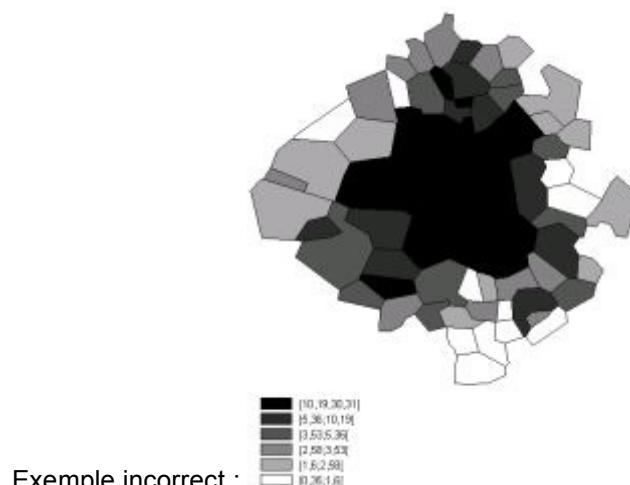
Dans les faits, cette liaison valeur/couleur, si elle est théoriquement indispensable, est assez difficile à réaliser lorsque les classes sont disparates.

On essaye de maximiser la différenciation de deux tons successifs, pour rendre la lecture de la carte plus simple.



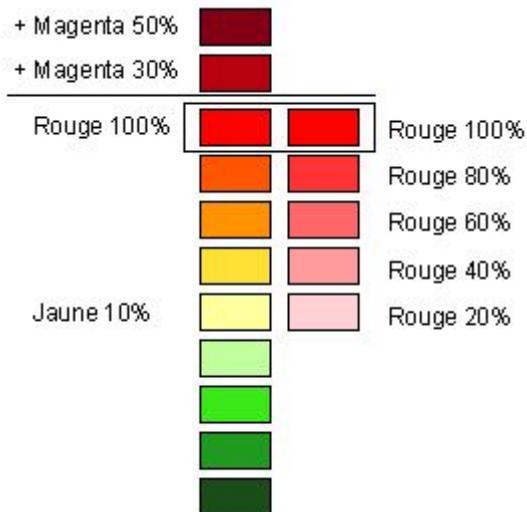
En général, on affecte donc la teinte la plus claire possible (mais distinguable sans ambiguïté du fond, blanc) à la classe la plus faible, et la teinte la plus foncée, sombre, dense (mais distincte du noir qui est réservé au dessin du fond et à l'écriture) à la classe la plus forte.

- Les couleurs de la variable doivent se distinguer clairement du blanc de la page et du noir du fond de carte et des textes.



Avec certains logiciels de CCAO il est possible de contourner le problème en changeant la

couleur des polygones dessinant le fond de carte pour les individus devant recevoir les couleurs les plus sombres de la palette. Ainsi, on peut utiliser des couleurs proches du noir sans que le fond de carte disparaîtse, voire utiliser un contour utilisant un gris.



Dans la réalisation de palettes en deux progressions, il existe une tradition, une habitude de présentation qui veut que l'on utilise une couleur chaude (rouge, jaune, marron) pour représenter les variations positives et une couleur froide (bleu, vert, gris) pour représenter les négatives.

Le jaune est souvent utilisé comme point de départ car il fait bien la liaison entre le rouge et le vert et permet de donner plus de luminosité aux teintes faibles.

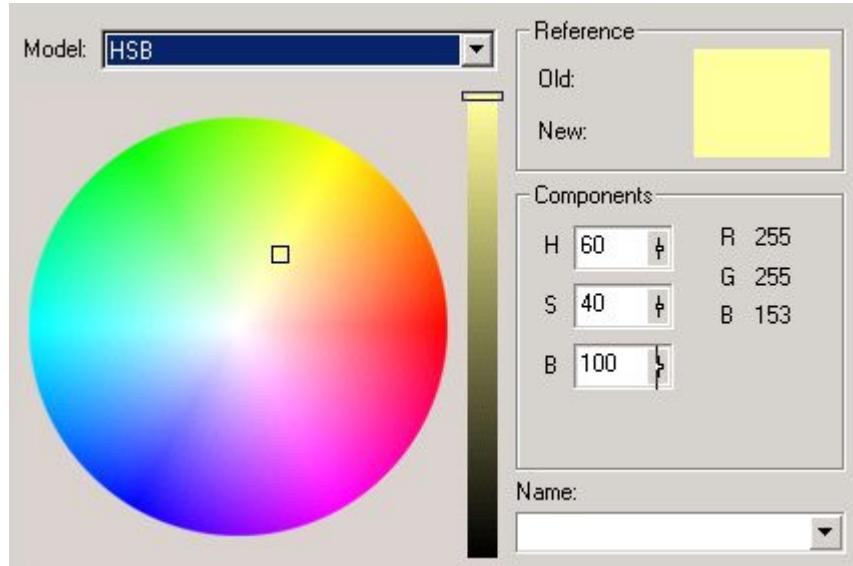
Le même principe de complément de la couleur primaire par une couleur secondaire est utilisé lorsque l'on veut renforcer la différenciation des teintes successives lorsque l'on étend la palette de couleurs. Ici on a créé deux classes supplémentaires en ajoutant du magenta au rouge.



Les pourcentages indiqués correspondent à la densité de la teinte principale. Vous pouvez retrouver ces valeurs dans les logiciels de dessin en utilisant une gamme colorimétrique de

type TSL : Teinte, Saturation, Luminosité (HSB, hue saturation brillance, en anglais).

Mode HSB (CSL basé sur la roue) dans CorelDraw© :



- Essayez d'utiliser toujours des couleurs simples et le moins de teintes différentes possibles : cela garantira une meilleure conformité des couleurs à l'impression (ainsi qu'un tarif moins cher chez l'imprimeur).

La contrainte du noir et blanc

Dans de nombreux cas, on ne sait pas exactement quel va être le moyen d'impression utilisé pour la carte, et la méthode de duplication. On doit alors utiliser la gamme la plus passe-partout, celle dont on est quasiment sur qu'elle passera à l'impression et à la photocopie : le noir et blanc tramé.

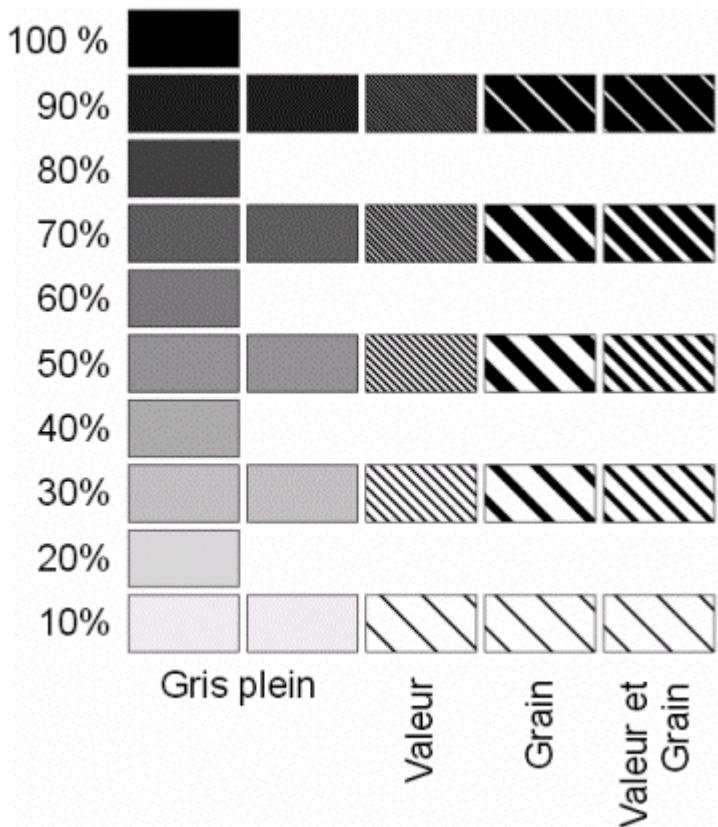
L'utilisation du noir et blanc suppose quelques règles supplémentaires et une contrainte majeure de réalisation, lorsque l'on veut représenter deux progressions.

Un dégradé de gris doit utiliser au maximum l'intervalle entre le gris très clair et le gris très foncé, sans être confondu avec le blanc de la page et le noir du fond. Or les techniques de reproduction les plus répandues ont une fâcheuse tendance à transformer les nuances de gris.

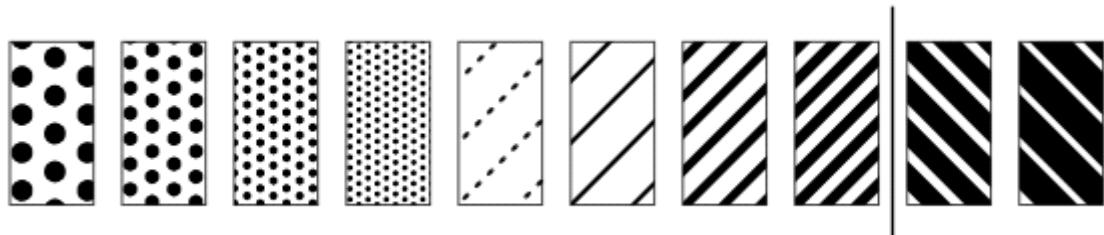
En effet, les techniques d'impression et de reproduction ont des effets différents sur les niveaux de nuance utilisés, et souvent ils ont des limites à ne pas franchir si l'on veut que les dégradés "passent" bien.

- N'utilisez pas de gris sous les 10% ni au-dessus des 85%.

Le pire des moyens de reproduction d'une carte, la photocopie, réduit fortement le contraste d'un document, surtout si les nuances de gris ne sont pas tramées, c'est à dire composée de motifs. La sémiologie graphique de Bertin nous apporte des moyens performants de réalisation de trames en noir et blanc, principalement la variation de valeur (% de figurés noirs sur le fond blanc) et de grain (largeur des figurés noirs sur le fond blanc).



Si l'on veut représenter un variable présentant des sous catégories marquées, il faut alors utiliser d'autres méthodes de dessin, voire d'autres figurés ("variables graphiques").



Tracés et aplats

Une autre contrainte de dessin assez méconnue est la largeur minimale du trait. En effet, selon la technique d'impression retenue, la largeur du plus fin trait imprimable varie beaucoup. Le problème est que là encore, les logiciels de CCAO et surtout les logiciels de dessin nous proposent par défaut des traits souvent trop minces pour l'impression.

Les imprimantes de bureau ont beaucoup évolué et sont capables de performances que peu d'imprimeurs réalisent aux tarifs normaux, sans supplément. Ainsi, un tramage de 0.1mm se verront sûrement très bien sur votre imprimante, mais une fois imprimés et/ou photocopiés, c'est la catastrophe.

- Utilisez des traits d'au minimum 0.2 mm, toujours en noir 100%.

Dans le même domaine, il existe une technique simple pour éviter que les nuances en aplats de gris ne soient transformées par la duplication de votre carte. Les photocopieuses, noir et blanc comme couleur, ont souvent des problèmes pour reproduire les aplats continus sans les maculer de moirés et de taches. Les photocopieurs noir et blanc sont de même optimisés pour le texte (2 couleurs) et pas pour les dégradés de gris.

La solution est simple en théorie, mais parfois difficile à mettre en oeuvre...

- Utilisez des aplats tramés et non pleins.

Voici la différence (le tramage est grossi pour être bien visible) :

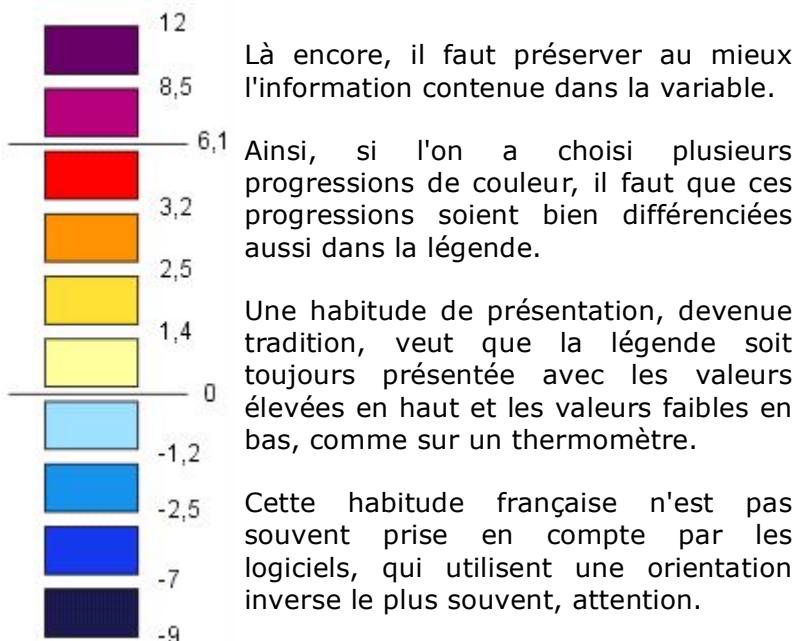
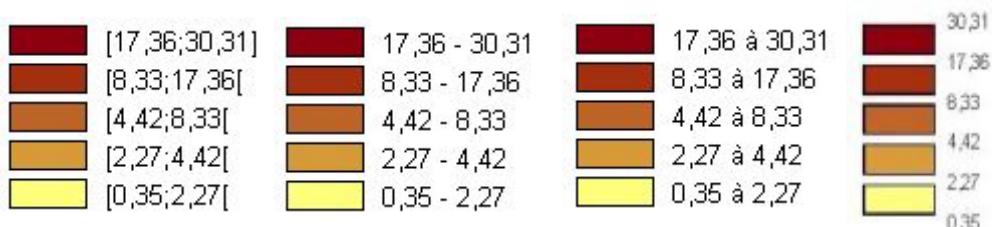


A l'œil, à distance normale, la différence est minime car on ne distingue plus les points formant le tramage, mais la photocopieuse, elle, reproduira bien les points, sans faire d'erreurs de contraste puisqu'on reste à deux couleurs : noir et blanc.

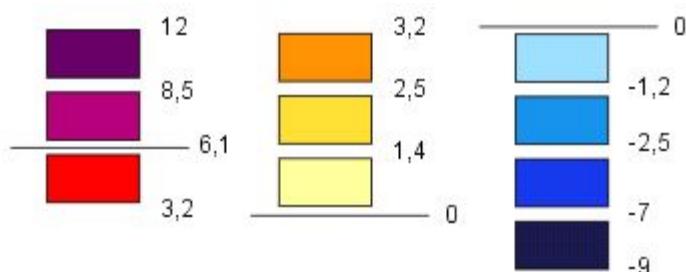
d – Présentation de la légende et mise en page du document cartographique.

En ce qui concerne les intitulés des caissons de légende (les rectangles colorés d'une légende de carte choroplète s'appellent des caissons), il existe plusieurs solutions de présentation.

On peut présenter ces intitulés de manière mathématiquement juste, mais c'est parfois un peu difficile à lire (selon le public et l'objectif donné à la carte), il faut alors simplifier :



En cas de manque de place pour la légende, on peut toujours la présenter horizontalement :

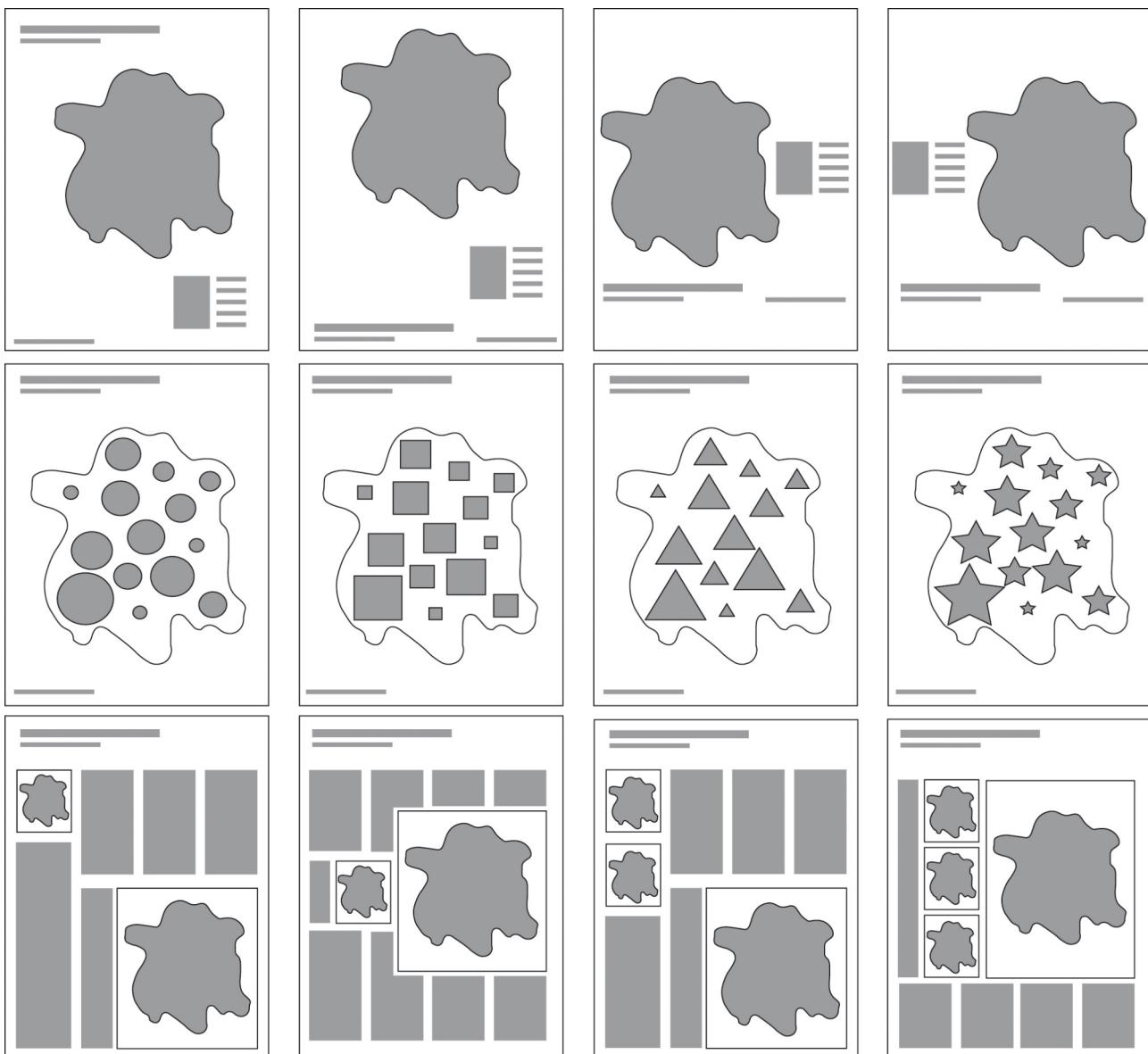


Mise en page générale

La mise en page d'une carte comprend certains éléments que l'on placera autour du fond de carte. Mais retenez qu'une carte claire et agréable à lire est une carte simple, sans fioritures.

Une carte doit présenter, autour du fond de carte :

- Un titre qui répond aux questions : Quoi, Où, Quand ?
- Une légende qui renseigne sur l'unité utilisée et le ratio variable/surface du symbole
- Une échelle graphique (cf. exemples ci-dessous)
- Une flèche indiquant l'orientation de la carte si celle-ci n'est pas évidente (nord)
- Une indication de la source de la carte (pour informer sur la validité des données et leur date de mise au point)
- Une indication de l'auteur de la carte si celui-ci n'est pas celui du document global ou si la carte ne fait pas partie d'une série.



VII – Cartes composées et autre types de représentations cartographiques

Les cartes composées

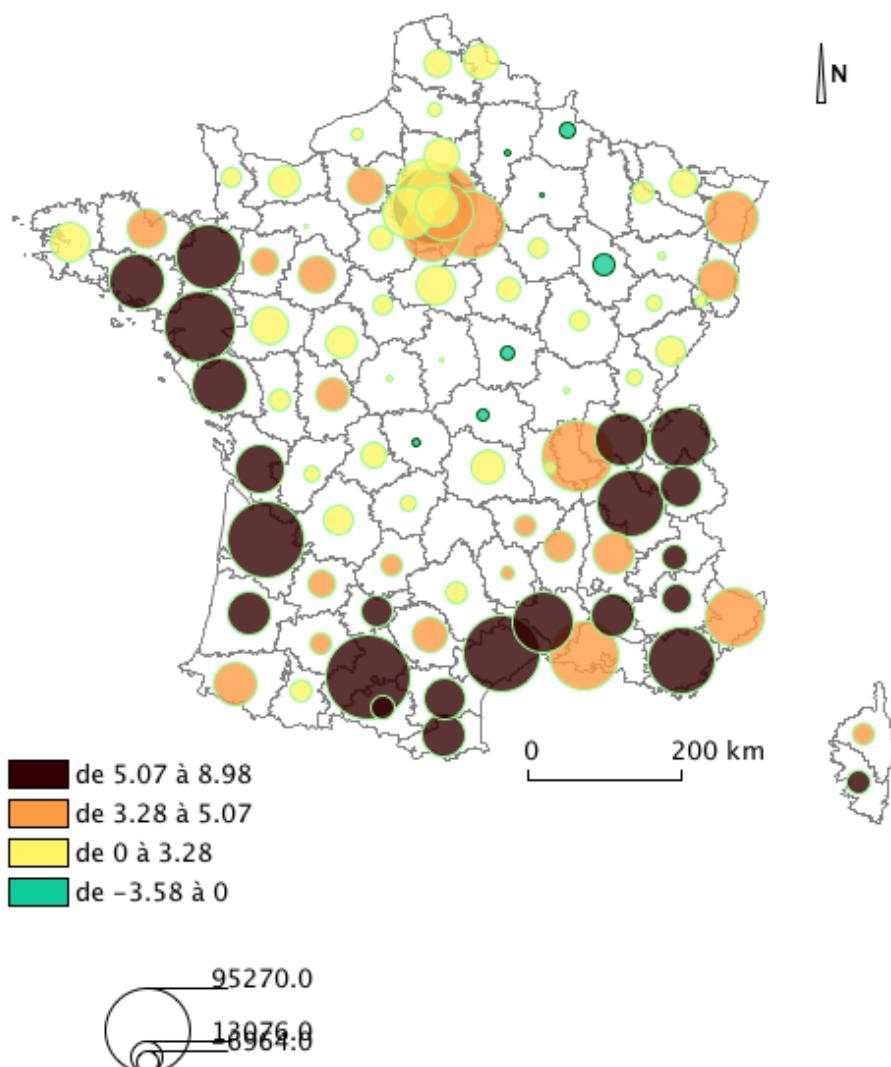
Les cartes dites composées ou doubles sont en fait des cartes présentant deux variables sur le même fond de carte.

Elles se présentent sous deux formes :

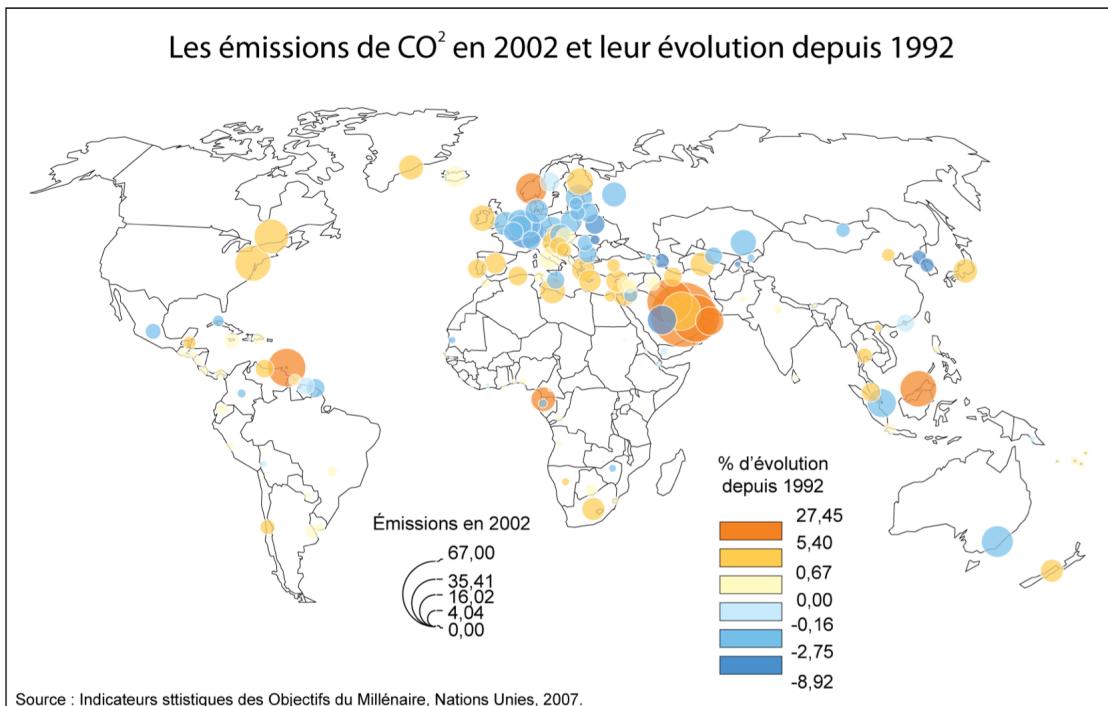
- Une symbolique par-dessus une carte en plages de couleurs
- Une chropoplèthe dont les couleurs sont en fait à l'intérieur des symboles d'une symbolique

Cela suppose naturellement que les deux phénomènes représentés ensemble sont complémentaires, ou au moins thématiquement reliés. C'est un type de carte particulièrement adapté lorsque l'on veut représenter l'aspect absolu ET l'aspect relatif d'une variable, par exemple la variation de population absolue entre deux dates et la part que cette variation représente par rapport à la période (taux d'évolution).

Évolution de la population entre 1995 et 2000



Mais attention, si ces cartes sont intéressantes pour gagner de la place et permettre des comparaisons rapides, elles cumulent les contraintes des deux types de cartes. Le problème principal est de préserver la visibilité de chaque représentation partout, donc de ne pas cacher la choroplèthe avec les symboles ou de ne pas utiliser de symboles trop petits pour qu'on puisse en voir la couleur.



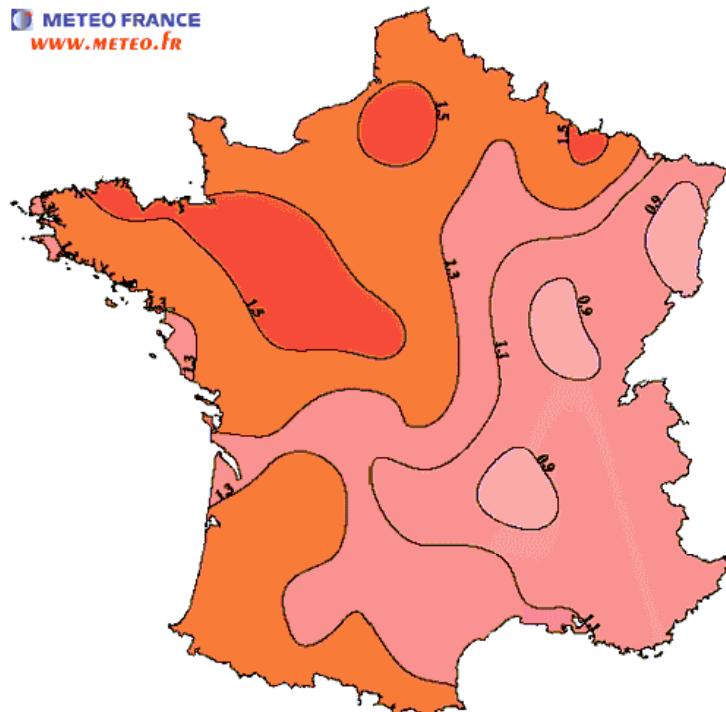
Les cartes en auréoles ou lissées

Les cartes lissées ont pour objectif principal de proposer une visualisation continue dans l'espace d'un phénomène alors que l'on n'en possède que des éléments ponctuels.

Cela sous-entend naturellement que l'on sache par avance que le phénomène à représenter est continu dans l'espace (qu'il possède une valeur distincte en tout point de l'espace), mais que l'on a pu faire qu'une quantité limitée de relevés sur le terrain.

C'est le type de carte intéressant pour représenter les phénomènes naturels généraux, comme en météorologie, en géologie, en biogéographie...

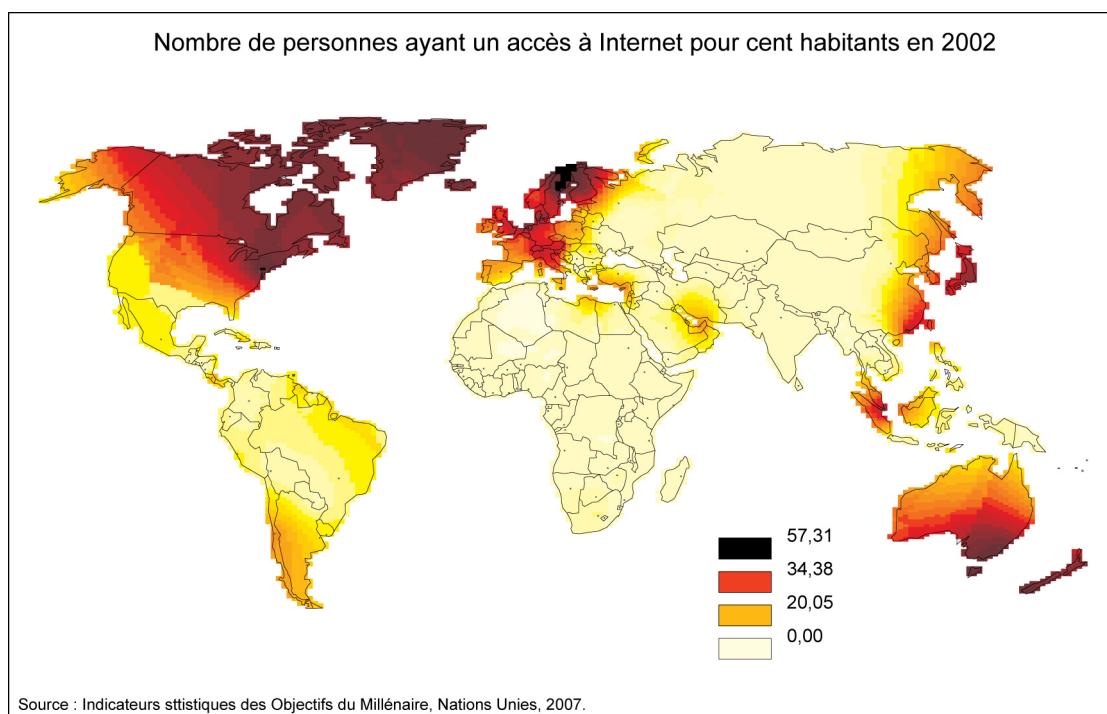
Par exemple, lorsque l'on veut cartographier les variations de température sur une zone, alors que l'on ne possède que les valeurs des stations météorologiques implantées dans cette zone, on est obligés de recourir à ce type de carte.



Evolution des températures minimales en France au cours du 20ème siècle
Tendances (en °C/siècle) 1901-2000 à partir de 70 séries de températures minimales

Ce type de représentation fonctionne bien dans le cas de phénomènes continus dans l'espace et de relativement faible variabilité spatiale. Mais les logiciels vous permettent de réaliser des cartes lissées à partir de toute variable quantitative, même si ses valeurs sont en fait discontinues dans l'espace, ou des moyennes par unité administrative.

Par exemple, ce lissage du nombre d'accès à Internet pour cent habitants, réalisé avec le logiciel PhilCarto (gratuit et proposant de nombreux types de représentations cartographiques, dont les lissages).



La coloration de la Californie aux USA est par exemple totalement faussée par la proximité de

la valeur plus faible du Mexique. De même pour le Groenland, qui prend la valeur du Danemark. Il faut donc faire attention à l'adaptation de la méthode à la réelle répartition spatiale des individus de la variable.

Pour pouvoir dessiner des plages de couleurs continues, ou des courbes concentriques du type courbes de niveau, il faut que le logiciel utilisé interpole les valeurs de la variable aux endroits où il n'y a pas de valeur connue. Cette création de valeurs est faite le plus souvent par interpolation, en utilisant des méthodes mathématiques plus ou moins complexes, sur une grille plus ou moins fine de points encadrant les localisations des valeurs connues.

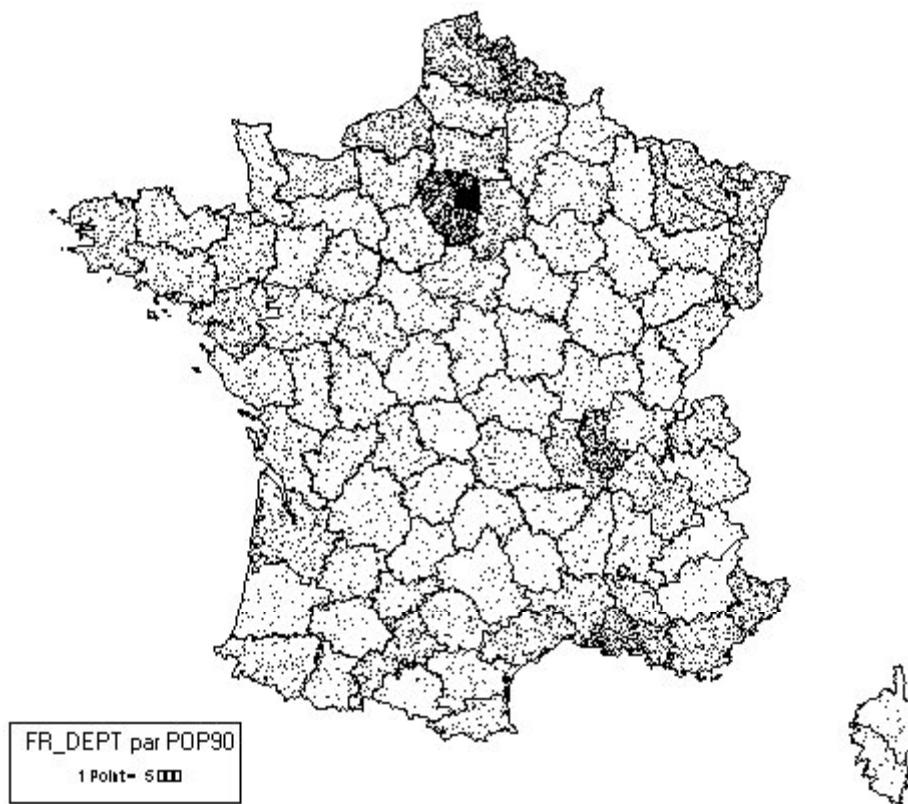
La méthode la plus simple consiste, pour un point donné de la grille d'interpolation, à faire une moyenne des valeurs des points connus, en pondérant cette moyenne par le carré de l'inverse de la distance entre chaque point connu et le point interpolé. Cette méthode, à cause de la formule employée, s'appelle la méthode gravitaire (la gravité exercée par un objet diminue comme l'inverse du carré de la distance à cet objet).

Il existe d'autres méthodes d'interpolation de valeurs sur une grille, faisant intervenir les notions de probabilités et de statistiques, qui ont des résultats plus intéressants.

Mais il faut garder à l'esprit que ces cartes présentent des données interpolées, donc fausses (non réelles) par définition, et que la méthode choisie influence fortement sur la valeur prise par un point précis. Donc c'est à la fois créer de l'information et la simplifier pour la représenter, ce qui nous éloigne un peu des valeurs réelles...

Les cartes en densités de points

Ce type de carte présente l'information sous la forme de semis de points réguliers ou aléatoires, dont la densité (le nombre par unité de surface) traduit la valeur de la variable dans le polygone observé.



On utilise cette méthode pour des variables absolues, car on définit une valeur pour chaque point. C'est assez étonnant de représenter une variable absolue sur une surface, mais en fait ce sont des points, des symboles ponctuels, qui sont répartis sur cette surface, et c'est leur nombre qui représente l'information graphiquement, pour donner une idée de densité.

Le problème majeur de ce type de carte est que l'on a tendance à inconsciemment associer la localisation d'un point avec celle de la variable, puisque c'est une carte. Or, dans ce cas précis, la localisation d'un point n'a pas de sens, elle est choisie au hasard sur une surface donnée, pour obtenir une dispersion maximale.

Dans les zones où la densité de points est faible, on se retrouve donc avec peu de points, et là la tentation est grande de donner un sens à la localisation de chaque point, comme on a l'habitude de voir les cartes en symboles proportionnels.

Exemple de carte en densités de points peu interprétable :

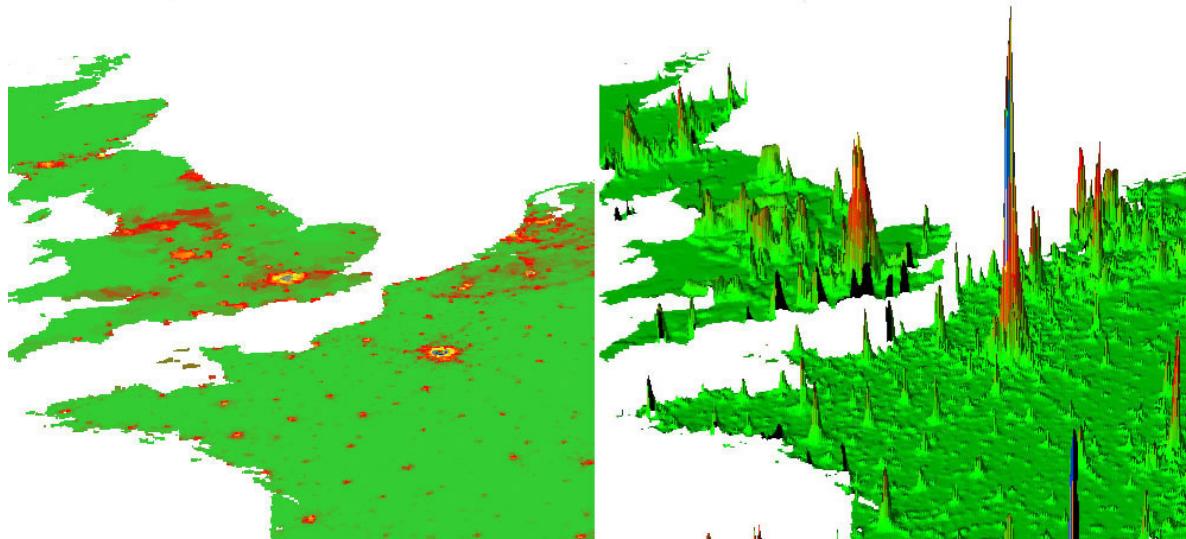


Les cartes en trois dimensions

L'utilisation de la troisième dimension pour représenter graphiquement une variable statistique est aujourd'hui rendue possible par certains logiciels. Le principe de construction est simple : au lieu d'utiliser une « variable visuelle » ou « variable rétinienne » (selon le terme de J. Bertin¹) pour figurer la variable, on va utiliser directement la troisième dimension, en concevant la représentation directement en relief. La restitution du résultat s'effectue ensuite en repassant de l'espace à trois dimensions à un support plan à deux dimensions, par des techniques de dessin classiques en architecture (perspective, isométrie...) ou par des techniques de modélisation informatiques.

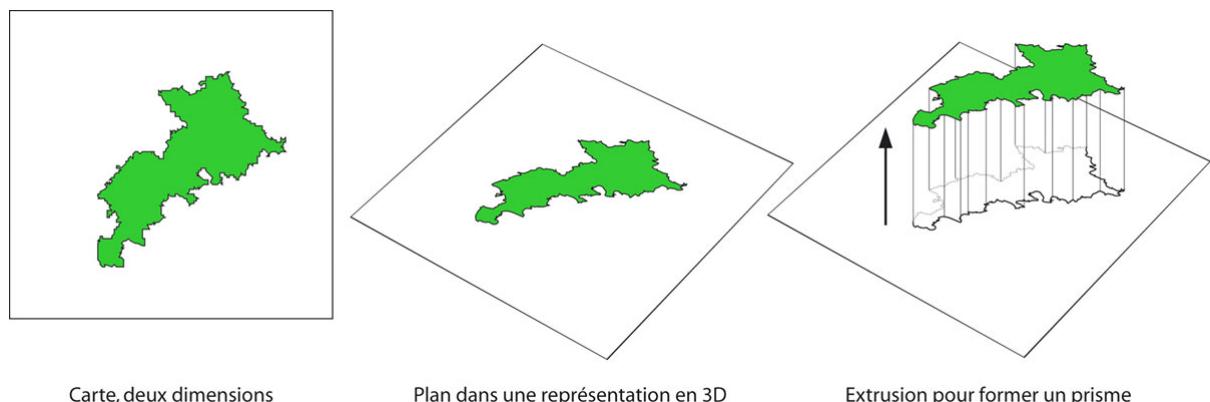
¹ Bertin J., 1977.

Exemple des densités de population en Europe en 2000, carte lissée en 2D, et carte en relief 3D.



Dans le cas d'une variable discontinue dans l'espace, on va alors extruder les polygones représentant les unités statistiques, pour former des prismes.

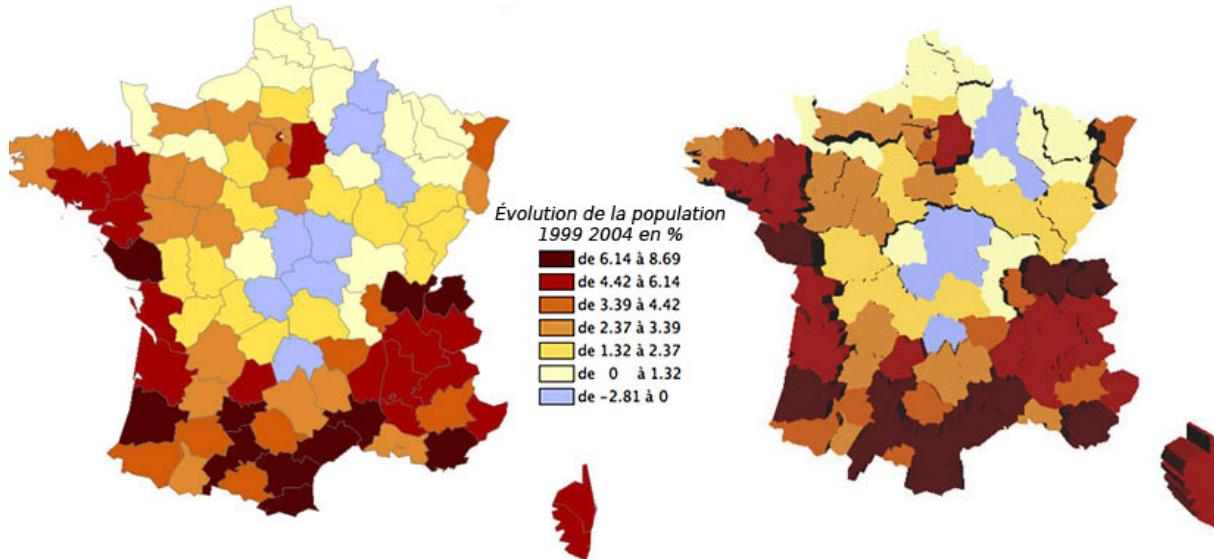
Procédé d'extrusion.



Si l'observateur de la carte en relief est situé à sa verticale et regarde dans la direction nadirale², l'image résultante ne se démarque pas beaucoup d'une carte 2D classique. On remarque quelques ombrages et distorsions, notamment sur les parties périphériques du fond de carte, la Corse notamment, cf. figures ci-après. Ce type de carte est appellé « carte en prismes ».

² L'observateur situé au-dessus de la carte regarde verticalement vers elle.

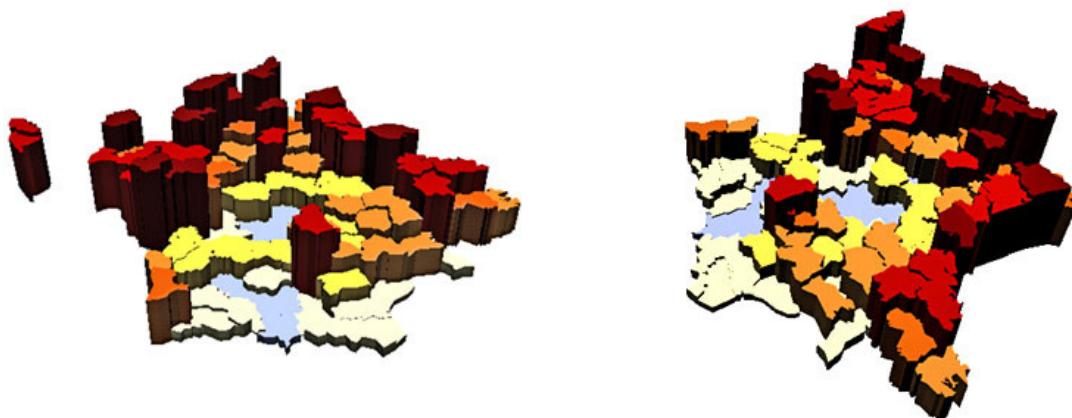
Extrusion d'une analyse thématique choroplète, l'exemple de l'évolution de la population des départements français entre 1999 et 2004 (estimations annuelles INSEE).



L'apport du passage aux trois dimensions réside tout d'abord dans une image, une représentation visuelle, qui est plus immédiatement perceptible. La vision d'un relief, d'un paysage, fait appel à des capacités d'interprétation beaucoup plus basiques que la lecture d'une carte en deux dimensions, abstraction qu'il faut interpréter.

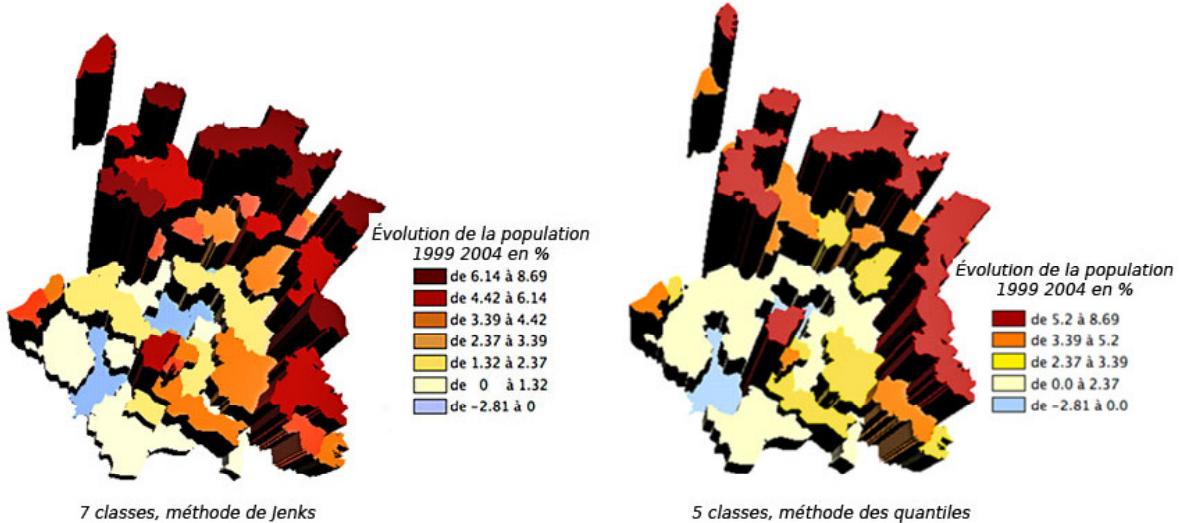
Un intérêt plus pratique de ce type de représentation apparaît lorsque l'on fait pivoter la carte prismatique autour d'un point central (figures ci-dessous).

Deux vues obtenues par pivot sur trois dimensions.



Comme la partie sud de la France, et plus précisément les départements situés sur les côtes et frontières, présente les taux d'évolution les plus importants, se forme une sorte de « mur », qui va gêner la visualisation si l'on utilise un positionnement « classique » du nord de la carte vers le haut de l'image. En pivotant le nord vers l'observateur, tout en abaissant le point de vue, on obtient une représentation beaucoup plus éloquente.

Cette représentation permet aussi d'apprécier visuellement de la perte d'information due à la discréétisation d'une variable, ou de comparer deux discréétisations différentes d'une même variable statistique.



Mais ce type de représentation reste difficile à réaliser, et surtout à rendre lisible à un public qui ne reconnaît pas forcément le fond de carte une fois celui-ci pivoté. Les cartes en prismes « manipulables » restent un bon outil pour le géographe, dans la représentation des variables statistiques sur ordinateur.