

UTILISATION DE LA CARTOGRAPHIE EN VUE DE LA CONSTRUCTION D'ÉCHANTILLONS

LE CAS DE L'ÉCHANTILLON MAÎTRE ET DE L'ÉCHANTILLON EMPLOI 1999

G. BOURDALLE

INSEE, Unité Méthodes Statistiques

Une des premières étapes en vue de la constitution d'échantillons de logements est la constitution des unités de tirage (Unités primaires pour l'Échantillon-Maître, unités de 1^{ère} phase pour l'Échantillon Emploi). L'utilisation d'un outil cartographique, de par ses possibilités de collecte, de gestion, de manipulation, d'analyse et d'affichage des données, nous a semblé adapté pour ce travail.

Trois types d'utilisation de l'outil ont été faites :

- Constitution des unités de tirage
- Aide au choix des échantillons tirés
- Diffusion des résultats

1. Échantillon-Maître 1999

Plusieurs étapes sont nécessaires à la constitution de l'Échantillon-Maître 1999.

Une des premières étapes est la constitution des unités de tirage appelées unités primaires (UP).

Ces unités primaires peuvent être :

- pour la strate de gestion 0 ($TU99=0$), des communes ou regroupements de communes
- pour la strate de gestion 1 ($TU99=1,2,3$), l'unité urbaine ou un regroupement d'unités urbaines
- pour les strates 2 ($TU99=4,5$) et 3 ($TU99=6,7$), l'unité urbaine
- pour la strate 4 ($TU99=8$), l'unité urbaine de Paris

La création de ces unités primaires se fait en fonction de règles bien précises. Une des règles concerne le nombre de résidences principales. Pour la strate de gestion 0, ce nombre doit être compris entre 1800 et 3600, et pour les autres strates le nombre de résidences principales doit être au moins égal à 1800.

Du fait que la plupart des communes rurales et que certaines unités urbaines de strate 1 ont moins de 1800 résidences principales, un travail de regroupement a été nécessaire.

Au recensement de la population de 1999, les communes de la France métropolitaine se répartissent comme suit:

| <i>Tranche de taille de l'unité urbaine</i> | <i>Nb de communes</i> | <i>Poids (en %)</i> |
|--|-----------------------|---------------------|
| Commune rurale (TU99=0) | 30611 | 83.7 |
| Unité urbaine de moins de 5 000 hab. (TU99=1) | 1547 | 4.2 |
| Unité urbaine de 5 000 à 9 999 hab. (TU99=2) | 931 | 2.5 |
| Unité urbaine de 10 000 à 19 999 hab. (TU99=3) | 689 | 1.9 |
| Unité urbaine de 20 000 à 49 999 hab. (TU99=4) | 602 | 1.6 |
| Unité urbaine de 50 000 à 99 999 hab. (TU99=5) | 495 | 1.4 |
| Unité urbaine de 100 000 à 199 999 hab (TU99=6) | 320 | 0.9 |
| Unité urbaine de 200 000 à 1 999 999 hab. (TU99=7) | 974 | 2.7 |
| Unité urbaine de Paris (TU99=8) | 396 | 1.1 |

D'un point de vue théorique, ces travaux de regroupement sont simples, d'un point de vue pratique et vu le nombre de communes à traiter, ils peuvent être extrêmement lourds si nous n'utilisons pas un outil cartographique.

1.1 Choix de l'outil cartographique

Le logiciel cartographique choisi est MapInfo. Quatre raisons importantes ont guidé notre choix : la convivialité du logiciel, l'existence d'un langage de programmation (MapBasic), l'exportation possible des données vers d'autres logiciels, en autres SAS, logiciel statistique nous ayant permis en amont de préparer les fichiers stratifiés et en aval de procéder aux différents contrôles, et enfin la disponibilité à la fois du logiciel et des compétences humaines dans les directions régionales. En effet, l'INSEE a choisi le logiciel MapInfo comme Système d'Information Géographique (S.I.G) pour les études et la diffusion.

L'apport d'un outil cartographique nous a permis à la fois de pouvoir effectuer ces opérations de regroupement en un temps relativement faible et surtout de pouvoir contrôler efficacement les travaux faits. Deux types de contrôles ont pu être faits. Premièrement, s'assurer que les 30611 communes de strate 0 et les 3167 de strate 1 sont bien attribuées à une unité primaire et, deuxièmement, s'assurer que les seuils en nombre de résidences principales ne sont pas dépassés.

L'autre application de cet outil cartographique a été de cartographier régionalement les unités primaires tirées, en visualisant en même temps les unités primaires de l'ancien Échantillon-Maître appelé EM90 et les unités de première phase de l'Enquête Emploi ; ceci permet aux directions régionales d'avoir une vision d'ensemble de leurs nouveaux échantillons afin de préparer au mieux leurs réseaux d'enquêteurs.

1.2 Descriptif de l'application

L'application informatique est écrite en MapBasic. Elle se présente sous la forme d'un CDROM contenant tous les fonds de cartes nécessaires (communal, cantonal, départemental, régional et national).

L'interface utilisateur est composée de deux fenêtres : la carte proprement dite et un tableau contenant différentes variables statistiques relatives aux cantons, aux communes ou aux unités urbaines. Les variables affichées seront le nombre de logements principaux ainsi que différentes autres variables auxiliaires (données de population ou fiscales).

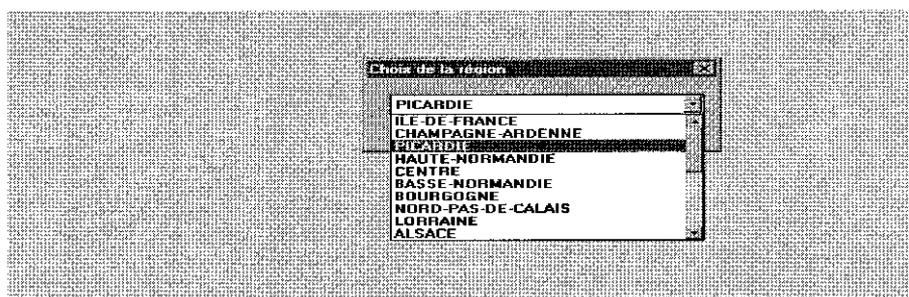
Pour rendre l'application plus conviviale et aussi aider l'opérateur dans les divers choix de regroupement ou de fractionnement, un fonds de couleur caractérise l'entité (canton ou unité urbaine) selon qu'elle contient moins de 1800 résidences principales, entre 1800 et 3600, ou plus de 3600 résidences principales. De plus, l'opérateur peut connaître à tout instant le nombre exact de logements principaux de l'unité primaire qu'il est en train de constituer.

Ce principe de fonds de couleur et de variable cumulative de nombre de résidences est aussi disponible lors des éventuelles manipulations des communes.

La prise en main de cette application par les équipes régionales en charge de la constitution des unités primaires a nécessité un jour de formation, y compris la formation à la méthodologie du regroupement.

1.3 Mise en oeuvre de l'application

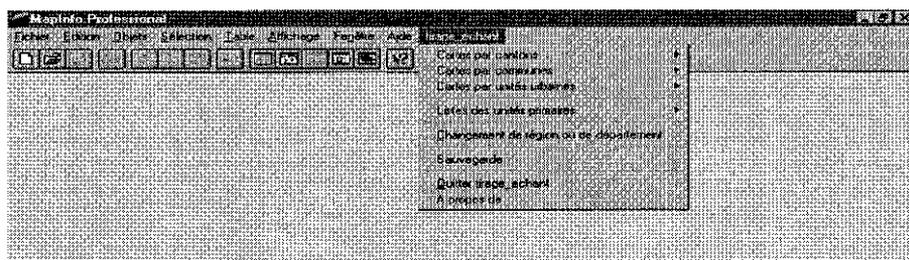
La première opération consiste à choisir la région.



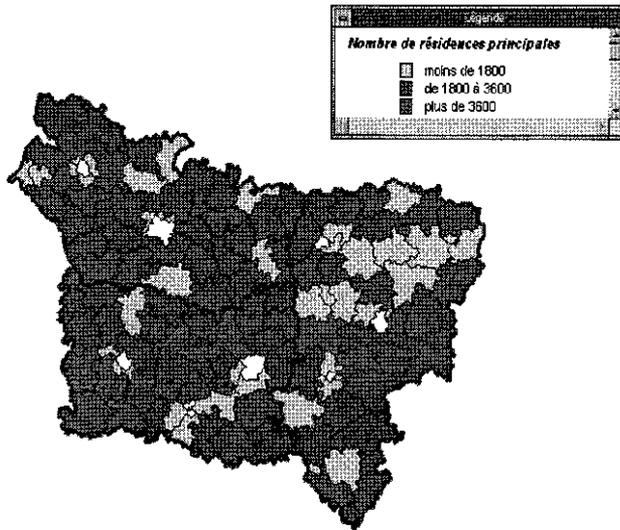
La création des unités primaires d'une région s'effectue de la façon suivante :

Les cantons sont d'abord traités, puis on passe aux communes, avec retour possible aux cantons. Le fonds communal permet de traiter les cantons devant être fractionnés (de taille supérieure à 3600 résidences principales). Les unités urbaines seront traitées de façon indépendante.

Une fois, la région choisie, les traitements suivants sont proposés.



La carte cantonale de la région choisie apparaît, avec une différenciation de couleur par type de cantons. Nous distinguerons quatre types de cantons : les cantons dont le nombre de résidences principales est inférieur à 1800, les cantons dont le nombre de résidences principales est compris entre 1800 et 3600, les cantons dont le nombre de résidences principales est supérieur à 3600 et enfin les cantons entièrement composés de communes autre que rurales. Il est à noter que ces derniers ne peuvent être traités, puisque hors champ des opérations de regroupement.



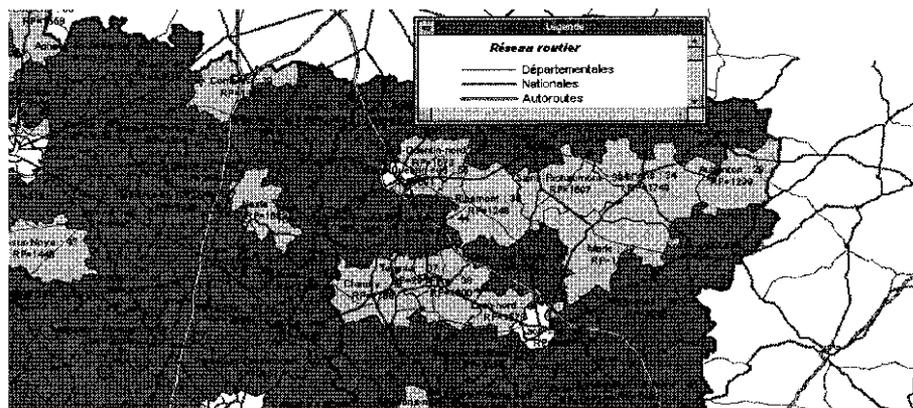
La création de l'unité primaire se fait en 3 étapes :

Etape 1 : Sélection de l'entité (cantons, communes ou unités urbaines) (création d'un zonage)

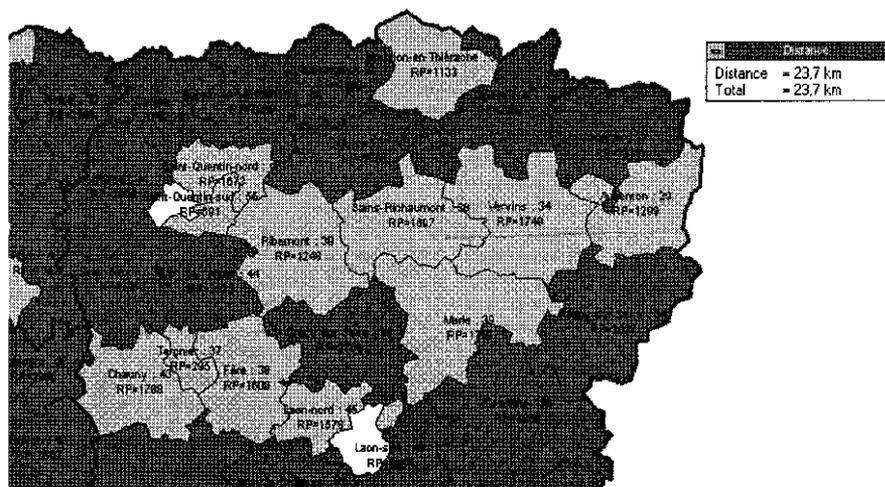
Etape 2 : Affichage du tableau statistique

Etape 3 : Validation du zonage

Un second outil permettant de faire apparaître les routes (autoroutes, nationales, départementales) afin de s'assurer de l'accessibilité.

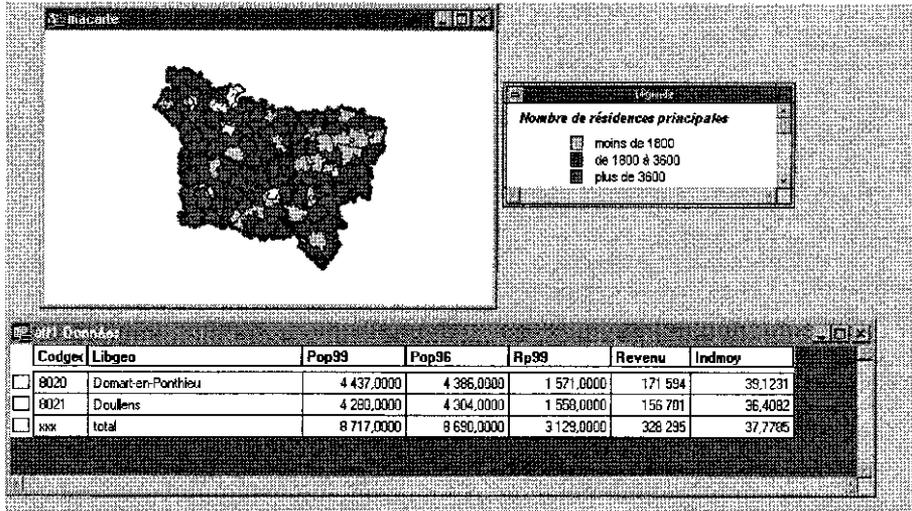


Un troisième outil permettant de calculer la distance (à vol d'oiseau) entre les extrêmes de l'unité primaire. Ce dernier outil est plus particulièrement utilisé lors de la création des unités primaires à partir des communes.

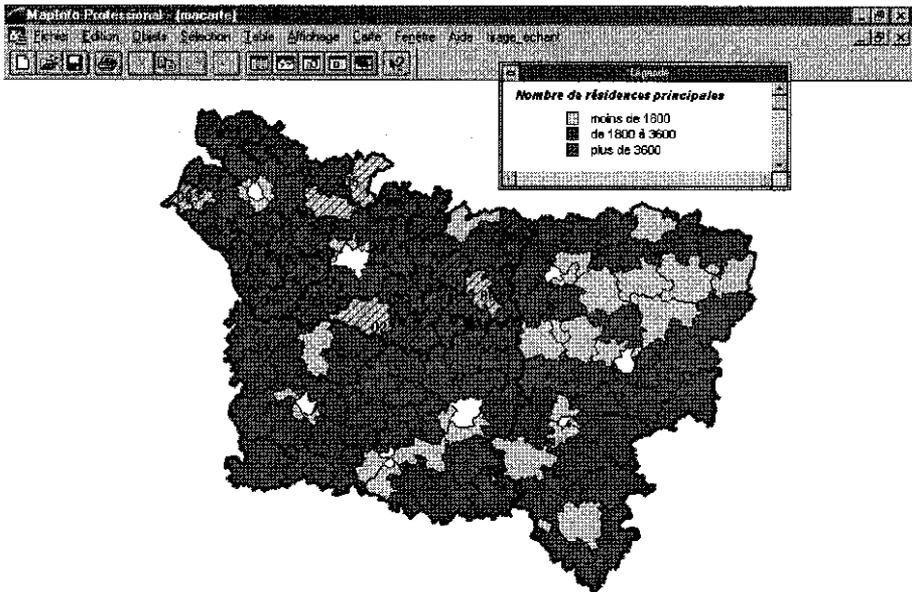


Etape 2 : Affichage du tableau statistique

Si le zonage sélectionné a un nombre de résidences principales inférieur à 1800 ou supérieur à 3600, apparaît un **WARNING** informant l'utilisateur du non-respect des seuils.



Etape 3 : Validation de la zone et création de l'unité primaire



1.4 Restitution des résultats

La sauvegarde des différents fichiers ayant permis la constitution des unités primaires se fait d'une façon automatique en utilisant l'outil *sauvegarde* de l'application. Ensuite, les fichiers sont envoyés par messagerie électronique à la Direction Générale (U.M.S) à fins de contrôle et validation.

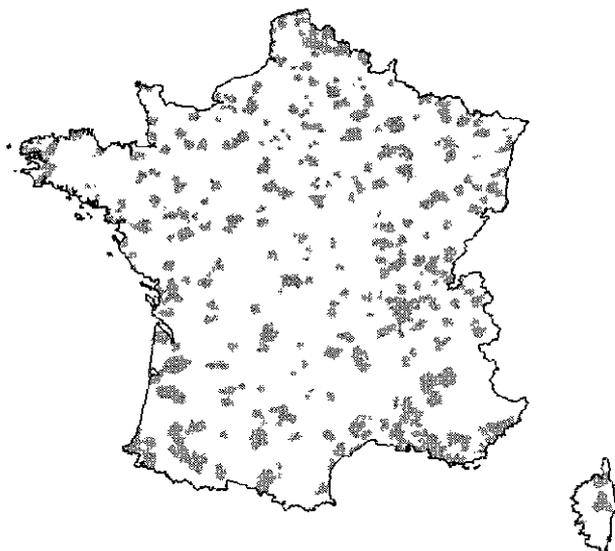
Certains fichiers reçus seront exportés vers SAS afin de procéder aux différents contrôles cités ci-dessus (respect des seuils de taille, vérification de l'exhaustivité...)

1.5 Quelques données numériques sur l'Échantillon-Maître

| | Strate 0 | Strate 1 | Strate 2 | Strate 3 | Strate 4 | Total |
|---|------------------|------------------|--------------------|--------------------|--------------------|-------------------|
| Nb d'unités primaires dans la base avant tirage | 2226 | 975 | 180 | 52 | 1 | 3434 |
| Nb d'unités primaires dans l'EM | 128 | 75 | 93 | 52 | 1 | 349 |
| Nb de communes dans l'EM | 2017 | 284 | 610 | 1294 | 396 | 4601 |
| NB total de logements | 7 380 663 | 4 934 454 | 3 839 693 | 988 099 | 4 549 934 | 28 692 843 |
| NB total de logements dans l'EM | 452348 | 445022 | Non connu * | Non connu * | Non connu * | |
| Poids de l'EM (%) | 6,13 | 9,01 | * | * | * | |

* En effet, dans ces strates, la constitution finale de l'EM inclut une étape de tirage des districts, qui sera réalisée en 2001.

1.6 Représentation cartographique des strates 0, 1, 2 et 3



2. Échantillon-Emploi 1999

De même que pour l'EM99, la constitution de l'Échantillon Emploi 1999 a nécessité des travaux préliminaires. Un des premiers travaux a été la constitution des unités de tirage, appelées unité de première phase.

Ces unités de première phase peuvent être de cinq types différents, soit des districts, soit des Iris, soit des regroupements d'Iris, soit des communes ou soit des regroupements de communes et elles doivent avoir au moins 120 logements (principaux, secondaires, vacants, occasionnels).

Du fait de cette dernière contrainte et suivant le type d'unités, rurales ou urbaines, nous avons dû procéder à des opérations de regroupement. Dans l'urbain, nous avons procédé à des regroupements d'Iris et dans le rural à des regroupements de communes.

Du fait du faible nombre d'Iris2000 de moins de 120 logements, aux alentours de 800, et surtout du fait que toutes les communes découpées en Iris ne sont pas numérisées, cette opération de regroupement s'est faite manuellement en région. En revanche, dans le rural où 11 900 communes ont moins de 120 logements, nous nous sommes servis de la cartographie.

La problématique était de rattacher les communes de moins de 120 logements avec des communes contiguës, faisant partie de la même zone d'emploi, de façon à former un agrégat d'au moins 120 logements.

Une des solutions envisagées était de cartographier, région par région, les communes, de différencier les communes de taille strictement inférieure à 120 logements et de procéder aux regroupements en utilisant l'application cartographique de l'Échantillon-Maître 1999.

Mais, du fait de l'existence d'un fichier de contiguïté des communes, nous avons procédé d'une manière plus automatique. Un algorithme de regroupement automatique des communes tenant compte des critères de regroupement, a été développé sous SAS par le CNIP (G.MORVAN)¹.

La sortie de cet algorithme se présente sous forme de tables SAS qui sont exportées sous MapInfo et cartographiées.

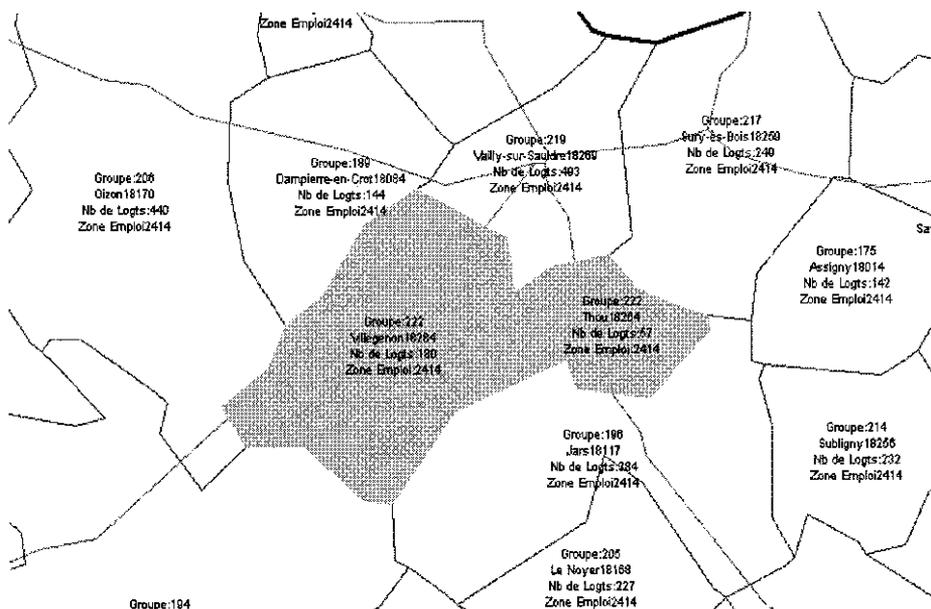
Les cartes ainsi créées, une par région, ont été transmises aux régions pour validation. Afin de permettre un travail de validation visuelle et rapide, un fonds de couleur caractérise les communes suivant qu'elles font partie ou non d'un regroupement.

Cette validation est rendue nécessaire du fait que les regroupements proposés par l'algorithme sont des regroupements basés sur une contiguïté plane faisant abstraction des problèmes de terrain (vallée, voie ferrée, cours d'eau.....)

De même, apparaissent sur les fonds de cartes différentes informations telles que le nom de la commune, son code communal, le nombre de logements, la zone d'emploi. Ces informations sont nécessaires lors d'une invalidation d'un regroupement proposé et de la constitution d'un nouveau regroupement.

¹ Cet algorithme n'introduit pas de critères de similitude comme celui qui a été présenté par M.CHRISTINE et M.ISNARD au cours des VIIèmes JMS (Session 3).

Exemple de cartographie en sortie de l'algorithme de regroupement



2.1 Quelques données numériques

Strate 0: Communes rurales (TU99=0)

Strate 1: Unités urbaines de moins de 10 000 habitants (TU99=1,2)

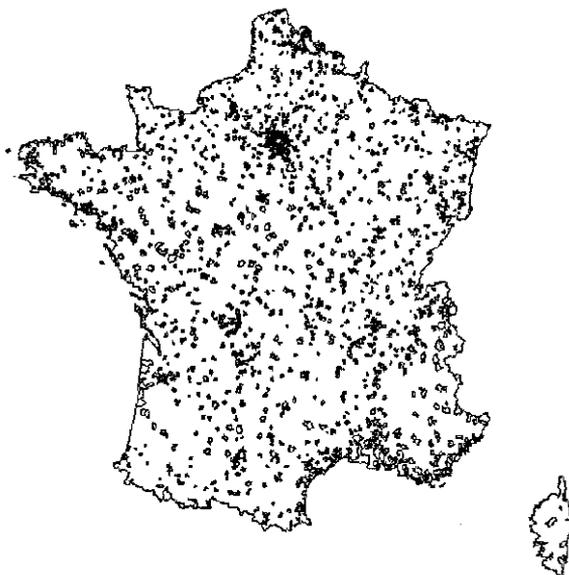
Strate 2: Unités urbaines de 10 000 à moins de 50 000 habitants (TU99=3,4)

Strate 3: Unités urbaines de 50 000 à moins de 200 000 habitants (TU99=5,6)

Strate 4: Unités urbaines de 200 000 habitants ou plus (TU99=7,8)

| | Strate 0 | Strate 1 | Strate 2 | Strate 3 | Strate 4 | Total |
|---|----------|----------|----------|----------|----------|-------|
| Nb d'unités de tirage dans la base avant tirage | 22327 | 4553 | 052 | 3722 | 12389 | 47043 |
| Nb d'unités de première phase | 668 | 304 | 396 | 527 | 1107 | 3002 |

2.2 Représentation cartographique de l'Échantillon Emploi 1999 (toutes strates confondues)



2.3 Représentation graphique des deux échantillons (Maître et Emploi) (toutes strates confondues)

