



TD 6 : Jointure et discrétisation

Thématique géographie

ArcGis/MapInfo/GvSig/QGIS

Temps de réalisation : 1 à 2 heures selon assimilation des notions théoriques.

Pré-requis : Environnement bureautique informatique classique. Environnement du logiciel SIG utilisé (TD1 et TD2 réalisés).

Documents nécessaires : fiche d'aide à l'utilisation (au choix : ArcGis/MapInfo/GvSig/QGIS).

Fonctionnalités observées :

Observation des données	1
Calcul de surface	2
Jointure à partir d'une colonne	2
Mise à jour d'un champ en fonction d'autres champs	2
Discrétisation des données	2
Supplément fusion.....	3

Données utilisées (Conditions d'utilisations dans \\DATAS\) :

- ©IGN – Géofla® 2009 – www.ign.fr
- ©INSEE – Recensement de la population 2007 – <http://www.insee.fr/> .

Contexte :

Nous voulons réaliser une carte simple de la densité de population française en 2007.

Objectif : Savoir joindre des données statistiques à des entités géographiques pour les représenter dans l'espace et avoir quelques notions de discrétisation des données.

Observation des données

Ouvrez le logiciel SIG choisi et paramétrez l'interface logiciel (projection) si nécessaire.

Toutes les fiches ArcGis A/ MapInfo A/ GvSIG A/ QSIG A si besoin

Au démarrage du logiciel vous vous retrouvez sur un document ou projet vide.

Ouvrez le fichier suivant (dans le dossier \\IGNGEOFLA\) :

Nom	Format	Type	Précision	Projection	description
Dps_fr.shp	ESRI Shape-SHP	Vecteur polygone	1 000 000	Lambert93/2154	Contour des départements

Fiche ArcGis_B01/ MapInfo_B01, G01/ GvSIG_B01/ QSIG_B01

Note de logiciel :

Sous MapInfo : Vous devez importer le fichier SHP (pour qu'il ne soit plus en lecture seule, puisque nous serons amenés à mettre ce fichier à jour ultérieurement), en précisant le système de projection à prendre en compte au moment de l'import (bouton **Projection**).

Le système de projection se choisit comme suit : Catégorie *Système français RGF93, Lambert II 93 (EPSG : 2154)*.

Fiche MapInfo_G01

Vous pouvez ensuite ouvrir le fichier *.TAB et travailler avec normalement.

Fiche MapInfo_B01

Ouvrez le fichier des données statistiques de l'INSEE concernant la population par département en France métropolitaine (Nb. : ce fichier est une extraction du fichier total fourni par l'INSEE).

Nom	Format	Type	Précision	Projection	description
Pop_dpt.dbf	Table DBF IV	Table attributaire	nd	nd	Population départementale

Fiche ArcGis_B01/ MapInfo_B01/ GvSIG_A03, B04/ QSIG_B01

Observez les données attributaires de la couche des départements et la table de l'INSEE (population).

Vous voulez représenter la densité de population :

- Vous disposez des contours des départements géoréférencés ; il est donc possible de calculer les surfaces.
- Vous disposez de la population par département dans la table INSEE ; il est possible de « rattacher » cette table aux contours des communes (jointure par champs).

Calcul de surface

Afin d'obtenir la surface des départements, il faut créer une colonne (ou champ) qui contiendra l'information, puis de calculer la donnée en km² si possible.

Créez la colonne « surface » au format numérique avec deux chiffres significatifs (ce qui est largement suffisant, puisque l'on travaille avec des données au 1 :1.000.000^e).

Fiche ArcGis_C02, C05/ MapInfo_C02, C05/ GvSIG_C06, C03/ QGIS_C02, C06

Créez une colonne « DENSITE » qui contiendra ultérieurement la densité de population des départements.

Quel type de champ devez-vous créer ?

Fiche ArcGis_C02/ MapInfo_C02/ GvSIG_C03/ QSIG_C03

Jointure à partir d'une colonne

Joignez les données de la population (INSEE) aux contours des départements (et pas l'inverse, sinon, vous perdez l'information géographique).

Quel est le champ de jointure dans la table attributaire de 'DEPARTEMENT' :

Quel est le champ de jointure dans la table attributaire de 'pop_dept' :

Pour quelle raison relative au format/à la nature de ces colonnes, la jointure est-elle possible ?

.....

Fiche ArcGis_D03/ MapInfo_D02/ GvSIG_D03/ QSIG_D03

Ouvrez la table attributaire de la couche 'DEPARTEMENT' pour observer que la jointure a bien été réalisée, et que vous disposez bien de la colonne du nombre d'habitants dans cette table.

Mise à jour d'un champ en fonction d'autres champs

Calculez la valeur de la colonne « DENSITE » créée précédemment, en divisant la colonne de population par la surface.

Fiche ArcGis_C05/ MapInfo_C05/ GvSIG_C06/ QSIG_C06

Discretisation des données

Pour réaliser une discretisation, il faut choisir le nombre de classes et les bornes de chaque classe. Pour réaliser une bonne discretisation, il faut justifier les classes. Il est donc indispensable d'observer l'histogramme de répartition des valeurs avant de fixer les classes. La classification conditionne l'interprétation de la carte finale !

Choix du nombre de classes : L'idéal pour une carte lisible est de trois ou quatre classes, avec un maximum de sept classes à ne jamais dépasser.

Choix de la méthode : Il n'existe malheureusement pas de méthode miracle, mais il existe des orientations, ou conseils pour orienter au mieux les choix.

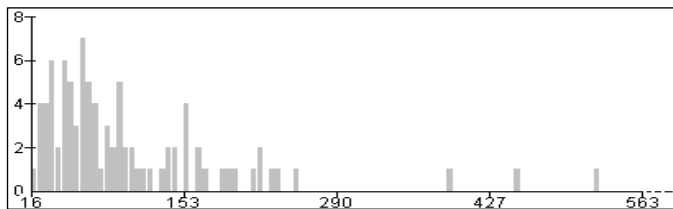
Nous retiendrons trois types essentiels de distribution des données, et la tendance des classifications qui leurs sont associées :

- **Symétrique** (courbe de Gauss) : Les données étant centrées autour de la moyenne, la méthode de la **moyenne et de l'écart-type** est souvent la plus appropriée.
- **Dissymétrique** : C'est la distribution la plus fréquemment rencontrée ; la méthode des **effectifs égaux** permet de bien représenter ces données, mais la représentation par les **seuils naturels** peut être indiquée. (la dissymétrie particulière où les données sont regroupées sur les valeurs faibles peut amener la méthode de **progression géométrique** qui apporte un regard intéressant sur les données).
- **Uniforme** : On retrouve ces histogrammes très rarement ; elle est associée à la méthode de seuils d'**amplitude égale**. On utilise l'amplitude égale pour mettre en évidence une classe par rapport à une autre (pourcentages, températures...).

Observez l'histogramme de répartition des données, et essayez plusieurs classifications, en modifiant le nombre de classes et la méthode de classification.

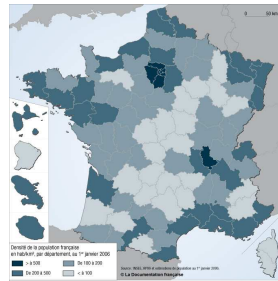
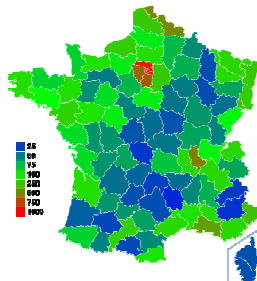
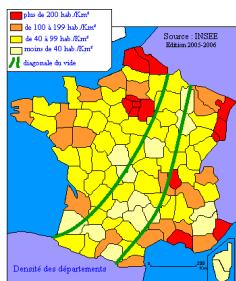
Fiche ArcGis_E02/ MapInfo_E01 à E03/ GvSIG_E02/ QSIG_E02

Exemple de représentation de l'histogramme des données :



NB. Cet histogramme est une partie de l'histogramme délivré par ArcView (ESRI) au moment du choix de classification des données du présent exercice. Des regroupements visuels ont déjà été réalisés par le logiciel.

On retrouve sur internet (ou ailleurs) de nombreux exemples de classification de cette donnée de densité par département :



Sources :

Figures gauche (2005) et centre (2008) :
Wikipedia. fr.wikipedia.org
Figure droite : Documentation Française (2006)
www.ladocumentationfrancaise.fr
Source des chiffres : INSEE

On remarque que la méthode de classification et le nombre de classes sont différents (sans parler des couleurs). Tout ceci dépend de ce que l'on veut mettre en avant dans la carte finale.

Adaptez l'affichage, la méthode et le nombre de classes pour faire ressortir les départements où se situe une grande agglomération, mais en essayant de distinguer au mieux la « diagonale du vide ».

Étiquetez chaque département avec son numéro de département.

Fiche ArcGis_E03/ MapInfo_B02/ GvSIG_E03/ QSIG_E03

Supplément fusion

Vous voulez représenter la densité de population à l'échelle supérieure (la région). L'action à réaliser est une fusion (ou agrégation, regroupement, selon les logiciels).

Fusionnez les départements appartenant à la même région. N'oubliez pas de conserver la densité et de l'agréger au niveau régional lors de la fusion.

Fiche ArcGis_D04/ MapInfo_D04/ GvSIG_D04/ QSIG_D04

Quel champ dans la table de la couche des départements permet la fusion ?

Représentez la densité au niveau régional comme précédemment.