

Philippe Waniez

Eclats

Mode d'emploi



<https://www.pexels.com/photo/sky-lights-night-new-year-s-eve-66277/>

Table des matières

Introduction : de la version 1 à la version 2	5
Première partie : le traitement des fichiers Excel .xlsx	9
Chapitre 1 : présentation générale du traitement des fichiers .xlsx	11
1. Objectifs	11
2. Données Excel à traiter	12
Chapitre 2 : l'agrégation des données statistiques	17
1. L'interface d'Eclats pour l'agrégation	17
2. L'agrégation et ses résultats	18
Chapitre 3 : l'éclatement des données statistiques	23
1. L'interface d'Eclats pour l'éclatement	23
2. L'éclatement et ses résultats	24
Deuxième partie : le traitement des shapefiles .shp	27
Chapitre 4 : présentation générale du traitement des shapefiles	29
1. Que sont les shapefiles	29
1. Objectifs	30
2. Données shapefile à traiter	30
Chapitre 5 : les différents modes de sélection dans les shapefiles	33
1. La sélection directe	33
2. La sélection indirecte avec fichier d'identifiants	38
3. Les sélections indirectes multiples avec fichier d'identifiants	41
Chapitre 6 : le recodage des identifiants d'un shapefile	47
Chapitre 7 : renommer tous les fichiers d'un même shapefile	53
Chapitre 8 : la conversion des shapefiles en fichiers .AI	55

Introduction

De la version 1 à la version 2

Eclats est un programme utilitaire destiné à être utilisé en complément de **Philcarto** (<http://philcarto.free.fr>). La version 1.xx permettait exclusivement d'exploiter des classeurs Excel® au format .xlsx pour réaliser des sélections ou des agrégations (additions) de lignes extraites d'un classeur. Autrement dit, la version 1.xx était seulement destinée à la préparation des données statistiques en vue de leur cartographie ultérieure.

La version 2.xx étend les fonctions du programme aux shapefiles, c'est-à-dire aux fonds de cartes enregistrés au format .shp. Sur des tels fichiers, **Eclats** réalise soit une sélection simple (le résultat est un seul nouveau shapefile contenant un extrait de celui fourni en entrée), soit une sélection multiple (le résultat consiste en plusieurs nouveaux shapefiles contenant chacun un extrait du shapefile fourni en entrée). Autrement dit, la version 2.xx permet de préparer un fond de carte correspondant précisément à l'espace devant être cartographié. Les utilisateurs « historiques » de **Philcarto** retrouveront dans cette extension une partie des fonctions du programme **ShapeSelect** qui, devenu obsolète, n'est plus diffusé.

Eclats est une application 64 bits (x64) conçue sous Windows 10 64 bits mais qui fonctionne aussi avec Windows 7 64 bits et les versions suivantes. Si la version de Windows installée sur votre ordinateur est une version 32 bits, **Eclats** ne peut pas fonctionner. Pour accéder au contenu des fichiers Excel, Eclats utilise la bibliothèque de fonctions SpreadsheetLight téléchargeable gratuitement sur Internet à l'adresse : <http://spreadsheetlight.com/>. Cette bibliothèque ne fonctionne correctement qu'en mode 64 bits ; d'où la limitation indiquée ci-dessus.

Le programme d'installation d'**Eclats** (Figure n° 1) est téléchargeable sur le site **Philcarto** (<http://philcarto.free.fr>). Ce programme installe tous les composants nécessaires au bon fonctionnement d'**Eclats**, y compris SpreadsheetLight.

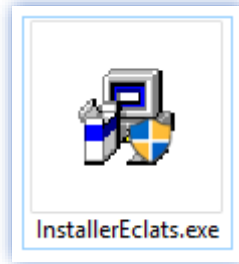


Figure n°1 : l'icône du programme d'installation d'**Eclats**.

Si **Eclats 1.xx** est déjà installé sur l'ordinateur, il est préférable de le désinstaller avec le panneau de configuration de Windows **avant l'installation de la version 2.xx**. Pour installer **Eclats 2.xx**, il suffit d'un double clic sur cette icône et de valider les options proposées sans rien changer. Si **Eclats 2.xx** est déjà installé sur l'ordinateur, le programme affiche le message suivant (fig. 2). Dans ce cas, cliquer sur le bouton Oui (ce qui a pour effet de désinstaller **Eclats**) et relancer ensuite le programme d'installation.

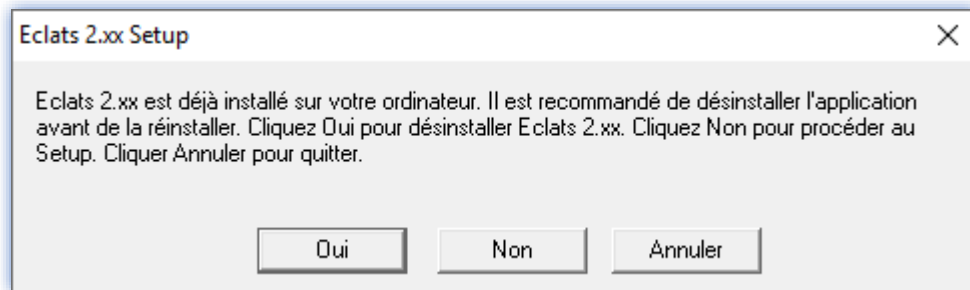


Figure n°2 : la désinstallation préalable d'**Eclats**.

Lorsque l'installation est achevée, on trouve l'icône du programme sur le bureau de Windows (fig. 3). Le U figurant sur cette icône signifie « Utilitaire ». Un double clic sur cette icône permet de lancer l'exécution d'**Eclats**. La fenêtre d'accueil s'affiche à chaque lancement du programme (fig. 4).



Figure n°3 : l'icône d'**Eclats** sur le bureau de Windows.



Figure n°4 : la fenêtre d'accueil d'Eclats.

Dans cette fenêtre, la langue d'utilisation peut être choisie par un clic sur le bouton ou sur le drapeau de la langue désirée. Par défaut, le programme fonctionne en Français. Après lecture de la licence d'utilisation (faire défiler le texte avec l'ascenseur situé à droite du texte), l'utilisateur peut soit l'accepter, soit la refuser. Si l'utilisateur s'engage à respecter les termes de la licence, il doit cliquer sur le bouton **J'accepte** et le programme passera à l'étape suivante, celle de l'ouverture des fichiers. Un clic sur le bouton **Je refuse** provoque la fin du programme.

Après avoir choisi la langue d'utilisation et accepté la licence d'utilisation, un dialogue permet d'indiquer au programme le type de traitement à réaliser (fig. 5). Ce dialogue, nouveau dans la version 2, comprend 6 fonctions accessibles par des boutons répartis en deux colonnes.

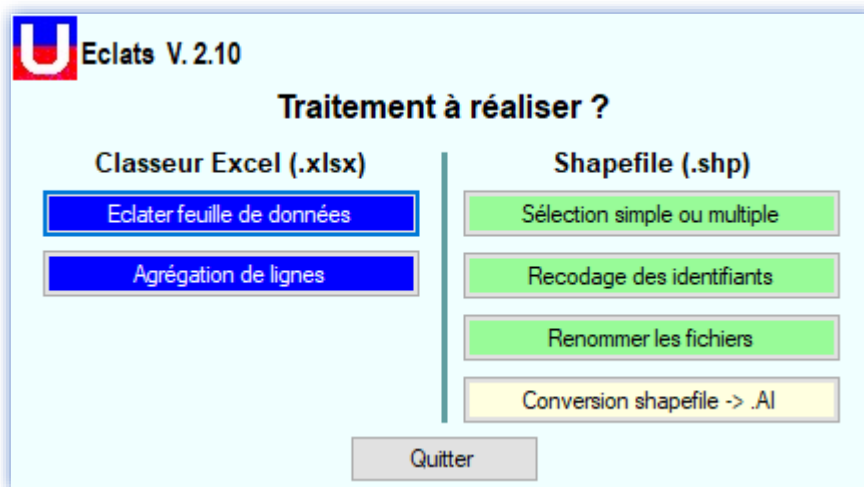


Figure n°5 : la fenêtre de choix du type de traitement.

La colonne de gauche est composée de deux boutons bleus destinés à traiter des classeurs Excel. Il s'agit des mêmes fonctions que celles assurées par la version 1 du programme, c'est-à-dire d'une part l'éclatement d'une feuille de données en plusieurs feuilles et d'autre part l'agrégation des lignes d'une feuille. Ces deux fonctions recourent aux modalités d'une variable nominale d'identification (codes de départements, par exemple) qui permet soit de séparer les lignes, soit de les additionner.

La colonne de droite comprend quatre boutons, trois verts et un jaune, destinés à traiter des shapefiles (autrement dit, des fonds de cartes). La première fonction (premier bouton vert) réalise des opérations de sélection simple ou multiple. Ce sont les traitements autrefois assurés par **ShapeSelect**, avec de nombreuses améliorations cependant. Le bouton situé au-dessous permet de recoder les identifiants sur la base d'un fichier de recodage à deux colonnes : ancien identifiant, nouvel identifiant. Le troisième bouton facilite le changement de nom, en une seule opération, des quatre fichiers constitutifs d'un shapefile (au lieu de quatre en utilisant l'explorateur de Windows). Enfin, le bouton jaune convertit un shapefile en fichier .AI directement utilisable avec **Philcarto**. Enfin, un clic sur le bouton gris conduit à l'abandon de l'exécution du programme.

Pour apprendre à utiliser **Eclats**, deux dossiers d'exemples sont fournis dans le dossier Eclats_Tests, l'un pour le traitement des fichiers Excel .xlsx (Eclats_Test_Brasil), l'autre pour le traitement des shapefiles .shp (Eclats_Test_Gironde). En fonction du type de traitement choisi, il faut ouvrir l'un ou l'autre de ces dossiers comme indiqué dans les pages à venir.

Première partie

Le traitement des fichiers Excel .xlsx

Chapitre 1

Présentation générale du traitement des fichiers .xlsx

Eclats est un programme utilitaire destiné à être utilisé en complément de **Philcarto** (<http://philcarto.free.fr>). **Eclats** permet d'exploiter des classeurs Excel® au format .xlsx (pour des raisons techniques, le format antérieur, .xls, n'est pas autorisé ; si vous souhaitez utiliser des fichiers au format .xls, il faut d'abord les convertir au format .xlsx).

1. Objectifs

Eclats réalise deux types d'opérations :

1. L'éclatement d'une feuille de calcul ; il s'agit de diviser une feuille de calcul Excel .xlsx en plusieurs fichiers Excel .xlsx différents en fonction de l'appartenance des lignes de la feuille de calcul initiale à un niveau de maillage supérieur au niveau dans lequel les données statistiques sont fournies au programme (niveau de base). Par exemple, si le département est le niveau de base, le maillage de niveau supérieur est la région. **Eclats** produit autant de fichiers qu'il y a de régions. Ces nouveaux fichiers renferment une copie des données du niveau de base, mais seulement pour chacune des régions. Autrement dit, si les 96 départements français métropolitains forment le niveau de base, **Eclats** produit 13 nouvelles feuilles de calcul, une feuille par région où se répartissent ces 96 départements et leurs données statistiques présentes dans la feuille de calcul initiale.

2. L'agrégation des données d'une feuille de calcul ; il s'agit d'additionner les quantités présentes dans une feuille de calcul Excel .xlsx en fonction de l'appartenance des lignes de la feuille de calcul initiale à un niveau de maillage supérieur au niveau dans lequel les données statistiques sont fournies au programme (niveau de base). Par exemple, si le département est le niveau de base, le maillage de niveau supérieur est la région. **Eclats** produit une seule nouvelle feuille de calcul avec autant de lignes qu'il y a régions. Autrement dit, si les 96 départements français métropolitains forment le niveau de base, **Eclats** produit une nouvelle feuille de calcul comprenant 13 lignes. Dans les cellules de cette feuille de calcul, on trouve l'addition, colonne par colonne des données statistiques qui caractérisent les départements dans la feuille de calcul initiale. Ces additions portent seulement sur des quantités ou des effectifs ; les autres types de données comme par exemple des pourcentages ou des modalités nominales sont ignorés dans la feuille contenant le résultat de l'agrégation.

La figure 6 (fig. 6) résume ces deux fonctions assurées par **Eclats** sur les fichiers .xlsx.

Bien entendu, **Eclats** ne fait rien qui ne puisse être fait avec Excel. Mais, lorsque les fichiers sont volumineux et le nombre de modalités d'éclatement ou d'agrégation est grand, le temps nécessaire pour réaliser ces opérations est prohibitif, et les risques d'erreur importants. Le principal intérêt d'**Eclats** est de libérer le temps du chercheur pour des tâches plus valorisantes...

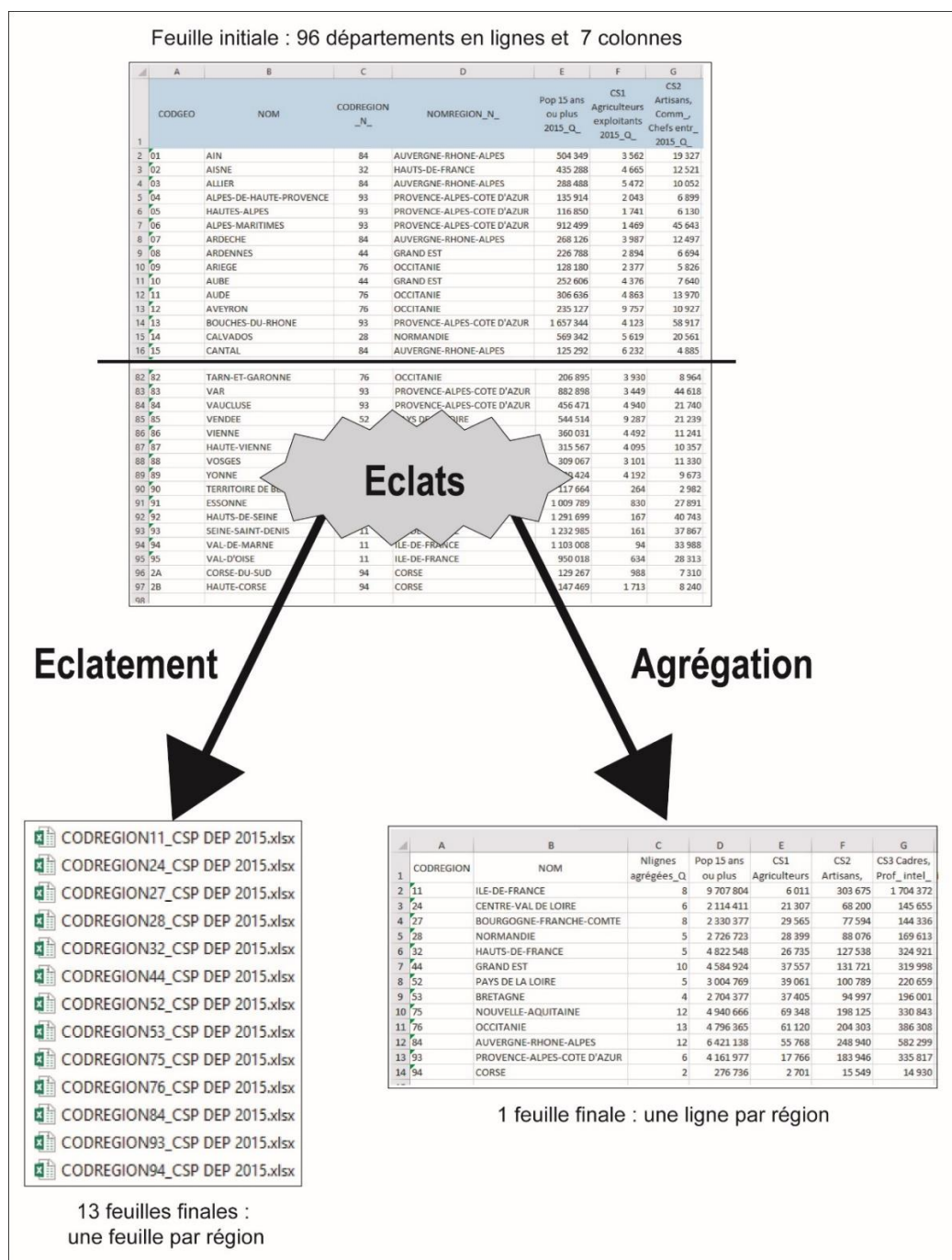


Figure n°6 : les deux fonctions d'Eclats pour le traitement des feuilles de calcul .xlsx.

2. Données Excel à traiter

Pour fonctionner, **Eclats** a besoin de deux feuilles de calcul différentes situées dans le même classeur Excel®, ou bien de deux classeurs Excel® différents contenant chacun l'une de ces deux feuilles de calcul.

1. Une première feuille de calcul (ou un premier fichier .xlsx) contient des données statistiques directement utilisables avec **Philcarto**, dans le format requis par **Philcarto** (voir la documentation de **Philcarto**). La première colonne elle contient les identifiants des éléments géographiques. Les autres colonnes renferment les données statistiques à traiter.

2. Une seconde feuille de calcul (ou un second fichier .xlsx) renferme les identifiants nécessaires ; les colonnes sont composées des codes d'agrégation ou de sélection des éléments géographiques. Une des colonnes de cette feuille (ou de ce fichier) doit contenir une

colonne identique à celle de la première colonne de la feuille (ou du fichier) de données statistiques. Cette colonne est appelée colonne de jointure ; elle n'est pas obligatoirement située en première position.

Par exemple, dans le cas des communes françaises, les colonnes (en dehors de la colonne de jointure) contiendront les codes des départements ou des régions auxquels ces communes appartiennent.

Considérons, par exemple, les résultats du Recensement Agropastoral du Brésil réalisé en 2017-2018 par l'Institut Brésilien de Géographie et de Statistique (IBGE). Dans le dossier Eclats_Test_Brasil on trouve un fichier Excel dénommé Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total qui donne la répartition du nombre d'exploitations agropastorales en 14 tailles de superficie totale. Dans ce fichier Excel, les lignes de la feuille de calcul EstAgro por grandes grupos área représentent les 5 563 communes du pays. La première ligne (fig. 7) contient les noms des colonnes (identifiants dans la colonne CODMUN, noms des communes dans la colonne NOM et variables statistiques). Les variables statistiques portent un nom complété par un suffixe :

Q pour les quantités ou les effectifs (un nombre de...) ;

R pour les rapports, indices, taux ;

N pour les données nominales et catégorielles.

Ces suffixes ne sont pas obligatoires, mais ils facilitent l'emploi du programme.

Notons que pour nommer les colonnes, il faut éviter les symboles entrant dans la définition des format d'Excel, tels que * # % + - € \$ £ . ; il est préférable de se limiter aux chiffres et aux lettres ainsi que _ pour indiquer les types des variables.

	A	B	C	D	E	F	G
	CODMUN	NOM	Estabelecimentos Total_Q_	Mais de 0 a menos de 2 ha_Q_	De 2 a menos de 5 ha_Q_	De 5 a menos de 20 ha_Q_	De 20 a mais de 200 ha_Q_
1							
2	1100015	Alta Floresta D'Oeste (RO)	2 884	178	511	748	
3	1100023	Ariquemes (RO)	2 925	122	234	625	
4	1100031	Cabixi (RO)	1 075	8	62	327	
5	1100049	Cacoal (RO)	3 805	226	590	1 076	
6	1100056	Cerejeiras (RO)	718	28	111	146	
7	1100064	Colorado do Oeste (RO)	1 672	40	134	372	
8	1100072	Corumbiara (RO)	1 488	14	89	343	
9	1100080	Costa Marques (RO)	1 495	43	150	183	
10	1100098	Espigão D'Oeste (RO)	1 999	33	139	320	
11	1100106	Guajará-Mirim (RO)	545	61	86	63	
12	1100114	Jaru (RO)	3 546	76	220	764	
13	1100122	Ji-Paraná (RO)	2 543	58	164	621	
14	1100130	Machadinho D'Oeste (RO)	4 228	41	210	526	
15	1100148	Nova Brasilândia D'Oeste (RO)	3 007	126	635	895	

Figure n°7 : un extrait de la feuille de calcul EstAgro por grandes grupos área total du fichier Excel Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total.

Il y a au Brésil au moins 4 niveaux géographiques ainsi hiérarchisés : communes, micro-régions, méso-régions et Unités de la Fédération (les Etats fédérés). Autrement dit, les communes appartiennent à 3 niveaux de maillage territorial supérieurs à celui des communes. C'est sur un cas de ce type que Eclats apparaît particulièrement puissant (beaucoup de communes et plusieurs niveaux de maillage). La feuille de calcul Divisão Territorial do Brasil du fichier Excel Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total donne, pour chaque commune (MUNicípio), son appartenance à chacun des niveaux supérieurs (fig. 8). On retrouve dans cette feuille la colonne dénommée CODMUN. Elle est strictement identique à la première colonne de la feuille de statistiques EstAgro por grandes grupos área. Les colonnes CODMICRO, CODMESO et CODUF donnent les codes des éléments géographiques

d'appartenance des communes aux trois niveaux supérieurs. Chacune de ces colonnes est accompagnée par une seconde colonne (NOMES MICRO, NOMES MESO et NOMES UF) donnant le nom de l'élément d'appartenance géographique de la commune qui figure en ligne.

	A	B	C	D	E	F	G	H
1	CODMUN	NOMES MUN	CODMICRO	NOMES MICRO	CODMESO	NOMES MESO	CODUF	NOMES UF
2	1100015	Alta Floresta D'Oeste	11006	Cacoal	1102	Leste Rondoniense	11	Rondônia
3	1100379	Alto Alegre dos Parecis	11006	Cacoal	1102	Leste Rondoniense	11	Rondônia
4	1100403	Alto Paraíso	11003	Ariquemes	1102	Leste Rondoniense	11	Rondônia
5	1100346	Alvorada D'Oeste	11005	Alvorada D'Oeste	1102	Leste Rondoniense	11	Rondônia
6	1100023	Ariquemes	11003	Ariquemes	1102	Leste Rondoniense	11	Rondônia
7	1100452	Buritis	11001	Porto Velho	1101	Madeira-Guaporé	11	Rondônia
8	1100031	Cabixi	11008	Colorado do Oeste	1102	Leste Rondoniense	11	Rondônia
9	1100601	Cacaulândia	11003	Ariquemes	1102	Leste Rondoniense	11	Rondônia
10	1100049	Cacoal	11006	Cacoal	1102	Leste Rondoniense	11	Rondônia
11	1100700	Campo Novo de Rondônia	11001	Porto Velho	1101	Madeira-Guaporé	11	Rondônia
12	1100809	Candeias do Jamari	11001	Porto Velho	1101	Madeira-Guaporé	11	Rondônia
13	1100908	Castanheiras	11006	Cacoal	1102	Leste Rondoniense	11	Rondônia
14	1100056	Cerejeiras	11008	Colorado do Oeste	1102	Leste Rondoniense	11	Rondônia
15	1100924	Chupinguaia	11007	Vilhena	1102	Leste Rondoniense	11	Rondônia
16	1100064	Colorado do Oeste	11008	Colorado do Oeste	1102	Leste Rondoniense	11	Rondônia

Figure n°8 : un extrait de la feuille de calcul « Divisão Territorial do Brasil » du fichier Excel « Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total ».

Comme indiqué au §2 l'utilisation d'**Eclats** nécessite l'ouverture de deux feuilles de calcul, l'une contenant les identifiants de sélection ou d'agrégation, l'autre renfermant les données statistiques à traiter. Ces deux feuilles peuvent être dans le même classeur Excel, ou bien se situer dans deux classeurs différents. C'est la raison pour laquelle deux dialogues d'ouverture de fichiers sont affichés lors du lancement du programme. Le premier dialogue concerne le fichier contenant la feuille des identifiants (fig. 9) ; le second dialogue a trait au fichier contenant la feuille des données statistiques (fig. 10). Si ces deux feuilles sont dans le même fichier, il faut ouvrir deux fois le même fichier (c'est le cas sur ces deux figures).

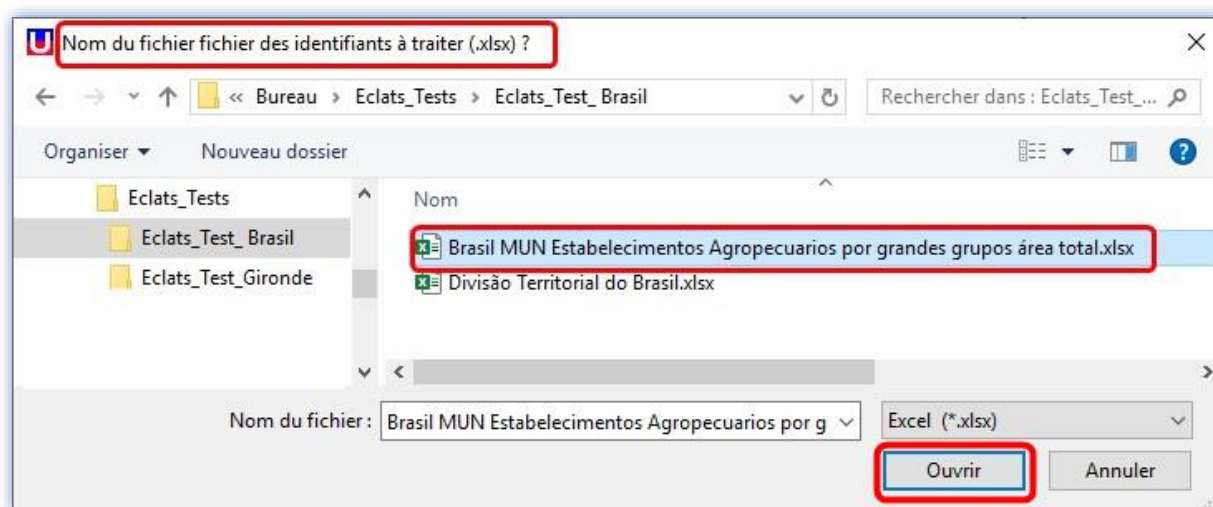


Figure n°9 : Sélection du fichier .xlsx contenant les identifiants pour l'agrégation.

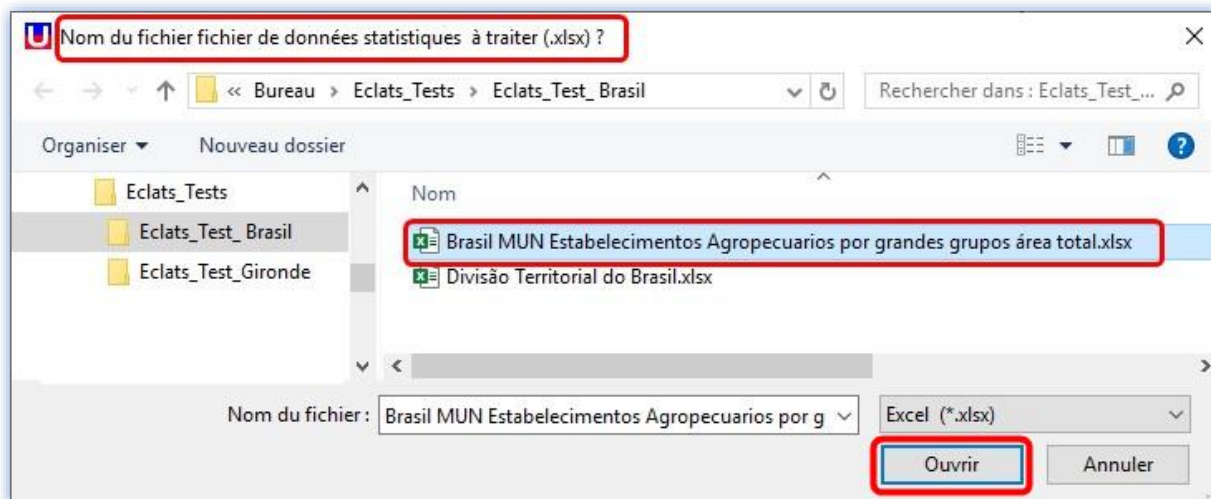


Figure n°10 : Sélection du fichier .xlsx contenant les données statistiques.

Le fichier Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total contient les deux feuilles de calcul nécessaires (fig. 11) : Divisão Territorial do Brasil (identifiants) et EstAgro por grandes grupos área total (statistiques).

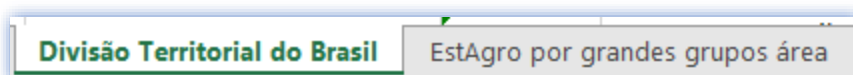



Figure n°11 : Les deux feuilles de calcul enregistrées dans le même classeur Excel.

Lors de l'ouverture du fichier des identifiants, on aurait pu aussi ouvrir le fichier Divisão Territorial do Brasil qui contient les mêmes informations que la feuille de calcul du même nom dans le fichier Excel Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total. Une bonne pratique consiste sans doute à séparer les deux feuilles dans des fichiers différents afin de rendre réutilisable la feuille des identifiants avec d'autres données statistiques.

Chapitre 2

L'agrégation des données statistiques

Ce traitement passe par un clic sur le bouton . Après avoir ouvert les deux fichiers nécessaires (identifiants et statistiques), **Eclats** affiche son interface dans une fenêtre unique dépourvue de menus.

Par exemple, il s'agit ici d'agréger le nombre d'exploitation agricoles au Brésil réparties en 14 tailles de superficie totale à partir du niveau « communes » vers le niveau « unités de la fédération ». Autrement dit de passer d'une feuille de calcul comprenant 5 564 lignes (y compris la première ligne contenant les noms des colonnes) à une seule nouvelle feuille de calcul contenant les 27 Unités de la Fédération (26 Etats plus le District Fédéral où se trouve la capitale Brasília).

1. L'interface d'Eclats pour l'agrégation

En plus du sélecteur de fonction (Agréger ou Eclater), l'interface comprend trois rangées de dialogues de contrôle (fig. 12). Les couleurs de chacun des dialogues correspondent aux étapes successives que l'utilisateur doit parcourir. Au stade initial, on peut soit agir sur les différents dialogues soit quitter le logiciel par un clic sur le bouton



Les dialogues **A**, **B** et **C** en bleu, concernent la feuille des identifiants. En **A**, on trouve la liste des feuilles trouvées dans le fichier ouvert en premier après le lancement du programme (ici les deux feuilles du fichier Excel Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total. L'utilisateur devra choisir la bonne feuille, ici Divisão Territorial do Brasil. Le contenu de cette feuille apparaîtra en **B** et en **C** après avoir cliqué sur le nom de la feuille en **A**. En **B** on sélectionne la variable de jointure (celle qui permet d'associer la feuille des identifiants avec celle des statistiques). Notons que lors de l'ouverture de la feuille de données statistiques, on recherche si le nom de la première colonne de la feuille statistique choisie est trouvé en **B** ; dans l'affirmative, le nom de la colonne commune qui permet la jointure est automatiquement sélectionné en **B**. En **C**, on choisit la variable d'agrégation.

Le dialogue **D**, en vert, permet de choisir la feuille des données statistiques. On y trouve la liste des feuilles présentes dans le fichier ouvert en premier après le lancement du programme. L'utilisateur devra choisir la bonne feuille, ici EstAgro por grandes grupos área total.

Eclats V.2.10

Identifiants pour agréger le fichier de statistiques

Sélectionner une feuille d'identifiants
Divisão Territorial do Brasil
EstAgro por grandes grupos área
A

Sélectionner la variable de jointure
B

Sélectionner variable d'agrégation
C

Statistiques à agréger

Sélectionner la feuille de statistiques
Divisão Territorial do Brasil
EstAgro por grandes grupos área
D

Variables à supprimer en sortie

E

☒ sélectionner tout ☒ dessélectionner tout

☒ Supprimer variable d'association

Noms des éléments agrégés

Sélectionner une variable de noms
F

Contrôles

Préfixe noms fichiers en sortie ?

Suffixe noms fichiers en sortie ?

☐ Enregistrer aussi au format .txt

G

Quitter

Figure n°12 : L'interface d'**Eclats** pour l'agrégation, après l'ouverture initiale des deux fichiers Excel.

Le dialogue **E**, en rose, permet d'éliminer les variables de quantité dont l'utilisateur n'a pas besoin. Les cases des variables qui ne sont pas quantitatives et qui possèdent les suffixes **_N_** et **_R_** sont automatiquement cochées ; les autres variables à éliminer doivent être cochées par l'utilisateur. Elles seront donc éliminées des résultats du traitement

Le dialogue **F**, en orange, sert à désigner la colonne de la feuille des identifiants qui contient les noms des éléments agrégés. En l'absence de choix, seuls les identifiants sont fournis dans la feuille contenant les résultats.

Enfin le dialogue **G** autorise divers contrôles sur le fichier de sortie.

Le traitement peut commencer lorsque l'utilisateur a fait son choix en **A**, **B**, **C** et **D**. Les autres dialogues sont optionnels : ils permettent d'affiner le contenu du fichier de sortie.

2. L'agrégation et ses résultats

Voici comment se présente l'interface d'Eclats pour l'agrégation au niveau des unités de la fédération (fig. 13) : ici, la variable d'agrégation est « COD UF » (en **B**) et les éléments agrégés seront nommés avec les valeurs de « NOMES UF » (en **H**). Lorsque les différents dialogues sont correctement définis, le bouton **Exécuter** fait son apparition. Un clic sur ce bouton déclenche le traitement.

Eclats V. 2.10

Identifiants pour agréger le fichier de statistiques

Sélectionner une feuille d'identifiants	Sélectionner la variable de jointure	Sélectionner variable d'agrégation
<div>Divisão Territorial do Brasil</div> <div>EstAgro por grandes grupos área</div>	<div>CODMUN</div> <div>NOMES MUN</div> <div>CODMICRO</div> <div>NOMES MICRO</div> <div>CODMESO</div> <div>NOMES MESO</div> <div>CODUF</div> <div>NOMES UF</div>	<div>NOMES MUN</div> <div>CODMICRO</div> <div>NOMES MICRO</div> <div>CODMESO</div> <div>NOMES MESO</div> <div>CODUF</div> <div>NOMES UF</div>

Statistiques à agréger

Sélectionner la feuille de statistiques

Divisão Territorial do Brasil

EstAgro por grandes grupos área

Variables à supprimer en sortie	Noms des éléments agrégés	Contrôles
<input type="checkbox"/> NOM <input type="checkbox"/> Estabelecimentos Total_Q_ <input type="checkbox"/> Mais de 0 a menos de 2 ha_Q_ <input type="checkbox"/> De 2 a menos de 5 ha_Q_ <input type="checkbox"/> De 5 a menos de 20 ha_Q_ <input type="checkbox"/> De 20 a menos de 50 ha_Q_ <input type="checkbox"/> De 50 a menos de 100 ha_Q_ <input checked="" type="checkbox"/> sélectionner tout <input checked="" type="checkbox"/> dessélectionner tout <input checked="" type="checkbox"/> Supprimer variable d'association	<div>Sélectionner une variable de noms</div> <div>NOMES MUN</div> <div>CODMICRO</div> <div>NOMES MICRO</div> <div>CODMESO</div> <div>NOMES MESO</div> <div>CODUF</div> <div>NOMES UF</div>	<div>Préfixe noms fichiers en sortie ?</div> <div></div> <div>Suffixe noms fichiers en sortie ?</div> <div></div> <input type="checkbox"/> Enregistrer aussi au format .txt <div>Exécuter</div> <div>Quitter</div>

Figure n°13 : L'interface d'**Eclats** pour l'agrégation, après le choix par l'utilisateur des différents paramètres de traitement.

Le calcul d'agrégation dure de quelques secondes à quelques minutes en fonction du volume de données à traiter. Un message s'affiche en bas et à droite de la fenêtre principale pour indiquer l'étape de traitement en cours. Quand le traitement se termine une alerte s'affiche ; elle indique l'heure de fin normale de l'agrégation (fig. 14). Ici, le temps de traitement est de 7 secondes ! Après avoir validé avec « OK », on quitte le programme. Une nouvelle utilisation implique de relancer le programme par un clic sur son icône.

Les résultats de l'agrégation sont enregistrés dans le même répertoire que celui du fichier Excel contenant les données statistiques. Le nom du fichier est « Agreg » suivi entre crochets du nom de la variable d'agrégation ; ici on obtient Agreg[COD UF].xlsx.

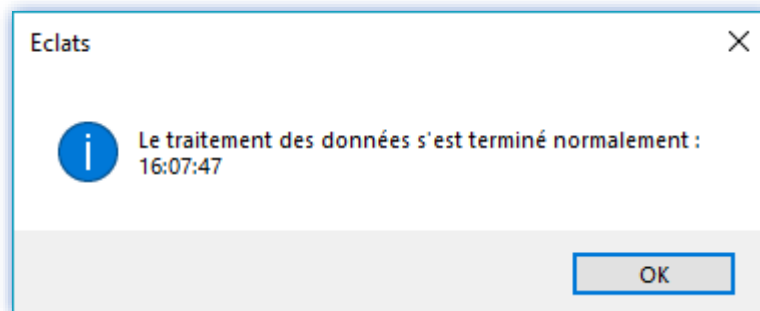


Figure n°14 : L'alerte de fin de traitement.

Voici comment se présente dans Excel la feuille Agrégation du fichier Agreg[COD UF].xlsx (fig. 15). La première colonne contient le code d'agrégation (ici COD UF) suivi du nom correspondant à chacun des codes d'agrégation (les noms de UF). On indique ensuite le nombre de lignes (le nombre de codes 'CODMUN' qui ont été agrégées pour chaque UF ; autrement dit, le nombre de communes ayant des données par UF. Enfin, les effectifs agrégés constituent les colonnes suivantes.

	A	B	C	D	E
1	CODUF	NOM	Nombre de 'CODMUN' agrégés_Q_	Estabelecimentos Total_Q_	Mais de 0 a men
2	11	Rondônia	52	91 125	
3	12	Acre	22	37 067	
4	13	Amazonas	62	77 750	
5	14	Roraima	15	16 579	
6	15	Pará	144	277 502	
7	16	Amapá	16	8 313	
8	17	Tocantins	139	62 952	
9	21	Maranhão	217	202 265	
10	22	Piauí	224	237 207	
11	23	Ceará	184	370 505	
12	24	Rio Grande do Norte	167	62 893	
13	25	Paraíba	223	161 607	
14	26	Pernambuco	185	279 344	
15	27	Alagoas	102	97 769	
16	28	Sergipe	75	92 709	
17	29	Bahia	416	756 406	
18	31	Minas Gerais	853	604 826	
19	32	Espírito Santo	78	107 355	
20	33	Rio de Janeiro	91	64 980	
21	35	São Paulo	640	187 805	
22	41	Paraná	399	304 110	
23	42	Santa Catarina	295	182 460	
24	43	Rio Grande do Sul	497	364 114	
25	50	Mato Grosso do Sul	79	70 557	
26	51	Mato Grosso	141	118 370	
27	52	Goiás	246	151 770	
28	53	Distrito Federal	1	5 238	
29					

Figure n°15 : Une copie d'écran partielle de la feuille de calcul résultat de l'agrégation.

La cohérence des résultats pour l'ensemble du Brésil peut être vérifiée en comparant les totaux des colonnes non agrégée et agrégées (fig. 16). Sur ce tableau, on observe que les sommes par variables sont identiques avant et après agrégation.

Nom des colonnes	Total avant agrégation	Total après agrégation
Nombre de communes	5 563	5 563
Estabelecimentos Total_Q_	4 993 578	4 993 578
Mais de 0 a menos de 2 ha_Q_	1 074 805	1 074 805
De 2 a menos de 5 ha_Q_	817 425	817 425
De 5 a menos de 20 ha_Q_	1 380 837	1 380 837
De 20 a menos de 50 ha_Q_	855 561	855 561
De 50 a menos de 100 ha_Q_	393 949	393 949
De 100 a menos de 500 ha_Q_	365 453	365 453
De 500 ha e mais_Q_	105 548	105 548

Figure n°16 : Le tableau des totaux des colonnes avant et après l'agrégation pour l'ensemble du Brésil.


On peut vérifier les calculs par Etat de la même façon, par exemple pour l'Etat du Minas Gerais, identifiant 31 (fig. 17). Sur ce tableau, on observe que les sommes par variables sont identiques avant et après agrégation. Sur ce tableau, on observe aussi que les sommes des variables sont identiques avant et après agrégation.

Nom des colonnes	Total avant agrégation	Total après agrégation
Nombre de communes	853	853
Estabelecimentos Total_Q_	604 826	604 826
Mais de 0 a menos de 2 ha_Q_	63 459	63 459
De 2 a menos de 5 ha_Q_	101 883	101 883
De 5 a menos de 20 ha_Q_	190 161	190 161
De 20 a menos de 50 ha_Q_	122 220	122 220
De 50 a menos de 100 ha_Q_	58 769	58 769
De 100 a menos de 500 ha_Q_	57 228	57 228
De 500 ha e mais_Q_	11 106	11 106

Figure n°17 : Le tableau des totaux des colonnes avant et après l'agrégation pour l'Etat du Minas Gerais.

Chapitre 3

L'éclatement des données statistiques

Ce traitement passe par un clic sur le bouton . Après avoir ouvert les deux fichiers nécessaires (identifiants et statistiques), **Eclats** affiche son interface dans une fenêtre unique dépourvue de menus.

Par exemple, il s'agit ici de faire éclater en 27 fichiers différents, un par unité de la fédération (UF), la feuille de calcul EstAgro por grandes grupos área total du fichier Excel Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total qui contient ces données pour l'ensemble des communes brésiliennes. Dans chacun des 27 fichiers doivent figurer toutes les communes de chacune des UF, et seulement ces communes-là.

1. L'interface d'Eclats pour l'éclatement

Il y a peu de différences entre les interfaces d'Eclats pour l'agrégation et pour l'éclatement des fichiers .xlsx (fig. 18). Les noms des fonctions changent, « éclater » au lieu de « agréger » mais leur fonctionnement reste souvent le même.

Les dialogues **A**, **B** et **C** en bleu, concernent la feuille des identifiants. En **A**, on trouve la liste des feuilles trouvées dans le fichier ouvert en premier après le lancement du programme (ici les deux feuilles du fichier Excel Brasil MUN Estabelecimentos Agropecuarios por grandes grupos área total. L'utilisateur devra choisir la bonne feuille, ici Divisão Territorial do Brasil. Le contenu de cette feuille apparaîtra en **B** et en **C** après avoir cliqué sur le nom de la feuille en **A**. En **B** on sélectionne la variable de jointure (celle qui permet d'associer la feuille des identifiants avec celle des statistiques). Notons que lors de l'ouverture de la feuille de données statistiques, on recherche si le nom de la première colonne de la feuille statistique choisie est trouvé en **B** ; dans l'affirmative, le nom de la colonne commune qui permet la jointure est automatiquement sélectionné en **B**. En **C**, on choisit la variable d'éclatement.

Les trois dialogues verts concernent les statistiques. Ici, il y a trois dialogues différents au lieu d'un seul en agrégation. Le dialogue **D**, permet de choisir la feuille des données statistiques. On y trouve la liste des feuilles présentes dans le fichier ouvert en premier après le lancement du programme. L'utilisateur devra choisir la bonne feuille, ici EstAgro por grandes grupos área total.

Les dialogues **E** et **F** se rapportent aux types de variables statistiques. Par défaut, le type est quantité (**_Q_**). Si l'on a affaire à des variables ayant déjà un suffixe **_R_** pour les variables de rapports et **_N_** pour les variables nominales, alors les cases correspondantes de **E** et de **F** sont automatiquement cochées. Dans le cas contraire, en l'absence de suffixe, il est indispensable de définir le type des variables qui ne sont pas des quantités, afin d'éviter qu'elles soient définies comme des quantités.

Le dialogue **G**, en rose, permet d'éliminer les variables qu'on ne souhaite pas conserver après éclatement.

Dans le dialogue **H** intitulé « Fichiers éclatés à enregistrer » sont listés les noms des modalités de la variable d'éclatement sous la forme de case à cocher. Par défaut, toutes les cases sont cochées ce qui signifie que tous les fichiers de sortie sont enregistrés. L'utilisateur peut décocher les cases des modalités dont il n'a pas besoin ; ainsi décochés, les fichiers ne

seront pas enregistrés. Cette option est utile lorsque le nombre de fichiers de sortie est grand et qu'on a besoin seulement de quelques-uns d'entre eux.

Enfin le dialogue **I** autorise divers contrôles sur le fichier de sortie.

Eclats V.2.10

Identifiants pour éclater le fichier de statistiques

Sélectionner une feuille d'identifiants
Divisão Territorial do Brasil
EstAgro por grandes grupos área
A

Sélectionner la variable de jointure
B

Sélectionner variable d'éclatement
C

Statistiques à éclater

Sélectionner la feuille de statistiques
Divisão Territorial do Brasil
EstAgro por grandes grupos área
D

Variables nominales ?
E

Variable de rapports ?
F

☒ sélectionner tout ☒ dessélectionner tout ☒ sélectionner tout ☒ dessélectionner tout

Variables à supprimer en sortie

G

☒ sélectionner tout ☒ dessélectionner tout

☒ Supprimer variable d'association

Fichiers éclatés à enregistrer

H

☒ sélectionner tout ☒ dessélectionner tout

Contrôles

Préfixe noms fichiers en sortie ?
Suffixe noms fichiers en sortie ?

☐ Enregistrer aussi au format .txt
☒ Ajouter le nom du fichier source

I

Quitter

Figure n°18 : L'interface d'**Eclats** pour l'éclatement, après l'ouverture initiale des deux fichiers Excel.

Le traitement peut commencer lorsque l'utilisateur a fait son choix en **A**, **B**, **C** et **D**. Les autres dialogues sont optionnels : ils permettent d'affiner le contenu du fichier de sortie.

2. L'éclatement et ses résultats

Voici comment se présente l'interface d'Eclats pour l'éclatement au niveau des unités de la fédération (fig. 19) : ici, la variable d'éclatement est « COD UF » (en **C**). Lorsque les différents dialogues sont correctement définis, le bouton **Exécuter** fait son apparition. Un clic sur ce bouton déclenche le traitement.

Eclats V.2.10

Identifiants pour éclater le fichier de statistiques

Sélectionner une feuille d'identifiants	Sélectionner la variable de jointure	Sélectionner variable d'éclatement
Divisão Territorial do Brasil EstAgro por grandes grupos área	CODMUN NOMES MUN CODMICRO NOMES MICRO CODMESO NOMES MESO CODUF NOMES UF	NOMES MUN CODMICRO NOMES MICRO CODMESO NOMES MESO CODUF NOMES UF

Statistiques à éclater

Sélectionner la feuille de statistiques	Variables nominales ?	Variable de rapports ?
Divisão Territorial do Brasil EstAgro por grandes grupos área	<input type="checkbox"/> De 2 a menos de 5 ha_R_ <input type="checkbox"/> De 5 a menos de 20 ha_R_ <input type="checkbox"/> De 20 a menos de 50 ha_R_ <input type="checkbox"/> De 50 a menos de 100 ha_R_ <input type="checkbox"/> De 100 a menos de 500 ha_R_ <input type="checkbox"/> De 500 ha e mais_R_ <input checked="" type="checkbox"/> UF_N_	<input type="checkbox"/> NOM <input type="checkbox"/> Estabelecimentos Total_Q_ <input type="checkbox"/> Mais de 0 a menos de 2 ha_Q_ <input type="checkbox"/> De 2 a menos de 5 ha_Q_ <input type="checkbox"/> De 5 a menos de 20 ha_Q_ <input type="checkbox"/> De 20 a menos de 50 ha_Q_ <input type="checkbox"/> De 50 a menos de 100 ha_Q_
	<input checked="" type="checkbox"/> sélectionner tout <input type="checkbox"/> dessélectionner tout	<input checked="" type="checkbox"/> sélectionner tout <input type="checkbox"/> dessélectionner tout

Variables à supprimer en sortie	Fichiers éclatés à enregistrer	Contrôles
<input type="checkbox"/> NOM <input type="checkbox"/> Estabelecimentos Total_Q_ <input type="checkbox"/> Mais de 0 a menos de 2 ha_Q_ <input type="checkbox"/> De 2 a menos de 5 ha_Q_ <input type="checkbox"/> De 5 a moins de 20 ha_Q_ <input type="checkbox"/> De 20 a moins de 50 ha_Q_ <input type="checkbox"/> De 50 a moins de 100 ha_Q_ <input checked="" type="checkbox"/> sélectionner tout <input type="checkbox"/> dessélectionner tout <input checked="" type="checkbox"/> Supprimer variable d'association	<input checked="" type="checkbox"/> 11 <input checked="" type="checkbox"/> 12 <input checked="" type="checkbox"/> 13 <input checked="" type="checkbox"/> 14 <input checked="" type="checkbox"/> 15 <input checked="" type="checkbox"/> 16 <input checked="" type="checkbox"/> 17 <input checked="" type="checkbox"/> sélectionner tout <input type="checkbox"/> dessélectionner tout	Préfixe noms fichiers en sortie ? <input type="text"/> Suffixe noms fichiers en sortie ? <input type="text"/> <input type="checkbox"/> Enregistrer aussi au format .txt <input checked="" type="checkbox"/> Ajouter le nom du fichier source <input type="button" value="Exécuter"/> <input type="button" value="Quitter"/>

Figure n°19 : L'interface d'**Eclats** pour l'éclatement, après le choix par l'utilisateur des différents paramètres de traitement.

L'éclatement dure de quelques secondes à quelques minutes en fonction du volume de données à traiter et du nombre de fichiers à créer. Un message s'affiche en bas et à droite de la fenêtre principale pour indiquer l'étape de traitement en cours. Quand le traitement se termine une alerte s'affiche ; elle indique l'heure de fin normale de l'éclatement. Ici, le temps nécessaire à l'enregistrement des 27 fichiers est de 22 secondes. Après avoir validé avec « OK », on quitte le programme. Une nouvelle utilisation implique de relancer le programme par un clic sur son icône.

En cours d'exécution du programme, l'opérateur peut annuler le traitement (par exemple s'il s'est aperçu qu'il s'est trompé de variable d'éclatement...). Il suffit de cliquer sur le bouton **Interrompre le traitement** qui apparaît dans l'angle supérieur droit de la fenêtre.

Les 27 fichier Excel résultant de l'éclatement sont enregistrés dans le même répertoire que celui du fichier Excel contenant les données statistiques (fig. 20). Le nom du fichier commence par le nom de la variable d'éclatement (COD UF) suivi de la modalité correspondante de cette variable (11, 12, 13...), suivi du caractère « _ », suivi du nom du fichier source (ce dernier est une option, cochée par défaut, dans le dialogue I).

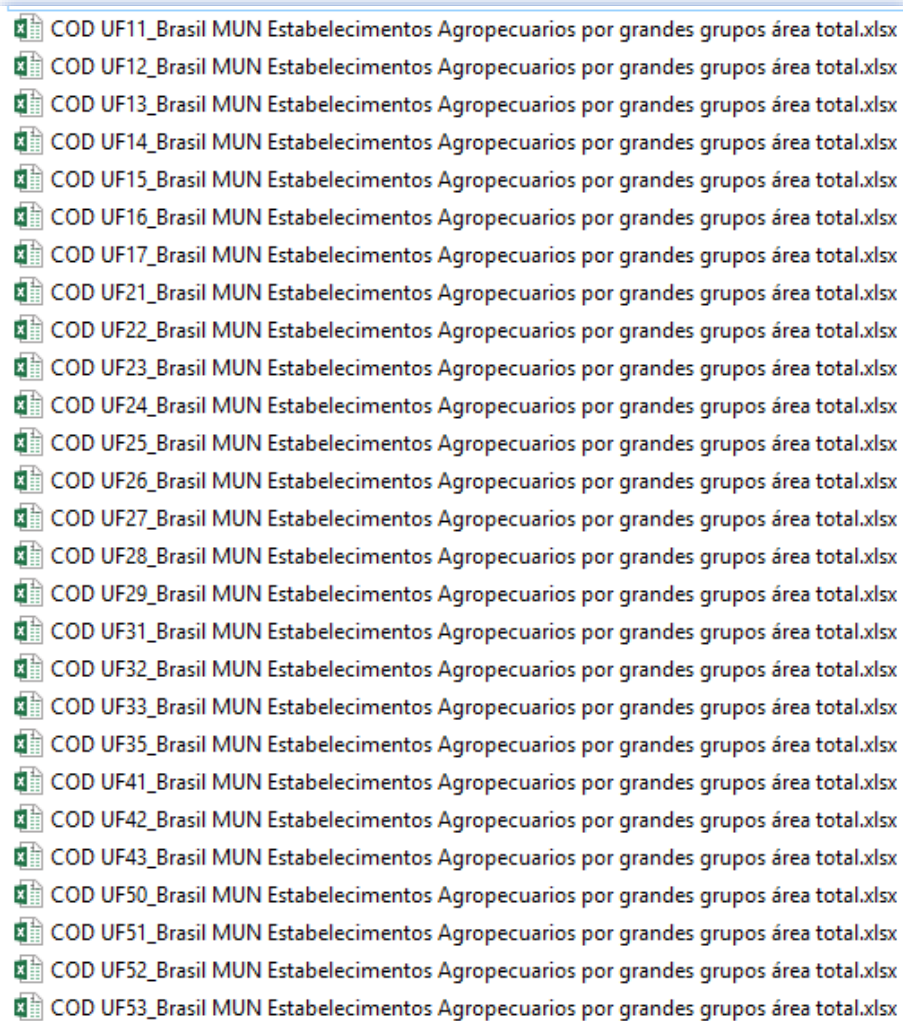


Figure n°20 : Les 27 fichiers résultant de l'éclatement.

Par exemple, ouvrons avec Excel le fichier du Minas Gerais (COD UF=31). On observe une feuille de calcul nommée Eclat composée de 854 lignes et de 18 colonnes (fig. 21). Les lignes correspondent aux 853 communes de cette UF auxquelles il faut ajouter la première ligne, celle des noms de colonnes. Les 18 colonnes sont celles de la feuille de données statistiques initiales. Autrement dit, chaque fichier produit par le traitement d'éclatement d'**Eclats** est une copie partielle de la feuille de données statistiques ouverte initialement.

	A	B	C	D	E	F	G
1	CODGEO	NOM	Estabelecimentos Total_Q_	Mais de 0 a menos de 2 ha_Q_	De 2 a menos de 5 ha_Q_	De 5 a menos de 20 ha_Q_	De 20 a meno
2	3100104	Abadia dos Dourados (MG)	1 056	111	92	224	
3	3100203	Abaeté (MG)	973	4	21	115	
4	3100302	Abre Campo (MG)	1 629	249	342	564	
5	3100401	Acalaca (MG)	290	73	66	85	
6	3100500	Açucena (MG)	748	61	106	199	
7	3100609	Água Boa (MG)	1 725	325	258	484	
8	3100708	Água Comprida (MG)	104	6	4	13	
9	3100807	Aguanil (MG)	326	6	45	136	
10	3100906	Águas Formosas (MG)	1 401	323	145	421	
11	3101003	Águas Vermelhas (MG)	613	21	85	157	
12	3101102	Aimorés (MG)	1 393	38	95	435	
13	3101201	Aiuruoca (MG)	554	33	54	135	
14	3101300	Alagoa (MG)	306	12	27	171	
15	3101409	Albertina (MG)	206	14	42	96	
16	3101508	Além Paraíba (MG)	365	6	19	60	
17	3101607	Alfenas (MG)	959	86	184	340	
18	3101631	Alfredo Vasconcelos (MG)	377	72	105	122	

Figure n°21 : Une copie d'écran partielle de la feuille de calcul du Minas Gerais.

Deuxième partie

Le traitement des Shapefiles .shp

Chapitre 4

Présentation générale du traitement des shapefiles

Eclats 2.xx est un programme utilitaire destiné à être utilisé en complément de **Philcarto** (<http://philcarto.free.fr>). Une des fonctions du programme consiste en l'exploitation des couvertures shapefiles. De nombreuses couvertures sont disponibles sur Internet. Elles se rapportent à des espaces géographiques variés et à des échelles allant du continent au quartier de ville.

1. Que sont les shapefiles ?

Ce format de fichier provient de la société ESRI, spécialiste des systèmes d'information géographique. Elle commercialise notamment le logiciel ArcGIS®. Le succès des logiciels ESRI a fait qu'au cours des ans, le format de fichier Shape, est devenu l'un des standards de fait de l'industrie de l'information géographique. On trouve ainsi de nombreuses bases cartographiques en téléchargement gratuit sur internet au format Shape.

Les shapefiles sont des couvertures géographiques, c'est-à-dire l'enregistrement d'un thème cartographique sous forme numérique, par exemple, les contours des départements français. Il peut y avoir plusieurs couvertures pour un même espace géographique comprenant des éléments cartographiques différents : routes, rivières, etc. Chaque couverture est homogène par rapport au thème qu'elle contient.

Une couverture est composée au minimum d'un triplet de fichiers portant le même nom, mais des types différents. Ces types sont les suivants :

.shp (shape) contient les formes des éléments cartographiques digitalisés sous la forme de séquences de points. Les coordonnées de chaque élément cartographique sont fournies soit sous forme projetée (coordonnées dans le plan de projection X,Y) soit sous forme non projetée (coordonnées angulaires en latitude, longitude).

.dbf (DBase file) renferme les données attributaires de chaque élément cartographique digitalisé figurant dans le fichier .shp. Il s'agit d'une base de données d'un format ancien mais toujours en vigueur dont les différents champs décrivent l'identifiant qui peut occuper plusieurs champs d'enregistrement, le périmètre et la superficie (dans le cas de polygones fermés), et parfois quelques variables statistiques prêtes à être cartographiées.

.shx est un fichier d'index qui permet d'associer les coordonnées du fichier .shp avec les données du fichier .dbf (notamment l'identifiant).

Pour une même couverture, ces trois fichiers suffisent pour une utilisation avec **Philcarto**. Il existe d'autres types de fichiers, comme par exemple **.prj** qui est indispensable pour un affichage du shapefile dans Google Earth. Consulter à ce sujet : <https://fr.wikipedia.org/wiki/Shapefile>

Dans le texte qui suit, on appelle shapefile l'ensemble des fichiers composant une couverture géographique. S'il s'agit d'un type de fichier particulier, comme par exemple .shp, ce type est indiqué ainsi.

La lecture des shapefiles par un logiciel autre que ArcGIS® ne pose pas de problème particulier : la société ESRI a en effet ouvert son format et publié un document technique détaillé. La seule (petite) difficulté réside dans le décodage de certaines valeurs enregistrées en binaire afin de compacter les données.

2. Objectifs

Certains fonds de cartes fournis au format shapefile sont extrêmement volumineux car ils se rapportent à des espaces géographiques très étendus et/ou divisés finement par un grand nombre d'éléments cartographiques. Par exemple, c'est le cas des bases cartographiques distribuées en France par l'IGN, comme « Admin-Express » (2019) téléchargeable à l'adresse : <http://professionnels.ign.fr/adminexpress> et Contours IRIS (2013) téléchargeable à l'adresse <http://professionnels.ign.fr/contoursiris>. Ces deux bases de données qui couvrent l'ensemble de la France métropolitaine contiennent, pour la première, la description de l'ensemble des communes françaises (plus de 36 000), et pour la seconde, l'ensemble des IRIS (Ilots Regroupés pour l'Information Statistique) de chaque département. Dans les deux cas, les recensements annuels donnent gratuitement et chaque année de nombreuses informations statistiques (voir <https://www.insee.fr/fr/information/3561862>) dont la cartographie regorge d'enseignements. Bien d'autres pays proposent des fichiers du même genre, avec des niveaux de maillage et une fréquence de distribution des données différents.

Cependant, on ne travaille pas toujours sur l'ensemble d'un pays ou sur tous les éléments cartographiques d'une de ses bases de données. Les Systèmes d'Information géographiques permettent des sélections d'entités géographiques. Mais quand on souhaite réaliser une telle opération sans recourir à un logiciel souvent complexe, **Eclats** offre une solution pratique et simple d'emploi. Son utilité est la suivante : il lit chaque élément cartographique dans le triplet de fichiers .shp, .shx et .dbf d'une couverture shapefile (auquel s'ajoute aussi un fichier .prj qui n'est pas traité mais simplement copié); ce faisant, il sélectionne chaque élément correspondant aux conditions de sélection choisies par l'utilisateur, puis l'écrit ensuite dans un nouveau triplet de fichiers contenant seulement les éléments sélectionnés. Autrement dit, **Eclats** lit en entrée le triplet composant une observation d'un shapefile, et enregistre en sortie dans un autre shapefile un triplet composé des seuls éléments cartographiques correspondant aux critères de sélection. Le triplet de sortie peut ensuite être aisément importé directement dans **Philcarto** ou **Phildigit**.

Eclats est surtout destiné au traitement des fonds de carte volumineux mais peut être utilisé sur n'importe quel triplet de fichiers .shp, .shx et .dbf pour y réaliser des sélections.

3. Données shapefile à traiter

Pour fonctionner, **Eclats** a besoin d'une couverture shapefile telle que celle des IRIS de la France métropolitaine et des départements d'Outre-Mer en 2013 téléchargeable à l'adresse : <http://professionnels.ign.fr/contoursiris>. L'IGN publie chaque année une couverture mise à jour.

Sur la page de l'IGN, cliquer sur l'onglet « Téléchargement ». Le fichier à télécharger est « Contours...Iris® France entière édition 2013 par territoires (7z de 33,5 Mo) » Télécharger ce fichier .zip France entière par territoires, puis décompresser ce fichier sur le bureau. Les données à utiliser après décompression du fichier téléchargé sont dans le dossier C:\Bureau\CONTOURS-IRIS_1-0_SHP_LAMB93_FXX_2013-01-01 si le dossier décompressé a été placé sur le bureau de Windows. Dans cette couverture, on trouve les données relatives à l'ensemble des communes de France métropolitaine.

Pour simplifier la présentation, on utilise ici un extrait de cette couverture relatif au seul département de la Gironde. Cet extrait est présent dans le dossier Eclats_Test\Eclats_Test_Gironde\ . Le dossier Eclats_Test est téléchargeable sur la même page de que **Eclats** sur le site **Philcarto**.

Le dossier CONTOURS-IRIS_1-0_SHP_LAMB93_D33-2013 figurant dans le dossier Eclats_Test\Eclats_Test_Gironde\ contient les contours des IRIS des communes du département 33, c'est-à-dire le Département de la Gironde. Pour les communes qui ne sont pas « irisée », c'est leur contour communal qui est donné.

Le dossier renferme 5 fichiers (fig. 22). Trois de ces fichiers sont utilisés par **Eclats** : CONTOURS-IRIS33.dbf, CONTOURS-IRIS33.shp et CONTOURS-IRIS33.shx. Le fichier CONTOURS-IRIS33.prj est utile avec d'autres logiciels (tel Google Earth). Le fichier Liste33.dbf contient la liste des communes « irisées » avec leur nombre respectifs d'IRIS. Les fichiers de types .dbf peuvent être visualisés et même modifiés par le module Classeur du logiciel OpenOffice.org

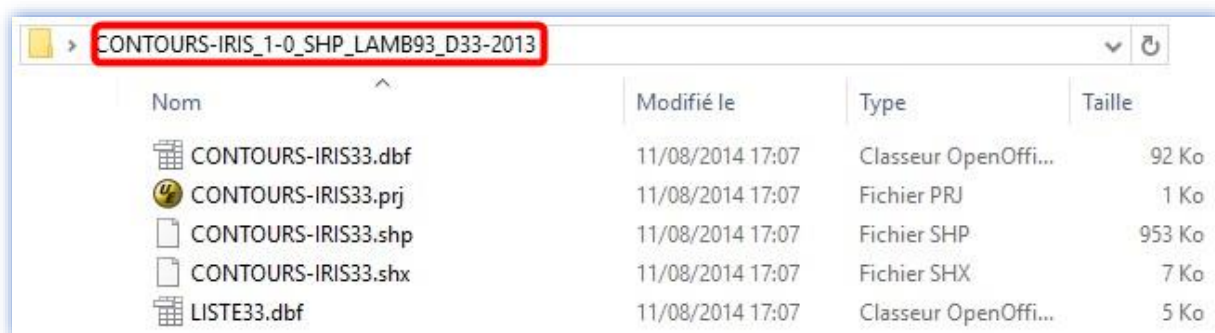


Figure n°22 : les fichiers du shapefile CONTOURS-IRIS_1-0_SHP_LAMB93_D33-2013.

Dès que l'utilisateur a accepté la licence d'utilisation d'**Eclats**, et cliqué sur le bouton **Sélection simple ou multiple**, il lui est demandé d'ouvrir le fichier .shp dans lequel il souhaite réaliser une sélection, ici CONTOURS-IRIS33.shp (fig. 23). Un clic sur le bouton Ouvrir déclenche la lecture des trois fichiers .dbf, .shp, .shx. Cette lecture est plus ou moins longue en fonction de la taille de ces fichiers.

La lecture des fichiers étant achevée, **Eclats** affiche son interface. Elle est composée d'une fenêtre de couleur bleu clair (fig. 24). Dans la partie supérieure est affiché en rouge le nom du fichier en cours de traitement. Au centre-gauche, on trouve une liste de champs. Il s'agit des noms des différentes colonnes du fichier .dbf. Les valeurs de ces champs permettent notamment d'identifier les éléments cartographiques enregistrés dans le fichier .shp. Pour opérer une sélection d'éléments cartographiques dans le fichier .shp, l'utilisateur doit sélectionner dans cette liste le champ qui identifie les éléments sélectionnés.

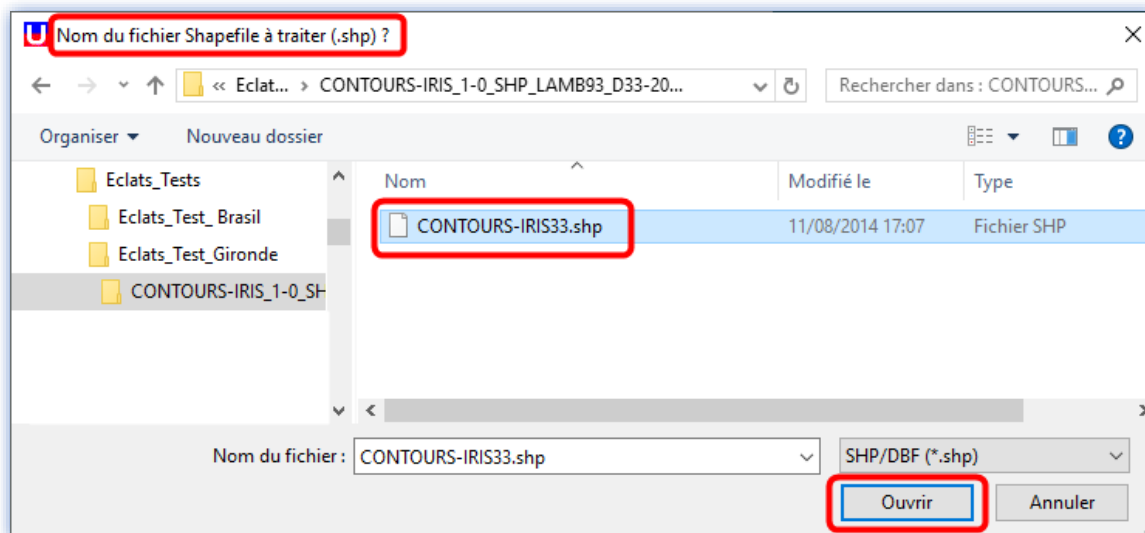


Figure n°23 : la sélection et l'ouverture du fichier CONTOURS-IRIS33.shp.

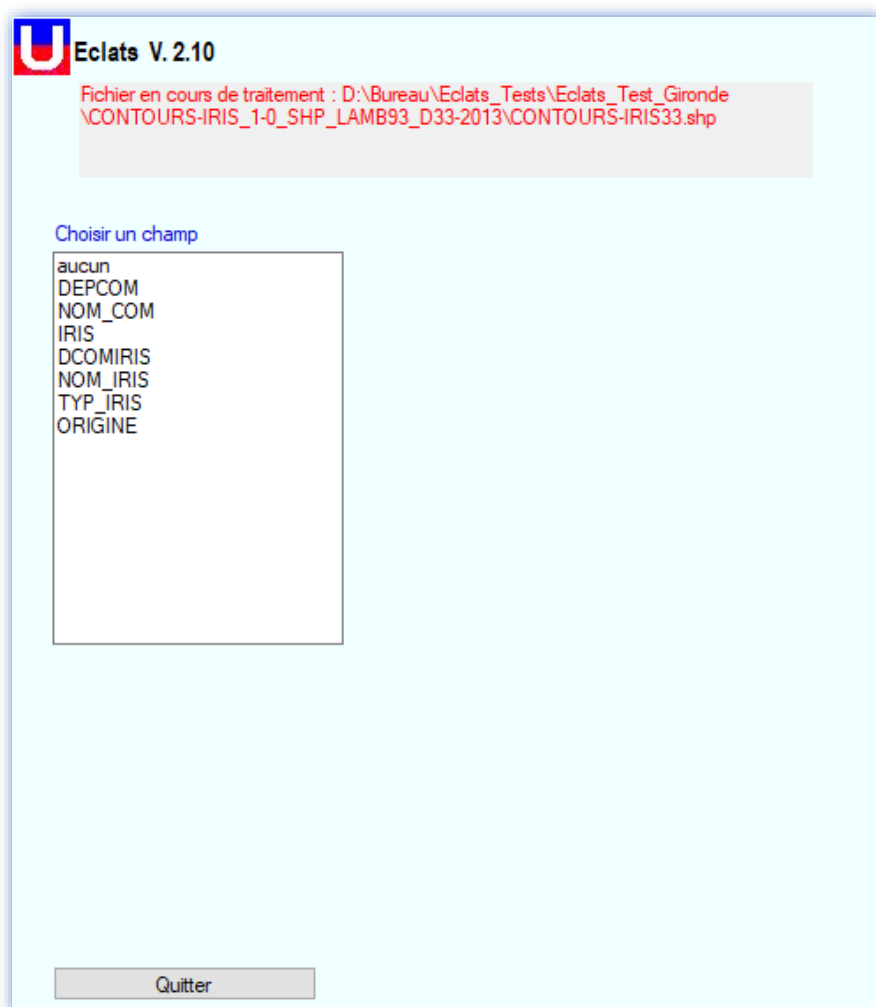


Figure n°24 : l'interface d'Eclats après l'ouverture de CONTOURS-IRIS33.shp.

Chapitre 5

Les différents modes de sélection dans les shapefiles

Avec **Eclats**, la sélection d'éléments cartographiques dans un shapefile peut être réalisée de trois façons différentes : directement en cliquant sur la liste des éléments à retenir après avoir sélectionné l'un des champs proposé dans la liste « Sélection de l'un de ces champs » ; indirectement en donnant au programme un fichier texte contenant la liste des éléments à retenir ; indirectement en opérant plusieurs sélections en donnant un fichier texte contenant les noms des différentes sélections à réaliser suivis des éléments composant chacune d'elles. Ces trois modes de sélection donnent au programme une flexibilité lui permettant de prendre en compte différentes situations.

1. La sélection directe

La communauté urbaine de Bordeaux dénommée « Bordeaux Métropole » est un Etablissement Public de Coopération Intercommunale (EPCI) comprenant 28 communes. La commune de Bordeaux est l'une d'elles ; son code géographique officiel est 33063 (tab. 1). Attention, il ne s'agit pas du code postal mais d'un code défini par l'INSEE (voir <https://www.insee.fr/fr/information/3363419>).

CODGEO	NOM	CODE_EPCI	NOM_EPCI
33003	AMBARÈS-ET-LAGRAVE	243300316	BORDEAUX MÉTROPOLE
33004	AMBÈS	243300316	BORDEAUX MÉTROPOLE
33013	ARTIGUES-PRÈS-BORDEAUX	243300316	BORDEAUX MÉTROPOLE
33032	BASSENS	243300316	BORDEAUX MÉTROPOLE
33039	BÈGLES	243300316	BORDEAUX MÉTROPOLE
33056	BLANQUEFORT	243300316	BORDEAUX MÉTROPOLE
33063	BORDEAUX	243300316	BORDEAUX MÉTROPOLE
33065	BOULIAC	243300316	BORDEAUX MÉTROPOLE
33069	LE BOUSCAT	243300316	BORDEAUX MÉTROPOLE
33075	BRUGES	243300316	BORDEAUX MÉTROPOLE
33096	CARBON-BLANC	243300316	BORDEAUX MÉTROPOLE
33119	CENON	243300316	BORDEAUX MÉTROPOLE
33162	EYSINES	243300316	BORDEAUX MÉTROPOLE
33167	FLOIRAC	243300316	BORDEAUX MÉTROPOLE
33192	GRADIGNAN	243300316	BORDEAUX MÉTROPOLE
33200	LE HAILLAN	243300316	BORDEAUX MÉTROPOLE
33249	LORMONT	243300316	BORDEAUX MÉTROPOLE
33273	MARTIGNAS-SUR-JALLE	243300316	BORDEAUX MÉTROPOLE
33281	MÉRIGNAC	243300316	BORDEAUX MÉTROPOLE
33312	PREMPUYRE	243300316	BORDEAUX MÉTROPOLE
33318	PESSAC	243300316	BORDEAUX MÉTROPOLE
33376	SAINT-AUBIN-DE-MÉDOC	243300316	BORDEAUX MÉTROPOLE
33434	SAINT-LOUIS-DE-MONTFERRAND	243300316	BORDEAUX MÉTROPOLE
33449	SAINT-MÉDARD-EN-JALLES	243300316	BORDEAUX MÉTROPOLE
33487	SAINT-VINCENT-DE-PAUL	243300316	BORDEAUX MÉTROPOLE

33519	LE TAILLAN-MÉDOC	243300316	BORDEAUX MÉTROPOLE
33522	TALENCE	243300316	BORDEAUX MÉTROPOLE
33550	VILLENAVE-D'ORNON	243300316	BORDEAUX MÉTROPOLE

Tableau n°1 : la liste des communes membres de « Bordeaux Métropole »

On souhaite extraire du shapefile CONTOURS-IRIS33.shp les IRIS de la commune de Bordeaux. Voici comment opérer : dans la liste « Choisir un champ » située à gauche dans la fenêtre d'**Eclats**, cliquer sur DEPCOM qui contient les identifiants des communes (code du département + code de la commune dans le département). A la suite de ce clic, la liste « Choisir un ou plusieurs codes de sélection » se remplit avec tous les identifiants des différentes communes trouvées dans le fichier CONTOURS-IRIS33.dbf. Parmi ces identifiants se trouve celui de Bordeaux (33063). Rechercher cet identifiant en faisant dérouler la liste avec l'ascenseur ; une fois 33063 localisé, cliquer sur la case à cocher correspondante (fig. 25). On pourrait, si nécessaire sélectionner ainsi d'autres communes voire toutes les communes (avec le feu vert situé à gauche de la liste de cases à cocher). Ici, on souhaite seulement les IRIS de Bordeaux, donc la seule case 33063 doit être cochée.

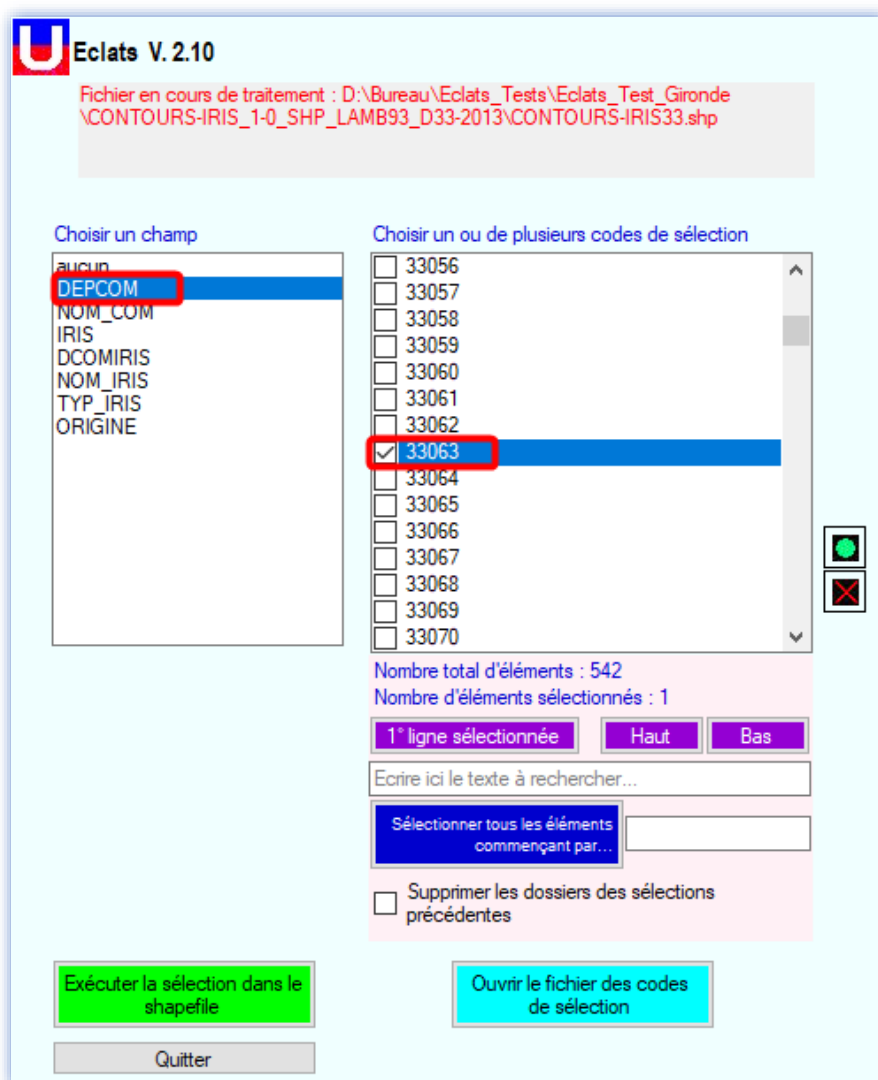


Figure n°25 : la sélection et l'ouverture du fichier CONTOURS-IRIS33.shp.

Le cadre violet situé au-dessous de la liste de cases à cocher (liste située à droite) permet de naviguer dans celle-ci. Différentes possibilités sont proposées, utiles notamment lorsque cette liste est longue (fig. 26) et qu'on effectue une sélection directe.

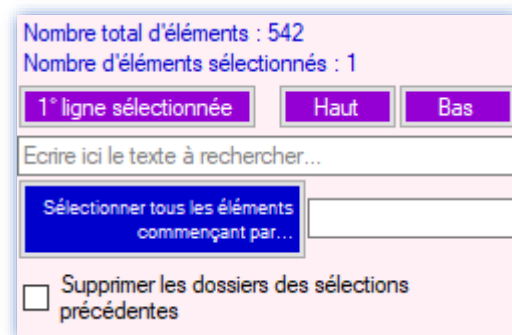


Figure n°26 : le dialogue de navigation dans liste des éléments à sélectionner

Deux informations sont données à l'utilisateur : le nombre total d'éléments différents dans la liste et le nombre d'éléments sélectionnés (cochés). Suivent trois boutons permettant d'afficher soit la première ligne sélectionnée (cochée), soit le haut de la liste à partir de la première ligne, soit le bas de la liste, à partir de la dernière ligne. L'utilisateur peut aussi donner un texte à rechercher parmi tous les éléments de la liste ; le premier élément trouvé est affiché en haut de liste. Enfin, il est possible de sélectionner (de cocher) tous les éléments commençant par une chaîne de caractère donnée (par exemple un numéro de

département...); une fois la chaîne écrite, un clic sur le bouton sélectionne (coche) tous les éléments de la liste commençant par la chaîne donnée. Il est possible d'annuler la sélection avec le bouton Enfin, une case à cocher (non cochée par

défaut) permet d'éliminer les dossiers des sélections précédentes s'il y en a dans le dossier où se trouvent le shapefile à traiter.

Dès qu'une case à cocher est cliquée, le bouton qui permet d'exécuter la sélection fait son apparition. Pour opérer la sélection, cliquer sur ce bouton vert... Le traitement a lieu et **Eclats** affiche un message de fin de sélection (fig. 27).

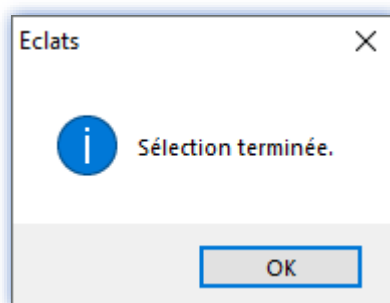


Figure n°27 : la sélection est terminée.

Le résultat est un nouveau shapefile enregistré dans un dossier créé par **Eclats** dans le dossier contenant le shapefile initial. Ce dossier s'appelle **Sélection**. Il contient les trois fichiers nécessaires à une importation dans **Phildigit** ou **Philcarto** : .dbf., .shp, .shx auxquels s'ajoute le fichier .prj utile, par exemple, pour importer cette nouvelle couverture dans GoogleEarth (fig. 28).

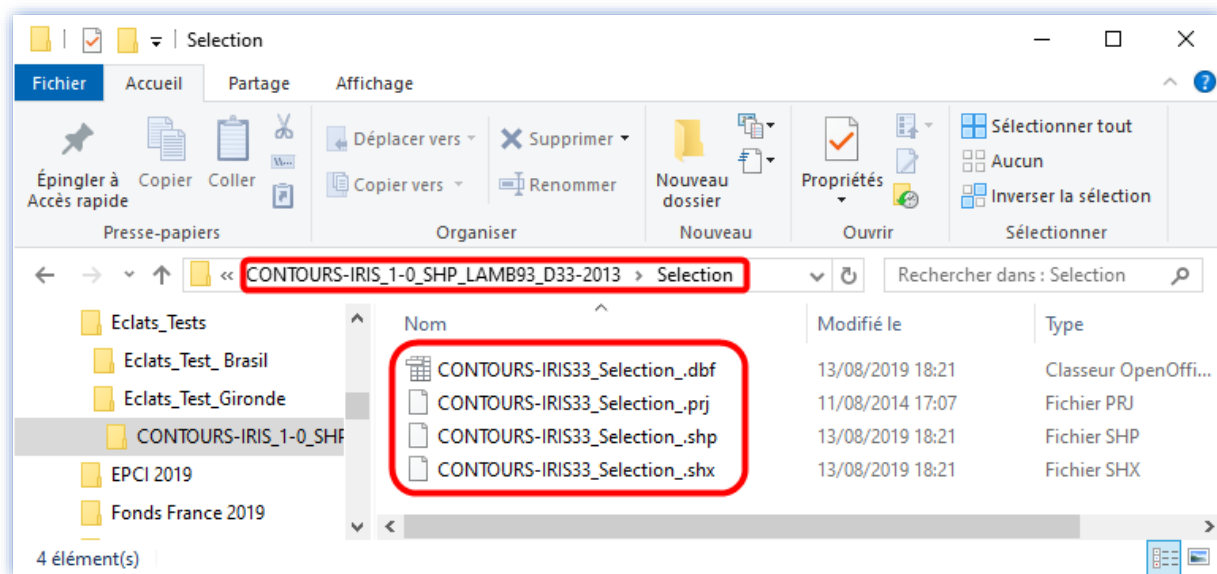


Figure n°28 : les quatre fichiers créés par la sélection directe.

Le type de chaque fichier est celui des fichiers d'origine ; mais au nom de chacun d'eux a été ajouté le suffixe `_selection_`. On pourrait renommer ces fichiers en Bordeaux-IRIS, par exemple, suivi de leurs types respectifs.

La visualisation de la sélection par importation dans **Philcarto** (fig. 29) montre les IRIS de la ville de Bordeaux. Mission accomplie pour **Eclats** !

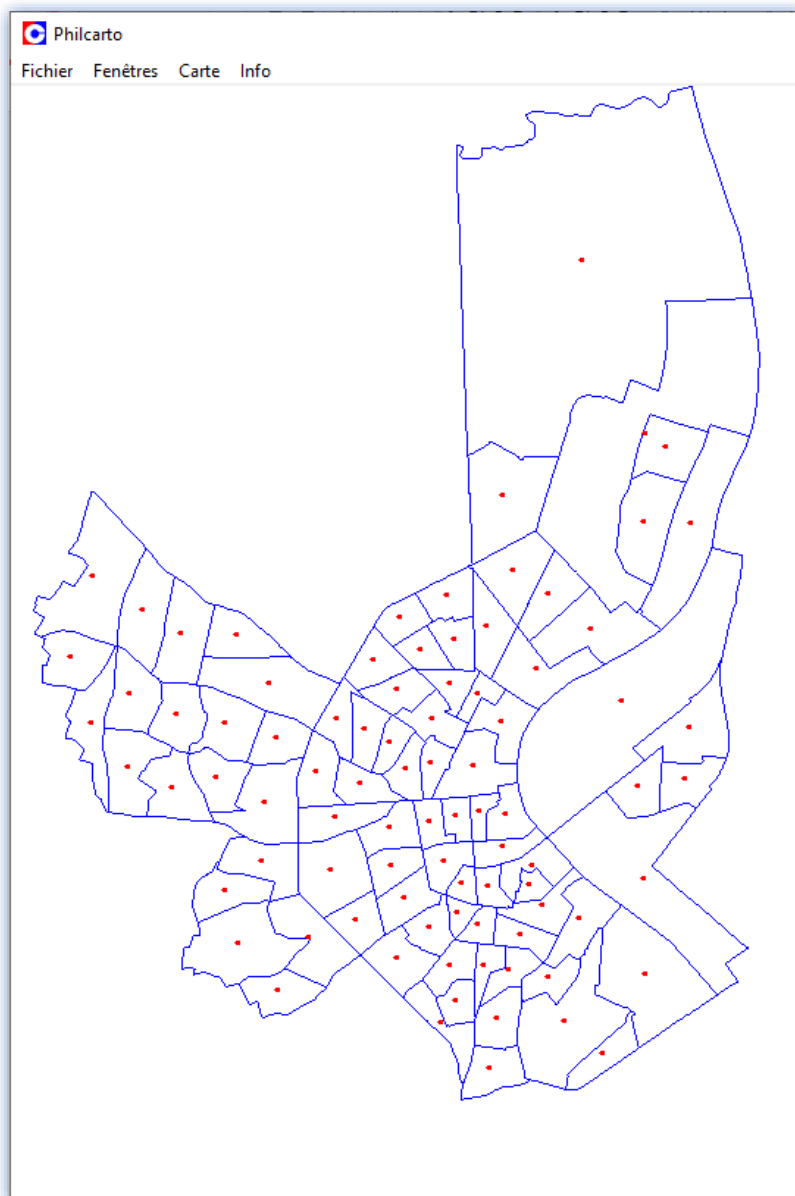
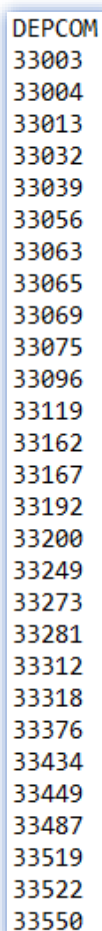


Figure n°29 : La visualisation de la sélection par **Philcarto**.
Les points rouges sont les centroïdes des IRIS ajoutés par **Philcarto** pour rendre possible la réalisation de cartes en cercles (centres des cercles).

2. La sélection indirecte avec fichier d'identifiants

Dans l'exemple précédent, ce sont les IRIS de la seule commune de Bordeaux qui ont été extraits du shapefile du département de la Gironde. Ici, on cherche à extraire les IRIS de l'ensemble des 28 communes de la métropole bordelaise. Comme précédemment, on pourrait sélectionner ces 28 communes l'une après l'autre en cochant leurs cases respectives dans la liste « Sélection d'une ou plusieurs de ces modalités : ». C'est faisable, mais fastidieux, et surtout sujet à erreurs surtout quand le nombre de communes à retenir est grand. C'est la raison pour laquelle **Eclats** propose d'exploiter le contenu d'un fichier texte dans lequel chaque ligne contient l'élément d'identification de chaque commune à retenir selon les valeurs du champ sélectionné dans la liste « Sélection de l'un de ces champs : ».

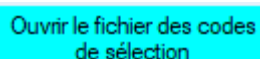
Avant toute chose, il faut créer un fichier texte (par exemple avec le Bloc-Notes de Windows, ou bien tout autre programme capable de créer un fichier .txt). Ce fichier contient les identifiants des communes à retenir, ici sur le champ DEPCOM (fig. 30). La première ligne renferme le nom de la colonne, ici DEPCOM.



```
DEPCOM
33003
33004
33013
33032
33039
33056
33063
33065
33069
33075
33096
33119
33162
33167
33192
33200
33249
33273
33281
33312
33318
33376
33434
33449
33487
33519
33522
33550
```

Figure n°30 : La liste des codes INSEE des 28 communes de Bordeaux Métropole enregistrée dans le fichier texte : Listelden_BordeauxMetropole.TXT.

Après avoir enregistré ce fichier texte et quitté l'éditeur de texte, on lance l'exécution d'**Eclats**. On ouvre la couverture shapefile, puis on sélectionne le champ DEPCOM dans la liste « Choisir un champ ». Au lieu de cocher les cases des identifiants de chaque commune,



Ouvrir le fichier des codes
de sélection

on clique sur le bouton . Un dialogue d'ouverture d'un fichier texte permet de sélectionner le fichier de sélection, ici Listelden_BordeauxMetropole.TXT (fig. 31).

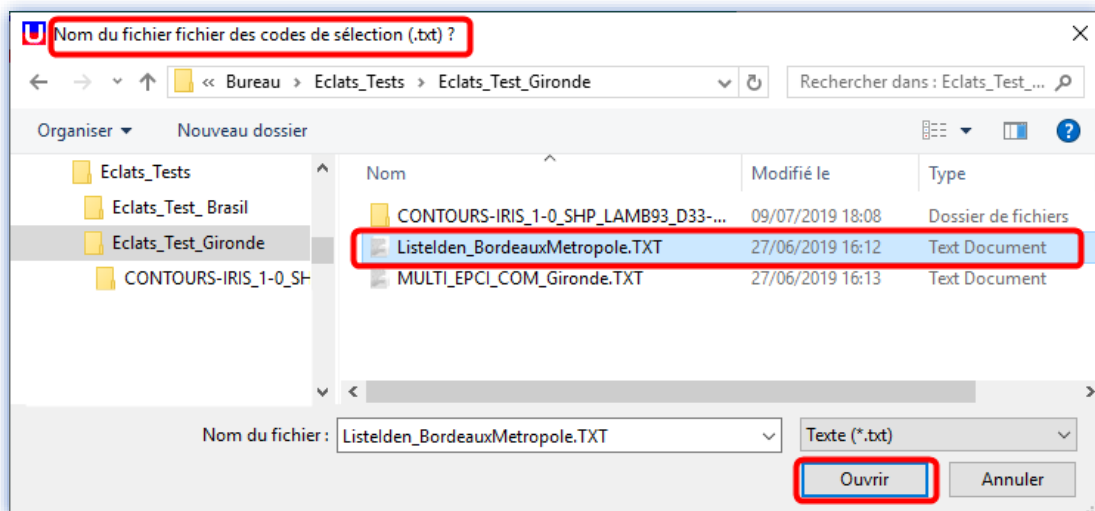


Figure n°31 : L'ouverture du fichier de sélection des communes Listelden_BordeauxMetropole.TXT.

Une fois ouvert le fichier de sélection des communes, **Eclats** coche, dans la liste « Choisir un ou plusieurs codes de sélection », les cases des communes à retenir (fig. 32).

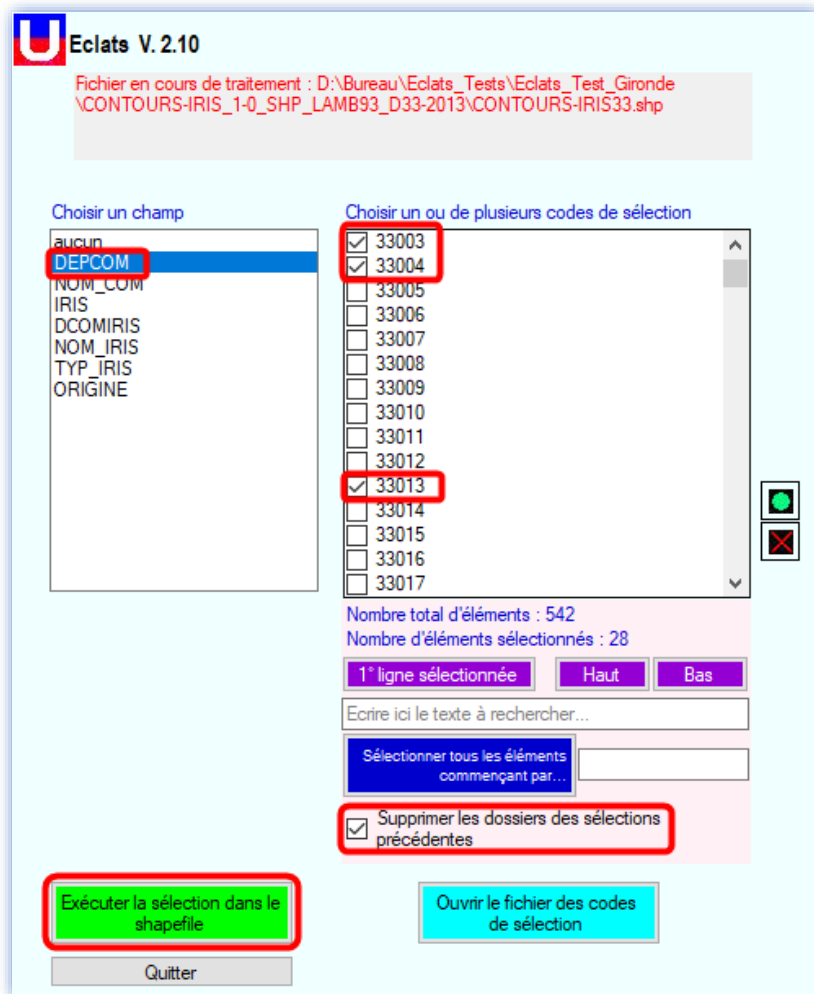


Figure n°32 : Les identifiants des communes figurant dans le fichier Listelden_BordeauxMetropole.TXT sont maintenant cochés. Noter que le nombre d'éléments sélectionnés est égal au nombre de communes de la Métropole.

Avant de lancer le traitement, on veille à supprimer d'anciennes sélections du dossier du shapefile à traiter ☒ Supprimer les dossiers des sélections précédentes . Le bouton Exécuter la sélection dans le shapefile qui permet d'exécuter la sélection a fait son apparition car au moins une case est cochée. Pour opérer la sélection des IRIS des communes de la métropole bordelaise, cliquer sur ce bouton vert... Le traitement a lieu et **Eclats** affiche un message de fin de sélection (fig. 33).

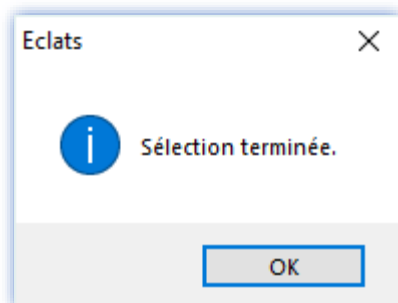


Figure n°33 : la sélection est terminée.

Le résultat est un nouveau shapefile enregistré dans un dossier créé par **Eclats** dans le dossier contenant le shapefile initial. Ce dossier s'appelle Selection. Il contient les trois fichiers nécessaires à une importation dans **Phildigit** ou **Philcarto** : .dbf., .shp, .shx (fig. 34).

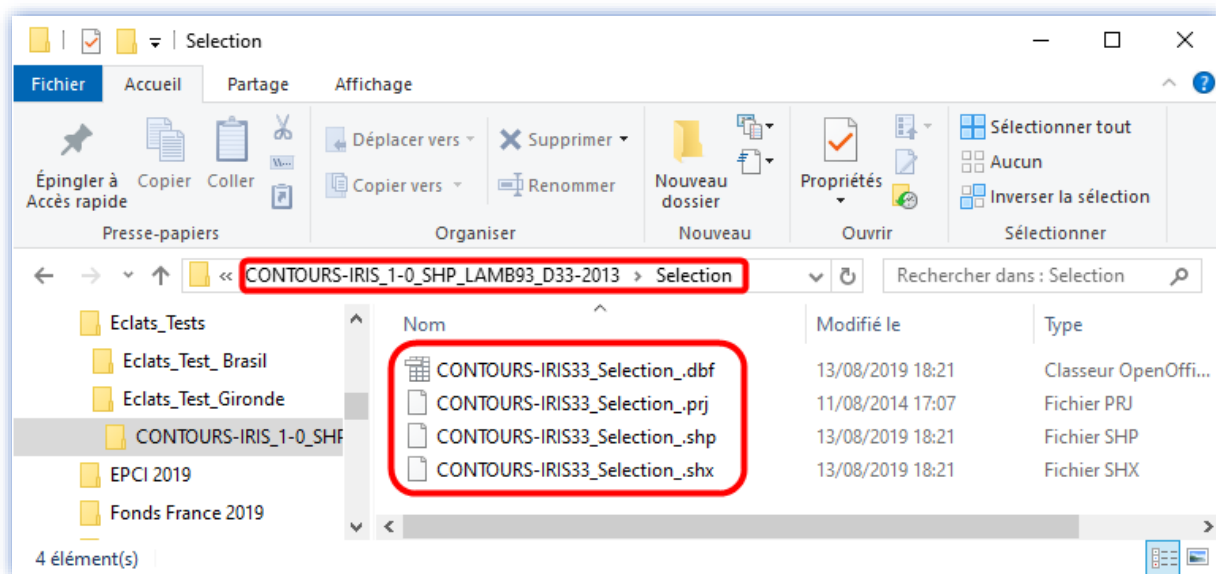


Figure n°34 : les quatre fichiers créés par la sélection indirecte.

Le type de chaque fichier est celui des fichiers d'origine ; mais au nom de chacun d'eux a été ajouté le suffixe _selection_. On pourrait modifier les trois noms en Bordeaux-IRIS, par exemple, suivi de leurs types respectifs.

La visualisation de la sélection par importation dans **Philcarto** (fig. 35) montre les IRIS de la ville de Bordeaux. Un passage par **Phildigit** permettrait d'ajouter les contours des communes (fonction « Regroupements ») pour habiller la carte des IRIS. Comme précédemment, la mission a été accomplie par **Eclats** !

Attention ! Lors de l'importation des IRIS sélectionnés dans **Philcarto** ou **Phildigit**, le champ d'identification des éléments cartographiques est l'identifiant des IRIS, appelé DCOMIRIS (code commune + numéro d'IRIS).

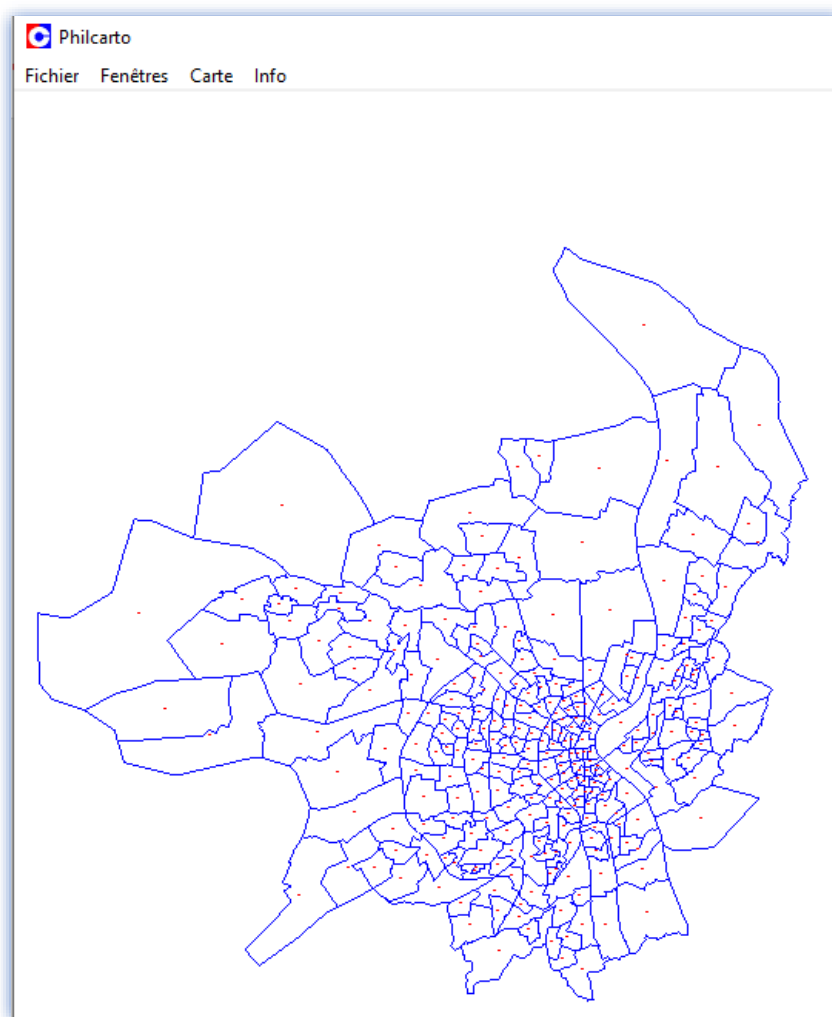


Figure n°35 : La visualisation de la sélection par **Philcarto**.
Les points rouges sont les centroïdes des IRIS ajoutés par **Philcarto** pour rendre possible la réalisation de cartes en cercles (centres des cercles).

3. Les sélections indirectes multiples avec fichier d'identifiants

Dans l'exemple précédent, ce sont les IRIS des communes de la seule Métropole de Bordeaux qui ont été extraits du shapefile du département de la Gironde. Ici, on cherche à extraire les IRIS des deux EPCI du département de la Gironde. La méthode présentée ci-après est généralisable à un nombre quelconque d'EPCI comme ceux d'un autre département, d'une autre région, ou bien encore des sous-ensembles géographiques multiples définis sur un autre découpage que celui des EPCI. Le résultat attendu est un shapefile avec ses trois fichiers enregistrés dans autant de dossiers qu'il y a de sous-ensembles géographiques définis dans la première colonne du fichier de sélection multiple.

Avant toute chose, il faut créer un fichier texte (par exemple avec le Bloc-Notes de Windows, ou bien tout autre programme capable de créer un fichier .txt). Le nom de ce fichier doit commencer par le mot MULTI (écrit obligatoirement de cette façon car il fait comprendre à **Eclats** qu'il va opérer des sélections multiples). Ce fichier contient deux colonnes pour chacune des communes à retenir. La première colonne, appelée NOM_EPCI, renferme le nom de chacun des sous-ensembles attendus, ici ceux deux des EPCI du département de la

Gironde, Bordeaux Métropole avec ses 28 communes et la Communauté d'Agglomération Bassin d'Arcachon Sud-Pôle Atlantique (COBAS) avec ses 4 communes. La seconde colonne contient les identifiants des communes, ici DEPCOM. Le fichier de sélection comprend donc en tout 33 lignes : 28 pour Bordeaux, 4 pour Arcachon et une pour les noms de colonne (fig. 36).

NOM_EPCI	DEPCOM
BORDEAUX MÉTROPOLE	33003
BORDEAUX MÉTROPOLE	33004
BORDEAUX MÉTROPOLE	33013
BORDEAUX MÉTROPOLE	33032
BORDEAUX MÉTROPOLE	33039
BORDEAUX MÉTROPOLE	33056
BORDEAUX MÉTROPOLE	33063
BORDEAUX MÉTROPOLE	33065
BORDEAUX MÉTROPOLE	33069
BORDEAUX MÉTROPOLE	33075
BORDEAUX MÉTROPOLE	33096
BORDEAUX MÉTROPOLE	33119
BORDEAUX MÉTROPOLE	33162
BORDEAUX MÉTROPOLE	33167
BORDEAUX MÉTROPOLE	33192
BORDEAUX MÉTROPOLE	33200
BORDEAUX MÉTROPOLE	33249
BORDEAUX MÉTROPOLE	33273
BORDEAUX MÉTROPOLE	33281
BORDEAUX MÉTROPOLE	33312
BORDEAUX MÉTROPOLE	33318
BORDEAUX MÉTROPOLE	33376
BORDEAUX MÉTROPOLE	33434
BORDEAUX MÉTROPOLE	33449
BORDEAUX MÉTROPOLE	33487
BORDEAUX MÉTROPOLE	33519
BORDEAUX MÉTROPOLE	33522
BORDEAUX MÉTROPOLE	33550
CA BASSIN D'ARCACHON SUD-PÔLE ATLANTIQUE (COBAS)	33009
CA BASSIN D'ARCACHON SUD-PÔLE ATLANTIQUE (COBAS)	33199
CA BASSIN D'ARCACHON SUD-PÔLE ATLANTIQUE (COBAS)	33527
CA BASSIN D'ARCACHON SUD-PÔLE ATLANTIQUE (COBAS)	33529

Figure n°36 : La liste des communes du département de la Gironde précédées du nom des EPCI dont elles sont membres. Le fichier s'appelle : MULTI_EPCI_COM_Gironde.TXT.

La seconde colonne est le code de chacune des 32 communes. Ces deux colonnes sont séparées par le caractère « tabulation » (TAB sur certains claviers). Le procédé de sélection est le suivant : grâce au préfixe MULTI du nom du fichier de sélection, **Eclats** scrute la première colonne afin de définir le nombre de sous-ensembles géographiques à sélectionner. Dans un second temps, il réitère la sélection des éléments cartographique pour chaque sous-ensemble géographique qu'il enregistre dans un shapefile placé dans un dossier dont le nom correspond à celui du sous-ensemble géographique (les noms figurant dans la première colonne du fichier de sélection).

Le mode opératoire est le même que pour la sélection indirecte avec fichier d'identifiants présentée en 4.2. La seule différence est qu'il faut ouvrir le fichier MULTI_EPCI_COM_Gironde.TXT (fig. 37) qui comprend deux colonnes au lieu du fichier Listelden_BordeauxMetropole.TXT qui n'a qu'une seule colonne.

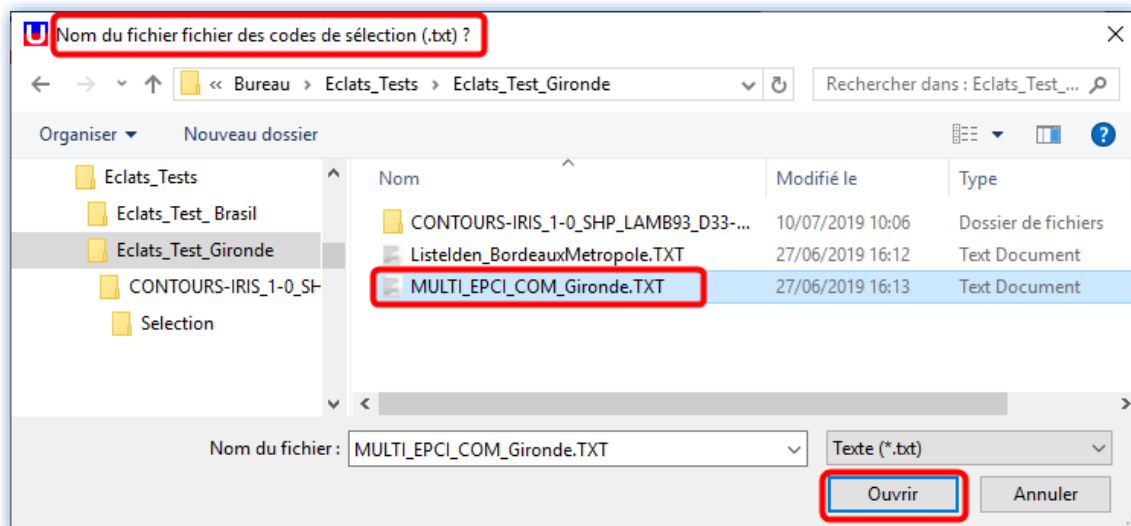


Figure n°37 : L'ouverture du fichier de sélection multiple des communes MULTI_EPCI_COM_Gironde.TXT.

En cours de traitement, le programme informe l'utilisateur sur l'état d'avancement du traitement et indique le nom de chaque dossier créé (fig. 38). Dans ce cadre, on trouve : 1. le nombre de sélections à réaliser ; 2. Le numéro de sélection en cours ; 3. Le nom du dossier en cours de création ; une barre d'avancement qui progresse à chaque sélection achevée. Le bouton **Interrompre le traitement** permet d'arrêter l'exécution du programme (utile en cas d'erreur manifeste). Le nombre de sélections à réaliser correspond au nombre de modalités différentes en 1^o colonne du fichier de sélection.

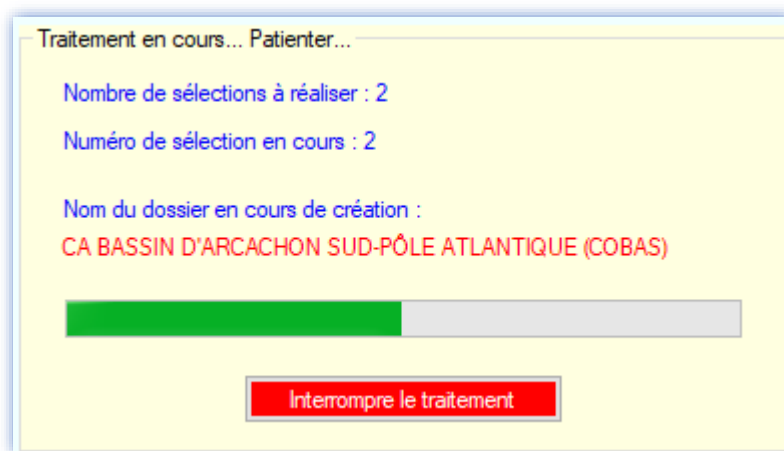


Figure n°38 : La fenêtre d'information sur les sélections multiples en cours de traitement.

Le traitement a lieu et **Eclats** affiche un message de fin de sélection (fig. 39) qui indique où ces sélections multiples ont été enregistrées, combien de dossiers ont été créés ainsi que le type d'éléments que contient chacun de ces dossiers.

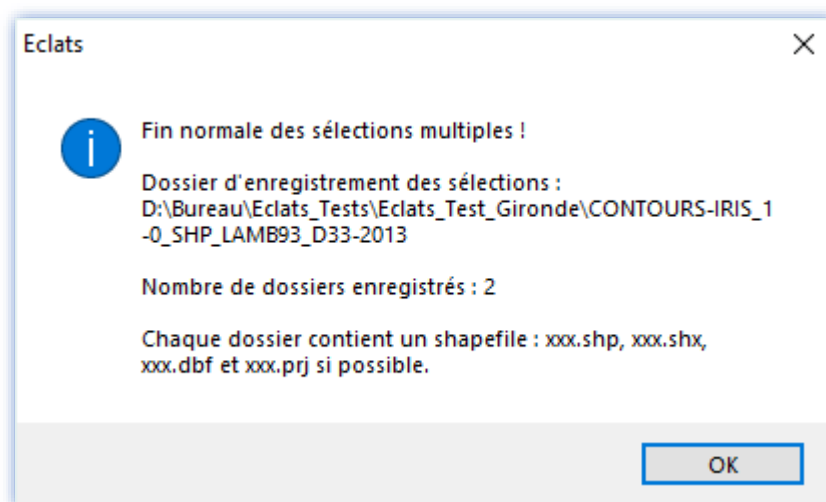


Figure n°39 : la sélection multiple est terminée.

Chacun des dossiers renfermant le shapefile correspondant à une sélection multiple est enregistré dans le dossier contenant le shapefile initial. Le nom de ce dossier correspond à la modalité correspondante de la première colonne du fichier de sélection multiple (fig. 40).

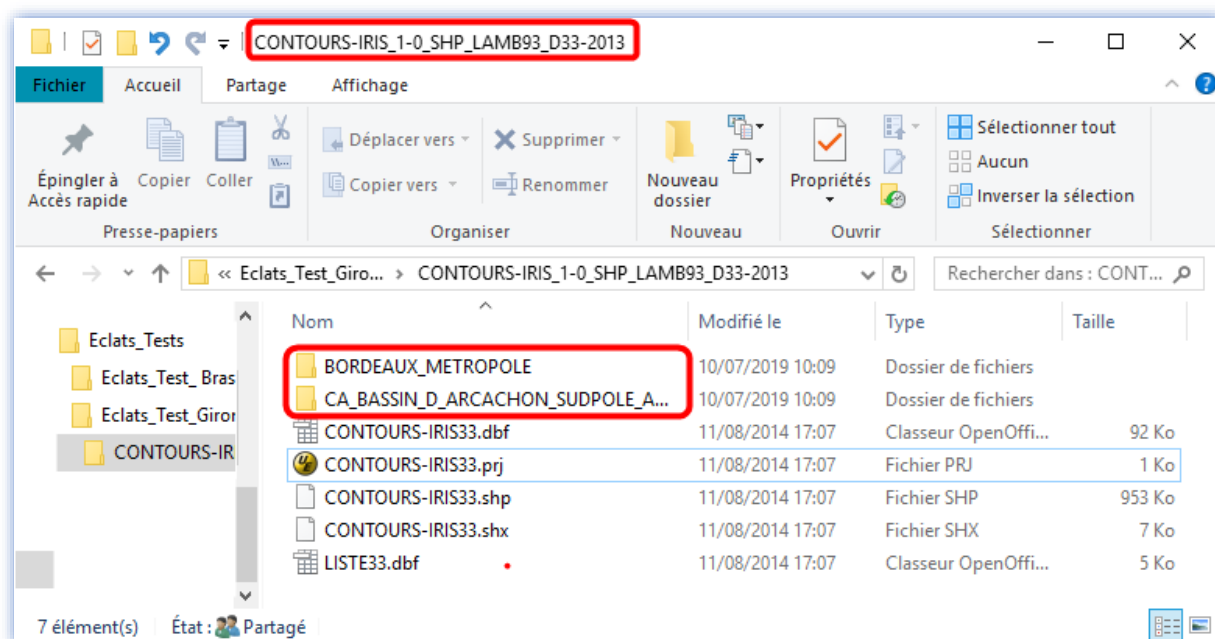


Figure n°40 : Les deux dossiers créés par la sélection multiple.

Dans chacun des dossiers de sélection multiple, on trouve un shapefile composé des trois fichiers nécessaires à une importation dans **Phildigit** ou **Philcarto** : .dbf., .shp, .shx (fig. 41).

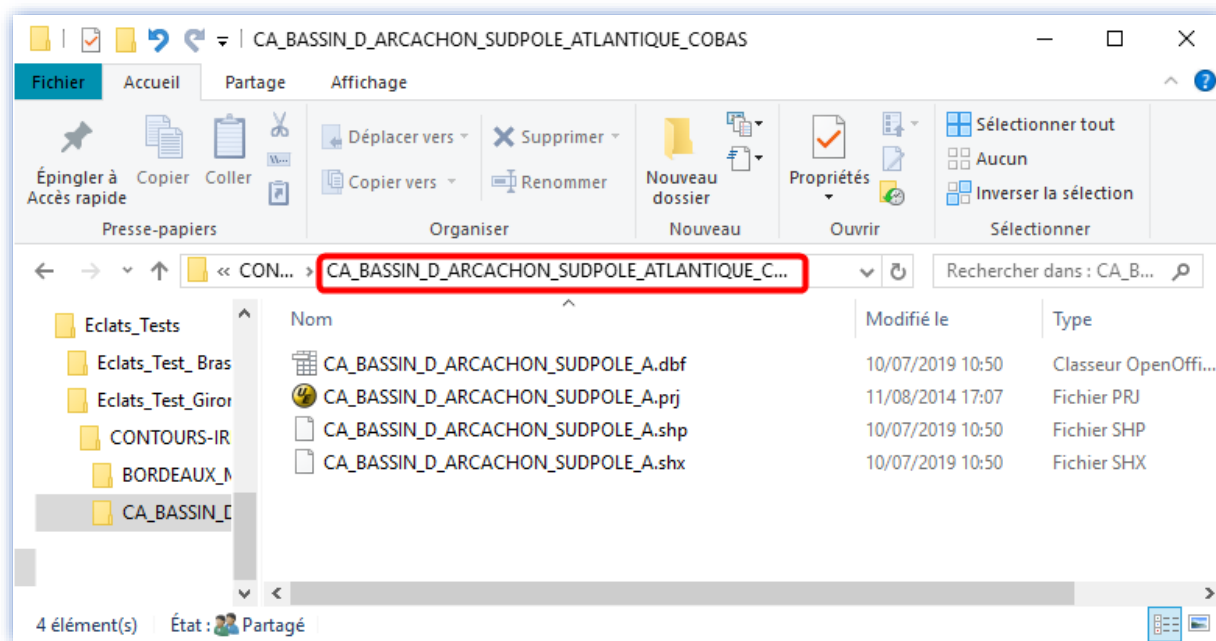


Figure n°41 : les quatre fichiers créés par la sélection multiple ; ici l'EPCI du Sud du bassin d'Arcachon.

Le type de chaque fichier est celui des fichiers d'origine ; le nom du shapefile, commun au triplet de fichiers, correspond à la modalité de la première colonne du fichier de sélection multiple. Pour nommer le shapefile, la modalité est modifiée de manière à supprimer les caractères accentués ; et si cette modalité comprend plus de 30 caractères, seuls les 30 premiers caractères sont retenus.

La visualisation de la sélection par importation dans **Philcarto** (fig. 42) montre les IRIS des communes du Sud bassin d'Arcachon. Un passage par **Phildigit** permettrait d'ajouter les contours des communes (fonction « Regroupements » pour habiller la carte des IRIS.

Attention ! Lors de l'importation des IRIS sélectionnés dans **Philcarto** ou **Phildigit**, ou lors de leur conversion en fichier .AI par **Eclats**, le champ d'identification des éléments cartographiques est l'identifiant des IRIS, appelé DCOMIRIS (code commune + numéro d'IRIS).

Comme précédemment, la mission a été accomplie par **Eclats** !

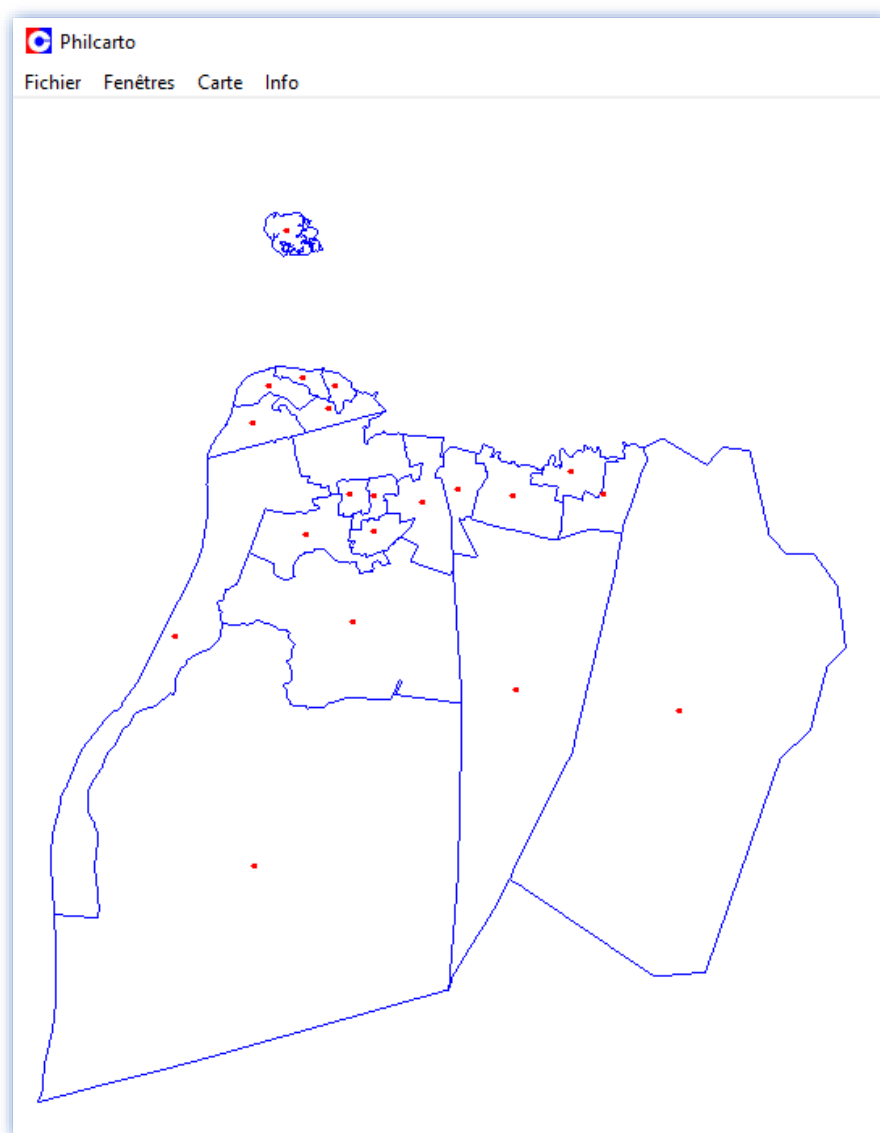


Figure n°42 : La visualisation par Philcarto des Iris des communes du Sud du Bassin d'Arcachon. Les points rouges sont les centroïdes des IRIS ajoutés par Philcarto pour rendre possible la réalisation de cartes en cercles (centres des cercles).

Chapitre 6

Le recodage des identifiants d'un shapefile

Certains travaux nécessitent le recodage des identifiants présents dans un shapefile. On sait que les identifiants des éléments d'un shapefile figurent dans le fichier .dbf. Pour modifier les identifiants, il suffit de saisir directement les nouveaux identifiants en face (ou à la place) des anciens. C'est possible avec OpenOffice Calc, par exemple, mais pas avec Excel qui, curieusement, ne sait pas ouvrir les fichiers .dbf. C'est possible, mais risqué, fastidieux et pas systématique. C'est la raison de la présence du bouton **Recodage des identifiants** dans la fenêtre de fonctions d'**Eclats**.

L'exemple présenté ici est le suivant : dans la base cartographique « Admin-Express-cog » de l'IGN pour l'année 2019, un shapefile dénommé DEPARTEMENT_CARTO donne les contours de départements métropolitains de la France. Mais alors qu'il existe un shapefile pour les chefs-lieux des communes (un point par commune identifié par le code de chaque commune) dénommé CHEF_LIEU_CARTO, il n'existe pas de shapefile pour les chefs-lieux des départements. Autrement dit, les contours de départements sont orphelins de leurs chefs-lieux ! Eclats permet de pallier ce manque en deux opérations : tout d'abord la sélection des chefs-lieux des communes chefs-lieux de départements (ces points restent identifiés par leur code communes) ; puis le recodage des codes communes des chefs-lieux par les codes des départements.

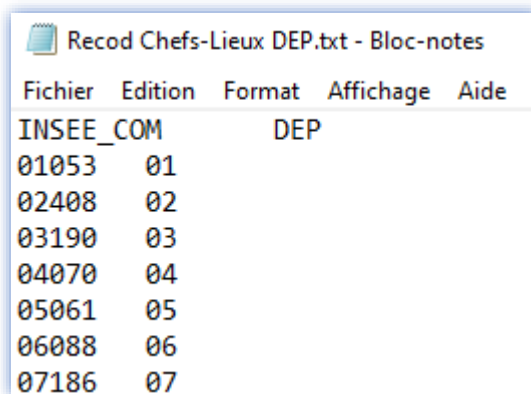
Pour réaliser ces opérations, on utilise les données du dossier Eclats_Test_Chefs-Lieux_Départs. La sélection des chefs-lieux des départements figure dans le dossier Selection du dossier Chefs-Lieux France. Le fichier .dbf du dossier Selection contient pour les 96 département les éléments suivants (fig. 43) :

	A	B	C	D
1	ID,C,24	NOM_CHF,C,60	STATUT,C,24	INSEE_COM,C,5
2	PAIHABIT0000000042284849	TARBES	PRéFECTURE	65440
3	PAIHABIT0000000080595713	FOIX	PRéFECTURE	09122
4	PAIHABIT0000000108056618	AURILLAC	PRéFECTURE	15014
5	PAIHABIT0000000088356464	AMIENS	PRéFECTURE	80021

Figure n°43 : la visualisation partielle dans OpenOffice du fichier CHEF_LIEU_CARTO_Selection. Noter la mauvaise prise en compte des caractères accentués (é au lieu de É).

Dans ce fichier .dbf, on observe que les identifiants des éléments cartographiques sont identifiés sur 5 caractères qui figurent dans la quatrième colonne. Or, si l'on souhaite utiliser ces identifiants avec les contours des départements, il faut remplacer les codes des communes chefs-lieux de départements par les codes des départements (ayant deux caractères seulement). Ainsi qu'il a déjà été dit, ce travail peut être fait avec OpenOffice ; on peut aussi demander à **Eclats** de le faire...

La première étape consiste à établir un fichier texte comprenant deux colonnes identifiées, les deux valeurs de chaque ligne étant séparées par un caractère de tabulation. La première est l'ancien identifiant (le code commune) ; la seconde est le nouvel identifiant (le code département). Le fichier Recod Chefs-Lieux DEP.txt donne les équivalents entre ces deux identifiants (fig. 44).



INSEE_COM	DEP
01053	01
02408	02
03190	03
04070	04
05061	05
06088	06
07186	07

Figure n°44 : la visualisation partielle dans le bloc note du fichier Recod Chefs-Lieux DEP.txt

Ce fichier une fois établi, le recodage est très simple à réaliser : lancer l'exécution d'**Eclats** ; choisir le bouton **Recodage des identifiants** ; puis ouvrir le shapefile à traiter, CHEF_LIEU_CARTO_Selection_.shp (fig. 45), et enfin ouvrir le fichier de recodage Recod Chefs-Lieux DEP.txt (fig. 46).

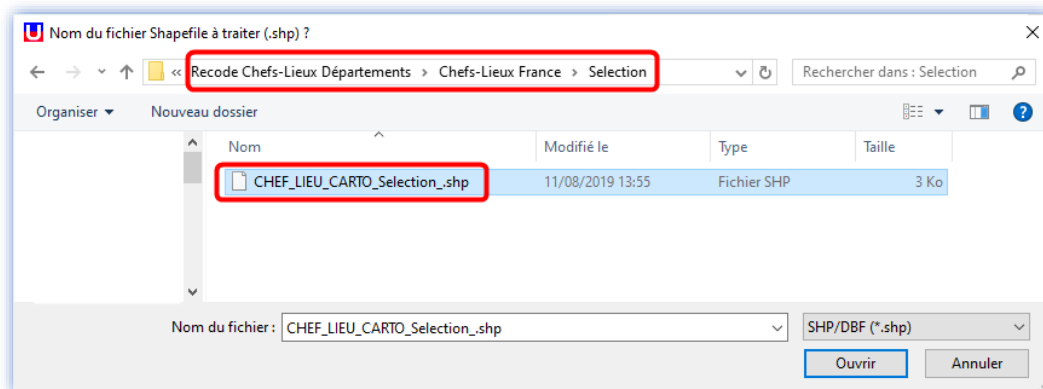


Figure n°45 : l'ouverture du shapefile à traiter CHEF_LIEU_CARTO_Selection_.shp.

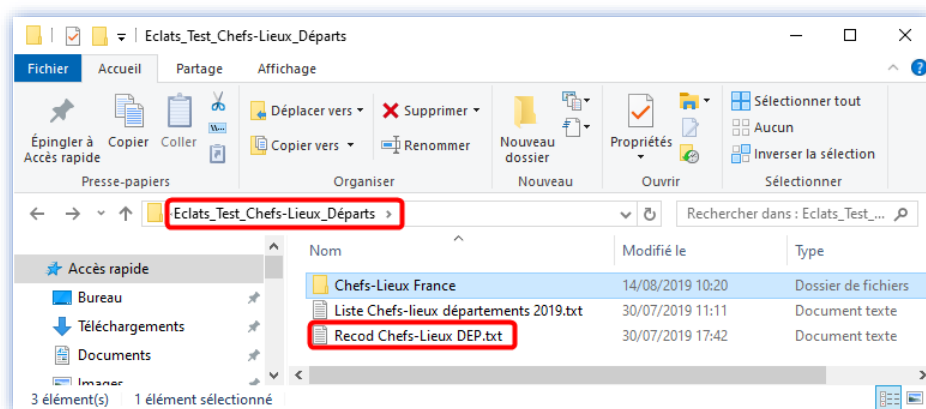


Figure n°46 : l'ouverture du fichier de recodage Recod Chefs-Lieux DEP.txt.

Après avoir ouvert le shapefile à recoder et le fichier de recodage, **Eclats** affiche une fenêtre avec deux listes (fig. 47) : à gauche les champs du fichier .dbf à recoder, à droite les deux champs du fichier de recodage. Ces deux fichiers ont un champ en commun INSEE_COM qui permet joindre les données des deux fichiers. Ici, il s'agit de recoder INSEE_COM avec les valeurs de DEP. La première est sélectionnée dans la liste de gauche et la seconde dans la liste de droite. Un clic sur le bouton **Recoder** réalise l'exécution du recodage.

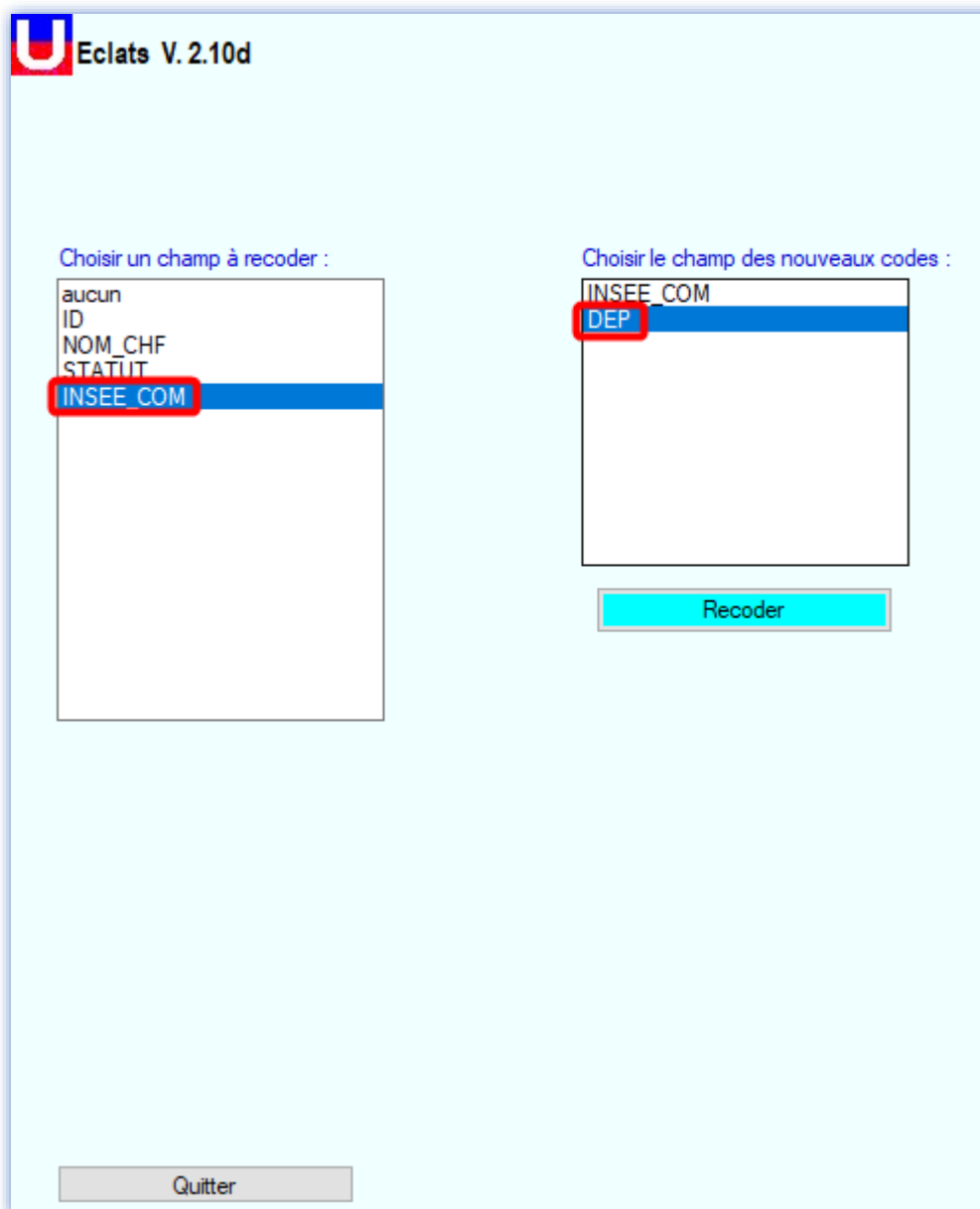


Figure n°47 : sélection des champs à recoder et de recodage.

Lorsque le recodage est terminé, un dossier est créé dans le dossier du shapefile à recoder. Ce dossier porte le nom Recodage (fig. 48).

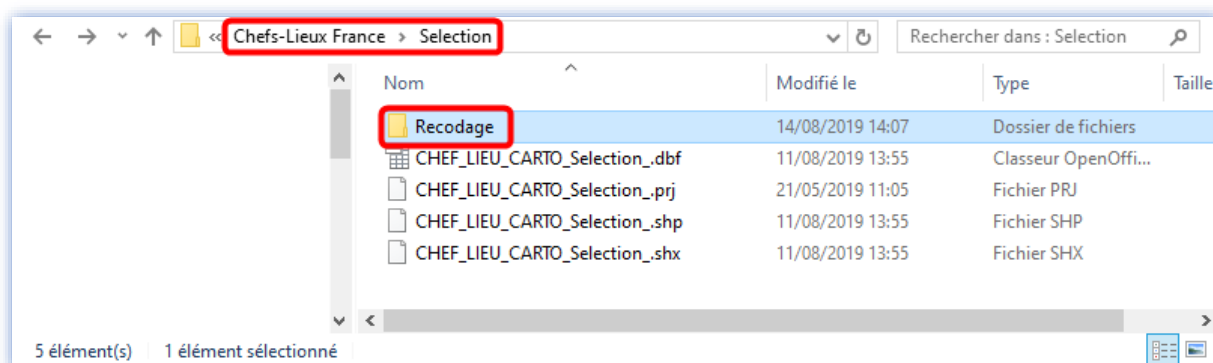


Figure n°48 : l'ouverture du fichier de recodage Recod Chefs-Lieux DEP.txt.

Dans le dossier Recodage, on trouve les mêmes noms de fichiers que dans le dossier Selection à une différence près : le suffixe _Recodage_ a été ajouté à ces noms de fichiers (fig. 49).

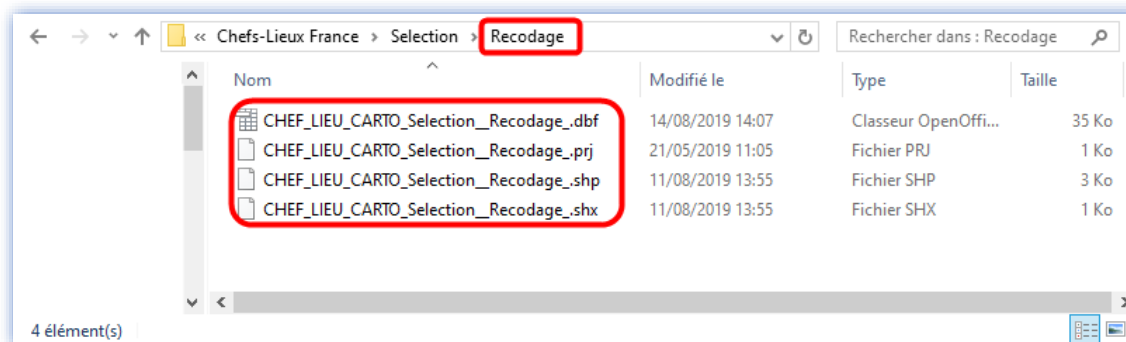


Figure n°49 : les fichiers créés par **Eclats** dans le dossier Recodage.

Parmi ces nouveaux fichiers, seul le fichier CHEF_LIEU_CARTO_Selection_Recodage_.dbf est différent du fichier d'origine CHEF_LIEU_CARTO_Selection.dbf. En effet, les autres fichiers sont inchangés (ce sont de simples copies) alors que CHEF_LIEU_CARTO_Selection_Recodage_.dbf contient le recodage proprement dit. Ouvert dans OpenOffice, on observe que le nouveau fichier .dbf contient un champ de plus dénommé DEP qui renferme précisément les codes des départements (fig. 50).

	A	B	C	D	E
1	ID,C,24	NOM_CHF,C,60	STATUT,C,24	INSEE_COM,C,5	DEP,C,254
2	PAIHABIT0000000042284849	TARBES	PréFECTURE	65440	65
3	PAIHABIT0000000080595713	FOIX	PréFECTURE	09122	09
4	PAIHABIT0000000108056618	AURILLAC	PréFECTURE	15014	15
5	PAIHABIT0000000088356464	AMIENS	PréFECTURE	80021	80
6	PAIHABIT0000000016748848	BOURGES	PréFECTURE	18033	18

Figure n°50 : la présence du champ DEP dans le fichier .dbf recodé.

Arrivé à ce stade, le recodage est terminé. Il suffit en effet d'utiliser le champ DEP en lieu et place du champ INSEE_COM dans l'importation du shapefile ou dans la création d'un fond de carte au format .AI (fig. 51).

Eclats V.2.10d Choisir le ou les champs d'identification des éléments dans le fichier .AI ...

Identifiant

ID
NOM_CHF
STATUT
INSEE_COM
DEP
SEQUENCE

→ ← DEP

Valider

Annuler

	ID	NOM_CHF	STATUT	INSEE_COM	DEP	Σ ^
	PAIHABIT0000000042284849	TARBES	PRÆFECTURE	65440	65	01
	PAIHABIT0000000080595713	FOIX	PRÆFECTURE	09122	09	02
	PAIHABIT00000000108056618	AURILLAC	PRÆFECTURE	15014	15	03
	PAIHABIT0000000088356464	AMIENS	PRÆFECTURE	80021	80	04
	PAIHABIT0000000016746848	BOURGES	PRÆFECTURE	18033	18	05

CHEF_LIEU_CARTO_SELECT Appliquer ☒ Appliquer les mêmes identifiants à tous les dossiers ☒ Ajouter un calque de centres (.P) à tout calque de surfaces (.S)

Figure n°51 : l'utilisation du champ DEP du fichier .dbf recodé pour la création d'un fond de carte au format .AI.

Chapitre 7

Renommer tous les fichiers d'un même shapefile

Tous les fichiers composant un shapefile peuvent être renommés sans difficulté en utilisant l'explorateur de Windows. Mais ce travail répétitif (il faut renommer au moins trois fichiers strictement de la même façon, en conservant néanmoins le type de fichier) est sujet à erreur et, dans tous les cas, fastidieux. **Eclats** permet de renommer tous les fichiers d'un shapefile en une seule fois ; le programme parcourt le dossier où se situe un fichier .shp sélectionné par l'utilisateur et change le nom de tous les fichiers portant le même nom (à l'exception du format) selon une chaîne de caractères donnée par l'utilisateur. Cette fonction est pratique et s'avère d'une grande efficacité pour renommer les fichiers de sélection produits par **Eclats**.

Pour renommer un shapefile, il faut d'abord choisir cette fonction par un clic sur le bouton **Renommer les fichiers** après avoir lancé l'exécution d'**Eclats**. Un dialogue permet ensuite de sélectionner un fichier .shp, par exemple le fichier CHEF_LIEU_CARTO_Selection_Recodage.shp enregistré au cours du chapitre précédent (fig. 52). Il s'agit des chefs-lieux des départements issus d'une sélection des chefs-lieux des communes qui sont aussi chefs-lieux de département. Le code de la commune commune y a été remplacé en code département. Le nouveau nom du shapefile à attribuer est CHEF_LIEU_DEP.

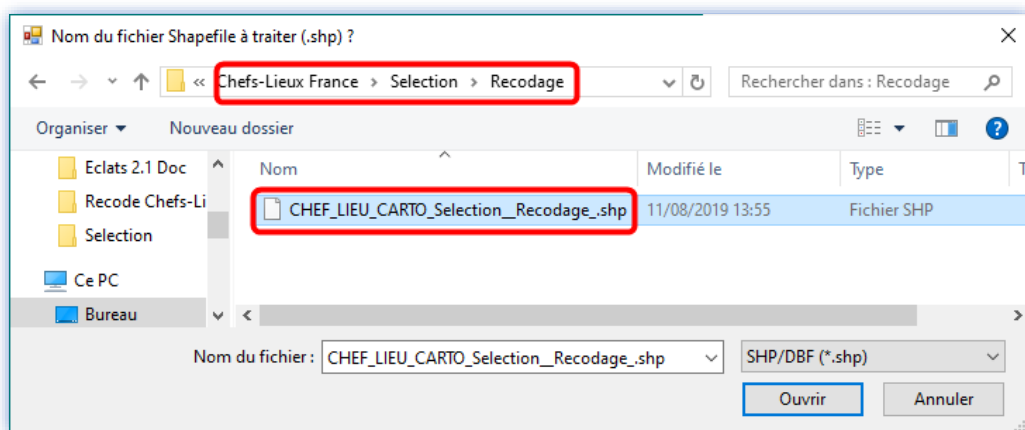


Figure n°52 : la sélection du fichier CHEF_LIEU_CARTO_Selection_Recodage.shp à renommer.

A la suite de l'ouverture du fichier à renommer s'affiche l'interface nécessaire à la réalisation de ce travail (fig. 53). Le nom du shapefile sélectionné (Ancien nom :) initialement sélectionné par l'utilisateur est indiqué mais n'est pas modifiable. Le nouveau nom (ici CHEF_LIEU_DEP) doit être entré par l'utilisateur. Selon les conventions imposées sur les noms de fichiers par Windows, les caractères \ / : * ? " < > | ne sont pas admis dans le nouveau nom ; le programme les ignore simplement. Lorsque le nouveau nom est saisi, un clic sur le bouton **Renommer** provoque les changements de noms.

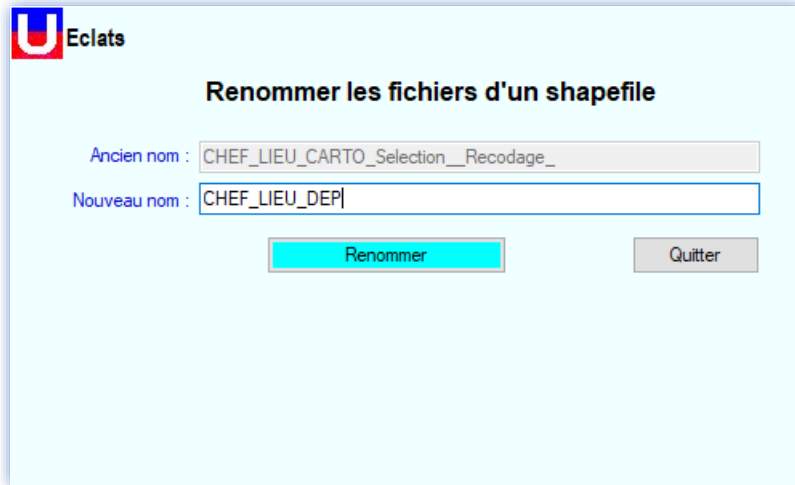


Figure n°53 : l'ancien nom doit être renommé avec le nouveau nom.

A la fin du recodage, la liste des fichiers renommés est donnée pour information (fig. 54). Ici, les quatre fichiers composant le shapefile ont été renommés.

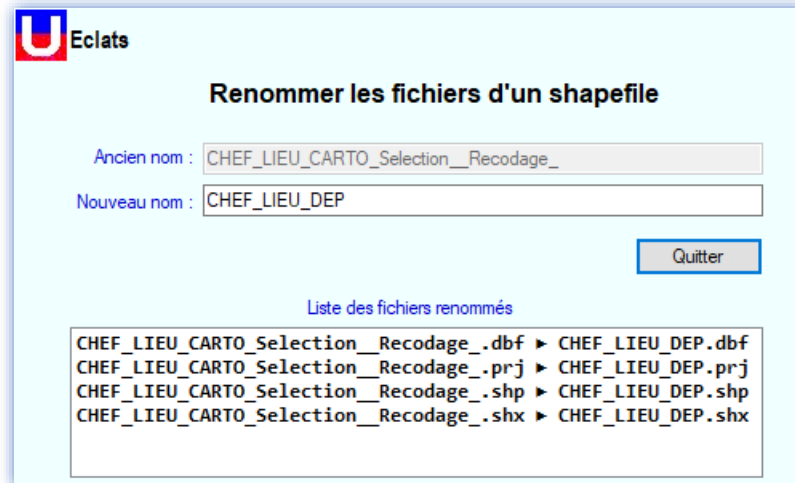


Figure n°54 : la liste des fichiers format le shapefile, avant le recodage et après le recodage.

Contrairement aux autres traitements proposés par **Eclats**, un clic sur le bouton **Quitter** ne fait pas quitter le programme, mais renvoie à la fenêtre qui propose tous les boutons de fonction du programme, ce qui permet, par exemple, de réitérer le recodage sur un autre shapefile.

Chapitre 8

La conversion des shapefiles en fichiers .AI

Les fonds de cartes utilisables par **Philcarto** sont au format shapefile ou .AI. Dans le cas des shapefiles, l'utilisateur doit former les identifiants des unités spatiales et éventuellement choisir une projection cartographique après chaque ouverture du shapefile par **Philcarto**, ce qui peut s'avérer fastidieux lors d'une utilisation fréquente du même fond de carte. Dans ce cas, il est préférable de recourir à un fichier .AI issu de la conversion du shapefile et qui contient le fond de carte dans sa forme définitive avec des éléments cartographiques déjà identifiés. Ceci peut être fait avec **Phildigit** dans le cas d'un fond de carte comprenant de nombreux calques et/ou nécessitant des opérations complémentaires (regroupements, calcul des centroïdes...). **Eclats** réalise ce genre de conversion dans les cas plus simples.

Le cas envisagé ici est celui d'un shapefile qu'on souhaite convertir en .AI pour l'utiliser comme fond de carte avec **Philcarto**. Il peut s'agir d'un shapefile provenant d'une banque de données comme « Admin-Express » de l'IGN, ou bien encore d'un shapefile issu d'une sélection par **Eclats** lors d'une utilisation précédente du programme. Le traitement est déclenché par un clic sur le bouton **Conversion shapefile -> .AI** de la fenêtre « Traitement à réaliser ? » Le programme demande alors de sélectionner le dossier contenant le shapefile à convertir (fig. 55).

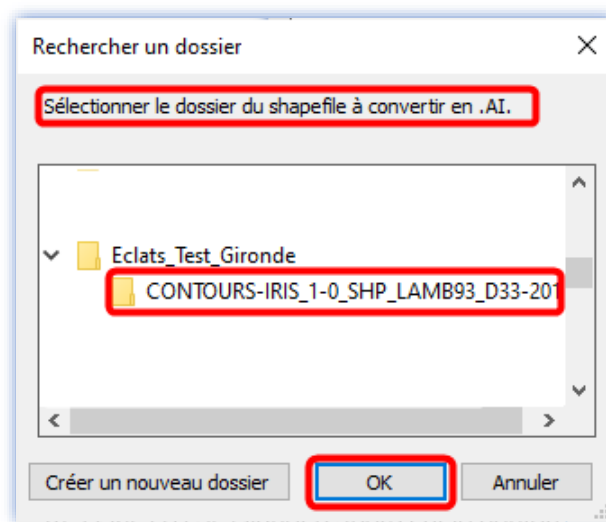


Figure n°55 : le dialogue de sélection du dossier contenant le shapefile à convertir en .AI.

Si le dossier sélectionné renferme plusieurs shapefiles, chacun d'eux constitue un calque du même fichier .AI.

Si le dossier sélectionné renferme un ou plusieurs autres dossiers renfermant chacun un ou plusieurs shapefiles ces dossiers sont successivement traités et les shapefiles sont convertis en fichier .AI comprenant un ou plusieurs calques, selon le nombre de shapefiles dans chaque dossier traité. Dans ce cas, les éventuels shapefiles du dossier initialement sélectionné ne sont pas traités.

A la suite de la sélection du dossier contenant le ou les shapefiles, l'utilisateur doit identifier les éléments cartographiques à enregistrer dans le fichier .AI. Une fenêtre s'affiche à cet effet (fig. 56). A gauche figure la liste des champs présents dans le fichier .dbf du shapefile en cours de traitement. L'utilisateur doit d'abord sélectionner le champ d'identification dans cette liste (ici DCOMIRIS qui comprend le numéro de département suivi du numéro de commune suivi du numéro d'iris dans chaque commune). Un clic sur la flèche jaune dirigée vers la droite indique au programme qu'il doit affecter les valeurs de DCOMIRIS à chaque élément cartographique du fichier .AI. Enfin, un clic sur le bouton **Valider** déclenche le traitement.

Identifiant

DEPCOM
NOM_COM
IRIS
DCOMIRIS
NOM_IRIS
TYP_IRIS

Valider
Annuler

DCOMIRIS

DEPCOM	NOM_COM	IRIS	DCOMIRIS	NOM_IRIS
33449	SAINT-MEDARD-EN-JALLES	0109	334490109	CORBIAC 2
33039	BEGLES	0108	330390108	ARGOUS
33039	BEGLES	0101	330390101	LA CASTAGNE-LA FERRADE
33039	BEGLES	0106	330390106	LE PRECHE
33281	MERIGNAC	1503	332811503	BEAUDESERT

BORDEAUX_METROPOLE_C Appliquer

☒ Appliquer les mêmes identifiants à tous les dossiers
☒ Ajouter un calque de centres (.P) à tout calque de surfaces (.S)

Figure n°56 : le dialogue d'identification des éléments cartographiques dans le fichier .AI.

Cette fenêtre d'identification des éléments cartographiques dans le fichier .AI possède en outre deux options sous forme de cases à cocher. La première de ces deux cases ☒ Appliquer les mêmes identifiants à tous les dossiers est utile s'il y a plusieurs dossier à traiter successivement avec le même schéma d'identification que celui du premier dossier traité. Dans ce cas la fenêtre d'identification est proposée une seule fois pour chacun des shapefiles du premier dossier et n'est pas réaffichée pour les dossiers suivants.

La seconde option concerne le contenu du fichier .AI. Dans la perspective d'une utilisation avec **Philcarto**, il est utile d'ajouter un calque contenant les centres des surfaces : un calque .P pour chaque calque .S. Cet ajout a lieu quand la case à cocher ☒ Ajouter un calque de centres (.P) à tout calque de surfaces (.S) est effectivement cochée. Cette opération est inutile lorsqu'il existe un shapefile de centres (de chefs-lieux par exemple) ; dans ce cas, il faut décocher la case.

A ce stade, deux cas de figure peuvent alors se présenter : soit il s'agit d'un fond de carte déjà projeté, soit le shapefile contient des coordonnées géographiques angulaires en latitude et longitude. Dans le premier cas, le traitement est directement réalisé : un fichier .AI est créé dans chaque dossier rencontré par **Eclats** contenant un shapefile.

Dans le second cas, celui où on a affaire à des coordonnées géographiques non-projetées, il faut indiquer au programme s'il faut appliquer une projection ou bien laisser les coordonnées géographiques telles quelles. Pour choisir entre la projection gnomonique ou pas de projection du tout, un dialogue (fig. 57) présente à l'utilisateur cette alternative. Par défaut, c'est la projection gnomonique qui est sélectionnée car elle donne le plus souvent de bons résultats sur le plan graphique. Les valeurs X et Y indiquent, pour information, les coordonnées du point de tangence avec le plan de projection. Lorsque l'utilisateur a fait son choix, un clic sur le bouton **Valider** lance la conversion du shapefile en .AI.

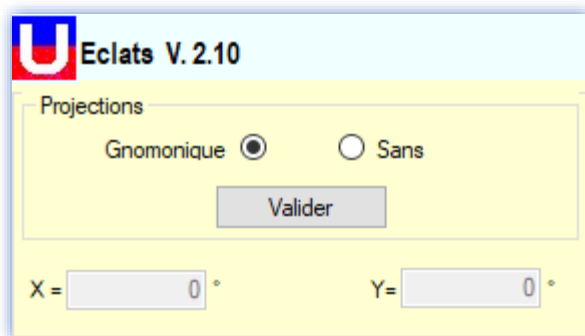


Figure n°57 : le dialogue pour projeter ou non le fond de carte.

Le résultat du traitement comprend deux fichiers (fig. 58) : tout d'abord un fichier .AI contenant le fond de carte à utiliser avec **Philcarto**. Le nom de ce fichier est identique au nom du shapefile, à l'exception du type .AI. A ce premier fichier s'ajoute un fichier texte formant un fichier des données pour tester le fond de carte (fig. 59). Sa première colonne contient les identifiants (ici DCOMIRIS) ; les autres colonnes (sauf les deux dernières) contiennent les données du fichier .dbf. Avec ce couple de fichiers, l'utilisateur peut immédiatement faire un essai du fond de carte au format .AI avec **Philcarto**.

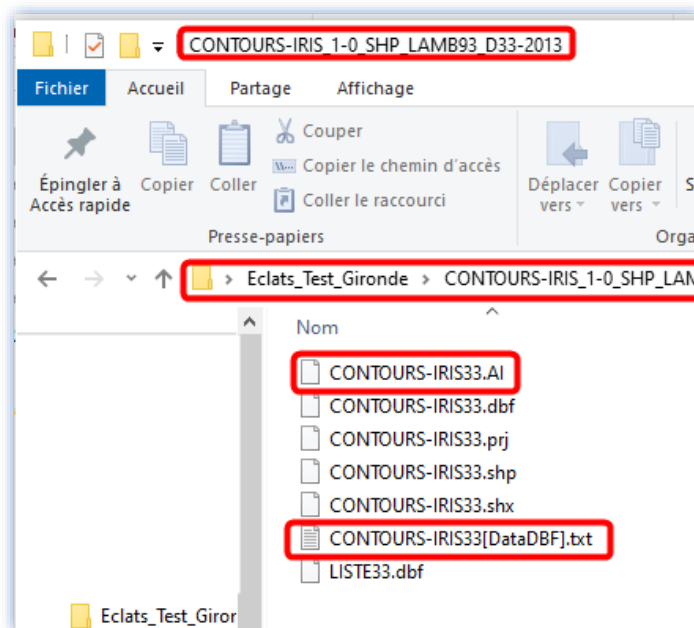


Figure n°58 : deux nouveaux fichiers sont créés : .AI et .txt.

CONTOURS-IRIS33[DataDBF].txt - Bloc-notes										
Fichier	Edition	Format	Affichage	Aide						
Iden	DEPCOM	NOM_COM	IRIS	DCOMIRIS	NOM_IRIS	TYP_IRIS	ORIGINE	SEQUENCE	PREFIXE/SEQUENCE	
332850000	33285	MOMBRIER	0000	332850000	MOMBRIER	Z	X	001	CONTOURS-IRIS33001	
331630000	33163	FALEYRAS	0000	331630000	FALEYRAS	Z	X	002	CONTOURS-IRIS33002	
331200000	33120	CERONS	0000	331200000	CERONS	Z	X	003	CONTOURS-IRIS33003	
335330000	33533	TOULENNE	0000	335330000	TOULENNE	Z	X	004	CONTOURS-IRIS33004	
333330000	33333	LE PORGE	0000	333330000	LE PORGE	Z	X	005	CONTOURS-IRIS33005	
335080000	33508	SAVIGNAC	0000	335080000	SAVIGNAC	Z	X	006	CONTOURS-IRIS33006	
333100000	33310	ORIGNE	0000	333100000	ORIGNE	Z	X	007	CONTOURS-IRIS33007	
330260000	33026	BALIZAC	0000	330260000	BALIZAC	Z	X	008	CONTOURS-IRIS33008	
334710000	33471	SAINT-SAUVEUR	0000	334710000	SAINT-SAUVEUR	Z	X	009	CONTOURS-IRIS33009	
334490109	33449	SAINT-MEDARD-EN-JALLES	0109	334490109	CORBIAC	2	H	X	010	CONTOURS-IRIS33010
330060000	33006	ANGLADE	0000	330060000	ANGLADE	Z	X	011	CONTOURS-IRIS33011	
331300000	33130	COIMERES	0000	331300000	COIMERES	Z	X	012	CONTOURS-IRIS33012	
330760000	33076	BUDOS	0000	330760000	BUDOS	Z	X	013	CONTOURS-IRIS33013	
332540000	33254	LOUPIAC-DE-LA-REOLE	0000	332540000	LOUPIAC-DE-LA-REOLE	Z	X	014	CONTOURS-IRIS33014	
334410000	33441	SAINT-MARTIN-LACAUSSADE	0000	334410000	SAINT-MARTIN-LACAUSSADE	Z	X	015	CONTOURS-IRIS33015	

Figure n°59 : un extrait du fichier .txt contenant des données tests provenant du fichier .dbf
(les décalages observables dans les colonnes sont dus aux caractères de tabulation).