

Université Ferhat Abbas, Sétif 1
Institut d'Architecture et des Sciences de la Terre
Département des Sciences de la Terre

SUPPORT DE COURS SGBD & SIG

Systèmes de gestion des bases de données &
Système d'Information Géographique

Dr. Bersi Mohand | mbersi@univ-setif.dz

Ce cours permettra dans un premier temps aux étudiants en 2^{ème} année Licence Géologie d'avoir des notions de base sur la création et la gestion des bases de données, ensuite l'affichage et le traitement de ces bases de données dans un système d'information géographique (ArcGis).

RÉFÉRENCES

Arctur, David, and Michael Zeiler. Designing Geodatabases: case studies in GIS data modeling. ESRI, Inc., 2004.
Dongbao, Zhao. "To Understand Further Geodatabase Spatial Data Model [J]." Computer & Digital Engineering 4 (2007): 022.
ENGREF, CEMAGREF ENGREF CEMAGREF. "Systèmes d'Information Géographique." (1996).
Jones, Chris B. Geographical information systems and computer cartography. Routledge, 2014.
<http://desktop.arcgis.com/fr/arcmap/10.3/guide-books/map-projections/projection-types.htm>
<https://www.forumactif.com/search/sig-pour-tous.forumactif.com>
http://www.emse.fr/tice/uved/SIG/Glossaire/co/SGBD_1.html

1.1 Notion de base de données

Une base de données est un grand ensemble de données structurées concernant un sujet (objets cibles), en science de la Terre ces objets cibles peuvent être des gisements, des forages d'eau, des ouvrages d'art ...etc. Un système de gestion de bases de données (SGBD) est un logiciel de haut niveau d'abstraction qui permet de manipuler ces informations (Oracle, Access...etc.). Les bases de données sont parfois classées selon la façon dont elles organisent les données. L'approche la plus courante est la base de données relationnelle, une base de type tabulaire dans laquelle les données sont définies pour qu'elles puissent être réorganisées et rendues accessibles à travers plusieurs méthodes.

Gisements	Type de minerai	Type d'encaissant	Année de découverte
G01	Fer	Carbonaté	1902
G02	Or	Quartzitique	1955
G03	Argent	basique	2001
G04	Platine	Doléritique	2016

Une base de données est constituée d'une entité (objet ou cible à étudier) dans cet exemple l'entité c'est les gisements, à qui on a attribué des attributs (type de minerai, type d'encaissant...etc.).

Entité : Une entité est un sujet, une notion en rapport avec le domaine d'activité pour lequel la base de données est utilisée, et concernant lequel des données sont enregistrées.

Attribut : Un attribut est une caractéristique d'une entité susceptible d'être enregistrée dans la base de données.

Enregistrement : Un enregistrement est une donnée composite qui comporte plusieurs champs dans chacun desquels est enregistrée une donnée.

1.2 Modèles de description des données

Le schéma ou modèle de données, est la description de l'organisation des données. Il se trouve à l'intérieur de la base de données, et renseigne sur les caractéristiques de chaque type de donnée et les relations entre les différentes données qui se trouvent dans la base de données. Il existe plusieurs types de modèles de données (relationnel, entité-association, objet, hiérarchique et réseau).

A | Modèle de données relationnel

Le modèle relationnel représente la base de données comme un ensemble de tables, sans préjuger de la façon dont les informations sont stockées dans la machine. **Le schéma conceptuel** consiste à présenter les entités avec leurs attributs en précisant la relation et **les cardinalités** (La cardinalité d'une relation est son nombre d'occurrences). Une clé primaire désigne l'attribut sur lequel la relation entre deux ou plusieurs bases de données va se faire. Les tables constituent donc la **structure logique** du modèle relationnel.

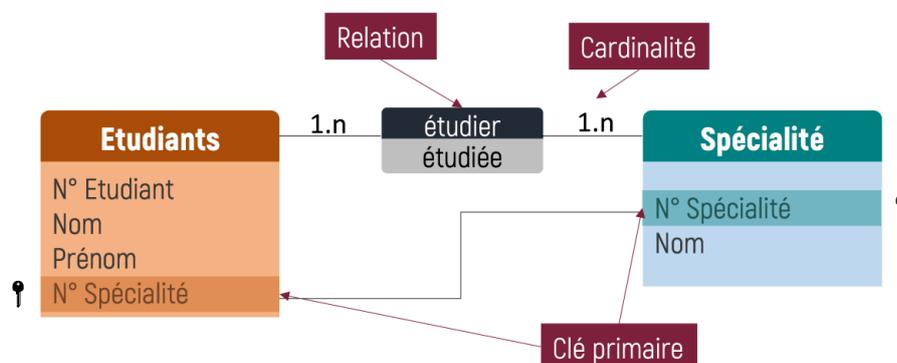


Schéma conceptuel de données pour deux bases de données.

Pour exécuter ce modèle et le rendre logique on utilise les requêtes. Le résultat sera une table de données avec commune entre les trois bases selon notre choix.

TD N°01 | Création d'une base de données relationnelle sous Access.

10 étudiants appartenant à 3 spécialités différentes, étudient chez 3 enseignants différents. Réalisez le schéma conceptuel puis exécutez le modèle logique en utilisant les requêtes.

Solution

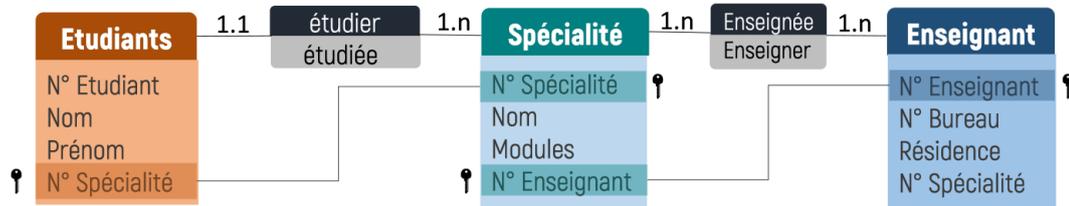


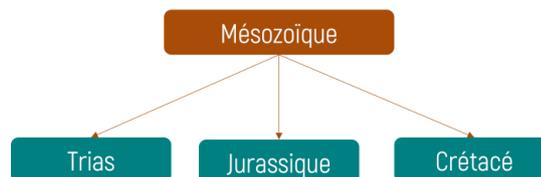
Schéma conceptuel de données pour trois bases de données.

N° Etudiant	Nom	Prénom	N° Spécialité	Nom	Nom Enseignant
20180002221	Badache	Mustapha	D05-001	RME	Pr. Bellouche
20180002345	Karich	Boualam	D05-002	Hydro	Dr. Bersi

Exécution du schéma conceptuel et obtention d'un modèle logique.

B | Modèle de données hiérarchique

Dans un modèle hiérarchique les données sont classées hiérarchiquement, selon une arborescence descendante. Ce modèle utilise des pointeurs entre les différents enregistrements. Il s'agit du premier modèle de SGBD.



Exemple de présentation de données selon un modèle hiérarchique.

C | Modèle de données orientées objets

Le modèle orienté objets utilise des objets graphiques avec des contenus programmables (scriptes) pour structurer la base de données et faire appel à l'information. Dans cet exemple on a utilisé un Combobox pour parcourir les âges du mésozoïque.



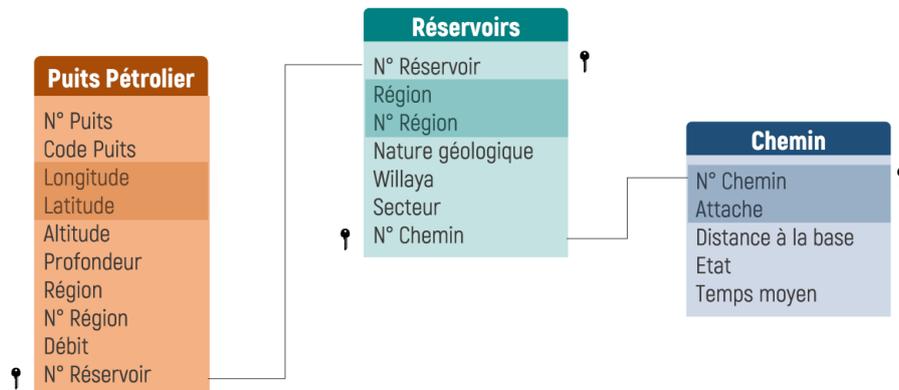
Modèle de données orientées objet.

1.3 Utilisation des bases de données

L'utilisation des bases de données est vaste, toutes les sociétés et les entreprises possèdent des bases de données concernant leurs employés, leurs équipements ...etc. les données stockées dans une bases peuvent être utilisées sous différents formes, les plus communs sont les formulaires. Si on prend l'exemple d'une base données qui décrit les minéraux, les formulaires sera sous forme d'une fiche technique qui décrit chaque minéral (Nom, caractéristiques en lumière naturelle et en lumière polarisé et analysée).

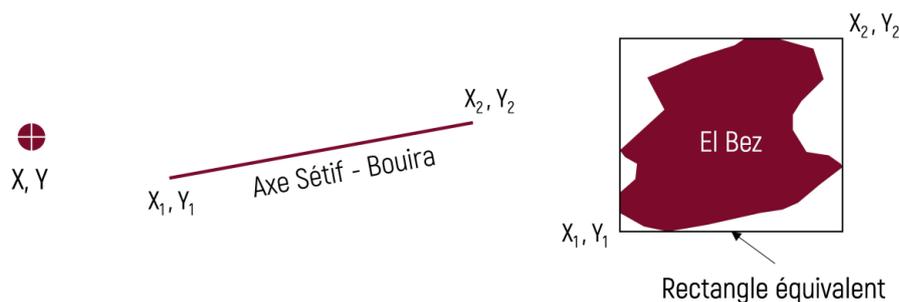
2.1 Bases de données géographiques

Une base de données géographique est une base de données décrivant des entités (classes) géographiquement localisées. Si on rajoute à une base de données normale des attributs comportant une information sur l'emplacement de ces entités elle deviendra donc une base de données géographique.



Exemple de trois bases de données géographiques.

L'information géographique dans une base de données sera dépendante du type de l'entité (objet cible). Ainsi, un objet ponctuel (puits pétrolier, gisement, source, forage d'eau, pont ...etc.) on lui attribue une information sous forme de coordonnées (x, y) , (longitude, latitude). Un objet linéaire on lui attribue une information qui contient le point de départ et d'arrivée [Sétif-Bouira, Est-Ouest, $x_1, y_1 - x_2, y_2$]. Enfin les objets surfaciques ils auront une information géographique par région ou bien les quatre points de son rectangle équivalent.



Type d'informations géographiques qu'on peut attribuer aux objets ponctuels, linéaires et surfaciques.

2.2 Système d'Information Géographique (SIG)

Système informatique permettant, à partir de diverses sources, de rassembler, organiser, gérer, analyser, combiner et de présenter des informations localisées géographiquement. Un SIG manipule deux types de données, il s'agit des données graphiques et des données alphanumériques. Il existe plusieurs systèmes informatiques qui manipulent les bases de données géographiques et prennent en compte la référence spatiale des données (ArcGis, QGIS, MapInfo, SuperGis, Surfer...etc.).

A | Données graphiques

Les données graphiques permettent de présenter les objets réels dans un SIG en utilisant trois formes de bases. Les objets ponctuels on les représente avec des points, les objets linéaires avec des lignes et les objets surfaciques avec des polygones. Ces types de données sont des **données vectorielles**.

Les points : Ils définissent des localisations d'éléments séparés pour des phénomènes géographiques trop petits pour être représentés par des lignes ou des surfaces (Mine, station météorologique, pont...etc.).

Les lignes : Les lignes représentent les formes des objets géographiques trop étroits pour être décrits par des surfaces (Failles, rivières, routes...etc.). Ces lignes sont souvent appelées dans les systèmes informatiques (ArcGis) **polylines**.

Les polygones : Ils représentent la forme et la localisation d'objets homogènes comme des pays, des couches géologiques, des types de sols, des plans d'eau, des forêts...etc.

Points



Polygones

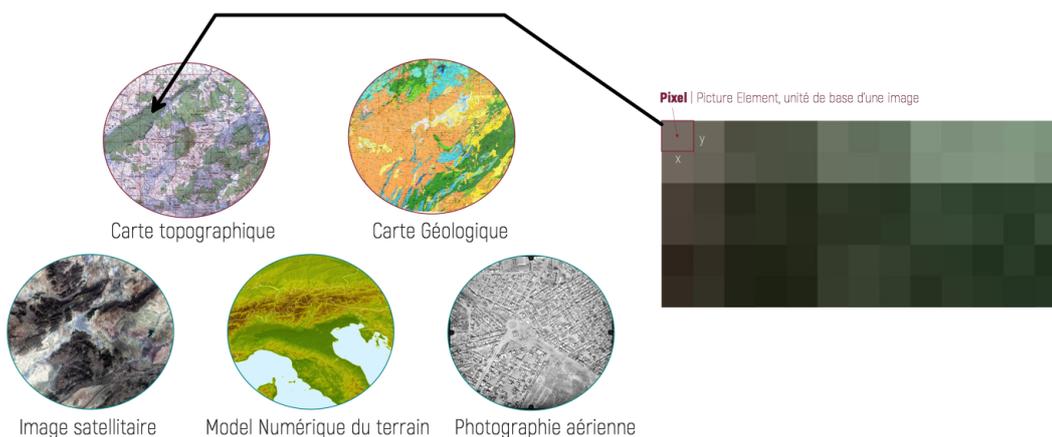


Polylines



Les trois types de données vectorielles qu'on utilise dans un SIG pour représenter les différentes entités.

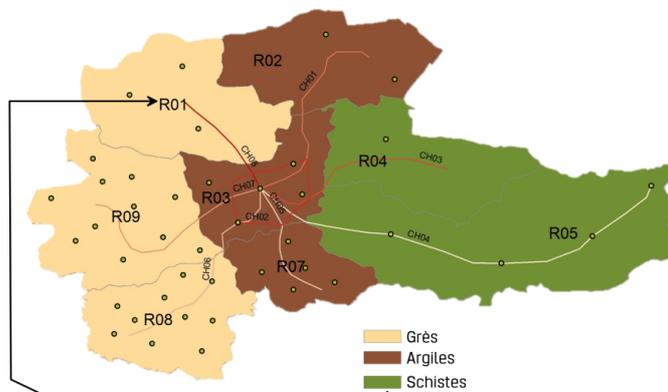
Les **données raster** permettent de représenter les objets réels selon une grille régulière et rectangulaire, organisée en lignes et en colonnes, chaque maille de cette grille (Pixel) ayant une intensité de gris ou une couleur (une couleur pour les cartes géologiques et topographique, un niveau de gris pour les images satellitaires et MNT).



Les différents types de données raster.

B| Données alphanumériques

Les données alphanumériques dans un SIG représentent les bases de données qu'on associe pour chaque entité vectorielle. Souvent on les appelle tables attributaires, elles contiennent toutes les informations sur l'entité sous forme d'attribut. Ces données peuvent être des chiffres, des textes, des dates... etc., d'où leur appellation Alphanumériques.



FID	Shape *	Id	N° Réser	Région	Géologie	Secteur	N° Chemin
0	Polygon	0	R01	Nord-Ouest	Grès	3	CH08
1	Polygon	0	R02	Centrale-Nord	Argiles	3	CH01
2	Polygon	0	R03	Centrale	Argiles	0	CH02
3	Polygon	0	R04	Est	Schistes	1	CH03
4	Polygon	0	R05	Sud-Est	Schistes	1	CH04
5	Polygon	0	R07	Centrale-Sud	Argiles	2	CH05
6	Polygon	0	R08	Sud-Ouest	Grès	3	CH06
7	Polygon	0	R09	Ouest	Grès	2	CH07

Propres au logiciel

Introduits par l'utilisateur

Exemple d'une table attributaire qui décrit les réservoirs pétroliers

Les systèmes de coordonnées permettent de localiser les cibles suivant des références, chaque système à ces propres origines. Parmi les systèmes de coordonnées les plus utilisés au monde « universels » on peut citer le système de coordonnées géographiques « WGS 84 » et le système Univers Transverse Mercator « UTM ». On Algérie en plus de ces deux systèmes WGS 84 et UTM on utilise souvent le système Lambert Conique Conforme « Lambert Conformal Conic ».

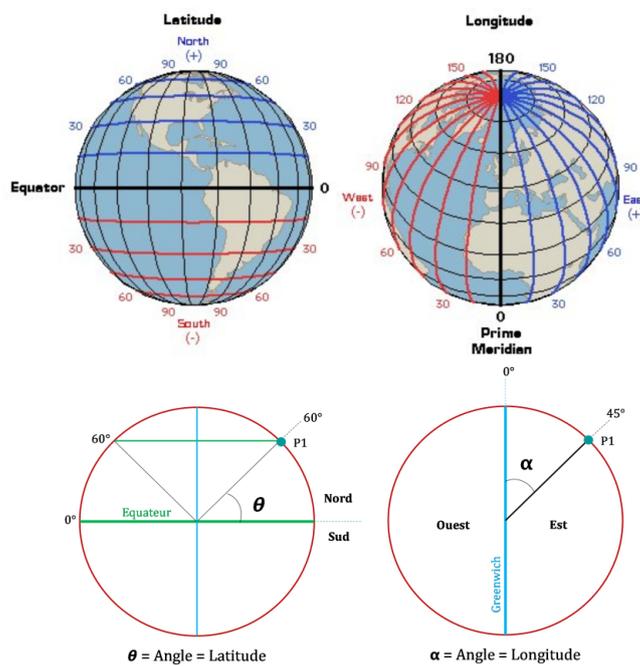
3.1 Les coordonnées géographiques

Les coordonnées géographiques ont comme origine le point d'intersection entre le méridien d'origine « Greenwich » et la parallèle d'origine « équateur ». Ce point se situe dans le golfe de Guinée aux larges de l'Atlantique.

La sphère de la Terre a été divisée en 360 méridiens (Chaque 1°) avec un méridien d'origine (Greenwich) et 180 parallèles, avec une parallèle d'origine (Equateur).

A | La latitude est une valeur angulaire, expression du positionnement nord ou sud d'un point sur Terre. D'un point de vue mathématique, la latitude d'un point est l'angle au centre que forme la normale (verticale) en ce point avec le plan équatorial. La latitude est une mesure angulaire s'étendant de 0° à l'équateur à 90° aux pôles.

B | La longitude est une valeur angulaire, expression du positionnement est ou ouest d'un point sur Terre. En géodésie, c'est l'angle au centre que forme le plan passant par ce point et par l'axe de rotation de la terre avec le plan du méridien de Greenwich.



Les coordonnées géographiques.

Tous les lieux situés à la même longitude forment un demi-plan limité par l'axe des pôles géographiques, coupant la surface de la terre sur un demi-cercle approximatif dont le centre est le centre de la Terre, l'arc allant d'un pôle à l'autre. Un tel demi-cercle est appelé méridien.

À la différence de la latitude (position nord ou sud) qui bénéficie de l'équateur et des pôles comme références, aucune référence naturelle n'existe pour la longitude.

La longitude est donc une mesure angulaire sur 360° par rapport à un méridien de référence, avec une étendue de -180° à +180°, ou respectivement de 180° ouest à 180° est.

Le méridien usuel de référence est le méridien de Greenwich (qui sert aussi de référence pour les fuseaux horaires).

La référence d'un point seront données comme suite :

Université Setif1, El Bez, Longitude : 5° 22' 3" E, Latitude : 36° 11' 52" N

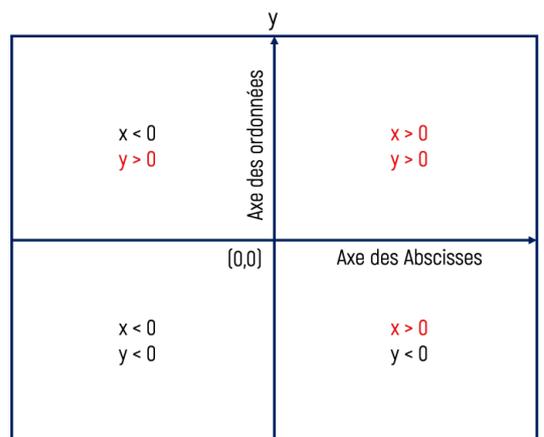
3.2 Les coordonnées projetées

Un système de coordonnées projetées se définit sur une surface plane, à deux dimensions. Contrairement à un système de coordonnées géographiques, un système de coordonnées projetées possède des longueurs, des angles et des surfaces constants dans les deux dimensions. Un système de coordonnées projetées est toujours basé sur un système de coordonnées géographiques, lui-même basé sur une sphère ou un ellipsoïde.

Dans un système de coordonnées projetées, des emplacements sont identifiés par des coordonnées x, y sur une grille, dont l'origine est située au centre de cette grille. Chaque position possède deux valeurs qui la situent par rapport à cet emplacement central. L'une précise sa position horizontale et l'autre, sa position verticale. Ces deux valeurs sont appelées la coordonnée x et la coordonnée y . Avec cette notation, les coordonnées à l'origine sont $x = 0$ et $y = 0$.

Sur une grille composée de lignes horizontales et verticales également espacées, la ligne horizontale au centre est appelée l'axe des x et la ligne verticale au centre est appelée l'axe des y . Les unités sont constantes et également espacées sur toute la plage des x et des y . Les lignes horizontales au-dessus de l'origine et les lignes verticales à droite de l'origine ont des valeurs positives; les lignes situées au-dessous ou à la gauche de l'origine ont des valeurs négatives. Les quatre quadrants représentent les quatre combinaisons possibles de coordonnées des X et des Y positives et négatives.

Parmi les systèmes de coordonnées projetées les plus utilisés en Algérie, on peut citer le système Lambert Conique Conforme et le système Universal Transverse Mercator (UTM). Comme ces deux systèmes sont kilométriques, et si on a une carte d'Algérie sans indication de type de projection on peut le déduire suivant le nombre de chiffre des coordonnées de l'axe des ordonnées « y ». Ainsi, si y possède 3 chiffres « en kilomètre » donc c'est le système Lambert et si y possède 4 chiffres donc c'est le système UTM.



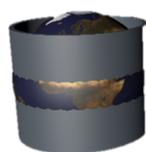
Origine des coordonnées projetées.

A | La Projection UTM (Projection cylindrique) :

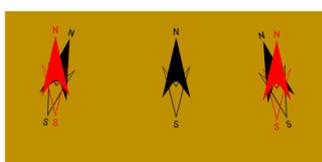
Le système UTM (Universal Transverse Mercator) est une application spécialisée de la projection de Mercator transverse. Le globe est divisé en soixante zones nord et sud, chacune d'elles s'étendant sur 6° de longitude. Chaque zone a son propre méridien central. Les zones 1N et 1S débutent à 180° Ouest. Les limites de chaque zone sont 84° N et 80° S, la division entre les zones nord et sud ayant lieu au niveau de l'équateur. Les régions polaires utilisent le système de coordonnées UPS (Universal Polar Stereographic).

B | Projection Lambert (Projection Conque)

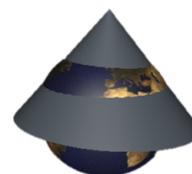
On projette l'ellipsoïde sur un cône tangent à un cercle ou sécant en deux cercles. Puis on déroule le cône pour obtenir la carte. Elle est souvent utilisée pour les zones placées au centre des deux hémisphères (Algérie, France, Belgique...).



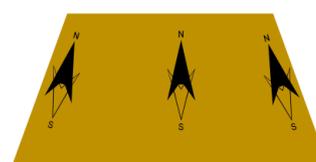
UTM



- Conservation des distances
- Fausse les orientations



Lambert



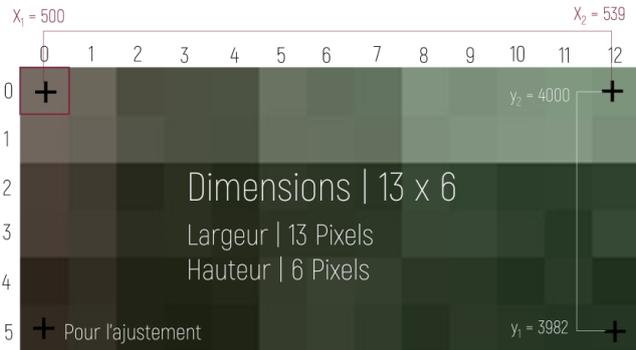
- Distorsion des distances
- Conservation des orientations

Comparaison entre les coordonnées UTM et Lambert

Dans un SIG on peut réaliser plusieurs manipulations sur les données, selon le type de données et l'objectif de l'utilisateur on peut citer ces manipulations les plus utilisées.

4.1 Le Calage (Géoréférencement)

Le géoréférencement est la technique de positionnement spatial d'une entité dans un cadre unique et une situation géographique bien définie dans un système de coordonnées et de références spécifiques. Il consiste en une opération courante dans le système d'information géographique (SIG) pour les deux objets raster (image bitmap pixels) à des objets vectoriels (points, lignes, polygones et polygones qui représentent des objets physiques).



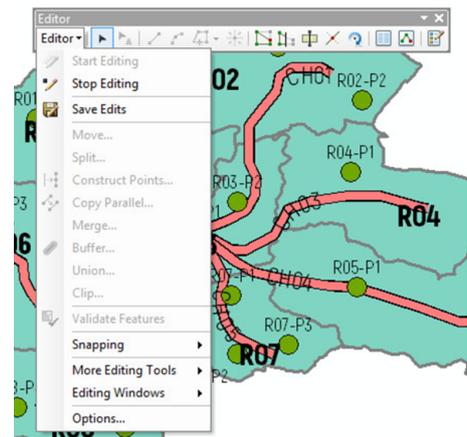
Le calage nécessite 4 points de contrôle répartis sur les 4 corners de la carte.

Pour géoréférencer une entité vectorielle il suffit de désigner le type de coordonnées pour les fichiers vecteurs (points, polygones et polygones).le type de fichiers vectorielles sous ArcGis est le type Shapefile (fichier de forme). Pour géoréférencer une carte raster il faut placer au moins quatre points (points de contrôle). Ces points on les place sur les quatre corners de l'image (carte). Pour chaque point on introduit les coordonnées suivant le système de projection avec lequel la carte a été réalisée.

4.2 La digitalisation

La digitalisation (numérisation) consiste à créer des fichiers vecteurs pour les entités présentes dans la carte raster (ce qu'on a fait dans le TD). Si la carte possède des points on va créer un fichier shapefile avec le type points et on positionne directement les points sur ceux de la carte. S'il s'agit de lignes on utilise un shapefile de type polygones pour tracer. Dans le cas de surfaces on utilise un fichier shapefile de type polygones pour tracer le périmètre de l'entité.

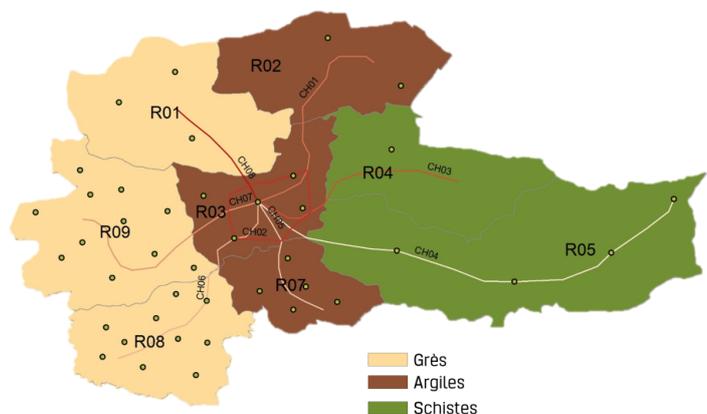
Pour commencer il faut aller sur Editor/ Start Editing
Pour terminer Editor/ Save Edits puis Editor/ Stop Editing



outil de digitalisation "Editor" sur ArcGis

4.3 L'analyse spatiale

Une fois on a créé nos fichiers shapefile on peut effectuer une analyse spatiale sur les données, cela est possible avec l'outil **Symbology** de ArcGis. Cet outil nous permet de faire une analyse rapide sur une des attributs de l'entité, dans cet exemple on a affiché les réservoirs suivant le type de lithologie, cela nous a permet de comprendre la répartition inégale des puits dans les différents réservoirs. Ainsi, on a constaté que les réservoirs de nature gréseuse grâce à leur forte porosité et perméabilité emmagasinent plus d'hydrocarbures donc il nous faut plus de puits pour l'exploitation.



Exemple d'utilisation de l'analyse spatiale.