

Programmation statistique - QGIS

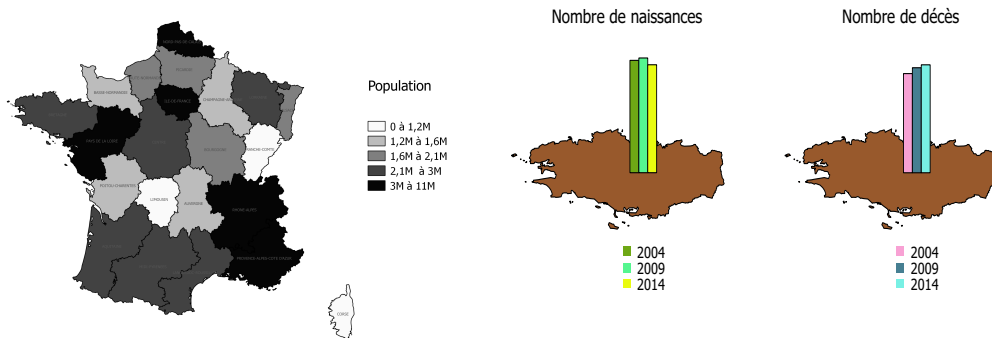
TD d'introduction

STID1 - 2021/22

A. Cuzol



www.qgis.org



Exercice - 1 Introduction - Informations à lire avant d'ouvrir QGIS

QGIS est un logiciel libre (gratuit). Il peut être téléchargé sur le site www.qgis.org. C'est un logiciel SIG ("Système d'information géographique") qui permet de créer, visualiser et analyser des informations géographiques. Un autre exemple de logiciel SIG est le logiciel payant MapInfo.

Les données d'un SIG sont regroupées en différents types :

- Données géographiques : relatives à la forme et la localisation des objets et phénomènes. Ces données géographiques peuvent être de deux types :
 - données "vectorielles" : points, lignes, polygones, etc (exemple : ensemble de villes) ;
 - données "rasters" (images) : image d'un plan ou photo numérisée.Ces données sont organisées dans le logiciel sous forme de couches superposées.
- Données descriptives (attributaires) : ensemble des attributs descriptifs des objets et phénomènes (exemple : population dans les villes observées).
- Données de style : paramètres d'affichage des objets (type du trait, couleur, etc).
- Métadonnées : données sur les données (date d'acquisition, nom du propriétaire, méthode d'acquisition, etc).

Quels fichiers peut-on ouvrir avec QGIS ? On peut ouvrir :

- la plupart des formats vecteurs (fichiers "shapefile" avec extension .shp, fichiers .tab de Mapinfo)
- les formats rasters : modèles numériques de terrain, photos aériennes/satellites, fonds cartographiques scannés de l'IGN, etc
- les données en ligne accessibles sur des serveurs géographiques (WMS : "Web Map Service", WFS : "Web Feature Service").

Le logiciel QGIS contient plusieurs éléments :

- La fenêtre "**Couches**" : liste et décrit les couches présentes dans la carte (données vectorielles et données rasters).
- La fenêtre "**Carte**" : affiche les différentes couches. Cette fenêtre est liée à la fenêtre "Couches" : si on modifie les couches ou leur style, la carte sera modifiée.
- La fenêtre "**Données attributaires**" : cette fenêtre est associée à une couche vectorielle. Cela permet de visualiser les données de la couche sous forme d'un tableau.

Notion de projet : QGIS sauvegarde le travail en cours dans un projet (extension .qgs). Ce projet enregistre :

- les couches ajoutées ;
- leurs propriétés ;
- la carte associée ;
- le composeur d'impression (mise en page).

Avant de faire les exercices, créer un répertoire "TD_QGIS" sur votre ordinateur et y copier les données du répertoire "Carte et données pour TD" disponible dans l'espace moodle "STID1 / Programmation Statistique - QGIS".

Attention : il faut bien copier tous les fichiers dans votre répertoire de travail. Par exemple, quand vous ouvrirez un fichier de type "shapefile" nommé "Carte.shp" dans QGIS, votre répertoire devra contenir tous les fichiers suivants qui stockent diverses informations utiles :

- Carte.shp
- Carte.cpg
- Carte.dbf
- Carte.qpj
- Carte.shx

Exercice - 2 Première ouverture de QGIS et création d'une première carte

1. Observer tous les menus et les différents boutons (en passant la souris sur un bouton, on obtient des informations sur ce bouton).
2. Où se trouve la fenêtre affichant la carte ?
3. Où se trouve la fenêtre décrivant les différentes couches ?
4. Comment peut-on ajouter une couche vectorielle ou une couche raster ?
5. Comment peut-on ouvrir la table des données attributaires ?
6. Ouvrir la couche vectorielle "Morbihan_Departements.shp" représentant le Morbihan (située dans le répertoire /Morbihan).
7. Tester des manipulations sur la carte : déplacement, zoom, etc. Utiliser la commande "Zoom sur l'emprise" (menu ou bouton) pour revenir à la configuration initiale.
8. La couche apparaît-elle bien dans la fenêtre "Couches" ?
9. Ouvrir la table d'attributs associée. Combien de variables contient cette table ?
10. Modifier le style de la couche : double-cliquer sur le nom de la couche et aller dans le menu "Symbologie".
11. Sauvegarder votre "projet" avec l'extension ".qgs" (par exemple "Exercice1.qgs").
12. Fermer le logiciel, le relancer et vérifier que vous retrouvez bien l'ensemble de votre travail.

Exercice - 3 Création d'une carte contenant plusieurs couches

1. Repartir du projet précédent ;
2. Ajouter la couche vectorielle des routes du Morbihan : "Morbihan_RoutierGeneral.shp" ;
3. Ajouter la couche vectorielle des villes du Morbihan : "Morbihan_Villes.shp".
4. Tester le masquage d'une des couches en décochant la case devant le nom de cette couche.
5. Modifier l'ordre des couches (en cliquant-déplaçant le nom de la couche). Que se passe-t-il lorsque la couche "Morbihan_Departements" est placée au sommet de la liste ?
6. Remettre les couches dans le bon ordre et modifier leur style (en double-cliquant sur chaque couche) afin de créer une carte ayant un rendu visuel qui vous satisfait :
 - épaisseur/couleur des traits ;
 - taille/couleur des points ;
 - couleur du fond de carte ;
 - affichage/taille/police/couleur des étiquettes (dans l'onglet correspondant) ;
 - etc.

Exercice - 4 Outil de mise en page

L'outil de mise en page (dans *Projet/Nouvelle mise en page*) permet de mettre en forme votre carte avant de la sauvegarder ou l'insérer dans un rapport.

1. Ouvrir la carte de l'exercice précédent.
2. Ouvrir une nouvelle mise en page.
3. Afficher la carte dans la nouvelle mise en page ("*Ajouter une carte*").
4. Ajouter un titre (en utilisant "*Ajouter une étiquette*". Choisir la police Calibri pour ce titre.
5. Sauver cette carte au format PDF dans votre répertoire personnel.
6. Ouvrir le PDF pour vérifier que l'exportation a bien fonctionné.

Quelques informations à lire avant de continuer les exercices.

Pour l'instant, nous n'avons représenté que des éléments géographiques (délimitations de départements, routes, etc). Dans la suite, nous allons chercher à utiliser ou représenter d'autres données disponibles (la population par exemple). Ces **données attributaires** peuvent contenir plusieurs variables de types différents :

- Si la variable est qualitative on utilisera les "symboles catégorisés" : par exemple plusieurs couleurs pour représenter les différents départements d'une région.
- Si la variable est quantitative on utilisera les "symboles gradués" : par exemple plusieurs couleurs pour différentes plages de valeurs de population.

Remarque : pour une variable quantitative on peut aussi utiliser des "symboles proportionnels" (par exemple avec une taille qui varie selon les valeurs de la variable). Exemple : taille d'un point qui varie pour représenter la population dans différentes villes. Cependant ces symboles proportionnels ont des défauts et sont donc à manier avec précaution :

- la proportionnalité des symboles n'est pas affichée dans la légende ;
- faire varier les symboles peut compliquer la lecture de la carte.

Exercice - 5 Données attributaires

1. Observer les données attributaires associées à la couche "Morbihan_Villes".
2. Combien d'individus contient cette table? Combien de variables? De quels types sont ces variables?
3. Sélectionner les lignes décrivant Vannes et Lorient. Que s'est-il passé sur la carte?
4. Inversement, sélectionner des villes sur la carte (utiliser le bouton proposant plusieurs types de sélections possibles) et observer l'effet sur le tableau.

Exercice - 6 Sélection selon les données attributaires

1. Dans un nouveau projet, ouvrir la couche vectorielle "France_departements.shp" (située dans le répertoire /France).
2. Ouvrir la table des données attributaires. Afficher uniquement les départements bretons dans cette table.
3. Afficher uniquement les départements ayant une population supérieure à 1 million d'habitants.
4. Afficher uniquement les départements ayant une densité de population supérieure à 100 habitant par km².
5. Représenter en bleu les départements sélectionnés à la question 4 sur la carte, en blanc les autres départements. Pour cela :
 - (a) Sauver la sélection comme une nouvelle couche que vous nommerez par exemple "Dep_densite.shp" :
 - aller dans /Couche/Sauvegarder sous;
 - cocher "N'enregistrer que les entités sélectionnées";
 - aller dans "Parcourir".
 - (b) Superposer les couches "France_departements.shp" et "Dep_densite.shp", modifier leurs couleurs.

Programmation statistique - QGIS

TD 2

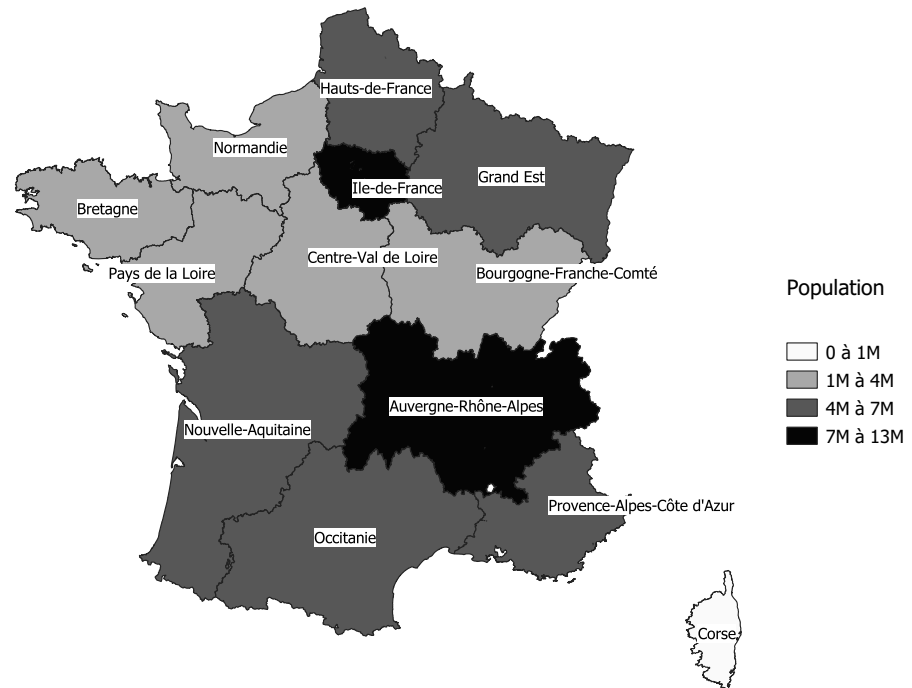
STID1 - 2021/22

A. Cuzol

Exercice - 1 Représentation de données attributaires

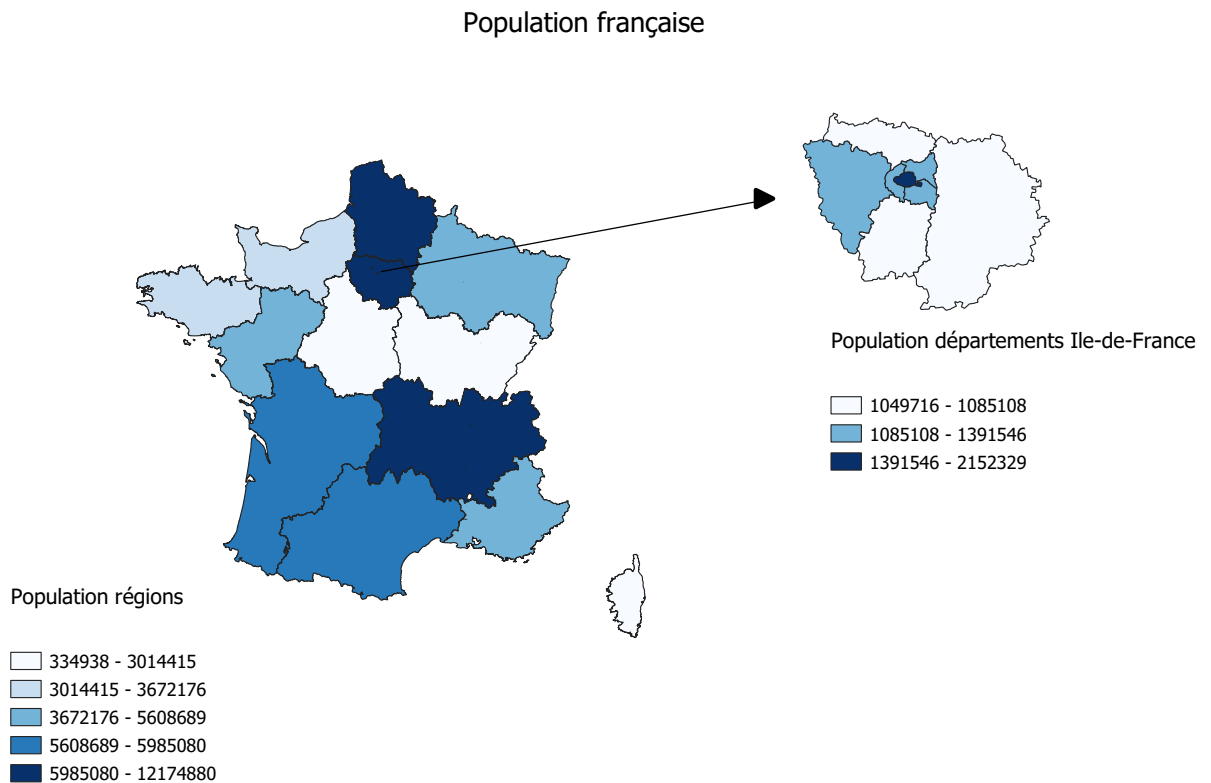
1. Ouvrir la couche vectorielle "France_Regions.shp".
2. Représenter les régions classées par noms. Pour cela :
 - aller dans *Propriétés de la couche/Symbologie/Symbole Catégorisé*
 - sélectionner la variable "Noms" dans *Valeurs*
 - choisir la palette "Couleurs au hasard"
 - cliquer sur *Classer*
3. Afficher la légende en cliquant sur le petit triangle en face du nom de la couche. Ce choix de représentation des noms de régions vous semble-t-il pertinent ?
4. Proposer une autre façon de représenter l'information "Nom de la région". Réaliser la carte correspondante.
5. Repartir de la carte vierge. Représenter les régions par classes de valeurs de la population (aller dans *Propriétés de la couche/Symbologie/Symbole Gradué*). Fixer les 5 classes suivantes :
 - 0 à 1 000 000
 - 1 000 000 à 4 000 000
 - 4 000 000 à 7 000 000
 - 7 000 000 à 13 000 000et les étiquettes de légendes suivantes :
 - 0 à 1M
 - 1M à 4M
 - 4M à 7M
 - 7M à 13M
6. Choisir une palette de couleurs en niveaux de gris (de blanc à noir).
7. Ajouter les étiquettes des noms de régions et faire en sorte que les noms soient tous lisibles : dans le menu *Étiquettes/Rendu*, cliquer sur "Afficher toutes les étiquettes pour cette couche". Redimensionner les étiquettes si c'est nécessaire. Ajouter un arrière-plan de couleur blanche.
8. Ouvrir une nouvelle mise en page et reproduire la mise en page ci-dessous (carte + légende). Sauvegarder la carte et sa légende au format .pdf. Remarque : pour modifier le titre et les éléments de la légende, décocher "Mise à jour auto" dans *Propriétés de l'objet / Éléments de la légende*.

Population des régions de France



Exercice - 2 Reproduction d'une carte

Le but de l'exercice est de reproduire la carte suivante au format pdf :



Démarche :

1. Afficher la carte de la population des régions françaises : choisir 5 classes et des effectifs égaux dans chaque classe.
2. Afficher la carte de la population des départements d'Ile-de-France en suivant les étapes suivantes :
 - (a) Ouvrir la couche des départements français.
 - (b) Sélectionner les départements d'Ile-de-France.
 - (c) Sauver cette sélection comme une nouvelle couche que vous nommerez "IF_Départements.shp".
 - (d) Représenter la population en Ile-de-France en choisissant 3 classes et le mode "ruptures naturelles".
3. Mise en page de la carte finale :
 - (a) Cocher uniquement la couche des régions dans la fenêtre des couches.
 - (b) Ajouter cette carte dans une nouvelle mise en page.
 - (c) Important : dans *Propriétés de l'objet*, cliquer sur "Verrouiller les couches".
 - (d) Ajouter la légende.
 - (e) N'afficher que la partie de légende correspondant aux régions en cliquant sur "Ne montrer que les entités à l'intérieur de la carte liée" dans *Propriétés de l'objet*.
 - (f) Décocher "Mise à jour auto" pour cette légende et supprimer le texte "France_regions".
 - (g) Modifier le titre de la légende.

- (h) Revenir dans QGIS et cocher uniquement la couche des départements d'Ile-de-France. Faire un zoom sur l'emprise totale (bouton loupe avec 3 flèches).
- (i) Dans la fenêtre de mise en page, suivre la même démarche que précédemment pour cette carte et sa légende.
- (j) Ajouter la flèche.
- (k) Ajouter le titre.
- (l) Exporter au format pdf. **Vérifier le résultat obtenu en ouvrant le pdf à partir de votre répertoire de travail.**

Exercice - 3 Sauvegarde de style

1. Afficher la carte des pays du monde.
2. Représenter la population des pays en 2005 (choisir 5 classes, ruptures naturelles).
3. Enregistrer le style de cette couche dans un fichier .qml (dans *Symbologie/Style* tout en bas).
4. Sélectionner les pays ayant une latitude positive et une longitude comprise entre -40 et 40.
5. Sauvegarder cette sélection dans une couche.
6. Masquer la couche des pays du monde et afficher la couche des pays sélectionnés.
7. Appliquer à cette couche le style sauvé précédemment en allant dans *Symbologie/Style/Charger le style*. Est-il judicieux de conserver les mêmes classes que pour la carte du monde?

Programmation statistique - QGIS

TD 3

STID1 - 2021/22

A. Cuzol

Exercice - 1 Importation de données attributaires par jointure

On a souvent besoin de compléter les données attributaires existantes avec d'autres variables, stockées par exemple dans un fichier Excel. Cette opération s'appelle une **jointure**.

Exemple : on dispose d'un fichier Excel "Présidentielle_2017_Resultat_Tour_2" contenant des informations sur le 2ème tour des élections présidentielles de 2017 : taux d'abstention dans chaque département, répartition des votes pour Emmanuel Macron et Marine Le Pen.

On souhaite mettre en place une **jointure** pour insérer ces nouvelles informations dans la table des données attributaires de la couche des départements "France_departements.shp". On pourra ensuite représenter ces informations sur des cartes.

1. Ouvrir la couche des départements "France_departements.shp".
2. Afficher la table des données attributaires pour voir quelles variables contient cette table.
3. Dans QGIS, ouvrir le fichier "Présidentielle_2017_Resultat_Tour_2" (même technique que pour une couche vectorielle). Afficher la table associée.
4. Faire la jointure pour ajouter les colonnes "Taux abs", "Voix Macron" et "Voix Le Pen" à la table des données attributaires de la couche "France_departements". Pour cela, aller dans *Propriétés de la couche* de la couche "France_departements" puis aller dans *Jointures*. Suivre les étapes suivantes :
 - (a) Cliquer sur "+" : "Ajouter une nouvelle jointure".
 - (b) "Joindre la couche" : choisir la couche issue du fichier Excel.
 - (c) "Champ de jointure" : choisir une variable du fichier Excel qui était déjà présente dans la table des données attributaires de la couche "France_departements". C'est cette variable qui permettra de faire la jointure entre la couche initiale et les nouvelles données du fichier Excel.
 - (d) "Champ de la couche cible" : choisir la variable contenant la même information que la variable choisie dans l'étape "Champ de jointure".
 - (e) Cocher "Champs joints" et sélectionner "Taux abs", "Voix Macron" et "Voix Le Pen".
 - (f) Remarque : on peut raccourcir le nom des variables qui seront créées dans la table des données attributaires en modifiant le préfixe dans "Personnaliser le préfixe du nom du champ".
5. Afficher la table des données attributaires de la couche "France_departements" et vérifier que les nouvelles colonnes ont bien été ajoutées.
6. Pourquoi certains départements contiennent des "NULL" ? Résoudre le problème.

7. Une fois que la jointure a bien fonctionné, sauver cette nouvelle couche avec le nom "France_departements_elections.shp".
8. Représenter le taux d'abstention dans les départements français à l'aide d'une représentation adaptée (avec un découpage en classes).

Exercice - 2 Remplacement de valeurs manquantes "NULL"

Dans le fichier "Origine_etudiants_STID_2015", on a les villes et départements d'origine des étudiants STID de la promotion arrivée en 2015. On souhaite représenter l'information des départements d'origine sur une carte.

1. Ouvrir le fichier "Origine_etudiants_STID_2015" avec Excel. A l'aide d'un tableau croisé dynamique, calculer le nombre d'étudiants par département.
2. Ouvrir un nouveau fichier Excel que vous nommerez "Nb_etudiants". Créer deux colonnes dans ce fichier : une colonne nommée "DEP" et une colonne nommée "NB_etud". Compléter ces deux colonnes avec le résultat du tableau croisé (utiliser un copier/collage spécial->valeurs).
3. Dans QGIS, ouvrir le fichier "Nb_etudiants". Afficher la table associée.
4. Faire la jointure pour ajouter la colonne "NB_etud" à la table des données attributaires de la couche "France_departements"
5. Sauvegarder le résultat comme une nouvelle couche "France_departements_STID_2015.shp"
6. Afficher la carte avec les départements d'origine des étudiants en fixant les classes suivantes :
 - 1 à 5
 - 6 à 10
 - 11 à 15
 - 16 à 20
 - plus de 20Quel problème constate-t-on?
7. On souhaite remplacer les valeurs "NULL" par des "0" (départements sans étudiant). Suivre les étapes suivantes :
 - (a) Dans la table, sélectionner les lignes contenant "NULL". Deux méthodes : manuellement ou avec l'expression "Nom_var is NULL" dans l'outil de sélection (où Nom_var est à remplacer par le bon nom de colonne).
 - (b) Rendre la couche éditable (cliquer sur le crayon dans la table).
 - (c) Ouvrir la calculatrice de champs (cliquer sur le boulier).
 - (d) Cocher "Ne mettre à jour que les entités sélectionnées".
 - (e) Cocher "Mise à jour d'un champ existant". Sélectionner la colonne contenant le nombre de candidats.
 - (f) Ecrire "0" dans "Expression", puis valider.
 - (g) Sauver la table modifiée en cliquant à nouveau sur le crayon.
 - (h) Vérifier que les "NULL" ont bien été remplacés par des "0" dans toute la table.
8. Afficher à nouveau la carte des nombres de candidats dans chaque département avec les classes suivantes :
 - 0
 - 1 à 5
 - 6 à 10
 - 11 à 15
 - 16 à 20
 - plus de 20
9. Ajouter le code de chaque département sur la carte.
10. Mettre en forme la carte en incluant la légende et un titre.

Programmation statistique - QGIS

TD 4

STID1 - 2021/22

A. Cuzol

Exercice - 1 Représentation de données attributaires avec des diagrammes

On peut utiliser un diagramme (camembert, barre) pour représenter la répartition d'une variable qualitative dans chaque entité géographique. La taille de chaque diagramme pourra être proportionnelle à une variable quantitative.

Exemple : On dispose du fichier Excel "Origine_etudiants_STID_2015" donnant des informations sur l'origine géographique des étudiants STID1 arrivés en 2015.

1. À l'aide d'un tableau croisé dynamique dans Excel, calculer :
 - le nombre d'étudiants dans chaque département ;
 - le nombre d'étudiants de sexe masculin dans chaque département ;
 - le nombre d'étudiants de sexe féminin dans chaque département.
2. Dans un fichier Excel qu'on appellera "Nb_candidats_sexe", créer les 4 colonnes suivantes et les remplir avec le résultat du tableau croisé :
 - CODE_DEP
 - NB
 - NB_M
 - NB_F
3. Ouvrir la couche vectorielle "France_departements.shp"
4. Faire la jointure pour ajouter les colonnes NB, NB_M et NB_F à la table de la couche des départements. **Sauver la couche après jointure en la renommant.**
5. Dans la table des données attributaires de la nouvelle couche obtenue par jointure, remplacer les "NULL" par des "0".
6. Aller dans *Propriétés de la couche/Diagrammes*.
 - (a) Choisir "Diagramme en camembert".
 - (b) Dans l'onglet "Taille", cocher "Taille variable" et choisir l'attribut NB. Dans "Valeur maximale", cliquer sur "Trouver". Dans "Taille", choisir 15 par exemple.
 - (c) Dans l'onglet "Attributs", choisir NB_M et NB_F. Valider.

Exercice - 2

Reprendre l'exercice 1 du TD 3 et représenter la répartition des votes pour Emmanuel Macron et Marine Le Pen dans chaque département. Tester 3 représentations différentes :

- Camemberts ;
- Barres empilées ;
- Diagrammes en barres (nommés de manière incorrecte "histogrammes" dans le logiciel).

TP noté sur le logiciel QGIS

STID1 - IUT Vannes

2019-2020

G. Copros et A. Cuzol

À lire avant de commencer :

- Le sujet et les données sont disponibles sur la plateforme pédagogique "Programmation statistique - QGIS".
- Tous les documents sont autorisés, mais **l'accès à tout type de ressource ou service en ligne est interdit**. Seule la connexion à la plateforme pédagogique est autorisée.
- Vous devez déposer **un unique fichier pdf** que vous nommerez "Nom1_Nom2.pdf" (ou "Nom1.pdf" si vous travaillez seul). Exemple : "Dupont_Durand.pdf". Ce fichier contiendra deux cartes sur une même page. Le rendu final doit être similaire à l'image donnée à la fin du sujet.
- Le fichier final devra être déposé sur la plateforme **au plus tard à l'heure de fin de l'évaluation**. Vous êtes invités à prendre vos précautions afin de ne pas dépasser cet horaire. Tout retard compris entre 0 et 5 minutes conduira à une note divisée par deux, et au-delà de 5 minutes les fichiers ne pourront plus être déposés et ne seront pas évalués.

Données à disposition :

- La carte des pays du monde : fichier "World.shp" et tous les fichiers associés.
- Le fichier "Commonwealth.ods" contenant la liste des pays appartenant au Commonwealth, et leur population en 2017.
- Le fichier "English_in_the_world.csv" contenant la liste des pays où l'anglais est utilisé de manière officielle, avec les différents statuts ci-dessous :
 - "Official" ;
 - "Official, in minority" ;
 - "Co-official" ;
 - "Co-official, in minority" ;
 - "Regional" ;
 - "Lingua franca".

Rendu attendu : vous devez rendre un unique fichier pdf reproduisant les deux cartes ci-après sur une même page. Quelques consignes à respecter :

Pour la carte des populations des pays du Commonwealth :

- Sur la carte du haut, respecter les classes utilisées, et utiliser la palette de couleurs "dégradés de bleus".
- Afficher les étiquettes en rouge, avec la police "Arial" et le style "Gras Italique".
- Pour afficher les pays n'appartenant pas au Commonwealth, afficher une couche séparée, en blanc, avec une opacité de 20%.

Pour la carte du statut de l'anglais :

- Respecter l'ordre utilisé dans la légende.
- Respecter à peu près les couleurs.

Cartes à reproduire :

