

Département Science des données  
Institut universitaire de technologie de Metz  
Université de Lorraine

Rapport de stage présenté par

**MARCEAU FINANCE**

Bachelor universitaire de technologie Science des données

2<sup>e</sup> année, parcours Exploration et modélisation statistique

# **ETUDE DU BUDGET DES COMMUNES FORESTIERES**

Stage réalisé au sein de

**BUREAU D'ECONOMIE THEORIQUE ET APPLIQUEE**

BUREAU D'ECONOMIE THEORIQUE ET APPLIQUEE  
14 rue Girardet  
54000, Nancy, France

Année universitaire 2023-2024



## Résumé

Le stage avait pour but de construire et d'analyser une base de données de panel sur le budget des communes forestières françaises. J'ai débuté le stage par une bibliographie pour m'imprégner de la comptabilité des communes et de l'économie des forêts. Ces recherches m'ont aidé à identifier des limites et problèmes potentiels de mon étude (identification des communes forestières, fusions et scissions de communes, données manquantes). Les principales données, les balances comptables des communes de 2010 à 2022, proviennent de la DGFIP<sup>1</sup>. Elles ont été complétées par des données de l'IGN<sup>2</sup>, de l'ONF<sup>3</sup> et de l'INSEE<sup>4</sup>. Le traitement des données a été réalisé avec Rstudio. L'analyse budgétaire montre que les communes françaises ayant un budget lié à la forêt sont de plus en plus nombreuses au fil des ans. La part des recettes et dépenses liée à la forêt augmente de 2010 à 2022 avec un léger creux en 2020 lié au Covid. Le stage se poursuit encore quelques semaines après la soutenance ce qui permettra d'affiner les résultats mais aussi de faire potentiellement de la cartographie.

---

<sup>1</sup> Direction Générale des Finances Publiques

<sup>2</sup> Institut national de l'information géographique et forestière

<sup>3</sup> Office National des Forêts

<sup>4</sup> Institut national de la statistique et des études économiques

## **Abstract**

The aim of my internship was to build and analyze a panel database on the budget of French forest municipalities. At first, I had to read several studies (technical and analytic reports, and scientific papers) about the forest economy in France and also studies about accounting for municipalities. This research helped me to understand the subject and identify some limits and potential problems, such as identification of forest municipalities, mergers and splits of municipalities, missing data. The data I used came from the official French government website « [data.economie.gouv.fr](http://data.economie.gouv.fr) ». This data includes the accounting records from all the French municipalities from the year 2010 to 2022, providing all the incomes and all the expenses. I also got side data to add other information about the municipalities and their forest such as the surface of the forest, which species they contain and how much surface they are on (main sources: IGN, ONF, INSEE). After all the manipulations, I finally got my dataset and did the data analysis.

## **Avant-propos**

Ce rapport de stage rentre dans le cadre de l'obtention de la seconde année de mon BUT SD (Science des données), IUT de Metz, dans le parcours Exploration et Modélisation Statistiques. Il étudiera les communes forestières ainsi que leur budget et les caractéristiques de certaines forêts communales.

Il est le fruit de 8 semaines de stage au sein du BETA à Nancy, du 15 avril au 7 juin 2024, qui a été encadré par Mme Niedzwiedz, M Garcia et Mme Montagné-Huck.

Le milieu de la recherche est un milieu très intéressant qui permet des perspectives de missions très variées.

Les difficultés étaient assez nombreuses notamment lors de la manipulation des différentes bases de données volumineuses à ma disposition, qui m'ont poussé à me renseigner davantage sur la manipulation de données sous R.

.

## Remerciements

Je tenais à remercier ma tutrice, Madame Niedzwiedz Alexandra, Gestionnaire de l'Observatoire pour l'économie de la Forêt, Référente données du laboratoire, pour toutes les aides lors de la création du code R, pour les apports en connaissance, fichiers, bibliographiques.

Je remercie également Monsieur Garcia Serge, Directeur de Recherche pour la proposition de stage au sein de son équipe, pour l'expression des attentes quant à de futures analyses économétriques, les discussions et conseils prodigués pour la construction de la base de données.

Je voudrais aussi remercier Monsieur Franck Gauzère pour s'être déplacé pour la visite de stage et pour ses conseils.

Enfin, je remercie Madame Montagné-Huck Claire, Ingénieure de d'études, pour ses connaissances partagées lors des différentes réunions et ses encouragements au cours du stage.

# Sommaire

Résumé .....	3
Abstract .....	4
Avant-propos .....	5
Remerciements .....	6
Sommaire .....	7
Lexique.....	9
Introduction .....	12
1. Présentation de l'organisation .....	12
1.1. Présentation synoptique de l'organisation.....	12
1.2. Historique de l'organisation .....	14
1.3. L'organisation dans son environnement.....	14
2. Mon stage et mes missions.....	15
2.1. Mes missions .....	16
2.1.1. Ma réappropriation des missions.....	16
2.1.2. Les outils à ma disposition .....	16
2.1.3. Les problèmes rencontrés et les solutions mises en place pour les résoudre et/ou les contourner.....	16
2.1.4. Les résultats obtenus .....	18
2.1.5. Mes préconisations .....	18
3. Importation et Transformation des données .....	19
3.1 Importation .....	19
3.2 Transformation .....	20
4. Analyses Statistiques.....	22
4.1 Évolution du nombre de communes par année.....	23

4.2	Répartition des communes en fonction de la surface de forêt sur la commune .....	24
4.3	Parts des dépenses et recettes forestières dans le budget communal.....	25
4.4	Distribution des SIREN (non-unique) par plage de dépenses .....	26
4.5	Distribution des SIREN (non-unique) par plage de recettes .....	27
4.6	Top 10 des départements avec le plus de SIREN .....	28
4.7	Evolution des recettes et dépenses moyenne.....	29
4.8	Moyenne des dépenses et recettes annuelle par classe de surface de forêt communale	30
4.9	Parts moyennes des recettes et dépenses par an par plage de surface communal .....	31
5	Conclusion.....	33
	Bibliographie.....	34
	Table des tableaux .....	37
	Table des figures.....	37
	Annexes .....	38
	Annexe 1 : Métadonnées des Fichiers Balances comptables. ....	39
	Métadonnées Fichier Balance comptables avant 2016.....	39
	Métadonnées Fichier Balance comptables après 2016.....	40
	Annexe 2 : Métadonnées Fichier Recensement .....	41
	Annexe 3 : Métadonnées Fichier Emplacement Forêt Communales .....	42
	Annexe 4 : Métadonnées Fichier IGN.....	43
	Annexe 5 : Description des comptes .....	44

## Lexique

**Forêt** : Une forêt est un milieu naturel où vivent une multitude d'espèces animales et végétales. (PEFC)

**Forêt communale** : Une forêt communale est un élément du patrimoine privé de la commune et aussi une composante du patrimoine forestier national, répondant à des enjeux d'intérêt général. (ONF)

**Commune forestière** : En l'absence de définition officielle, nous proposons ici deux manières de définir une commune forestière. La première est basée sur la présence d'une surface de forêt appartenant à la commune ; la seconde est basée sur l'observation, au moins une fois sur la période étudiée, de dépenses ou recettes forestières.

**Hectares (ha)** : Mesure de superficie qui correspond à 10 km<sup>2</sup> (Le Robert)

**Conifère** : Arbre dont les organes reproducteurs sont en forme de cônes (pomme de pin) et qui porte des aiguilles persistantes (ordre des Conifères ; ex. cèdre, pin, sapin...). (Le Robert)

Les conifères se trouvent essentiellement en zone montagneuse, dans le massif landais et dans les plantations assez récentes de l'ouest de la France.

**Forêt Mixte** : Forêt tempérée constituée d'un mélange de feuillus et de conifères. Les peuplements mixtes se rencontrent souvent en moyenne montagne ou dans les massifs forestiers accueillant les deux autres types de peuplements (Sologne, Dordogne, Bretagne). (lalanguefrancaise)

**Feuillu** : Arbre généralement à feuilles, caduques, c'est-à-dire perdant leurs feuilles à l'automne. Exemples de feuillus : Chêne, Merisier, Bouleau, etc. Les peuplements de feuillus se situent surtout dans les plaines ou à moyenne altitude. (PEFC)

**Sans-couvert** : Forêt ayant subi un changement brutal de couverture du sol suite à une perturbation anthropique (coupe rase) ou un incident (tempête, incendie...). (IGN) L'absence de couvert ne change pas l'occupation du sol qui reste catégorisé comme forêt.

**Budget** : Acte par lequel sont prévues et autorisées les recettes et les dépenses annuelles de l'État ou d'autres services assujettis aux mêmes règles. (Le Robert)

Dataframe : Un DataFrame est une structure qui organise les données en tables bidimensionnelles de lignes et de colonnes, comparables à une feuille de calcul. (databricks)

Données en panel : Les données de panel (ou données longitudinales) sont représentatives d'une double dimension : individuelle et temporelle. Un panel équilibré (balanced panel) a le même nombre d'observations pour tous les individus, un panel déséquilibré (unbalanced panel) est un panel où il manque des observations pour certains individus. (Cairn)



# Introduction

## 1. Présentation de l'organisation

### 1.1. Présentation synoptique de l'organisation

L'Institut des sciences et industries du vivant et de l'environnement, AgroParisTech, est une grande école française spécialisée dans l'agronomie, l'agroalimentaire, les forêts, l'environnement et le développement durable, placée sous la tutelle du ministère de l'Agriculture et de la Souveraineté alimentaire (MASA). AgroParisTech est réparti sur 8 sites, dont cinq campus en France métropolitaine (Paris-Saclay, Nancy, Montpellier, Clermont-Ferrand) et à Kourou. AgroParisTech a deux missions principales : former des ingénieurs et des chercheurs capables de répondre aux défis environnementaux et agricoles actuels, et mener des activités de recherche qui permettront à la société de s'adapter et de devenir meilleure. AgroParisTech est ainsi organisé en départements de formation et de recherche.

Tableau 1 : Informations juridiques de l'organisation

<b>Identité de l'organisation</b>	
<b>Dénomination</b>	<b>Institut des sciences et industries du vivant et de l'environnement (AgroParisTech)</b>
<b>Forme juridique</b>	<b>Établissement public à caractère scientifique, culturel et professionnel Sous tutelle du ministre chargé de l'agriculture</b>
<b>SIREN (siège)</b>	<b>130 002 850 00134</b>
<b>Activité principale</b>	<b>Enseignement Supérieur</b>
<b>Date d'immatriculation</b>	<b>13/12/2006</b>
<b>Code APE</b>	<b>8542Z</b>
<b>Début d'activité</b>	<b>13/12/2006</b>
<b>Adresse du siège</b>	<b>22, place de l'Agronomie, CS20040 91123 Palaiseau Cedex</b>
<b>Département du siège</b>	<b>91</b>

Figure 1 : Organigramme AgroParisTech

3000 étudiants (cycle ingénieur, master, maîtrise spécialisée, doctorat)  
70 élèves du corps des ponts, des eaux et des forêts  
250 enseignants-chercheurs, enseignants et cadres scientifiques  
500 agents administratifs et techniques



## ***1.2. Historique de l'organisation***

Si AgroParisTech a été officiellement créé fin 2006 par la fusion de 3 écoles d'ingénieurs, l'histoire de son campus nancéien remonte au XIX<sup>ème</sup> siècle. Le campus de Nancy est le site historique de l'école royale forestière fondée en 1824, devenue l'École nationale des eaux et forêts (ENEF) par décret du 15 octobre 1898. En 1965, l'ENEF fusionne avec l'École nationale du génie rural pour former l'École nationale du génie rural, des eaux et des forêts (ENGREF). Enfin, en vue d'améliorer la visibilité des grandes écoles à l'international, l'ENGREF devient AgroParisTech en 2007 en fusionnant avec deux autres écoles historiques : l'Institut National Agronomique (INA, créé en 1848) et l'École nationale supérieure des industries agricoles et alimentaires (ENSIA, créée en 1893).

## ***1.3. L'organisation dans son environnement***

Le centre Nancéien se concentre principalement sur les sciences forestières et les écosystèmes naturels. Il accueille deux unités mixtes de recherche : Silva et le Bureau d'économie théorique et appliquée (BETA).

Le laboratoire dans lequel j'ai effectué mon stage, le BETA, est une unité mixte de recherche des universités de Strasbourg et de Lorraine, du CNRS<sup>5</sup>, d'INRAE<sup>6</sup> et d'AgroParisTech. C'est le principal laboratoire en sciences économiques et de gestion du Grand Est. Il développe des outils de modélisation, de cliométrie, d'économétrie, des méthodes d'évaluation et de prospective. Ses activités sont organisées autour de six groupes de recherche et six enjeux sociétaux pour le développement durable :

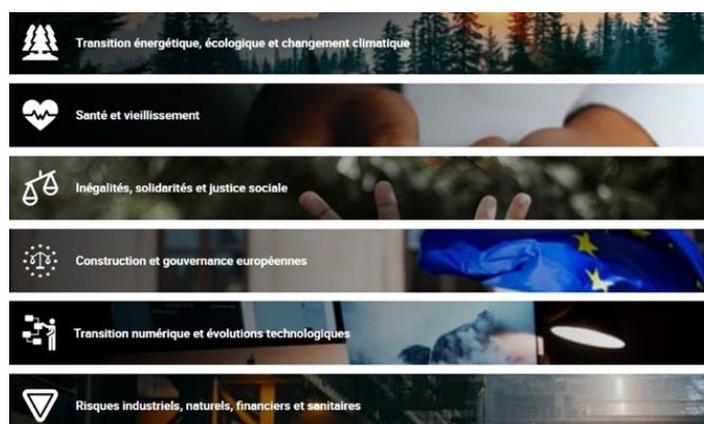
---

<sup>5</sup> Le Centre National de la Recherche Scientifique (CNRS) a pour mission de « mobiliser toutes les sciences pour affronter les défis du monde contemporain. <https://www.cnrs.fr/>

<sup>6</sup> INRAE, l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement, a pour mission de « produire et mobiliser des connaissances scientifiques, et de mobiliser ces connaissances au service de l'innovation, de l'expertise et de l'appui aux politiques publiques. » <https://www.inrae.fr/>



Figure 2 : Organisation scientifique du BETA



## 2. Mon stage et mes missions

L'objectif du stage était de construire une base de données de panel à partir des données comptables des communes, dans le but d'analyser le budget des communes forestières. Dans un premier temps, j'ai dû comprendre les variables de la base de données mise à disposition par la Direction Générale des Finances Publiques (DGFIP). Il m'a fallu agréger les différentes données des balances comptables annuelles des communes pour la période (2010-2022). Ensuite, j'y ai joint des données complémentaires permettant de caractériser les communes (surface, population) mais aussi les forêts qui leur appartiennent (surface totale de forêt, surface de forêt de feuillus, de conifères, etc.).

Une fois toutes les données agrégées, j'ai pu commencer les analyses. J'ai séparé les communes en fonction de leur surface forestière. En effet, le niveau de dépenses et de recettes liées à la forêt est directement lié à la superficie forestière détenue par la commune.

## 2.1. Mes missions

### 2.1.1. *Ma réappropriation des missions*

Les données relatives à la forêt sont très nombreuses, riches, complexes et trop peu exploitées, alors que les enjeux environnementaux et de politique publique sont très importants.

J'ai en premier lieu regardé les métadonnées en profondeur afin de mieux comprendre les différents jeux de données que j'allais traiter. Cela m'a permis de comprendre l'articulation entre les variables, et d'identifier les variables clés pour notre étude.

### 2.1.2. *Les outils à ma disposition*

Les différents outils à ma disposition sont les suivants :

- RStudio et les package Dplyr, Read\_Excel, ggplot2, gt

Dplyr, sert à la manipulation de base de données

Read\_excel, sert à lire les fichiers Excel

Ggplot2, sert à faire des graphiques de meilleure qualité

Gt, sert à faire de meilleurs tableaux

Ces outils m'ont permis dans un premier temps d'importer et dans un second temps de manipuler les différentes données, massives, que j'avais à ma disposition. J'ai ensuite pu utiliser R pour créer différents graphiques et dataframes afin d'analyser plus simplement les données.

- Plan comptable M14 des communes, il sert à avoir tous les numéros de compte utilisés pour la comptabilité des communes. Il permet de catégoriser les dépenses et les recettes.

- Guide R pour package « dplyr »

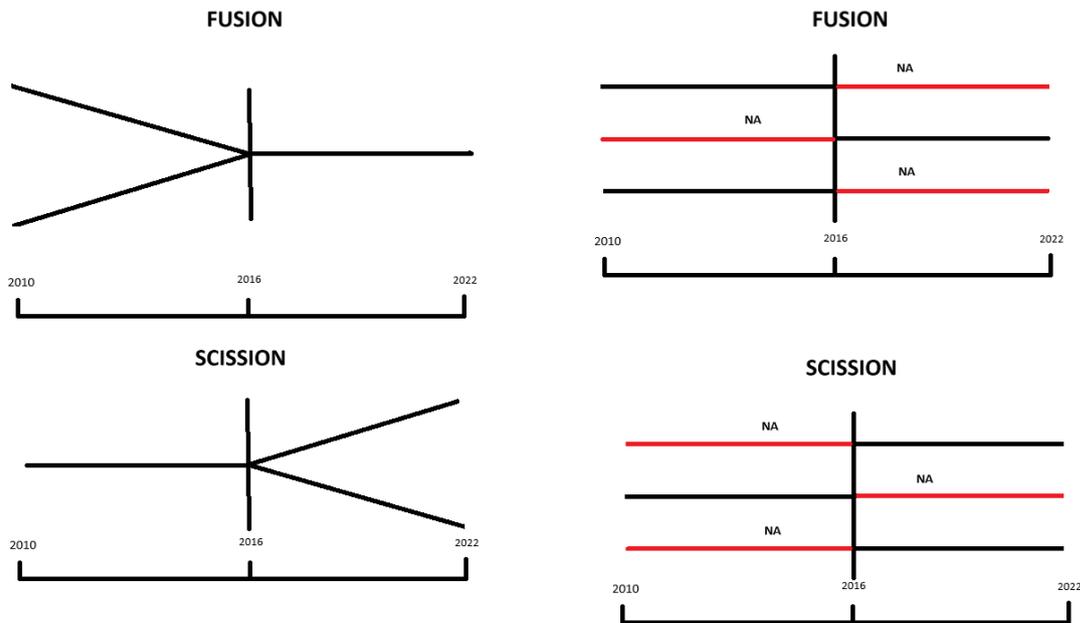
- Bibliographie : documents de la cour des comptes, sur les scolytes, sur les communes forestières m'ayant aidé à comprendre le contexte de l'étude.

### 2.1.3. *Les problèmes rencontrés et les solutions mises en place pour les résoudre et/ou les contourner*

Un des problèmes majeurs rencontré lors de l'importation des données est la fusion ou la scission des différentes communes, la première solution était de faire en sorte d'agréger les communes avant la fusion ou après la scission pour n'avoir qu'un seul individu. C'était possible grâce à des jeux de données décrivant les scissions et les fusions annuelles des différentes

communes. Cependant après réunion les encadrants du stage, il a été choisi de les considérer comme des individus à part entière et de ne pas les regrouper afin de pouvoir étudier l'ensemble des individus sur toute la période de 2020 à 2022.

Schématisation de fusion et scission de commune



*Figure 3 : Schéma fusion et scission*

Le second problème était de disposer de données en panel, donc il fallait qu'on ait les 12 années pour chaque commune même pour celles qui n'étaient plus là lors des années suivantes du fait de regroupements de communes, j'ai donc dû utiliser la fonction « complete » de dplyr pour rajouter toutes les années manquantes à toutes les communes, en mettant des NA dans les dépenses et les recettes.

Le troisième problème était la prise en compte des données de recensement (INSEE) pour les communes. Ces données sont en effet disponibles pour tous les 5 ans, j'ai donc décidé de prendre les données de 2008 et les affecter aux années 2010 à 2012, celle de 2013 aux années de 2013 à 2018 et enfin celles de 2019 aux années de 2019 à 2022.

#### *2.1.4. Les résultats obtenus*

Nous pouvons remarquer que les recettes augmentent. De nombreux facteurs peuvent l'expliquer. Une des raisons est le changement climatique qui pousse de plus en plus la société à réduire son empreinte carbone, et donc à utiliser les matériaux comme le bois qui stocke du carbone et dont le processus de transformation génère des émissions limitées de gaz à effet de serre.

Les dépenses ont plus tendance à varier d'une année à l'autre comparées aux recettes qui ont une variation plus faible.

Le nombre de communes forestières est en augmentation sur les 12 années étudiées.

Voir partie 4 de ce rapport pour une description détaillée des analyses réalisées.

#### *2.1.5. Mes préconisations*

Ma préconisation la plus importante est de souligner l'intérêt primordial d'effectuer une étude bibliographique avant de s'attaquer à l'analyse statistique dans un domaine que l'on ne connaît que très peu, cela permet de mieux d'appréhender le sujet mais aussi d'éviter d'aller vers des voies sans issue.

Vu la durée limitée du stage, il a fallu consacrer le temps nécessaire à la création d'une base de données fiable, qui facilite non seulement mes traitements mais aussi ceux qui seront réalisés après la fin du stage. En étant à l'écoute de mes encadrants, j'ai pu proposer des solutions répondants à leurs attentes.

J'ai effectué les premières analyses statistiques sur les données, qui permettront d'engager des analyses approfondies dans le but de mieux comprendre les comportements des communes forestières en matière de gestion forestière.

Dans les analyses, il faut bien faire attention aux variables prises en compte. En particulier, l'utilisation de la surface forestière nous oblige à être précis dans les variables utilisées. Par commune, nous connaissons la surface totale de forêt présente sur le territoire (IGN) par groupe d'essences (feuillu, résineux, mixte) mais sans distinction de la propriété (privée ou publique). Pour distinguer les forêts appartenant aux communes, sources de dépenses et de revenus pour les communes, nous devons passer par les données de l'ONF, qui ne distinguent pas les groupes d'essences. Si on travaille sur les données de l'IGN, il faut garder en mémoire qu'elles couvrent toutes les forêts (et pas uniquement les forêts communales). Quand on distingue le type de forêt majoritaire de la commune (feuillu, résineux, mixte), on fait l'hypothèse que la répartition par

groupe d'essences des forêts appartenant à la commune est la même que la répartition de l'ensemble de la forêt présente sur la commune.

### **3. Importation et Transformation des données**

#### ***3.1 Importation***

##### **Importation de la base de données principale**

Pour débiter ma boucle d'importation des données j'ai fait en sorte d'installer les fichiers dans le bon répertoire et de tous les avoir nommés de la même manière, seule l'année change d'un fichier à l'autre. Cela me permettra de créer une boucle avec une variable « i » variant de 2010 à 2022. Ensuite j'ai créé une variable texte qui recevra ma chaîne de caractère avec le paramètre « i » en son centre qui changera l'année en fonction de l'itération de la boucle. Une première data frame recevra les fichiers bruts, c'est sur cette dataframe que l'on va effectuer les premiers changements de type de variables, transformer les variables « OBNETCRED », « OBNETDEB » en numérique, ce sont les variables contenant les valeurs des comptes associés. On va aussi modifier le nom de la colonne « siren » pour la mettre en majuscule cela facilitera l'appel de colonne dans la suite du code. Une fois ces transformations basiques effectués je vais utiliser le package « dplyr » de « tidyverse » pour pouvoir effectuer les premières agrégations avec des « group\_by » sur le SIREN et le numéro de compte. On sommerà en même temps les « OBNETDEB » dans une variable « SD » et les « OBNETCRED » dans une variable « SC ». On enverra ensuite cette dataframe dans la dataframe Total, puis on importera chaque ligne dans une dernière dataframe « BudgetTotal ».

C'est dans cette dernière que nous associeront des niveaux d'importance à chacun de nos comptes : « type 1 » pour les comptes (7022, 7023, 7025, 61524), « type 2 » pour les comptes (7021, 7024, 7028, 7035, 7037, 6282) et « type 3 » pour les autres comptes.

Nous souhaitons en effet distinguer les dépenses et recettes liées à la forêt des autres. Certains numéros de compte ne concernent que les activités de Sylviculture et d'exploitation forestière (type 1), certains ne concernent que partiellement la forêt (type 2) et les autres ne les concernent pas du tout (type 3).

Nous allons maintenant devoir récupérer les bons noms de communes, les fichiers sont structurés de la même manière de 2010 à 2015 et d'une manière différente de 2016 à 2022, en effet il peut y avoir plusieurs noms de commune pour un SIREN, il peut y avoir les noms du

traitement des eaux ou d'autre structure appartenant à la commune mais qui ne possède pas le même nom que la commune, c'est pourquoi il y a des critères correspondant à la ligne avec le bon nom de commune correspondant.

Pour les fichiers allant de 2010 à 2015, les critères pour que le nom de commune récupéré soit le bon sont que la colonne Budget soit égale à « BP » qui signifie budget principal et pour les communes allant de 2016 à 2022 et il fallait que la colonne « CTYPE » soit égale à « 101 » (voir annexe).

Une fois ces conditions vérifiées, nous pouvons envoyer le nom de la commune « LBUDG » avec le numéro de département, le code Insee et le numéro de SIREN dans la data frame « unique\_communes\_df ».

### **Importation des bases de données annexes.**

J'ai importé un fichier annexe « surface-boise.xlsx » correspondant au fichier possédant toutes les caractéristiques forestières notamment la surface de la forêt mais aussi la composition en hectares de la forêt avec la surface des feuillus, mixtes, conifères, sans couvert et j'ai passé en caractère le numéro de SIREN.

Le second fichier annexe « emplacements-forêts-communales.csv », contient l'emplacement des forêts communales, un indicateur qui nous dira s'il s'agit d'une forêt communale ou non avec les variables géographiques, pour pouvoir effectuer des cartographies, nous aurons aussi la surface communale en hectares.

Le troisième et dernier fichier annexe « bdc21.csv », est un fichier de recensement ils nous renseignera sur la population de la commune. J'ai regroupé dans ce fichier tous les arrondissements de Paris en une seule commune étant donné que le découpage de Paris par arrondissement n'apparaît pas dans les autres fichiers.

## ***3.2 Transformation***

### **Transformations sur la base de données principale**

La première transformation sur les données à effectuer et celle sur la dataframe des noms de communes « unique\_communes\_df », on va utiliser la fonction « distinct » sur le numéro de SIREN étant donné que ce dernier est unique, pour n'avoir qu'une seule commune avec un numéro de SIREN.

J'ai ensuite créé l'INSEE de commune en agrégeant le code INSEE et le numéro de département, puis j'ai utilisé la fonction « ajouter\_zero ». Cette fonction rajoute le nombre de « 0 » suffisants pour que la longueur du code INSEE soit égale à 6 caractères ce qui est très important étant donné que l'INSEE de commune sera notre clé de jointure sur beaucoup de nos fichiers annexes.

On va maintenant réutiliser notre dataframe « BudgetTotal », on va grouper par SIREN, Année et importance et sommer « SC » dans la colonne Recettes et « SD » dans la colonne Dépenses. On va ensuite venir utiliser « pivot\_wider », ce qui va permettre d'avoir des colonnes Dépenses\_1, Dépenses\_2, Dépenses\_3 et Recettes\_1, Recettes\_2, Recettes\_3 (variables reflétant le budget par niveau d'importance) et on groupe par SIREN.

On va effectuer un premier filtrage pour n'avoir que des communes ayant des dépenses ou des recettes de type 1 ou de type 2.

Étant donné que nous voulons des données en panel, nous allons créer des années manquantes pour les communes qui ont fusionnées ou scissionnées en utilisant la fonction « complete » avec « fill » pour rajouter des valeurs NA pour les années créées.

### **Transformations sur la base de données complémentaires**

J'ai joint « surface-boise » avec les colonnes suivantes « `INSEE commune`, `Taux boisement`, `Surface commune`, `Surface forêt`, feuillus, conifères, mixtes, `sans couvert`, `annee PVA` » et la clé sera « Insee Commune », on utilise « left\_join », ainsi les communes qui ne correspondent pas ne seront pas supprimées comme avec un « right\_join » et « inner\_join ».

J'ai renommé les colonnes avec les unités correspondantes à la fin et on a décidé de « subset » toutes les communes où la surface de forêt est supérieure à 0.5 hectares.

J'ai joint « emplacement-forêts-communales.csv » avec comme clé « INSEE Commune », étant donné que l'on ne veut que les forêts communales, on va trier sur la colonne « frt\_com » qui est une colonne binaire qui permet de savoir s'il s'agit d'une forêt communale ou non. On choisira toutes les lignes où « frt\_com » qui vaut 1.

Les colonnes jointes sont les suivantes : « Nb\_forêt\_com, sum\_area\_ha\_fr\_com, iidtn\_fr, coordX, coordY, geo\_shape », les 3 dernières colonnes correspondent à des variables pour la

cartographie, « sum\_area\_frt\_com » correspond à la somme des hectares de forêts communales, et « iidtn\_frt » correspond à un identifiant de forêt.

Pour le dernier fichier de recensement, on a joint 3 colonnes « PTOT08, PTOT13, PTOT19 ». Ces colonnes correspondent à la population de la commune de 2008, 2013, 2019 ; la clé de jointure est « INSEE commune ».

Nous avons les années de 2010 à 2022, nous avons donc associé la PTOT08 à 2010 à 2012, puis PTOT13 de 2013 à 2018 et PTOT19 de 2019 à 2022.

Enfin nous avons créé 4 colonnes supplémentaires dans la data frame principale, « Dépenses\_Total » où l'on va sommer toutes les dépenses, et faire de la même manière pour les recettes.

On a créé « Part\_Dépenses\_Forêt » qui correspond au pourcentage de la dépense forêt par rapport à la dépense totale, on fait de même avec les recettes.

Et pour finir on extrait la dataframe dans un fichier csv pour pouvoir effectuer les analyses statistiques.

## **4. Analyses Statistiques**

Nous avons 26 140 communes uniques dans notre jeu de données ou l'on observe au moins une dépense ou une recette sur la période de 2010 à 2022. Les communes sont en moyenne boisées à hauteur de 31%. Les communes de notre jeu de données comprennent en moyenne 2 129 habitants. Il y a 10 637 communes qui possède de la forêt communale.

Nous présentons ici les principales caractéristiques des communes au vu de leurs données budgétaires.

#### 4.1 Évolution du nombre de communes par année

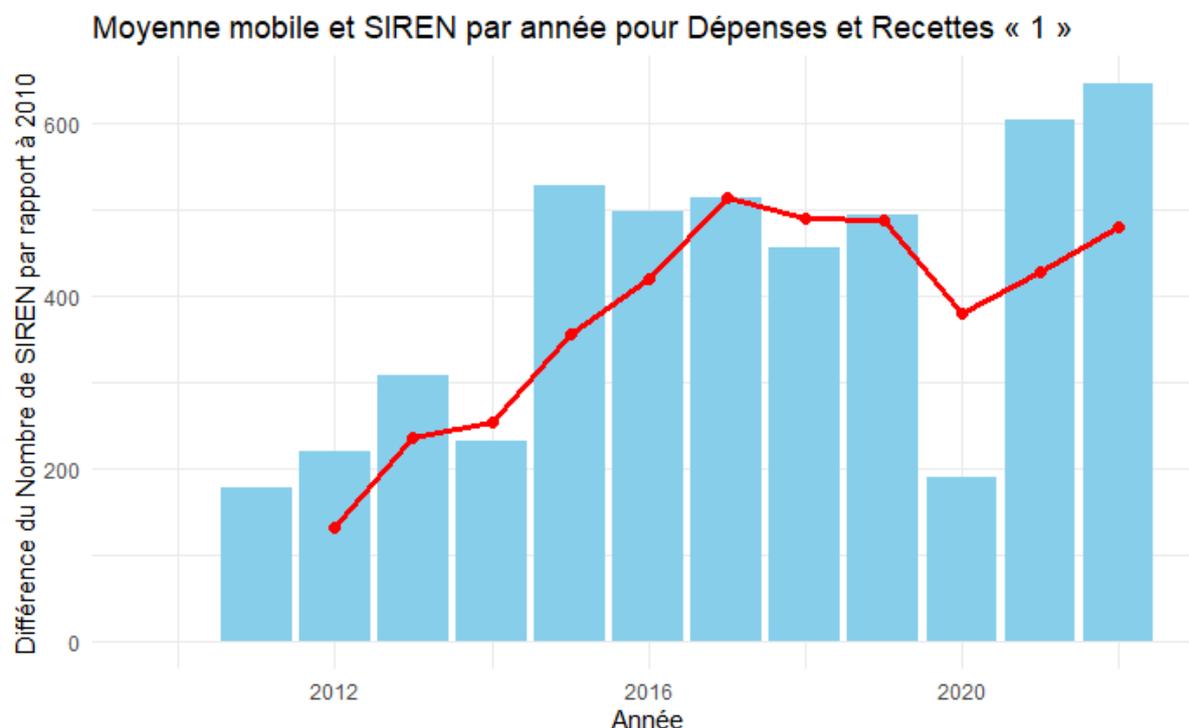


Figure 4 : Nombre de communes enregistrant des dépenses et recettes forestières

Ce premier graphique nous montre l'évolution du nombre de communes ayant des dépenses et des recettes forestières dans les comptes d'importance « 1 » par an, par rapport à notre année de référence qui est l'année 2010.

Le choix de faire apparaître une moyenne mobile n'est pas anodin. En effet cela permet aussi de rendre compte que sur toutes ces dernières années, hormis 2020 -à cause du Covid-, le nombre de communes ayant des dépenses et des recettes d'importance « 1 » ne fait qu'augmenter.

En 2020, à cause du Covid-19, les communes enregistrant un budget forestier est moins important que les autres années. En d'autres termes, les communes qui enregistrent des dépenses sylvicoles (liées à l'entretien des forêts) et des recettes forestières (coupes sur pied et ventes de bois abattus, ventes de menus produits forestiers, taxes d'affouages) sont de plus en plus nombreuses. En 2022, on observe 11 834 communes avec des dépenses et recettes forestières dans leur balance comptable.

## 4.2 Répartition des communes en fonction de la surface de forêt sur la commune

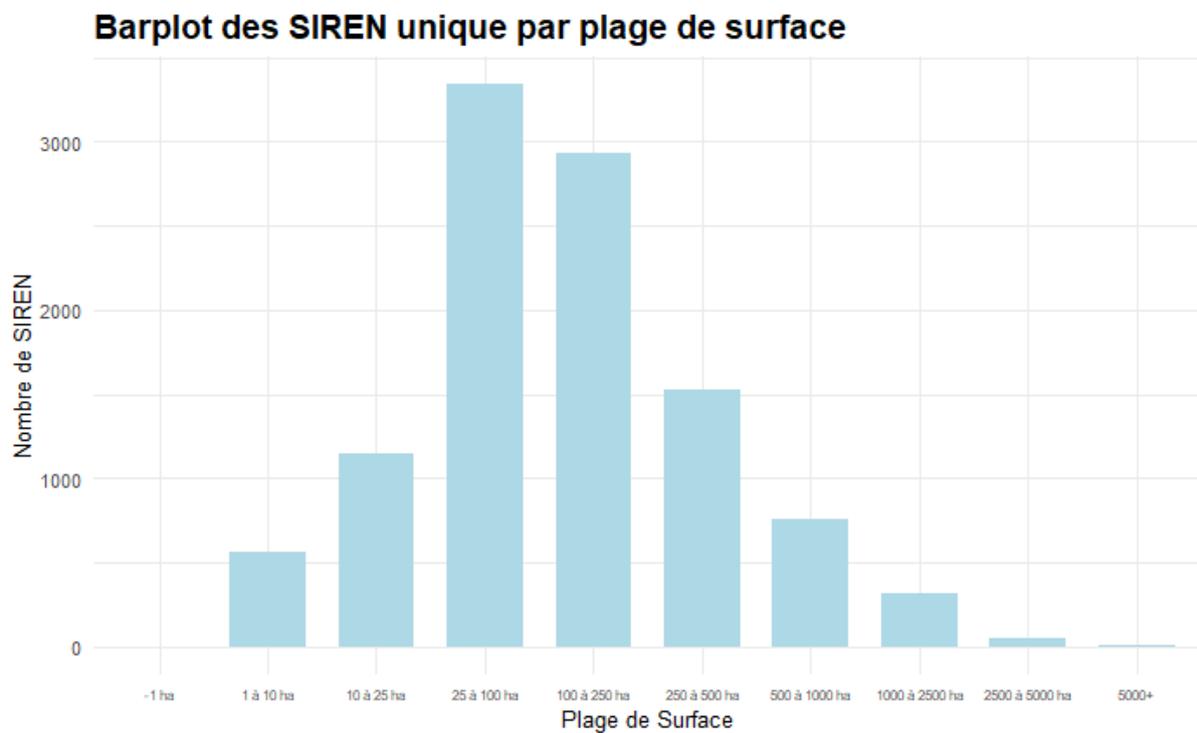


Figure 5 : Barplot des SIREN unique par plage de surface

Ce graphique montre le nombre de communes par classe de surface de leur forêt communale (forêt appartenant à la commune). La distribution du nombre de commune ressemble à une distribution normale, la surface de forêt communale moyenne est de 243 hectares. Il y a 5 communes qui possèdent une forêt communale supérieur à 5 000 hectares (Bayons, Asco, Laruns, Bédoin, Tende) et une seule qui en possède une de moins de 1 hectare (Propières).

### 4.3 Parts des dépenses et recettes forestières dans le budget communal

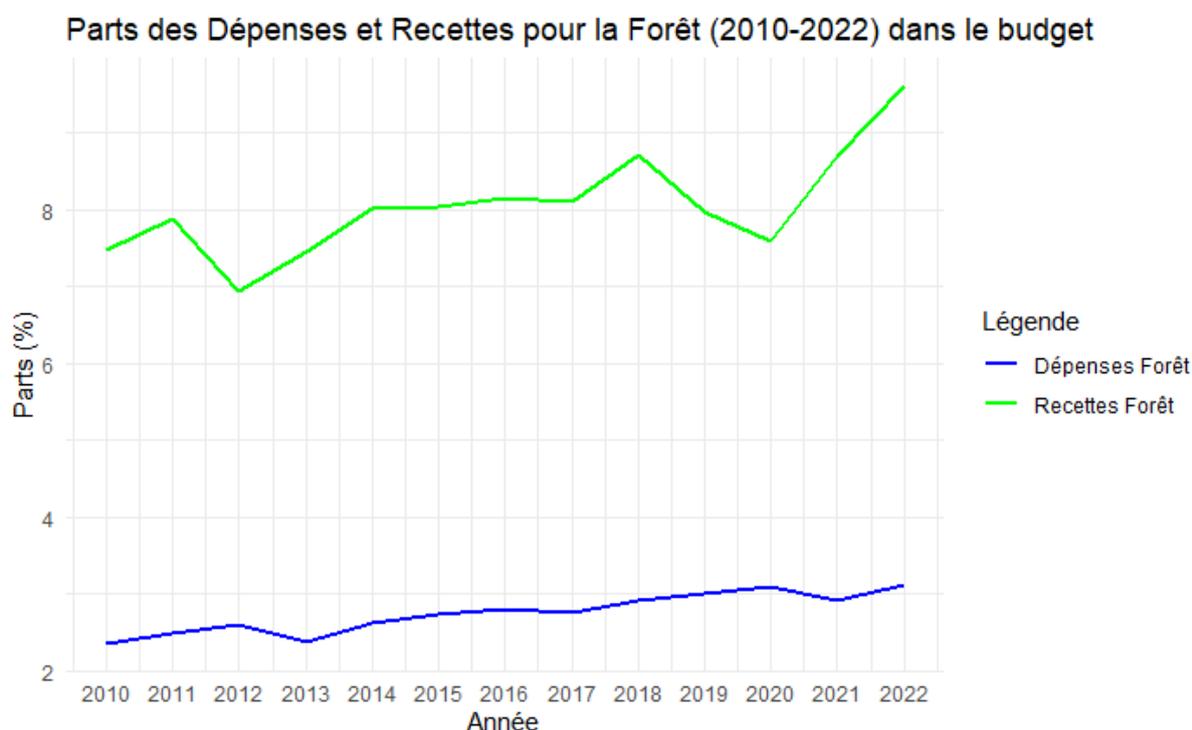


Figure 6 : Courbe des parts de recette et dépense forestières

Sur ce graphique on peut observer deux courbes, la part des recettes et des dépenses liées aux forêts par rapport aux dépenses et recettes totales. On remarque une chute abrupte du poids des recettes forestières en 2020, alors que le poids des dépenses forestières est stable. Cela pourrait traduire le fait que même si les communes vendent moins de produits forestiers, elles se doivent d'entretenir leurs forêts.

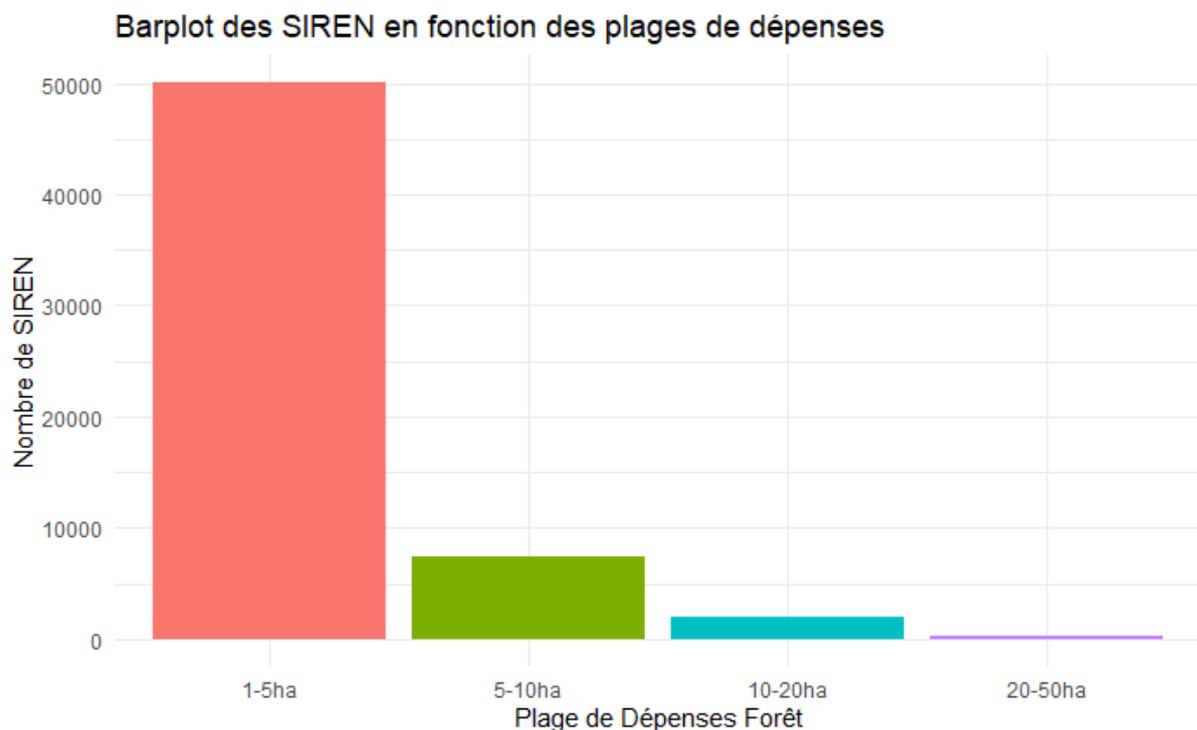
Malgré cette chute en 2020, on observe que la part des recettes et la part des dépenses ne font qu'augmenter (+0.12% en 2020 par rapport à 2010 pour les recettes, +0.7% en 2020 par rapport à 2010 pour les dépenses). Cela pourrait signifier plusieurs choses :

- Les communes considèrent de plus en plus les forêts comme une source économique intéressante. Les prix croissants des bois permettent de mieux valoriser les ressources des communes forestières ;
- Les communes exploitent davantage leurs forêts pour compenser la baisse d'autres recettes (dotations, subventions, etc.) ;
- Les coûts d'entretien sont de plus en plus élevés.

De plus la transition écologique pousse de plus en plus à utiliser des matériaux à faible empreinte carbone comme le bois. Cependant, l'exploitation des forêts, est parfois mal perçue par les usagers de la forêt (promeneurs) et associations environnementales, et source de tensions locales.

#### ***4.4 Distribution des SIREN (non-unique) par plage de dépenses***

Le graphique suivant montre comment se répartissent les individus (un individu étant une commune sur une année) en fonction du niveau du poids des dépenses forestières dans le budget global de la commune. Les communes sont donc comptabilisées autant de fois qu'elles enregistrent des dépenses forestières sur la période. Si la part du budget forestier d'une même commune varie dans le temps, la commune sera représentée dans les différentes classes correspondant à la part des dépenses forestières observées chaque année.



*Figure 7 : Barplot des SIREN en fonction de la plage*

Nous pouvons voir qu'une grande majorité des communes ont des dépenses forestières ont des dépenses qui sont très faibles par rapport à leurs budgets totaux.

Cependant, il y a quand même des communes où entre 20 et 50% de leurs dépenses sont liées aux dépenses forestières. Ces communes comprennent souvent des localités historiques où la gestion et l'entretien des forêts jouent un rôle central dans la préservation de leur patrimoine et de leur identité culturelle. De plus, ces communes peuvent dépendre fortement des ressources forestières pour leur économie locale, y compris le tourisme, l'exploitation du bois et la protection de l'environnement. Il sera intéressant dans la poursuite du travail de stage de proposer une cartographie de ces communes.

#### 4.5 Distribution des SIREN (non-unique) par plage de recettes

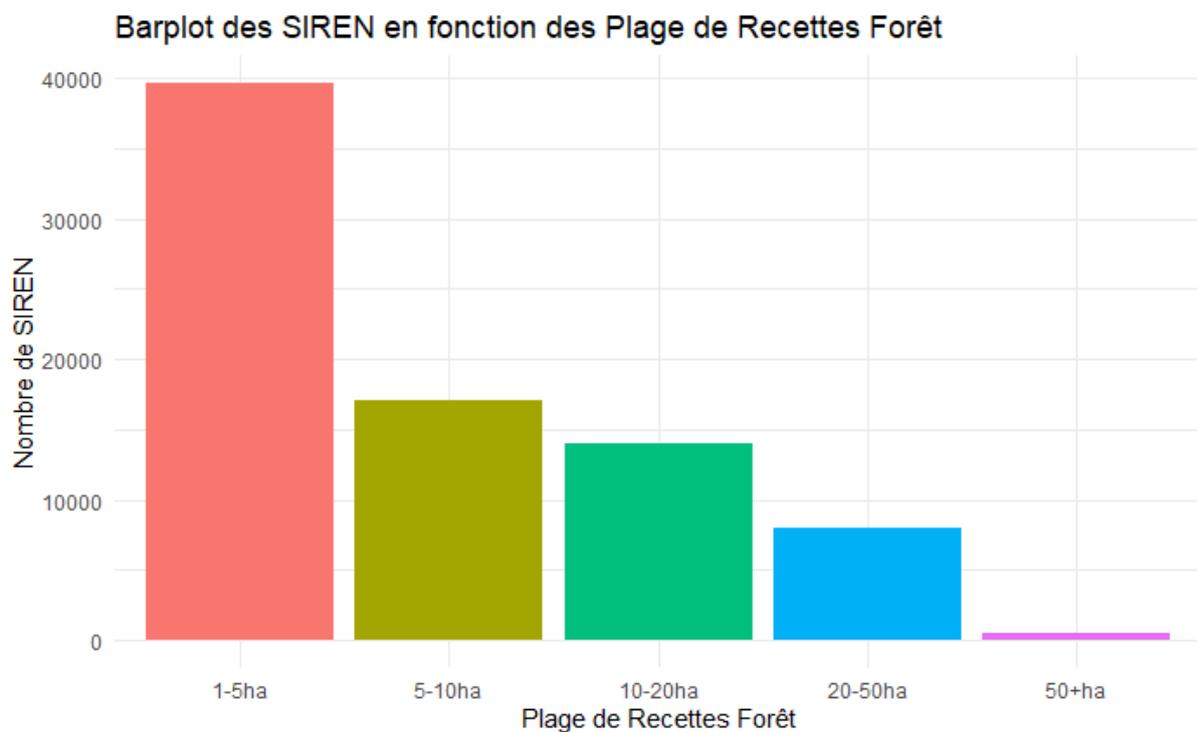


Figure 8 : Barplot des communes en fonction des plages de recettes

On peut voir malgré une domination de recettes très faible, qu'il y a quand même plus de communes qui effectue des recettes importantes par rapport à leurs recettes totales.

Il y a 405 SIREN dont plus de 50% des recettes dépendent de la forêt, les communes ne sont pas uniques sur ce graphique, en effet une ville peut apparaitre plusieurs fois car on prend toutes les années en comptes sauf celles où il n'y a pas eu de revenus ou de dépenses d'importance « 1 ».

#### 4.6 Top 10 des départements avec le plus de SIREN

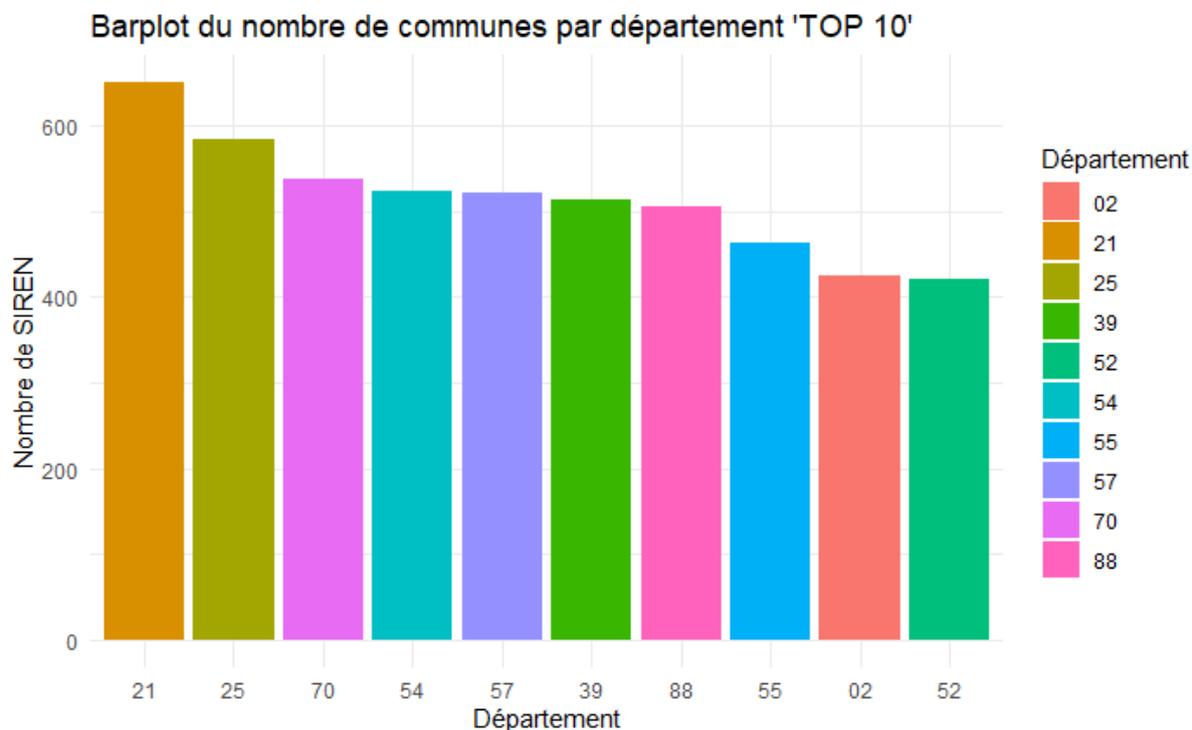


Figure 9 : Barplot du nombre de commune par département

Ce graphique n'affiche que les 10 départements avec le plus de communes forestières de mon fichier. Nous pouvons remarquer que la plupart des départements sont des départements du quart Nord-Est de la France, ce qui est cohérent avec la littérature consultée (sources).

Il y a donc plus de communes qui dépendent de leurs forêts dans le Nord-Est que dans le reste de la France. La région Grand Est est une région particulière dans le sens où, contrairement à la moyenne nationale, la forêt publique est plus importante que la forêt privée.

De plus les forêts du Nord-Est de la France sont très hétérogènes, en effet nous avons des forêts de plaines, plutôt feuillues, ainsi que des forêts de montagne, plutôt résineuses.

D'après le tableau ci-dessous, nous pouvons voir que les forêts qui se trouvent sur le territoire des communes sont des forêts de feuillus (Charmes, Bouleau, Chêne, etc.). Il s'agit des départements qui, comme dit précédemment, dépendent le plus de la forêt et le fait que le type majoritaire soit les feuillus (qui possède les types d'arbres les plus commercialisé en France), bien que dans les départements comme le 88 qui possèdent beaucoup de montagne, la part des

résineux et des feuillus reste assez similaire, puisqu'une partie du département est situé en plaine et une autre en montagne.

Tableau 2 : Nombre de communes parmi les types de forêts

Tableau des Types de Forêt des communes françaises		
Distribution des types de forêt avec les sommes des hectares		
Type_de_foret	Nombre d'Occurences	Surface_ha
coniferes	2,331	2,281,142
feuillus	22,885	8,701,339
mixtes	875	367,601
sans_couvert	49	67,087

De plus nous pouvons voir que le phénomène de domination est présent dans toute la France, avec à peu près 76,21% de la surface forestière recouverte de feuillus : il y a 22 885 communes qui ont une surface majoritairement feuillue.

#### 4.7 Evolution des recettes et dépenses moyenne

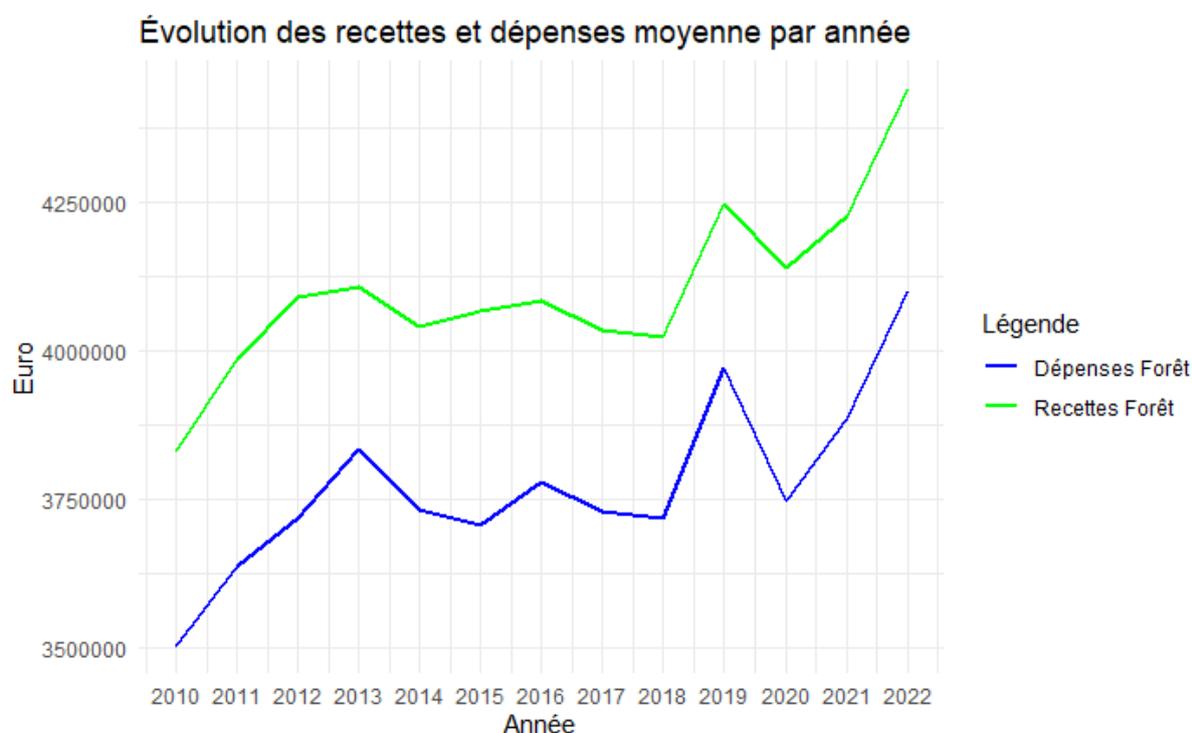


Figure 10 : Courbes des recettes et dépenses moyenne par année

La première chose que nous pouvons remarquer est l'allure très similaire des deux courbes. En effet, toutes les deux évoluent à la hausse et ont subi des baisses lors du covid sur l'année 2020. La courbe des dépenses a la même allure que celle des recettes, car les communes essaient de dépenser moins que ce qu'elles encaissent pour un budget équilibré ou bénéficiaire. Cela montre une gestion prudente des finances locales. Les communes cherchent à maintenir un équilibre budgétaire sur toute la période donnée.

Sur la période étudiée, les recettes annuelles sont supérieures aux dépenses. En 2022, les communes ont enregistré 332 millions d'euros de recettes forestières (recettes de type 1) et 74,8 millions d'euros de dépenses forestières (dépenses de type 1).

#### 4.8 Moyenne des dépenses et recettes annuelle par classe de surface de forêt communale

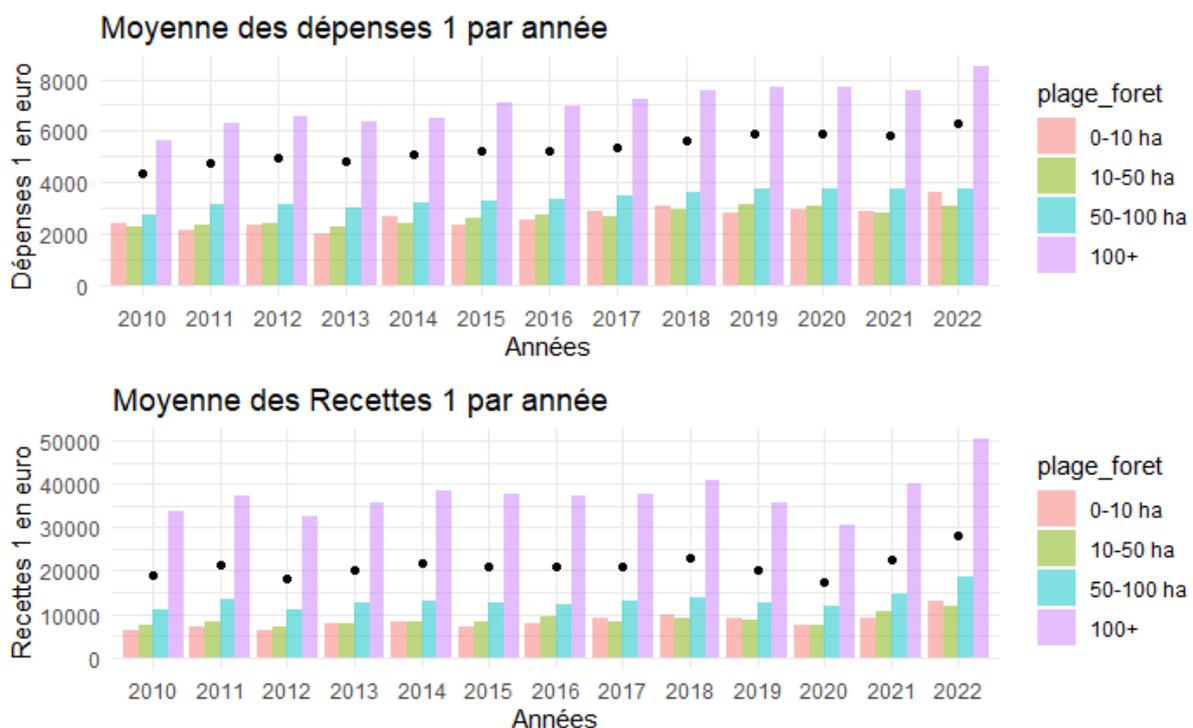


Figure 11 : Histogrammes des dépenses et recettes moyenne par an

Le point représente les dépenses et recettes moyennes par an de toutes les communes de notre jeu de données, alors que les plages ne sont que pour les communes ayant une forêt communale.

Sur les deux graphiques nous pouvons nous rendre compte que la plage qui domine le plus est celle ayant une surface de forêt communale supérieure à 100 hectares.

Nous pouvons aussi remarquer que la moyenne des dépenses et des recettes pour la classe des 100 hectares et plus a atteint le plus haut niveau jamais atteint en 2022 avec 50 275 euros. Le niveau maximum est également atteint en 2022 pour les autres classes de surface. Pour les dépenses, nous sommes face à une croissance modérée.

Il conviendrait de reproduire l'analyse sur des dépenses et recettes en euros constants afin d'éliminer l'effet de l'inflation.

#### ***4.9 Parts moyennes des recettes et dépenses par an par plage de surface communale***

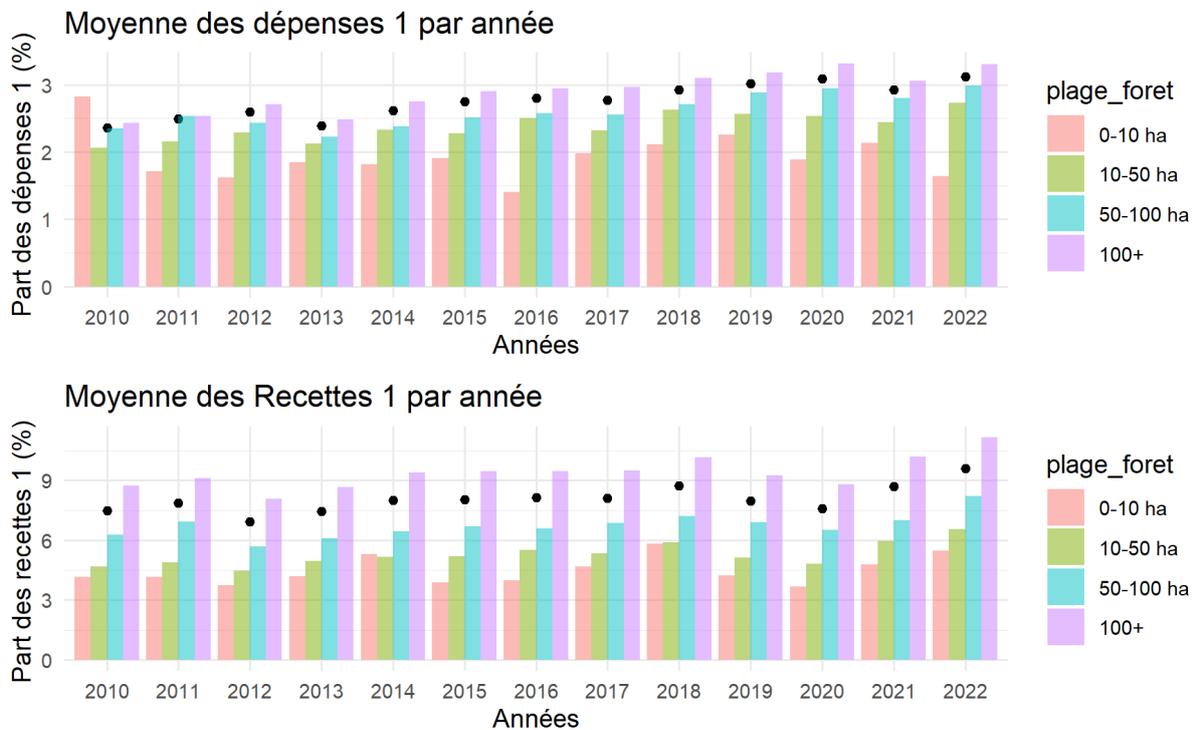


Figure 12 : Histogramme des parts des dépenses et recettes moyenne par an

En reprenant le même graphique mais cette fois en étudiant la part du budget forestier dans le budget total, cela nous permet de mieux visualiser à quel point la forêt est une partie importante dans le budget des communes.

On remarque qu'en 2010, les communes ayant moins de 10 hectares de forêt communale sont celles pour lesquelles la part moyenne des dépenses dans le budget total est la plus élevée, mais l'année suivante la part des dépenses est quasiment divisée par deux. Une des potentielles

explications à cette baisse pourrait être la tempête Klaus ayant eu lieu dans le sud de la France en 2009 ayant énormément impacté les communes du sud-ouest en causant beaucoup de dégâts dans les forêts et ayant engendré d'importantes dépenses de remise en état des forêts.

## **5 Conclusion**

En conclusion ce stage m'a permis de me rendre compte de comment se passe le travail au sein d'un laboratoire, ce qui est très différent d'une entreprise. En effet en entreprise le travail consiste plutôt à remplir une mission bien définie avec des compétences requises dans un domaine bien précis, dans un laboratoire de recherche ce n'est pas le cas la mission peut évoluer durant la durée du stage, mais peut aussi évoluer en fonction des différentes analyses et conclusion que nous voulons tirer, c'est pourquoi il faut souvent effectuer des points et des réunions pour constater l'avancement et voir si le sujet et les attentes n'ont pas changé.

Ce stage m'a aussi permis de me rendre compte que travailler derrière un bureau toute la journée n'est pas quelque chose de simple lorsqu'on apprécie le contact social autant que le fait de ne pas en avoir, cela m'a donc fait comprendre que j'ai besoin d'un environnement de travail où il y a plus de contact avec les personnes avec qui je travaille.

## Bibliographie

- ONF, 2024, « Frais de garderie de l'ONF payés par les communes ». <https://www.onf.fr/aux-cotes-des-territoires/lonf-et-les-communes-forestieres/+42::preserver-votre-forêt-avec-regime-forestier.html>

- CNFPT, 2018, « Notions clés sur les documents budgétaires »

[https://www.cnfpt.fr/sites/default/files/fiche\\_41\\_vademecum\\_kit\\_elections\\_2014\\_v1\\_0\\_2.09.2014.pdf](https://www.cnfpt.fr/sites/default/files/fiche_41_vademecum_kit_elections_2014_v1_0_2.09.2014.pdf)

- Sylvie Escande-Vilbois et Philippe Lacroix, 2002, « Les Budgets des communes forestières en crise »
- <https://hal.science/hal-03449424/document> IMF, 2004, « Manuel statistiques des finances publiques chapitre 5 Recettes »

<https://www.imf.org/external/pubs/ft/gfs/manual/fra/pdf/ch5.pdf>

- CGAAER, 2015, « Le régime forestier mise en œuvre par l'office national des forêts dans les forêts des collectivités »

<https://igedd.documentation.developpement-durable.gouv.fr/notice?id=Affaires-0008900&reqId=773b00ad-0376-4842-8ecb-e8a9cb0bc300&pos=1>

- INET & CNFPT, 2020, « Déchiffrer les comptes des collectivités locales : des clés pour agir »

[https://documentation.insp.gouv.fr/insp/doc/SYRACUSE/106463/dechiffrer-les-comptes-des-collectivites-locales-des-cles-pour-agir-un-cahier-des-eleves-administrat?\\_lg=fr-FR](https://documentation.insp.gouv.fr/insp/doc/SYRACUSE/106463/dechiffrer-les-comptes-des-collectivites-locales-des-cles-pour-agir-un-cahier-des-eleves-administrat?_lg=fr-FR)

- CGAAER, 2021, « Impact des scolytes sur le budget des communes forestières dans l'EST de la France et adaptation de la gestion des forêts communales au changement climatique »

<https://agriculture.gouv.fr/impact-des-scolytes-sur-le-budget-des-communes-forestieres-de-lest-de-la-france-et-adaptation-au>

- Communes forestières Puy-de-Dôme, 2021, « Questions fréquentes en lien avec les forêts des communes et des sections de communes »

([https://www.communesforestieres-aura.org/documents/ressources/63/2021-CoFor63\\_Guide\\_foncier\\_reediton.pdf](https://www.communesforestieres-aura.org/documents/ressources/63/2021-CoFor63_Guide_foncier_reediton.pdf))

- Cour des comptes, 2024, « La gestion durable de la forêt métropolitaine, quelle adaptation au changement climatique ? »

(<https://www.ccomptes.fr/sites/default/files/2024-03/20240312-RPA-2024-ENPA-gestion-durable-foret-metropolitaine.pdf>)

- Forestopic, 2024, « Des réactions mi-figue mi-raisin au rapport de la Cour des Comptes sur la forêt »

- LinkedIn ( Ronan Léonard), 2024, « Actualité rapport cour des comptes »

(<https://www.linkedin.com/feed/update/urn:li:activity:7183335364089028608/>)

- Légibase, 2024, « Compte 2117 – Bois et forêts »

(<https://compta-finances-locales.collectivites.legibase.fr/base-de-connaissances/m57-tome-1-le-cadre-comptable-la-nomenclature-par-nature/compte-2117-bois-et>)

- Jean-Claude Monin ( Annales des Mines – Responsabilité et environnement ),2009, « La forêt communale, au service de la gestion forestière durable et du développement local »

(<https://www.annales.org/re/2009/resumes/janvier/03-re-resum-FR-AN-AL-ES-janvier-2009.html>)

- D. Mathieu & A. Robert, 2021, « Forêts et finances communales dans le Doubs : incidence des revenus monétaires bruts des forêts communales sur le budget des collectivités locales »

(<https://hal.science/hal-03423916/document> )

- Collectivités locales, 2023, « Plan comptable M14 des communes »

(<https://www.collectivites-locales.gouv.fr/finances-locales/instruction-m14-comptabilite-des-communes> )

- Larmarange, 2023, « Manipuler les données avec dplyr ». <https://larmarange.github.io/analyse-R/manipuler-les-donnees-avec-dplyr.html>
- Bourbonnais, Régis. « Chapitre 13. Introduction à l'économétrie des données de panel », Économétrie. sous la direction de Bourbonnais Régis. Dunod, 2018, pp. 371-387. <https://www.cairn.info/econometrie--9782100773459-page-371.htm>

## **Table des tableaux**

Tableau 1 : Informations juridiques de l'organisation .....	12
Tableau 2 : Nombre de communes parmi les types de forêts .....	29

## **Table des figures**

Figure 1 : Organigramme AgroParisTech.....	13
Figure 2 : Organisation scientifique du BETA .....	15
Figure 3 : Schéma fusion et scission .....	17
Figure 4 : Nombre de communes enregistrant des dépenses et recettes forestières.....	23
Figure 5 : Barplot des SIREN unique par plage de surface .....	24
Figure 6 : Courbe des parts de recette et dépense forestières.....	25
Figure 7 : Barplot des SIREN en fonction de la plage .....	26
Figure 8 : Barplot des communes en fonction des plages de recettes .....	27
Figure 9 : Barplot du nombre de commune par département.....	28
Figure 10 : Courbes des recettes et dépenses moyenne par année .....	29
Figure 11 : Histogrammes des dépenses et recettes moyenne par an.....	30
Figure 12 : Histogramme des parts des dépenses et recettes moyenne par an .....	31

## **Annexes**

## *Annexe 1 : Métadonnées des Fichiers Balances comptables.*

### *Métadonnées Fichier Balance comptables avant 2016*

Variables	Définition	Observations
NDEPT	Numéro de département	Code département de région antérieur à 2016
LBUDG	Libellé du budget	On récupérera le nom des communes dans cette variable
INSEE	Numéro d'INSEE	Nous servira à créer l'INSEE de commune
BUDGET	Type de Budget : BP = Budget Principal BA = Budget Annexe	Nous servira à choisir les bonne ligne pour récupérer le nom de commune
SIREN	Numéro de SIREN	Servira de clé d'identification et de jointure
COMPTE	Numéro du compte	Nous permet d'identifier les bons numéro de compte pour notre variable « importance »
OBNETDEB	Opération Budgétaire Débit (montant net des annulations)	Nous permet d'avoir la valeur monétaire du débit.
OBNETCRED	Opération Budgétaire Crédit (montant net des annulations)	Nous permet d'avoir la valeur monétaire du crédit.

*Métadonnées Fichier Balance comptables après 2016*

Variables	Définition	Observations
NDEPT	Numéro de département	Code département
LBUDG	Libellé du Budget	On récupérera le nom des communes dans cette variable
INSEE	Numéro d'INSEE	Nous servira à créer l'INSEE de commune
CTYPE	Type d'établissement	Pour pouvoir identifier les lignes sur lesquelles nous récupéreront les noms de communes.  « 101 » pour Budget principal de commune.
SIREN	Numéro de SIREN	Servira de clé d'identification et de jointure
COMPTE	Numéro du compte	Nous permet d'identifier les bons numéro de compte pour notre variable « importance »
OBNETDEB	Opération Budgétaire Débit (montant net des annulations)	Nous permet d'avoir la valeur monétaire du débit.
OBNETCRED	Opération Budgétaire Crédit (montant net des annulations)	Nous permet d'avoir la valeur monétaire du crédit.

## *Annexe 2 : Métadonnées Fichier Recensement*

Variables	Définition	Observations
NCC	Nom en clair de la commune	
COM	Commune	Numéro d'INSEE de commune
PTOT19	Population totale 2019	
PTOT13	Population totale 2013	
PTOT08	Population totale 2008	

### *Annexe 3 : Métadonnées Fichier Emplacement Forêt Communales*

Variables	Définition	Observations
area_ha	Surface en hectares de la commune	
frt_com	Identifiant : 1 = forêt communale, 0 sinon	Sert dans l'importation pour agréger des données
CodeINSEE	Numéro d'INSEE	Servira comme clé de jointure
coordX	Coordonnées par rapport à l'axe des X	Servira pour la cartographie
coordY	Coordonnées par rapport à l'axe des Y	Servira pour la cartographie
geo_shape	Variables des différentes couches	Servira pour la cartographie

#### ***Annexe 4 : Métadonnées Fichier IGN***

Variables	Définition	Observations
INSEE commune	Numéro d'INSEE	Servira de clé de jointure
Surface commune	Surface de la commune en hectares	
Surface forêt	Surface de la forêt en hectares	
Taux boisement	(Surface forêt / Surface commune) en hectares	
Feuillus	Surface des feuillus en hectares	
Conifères	Surface des conifères en hectares	
Mixtes	Surface des forêt mixtes en hectares	
Sans couvert	Surface sans couvert en hectares	
annee PVA	Année de la dernière prise de vue aérienne	
iidtn_frt	Identifiant de forêt	

### *Annexe 5 : Description des comptes*

Numéro de compte	Description	Importance	Recette ou Charge
7022	Coupes de bois	1	Recette
7023	Menus produits forestier	1	Recette
7025	Taxes d'affouages	1	Recette
61524	Bois et forêts (élagage par des entreprises)	1	Charge
7021	Ventes de récoltes	2	Recette
7024	Remboursement forfaitaire TVA	2	Recette
7028	Autres produits agricoles et forestiers	2	Recette
7035	Locations de droits de chasse et de pêche	2	Recette
7037	Contribution pour dégradation des voies et chemins	2	Recette
6282	Frais de Gardiennage	2	Charge