

LES TERRITOIRES DU BADMINTON EN FRANCE

Mémoire de Master de géographie parcours
géomatique et modélisation spatiale

Maxime Hoyer

Directeur de Mémoire : **Samuel Carpentier Postel**

Année universitaire 2018-2019

Remerciements

Il m'est agréable d'adresser mes remerciements à tous ceux qui m'ont apporté de près ou de loin, aide et conseils lors de l'élaboration de mon mémoire.

Je voudrais remercier en particulier Monsieur Samuel Carpentier-Postel, Maître de conférences à Aix-Marseille Université, pour son encadrement et sa confiance dans un sujet peu orthodoxe. Ses conseils continus et variés m'ont permis d'avancer au mieux dans mon travail.

Je tiens à remercier également Monsieur Sébastien Bourdin, Professeur à l'EM Normandie Business School, pour les encouragements qu'il m'a apporté tout au long de cette année.

Je voudrais également remercier chaleureusement David Courbet, responsable informatique à la Fédération Française de Badminton, pour son aide et son amitié aussi bien dans le travail que sur les terrains de badminton.

Un grand merci à Léa Dapremont, mon amie, pour son travail sur la couverture de mon mémoire.

Enfin un immense merci à toutes les personnes qui m'ont aidé et soutenu de près ou de loin, que ce soit par leurs relectures ou leurs conseils.

Avant-propos

Ce mémoire est l'aboutissement de mon cursus universitaire de fin d'études en Master Géographie – Géomatique et Modélisation Spatiale à l'Université d'Aix-Marseille. Il présentera la répartition spatiale du badminton en France métropolitaine depuis 2008. Cette étude s'inscrit dans la continuité d'un travail centré sur le badminton et commencé lors de ma Licence MIASH¹.

Badiste² en Club depuis plus de 10 ans, j'ai déjà profité de mon engouement pour ce sport en fin de licence 3 où j'ai proposé un état des lieux du Badminton en région PACA. Dans la continuité, je me suis engagé dans un service civique (en accord avec la Ligue PACA de Badminton), pour continuer et approfondir mes recherches sur le développement du badminton en PACA. Et c'est donc naturellement que mon choix de mémoire de fin d'étude s'est porté sur le badminton.

Cette étude se veut être une contribution à la compréhension spatiale et sociale du badminton en France, sujet encore quasi inexploré. Ainsi cartes et statistiques permettront, tout au long du document, d'analyser la répartition spatiale de ce sport et d'en comprendre les possibles raisons.

Au cours de mes recherches j'ai toutefois rencontré un certain nombre de difficultés, notamment la fiabilité d'une partie des données des joueurs de badminton. D'autre part, il m'a fallu compenser mes lacunes en méthodes d'analyse géographique (formation préalable en MASTER MIAHS) pour étudier cette répartition spatiale

¹ Mathématiques et Informatiques Appliquées aux Sciences Humaines et Sociales.

² « Désigne un sportif qui pratique le badminton » ; L'internaute.fr.

Note : Tout le long de ce document, les termes « joueur », « badiste », « badminton » ou encore « club » seront en liens avec la Fédération Française de Badminton.

Table des matières

Remerciements	1
Avant-propos	2
Liste des figures	4
Introduction générale	5
Partie I : Le badminton en France, entre histoire et enjeux	7
Le sport dans les sociétés modernes	9
Les recherches dans le sport	11
Les recherches en géographie du sport	12
Le badminton comme objet d'étude	13
Les recherches sur le badminton	15
Le cadre et les hypothèses de travail	16
Partie II : Méthodologie, méthodes et outils d'interprétations	20
Les données	22
Analyses univariées	26
Analyses spatiales	28
Modèles bivariés et tests statistiques	31
Méthodes multivariées	39
Partie III : Le badminton, structures sociales et territoriales	44
Structure et répartition spatiale des badistes français	46
Diffusion spatiale du badminton	61
Caractéristiques sociodémographiques des badistes	66
Conclusion générale	79
Bibliographie	81
Annexe	85

Liste des figures

Listes des graphiques

Limite du taux de pénétration.....	33
Exemple du taux de variance expliquée	40
Exemple de dendrogramme.....	41
Exemple de classification.....	42
Nombre de licenciés par saison	46
Effectifs des badistes pour les saisons 2008-2009 et 2017-2018	47
Evolution du taux de pénétration des badistes par commune	58
Variances expliquées par les axes factoriels	67
Sauts d'inertie du dendrogramme	71
Partition en 2, 3 ou 5 classes	72

Listes des cartes

Nombre de badistes par département	49
Taux de pénétration par département	50
Evolution du taux de pénétration entre 2008-2009 et 2017-2018	52
Taux de pénétration des badistes par commune en 2008-2009.....	54
Communes ayant eu une création et/ou suppression de club entre 2008 et 2018	55
Taux de pénétration des badistes par commune en 2017-2018.....	56
Evolution du taux de pénétration par commune entre 2008-2009 et 2017-2018	57
Evolution du taux de pénétration entre 2016-2017 et 2017-2018.....	60
Indicateurs spatiaux des saisons 2008-2009 et 2017-2018	62
Concentration spatiale pour les saison 2008-2009 et 2017-2018.....	64
Classes issues de la CAH en 2017-2018	73

Listes des formules

Centre moyen	29
Distance standard	29
Equation du bilan démographique.....	31
Taux de pénétration.....	32
Equation de droite	34
Taux de variation.....	51

Introduction générale

Depuis la fin de la Seconde Guerre Mondiale, les sociétés développées ont consacré aux loisirs de plus en plus de moyens et de temps. A partir de ce moment, les loisirs ont évolué constamment et ont intégré des notions d'activités culturelles, touristiques, sportives etc. C'est de sport et notamment de l'évolution spatiale et temporelle d'une discipline, qu'il sera question dans ce document.

Le terme « sport » et sa définition n'ont cessé d'évoluer au cours du temps. On retrouve sa première origine au XII^{ème} siècle avec le verbe *desporter* qui signifiait, en ancien français, « s'ébattre ». Il désignait alors différentes pratiques de passe-temps. Quatre siècles plus tard, le terme *desporter* devient un synonyme de « s'amuser ». Puis au XVII^{ème} siècle, en Angleterre, le terme *desporter* se transforme en *disport* puis en *sport*. Enfin, c'est à partir du XIX^{ème} siècle que le mot « sport » commence à prendre son sens de « compétition ».

Selon Pierre de Coubertin (1863-1937), « *le sport est le culte volontaire et habituel de l'exercice musculaire intensif, appuyé sur le désir de progrès et pouvant aller jusqu'au risque* ». Mais, cette définition ne convient plus au sport moderne. Il existe, par ailleurs, de nombreuses définitions du mot « sport » :

- « Ensemble des exercices physiques, jeux individuels ou collectifs, donnant généralement lieu à des compétitions codifiées. » (Source : Larousse)
- « Toute activité physique régulière dont le principal but ou le seul but est la pratique corporelle. » (Source : Enquête Emploi du temps ; INSEE ; 1986)

- « Le sport c'est ce que font les gens quand ils pensent faire du sport. »³

Définir le sport n'est donc pas aisé. Il est plus pratique de le caractériser par ses attributs. Comme l'indique Grégory Martin dans sa thèse, « *le sport moderne se définit par quatre éléments indispensables :*

- *La motricité avec la mise en œuvre d'une ou plusieurs qualités physiques : d'endurance, de résistance, de force, de coordination, d'adresse, de souplesse, etc.*
- *Une activité institutionnalisée dont les règles tendent à être identiques pour l'ensemble de la planète.*
- *Une pratique majoritairement orientée vers la compétition.*
- *Une pratique organisée sous la tutelle d'une fédération. »*

Le sport a pris de plus en plus d'importance dans la vie quotidienne des français depuis plusieurs décennies. Pour preuve, le nombre de licences sportives a été multiplié par huit en près de soixante-dix ans (passant de deux millions en 1950 à seize millions en 2016⁴). Cependant ces chiffres ne prennent pas en compte les personnes pratiquant un sport hors fédération. Selon une enquête de l'INSERM en 2015, 66% de personnes interrogées⁵ pratiqueraient une activité physique régulièrement (au moins une fois par semaine).

Cette augmentation de la pratique sportive en France résulte donc de l'augmentation du temps alloué aux loisirs par les individus depuis plusieurs décennies. Dans ce contexte, même si tous les sports ne suscitent pas le même engouement, toutes les pratiques sportives sont amenées à évoluer et à se développer.

³ Institut National du Sport, de l'Expertise et de la Performance ; *Les pratiques sportives en France* ; 1986.

⁴ Ministère des Sports ; *Les Chiffres clés du sport 2017* ; 2017.

⁵ Individus âgés de 15 ans ou plus (France Métropolitaine + DOM).

Partie I : Le badminton en France, entre histoire et enjeux

Introduction de la première partie

Etudier un sport et notamment sous le prisme de la géographie, nous pousse à comprendre les différents questionnements qui en résultent. Le sport s'est grandement développé ces dernières décennies grâce à l'augmentation du temps alloué aux loisirs. De plus, il s'est imposé comme un phénomène mondial au cours du temps. En effet, la quasi-totalité des Etats au monde ont participé au moins une fois aux Jeux Olympiques⁶. De ce fait, le sport est devenu, en quelques décennies, un moteur économique et social.

C'est dans ce contexte que j'étudierai le badminton en France, que ce soit à l'échelle départementale ou communale. Ce document portera essentiellement sur les aspects géographiques mais essaiera tout de même d'élargir son point de vue grâce à des approches statistiques et sociologiques.

Comme nous le verrons, les liens entre sport et territoire sont évidents, mais la plupart des études portent sur les pratiques sportives en générale. Peu de recherches s'attardent sur un sport en particulier. Or, étudier une activité sportive précisément permet d'entrevoir les territoires avec un œil nouveau. Dans cet objectif, l'étude des territoires du badminton servira à identifier voire à comprendre les logiques d'implantation de ce sport de raquette.

Cette première partie va permettre de poser le contexte dans lequel s'inscrit l'étude ainsi que de présenter le cadre de travail. Dans cette partie, il sera donc question de sport, ainsi que de sa place dans le monde et dans la recherche. Par la suite nous évoquerons le badminton plus précisément. Enfin, cette partie posera le cadre de travail de cette étude.

⁶ Source : olympic.org

Le sport dans les sociétés modernes

Le sport « moderne » trouve son origine en Grande Bretagne à partir des années 1850. C'est à cette époque, que certaines activités physiques seront inventées tel que le football, le rugby, l'athlétisme ou encore le badminton etc. Ces sports seront ensuite exportés en Europe. Mais, les Anglo-Saxons ne se sont pas limités à la seule création d'activités physiques, ils inventent également les formes d'organisation de ces sports :

- Invention de règles
- Création d'associations sportives
- Elaboration d'un code éthique

Pour comprendre la fonction du sport dans les sociétés modernes, il faut mettre en avant les rôles qu'il induit.

Un Rôle **éducatif** : il joue un grand rôle éducatif notamment auprès des plus jeunes. L'apprentissage des règles participe au développement de la mémoire. Le respect de ces règles ainsi que les codes de conduite font partie du rôle éducatif du sport. De plus, le sport transmet des valeurs telles que le respect des gens, l'esprit de compétition, le dépassement de soi, etc.

Un rôle **social** : il favorise l'intégration sociale. Des personnes de milieux sociaux très différents peuvent se retrouver au sein d'une même association et partager des liens. Liens qui sont uniquement dus à leurs pratiques sportives communes. Le même phénomène intervient au niveau générationnel. Par ailleurs, le sport facilite également l'insertion de certaines personnes en situation de handicap. De plus, le bénévolat sportif permet un engagement de la personne envers son association et ses membres. Il participe également à une cohésion sociale internationale grâce à des grands événements sportifs tel que les Jeux Olympiques et Paralympiques (d'été et d'hiver), la Coupe du Monde du Football, etc. Selon la Fifa⁷, environ 3,572 milliards de personnes auraient suivi la Coupe du Monde de Football en Russie.

⁷ Fédération Internationale de Football Association.

Un rôle **économique** : il est un puissant moteur économique, créateur de richesses et d'emplois. Le secteur du sport génère, en 2016, près de 2% du PIB mondial⁸. Il se place donc entre des secteurs économiques tels que la culture (3% du PIB mondial) et le luxe (1%). Il serait estimé aujourd'hui, selon l'INSEEC⁹, à plus de 800 milliards €. De plus selon l'OCDE¹⁰ (Organisation de coopération et de développement économique), le sport est présenté comme une nouvelle source de croissance au vu de sa résistance aux chocs conjoncturels. En 2013 en France le poids économique du sport était estimé à 38 milliards € soit 1,8% du PIB national. Par ailleurs, le secteur privé du sport emploierait en 2017 plus de 120 000 personnes¹¹. Ainsi le marché du sport représente une part non négligeable de l'économie et de l'emploi en France. Ce rôle économique est l'enjeu le plus considérable du sport moderne.

La place du sport dans les sociétés modernes est donc très présente et en fait un élément incontournable. En France, le sport est enraciné dans la vie quotidienne des français. Mais quels sont les chiffres importants du sport en France ?

Fin 2017, le Ministère des Sports en partenariat avec le Ministère du Travail a publié un article intitulé *Le sport en quelques chiffres*. Cette étude présente notamment les chiffres du mouvement sportif en France :

- **107** : le nombre de fédérations affiliées au CNOSF (Comité National Olympique et Sportif Français).
- **307 500** : nombre d'associations sportives en France.
- **3,5 millions** : le nombre de bénévoles dans les associations sportives (CNOSF).
- **16 millions** : le nombre de licences délivrées par les fédérations en 2015.
- **34 millions** : le nombre de français de plus de 15 ans qui déclarent pratiquer une activité physique ou sportive au moins une fois par semaine en 2010.

Le sport est donc un élément majeur des sociétés contemporaines. Il cumule plusieurs facettes ; il est à la fois un loisir, une organisation institutionnelle, un enjeu économique majeur, etc. Il participe de ce fait, aux décisions politiques, que ce soit au niveau international, national ou communal. Il est directement concerné par l'espace géographique et participe à son organisation. Le sport apparaît donc comme un concept que la science ne peut omettre.

⁸ Ministère de la Ville, de la Jeunesse et des Sports ; *Les enjeux de l'économie du sport* ; 2016.

⁹ Grande école de commerce de management.

¹⁰ Organisation de la Coopération et de Développement Economique.

¹¹ Ministère du Travail & Ministère des Sports ; *Le sport en quelques chiffres* ; 2017.

Les recherches dans le sport

Le sport détient une place très importante dans le monde moderne, c'est pourquoi la recherche scientifique s'y intéresse de plus en plus. Le nombre d'études portant sur le sport n'a fait que croître en 50 ans. Aujourd'hui, les recherches sur le sport se retrouvent dans divers champs d'études scientifiques. De l'étude de la motricité à l'analyse psychologique des performances, le sport ne passionne plus seulement par sa pratique, mais à présent aussi par son étude.

Dans le cadre de ce mémoire, ce sont les recherches en sciences sociales qui vont être notamment intéressantes. Si la sociologie va se pencher sur les interactions sociales engendrées par le sport, la démographie va quant à elle se préoccuper des sportifs eux-mêmes.

Lara Muller en 2006 fait le lien entre caractéristiques sociodémographiques et pratiques sportives¹². Pour elle, les caractéristiques des individus ont un impact sur la pratique sportive. Elle va mettre en avant cinq variables qui influent la pratique physique :

- Le sexe
- L'âge
- Le niveau de diplôme
- Le niveau de vie
- La catégorie socio-professionnelle

Son étude démontre en effet une pratique sportive plus importante chez les hommes jeunes (15-24 ans) issus de milieux favorisés. Ce phénomène est d'autant plus important dans des sports comme le ski, l'athlétisme ou encore les sports de raquettes (tennis, tennis de table, badminton, squash).

Par ailleurs, des études démographiques permettent de mettre en avant l'importance de certaines caractéristiques sociodémographiques dans le choix de la pratique sportive. Cependant, elles ne prennent pas en compte le territoire. C'est dans ce contexte que l'apport de la géographie prend tout son sens.

¹² L. Muller ; *La pratique sportive en France, reflet du milieu social* ; Vie social ; 2006.

Les recherches en géographie du sport

Les recherches en géographie du sport ne sont pas nouvelles. On retrouve les premières traces à la fin des années 1870. Les études ont montré la place importante du sport dans les activités culturelles et économiques mais aussi son lien avec l'espace, les localisations et les flux. Ainsi ce champ d'étude permet de comprendre la notion de « territoire sportif ».

Les études en géographie du sport s'appuient généralement, comme l'explique Jean-Pierre Augustin, Philippe Bourdeu & Loic Ravenel¹³, sur une double démarche. « *La première interpelle l'ensemble des sciences sociales : la sociologie, l'histoire, l'économie, la science politique qui contribuent à l'explication, dans la mesure où leurs objets s'inscrivent dans l'espace et le différencient. Telle localisation sportive est expliquée par les variations dans la composition sociale des populations, telle autre par une présence historique ancienne, par des enjeux économiques ou par l'influence d'une volonté politique pour son développement. Le géographe interprète alors la spatialisation des faits sociaux et les combine en fonction de leurs localisations. La deuxième démarche interroge plus directement la discipline géographique en posant comme fondement explicatif la localisation elle-même. Celle-ci n'est plus seulement la conséquence d'autres processus, mais souligne les propriétés de l'espace comme la distance, la contiguïté, la superficie ou bien encore la situation.* »

Autrement dit soit les phénomènes économiques et sociaux contribuent à la différenciation de l'espace, soit ce sont les caractéristiques spatiales des territoires qui induisent des espaces différenciés.

Comme l'indique les auteurs de la *Géographie des Sports en France*, la répartition de la population est l'un des principaux facteurs d'analyse en géographie. En effet les villes, les régions côtières, les axes de communication (tel que les autoroutes ou les fleuves) sont des zones plus peuplées. Or, pour les auteurs, ces zones densément peuplées permettent la concentration des biens et des personnes, favorisant la diversité de l'offre sportive. De plus pour Augustin, Bourdeu et Ravenel, les territoires fortement peuplés permettent souvent l'émergence de nouvelles pratiques sportives et de distinction sociale via le sport.

¹³ J-P. Augustin, P. Bourdeu et L. Ravenel ; *Géographie des Sports en France* ; 2008.

Pour Guy Di Méo¹⁴, il est important de dissocier le territoire et le lieu. Il explique que le territoire est généralement abstrait, il tient plus d'une idée et d'un ressenti que de limites précises. A l'inverse, le lieu se caractérise par son usage ainsi que son contour géographique. Selon Di Méo, le territoire regroupe et associe des lieux. Dans ce cas, parler de territoire à propos d'une activité physique et sportive revient à mettre en exergue les secteurs où la pratique y est intensive ou absente.

A partir des composants inhérents à la géographie sportive, des analyses de répartition et de dynamiques spatiales d'une pratique sportive sont donc possibles.

Le badminton comme objet d'étude

Au V^{ème} siècle en Chine et en Amérique était pratiqué le « jeu du volant ». Cette pratique consistait à se renvoyer un volant avec le pied. Cet exercice ressemblait plus à notre football actuel qu'à du badminton. Pourtant on a vu l'apparition d'un volant. Plus tard le « jeu du volant » évoluera avec l'utilisation d'un battoir pour frapper le volant. Cette pratique se développera également en Europe au XIV^{ème} siècle où il sera si populaire qu'à Leicester, le Mardi Gras sera baptisé « La journée du volant ». Cette notoriété sera renforcée par François I^{er}, Charles IX, Louis IX, ainsi que toute l'aristocratie de cette époque qui en feront leur passe-temps privilégié. Au point même que le « jeu du volant » est recommandé par des pédagogues et des philosophes au début du XVIII^{ème} siècle.

Pourtant c'est en Inde que le badminton moderne va voir le jour. Ce sont les officiers britanniques en place là-bas qui vont l'importer en Angleterre avant de codifier et d'institutionnaliser le badminton tel qu'il existe aujourd'hui. En septembre 1893, une fédération anglaise de badminton est créée. L'essor de ce nouveau sport devient donc inéluctable avec la création de nombreux clubs ainsi que l'officialisation des règles. Pourtant, ce n'est qu'en 1934 que des représentants de fédérations nationales décident de créer une fédération internationale (IBF). Cette naissance est bien tardive en comparaison à d'autres disciplines telles que le tennis ou encore le football.

¹⁴ Guy Di Méo ; *Géographie sociale et territoires* ; 1998 ; Paris, Nathan

Malgré le succès de grandes compétitions internationales (Championnat du Monde, Coupes Thomas, etc.) ainsi que le nombre croissant de pratiquants dans le monde, il faudra attendre les Jeux Olympiques de Barcelone en 1992 pour que le badminton intègre les olympiades. A cette époque-là, la domination asiatique sur ce sport de raquette est déjà totale. Le premier à les bousculer sera le Danois Poul Erik Hoyer Larsen qui remportera la médaille d'or en simple messieurs au Jeux Olympiques de 1996.

On retrouve les premiers pas du badminton en France, en Bretagne au début du XX^{ème} siècle. Mais c'est à Dieppe que le badminton va prendre de l'envergure pour ensuite se diffuser en Europe Occidentale.

Le 18 janvier 1934 est créée la Fédération Française de Badminton (FFBad). De nombreux sports vont prendre leur envol durant l'entre-deux-guerres. Pourtant le badminton lui n'arrivera pas à percer sous l'ombre du tennis et du lawn-tennis¹⁵. La FFBad sera dissoute en 1941 par le régime de Vichy, puis rattachée à la Fédération Française de Tennis en 1944. La Fédération Française de Badminton (FFBad) comme nous la connaissons actuellement sera créée le 7 décembre 1978.

En 2010 Paul-André Tramier, alors président de la FFBad, expliquait dans une interview¹⁶ que le badminton était le sport numéro un en Asie. Pourtant, en 2017, la FFBad est encore à la 14^{ème} place des fédérations françaises en nombre de licenciés. Malgré cela, la FFBad a connu une très forte progression depuis 20 ans, avec une augmentation d'environ 218% de ses effectifs. Cependant avec un peu plus de 180 000 licenciés en 2017¹⁷, elle reste très loin de la Fédération Française de Football (plus de deux millions de licenciés) et de la Fédération Française de Tennis (plus d'un million de licenciés). Le badminton a donc encore des possibilités de progression sur le territoire français.

Ce potentiel de croissance peut être expliqué par plusieurs phénomènes. Paul-André Tramier expliquait en 2010 que c'est l'attrance des jeunes à l'école pour le badminton qui va permettre au sport de poursuivre son expansion : *« il est vrai que les pays latins comme la France sont,*

¹⁵ Julie Grall ; *Sport ou pratique infantine ? La construction manquée du badminton en tant que « vrai sport » en France dans l'entre-deux-guerres* ; 2015

¹⁶ Interview du journal 20 Minutes à l'occasion des Championnats du Monde de badminton à Paris.

¹⁷ Chiffres du Ministère des Sports ne prenant pas en compte les fédérations multisports.

ou plutôt étaient, en retard. Mais il faut voir qu'on est aussi le sport le plus pratiqué à l'école. Le potentiel du badminton est donc très important. [...] Le succès du badminton à l'école est simple à expliquer : c'est un sport où on s'amuse et qui a une bonne rationalisation de l'espace ».

De plus, le badminton prend peu à peu place dans les médias français et notamment à la télévision. Après avoir diffusé, en 2015, pour la première fois de son histoire les Yonex Internationaux de France, la chaîne L'Equipe 21 a signé en 2016 un contrat de diffusion avec la FFBad. C'était la première fois qu'une compétition de badminton était diffusée en clair à une heure de grande écoute. Au vu du succès, la chaîne a depuis renforcé la présence de ce sport olympique à la télévision en rajoutant à sa grille les Championnats d'Europe et les Championnats du Monde.

Si les études en géographie du sport ne sont pas récentes en France, aucune ne traite du badminton malgré son potentiel de croissance. C'est dans ce contexte qu'il m'a paru pertinent et intéressant d'en étudier les contours

Les recherches sur le badminton

Il existe un certain nombre de recherches sur le badminton, que ce soit en France ou dans d'autres pays, mais une grande partie de ces études portent sur le sport en lui-même. Elles analysent certains mouvements techniques, l'impact de certains déplacements sur le corps ou encore la pédagogie du sport, etc.

On trouve cependant quelques études plus orientées dans les sciences sociales. Notamment une thèse soutenue en 2018 par Julie Grall, qui traite du badminton en France de la fin du XIX^{ème} siècle à 1979. Cette thèse mettait en évidence les pratiques et les représentations du badminton à cette époque-là en France. Pour y parvenir, Julie Grall s'est servi d'articles de journaux publiés pendant près de cent ans afin de comprendre comment a évolué la perception de ce sport de raquette durant cette période.

De plus, un mémoire sociodémographique des pratiquants de Badminton en France a été réalisé en 2016 par Simon Maneo¹⁸. Il s'appuie sur un questionnaire réalisé par le Centre National de Développement du Sport portant sur les pratiques physiques et sportives ainsi que sur des questionnaires clients passés dans des boutiques spécialisées dans le badminton. Simon Maneo tente d'identifier qui sont les joueurs de badminton en France grâce à un échantillon de pratiquants de badminton. Cependant, cette population comprend les pratiquants de badminton en général (affilié ou non à la FFBad). De ces questionnaires clientèles, il en tirera certains résultats :

- La clientèle est majoritairement masculine (environ 60 %).
- Les clients sont relativement jeunes, entre 20 et 35 ans.
- La clientèle est en grande partie affiliée à la FFBad (76%).

Les sciences sociales et notamment la géographie ont donc très peu abordé le sujet du badminton.

Le cadre et les hypothèses de travail

Avant de parler de méthodes et d'analyses, il est important de poser et de définir un cadre à notre étude. Ce document va traiter des territoires du badminton en France. Cependant, même si les raisons du choix du badminton ont déjà été présentées précédemment, nous n'avons pas parlé des bornes spatiales et temporelles ou encore de la population. C'est-à-dire qu'entend-on par « la France », sur quelle période l'étude portera-t-elle, et quelle sera la population étudiée précisément ?

LE TERRITOIRE

Ce document abordera le badminton en France, mais uniquement la France métropolitaine. Les départements et territoires d'outre-mer sont donc écartés des analyses. Ces territoires se trouvant loin de la France métropolitaine, sur d'autres continents, ils possèdent des pratiques et des cultures quelques peu différentes de l'hexagone. Il serait donc difficile de les analyser de la même manière et mériterait des études spécifiques.

¹⁸ S. Maneo ; *Etude sociodémographique des pratiquants de badminton en France* ; 2016.

LA PERIODE

En ce qui concerne la période d'étude, nous parlerons de saisons sportives et non d'années calendaires. Une saison sportive se déroule du 1^{er} septembre au 31 août. Par exemple, la saison 2008-2009 s'étale du 1^{er} septembre 2008 au 31 août 2009.

L'étude portera sur dix saisons sportives de la saison 2008-2009 à 2017-2018 ; période comprenant trois saisons post-olympiques (Pékin en 2008, Londres en 2012 et Rio de Janeiro en 2016). Par ailleurs, ce choix de saison va de pair avec une fiabilité des données. D'autres part, la saison 2018-2019 étant toujours en cours, la FFBad ne dispose pas encore des données définitives.

LA POPULATION

Les populations étudiées seront ici les joueurs et les clubs affiliés à la Fédération Française de Badminton. Sont exclus tous les pratiquants de badminton dans les milieux scolaires, les clubs non affiliés ainsi que les clubs multisports, etc. Cependant, il n'est pas possible d'étudier les caractéristiques sociodémographiques des personnes affiliées à la Fédération de badminton. Le seul moyen serait de réaliser un questionnaire. Mais sa conception serait difficile à réaliser si l'on veut étudier le badminton dans la France entière. C'est pourquoi, ce seront les caractéristiques (effectifs selon l'âge, le sexe, etc.) des territoires qui seront analysés.

LES HYPOTHÈSES

Pour finir cette première partie, il est important d'évoquer les questionnements qui sont au cœur du choix de ce sujet. Cela passe notamment par le choix d'une problématique ainsi que de diverses hypothèses.

Tout d'abord, pour l'IH2EF¹⁹ une problématique est : « un processus qui permet de passer d'un thème de recherche général, ou d'une question de départ classiquement qualifiée de naïve à des hypothèses de travail précises. Elle procède par étapes successives d'exploration des possibles (cadre théorique, aspect de la question, hypothèses, etc.) et de choix d'options pour progressivement délimiter le travail à entreprendre ».

¹⁹ Institut des Hautes Etudes, de l'Education et de la Formation.

A partir recherches documentaires effectuées, nous avons remarqué que le badminton avait de grosses hausses d'effectifs depuis plusieurs dizaines d'années. Cependant il est possible de se demander si ce phénomène est commun à l'ensemble du territoire français.

Comment s'est spatialisé le badminton en France métropolitaine depuis 10 ans?

A la suite de la problématique, viennent donc les hypothèses de recherche. D'après le SPRER²⁰, une hypothèse est « une proposition que l'on veut défendre ou discuter. Cette proposition est la réponse à votre questionnement lors de l'élaboration de la problématique. L'hypothèse, pour être crédible, doit s'appuyer sur des faits réels et elle doit être vérifiable selon des données ».

Dans notre cas, quatre grandes hypothèses de travail permettent de structurer le document et répondre à la problématique qui a été posée :

- La pratique du badminton sur le territoire français est concentrée spatialement
- La pratique du badminton dépend des caractéristiques sociodémographiques des lieux où elle est exercée.
- Les caractéristiques sociodémographiques des lieux de pratiques du badminton ont évolué en dix ans.

Même si des études sur le badminton existent dans différents domaines, presque aucune étude géographique ne traite du badminton. Un sport qui comme vu précédemment a un fort potentiel de croissance et de développement.

A présent que le cadre théorique est posé, il est maintenant possible de passer à la partie méthodologique. Cette seconde partie permettra d'étudier toutes les procédures et les formules nécessaires à la réalisation des analyses.

²⁰ Service des Programmes, de la Réussite Educative et de la Recherche.

Conclusion de la première partie

Mêlant analyse théorique et recherche documentaire, la première partie a permis de mieux comprendre le contexte propre à l'étude du badminton en France.

Il fallait toutefois, avant d'aborder le badminton en tant que tel, présenter le sport dans sa définition, ses enjeux, sa géographie, ... et mettre ainsi en lumière plusieurs aspects comme :

- La construction d'une étude en géographie du sport
- Les caractéristiques sociodémographiques des sportifs
- L'importance du territoire.

C'est au travers de ces recherches que le badminton a été étudié ; son histoire, son rôle dans le monde, ainsi que ses enjeux en France.

Enfin, le cadre de travail a pu être mis en place et explicité. Il semble désormais nécessaire de poursuivre plus au cœur de l'étude. Pour ce faire, il est à présent temps de se consacrer à la méthodologie.

Partie II : Méthodologie, méthodes et outils d'interprétations

Introduction de la deuxième partie

La première partie de ce mémoire a permis d'aborder plusieurs thèmes notamment le sport dans la société, les recherches en géographie du sport ou encore l'« histoire » du badminton. A travers eux, l'importance sociodémographique des territoires a été identifiée.

Cette seconde partie consacrée à la méthodologie est un aspect inhérent à tout travail de recherche. La méthodologie se définit comme un « ensemble de règles et de démarches adoptées pour conduire une recherche »²¹. Elle abordera, dans un premier temps, les données ainsi que leurs structures. Puis dans un second temps, seront abordées des méthodes statistiques et des analyses spatiales. Enfin, il sera question de méthodes multivariées.

L'objectif global de cette seconde partie sera de présenter ainsi les différentes méthodes sollicitées pour répondre aux hypothèses de travail.

²¹ Source : Centre National de Ressources Textuelles et Lexicales (CNRTL).

Les données

Les données²² sont devenues un élément indispensable pour toute recherche en science sociale ou statistique. Mais les données ne sont pas seulement utiles dans ces domaines, elles sont aujourd'hui essentielles dans tous les secteurs et sont désormais au cœur de l'économie et de la société. En 2013 le journal *Le Monde* dans un article intitulé *Les données, puissance du futur*, compare les données à une matière première. Pour eux, les données sont semblables à des matières comme le charbon ou le fer et dont l'importance économique dépassera celle du pétrole.

Ce mémoire est réalisé avec la participation de la Fédération Française de Badminton. Avec notamment l'aide de David Courbet (responsable informatique à la FFBad), il m'a été possible d'accéder à certaines données issues des formulaires d'adhésion. Ce formulaire²³ permet l'obtention de la licence FFBad. Même si les informations demandées varient très peu, ce document est tout de même mis à jour à chaque nouvelle saison sportive. Les principaux renseignements sont :

- Les informations personnelles (nom, prénom, sexe, date de naissance, adresse, etc.)
- Les informations sportives (le nom du club, la Ligue, etc.)

Depuis le 25 mai 2018, le Règlement Général sur la Protection des Données (RGPD), voté par la Commission Européenne, stipule de nouvelles réglementations sur les données personnelles.

Le RGPD stipule notamment que :

- « Les données à caractère personnel doivent être traitées de manière licite et transparente, en garantissant la loyauté envers les personnes dont les données à caractère personnel sont traitées ;
- Il doit y avoir des finalités spécifiques pour traiter les données ;
- L'entreprise/organisation ne peut plus utiliser les données à caractère personnel pour d'autres finalités qui ne sont pas compatibles avec la finalité pour laquelle elles ont été initialement collectées ;

²² « Ensemble des indications enregistrées en machine pour permettre l'analyse et/ou la recherche automatique des informations » ; CNRTL.

²³ cf Annexe p.86

- L'entreprise/organisation doit s'assurer que les données à caractère personnel ne sont pas conservées plus longtemps que nécessaire ;
- L'entreprise/organisation doit mettre en place des mesures techniques et organisationnelles appropriées qui garantissent la sécurité des données personnelles »²⁴.

Or dans notre cas, la Fédération n'a pas demandé explicitement aux adhérents si leurs informations personnelles pouvaient être utilisées à des fins d'études. La Fédération n'a donc aucun droit de transmission sur ces données. Cependant, une partie des informations sportives était quant à elle accessible (Nom, prénom, sexe, date de naissance, club). La ville de résidence des joueurs n'étant pas accessible, ces derniers seront localisés en fonction de la ville de leur club.

La FFBad a procédé à une anonymisation de toute la base de données. Chaque joueur détient un numéro de licence qui lui est propre. De ce fait, même une personne ayant arrêté le badminton pendant 10 ans, retrouvera son ancien numéro de licence au moment de sa réinscription.

L'anonymisation (faites par la FFBad) a ici consisté à changer ce numéro de licence en un numéro aléatoire. Cette manipulation a pour but de rendre impossible le traçage d'un individu.

Ainsi, il m'a été possible de récupérer les informations suivantes pour chaque joueur pour chacune des différentes saisons :

- Un numéro d'anonymat
- Le sexe
- L'année de naissance
- Le nom du club
- La ville résidante du club
- Le code postal de la ville

²⁴ Source : Commission Européenne – Politique, Information et Service.

Ces données couvraient l'ensemble des pratiquants en France Métropolitaine sur la période 1996-1997 à 2017-2018 ; une base de données très dense (près de trois millions de lignes). Le choix a donc été fait de ne garder que les dix dernières années afin de réduire le nombre de données.

Plusieurs étapes ont été nécessaire afin d'étudier ces données :

- La mise forme des données (division des données par année)
- La correction de certaines erreurs (par exemple des clubs qui ne se trouvait pas dans la bonne commune)
- L'attribution d'un code INSEE à chaque commune. Le code INSEE est « un code numérique élaboré par l'Institut National de la Statistique et des Etudes Economiques, service public chargé de la production et de l'analyse des différentes données statistiques »²⁵. Ce dernier est unique pour chaque commune de France. Contrairement au code postal où un même code peut servir pour plusieurs communes. « Il existe un code postal pour chaque commune qui disposait en 1972 d'un bureau distributeur du courrier. Les communes n'ayant pas de tel bureau se sont vu attribuer les codes des bureaux distributeurs auxquelles elles étaient rattachées »ⁿ. Un code INSEE était donc nécessaire afin d'éviter des erreurs entre des communes ayant ou non des clubs de badminton ; un code postal servant à plusieurs communes, il pouvait engendrer une certaine ambiguïté.

Cependant dans ce type d'étude, une seule source de données est rarement suffisante.

En effet si l'on cherche à expliquer un phénomène, il est nécessaire d'avoir à disposition des données dites « exogènes »²⁶. Ce sont des données permettant d'ouvrir la compréhension de données dites « endogènes »²⁷.

²⁵ Source : Dictionnaire.Sensagent.fr.

²⁶ « Qui provient de l'extérieur, qui a une cause externe » ; CNRTL.

²⁷ « Qui provient de l'intérieur, qui une cause interne » ; CNRTL.

Dans notre cas, les données endogènes seront donc celles provenant de la FFBad. C'est-à-dire les informations concernant les joueurs et les clubs de badminton.

Les données exogènes retenues sont celles tirées d'études et d'articles scientifiques sur le sport en France. En effet, la littérature a permis d'aiguiller la sélection des variables exogènes. Ces dernières proviendront toutes de l'INSEE²⁸. Elles porteront sur quatre thèmes distincts :

- L'âge en 2010 et 2015
- Le diplôme en 2015
- Le revenu médian en 2015
- La catégorie sociale professionnelle en 2015

Comme précédemment, une étape de mise en forme des données a été nécessaire.

L'INSE met à disposition de nombreuses données issues d'enquêtes ou d'études. Celles qui nous intéressent sont les suivantes :

- Le nombre d'individus par classe d'âge, par commune en France²⁹
- Le nombre d'individus par le plus haut diplôme obtenu, par commune en France³⁰
- Le revenu médian par commune en France
- Le nombre d'individus par CSP³¹

Ces données sont fournies pour toutes les communes françaises. A partir de là, il est possible de calculer des taux pour chacune des variables. De plus, concernant le revenu annuel médian, il est donné également par commune. Il correspond à la médiane des revenus annuels médian des actifs d'une commune.

Enfin il reste un dernier type de données indispensables dans toutes études géographiques, ce sont les données géographiques (en format shp, ai, etc.). Elles permettent la spatialisation des données.

²⁸ Institut National de la Statistique et des Etudes Economiques.

²⁹ Liste des variables relatives à l'âge en annexe p...

³⁰ Liste des variables relatives au diplôme en annexe p...

³¹ Catégories Socio-Professionnelles ; Liste des catégories en annexe p...

Elles contiennent toutes les informations liées à la géométrie des objets représentés :

- Points
- Lignes
- Polygones

Ces données sont traitées principalement par des Systèmes d'Informations Géographiques (SIG).

Un SIG est « un logiciel informatique capable d'organiser et de présenter des données alphanumériques spatialement référencées. Il permet d'acquérir, d'organiser, de gérer, de traiter et de restituer des données géographiques sous forme de plans et cartes »³².

Dans notre cas, les données en format « shp » contiennent les limites départementales et communales de la France métropolitaine. Grâce à elles, il devient possible de cartographier les différentes données (endogènes et exogènes) présentées précédemment.

Analyses univariées

Les données utilisées dans tout travail de recherche sont généralement fournies de façon brute. Il est nécessaire par la suite de traiter ces données. La statistique est la branche des mathématiques consacrée au traitement de l'information. C'est un outil commun à toutes les sciences de l'observation.

En effet, les sciences sociales ne peuvent vérifier leurs hypothèses par des expériences. En statistique, il existe deux grandes méthodes : la statistique descriptive et la statistique inductive. La première permet de résumer l'information au moyen de graphiques, d'indicateurs, etc. La seconde quant à elle est plus utilisée dans l'interprétation des données.

³² Source : INGEO.fr

Si les statistiques ne sont relatives qu'à une seule variable, on parle de statistiques univariées. Dans le cas, où l'on se concentre sur deux variables simultanément, on parle alors de statistiques bivariées. Enfin si l'on souhaite se pencher sur plusieurs variables, on utilise alors de la statistique multivariée.

Il existe deux termes essentiels en statistique :

- La population : la population est l'ensemble des individus (ou unités statistiques) étudiés.
- Les variables : une variable est une information dont on recueille (ou observe ou mesure) la valeur pour chaque individu. C'est à partir de ces valeurs que vont être réalisées des statistiques.

On trouve différents indicateurs décrivant efficacement une série statistique :

- La moyenne : rapport entre la somme des valeurs d'un caractère et l'effectif total.
- La médiane : partage la série statistique en deux groupes d'effectifs égaux.
- Le minimum : plus petite valeur de la série.
- Le maximum : plus grande valeur de la série.
- L'écart-type : caractérise les écarts à la moyenne.
- La variance : est la moyenne des écarts à la moyenne au carré.

Différentes façons de représenter graphiquement une statistique univariée existent :

- Histogramme
- Boxplot (boite à moustache)
- Diagramme
- Cartes

Ces indicateurs et graphiques permettent de réaliser un grand nombre d'analyses en se concentrant sur une seule variable. Mais, ils interviennent également dans des analyses plus poussées sur plusieurs variables.

→ Pour mon étude, ces indicateurs et graphiques permettront de décrire nos données et de comparer différentes saisons.

Analyses spatiales

De nombreuses études en statistique appliquée présentent des méthodes statistiques destinées à des tableaux de données. Cependant, ces méthodes mathématiques ne permettent pas d'analyser les configurations spatiales.

En effet, l'analyse spatiale nécessite de prendre en compte l'espace comme caractère d'intérêt. Cette méthode d'analyse permet donc de lier la géographie et les statistiques. « L'analyse spatiale englobe l'ensemble des procédures quantitatives utilisées dans l'étude des localisations quelles que soient leur nature géométrique et les thématiques retenus. »³³

L'analyse spatiale repose sur trois grands objectifs :

- « Mettre en évidence les formes d'organisation spatiale
- Caractériser les processus de spatialisation qui sont à l'origine de ces structures (espacement, connexion, concentration, diffusion, etc.)
- Cerner les dynamiques des systèmes spatiaux. »³⁴

STATISTIQUES SPATIALES DESCRIPTIVES

La statistique spatiale descriptive permet de résumer la structure d'un semis de points et d'identifier une ou plusieurs localisations optimales. Elle s'organise notamment autour de deux indices :

- Le centre géométrique (qui est un indicateur de tendance centrale)
- Un indicateur de dispersion autour du centre.

Avant de parler d'indicateur, il a fallu construire un semis de points.

Un point géographique est « une position définie précisément par ses trois coordonnées de longitude, latitude et altitude. Un point est un objet abstrait de surface nulle. »³⁵

³³ C. Cauvin.

³⁴ T. Saint-Julien et C. Grasland, 2006.

³⁵ C. Grasland ; *Organisation de l'espace* ; Université Paris VII ; 2001.

Un semis de points est donc un ensemble de points définis par leurs coordonnées (X_i, Y_i) .

Dans notre cas, un point représentera un club de badminton et sera placé sur le centroïde de sa commune. Un centroïde peut être défini en tant que centre géométrique d'un polygone. Chaque point étant associé à un club, une commune possédant plusieurs clubs aura autant de points. Les points seront, ici, seulement définis par leur latitude et leur longitude. La création du semis de points aura été possible grâce à un SIG (le logiciel QGIS).

Les diverses analyses spatiales présentées dans ce document ont été réalisées à l'aide du logiciel CrimeStat. Il permet la réalisation d'un grand nombre de statistiques spatiales exploratoires, notamment les statistiques spatiales descriptives.

Pour décrire efficacement un semis de points, l'analyse spatiale propose trois indicateurs performants.

Le premier, le **centre moyen** (ou barycentre) est le premier paramètre descriptif d'une structure spatiale. C'est un indicateur de tendance centrale du semis à l'instar d'une moyenne arithmétique. Il peut être non pondéré ou pondéré. Dans notre cas, la pondération sera les effectifs de chaque club. Pour un caractère Z et d'effectif N , on définit un barycentre, noté G_Z , tel que :

$$G_Z = \min \sum n_j d_{ij}^2 \Leftrightarrow x_G = \frac{1}{N} \sum_i n_i x_i, \quad y_G = \frac{1}{N} \sum_i n_i y_i$$

Le second est la **distance standard**. Il représente l'écart-type d'une distribution spatiale. C'est donc une mesure de la dispersion absolue de la population autour de son barycentre. La distance standard s'exprime en distance :

$$\sigma_A = \overline{d_{GA}} = \sqrt{\frac{1}{N} \sum_i n_i d_{iG}^2}$$

Sources des formules : JM. Zaninetti ; *Statistique Spatiale, méthodes et applications géomatiques*.

Le troisième et dernier indicateur est l'ellipse de déviation standard³⁶. La distance standard définit un cercle autour du barycentre. Cependant dans le monde réel, la distribution d'un semis est généralement déformée dans une direction ou une autre. L'ellipse donne une indication indirecte de l'aplatissement de la distribution.

→ Toujours dans le cadre de notre étude, ces indicateurs permettront de décrire la spatialisation du badminton en France. Leur comparaison en fonction des saisons mettra en exergue, ou non, les directions prises au fil du temps.

HOTSPOT

Le HotSpot (ou point chaud) est un outil d'analyse spatiale utilisé pour mettre en exergue une concentration d'un phénomène ou d'une pratique. Il est notamment utilisé dans les recherches criminologiques.

Pour se faire, la méthode de classification spatiale hiérarchique des plus proches voisins est utilisée. Cet algorithme regroupe les points sur la base de leur proximité spatiale. Une distance de seuil entre deux points ainsi que le nombre minimal de points par classe sont ainsi définies préalablement. Le programme identifie des groupes d'ordre 1 dont les points sont plus rapprochés que la distance seuil et dans lesquels le nombre minimal de points est respecté.

Ici, la notion d'ordre 1 fera référence au plus proche voisin d'un point. Si on parle d'ordre 2, on désignera le deuxième plus proche voisin d'un point.

La classification est dite hiérarchique car les classes issues de l'ordre 1 vont être traitées de manière distincte. Pour une classification d'ordre 2, on reprendra tous les points séparément. Et ainsi de suite, si on augmente l'ordre de voisinage.

L'algorithme de classification va générer six résultats :

- L'ordre hiérarchique ainsi que le numéro du groupe
- Le centre moyen du groupe (avec ses coordonnées x et y)
- L'ellipse de déviation standard du groupe
- Le nombre de points dans le groupe
- La densité du groupe (nombre de points divisé par la surface du groupe).

³⁶ Formule en annexe p.. .

Ces résultats peuvent être exportés sur un SIG (type QGIS) pour être traités et visualisés.

→ L'objectif d'un tel algorithme est de mettre en lumière des zones de France où la pratique du badminton sera très concentrée et intense (en termes d'effectif de joueurs). On aura alors des petits groupes de communes où le badminton sera très pratiqué par rapport au reste du territoire.

Modèles bivariés et tests statistiques

Une grande partie des mathématiques appliquées consiste à la création de modèles statistiques. L'objectif de la modélisation statistique est donc d'étudier les propriétés d'un phénomène. Pour se faire, de nombreux modèles bivariés peuvent-être utilisés.

PYRAMIDE DES ÂGES

La pyramide des âges est une représentation graphique bivariée très courante en science sociale, notamment en démographie. Elle « représente la répartition par sexe et âge de la population à un instant donné. Elle est constituée de deux histogrammes, un pour chaque sexe, où les effectifs sont portés horizontalement et les âges verticalement. Les effectifs par sexe et âge dépendent de la fécondité, de la mortalité et des migrations. Mais la forme de la pyramide et les variations de celle-ci avec les années sont avant tout tributaires des variations de la fécondité. »³⁷ D'où l'équation du bilan démographique :

$$P_s = P_0 + (RA_{0,s} + NA_{0,s} - D_{0,s})$$

³⁷ Source: INSEE.fr.

Où :

- P_0 = Population au départ
- P_s = Population à un temps s
- $RA_{0,s}$ = Nombre de ré adhésions de licenciés entre 0 et s
- $NA_{0,s}$ = Nombre de nouvelles adhésions de licenciés entre 0 et s
- $D_{0,s}$ = Nombre de départs de licenciés entre 0 et s

→ Dans le cadre de la Fédération de badminton, on ne parlera pas de fécondité ou de mortalité. On évoquera la population sportive plutôt en termes d'entrée et de sorties. Les fluctuations d'une telle population se rapprochant d'un effet migratoire.

La pyramide des âges représente alors les effectifs suivant le sexe et l'âge des adhérents.

TAUX DE PÉNÉTRATION

Observer un effectif, par commune par exemple, permet de visualiser les concentrations importantes de personnes. Cependant, une analyse univariée d'un effectif peut présenter un certain biais. Par exemple, un effectif important de badistes dans une commune peut être simplement lié à une population importante dans cette commune. Pour éviter ce problème, il est possible de construire un ou des indicateurs. Dans le cadre de l'étude, cet indicateur sera le taux de pénétration.

Un taux de pénétration est « le rapport entre le nombre d'utilisateurs d'un produit ou d'un service et la population cible tout entière. »³⁸ Ici, le « produit » sera l'adhésion prise à la FFBad.

De ce fait, le taux de pénétration se calcule comme-ci :

$$\text{Taux de pénétration} = \frac{\text{Nombre de badiste d'un espace}}{\text{Population totale de l'espace}}$$

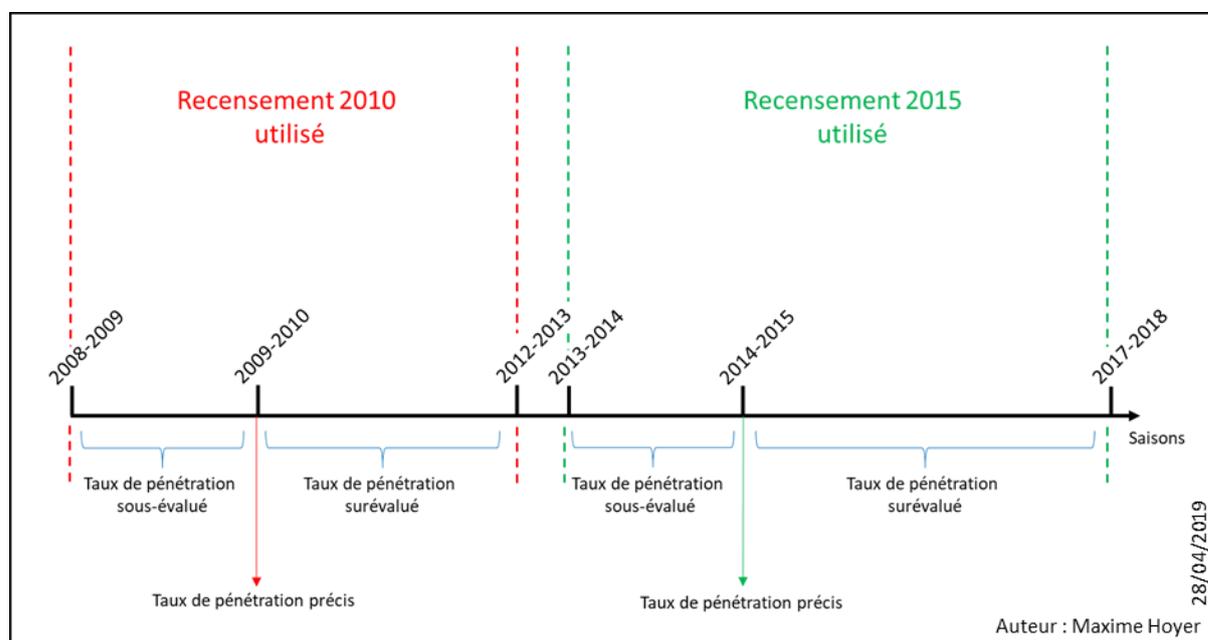
Grâce à cet indicateur il devient possible de comparer deux communes avec des populations différentes.

³⁸ Source: Marketing-Etudiant.fr.

Le calcul du taux de pénétration pour les dix années de notre période nécessite de connaître le nombre d'habitants par commune pour chaque année. Or, le recensement de la population n'est effectué que tous les cinq ans. Deux recensements tombent pendant la période 2008-2018. Il s'agit du recensement de 2010 ainsi que celui de 2015. Ce sont ces deux recensements qui vont être utilisés pour le calcul du taux de pénétration sur toute la période 2008-2018. Cependant, certains taux seront calculés sur des années différentes. Par exemple le taux de pénétration de la ville de Marseille pour la saison 2017-2018 sera :

$$Tx \text{ de pénétration}_{Marseille_{2017-2018}} = \frac{\text{Nombre de licenciés FFBad}_{Marseille_{2017-2018}}}{\text{Pop totale}_{Marseille_{2015}}}$$

Sous le postulat que la population en France a tendance à augmenter, le taux calculé pour la commune de Marseille peut être légèrement surestimé. En effet, on peut supposer que la population de Marseille est légèrement plus importante en 2018 qu'en 2015. De ce fait, le taux de pénétration peut être surévalué. Il est donc possible de représenter ce phénomène



Limite du taux de pénétration

Même si ce phénomène n'a que potentiellement peu d'impact, il est important de le prendre en compte lors des interprétations.

REGRESSION LINÉAIRE SIMPLE

La régression linéaire simple est une modélisation linéaire. C'est une méthode statistique utilisée pour modéliser une variable y en fonction d'une variable x , et ainsi faire des estimations et des prédictions. Elle peut être juste une approche exploratoire mais elle permet également de répondre à des questions : la variable X a-t-elle une relation linéaire avec la variable Y . En d'autres termes, la valeur de Y augmente-t-elle si la valeur de X augmente ?

Cette méthode aboutie à ce que l'on appelle une droite de régression. Cette dernière décrit le modèle par une droite et donc par l'équation de cette droite.

« Modèle de régression linéaire simple :

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Où :

- Y est la variable à expliquer (variable endogène).
- X est la variable explicative (variable exogène).

Ce modèle de régression suit des hypothèses :

- i. La distribution de l'erreur ε est indépendante de X ;
- ii. $\forall i = 1, \dots, n$, l'erreur ε est suit une loi normale : $E(\varepsilon_i) = 0$, $Var(\varepsilon_i) = \sigma^2$;
- iii. β_0 et β_1 sont constants. »³⁹

Si le modèle ne respecte pas l'une de ces hypothèses, il n'est pas valide.

→ Une série de régression linéaire simple a été réalisée dans ce document. L'objectif étant de mettre en évidence, ou non, une relation linéaire entre notre variable endogène (le taux de pénétration par commune) et les variables exogènes (âge, diplôme, CSP, niveau de vie).

³⁹ Source: Univ-Toulouse.fr.

ANOVA

L'ANOVA (*Analysis of Variance*) est un des modèles statistiques les plus utilisés en sciences sociales. Cette analyse de variance, dite paramétrique, permet de tester la relation entre une variable endogène continue⁴⁰ et une ou des variables exogènes nominales⁴¹. En d'autres termes, une ANOVA vise à tester les différences significatives des moyennes de diverses catégories. Pour qu'une ANOVA soit valide, les données utilisées doivent respecter deux conditions :

- i. La normalité des résidus
- ii. L'homoscédasticité des résidus

Si ces conditions sont respectées, il est alors possible de réaliser une ANOVA.

Dans un premier temps, il est important de fixer les hypothèses avant de les tester. Un test d'hypothèse permet d'accepter ou de rejeter une affirmation faite sur notre population. Il examine deux hypothèses concernant une population : l'hypothèse nulle (H_0) et l'hypothèse alternative (H_1). Généralement, l'hypothèse nulle spécifie l'absence d'effet ou l'absence de différence. C'est cette hypothèse que l'on teste le plus souvent. L'hypothèse alternative est celle qui spécifie l'effet que l'on souhaite trouver.

Dans le cas d'une ANOVA, les hypothèses sont les suivantes :

- H_0 : Toutes les moyennes sont égales
- H_1 : Au moins deux moyennes sont différentes

Il existe trois grandes étapes dans la réalisation d'une ANOVA :

- i. Calculer des variances intra-classes (variances à l'intérieur de chaque classes) et des variances inter-classes (variances entre les classes).
- ii. Calcul du rapport $F = \text{variance inter} / \text{variance intra}$
- iii. Test de significativité du modèle (test de Fisher)

Grâce au test de Fisher, il est possible de conclure de la validité d'une des hypothèses, soit H_0 ou soit H_1 . Si la p-valeur du test de Fisher est supérieure au seuil de significativité 5%, alors on accepte l'hypothèse nulle. À l'inverse si la p-valeur est inférieur à 5%, alors on rejette l'hypothèse nulle et conclut qu'au moins deux moyennes sont significativement différentes.

⁴⁰ « Variable numériques ayant un nombre infini de valeurs dénombrables entre deux valeurs » ; Minitab.com

⁴¹ « Variables présentant des catégories nommées par un nom » ; Minitab.com

→ Dans notre cas, l'ANOVA aura pour but de tester la relation entre des groupes de communes (créer ultérieurement) et leurs taux de pénétration. En d'autres termes, un ou des groupes ont-ils des taux de pénétration significativement différents des autres ?

TEST DE FISHER

En statistique, le test de Fisher désigne tout test dans lequel la statistique de test suit une loi de Fisher⁴². Ce type de test est généralement utilisé pour la comparaison de modèles statistiques. Comme pour n'importe quel test statistique, il faut établir les hypothèses H_0 et H_1 .

Le principe du test est de comparer le F (variance inter-classe / variance intra-classe) observé au F théorique suivant une loi de Fisher. On cherche donc à savoir si les résultats observés sont ceux attendus théoriquement.

Si le $F_{observé}$ est plus grand que le $F_{théorique}$, alors on rejette l'hypothèse nulle H_0 .

Si le $F_{observé}$ est plus petit que le $F_{théorique}$, alors on accepte l'hypothèse nulle H_0 .

→ Le F du taux de pénétration des communes de chaque groupe est-il plus grand ou plus petit que le $F_{théorique}$ (suivant une loi de Fisher) ?

TEST DE SHAPIRO-WILK

En statistiques, l'étude des résidus est primordial dans de nombreuses situations. Dans un repère (x ; y) on dispose des données numériques. Chaque point sera défini par ses coordonnées x et y. L'ensemble de ces points forme un nuage de points. Ensuite, une droite passant par le plus de points possibles est tracée. Les résidus vont correspondre à l'écart entre les points et la droite. Plus les résidus sont proches de zéro et plus les points seront proches de la droite. Les résidus sont donc un indice mesurant la qualité d'ajustement de la droite. Seules des données avec des résidus suivant une loi normale peuvent faire l'objet d'une analyse de variance paramétrique (type ANOVA).

→ Les résidus des taux de pénétration des communes de chaque groupe suivent-ils une loi normale ?

⁴² cf Annexe p.88

La significativité de ce test est donc une condition nécessaire à la réalisation d'une ANOVA.

Hypothèses du test :

- H_0 : Les résidus suivent une loi normale
- H_1 : Les résidus ne suivent pas une loi normale

Plus la statistique de test de Shapiro-Wilk⁴³ est élevée, et plus la compatibilité avec la loi normale sera vraisemblable. En d'autres termes, si la p-valeur est supérieure au seuil de significativité 5%, alors on accepte H_0 . De ce fait, les résidus suivent une loi normale.

Dans le cadre de la validité d'une ANOVA, on cherche à accepter l'hypothèse nulle H_0 .

TEST DE LEVENE

Le test de Levene⁴⁴ comme le test de Bartlett permet de tester l'homoscédasticité des résidus. Lorsque les résidus des données sont homogènes, on parle alors d'homoscédasticité. En d'autres termes, les résidus correspondent bien à un aléa. Dans le cas contraire, on parle d'hétéroscédasticité. Dans le cadre d'une ANOVA, on cherche l'homoscédasticité des résidus. Le test de Levene permet donc de tester cette homoscédasticité dans le cas où les résidus ne suivent pas une loi normale. Contrairement au test de Bartlett qui impose la normalité des résidus.

→ Les résidus des taux de pénétration des communes de chaque groupe sont-ils homogènes ?

Hypothèses du test :

- H_0 : Homoscédasticité des résidus
- H_1 : Hétéroscédasticité des résidus

Plus la statistique de test de Levene est élevée, et plus l'homoscédasticité des résidus sera vraisemblable. En d'autres termes, si la p-valeur est supérieure au seuil de significativité 5%, alors on accepte H_0 . De ce fait, les résidus sont homogènes. Dans le cadre de la validité d'une ANOVA, on cherche à accepter l'hypothèse nulle H_0 .

⁴³ cf Annexe p.89

⁴⁴ cf Annexe p.90

TEST DE KRUSKAL-WALLIS

Le test de Kruskal-Wallis⁴⁵ est un test semblable à l'ANOVA. Il est d'ailleurs utilisé comme substitut à l'analyse de variance. Quand les hypothèses de validité d'une ANOVA (normalité et homoscedasticité des résidus) ne sont pas respectées, il est possible d'utiliser ce test non-paramétrique.

→ Un ou des groupes de communes ont-ils des taux de pénétration significativement différents des autres ?

Si une ANOVA s'intéresse aux moyennes, le test de Kruskal-Wallis quant à lui se concentre sur les médianes. Il vise donc à tester les différences significatives entre les médianes de diverses catégories. D'où les hypothèses :

- H_0 : Toutes les médianes sont égales
- H_1 : Au moins deux médianes sont différentes

Si la p-valeur du test est supérieure au seuil de significativité 5%, alors on accepte l'hypothèse nulle. À l'inverse si la p-valeur est inférieur à 5%, alors on rejette l'hypothèse nulle et conclut qu'au moins deux médianes sont significativement différentes.

⁴⁵ cf Annexe p.90

Méthodes multivariées

ANALYSE EN COMPOSANTE PRINCIPALE

« L'analyse en composante principale (ACP) permet d'analyser et de visualiser un jeu de données contenant des individus décrits par plusieurs variables.

C'est une méthode statistique qui permet d'explorer des données dites multivariées (données avec plusieurs variables).

Elle synthétise l'information de plusieurs variables en seulement quelques nouvelles, appelées composantes principales (ou axes factoriels). Ces nouvelles variables correspondent à une combinaison linéaire des variables originelles. Le nombre de composantes principales est inférieur ou égal au nombre de variables d'origine »⁴⁶.

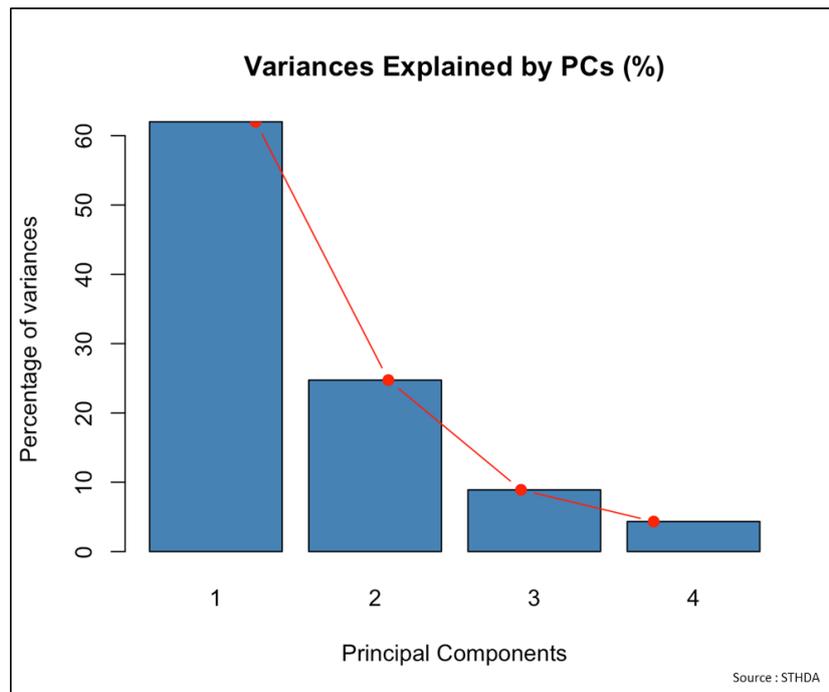
L'information importante contenue dans un jeu de données correspond à la variance (ou inertie totale⁴⁷) qu'il contient. L'objectif d'une ACP est donc d'identifier les directions (c'est-à-dire les composantes principales) où la variation des données maximales. En d'autres termes, l'ACP réduit les dimensions d'un jeu de données multivariés à deux ou trois composantes principales.

→ L'ACP portera sur l'ensemble des variables exogènes (31 variables au total) pour les communes possédant au moins un club de badminton. L'objectif étant de résumer l'information contenue dans ces 31 variables. Ce qui permettra par la suite la création de groupes de communes semblables statistiquement.

L'étape la plus importante d'une ACP est celle de la sélection du nombre d'axes factoriels. Cette décision se fait à partir de la variance expliquée. On cherche à prendre le minimum d'axes factoriels tout en maximisant notre variance expliquée.

⁴⁶ Source: Statistical Tools for High-Throughput Data Analysis.com (STHDA)

⁴⁷ Mesure la dispersion totale d'un nuage de points. L'inertie est donc égale à la somme des variances des variables étudiées.



Exemple de variance expliquée par les axes factoriels

Cet exemple présente la variance expliquée par 4 axes factoriels. Ici, seulement les 2 premiers axes seront retenus pour l'interprétation car ils expliquent à eux seuls environ 85% de l'information. Prendre plus d'axes n'augmenterait pas de façon significative la part de variance expliquée mais serait contraignant pour l'interprétation.

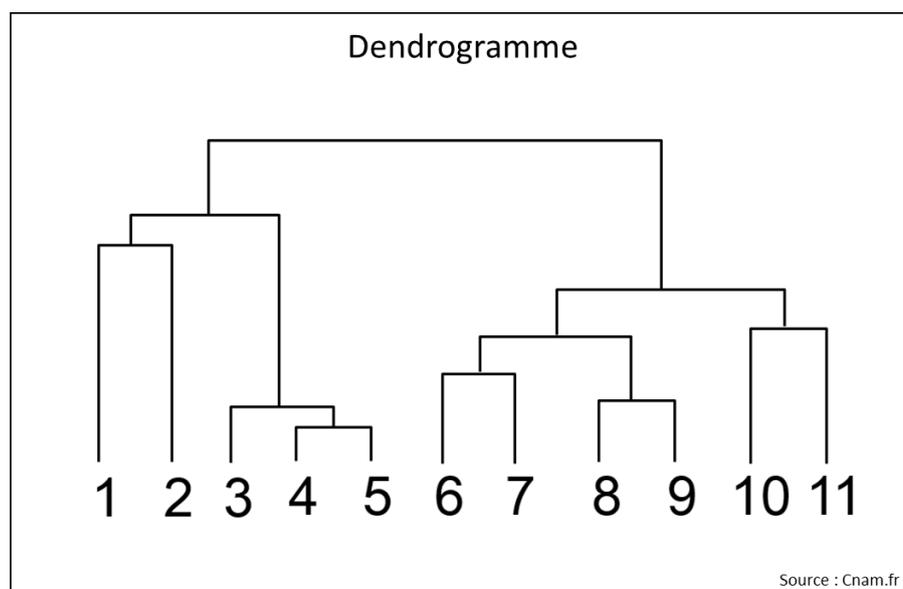
Ensuite vient l'interprétation de chaque axe. Comme expliqué précédemment, le principe d'une ACP est de résumer l'informations. De ce fait, de « n » variables on peut obtenir « n-1 » axes factoriels. Il est donc primordial de comprendre chaque axe sélectionné et trouver l'information qu'il résume. Enfin, il devient possible d'interpréter les caractéristiques des individus en fonction de leur position sur chaque axe factoriel.

Néanmoins, il est difficile de caractériser tous les individus d'un jeu de données sur 1, 2 voire 3 axes factoriels ; surtout si on a plusieurs centaines d'individus. Il faut donc utiliser une méthode pour classer nos individus en fonction de leur proximité sur les axes factoriels.

CLASSIFICATION ASCENDANTE HIÉRARCHIQUE

En statistique, il existe plusieurs techniques permettant de découper une population en classe. La classification ascendante hiérarchique (CAH) est la plus courante. Le but d'une CAH est donc de regrouper au sein d'une même classe, les individus les plus semblables (homogénéité intra-classe) tout en dissociant les classes au maximum (homogénéité inter-classe).

Le principe de la CAH est de rassembler des individus selon un critère de ressemblance (dans notre cas il est défini par l'ACP). Ce critère s'exprimera sous la forme d'une matrice de distance (tableau de m lignes et n colonnes). Elle permet d'exposer la distance entre chaque individu pris deux à deux. Plus deux individus seront dissemblables et plus ils seront éloignés l'un de l'autre (on parle de distance statistique). Ensuite, la CAH va rassembler les individus de manière itérative. La classification est dite ascendante car elle part des individus. Au départ chaque individu correspondra à une classe, puis au fur et à mesure des itérations, le nombre de classe diminuera et le nombre d'individus par classe lui augmentera. Il est possible de visualiser ce processus grâce à un dendrogramme.

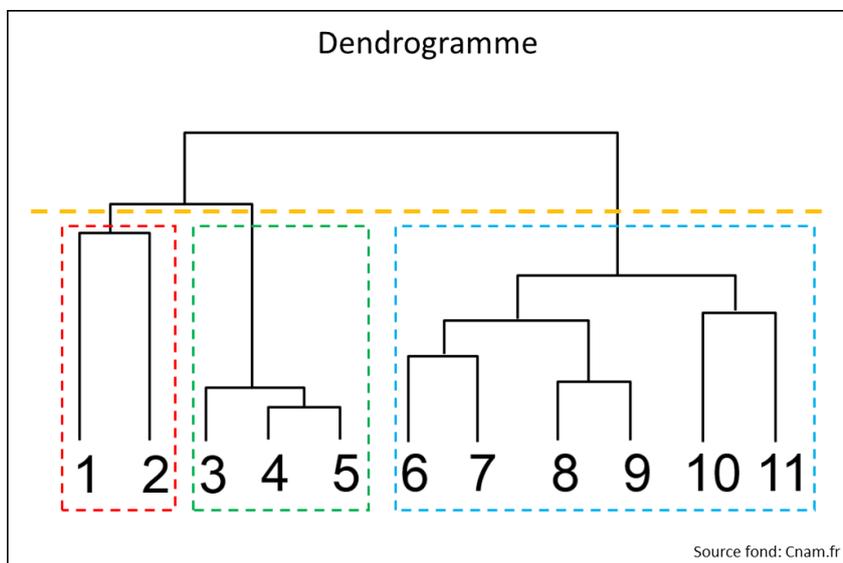


Exemple de dendrogramme

Cet exemple présente onze individus dans un dendrogramme représentant une partition d'une CAH. Au départ, on retrouve une partition en onze classe (une pour chaque individu). Mais au fur et à mesure des itérations, le nombre de classe diminue pour arriver à une classe unique.

Ensuite, il est nécessaire de sélectionner le bon nombre de classes pour la suite de la CAH. Pour cela, il faut retenir la partition qui minimise le plus la variance intra-classe et qui maximise le

plus la variance inter-classe. En d'autres termes, on cherche la partition où les individus dans les classes sont les plus semblables et les classes les plus différentes entre elles. Cette partition peut être également visualiser sur un dendrogramme.



Exemple de classification

Avec notre exemple précédent, une partition qui paraît satisfaisante serait celle ci-dessus. On a donc une partition en trois classes :

- La classe rouge (les individus 1 et 2)
- La classe verte (les individus 3, 4 et 5)
- La classe bleue (les individus 6 à 11)

La dernière étape de la CAH, comme pour de nombreuses analyses statistiques, est l'interprétation. Il faut comprendre les caractéristiques des classes grâce à notre connaissance des axes factoriels issus de l'ACP.

→ La CAH utilisée pour notre étude sera le prolongement de l'ACP. Le but de cette classification est donc de créer des groupes de communes semblables statistiquement. En d'autres termes, la CAH permettra de faire émerger, par exemple :

- Des communes avec un fort taux de jeunes par rapport aux autres communes
- Des villes avec un fort taux de cadres par rapport aux autres villes

Conclusion de la deuxième partie

Cette seconde partie a apporté des éléments de compréhension essentiels à tout travail scientifique. La partie méthodologique est la section qui doit apporter un éclairage sur les éléments qui seront employés ultérieurement. Elle va également permettre d'amorcer les réponses aux questionnements émis dans la première partie.

Ce chapitre a donc passé en revue divers aspects statistiques et géographiques tels que :

- Les analyses univariées
- Les analyses spatiales
- Les analyses bivariées
- Les analyses multivariées.

A présent, il est temps d'utiliser cette méthodologie à notre sujet et d'appliquer ces méthodes à nos données. Ce sera l'objectif de la troisième partie.

Partie III : Le badminton, structures sociales et territoriales

Introduction de la troisième partie

Une première partie avait permis d'identifier les enjeux du badminton en France et de donner un aperçu de variables sociodémographiques pertinentes, dans l'étude des populations sportives. Enfin, ce chapitre présentait les hypothèses de recherches relatives au sujet d'étude :

- La pratique du badminton sur le territoire français est concentrée spatialement
- La pratique du badminton dépend des caractéristiques sociodémographiques des lieux où elle est exercée.
- Les caractéristiques sociodémographiques des lieux de pratique du badminton ont évolué en dix ans.

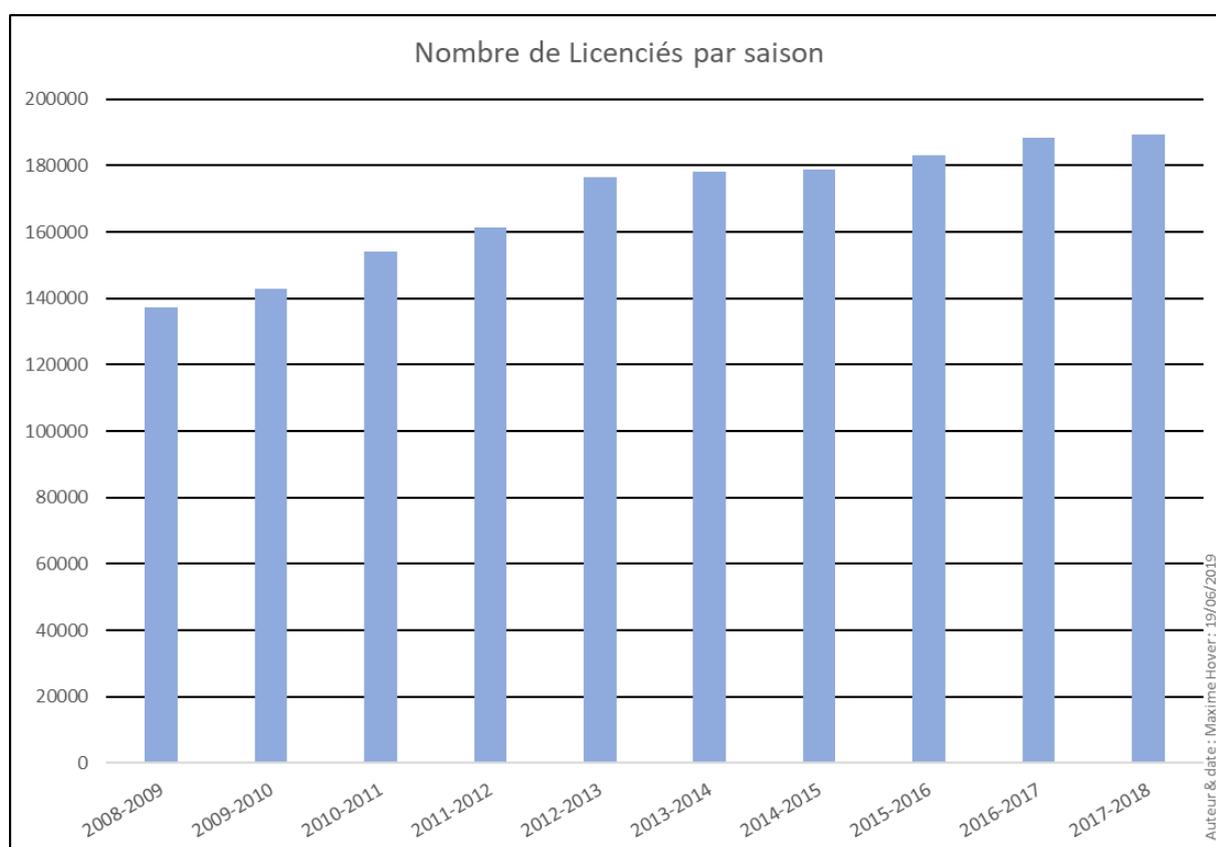
La deuxième partie présentait quant à elle les différentes méthodes et théories applicables à l'étude et aux réponses des hypothèses.

Cette troisième et dernière partie appliquera la méthodologie expliquée précédemment et conclura sur la validité ou non de ces hypothèses.

Structure et répartition spatiale des badistes français

STRUCTURE DE LA POPULATION

Le sport est devenu en 50 ans une pratique très courante. Elle permet le développement de soi, autant physiquement que socialement. De ce fait, le sport touche une grande partie de la population, quel que soit le sexe, l'âge, etc. Il n'est donc pas étonnant de trouver des fédérations sportives comme le football avec plus de deux millions d'adhérents. Une population relative à une Fédération (sportive par exemple) peut par conséquent être étudiée par le prisme démographique à l'image de la population d'un pays. Comme étudié précédemment, les notions de natalité et de mortalité n'interviennent pas dans l'équation du bilan démographique. Le seul facteur important concerne les migrations (apparentées à des entrées et des sorties). Une population sportive peut donc être étudiée avec les outils démographiques. Une première analyse porte donc sur la structure de notre population.

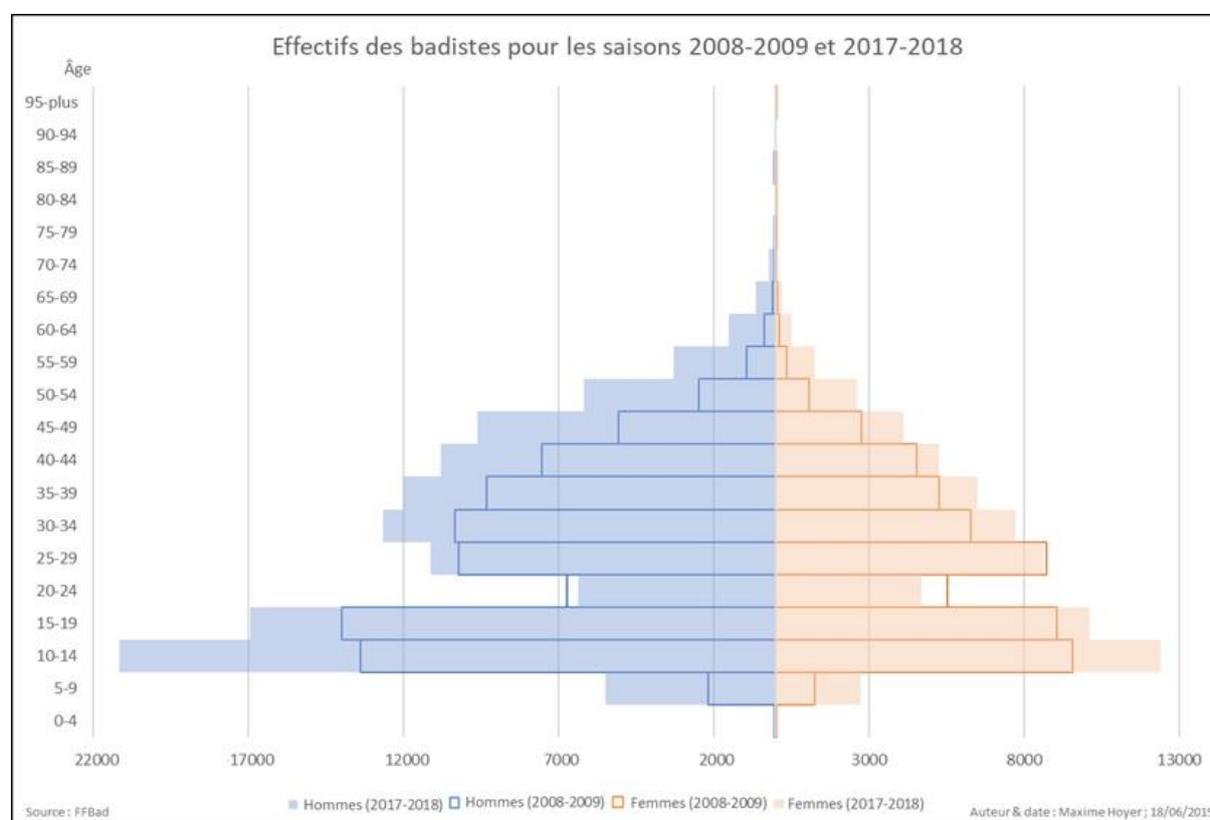


Nombre de licenciés par saison

Cet histogramme présente le nombre de licenciés à la FFBad pour les saisons de 2008-2009 à 2017-2018. Si ce graphique apporte peu d'informations, il est néanmoins intéressant car il met en évidence la constante augmentation du nombre d'adhérent au cours de ces années. On observe une augmentation de presque 40% (37,8%) de licenciés entre 2008-2009 et 2017-2018.

A l'origine, les données fournies portaient sur dix saisons (de la saison 2008-2009 à 2017-2018). Pour des raisons de simplification, seules la première et la dernière saison seront analysées en détail (la saison 2008-2009 et la saison 2017-2018).

Un des outils principaux en démographie est la pyramide des âges. Elle permet d'obtenir une bonne représentation de notre population selon l'âge et le sexe.



Effectifs des badistes pour les saisons 2008-2009 et 2017-2018

Ce graphique permet de mettre en lumière plusieurs phénomènes.

1. Le premier élément est la structure de la pyramide. En effet, elle est sensiblement identique entre les deux saisons. La structure de la population évolue donc peu entre la saison 2008-2009 et la saison 2017-2018.

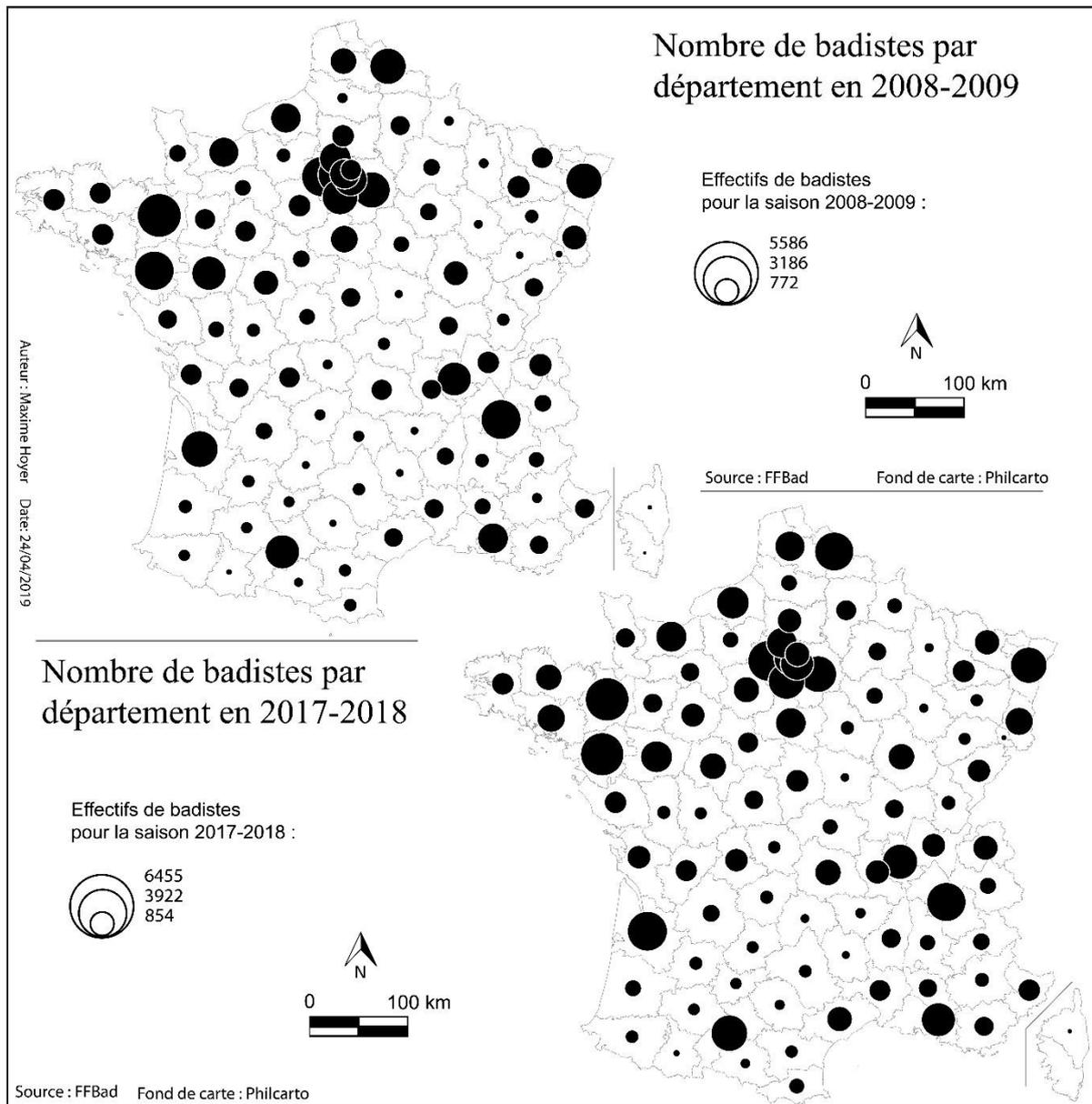
2. Le second phénomène est la place importante des femmes. Même si les hommes restent majoritaires, le nombre de femmes est élevé (environ 40%) rapport à de nombreux autres sports (5% dans le football, 29% dans le tennis, ...⁴⁸). Ce taux s'explique par la mixité de ce sport. En effet, le badminton, à l'instar des autres sports de raquette, est un des rares sports où des hommes peuvent affronter des femmes (en double mixte, un homme et une femme affrontent un homme et une femme). À noter que le double mixte est beaucoup plus présent dans le badminton qu'au tennis ou au tennis de table. Enfin, la pratique plus masculine du badminton ne fait que confirmer les éléments mis en avant par les études déjà menées sur les populations sportives.
3. Le troisième phénomène mis en lumière concerne l'âge des adhérents. En effet, on peut noter la grande hétérogénéité des classes d'âge. Les plus jeunes joueurs ont 4 ans (très ludique, le badminton peut être commencé à partir de cet âge) et les plus âgés ont plus de 95 ans. Par ailleurs les classes d'âges les plus représentées sont les 10-14 ans et les 15-19 ans (classes d'âges les plus sportives), que ce soit chez les hommes ou les femmes. Mais cela n'empêche pas le badminton d'attirer des personnes plus âgées. Le graphique met en avant des effectifs importants pour les hommes comme pour les femmes entre 25 et 49 ans. C'est un sport très bénéfique pour la santé car il sollicite tous les muscles, ainsi que l'équilibre, les réflexes ou encore la coordination. Ces bienfaits expliquent la pratique du badminton par ces classes d'âges. Des programmes pilotes menés par la FFBad dans des EHPAD ont mis en avant les bénéfices du badminton sur l'équilibre et la coordination motrice des personnes âgées. Raison pour laquelle on retrouve près d'une centaine de badistes de plus de 70 ans. A contrario on observe la très faible représentation des 20-24 ans, présentée comme une catégorie d'âge très sportive. Il semblerait que cette catégorie pratique plus le sport hors fédération ou soit plus attirée par d'autres disciplines.

Ce premier graphique permet donc de conforter certaines informations issues de la littérature (le sport est pratiqué en majorité par des hommes jeunes). Cependant, cette pyramide des âges apporte également de la nuance concernant ces indications. En effet, la littérature présentait les 15-24 ans comme la population la plus sportive (dans et hors Fédérations sportives). Pourtant, les catégories les plus représentées dans la FFBad se trouvent être les 10-14 ans et les 15-19 ans, très loin devant les 20-24 ans.

⁴⁸ Ministère des Sports ; *Les Chiffres clés du sport 2017* ; 2017.

LICENCIES PAR DÉPARTEMENT

Après avoir analysé la population de badistes français dans son ensemble, nous étudierons comment elle se spatialise. Pour cela ont été réalisées deux cartes ; une pour la saison 2008-2009 et une pour la saison 2017-2018. Elles présentent toutes deux le nombre de licenciés par département, représenté par des cercles proportionnels.

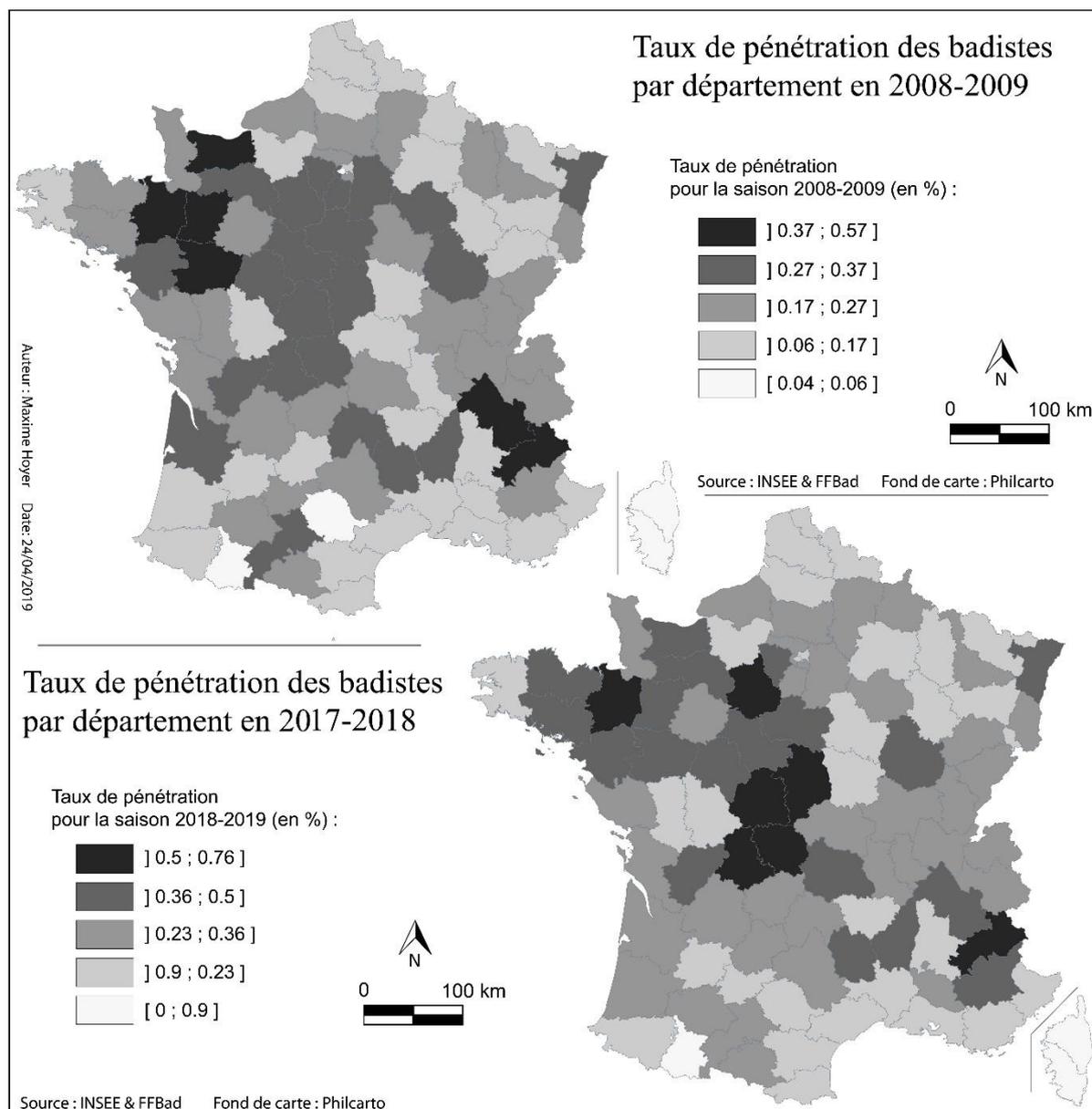


Nombre de badistes par département

Ces deux cartes présentent des structures spatiales quasi-identiques. En effet, les départements ayant le plus de joueurs ont généralement la population la plus importante (comme les départements d'Île-de-France, la Gironde, ou encore le Rhône). Cependant certains

départements un peu moins peuplés comme l'Isère ou l'Ille et Vilaine ont également de nombreux badistes.

Dans la grande majorité, le nombre de joueurs dépend en grande partie de la population de son département. Se concentrer seulement sur les effectifs ne suffit donc pas. Il faut aller plus loin et étudier directement le nombre de joueurs en fonction du nombre d'habitants. Pour cela, il est nécessaire d'utiliser un indicateur créé spécialement pour cette analyse, le taux de pénétration.



Taux de pénétration par département

Cette carte met en évidence les forts taux de pénétration dans les régions Centre, Normandie, Val de Loire, Bretagne et Rhône-Alpes en 2008-2009. On remarque que le taux de pénétration

de joueurs dans ces départements a augmenté en 2017-2018. La part de joueurs s'intensifie donc dans ces zones.

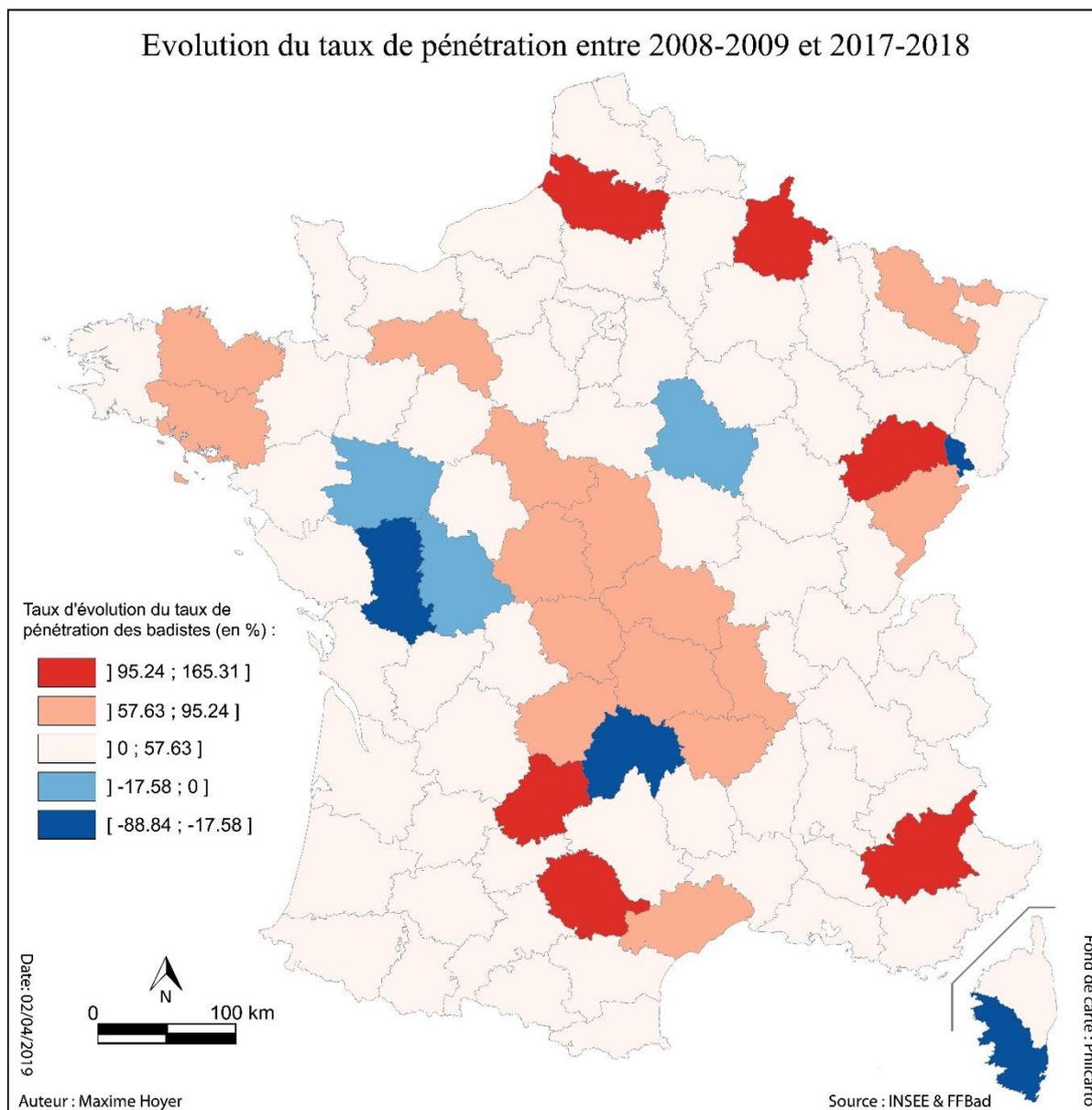
A l'inverse, on peut noter de faibles taux de pénétration dans le Nord-Pas-de-Calais, la Champagne-Ardenne, et sur toute la côte Méditerranéenne (dont la Corse). Des zones où il semble y avoir peu d'évolution en dix ans.

Finalement, ces cartes mettent en évidence une large zone (allant du Calvados à la Creuse en passant par la région Parisienne) avec des taux de pénétration plus élevés que sur le reste du territoire français. La pratique importante du badminton sur ces territoires peut paraître surprenante. Mais elle peut être en lien avec l'histoire du badminton en France. En effet comme l'explique Julie Grall dans sa thèse, « dans les années 70 l'essentiel de l'activité du badminton en France demeure partagée entre la Normandie et l'Île-de-France ». Néanmoins, on remarque que ce taux a quelque peu évolué en dix ans dans de nombreux départements. Pour mieux visualiser ce phénomène, il est nécessaire d'étudier le taux d'évolution.

Un taux d'évolution se calcul de la manière suivante :

$$\left(\frac{\text{valeur d'arrivée} - \text{valeur de départ}}{\text{valeur de départ}} \right) * 100$$

Le taux d'évolution du taux de pénétration, entre la saison 2008-2009 et la saison 2017-2018, a donc été calculé par département. Après son calcul, il a été cartographié pour identifier les zones qui ont connu des augmentations et celles qui ont connu des diminutions.



Evolution du taux de pénétration entre 2008-2009 et 2017-2018

Comme étudié précédemment, on remarque que la région Auvergne et une partie du Centre et du Limousin ont vu leur taux de pénétration augmenter en 10 ans. Par ailleurs, la majorité des départements de France métropolitains a vu son taux de pénétration soit stagner, soit augmenter de 50% maximum. Malgré cette générale augmentation, on remarque des diminutions du taux de pénétration comme dans les Deux-Sèvres, le Cantal ou encore la Corse du Sud. Les premières cartes présentaient des effectifs de joueurs et des taux de pénétration très faibles dans ces départements. La perte d'un club, par exemple, entre 2008-2009 et 2017-2018 peut fortement diminuer leur taux de pénétration. Cependant, ce genre d'information n'est pas facilement identifiable à l'échelle départementale.

Même si une étude départementale présente des avantages, limiter notre analyse à cette échelle ne serait pas pertinent. Comme vu précédemment, une étude départementale induit une grande perte d'informations. A cette échelle, il n'est pas possible de comprendre si, par exemple, les joueurs d'un département sont concentrés dans une seule commune ou répartis sur l'ensemble des communes.

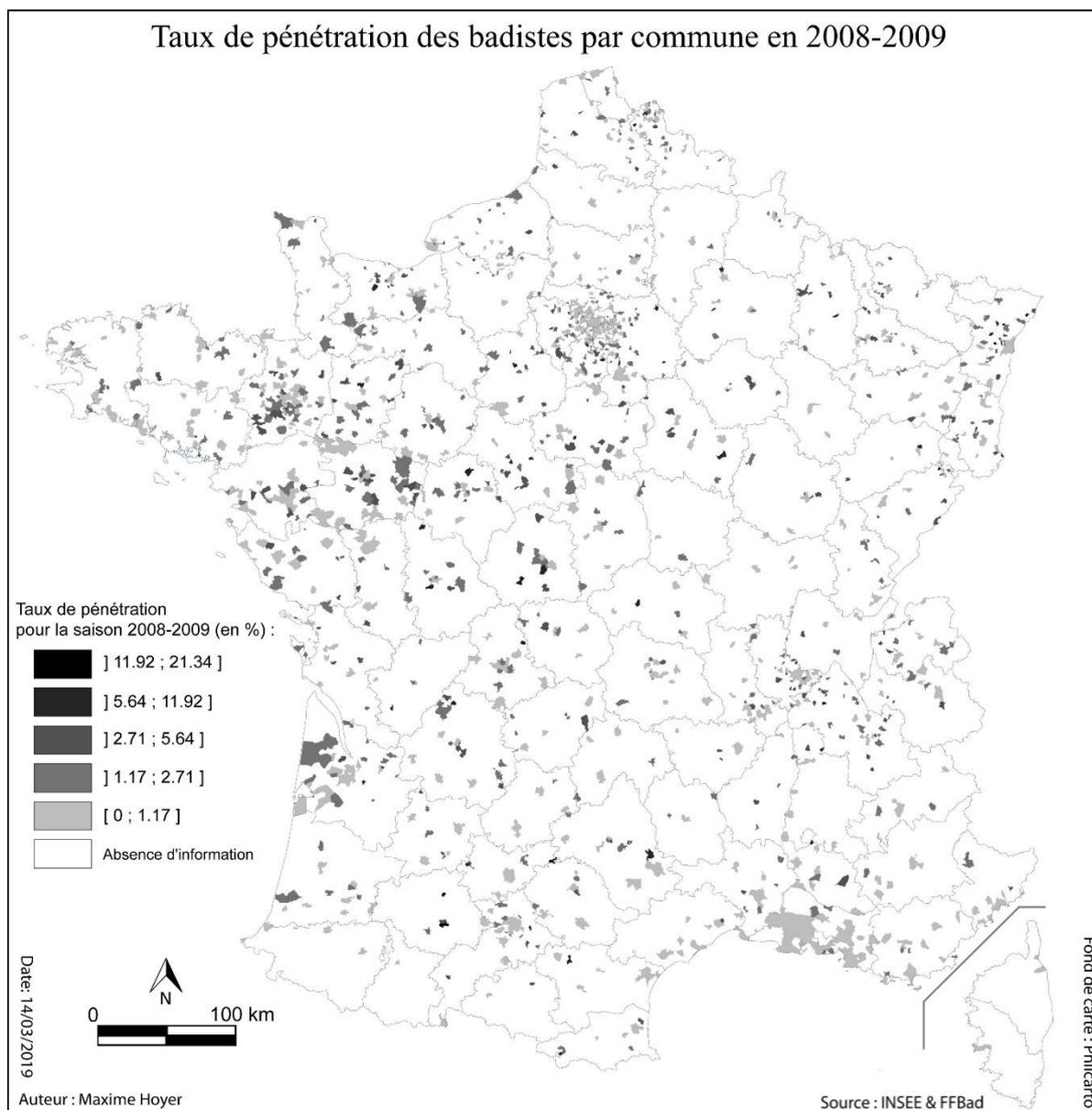
Il est donc nécessaire après cette analyse globale, de réaliser une étude communale pour obtenir une analyse plus fine.

LICENCIÉS PAR COMMUNE

Il est à présent temps d'étudier notre population de badiste à l'échelle communale. Pour rappel, les licenciés sont rattachés à la commune de leur club et non pas à leur lieu de résidence (ces informations n'étant pas accessibles).

Des cartes présentant le nombre de joueurs par commune (en cercle proportionnel) ont été réalisées pour les saisons 2008-2009 et 2017-2018 ; mais peu exploitables, ces cartes n'ont pas été retenues dans mon analyse.

L'indicateur du taux de pénétration, plus pertinent, a donc été retenu et calculé sur l'ensemble des communes françaises pour les saisons 2008-2009 et 2017-2018.



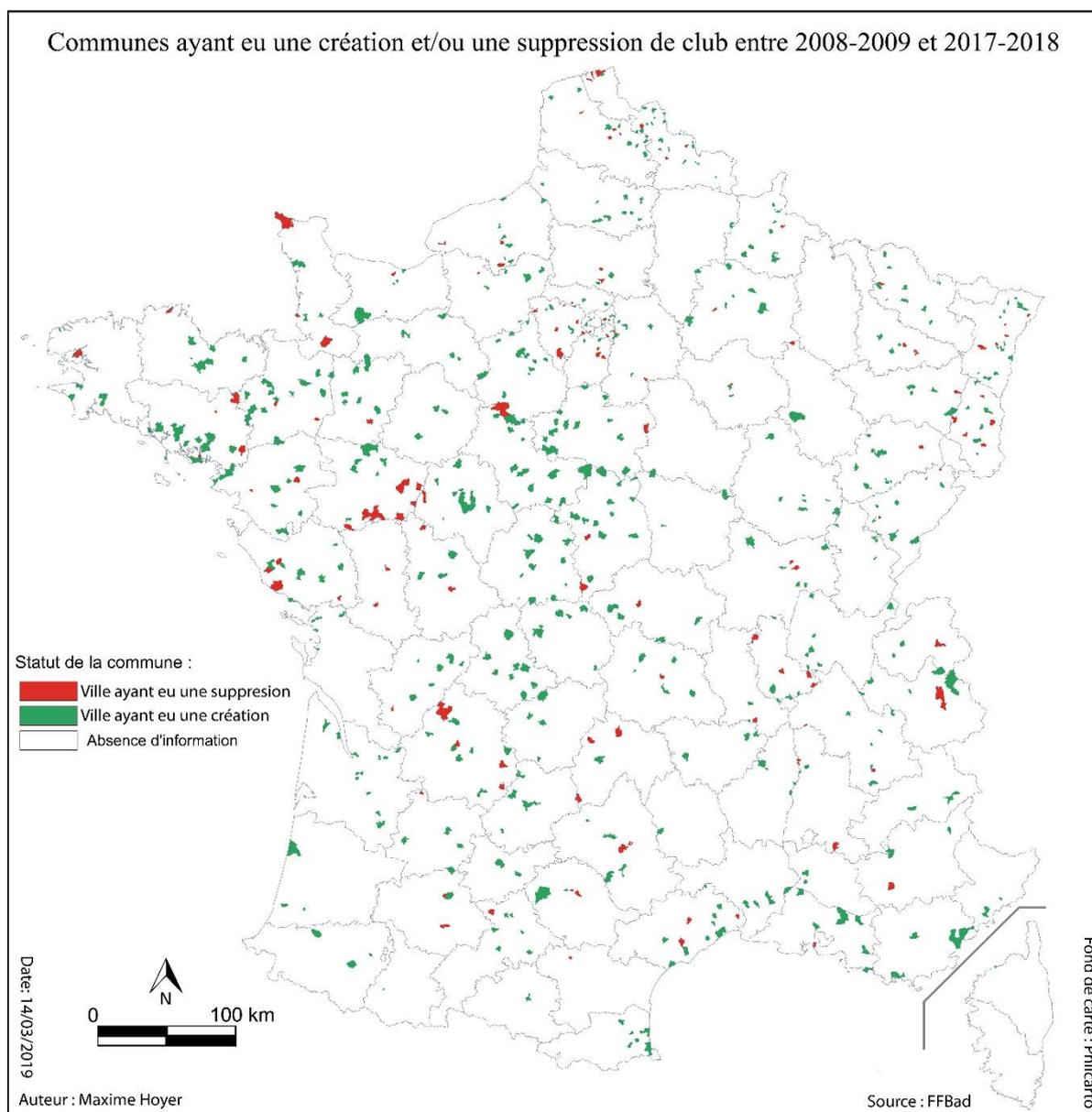
Taux de pénétration des badistes par commune en 2008-2009

Interprétation : On remarque le faible nombre de communes possédant un ou des clubs de badminton. Passer à l'échelle communale permet donc de mettre en avant de nouveaux phénomènes occultés par l'analyse départementale. Par ailleurs, seulement 1395 communes sont directement concernées par le badminton sur les plus de 36 000 de la France métropolitaine (elles représentent donc 3% des communes). Mais, ce nombre n'est cependant pas à négliger. En effet, la plupart des grandes villes françaises appartiennent à ces 3%.

Pour aller plus loin dans l'analyse, on retrouve des éléments communs avec les cartes vues précédemment. On remarque que de nombreuses communes de la Bretagne et du Val de Loire possèdent de forts taux de pénétration. La population de badistes est donc dispersée dans ces départements. À l'inverse, un département comme les Hautes-Alpes, avec également un fort

taux de pénétration, ne possède en réalité que cinq communes possédant un ou des clubs sur son territoire. Les badistes sont donc concentrés dans ce département.

Il est important de noter que l'analyse se faisant ici à l'échelle communale, des clubs ont pu apparaître et d'autres ont pu disparaître en 10 ans (information diluée à l'échelle départementale). En effet si 1395 communes possédaient un ou des clubs en 2008-2009, elles sont plus de 1770 en 2017-2018.

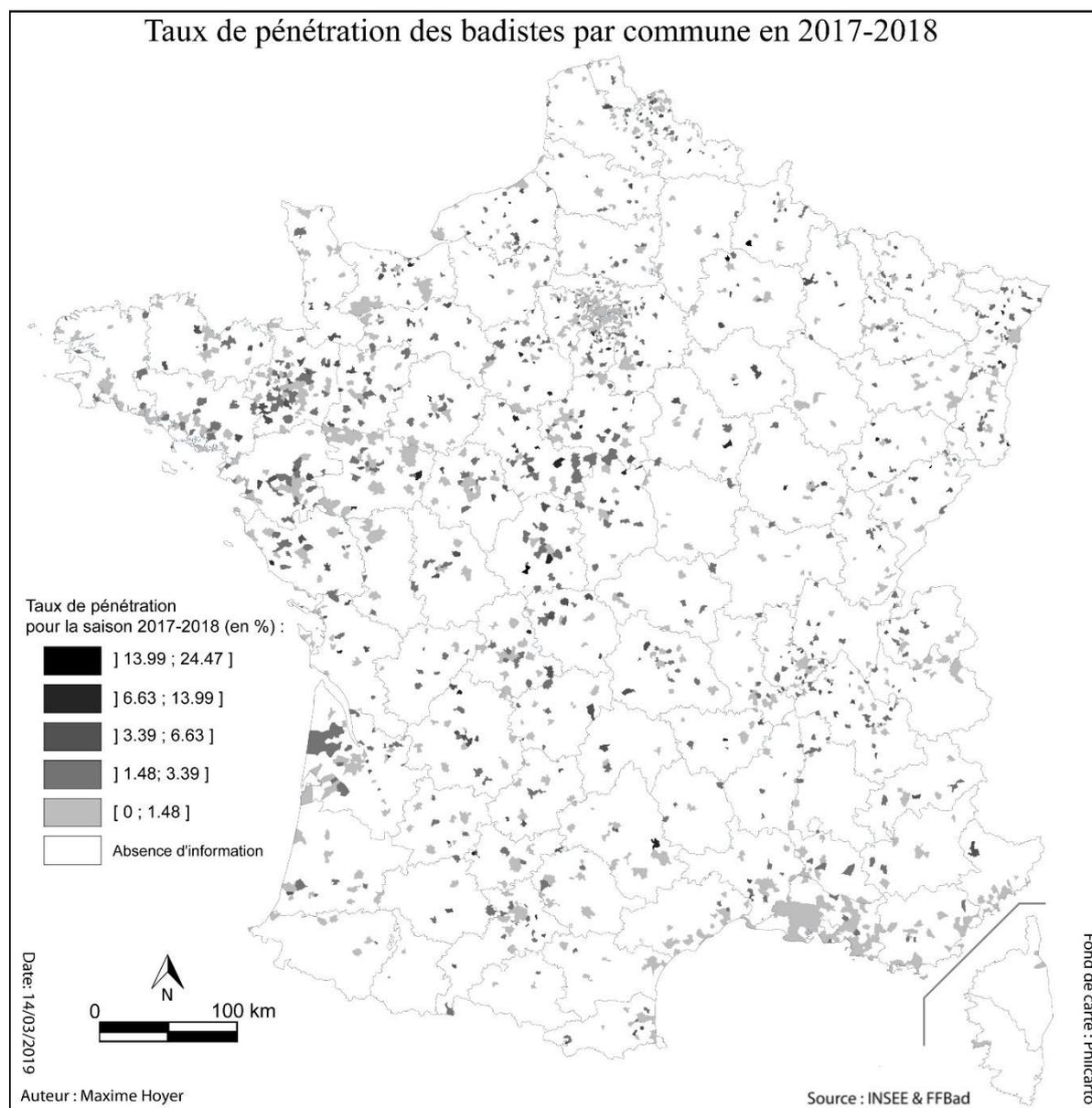


Communes ayant eu une création et/ou suppression de club entre 2008 et 2018

Interprétation : un peu plus de 500 communes ont vu un ou plusieurs clubs de badminton se créer. Le plus grand nombre de ces villes se situent dans la zone à forte augmentation taux de pénétration (cf p.50). A contrario, cette carte nuance ce phénomène. Même si un nombre

important de clubs ont été créés, d'autres ont également disparus (environ 140) entre 2008 et 2018. Par ailleurs, elle permet également de comprendre les raisons de la baisse du taux de pénétration des Deux-Sèvres et Maine-et-Loire en dix ans. En effet, on remarque que certaines communes de ces départements ont subi des suppressions de clubs.

Mais qu'en est-il du taux de pénétration par commune en 2017-2018 ?

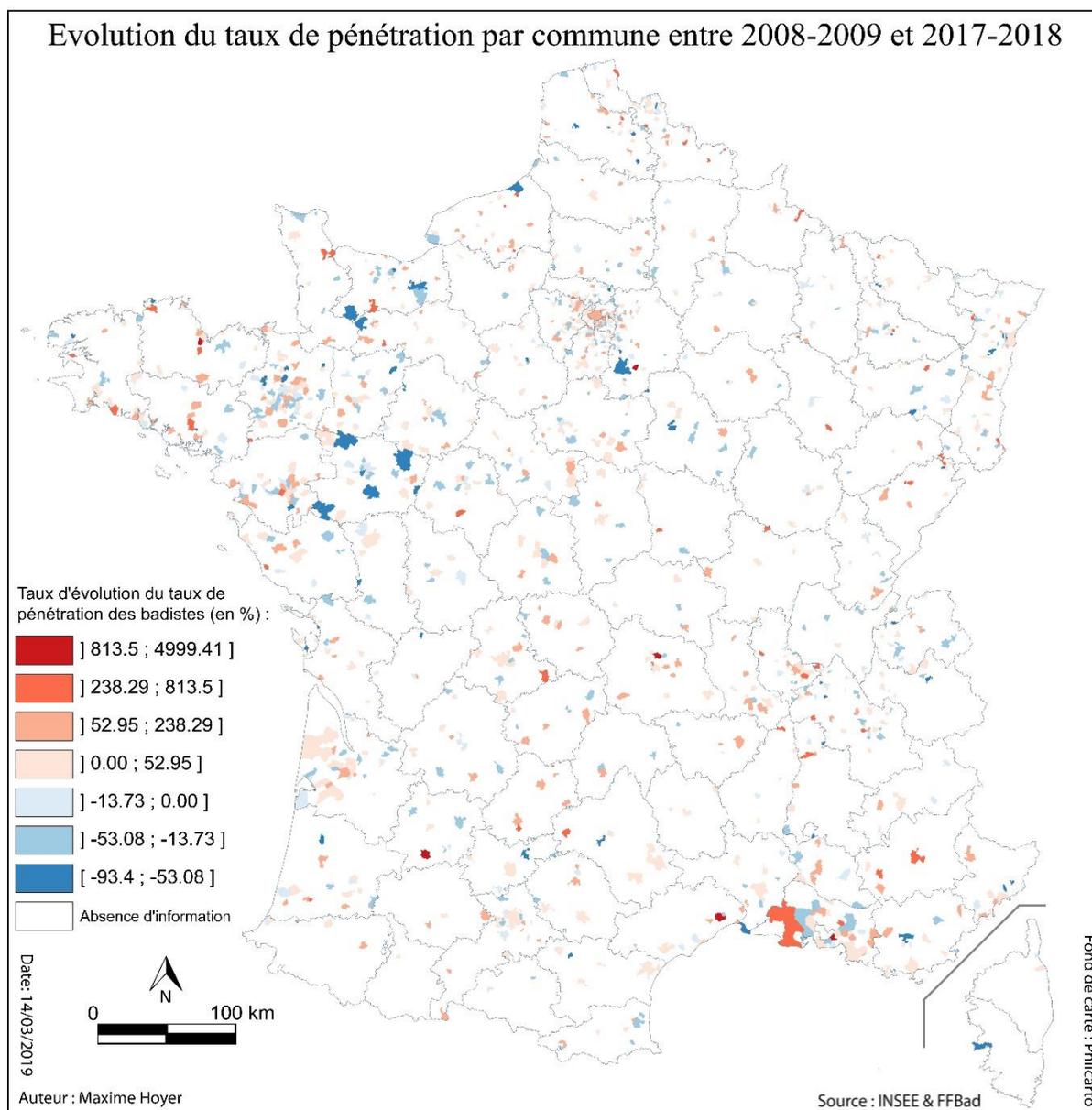


Taux de pénétration des badistes par commune en 2017-2018

Interprétation : Dans un premier temps, cette carte souligne l'augmentation du nombre de communes ayant un ou des clubs de badminton. De plus, elle permet de visualiser également le taux de pénétration des communes où un club a été créé entre 2008 et 2018. Comme on pourrait s'y attendre, le taux de pénétration dans ces communes est souvent bas (entre 0 et 1,48 %).

Néanmoins, on remarque que la Bretagne et les Pays de la Loire ont des taux de pénétration relativement élevés (entre 3,3% et 6.6%). Le badminton est toujours prédominant dans cette zone.

La carte ci-après analysera quant à elle l'évolution des communes entre 2008-2009 et 2017-2018.

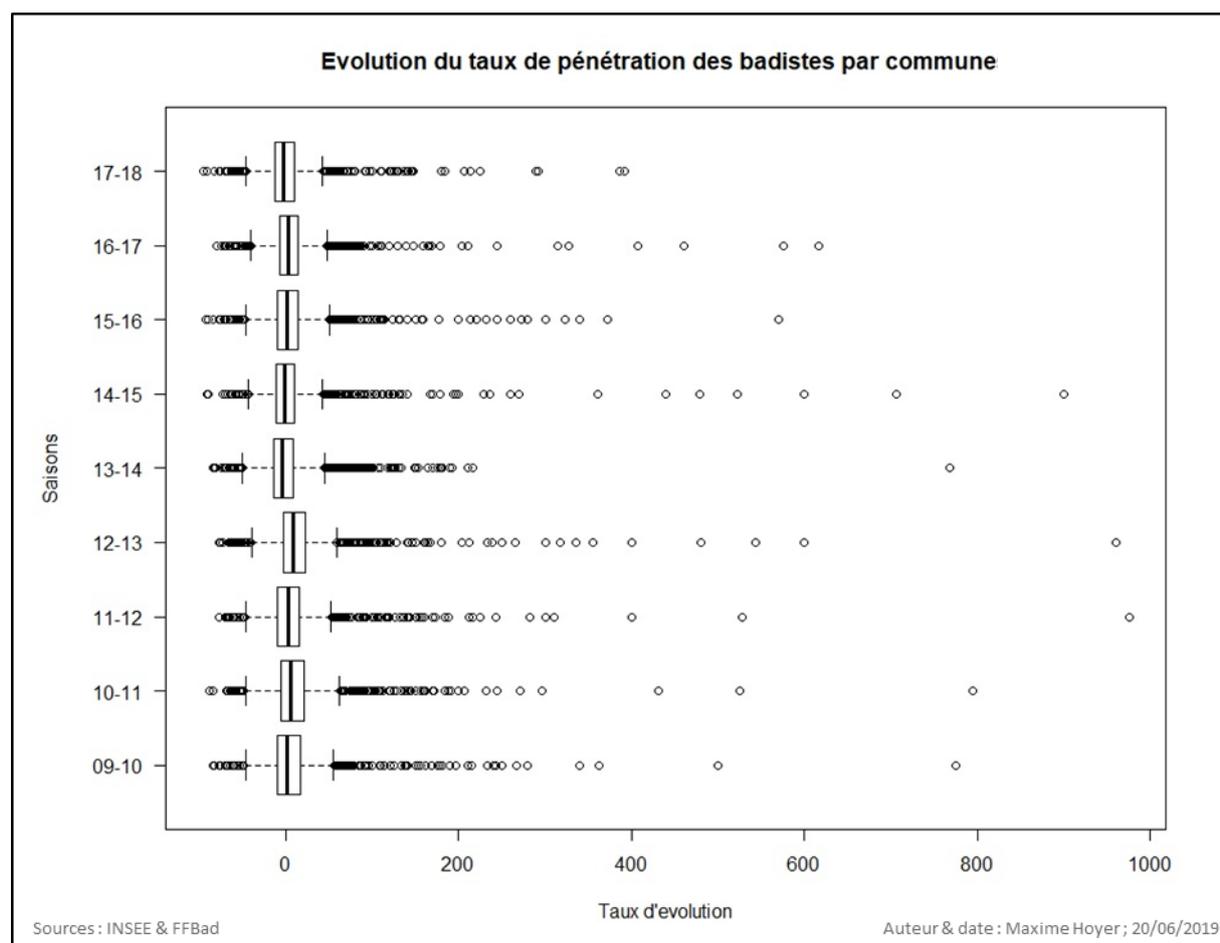


Evolution du taux de pénétration par commune entre 2008-2009 et 2017-2018

Dans la majorité des cas, le taux de pénétration semble stagner ou peu augmenter (maximum 50%). Cependant, des communes présentent des valeurs extrêmes. Le Maine-et-Loire présente, comme à l'échelle départementale, un taux de pénétration en baisse. A l'échelle communale on

remarque que seules quatre communes diminuent le taux global du département. De même, éparpillées sur l'ensemble du territoire, on retrouve des valeurs extrêmes positives (notamment dans les Côtes d'Armor, le Morbihan et les Bouches-du-Rhône). Si la tendance générale (médiane à 12%) tend vers une légère augmentation du taux de pénétration, il est toutefois nécessaire d'aborder cette évolution « statique » (c'est-à-dire une comparaison de deux années sur une longue période) avec du recul. Il est essentiel d'étudier les fluctuations de ce taux de pénétration.

Pour cela, a été réalisé une série de boxplot (aide à la lecture en Annexe p. ...). Ils portent sur l'évolution du taux de pénétration par commune entre une saison et la saison précédente. C'est pourquoi le graphique ne contient que neuf boxplots (la saison 2008-2009 ne pouvant être comparée avec la saison antérieure).



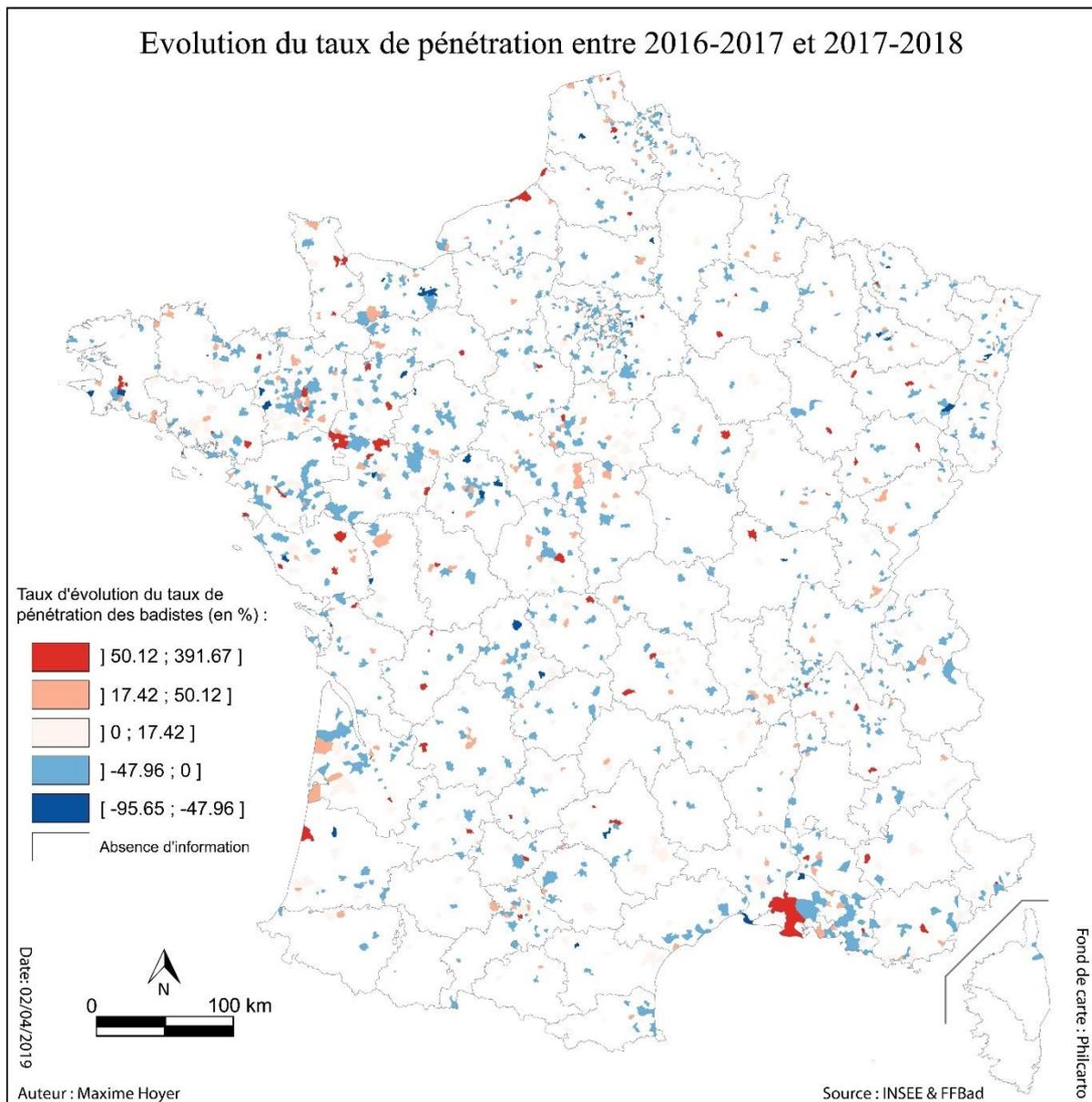
Evolution du taux de pénétration des badistes par commune

Les boxplots présentent une tendance à l'augmentation des taux de pénétration. Effet, la majorité des médianes sont positives. Néanmoins, trois saisons voient leur médiane en dessous de 0 :

- La saison 2013-2014
- La saison 2014-2015
- La saison 2017-2018

Ce graphique permet de nuancer l'évolution ; qui paraît linéaire si on se penche seulement sur les saisons de départ (2008-2009) et d'arrivée (2017-2018). Ce n'est pourtant pas le cas ; les deux saisons les plus récentes paraissent vraiment intéressantes. En effet, le faible taux d'évolution générale de la saison 2013-2014 peut être dû à des effets de structures. Les limites du taux de pénétration ont été présentées dans la partie 2 (cf p.33). Or il se trouve que la saison 2013-2014 est la première saison où le taux de pénétration a été calculé à l'aide du recensement 2015. L'évolution se fait donc entre la saison 2012-2013 (avec un taux de pénétration qui peut être sur-évalué) et la saison 2013-2014 (avec un taux de pénétration qui peut être sous-évalué). L'évolution entre ces deux saisons peut donc comporter un certain biais. C'est pourquoi, elle ne sera pas plus approfondie.

Néanmoins, il est intéressant de cartographier le taux d'évolution d'une de ces saisons pour visualiser les structures spatiales qui découlent de cette baisse. Le choix se portera donc sur la saison 2017-2018 qui présente une baisse relativement importante avec un taux d'évolution médian de -2,31%.



Evolution du taux de pénétration entre 2016-2017 et 2017-2018

Interprétation : Cette carte ne présente pas de structure spatiale évidente. Pourtant, il est à noter que la baisse du taux d'évolution est commune à la majorité des villes de France. Peu de communes ont un taux d'évolution qui baisse fortement (entre -47% et -96%). La plupart des villes qui voient leur taux de pénétration diminuer d'une saison sur l'autre ont un taux d'évolution compris en -47% et 0%. Néanmoins, certaines communes tirent tout de même leur épingle du jeu. En effet, même si la tendance générale est à une baisse du taux d'évolution, certaines communes voient ce dernier augmenter entre 2016-2017 et 2017-2018. C'est le cas de la ville d'Arles, dans les Bouches-du-Rhône.

Dans le début de cette troisième partie, nous avons abordé la structure de notre population et sa répartition spatiale. Cependant, même si certaines zones semblent plus concernées par le badminton, il est difficile de mettre en lumière une vraie structure spatiale. Il est donc nécessaire de trouver de nouveaux moyens pour mettre en avant des structures bien définies.

Pour ce faire, nous aurons besoin d'outils et de méthodes liés à l'analyse spatiale.

Diffusion spatiale du badminton

La diffusion spatiale se définit à la fois comme l'action de propagation et comme le résultat de cette action. Cette notion recouvre tous les processus en lien avec le déplacement dans l'espace géographique, et aux effets engendrés par ce dernier.

Comme vu précédemment, le badminton est réparti sur l'ensemble du territoire de la France métropolitaine. Cette distribution n'est pas homogène. En effet, les cartes précédentes mettaient en évidence des zones où la pratique du badminton était plus ou moins importante. Il est nécessaire :

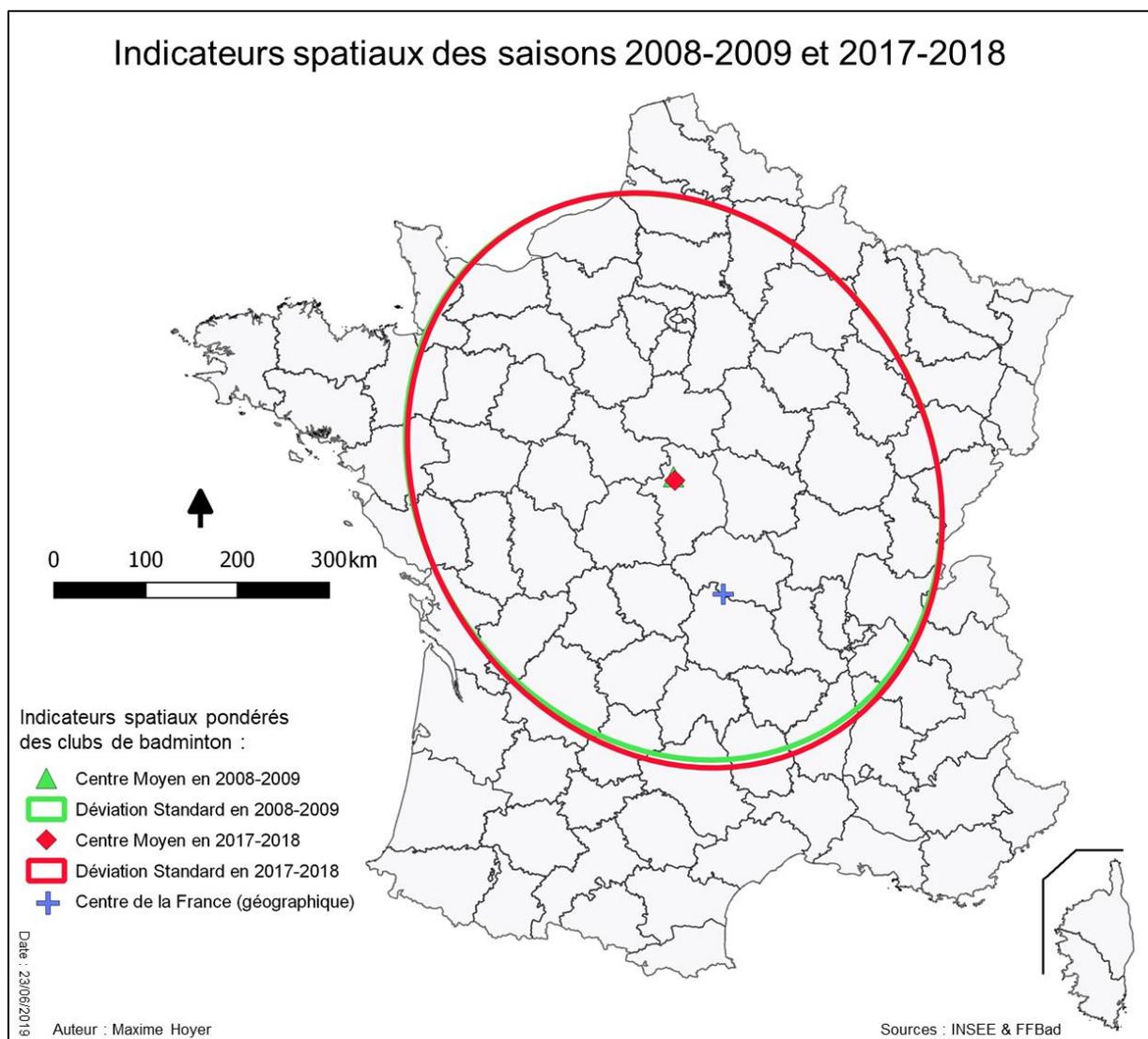
- Dans un premier temps, d'étudier les directions prises par cette diffusion ainsi que les zones de concentrations.
- Dans un deuxième temps, d'analyser ces phénomènes de manière temporelle.

DISPERSION SPATIALE

L'étude de la dispersion et des distances à un but essentiellement exploratoire. En effet, l'objectif de ces méthodes est de mettre en avant les tendances et les directions spatiales qu'a pris le badminton depuis 10 ans. Pour cela, nous allons utiliser les indicateurs d'analyses spatiales vus dans la partie méthodologie. Ces derniers ont donc été calculés à partir du semis de points des clubs de badminton. Les calculs ont été réalisés à la fois sur des points pondérés (c'est-à-dire que l'effectif de chaque club est associé à son point) et sur des points non pondérés.

Les résultats des indicateurs entre le semis de points pondéré et celui non pondéré étant presque identique, nous nous intéresserons essentiellement au semis pondéré. En effet, le fait que les

indicateurs pondérés et non pondérés soient semblables signifie que l'effectif n'a pas un grand impact sur la dispersion. En d'autres termes, le nombre moyen de joueurs par club est presque identique sur l'ensemble du territoire. C'est donc le nombre de clubs qui va impacter les indicateurs spatiaux et donc la direction spatiale prise.



Indicateurs spatiaux des saisons 2008-2009 et 2017-2018

Interprétation : Deux indicateurs principaux se chevauchent presque intégralement :

- Le centre moyen (\approx point d'équilibre du semis de points)
- L'ellipse standard (direction de la distribution)

En effet, à cette échelle les centres moyens sont confondus. Si les points ne sont pas exactement au même endroit géographique, leur distance est négligeable. On peut donc dire que l'équilibre du semis de points entre les deux saisons ne varie pas. La création de nouveaux clubs et la

disparition de certains en 10 ans ne change pas la position du centre moyen pour la saison 2017-2018.

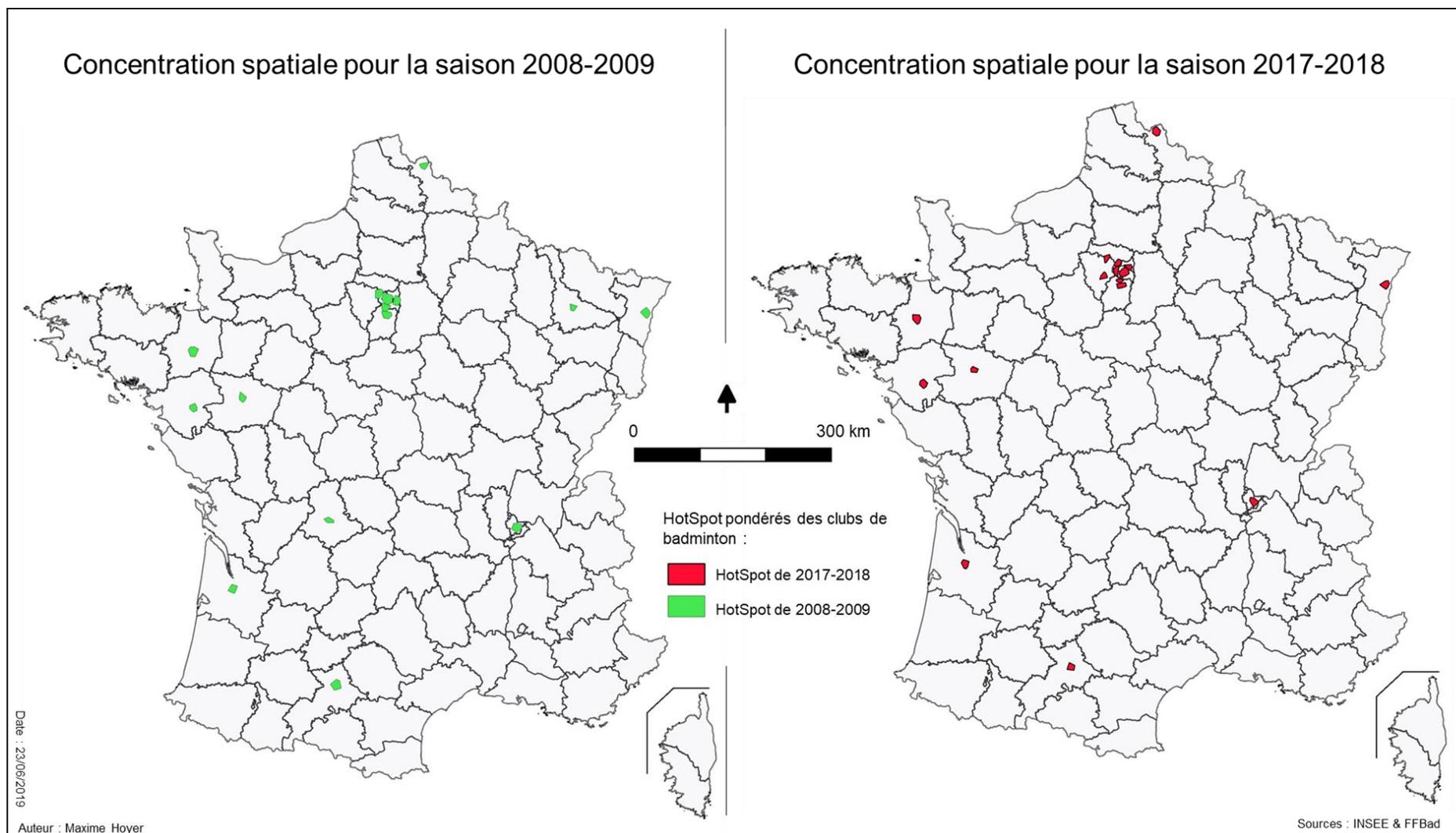
En ce qui concerne l'ellipse, on observe le même phénomène. En effet, les deux ellipses sont, comme pour les centres moyens, presque confondues. Donc, l'intervalle de 10 ans n'a que peu d'impact sur l'étalement de la distribution.

Par ailleurs, si l'on s'attarde également sur l'analyse des plus proches voisins (cf Annexe p.91) le même phénomène semble se dévoiler. En effet les distances moyennes des plus proches voisins entre les deux saisons sont très similaires (7735m en 2008-2009 et 7461m en 2017-2018). Malgré ce faible écart, il semblerait que cette distance tend à diminuer au fil du temps. La distance moyenne entre les clubs semble donc diminuer grâce notamment à l'augmentation du nombre de clubs.

On a donc une dispersion spatiale qui varie peu entre les saisons 2008-2009 et 2017-2018 même si les écarts de distance entre les clubs diminuent.

CONCENTRATION SPATIALE

Dans le point précédent, la dispersion spatiale a été évoquée. Ici, on parlera de concentration spatiale. Même si le badminton est réparti sur l'ensemble du territoire métropolitain (100% des départements possèdent au moins un club), certains territoires concentrent plus de clubs de badminton. Le but d'une analyse de hotspot est donc de mettre en avant ces zones.



Concentration spatiale pour les saison 2008-2009 et 2017-2018

Interprétations : Ces cartes mettent en avant plusieurs phénomènes :

1. Tout d'abord, il existe peu de différence entre les saisons 2008-2009 et 2017-2018. Pourtant, on remarque que certains hotspots existants sur la première carte ne sont plus présents sur celle de 2017-2018. Les hotspots présents dans les départements des Hautes-Viennes et de la Meurthe-Et-Moselle disparaissent entre les deux saisons. Il est difficile de comprendre les raisons de ce phénomène. Cependant, on observe un grand nombre de nouveaux clubs entre ces saisons dans les Hautes-Viennes (cf carte p.55). L'effectif des badistes sur le département s'est donc peut-être dispersé en dix ans.
2. De plus, les cartes affichent neuf zones de hotspots fixes entre 2008-2009 et 2017-2018. Certaines d'entre elles étaient prévisibles car elles se trouvent dans les plus grandes communes françaises :
 - Paris
 - Lyon
 - Bordeaux
 - Toulouse
 - Nantes
 - Lille

En revanche on ne retrouve pas de ville comme Marseille, qui comme vu précédemment, n'a pas une pratique du badminton très accrue.

3. D'autres villes moins peuplées présentent quant à elles des hotspots. C'est le cas de Strasbourg, d'Angers ou de Rennes. La région Ouest de la France possède donc plus de points de concentration du badminton. Hormis l'Île-de-France, seuls trois départements limitrophes possèdent un hotspot (Ille-et-Vilaine, Loire-Atlantique et Maine-et-Loire). On a donc une zone assez dense de la pratique du badminton. Contrairement au Sud-Est de la France qui en est dépourvu (hormis Lyon).

Pour répondre à la première hypothèse, même s'il n'existe pas de grandes zones de concentration spatiale du badminton (opposition Nord/Sud par exemple), la pratique du badminton est tout de même concentrée dans les grands bassins de populations et dans quelques bassins plus petits. De ce fait : **La pratique du badminton sur le territoire français est bien concentrée spatialement.**

Caractéristiques sociodémographiques des badistes

Après avoir étudié spatialement le badminton, nous allons à présent l'analyser par un prisme plus sociologique. Les informations sociodémographiques des individus n'étant pas accessible, on s'intéressera à celles des territoires. Ce sont donc les communes possédant au moins un club de badminton lors de la saison 2017-2018 qui seront étudiées. Le but ici est d'analyser **si les caractéristiques sociodémographiques des communes ont un impact sur le taux de pénétration**. Pour cela, plusieurs étapes vont être nécessaires :

1. Synthétiser l'information de notre jeu de données (ACP) ;
2. Regrouper les communes dans des classes (CAH) ;
3. Tester si le taux de pénétrations diffère en fonction des classes (tests statistiques).

CLASSIFICATION DES COMMUNES

Tout d'abord, il est nécessaire de créer une classification des communes. Cette catégorisation va se faire à l'aide des variables sociodémographiques suivantes :

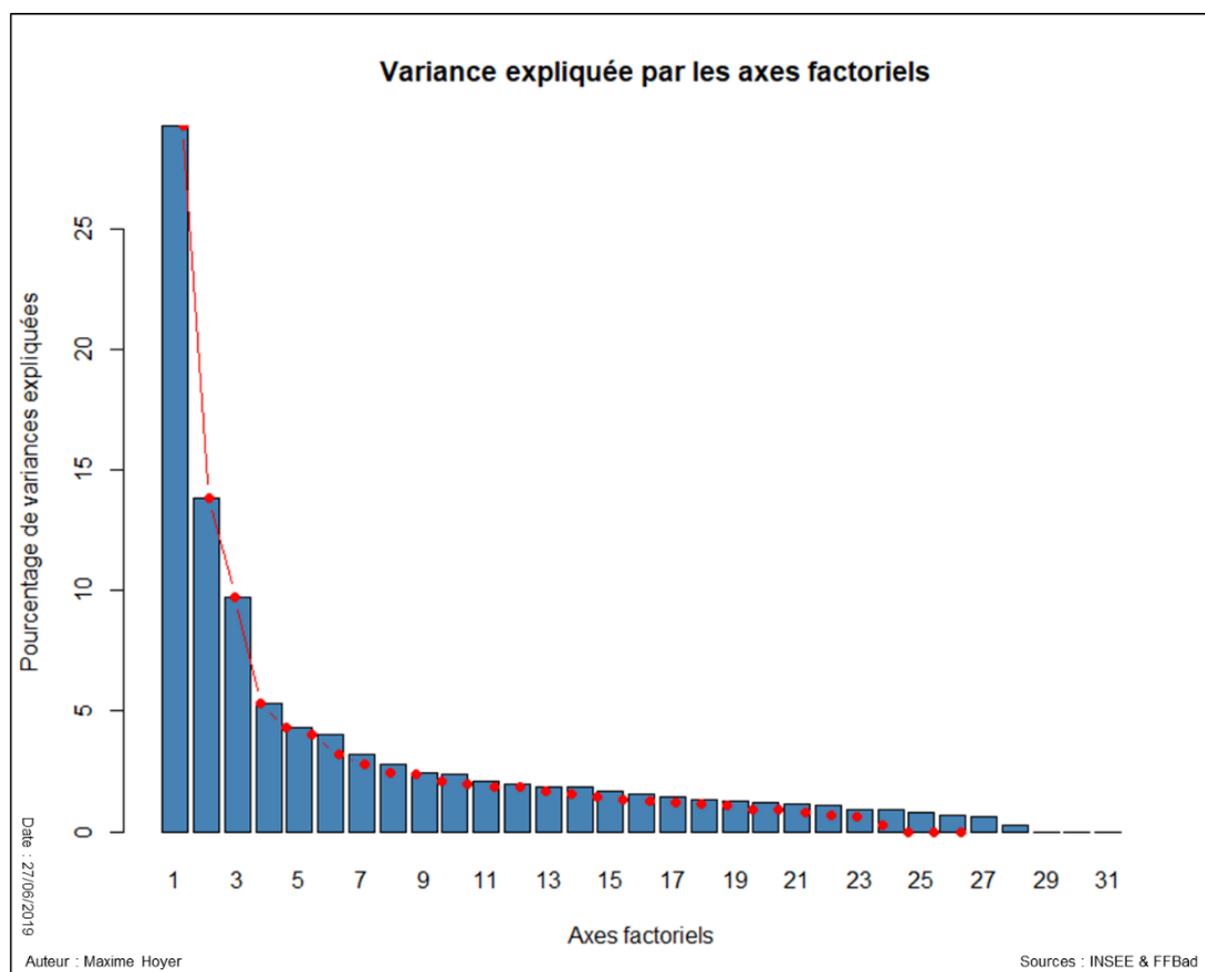
Liste des variables	
0-4 ans	80-84 ans
5-9 ans	85-89 ans
10-14 ans	90-94 ans
15-19 ans	95 ans et plus
20-24 ans	Aucun diplôme ou au mieux BEPC, brevet des collèges, DNB
25-29 ans	Diplôme de niveau CAP, BEP
30-34 ans	Diplôme de niveau Baccalauréat (général, techno, pro)
35-39 ans	Diplôme d'études supérieures
40-44 ans	Agriculteurs
45-49 ans	Artisans, commerçants, chefs d'entreprise
50-54 ans	Cadres et professions intellectuelles supérieures
55-59 ans	Professions intermédiaires
60-64 ans	Employés
65-69 ans	Ouvriers
70-74 ans	Revenu annuel médian
75-79 ans	

L'ensemble des 31 variables se trouvent sous la forme de taux (pourcentage), excepté pour le revenu annuel médian qui lui est exprimé en euros. En d'autres termes si une commune à un

taux de « 0-4 ans » de 10%, cela veut dire que 10% des habitants de cette commune ont entre 0 et 4 ans.

Dans un premier temps, à partir de ces variables, une ACP (Analyse en Composantes Principales) sera réalisée à l'aide du logiciel R. Cette méthode statistique très utilisée en science sociale a pour but de résumer l'information. Elle va permettre de condenser l'information contenue dans les 31 variables en seulement quelques nouvelles variables (axes factoriels).

Une fois l'ACP réalisée, il est nécessaire de choisir le nombre d'axes à analyser. Comme vu précédemment, le but est d'obtenir le maximum de variance expliquée tout en minimisant le nombre d'axes retenus.



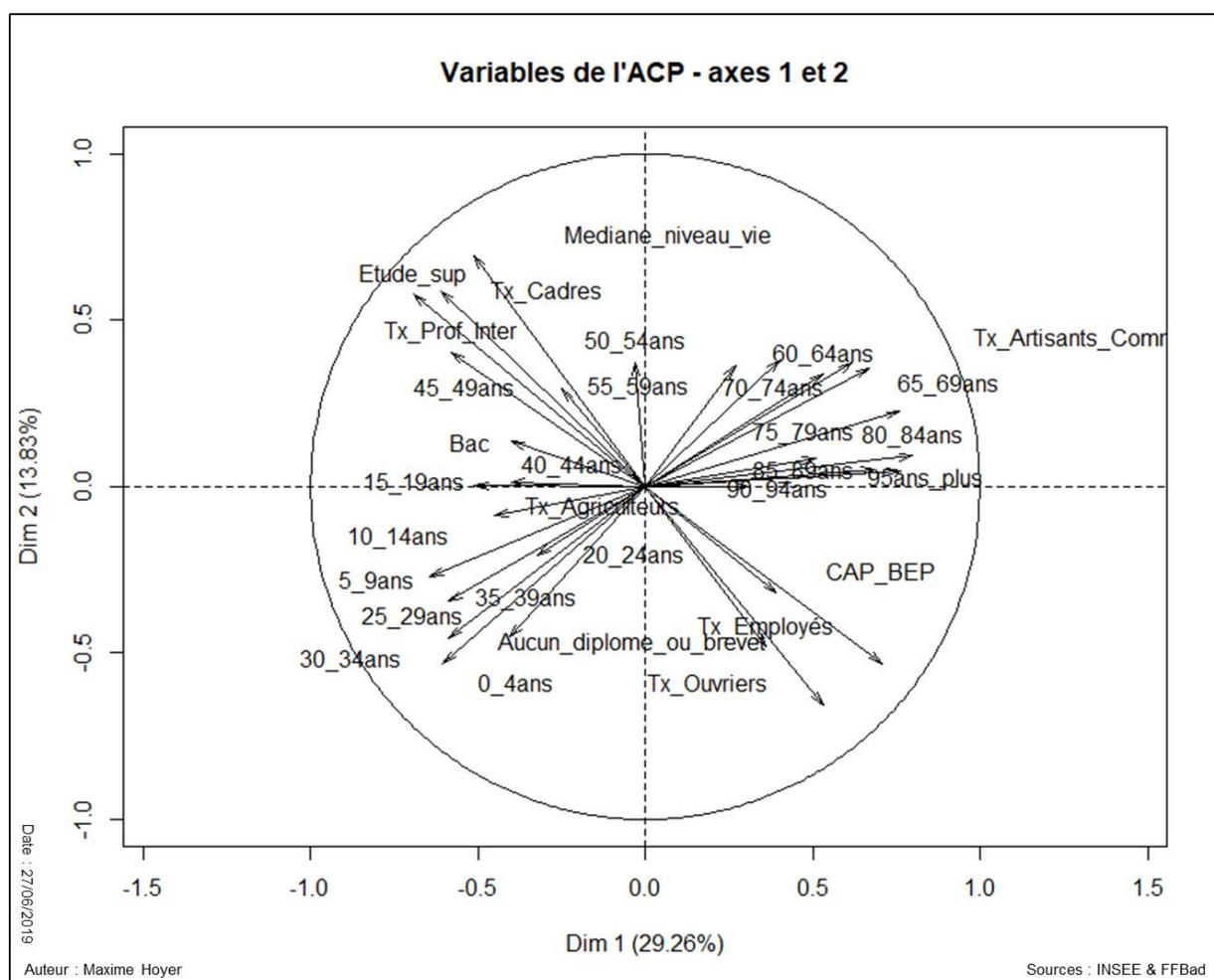
Variance expliquée par les axes factoriels

Interprétations : Dans notre cas, nous allons chercher lesquels de ces 31 axes factoriels sont pertinents. Les 3 premiers axes sont intéressants car ils représentent respectivement 29,26%, 13,83% et 9,74% de variance expliquée. Néanmoins à partir du 4^{ème} axe, les taux de variances

expliquées sont peu élevés (environ 5%) et sont très proches les uns des autres. L'axe 4 apporte donc peu d'informations par rapport à l'axe 5 et ainsi de suite. Seuls les 3 premiers axes seront retenus pour l'analyse. Ils contiennent à eux seuls 52,83% de l'information de nos 31 variables de départ.

La prochaine et dernière étape de l'ACP est l'interprétation des 3 axes retenus.

Il est nécessaire d'étudier les contributions de chaque variable aux 3 axes. Pour faciliter la visualisation, ces contributions vont être présentées sous forme de graphe.



Contributions des variables – Axes 1 et 2

Ce premier graphe permet de comprendre l'information contenue dans les deux premiers axes. L'axe des abscisses (axe 1) regroupe beaucoup d'informations. En effet, il met en avant deux oppositions :

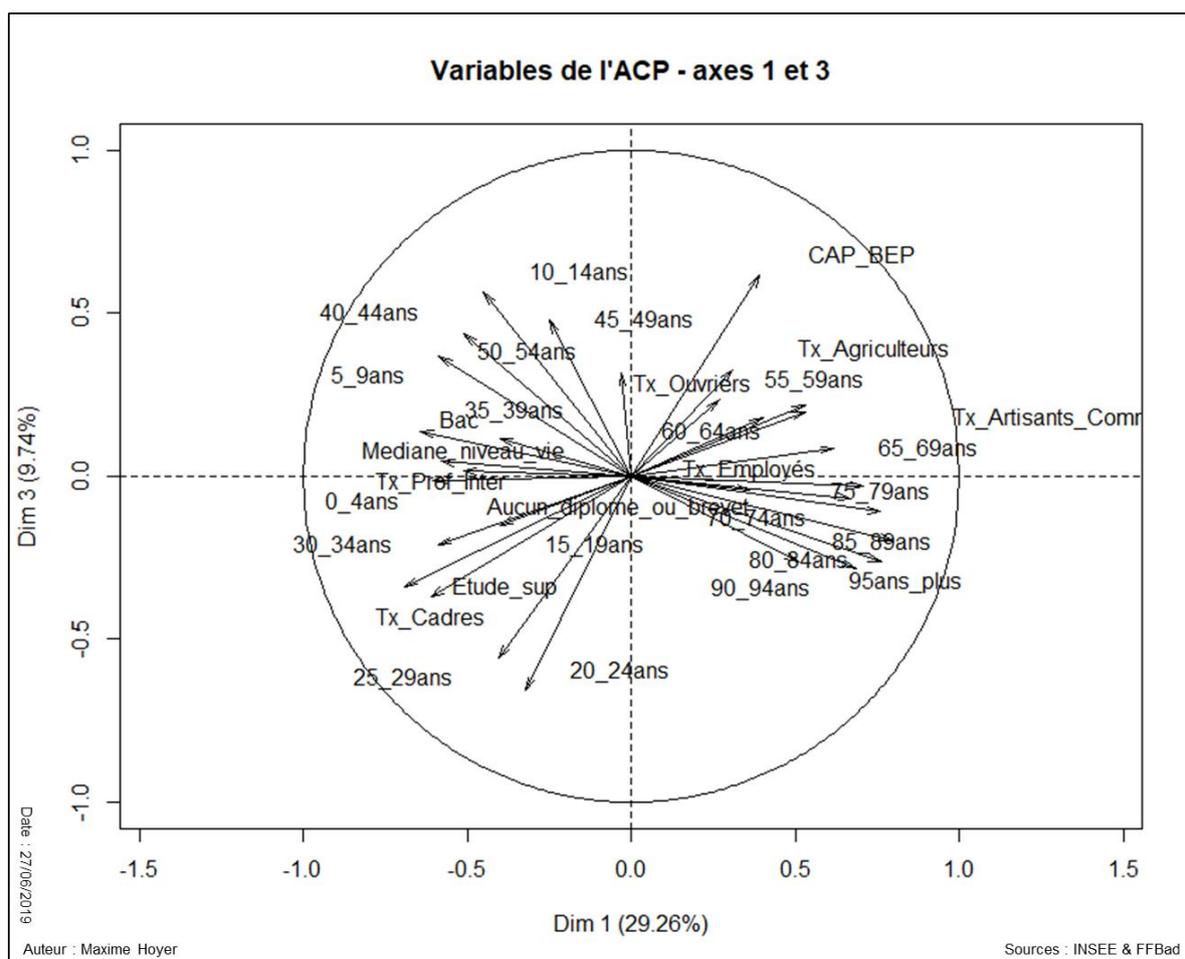
- **L'âge** : les moins de 55 ans se trouvent tous du côté négatif de l'axe (à gauche sur le graphe) alors que les 55 ans et plus se situent tous du côté positif.

- **Les CSP** : les Cadres et les Professions Intermédiaires ont des valeurs négatives contrairement aux Agriculteurs, Artisans, Employés et Ouvriers.

Par ailleurs la médiane du niveau de vie est très proche des cadres et des professions intermédiaires. L'axe 1 met donc en opposition les communes « jeunes » avec des cadres et des professions intermédiaires aux communes « âgées » avec les autres catégories socio-professionnelles.

L'axe des ordonnées quant à lui met en dualité une autre variable ; le plus haut diplôme obtenu. On observe que les variables « Aucun Diplôme et CAP/BEP » sont négatives sur l'axe 2 contrairement au « Baccalauréat et Études Supérieures » qui sont, elles, positives. Ce second axe met donc en opposition les niveaux de diplômes supérieurs ou égal au Baccalauréat aux diplômes en dessous du Baccalauréat.

Il reste maintenant à étudier le dernier axe retenu, l'axe 3.



Contributions des variables – Axes 1 et 3

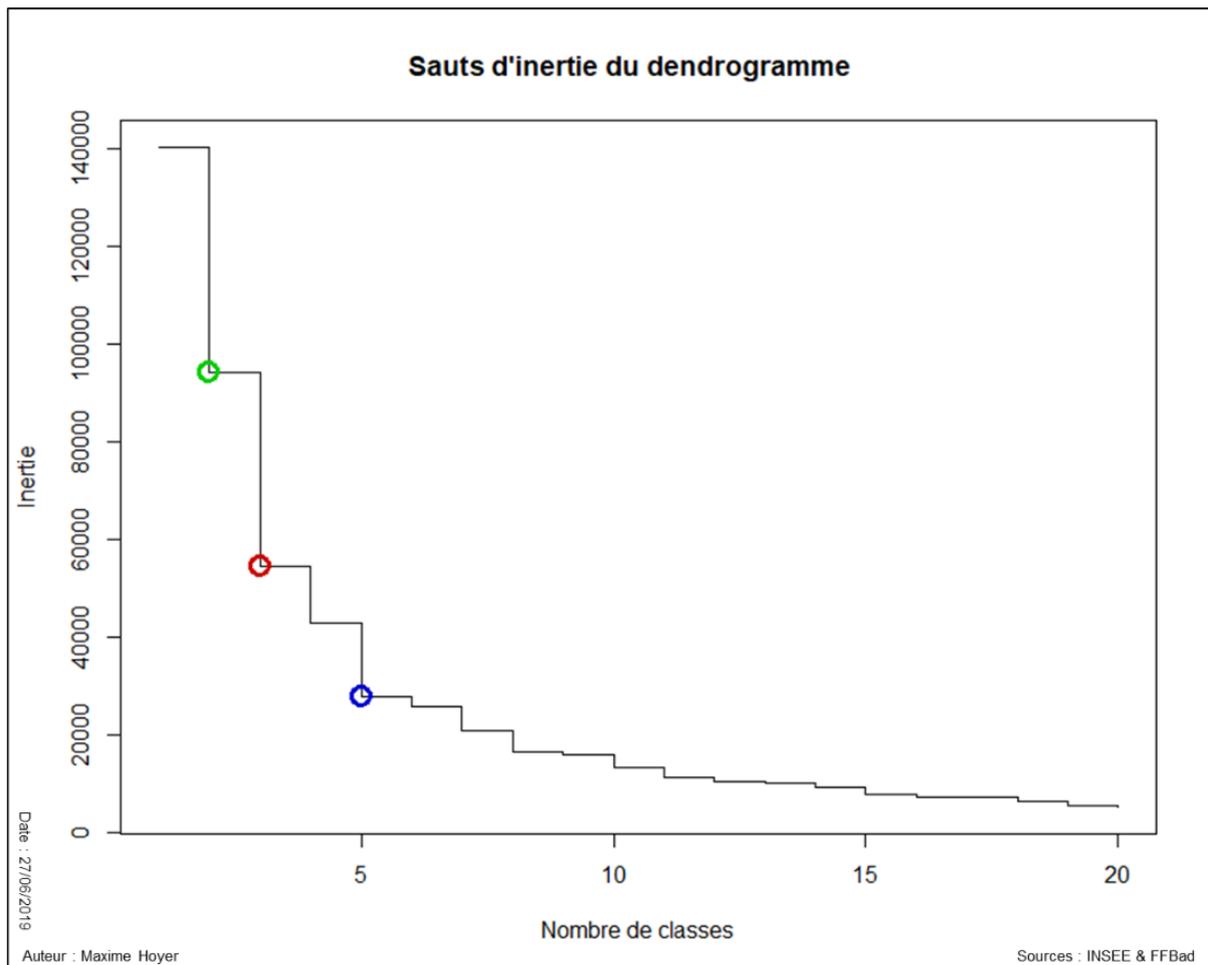
Interprétation : Plus on avance dans les axes et moins l'information qu'ils contiennent est évidente. Néanmoins, il semblerait que l'axe 3 mette en opposition la position dans le cycle de la vie :

- En positif : les parents de moins de 35 ans avec des enfants entre 0 et 4 ans et donc des grands-parents âgés de 70 ans ou plus.
- En négatif : Les parents un peu plus âgés (entre 35 et 70 ans) avec leurs enfants entre 10 et 20 ans.

Cette interprétation des axes sera très importante dans l'étude sur la classification des communes. En effet, toutes les communes des communes vont être projetés sur ces mêmes axes factoriels, puis divisées en classes. L'interprétation de ces classes reposera sur celle des axes ci-dessus.

A présent, il est temps de passer à la classification à proprement parler des communes. Comme expliqué en méthodologie, une Classification Ascendante Hiérarchique (CAH) va être utilisée. Elle se base sur la matrice des distances issues de l'ACP. Lors de l'analyse précédente, une matrice de distance de tous les points vers tous les autres a été calculée. C'est cette matrice qui sera la base des analyses faites dans cette CAH.

L'algorithme va partir de chaque point et va les agréger au fur et à mesure des itérations. Si bien que la dernière itération ne donne qu'une seule et unique classe. L'enjeu est donc d'identifier le nombre de classes adéquat. Pour cela, une analyse des sauts d'inertie est indispensable.

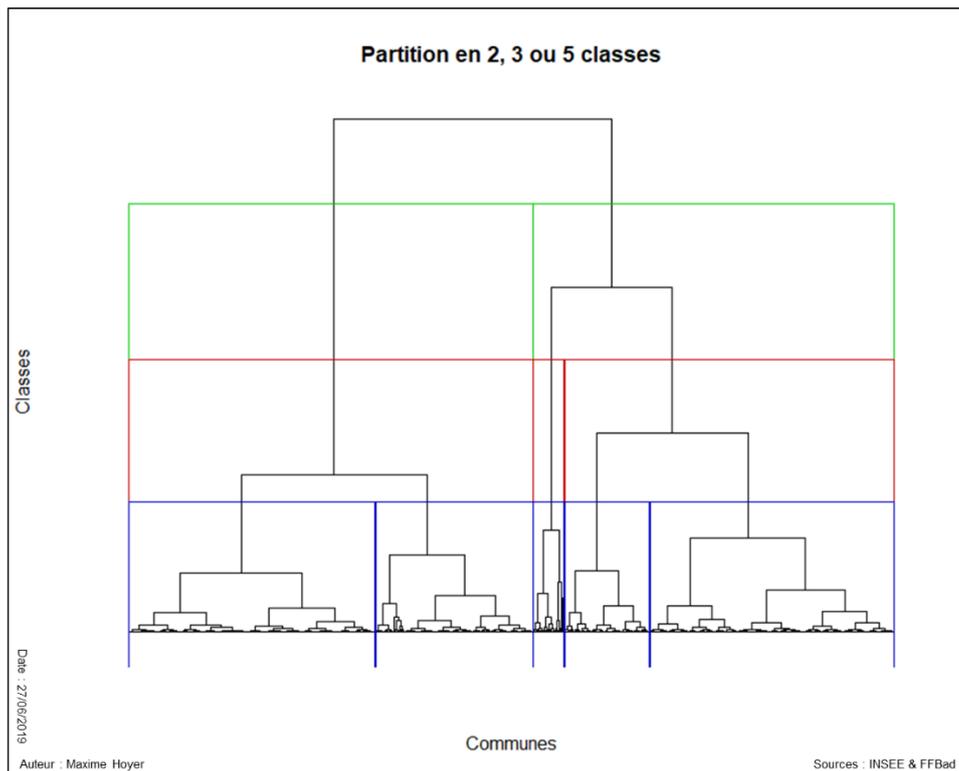


Sauts d'inertie du dendrogramme

Interprétation : Le graphique ci-dessus montre l'inertie obtenue par les premières répartitions (la répartition numéro 20, présente une classification en 20 classes). On cherche les plus grands sauts d'inertie entre 2 points. En d'autres termes, ce sont les grandes « chutes » sur l'axe des ordonnées qui seront à relever.

À la vue de cette courbe, seules 3 répartitions semblent pertinentes : en 2 classes (rond vert), en 3 classes (rond rouge) et en 5 classes (rond bleu).

Il est maintenant possible de visualiser le dendrogramme de la CAH en projetant en même temps les classes mises en avant grâce aux sauts d'inertie.



Partition en 2, 3 ou 5 classes

Interprétation : Le dendrogramme ci-dessus présente les différentes classifications. D'après le dendrogramme, une partition en 5 classes semble être la plus pertinente. À présent, chaque commune est associée à un groupe.

Ensuite, l'interprétation des classes se fait grâce à la projection de chaque communes (associé à leur classe par une couleur) sur les axes factoriels⁴⁹. Il est ainsi possible de lier chaque groupe aux différentes variables utilisées lors de l'ACP.

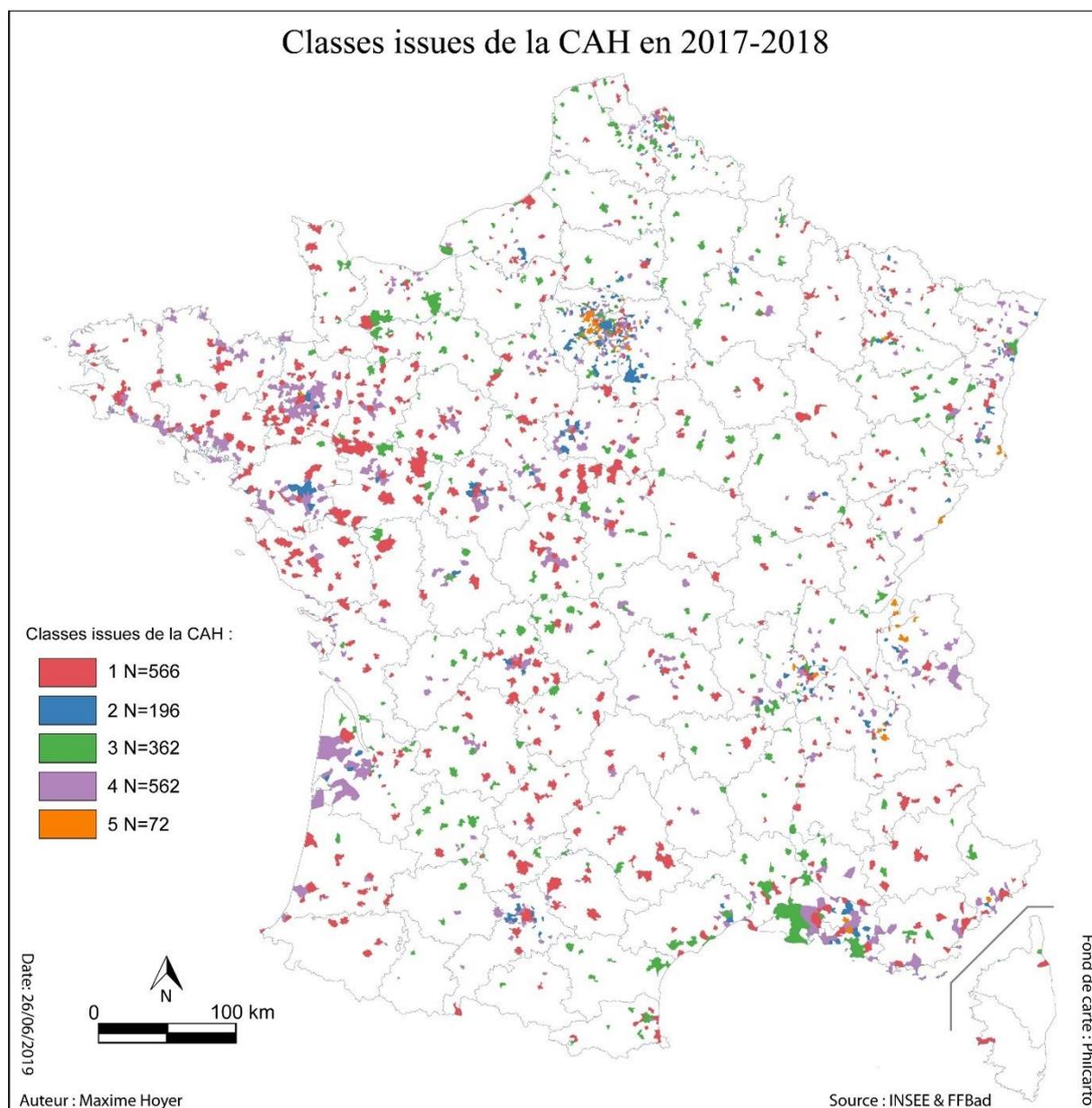
Variables associées à chaque classe :

- Classe 1 : 20-40 ans / 60-80 ans
- Classe 2 : 40-55 ans / Artisans / Agriculteurs
- Classe 3 : Employés / Ouvriers / CAP-BEP / 80 ans et plus
- Classe 4 : 0-19 ans / 55-60 ans / Baccalauréat
- Classe 5 : Cadres / Professions Intermédiaires / Études Supérieurs

⁴⁹ cf graphe en Annexe p.93

À présent, l'ensemble des communes avec un ou des clubs de badminton sont réparties en 5 classes. Il est maintenant possible de cartographier ces classes. Cette étape va permettre deux choses :

- Vérifier, si possible, l'interprétation de certaines classes
- Faire le lien entre l'analyse spatiale et l'analyse sociodémographique



Classes issues de la CAH en 2017-2018

Interprétation : Dans un premier temps, cette carte semble confirmer les interprétations des classes. En effet, on remarque que de nombreuses communes des Yvelines (département avec un bon niveau de vie) se trouvent dans la 5^{ème} classe (classe à forts taux de cadres et de

professions intermédiaires). Par ailleurs, la 3^{ème} classe regroupe des villes plus tournées vers l'agriculture ou l'industrie par rapport à l'ensemble des communes de notre étude (comme par exemple la commune d'Arles dans les Bouches-du-Rhône).

L'interprétation des classes paraît donc cohérente.

Dans un second temps, la carte met en évidence certaines zones (notamment où la pratique du badminton s'est avérée importante) concentrant des communes de même classe. En effet, on remarque que les départements de l'Ille-et-Vilaine, de la Loire-Atlantique et du Maine-et-Loire concentrent de nombreuses communes des classes 1 (20-40 ans / 60-80 ans) et 4 (0-19 ans / 55-60 ans / Baccalauréat). Cependant, la ville de Strasbourg, appartenant à la 3^{ème} classe (Employés / Ouvriers / CAP-BEP / 80 ans et plus), possède également une pratique du badminton importante.

Même si la carte permet de visualiser notre phénomène, elle ne permet pas de dire si le taux de pénétration dépend des caractéristiques sociodémographiques des communes.

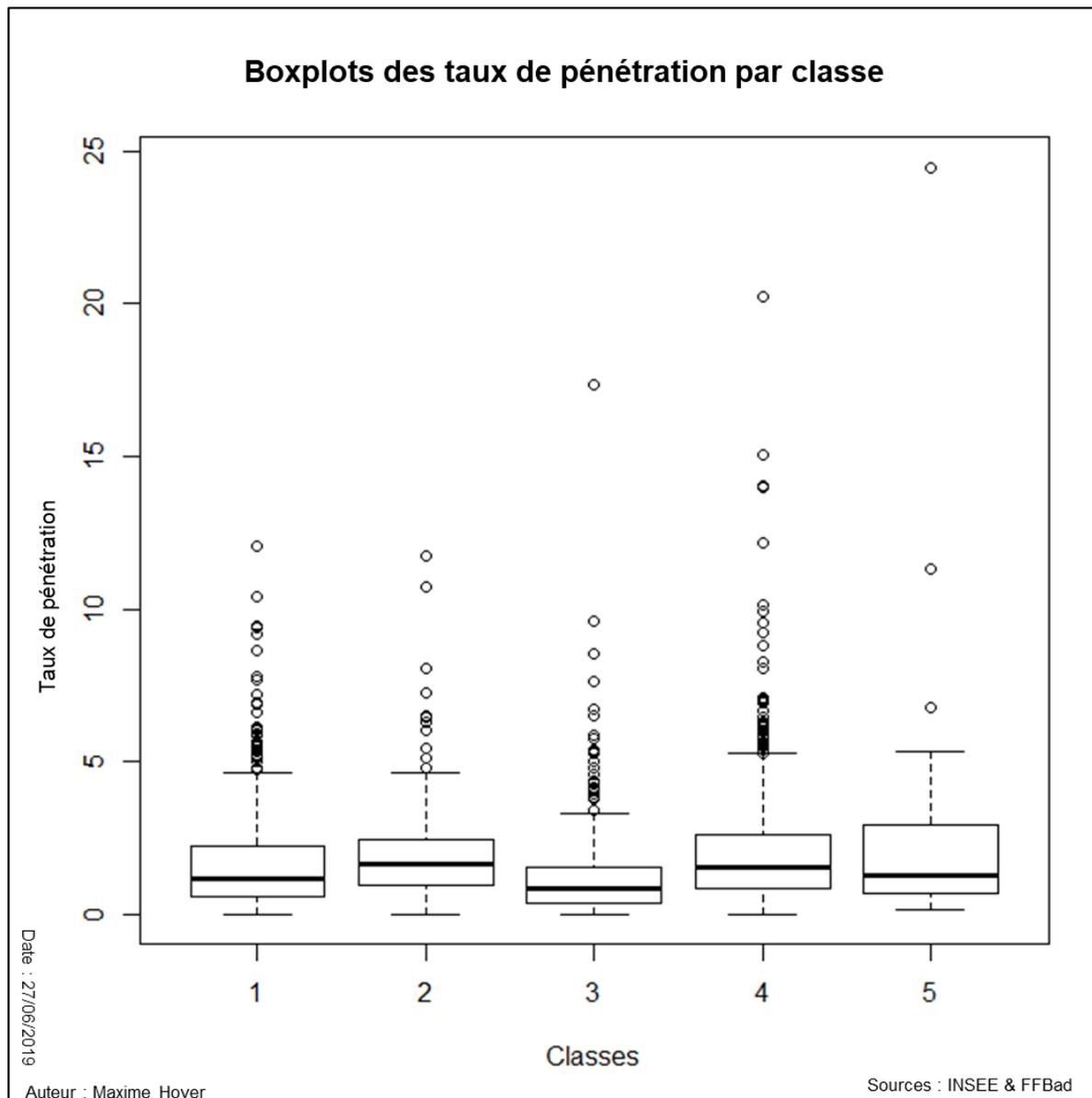
De ce fait, des tests statistiques vont être utilisés pour parvenir à cette conclusion.

LES CARACTÉRISTIQUES SOCIALES INFLUENCENT LA PRATIQUE

Cette dernière partie est consacrée à notre seconde hypothèse : « La pratique du badminton dépend des caractéristiques sociodémographiques des lieux où elle est exercée. »

L'objectif est d'identifier si le taux de pénétration des communes est significativement différent en fonction de leur classe. En d'autres termes on teste si, en moyenne, les taux de pénétration d'une classe sont différents des autres.

Dans un premier temps, l'analyse graphique nous aiguille dans les recherches. Pour cela, on compare les boxplots du taux de pénétration des différentes classes.



Interprétation : Ce graphique met en lumière des dispersions différentes. En effet, on remarque que la moyenne et la médiane du groupe 3 sont inférieures aux autres. A l'inverse, celles des classes 2 et 4 semblent supérieures aux autres. Trois groupes semblent se distinguer des autres. Cependant même si ce graphique permet de visualiser le phénomène escompté, **il n'atteste pas de la significativité de ces différences.**

C'est pourquoi, un test statistique très courant en science sociale va être utilisé, l'ANOVA. Comme expliqué en méthodologie, une ANOVA vise à tester les différences significatives des moyennes de différents groupes ; tout à fait approprié ici.

Néanmoins pour qu'une ANOVA soit réalisable, il est nécessaire de valider deux hypothèses de test :

- i. La normalité des résidus
- ii. L'homoscédasticité des résidus

Cependant, ces deux tests n'ont pas été concluants⁵⁰. En effet, les deux p-valeurs des tests sont inférieures à 5%, il n'y a donc pas de normalité et d'homoscédasticité des résidus. **Par conséquent il n'est pas possible de réaliser une ANOVA.**

Néanmoins, nous avons vu précédemment qu'il existait une alternative à l'ANOVA, le test de Kruskal-Wallis. Ce dernier s'intéresse aux médianes, quand une ANOVA se concentre sur les moyennes. Un test de Kruskal-Wallis a donc été réalisé : p-value < 2.2e-16. La p-valeur est ici inférieure au seuil de significativité 5%. Il existe donc une ou plusieurs classes avec des taux de pénétration significativement différentes des autres classes. Pour mettre en avant lesquelles, il est nécessaire d'analyser les sorties⁵¹ de ce test. Elles mettent en commun des groupes :

- Les classes 1, 2 et 5
- Les classes 2, 4 et 5
- La classe 3

Deux résultats ressortent des sorties. La première est qu'il existe une différence significative entre les classes 1 (20-40 ans / 60-80 ans) et 4 (0-19 ans / 55-60 ans / Baccalauréat). A la vue du graphique précédent, les communes de classe 4 ont donc tendance à avoir des taux de pénétration plus importants. Cette conclusion vient appuyer le fait que ceux sont les 10-19 ans qui pratiquent le plus le badminton.

La seconde conclusion est la plus marquante. En effet la classe 3 (Employés / Ouvriers / CAP-BEP / 80 ans et plus) est significativement différente des autres. Les communes avec des taux d'employés, d'ouvriers, de CAP-BEP plus importants ont des taux de pénétration significativement plus faibles que les autres communes.

On peut donc finalement conclure et répondre à notre seconde hypothèse : **les caractéristiques sociodémographiques des territoires du badminton ont donc bel et bien une influence sur sa pratique.**

⁵⁰ cf Annexe p.94

⁵¹ cf Annexe p.94

Enfin de par les analyses précédentes, il est possible de répondre à la dernière hypothèse. En effet, nous avons vu que les lieux de pratique du badminton n'avaient que peu évolué en 10 ans. Même si cette décennie a vu arriver de nouveaux clubs dans de nouvelles communes, la spatialisation du badminton n'en fût pas bouleversée. De ce fait, il est possible d'en déduire que **les variables sociodémographiques des territoires qui influent sur la pratique du badminton, n'ont que très peu changé en 10 ans.**

Conclusion de la troisième partie

Cette 3^{ème} partie aura mis l'accent sur la spatialisation de badminton en France : sa dispersion et sa concentration. Des analyses auront ensuite mis en avant quelles caractéristiques sociodémographiques avaient un impact sur la pratique du badminton.

Les analyses, cartes et graphes auront permis, par ailleurs, d'apporter un éclairage sur les hypothèses formulées et de conclure :

- Qu'il existe une concentration spatiale du badminton en France.
- Que les variables sociodémographiques des territoires influent sur la pratique du badminton.
- Que les caractéristiques sociodémographiques des territoires influant sur la pratique du badminton n'ont pas évolué en dix ans.

Venons-en maintenant à la conclusion finale de cette étude

Conclusion générale

Ce mémoire intitulé « Les territoires du badminton en France » visait un triple objectif :

- i. Apporter, tout d'abord, un état de lieux sur ce sport depuis plus de 10 ans.
- ii. Comprendre les logiques de spatialisations des clubs sur le territoire français.
- iii. Mettre en évidence l'importance des caractéristiques sociodémographiques dans la pratique du badminton.

Au travers de ces objectifs, ce mémoire présentait trois parties. La première permettait de situer le contexte de notre étude ainsi que les enjeux liés au sport et plus particulièrement au badminton en France. Cette partie introductive amenait également les questionnements et hypothèses. La deuxième partie quant à elle mettait en évidence les méthodes et les analyses utiles à nos questionnements. Enfin, la dernière partie présentait à la fois les résultats des analyses ainsi que leurs interprétations et répondait à nos hypothèses.

En définitif, l'étude présentée aura mis en lumière des phénomènes liés à la pratique du badminton en France :

- Des zones de concentrations spatiales du badminton.
- La faible évolution spatiale de la pratique entre 2008-2009 et 2017-2018.
- L'impact de certaines variables sociodémographiques sur la pratique du badminton.
- La faible évolution de l'impact de ces variables sociodémographiques entre 2008-2009 et 2017-2018.

Cette étude aura, par ailleurs, ouvert la voie à de nouvelles interrogations.

En effet, même si de nombreuses variables ont été utilisés dans les analyses multivariés, elles n'ont permis que d'expliquer un peu plus de 50% de la répartition spatiale du badminton en France. **Il serait donc intéressant de trouver quelles peuvent être les autres variables qui influent sur cette répartition.**

Par ailleurs ce mémoire, de géographie, s'est donc naturellement intéressé aux territoires plutôt qu'aux individus. **Une étude (par questionnaires par exemple) pourrait donc être particulièrement pertinente pour la compréhension de notre population de badistes.** Mais bien d'autres analyses seraient possibles sur ce sujet (exemple : se servir des communes de résidences au lieu de rattacher les licenciés à la commune de leur club ; si de nouveaux joueurs arrivent chaque année, qu'en est-il des départs ? ; ...).

Les perspectives et les intérêts de sont donc très nombreuses et j'espère, permettront au badminton de pouvoir continuer son expansion en France.

Bibliographie

- Annales de géographie ; *Qu'est-ce que le sport ? Cultures sportives et géographie* ; Annale n° 680 ; 2011.
- A. Pedro ; *Paul-André Tramier : « Le badminton préfigure ce que sera la domination asiatique du sport »* ; 20 Minutes ; 2014.
- D. Mathieu, J. Praicheux, J-P Volle ; *Sports en France* ; MappedMonde ; 1987.
- F. Bruyn, & A. Bringé ; *Un prolongement de la démographie sportive : l'analyse de durée appliquée aux populations de sportifs licenciés* ; Population Vol. 61 ; 2006.
- F. Pelletier & T. Spoorenberg ; *Aperçu sur les méthodes de projection* ; 2016.
- G. Di Méo ; *Géographie sociale et territoires* ; Fac-géographie ; 1998.
- G. Martin ; *Essai sur les territoires du judo en France* ; Université Michel de Montaigne ; 2011.
- Institut National de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture ; *Les indices d'auto-corrélation spatiale* ; OASIS ; 2015.
- Institut National de la Santé et de la Recherche ; *Évolution des pratiques physiques et sportives* ; 2008.
- Institut National de la Santé et de la Recherche ; *Activité physique* ; 2008.
- J. Bale & D. Trudo ; *Editorial. Sports Geography: An Overview* ; *Belgeo* n° 2 ; 2008.
- J. Cazeneuve ; *Vers une civilisation du loisir ?* *Revue française de sociologie* 3, n° 4 ; 1962.
- J. Harduin ; *La transformation stratégique du sport face au développement durable : le cas du badminton en France* ; Rouen Business School.

- J. Gingembre ; *La spécialisation résidentielle dans le Sillon Lorrain* ; Revue Géographique de l'Est ; Vol. 56 ; 2016.
- J. Grall ; *Sport ou pratique enfantine ? La construction manquée du badminton en tant que « vrai sport » en France dans l'entre-deux-guerres* ; Staps n° 107 ; 2015.
- J. Grall ; *Histoire du badminton en France (fin XIXe siècle – 1979) : pratiques et représentations* ; Université Rennes 2 ; 2018.
- J-M Zaninetti ; *Statistique spatiale : méthodes et applications géomatiques* ; Édition Lavoisier ; 2005.
- J-P Augustion ; *Le sport, une géographie mondialisée* ; Documentation géographique n°8112 ; 2016.
- J-P. Augustin ; *Les jeunes dans la ville : institutions de socialisation et différenciation spatiale dans la Communauté urbaine de Bordeaux : recherche de géographie sociale et politique* ; 1989.
- J-Y. Guillan ; *Histoire du badminton : Du jeu de volant au sport olympique* ; Édition Broché , 2002.
- L. Muller ; *La pratique sportive en France, reflet du milieu social* ; Vie social ; 2006.
- Ministère de la ville, de la jeunesse et des sports ; *Les enjeux de l'économie du sport* ; 2016.
- M. Douet Guérin ; *Pratiques sportives, normes et socialisation* ; Université Paris Descartes ; 2013.
- Ministère des Sports ; *Les Chiffres clés du sport 2015* ; 2015.
- Ministère des Sports ; *Les Chiffres clés du sport 2017* ; 2017.

- O. Donnat ; *Pratiques culturelles 1973-2008* ; Culture Études ; 2011.
- P. Langlois & É. Daudé ; *Concept et modélisations de la diffusion géographique* ; Cybergeog : European Journal of Geography ; 2007.

- P. Mignon ; *Point de repère - La pratique sportive en France : évolutions, structuration et nouvelles tendances* ; Informations sociales n° 187 ; 2015.

- P. Surault ; *Pour une estimation de la population « sportive »* ; Population Vol . 46 ; 1991.

- R. Rakotomalala ; *Comparaison de population*; Université de Lyon 2 ; 2008.

- R. Rakotomalala ; *Test de normalité* ; Université de Lyon 2 ; 2011.

- S. Carpentier-Postel ; *Analyse spatiale* ; Aix-Marseille Université ; 2018-2019.

- S. Carpentier-Postel ; *Analyse spatiale et interactions individus, espace et société* ; Aix-Marseille Université ; 2017-2018.

- S. Maneo ; *Étude sociodémographique des pratiquants du Badminton en France* ; Université Lyon 1 ; 2016.

- S. Oliveau ; *Autocorrélation spatiale : leçons du changement d'échelle* ; L'Espace géographique Vol. 39 ; 2010.

- T. Raymond ; *Sociologie du sport* ; Édition Fénixx ; 2002.

Annexe

Formulaire de demande de licence à la FFBaD pour la saison 2018-2019



#ES PRIT BAD

DEMANDE DE LICENCE 2018/2019

2

Remplissez en lettres capitales et signez en 2 exemplaires ce formulaire. Conservez un exemplaire et remettez le second à votre Président(e) de club.

RENOUVELLEMENT LICENCE N° (8 chiffres) **Nouvelle licence**

CLUB

LIGUE **DÉPT**

CIVILITÉ **NOM** **PRÉNOM**

SEXE : FÉM. MASC. **NÉ(E) LE** **NATIONALITÉ**

ADRESSE

CODE POSTAL **VILLE** **PAYS**

TÉL. FIXE : + 33 (0) **TÉL. MOBILE** : + 33 (0)

E-MAIL

DANS L'INTÉRÊT DU DÉVELOPPEMENT DU BADMINTON FRANÇAIS, J'ACCEPTÉ QUE CES INFORMATIONS PUISSENT ÊTRE UTILISÉES PAR LA FÉDÉRATION À DES FINS D'INFORMATION.
 J'ACCEPTÉ QUE LA FÉDÉRATION TRANSMETTE MES COORDONNÉES À DES TIERS AFIN NOTAMMENT DE RECEVOIR DES OFFRES PRIVILÉGIÉES ET SÉLECTIONNÉES PAR LA FÉDÉRATION.
 Vous disposez d'un droit d'accès, de modification, de rectification et d'opposition à la divulgation des données vous concernant (loi « Informatique et Libertés » du 6 janvier 1978 et l'article 21 du RGPD).
 Pour toute demande, adressez-vous à la FFBaD : mail@ffbad.org
 Je ne souhaite pas que mes résultats sportifs soient accessibles sur les sites dédiés.

PREMIÈRE PRISE DE LICENCE ➤

RENOUVELLEMENT DE LICENCE ➤

JOINDRE LE CERTIFICAT MÉDICAL DE NON CONTRE-INDICATION ➤➤➤

REPLIR LE QUESTIONNAIRE DE SANTÉ ET JOINDRE UNE ATTESTATION ➤➤➤

Téléchargez **JCI** le formulaire unique et obligatoire du certificat médical

Téléchargez **JCI** le questionnaire de santé (formulaire Cerfa N°15699*01) et l'attestation

Rappel : Si vous avez répondu OUI à une ou plusieurs questions du questionnaire de santé, un certificat médical datant de moins d'un an doit obligatoirement être fourni pour le renouvellement.



LE MAGAZINE OFFICIEL DE LA FFBAD À SEULEMENT 2€ POUR 4 NUMÉROS

OUI, JE SOUHAITE RECEVOIR 100% BAD EN VERSION PAPIER ET AU FORMAT NUMÉRIQUE TÉLÉCHARGEABLE DEPUIS MA BOITE MAIL
 OUI, JE SOUHAITE RECEVOIR 100% BAD UNIQUEMENT SOUS FORMAT NUMÉRIQUE TÉLÉCHARGEABLE DEPUIS MA BOITE MAIL
Attention l'adresse mail est obligatoire sinon vous ne recevrez que la version papier par défaut

ASSURANCE :
 Je soussigné, atteste avoir été informé de l'intérêt que présente la souscription d'un contrat d'assurance de personne couvrant les dommages corporels auxquels la pratique sportive peut m'exposer (art. L321-4 du code du sport). Dans ce cadre, je déclare avoir reçu, pris connaissance et compris les modalités d'assurances présentées dans la notice d'information assurance FFBaD ci-annexée, et de la possibilité de souscrire à l'une des garanties d'assurance Accident Corporel facultatives proposées avec ma licence.
 Le prix de l'option de base incluse dans ma licence est de 0,35€ TTC. Je peux ne pas y adhérer et le signifier par écrit auprès de mon club à l'aide du formulaire de refus mis à ma disposition (procédure obligatoire).
 La souscription d'une option complémentaire A ou B se fait à l'aide du bulletin d'adhésion individuelle téléchargeable depuis le site internet de la FFBaD - rubrique assurance.

CONTRÔLE ANTIDOPAGE :
 Considérant le Code du sport - Titre III - Santé des sportifs et lutte contre le dopage,
 - Article L232-1 0-3 : « Il est interdit à toute personne de s'opposer par quelque moyen que ce soit aux mesures de contrôle prévues par le présent titre. »
 - Article L232-12 : « Les opérations de contrôle sont diligentées par le directeur du département des contrôles de l'Agence française de lutte contre le dopage. Les personnes mentionnées à l'article L. 232-11 ayant la qualité de médecin peuvent procéder à des examens médicaux cliniques et à des prélèvements biologiques destinés à mettre en évidence l'utilisation de procédés prohibés ou à déceler la présence dans l'organisme de substances interdites. Les personnes agréées par l'agence et assermentées peuvent également procéder à ces prélèvements biologiques. Seules les personnes mentionnées à l'article L. 232-11 et qui y sont autorisées par le code de la santé publique peuvent procéder à des prélèvements sanguins. »

Signature du licencié ou de son représentant légal
 Faire précéder la signature de la mention :
 "lu, compris et accepté"

DROIT A L'IMAGE :
 Lors de manifestations organisées par la FFBaD, une ligue, un comité et/ou un club affilié à la fédération, l'image et la voix du licencié, sont susceptibles d'être captées par tous moyens vidéos dans le cadre des activités de la FFBaD et ce, pendant toute la durée de la saison. Le signataire du présent formulaire autorise la FFBaD à procéder à des captations d'image et de voix et à utiliser et diffuser, pour la promotion du badminton, les images et les voix ainsi captées, sur tous supports de communication quels qu'ils soient. Cette autorisation est consentie pour la promotion de la FFBaD et du badminton à titre gratuit pour une durée de 70 ans et vaut pour le monde entier.

LES RESPONSABLES LÉGAUX DES ENFANTS MINEURS SONT TENUS DE REMPLIR L'AUTORISATION D'INTERVENTION CI-DESSOUS

JE SOUSSIGNÉ(E) : M. , MME , Mlle , **NOM**
AGISSANT EN QUALITÉ DE : PÈRE, MÈRE, TUTEUR, TUTRICE,
 autorise pour mon fils (ma fille, mon pupille, ma pupille) le responsable du club à faire intervenir les services de santé publics en cas d'accident corporel de l'enfant.

NOM DE L'ENFANT **PRÉNOM**
LE / /
SIGNATURE DU REPRÉSENTANT LÉgal

Liste des variables relatives à l'âge

0_4_Hommes	0_4_Femmes
5_9_Hommes	5_9_Femmes
10_14_Hommes	10_14_Femmes
15_19_Hommes	15_19_Femmes
20_24_Hommes	20_24_Femmes
25_29_Hommes	25_29_Femmes
30_34_Hommes	30_34_Femmes
35_39_Hommes	35_39_Femmes
40_44_Hommes	40_44_Femmes
45_49_Hommes	45_49_Femmes
50_54_Hommes	50_54_Femmes
55_59_Hommes	55_59_Femmes
60_64_Hommes	60_64_Femmes
65_69_Hommes	65_69_Femmes
70_74_Hommes	70_74_Femmes
75_79_Hommes	75_79_Femmes
80_84_Hommes	80_84_Femmes
85_89_Hommes	85_89_Femmes
90_94_Hommes	90_94_Femmes
95_plus_Hommes	95_plus_Femmes

Liste des variables relatives au diplôme

Aucun diplôme ou au mieux BEPC, brevet des collèges, DNB Hommes 16 à 24 ans	Aucun diplôme ou au mieux BEPC, brevet des collèges, DNB Femmes 16 à 24 ans
Aucun diplôme ou au mieux BEPC, brevet des collèges, DNB Hommes 25 ans ou plus	Aucun diplôme ou au mieux BEPC, brevet des collèges, DNB Femmes 25 ans ou plus
Diplôme de niveau CAP, BEP Hommes 16 à 24 ans	Diplôme de niveau CAP, BEP Femmes 16 à 24 ans
Diplôme de niveau CAP, BEP Hommes 25 ans ou plus	Diplôme de niveau CAP, BEP Femmes 25 ans ou plus
Diplôme de niveau Baccalauréat (général, techno, pro) Hommes 16 à 24 ans	Diplôme de niveau Baccalauréat (général, techno, pro) Femmes 16 à 24 ans
Diplôme de niveau Baccalauréat (général, techno, pro) Hommes 25 ans ou plus	Diplôme de niveau Baccalauréat (général, techno, pro) Femmes 25 ans ou plus
Diplôme d'études supérieures Hommes 16 à 24 ans	Diplôme d'études supérieures Femmes 16 à 24 ans
Diplôme d'études supérieures Hommes 25 ans ou plus	Diplôme d'études supérieures Femmes 25 ans ou plus

Listes des catégories socio-professionnelles

Agriculteurs
Artisans, commerçants, chefs d'entreprise
Cadres et professions intellectuelles supérieures
Professions intermédiaires
Employés
Ouvriers

Ellipse de déviation standard

Calcul de l'ellipse de déviation standard :

$$\theta = \arctan \left[\left(\sum_i n_i (x_i - \bar{x})^2 - \sum_i n_i (y_i - \bar{y})^2 \right) + \frac{\sqrt{(\sum_i n_i (x_i - \bar{x})^2 - \sum_i n_i (y_i - \bar{y})^2)^2 + 4 \sum_i (x_i - \bar{x})(y_i - \bar{y})}}{2 \sum_i (x_i - \bar{x})(y_i - \bar{y})} \right]$$

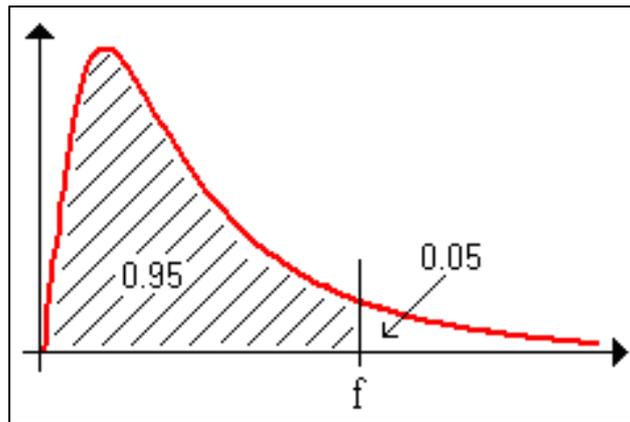
Source : JM. Zaninetti ; *Statistique Spatiale, méthodes et applications géomatiques* ; 2005.

Loi de Fisher

Définition : « On dit qu'une variable aléatoire X suit la loi de Fisher-Snedecor de paramètres m et n si elle admet une densité qui vaut : »

$$f(x) = \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} n^{n/2} m^{m/2} \frac{x^{m/2-1}}{(mx+n)^{m+n/2}}$$

Courbe de densité de la loi de Fisher :



(Source : Bibmath.net)

Test de Shapiro-Wilk

« La statistique de test s'écrit :

$$W = \frac{[\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} a_i(x_{(n-i+1)} - x_{(i)})]^2}{\sum_i (x_i - \bar{x})^2}$$

où

- $x_{(i)}$ correspond à la série des données triées
- $\lfloor \frac{n}{2} \rfloor$ est la partie entière du rapport $\frac{n}{2}$
- a_i sont des constantes générées à partir de la moyenne et de la matrice de variance covariance des quantiles d'un échantillon de taille n suivant une loi normale. Ces constantes sont fournies des tables spécifiques.

La statistique W peut donc être interprétée comme le coefficient de détermination entre la série des quantiles générées à partir de la loi normale et les quantiles empiriques obtenues à partir des données. »

Source : R. Rakotomalala ; *Test de normalité* ; Université de Lyon 2 ; 2011.

Test de Levene

« La statistique du test de Levene est :

$$W = \frac{n - K}{K - 1} * \frac{\sum_{k=1}^K n_k (\bar{Z}_k - \bar{Z})^2}{\sum_{k=1}^K \sum_{i=1}^{n_k} ((Z_k)_i - \bar{Z}_k)^2}$$

Avec $(Z_k)_i = \left| (X_{Y=k})_i - \bar{X}_k \right|$ et $\bar{Z} = \frac{\sum_{k=1}^K \sum_{i=1}^{n_k} ((Z_k)_i)}{K}$

La statistique de test de Levene suit une loi de Fisher-Snedecor. »

Source : Wordpress.com

Test de Kruskal-Wallis

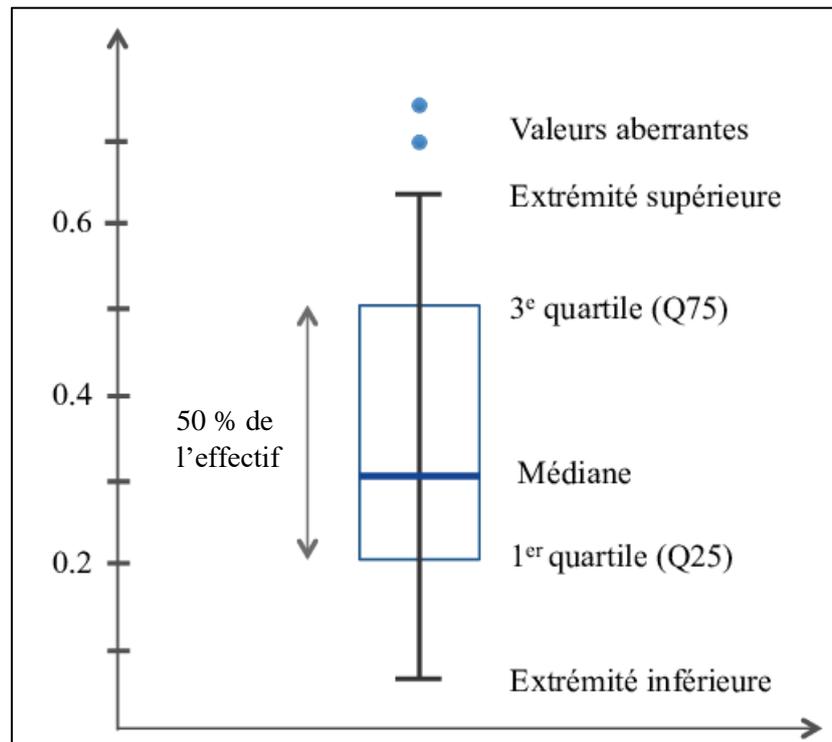
« La statistique du test de Kuskal-Wallis est définie de la manière suivante :

$$H = \frac{12}{n(n+1)} \sum_{k=1}^K n_k (\bar{r}_k - \bar{r})^2$$

Avec \bar{r} la moyenne globale des rangs, et \bar{r}_k la moyenne des rangs pour les observations du groupe k. »

Source : R. Rakotomalala ; *Comparaison de population*; Université de Lyon 2 ; 2008.

Explication Boxplot



Source : Researchgate.net.

Distance des plus proches voisins

Nearest neighbor analysis (2008-2009)

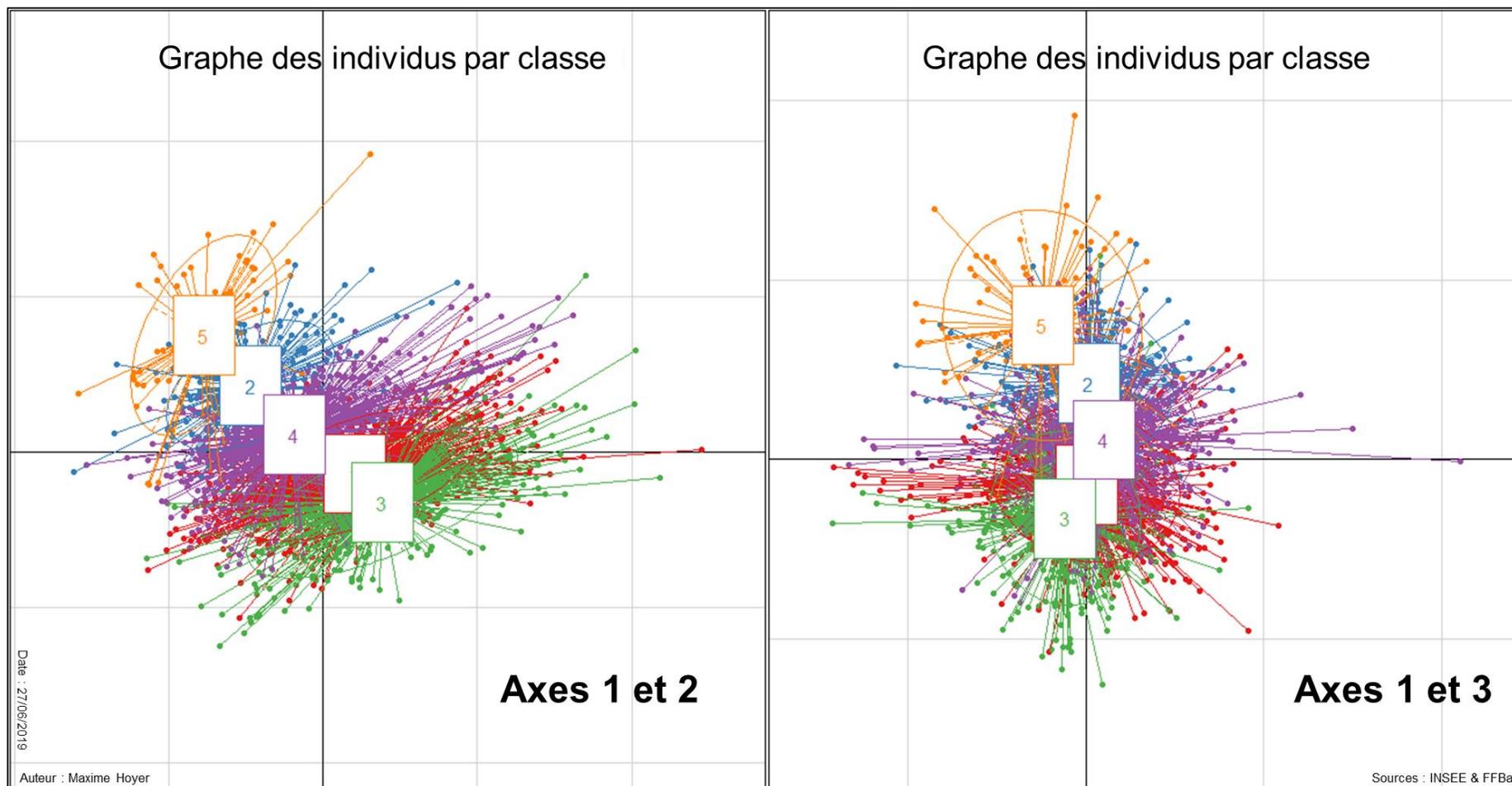
Mean Nearest Neighbor Distance ..: 7735.62 m

Standard Dev of Nearest

Neighbor Distance: 8022.54 m

Minimum Distance: 0.00 m

Maximum Distance: 1282059.58 m



Test de Shapiro-Wilk

Shapiro-Wilk normality test

W = 0.70884, p-value < 2.2e-16

Test de Levene

Levene's Test for Homogeneity of Variance

p-value < 0.00008726 ***

Test de Kruskal-Wallis

Kruskal-Wallis rank sum test

Kruskal-Wallis chi-squared = 95.809, df = 4, p-value < 2.2e-16

Linear Hypotheses:

	Estimate	Std. Error	t	value	Pr(> t)
2 - 1 == 0	0.29077	0.15625	1.861	0.32208	
3 - 1 == 0	-0.41094	0.12688	-3.239	0.00958	**
4 - 1 == 0	0.37439	0.11227	3.335	0.00703	**
5 - 1 == 0	0.53628	0.23590	2.273	0.14314	
3 - 2 == 0	-0.70171	0.16719	-4.197	< 0.001	***
4 - 2 == 0	0.08362	0.15640	0.535	0.98247	
5 - 2 == 0	0.24551	0.25981	0.945	0.87136	
4 - 3 == 0	0.78533	0.12706	6.181	< 0.001	***
5 - 3 == 0	0.94722	0.24328	3.893	< 0.001	***
5 - 4 == 0	0.16189	0.23599	0.686	0.95649	

Linear Hypotheses:

	Estimate	lwr	upr
2 - 1 == 0	0.29077	-0.13083	0.71237
3 - 1 == 0	-0.41094	-0.75329	-0.06858
4 - 1 == 0	0.37439	0.07146	0.67732
5 - 1 == 0	0.53628	-0.10021	1.17278
3 - 2 == 0	-0.70171	-1.15283	-0.25059
4 - 2 == 0	0.08362	-0.33837	0.50560
5 - 2 == 0	0.24551	-0.45552	0.94653
4 - 3 == 0	0.78533	0.44250	1.12815
5 - 3 == 0	0.94722	0.29079	1.60364
5 - 4 == 0	0.16189	-0.47486	0.79864

1 2 3 4 5

"b" "bc" "a" "c" "bc"