

Master 2 Géomarketing – Université Paris Est Créteil Val-de-Marne
Année universitaire 2013 – 2014

Mémoire de recherche appliquée



Nouvelle concurrence : quand l'assurance adapte sa stratégie d'implantation

*Présenté et soutenu par
Sophie STAËS*



Maitres de stage :
Marc Giannasi
Jean-François Jobert

Mémoire dirigé par Albert Da Silva
IAE Gustave Eiffel, Université Paris Est Créteil

AXA France
Direction Distribution
Administration Réseaux

REMERCIEMENTS

Avant tout je tiens à remercier chaque personne sans qui la réalisation de ce mémoire de recherche appliquée n'aurait pas été possible.

J'aimerais remercier en premier lieu Monsieur Christophe Davoine (Responsable Administration Réseaux), Madame Céline Commissaire (Responsable pilotage et Administration Réseaux), ainsi que Monsieur Marc Giannasi (Responsable du service Géomarketing et Chiffres Réseaux), pour m'avoir intégrée à leur équipe et m'avoir permise de bénéficier d'un encadrement ainsi que de précieux conseils et explications. Je n'oublie pas Monsieur Jean-François Jobert qui a pris la suite de Monsieur Giannasi et avec qui j'ai pris plaisir à travailler.

Mes remerciements s'adressent également à mes collègues de travail : Joëlle, Dominique, Claudine, Annie et Evelyne

Je tiens à remercier Monsieur Albert Da Silva, directeur de mémoire, pour s'être toujours montré disponible et à l'écoute mais aussi pour le temps qu'il a accepté de me consacrer.

J'adresse un remerciement particulier à mes camarades de promotion du Master 2 Géomarketing avec qui il a été possible d'échanger et de s'entraider.

Je n'oublie pas mon compagnon qui m'a toujours soutenue et encouragée au cours de la réalisation de ce travail ainsi que pour son aide précieuse.

Enfin, mes dernières pensées vont vers l'ensemble des collaborateurs AXA France Particuliers/Professionnels qui m'ont accueillie dans les meilleures conditions durant cette année universitaire mais aussi toutes les personnes ayant gentiment accepté de m'aider à la relecture et à la correction de ce mémoire.

CLAUSE DEONTOLOGIQUE

« Je certifie sur l'honneur que le présent mémoire est le fruit d'un travail personnel et que toute référence directe ou indirecte aux travaux de tiers est expressément indiquée. Je demeure seul(e) responsable des analyses et opinions exprimées dans ce document : l'université Paris-Est Créteil n'entend y donner aucune approbation ni improbation ».

RÉSUMÉ

Les sociétés d'assurance évoluent sur un marché non seulement attractif mais aussi totalement saturé. En effet, aujourd'hui chacun dispose déjà de ses différents contrats d'assurance, certains étant notamment obligatoires. Il est donc difficile d'acquérir de nouveaux clients et l'apparition progressive de nouveaux acteurs sur le marché ne facilite pas la tâche. Une véritable bataille se met alors en place pour gagner en parts de marché, notamment par l'optimisation des réseaux de distribution qui constituent le point de contact entre le client et la société. C'est grâce à la connaissance du territoire, ainsi que des différences économiques et comportementales de la population, que les sociétés d'assurance répondent en partie à cette vague de nouveaux concurrents. Le but final étant de mieux répartir leurs agents généraux sur le territoire francilien. Cependant, si certains modèles d'implantation existent aujourd'hui, il est encore difficile de n'en prendre qu'un en compte dans la prise de décision d'une implantation. Bien souvent le couplage de plusieurs d'entre eux permet à une étude géomarketing d'être vraiment efficace.

ABSTRACT

Insurance companies operate on an attractive market but also completely saturated. Indeed, today everybody has already their insurance contracts, some of them are mandatory. So, it is difficult to acquire new customers and the gradual emergence of new players in the market doesn't facilitate the task. A real battle begins to gain market share, by optimization of networks distribution which are the point of contact between the customer and the company. Through the expertise of the territory as well as economic and behavioural differences of population, those insurance companies respond in part to arrival of new competitors. The purpose is to allocate their general agents in the Paris area. However, if some implementation models exist today, it is still difficult to use only one in the decision making of an implementation. Often connect several of them allows a geomarketing study be truly effective.

SOMMAIRE

Remerciements	2
Clause déontologique	3
Résumé / Abstract	4
Sommaire	5
Introduction	6
1. La situation du marché de l'assurance : AXA face à ses concurrents	8
1.1 Introduction	8
1.2 Le réseau de points de vente comme vitrine historique des sociétés d'assurance	9
1.3 La pression concurrentielle	12
1.4 Conclusion	21
2. La prise en compte des réalités géographiques : contexte d'une étude de territoire en région	23
2.1 Introduction	23
2.2 Les données disponibles	23
2.3 L'Ile-de-France	27
2.4 Le maillage du territoire francilien	36
2.5 Conclusion	39
3. L'optimisation du territoire grâce à l'analyse du marché	40
3.1 Introduction	40
3.2 Le profil socio-économique de la région	41
3.3 La consommation de produits d'assurance	52
3.4 Conclusion	64
4. L'intérêt du modèle p-médian dans l'amélioration des études géomarketing	65
4.1 Introduction	65
4.2 Le modèle p-médian	65
4.3 Conclusion	73
Conclusion	74
Bibliographie / Webographie	76
Table des illustrations, tableaux et figures	77
Table des annexes	79
Annexes	80
Table des matières	82

INTRODUCTION

En 2013, le chiffre d'affaire du secteur de l'assurance en France s'élevait à 180,7 milliards d'euros¹, soit 9,3 milliards de moins qu'en 2011 (190Md€). Si les agents généraux sont largement majoritaires en assurance dommages, leur part de marché a diminué de 10% depuis les années 90 en passant de 45% en 1992² à 35% en 2010³. Ils ont plus de mal à exister sur le marché des produits financiers, et principalement celui de l'assurance vie pour lequel ils ne détiennent aujourd'hui que 8% (17% en 1992).

A l'origine de cette perte de marché ; l'apparition et le développement de nouveaux canaux de distribution tels que les mutuelles sans intermédiaires ou encore la bancassurance. En effet, alors que les premiers bancassureurs sont le Crédit Mutuel et le Crédit Agricole dans les années 80, en 2009 les bancassurances détiennent 36% du marché des produits d'assurance dans leur globalité⁴, les courtiers 18%, les mutuelles d'assurance 13% et seulement 11% pour les agents généraux. Si en 1992, la part de marché des réseaux bancaires est déjà de 46% pour l'assurance vie, elle semble aujourd'hui se stabiliser aux alentours de 60% depuis 1997. Enfin pour ce qui est de l'assurance dommages, bien que les bancassureurs ne disposent que de 14.5% de part de marché en 2010⁵, leur ascension est très régulière ce qui est remarquable sur un marché saturé.

Ce sont les réseaux traditionnels et notamment les agents généraux qui sont victimes de cette croissance. Ainsi il y a une véritable nécessité de s'adapter. On assiste alors à l'apparition des assurbanques⁶ en 2000 avec Allianz Banque⁷, en réponse aux bancassurances. Contrairement à cette dernière l'assurbanque doit encore faire ses preuves. Mais ce sont surtout les problématiques d'implantation des agences qui sont de plus en plus importantes. Il y a un réel besoin de déterminer les zones à fort potentiel ainsi que d'identifier les

¹ SCHINDLER Patrick (2013), *Analyse du secteur de l'assurance*

² [http://www.ffsa.fr/webffsa/risques.nsf/b724c3eb326a8defc12572290050915b/e3bf7156c14eb30fc12573a70042009a/\\$FILE/Risques_60_0027.htm](http://www.ffsa.fr/webffsa/risques.nsf/b724c3eb326a8defc12572290050915b/e3bf7156c14eb30fc12573a70042009a/$FILE/Risques_60_0027.htm) (10 avril 2014)

³ <http://www.linternaute.com/argent/assurance/conseil/agents-generaux-de-l-assurance/l-agent-general-s-occupe-avant-tout-de-la-sante-ou-de-l-automobile-et-peu-d-assurance-vie.shtml> (12 avril 2014)

⁴ <http://www.latribune.fr/vos-finances/assurance/20110728trib000639368/le-palmars-des-assureurs-les-plus-importants-de-france.html> (12 avril 2014)

⁵ <http://www.argusdelassurance.com/tendances/l-irresistible-ascension-des-bancassureurs.56949>

⁶ « L'assurbanque est un néologisme qui désigne le fait pour certaines sociétés d'assurance de proposer à leurs clients une palette de services bancaires » d'après www.definition-marketing.com

⁷ <http://www.argusdelassurance.com/distribution/diversification-l-assurbanque-doit-encore-faire-ses-preuves.34365> (12 avril 2014)

emplacements intéressants et les zones déjà saturées. Le sujet de ce mémoire de recherche est donc dans un premier temps d'expliquer le besoin de retravailler la couverture du réseau d'agences d'assurance dans le but d'être le plus visible possible et de toucher un maximum de clients.

La géomercatique, ou géomarketing en anglais, terme qui sera répété dans le mémoire suivant car il s'agit du plus utilisé, est une discipline capable de répondre à une telle problématique de restructuration des réseaux puisqu'il s'agit d'une branche du marketing qui consiste en l'intégration des notions d'espace dans les études. Dans le cas présent, les études sont de plus en plus complexes car il faut être partout, tout en optimisant les nouvelles implantations de façon à ne pas cannibaliser le réseau déjà existant.

Ce mémoire propose d'apporter une contribution dans l'analyse de la consommation des produits d'assurance et ainsi dans la recherche d'un modèle d'étude d'implantation en Ile-de-France.

Ce travail sera donc organisé en quatre chapitres. Le premier développera le fonctionnement et la situation du secteur de l'assurance de nos jours, il placera le cadre du mémoire en présentant l'intérêt d'un réseau de points de vente ainsi que la diversité de la pression concurrentielle, ceci du point de vue d'AXA, premier assureur mondial en terme de chiffre d'affaire⁸ (91 249 millions d'euros⁹ en 2013). Le second chapitre présentera le contexte d'une étude de territoire en Ile-de-France ainsi que les données disponibles en région pour mener de telles recherches. Le troisième étudiera les méthodes d'optimisation de maillage territorial, par l'analyse de la consommation des produits d'assurance en Ile-de-France grâce notamment au logiciel Tanagra. Enfin le dernier chapitre traitera de l'utilisation du modèle p-médian dans l'amélioration des études d'implantation ainsi que les limites de celui-ci.

⁸ <http://www.journaldunet.com/economie/finance/assureurs/> (12 avril 2014)

⁹ <http://www.axa.com/fr/groupe/profil-chiffres-cles/chiffre-affaires/> (02 septembre 2014)

1. LA SITUATION DU MARCHÉ DE L'ASSURANCE : AXA FACE À SES CONCURRENTS

Au sein même des grandes sociétés d'assurance, il existe un certain nombre d'acteurs qui interviennent chacun dans des modes de distributions différents. On retrouve ainsi le courtier, le collaborateur salarié, l'agent mandataire et l'agent général (avec ses collaborateurs d'agence).

1.1. Introduction

L'agent général d'assurance est défini comme une personne mandataire d'une société d'assurance. Celui-ci va représenter l'assureur mais aussi lui réserver l'exclusivité de sa production. L'agent général est lié à la société par un traité de nomination. Intermédiaire entre les clients et l'entreprise, il est responsable d'un point de vente et c'est principalement autour de celui-ci que s'orientent les questions d'implantation. On compte aujourd'hui, toutes compagnies confondues, environ 26 000¹⁰ points de vente d'assurance en France, et ce sont ces agences qui constituent la première visibilité d'une société auprès de ses clients, mais aussi la première porte ouverte aux clients potentiels. C'est pour cette raison que le travail d'optimisation du réseau se fera sur ce mode de distribution.

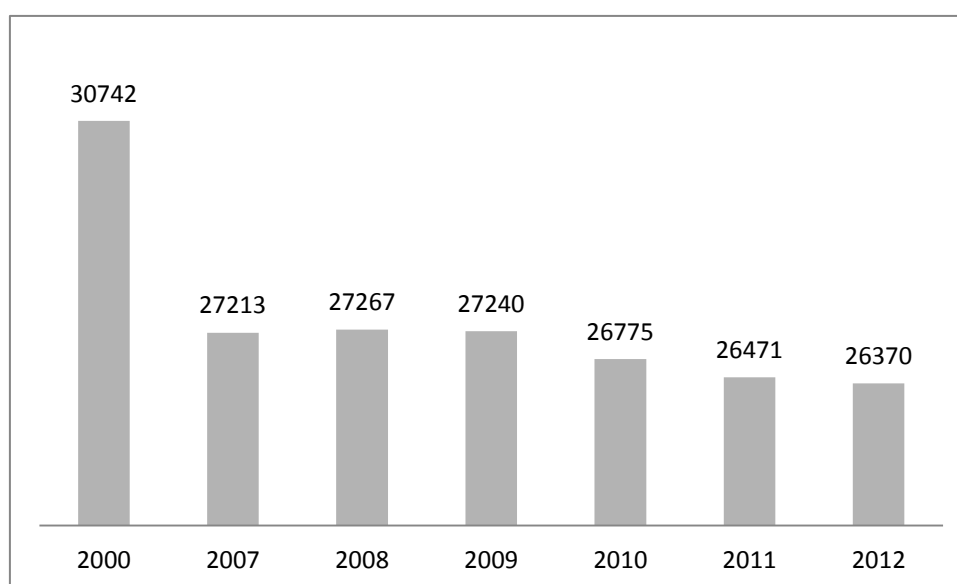
Dans cette première partie vont être présentés tout d'abord les intérêts d'un point de vente afin, plus tard, d'expliquer pourquoi il est primordial de bien travailler sur leur lieu d'implantation. De plus la concurrence dans le secteur de l'assurance est de plus en plus forte, en effet les acteurs sur ce marché sont de plus en plus nombreux, il est donc nécessaire de bien connaître celle-ci avant de se lancer dans une étude marketing ou géomarketing. Cette partie propose donc de présenter chacun des acteurs du marché de l'assurance avant de mieux en comprendre tous les aspects.

¹⁰ <http://www.infostat-marketing.com/implantation-des-assurances-analysez-la-concurrence.html>

1.2. Le réseau de points de vente comme vitrine historique des sociétés d'assurance

L'agent général fait son apparition au XIXe siècle¹¹ quand les sociétés comprennent le véritable intérêt stratégique d'être directement sur le terrain, c'est à cette époque que les agences physiques voient le jour. L'agent a alors pour vocation de créer un lien entre la société et l'assuré, de mettre à disposition son expertise et de faire de la prospection.

Figure 1 : Nombre de points de vente d'assurance en France



Infographie Sophie STAËS, d'après infostat-marketing.com¹², réalisé sous Excel

Comme il est lisible sur la figure 1, au début des années 2000 le nombre de points de vente d'assurance était de plus de 30 000. On remarque qu'ils sont de moins en moins nombreux au fil des années puisqu'en 2012 on n'en compte plus que 26 000. En effet, alors que les agents généraux étaient encore 25 000 au début des années 80, ils sont environ moitié moins (12 700)¹³ aujourd'hui (l'évolution des nombres d'agents généraux et de points de vente ne sont pas proportionnels car un agent peut tout à fait être propriétaire de plusieurs points de vente).

Une lente diminution due à la saturation du marché sur lequel ils sont positionnés, qui s'associe à l'entrée de nouveaux canaux de distribution, a eu pour effet de réduire progressivement leurs parts de marché. Et puisqu'ils sont moins nombreux, il est donc primordial d'optimiser leurs implantations car ils ont de nombreux avantages.

¹¹ MARTIN Pierre, *Petite histoire des agences d'assurance : au service du risque*

¹² <http://www.infostat-marketing.com/implantation-des-assurances-analysez-la-concurrence.html> (4 mai 2014)

¹³ <http://www.eurogrouppconsulting.fr/IMG/pdf/P10-2010-2.pdf> (4 mai 2014)

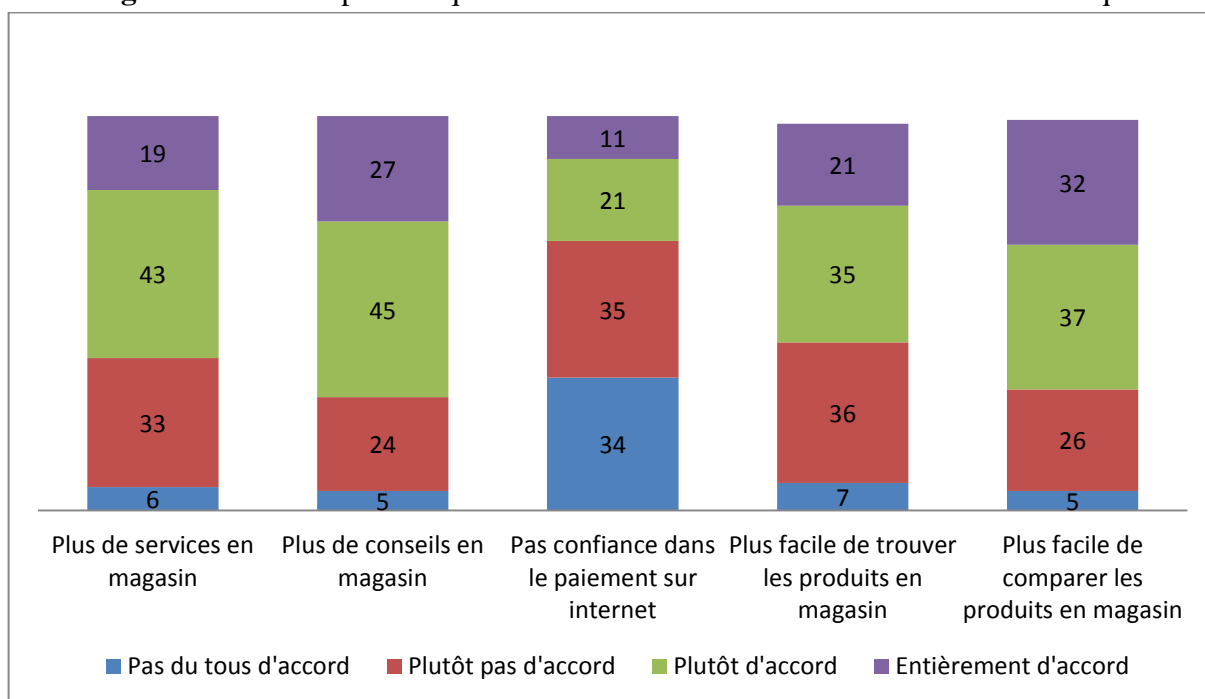
1.2.1. L'avantage de la présence physique

Ce réseau a de nombreux atouts pour la société d'assurance :

- Il s'agit du réseau le plus expert en assurance (suivi par les courtiers), ainsi l'agent est en mesure de répondre aux besoins du client.
- Il dispose d'une présence physique sur le terrain qui lui permet de créer une relation avec ses clients et futurs assurés. De cette façon il développe son portefeuille de contrats par la prospection.
- Il rassure le client par sa proximité, veille à la souscription des contrats, la gestion des sinistres et au règlement des indemnités.

L'avantage de la présence physique pour une agence d'assurance est approximativement la même que pour une boutique lambda. Le client y recherche le contact humain et selon le sondage Ipsos du 29 avril 2013, 72% des français préféreraient effectuer leurs achats en boutique plutôt que sur internet.

Figure 2 : Raisons pour lesquelles un client effectue ses achats dans une boutique



Infographie Sophie STAËS (2014), d'après les données du site Le café du commerce¹⁴, réalisé sous Excel

¹⁴<http://www.marquesetreseaux.com/2009/06/e-commerce-et-points-de-vente-linternet-de-la-distribution-multicanale/> (12 avril 2014)

La figure 2 est réalisée à partir de réponses non adaptées aux boutiques d'assurance, il n'a donc été gardé que certains critères. On remarque donc que la moitié des clients (45%) pense que le conseil est meilleur dans un point de vente que sur internet, de même que les services (43%). De plus la majorité d'entre eux (69%) ne sont pas effrayés à l'idée de faire des achats sur internet. Il est donc imaginable qu'il en est de même pour les souscriptions de contrats d'assurance. On en déduit donc que les clients se rendant en agence viennent y trouver de précieux conseils sur les produits, ils sont en effet convaincus qu'il sera plus facile d'y trouver le contrat d'assurance le plus adapté à leur besoins, mais aussi des services.

Aujourd'hui les points de vente sont organisés en réseaux dont le but est de couvrir l'ensemble du territoire pour prospecter un maximum de clients, ils sont un véritable atout et chacun redouble d'ingéniosité pour les perfectionner et choisir leurs implantations. Ainsi chaque société d'assurance a développé le sien avec un nombre d'agents généraux plus ou moins élevé.

1.2.2. La visibilité

L'un des buts premier d'une bonne implantation d'agence est la visibilité de l'agent général mais aussi de l'image de la compagnie. Non seulement l'implantation doit être extrêmement bien choisie mais l'intérêt est aussi d'être plus visible que ses concurrents. Ainsi les réseaux de points de vente sont particulièrement étudiés et chaque compagnie cherche à être présente partout.

Tableau 1 : Classement par nombre d'agents et de points de vente en France

1 ^{er}	AXA	3 495 agents et 3 757 points de vente
2 ^{ème}	Allianz	1 962 agents et 2 539 points de vente
3 ^{ème}	MMA	1 318 agents et 1 847 points de vente
4 ^{ème}	Gan assurance	960 agents et 1 046 points de vente
5 ^{ème}	Generali	952 agents et 1 060 points de vente

Réalisé d'après lalettredeassurance.fr (2014)¹⁵

Premier assureur en nombre d'agents généraux, AXA est loin devant Allianz, le second, ce qui lui offre un véritable avantage en termes de visibilité. Cette avance est notamment ressentie dans le chiffre d'affaire de l'assureur puisqu'il est de 91 249 Millions d'euros en 2013, ce qui maintient AXA au rang de premier assureur européen (vie et non vie

¹⁵ <http://www.lettredeassurance.com/2012/04/zoom-sur-les-reseaux-d%E2%80%99agents-generaux/> (12 avril 2014)

confondus). Alors que depuis 2010, le nombre de points de vente qui constituent le réseau d'AXA a légèrement diminué (en 3 ans perte de 95 agents et 89 points de vente¹⁶), son chiffre d'affaire France, lui, diminue peu (22.2 milliards d'euros en 2010, contre 21.2 milliards¹⁷ en 2012). Cependant au niveau mondial, celui-ci a considérablement augmenté (environ 85 000 millions d'euros¹⁸ en 2010, contre plus de 90 000 millions en 2013). Si bien entendu le chiffre d'affaire dans sa globalité ne provient pas exclusivement des agents généraux, il est tout de même utile de savoir que depuis quelques années AXA restructure activement son réseau de points de vente en travaillant sur leurs implantations afin d'être plus visible, plus présent et plus rentable. En effet les agents généraux sont à l'origine de 40% du chiffre d'affaire d'AXA France¹⁹, soit approximativement 7 milliards d'euros.

Avec la saturation du marché, l'entrée de nouveaux acteurs et par conséquent la tendance à la diminution du nombre de points de vente, l'objectif n'est plus aujourd'hui d'avoir le réseau le plus dense, mais d'être le plus visible possible avec moins d'agences. En effet, on le sait, le comportement d'achat est influencé par la visibilité et l'image de la marque.

Ainsi la recherche non seulement d'une forte présence dans les agglomérations mais aussi d'une présence efficace est indispensable au développement d'une société, c'est pourquoi l'étude d'implantation semble primordiale afin de se positionner face à ses concurrents.

1.3. La pression concurrentielle

De façon générale, le marché de l'assurance dispose d'une large gamme de produits et de prix, ce qui en fait un marché très en proie à la concurrence. Si avant les années 60 et l'entrée des nouveaux acteurs sur le marché des assurances, la concurrence était bien moins sévère qu'aujourd'hui, elle existait cependant entre les grands groupes d'assureurs.

Depuis les années 60, ce secteur a considérablement évolué notamment par l'accroissement du nombre d'acteurs sur ce marché. On assiste alors et au fur et à mesure des

¹⁶ <http://www.lalettredelassurance.com/2012/04/zoom-sur-les-reseaux-d%E2%80%99agents-generaux/> (17 avril 2014)

¹⁷ Rapports annuels AXA France au 31dec 2010 et 31dec 2012

¹⁸ <http://www.argusdelassurance.com/a-la-une/la-france-4eme-assureur-mondial-et-axa-1er-groupe-europeen.54428> (17avril 2014)

¹⁹ <http://www.lalettredelassurance.com/2011/04/12-000-agents-generaux-d%E2%80%99assurances/> (18avril 2014)

années à la montée d'une concurrence accrue multicanale. Alors qu'au début de la distribution de produits d'assurance les grandes compagnies spécialisées se concurrençaient uniquement entre elles, elles le sont aujourd'hui par de nouveaux entrants, étrangers au marché tels que les mutuelles sans intermédiaires (dans les assurances dommages), la bancassurance ou la vente directe sur internet. C'est donc un bouleversement dans l'évolution des parts de marché.

1.3.1. Les mutuelles sans intermédiaires

Les mutuelles d'assurance (aussi appelée assurances mutuelles) font partie de la grande famille des mutuelles au même titre que les mutuelles de santé et ont fait leur apparition dans les années 30 avec la création de la MAIF²⁰. C'est cependant depuis les années 60 qu'elles se sont réellement développées (1960 création de la Macif²¹, 1983 la mutuelle des motards²²...). Ces mutuelles sont dites sans intermédiaires lorsqu'elles ne rémunèrent ni agents généraux ni courtiers car elles sont propriétaires et responsables de leurs propres agences. On dit alors qu'elles ont un réseau intégré : elles contrôlent l'ensemble de leurs points de vente (à la différence des sociétés d'assurance qui vendent leurs contrats par l'intermédiaire de salariés, courtiers ou agents généraux à qui appartiennent les locaux : réseau non intégré). Les mutuelles sans intermédiaires sont des sociétés civiles soumises au code des assurances et proposent des assurances dommages.

En 1964 on assiste à la création du groupe GEMA, le Groupement des Entreprises Mutuelles d'Assurance (qui s'appelait à l'époque GSACM²³) par les sociétés GMF²⁴, MAAF²⁵, MAIF, Macif et Matmut suite au refus de leur demande d'adhésion à la Fédération Française des Sociétés d'Assurances (FFSA). Le principe d'une mutuelle d'assurance est d'être une société à but non lucratif qui propose un fonctionnement solidaire en assurant les risques grâce aux cotisations de ses sociétaires. Les clients ne sont pas dits « assurés » mais « sociétaires » car ils ont la possibilité de s'investir dans la vie de la société. En effet ils peuvent voter pour les délégués qui vont les représenter à l'assemblée générale ou se porter candidat pour être eux même délégués.

²⁰ Mutuelle d'assurance des instituteurs de France

²¹ Mutuelle assurance des commerçants et industriels de France

²² http://www.alternatives-economiques.fr/mutuelle-d-assurances_fr_art_223_31283.html (16 avril 2014)

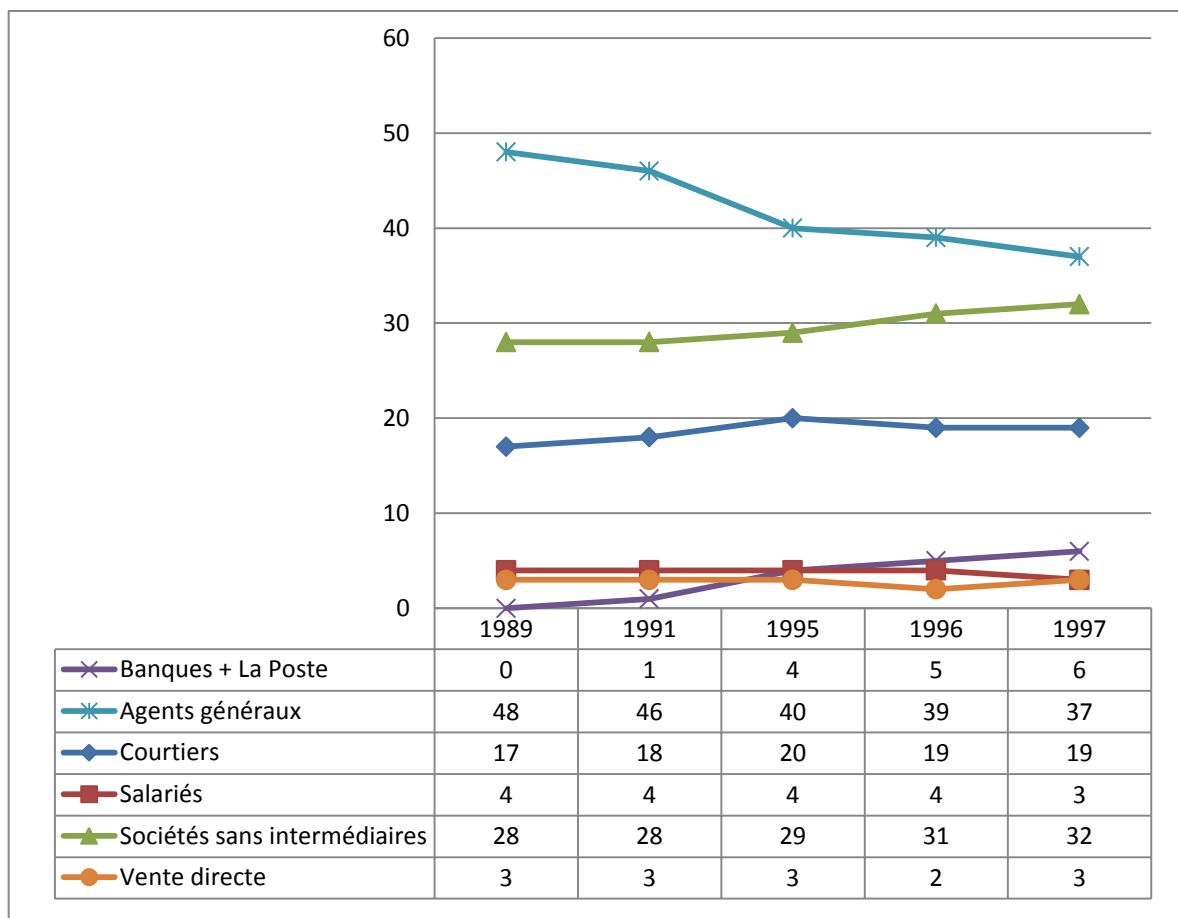
²³ Groupement des Sociétés d'Assurance à Caractère Mutuel

²⁴ Garantie mutuelle des fonctionnaires

²⁵ Mutuelle d'assurance des artisans de France

Grâce à leur mode de fonctionnement, les mutuelles sans intermédiaires proposent des tarifs de 20 à 30% inférieurs aux compagnies d'assurance traditionnelles. C'est donc sans difficulté qu'elles se sont approprié des parts de marché en constante évolution. Ce sont d'ailleurs les seules sur le marché de l'assurance dommages à voir leurs parts de marché augmenter depuis le début des années 90 (avec les banques qui débutent cependant sur ce marché), comme le montre la figure 3. De plus, cela semble être au détriment des agents généraux qui présentent les plus fortes diminutions de parts de marché.

Figure 3 : Evolution des parts de marché dans l'assurance dommages



Infographie Sophie STAËS (2014), d'après [senat.fr](http://www.senat.fr)²⁶, réalisé sous Excel

Les Mutuelles d'assurance viennent ainsi de dépasser les 20 millions de sociétaires²⁷ en assurance dommages et elles se développent aujourd'hui de plus en plus sur le marché de l'assurance vie dont elles ne détiennent pour le moment que 3% du marché²⁸.

²⁶ <http://www.senat.fr/rap/r98-045/r98-0458.html> (16 avril 2014)

²⁷ Chiffres du GEMA au 31 décembre 2013

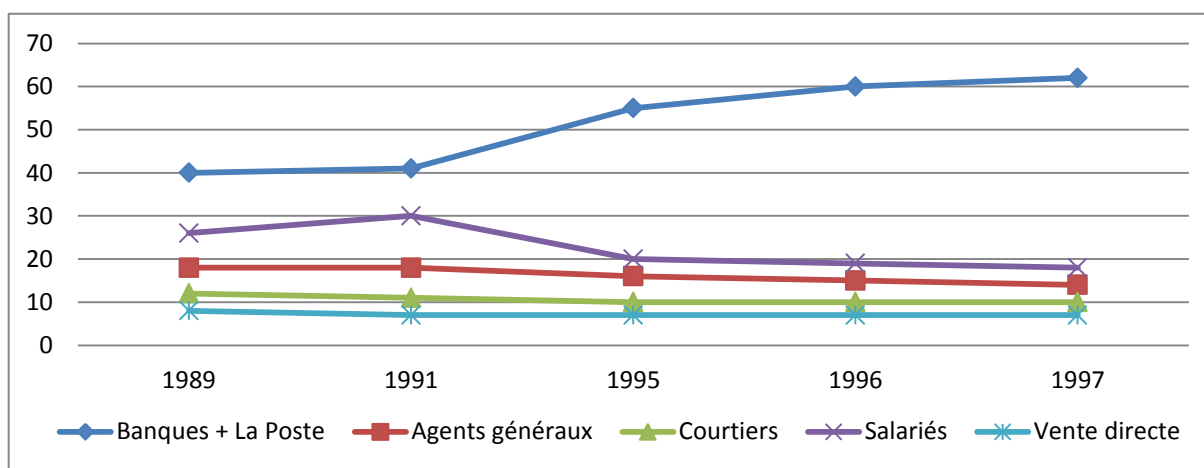
1.3.2. Le développement de la bancassurance

La bancassurance est un terme qui désigne la commercialisation de produits d'assurance par les réseaux bancaires. Ce mode de distribution existe depuis 1972 avec la création de la filiale « Les Assurances du Crédit Mutuel » (assurance vie et capitalisation) par le Crédit Mutuel de l'Est. En 1984, l'article 5 de la loi bancaire n° 84-46 autorise la commercialisation par les banques de produits d'assurance vie²⁹. C'est alors que va débuter une forte croissance de la bancassurance en France, les produits bancaires étant à cette époque moins rentables que le marché des produits d'assurance vie.

L'assurance vie permet d'épargner, de transmettre son capital, de préparer un projet ou de disposer d'un complément de revenus (définition AXA). Le contrat de capitalisation est complémentaire à l'assurance vie, il s'agit d'un système de placement financier.

En quelques années la part de marché des bancassurances dans les produits d'assurance vie-capitalisation a explosé. Grâce à la figure 4, il est remarquable que les banques sont donc non seulement celles qui détiennent les plus fortes parts de marché dans le domaine de l'assurance vie-capitalisation, mais aussi les seules à avoir connu une évolution positive entre 1989 et 1997. Cela signifie donc qu'à leur arrivée sur le marché, les bancassureurs se sont approprié une très forte part de marché au détriment des concurrents.

Figure 4 : Evolution des parts de marché dans l'assurance vie-capitalisation



Infographie Sophie STAËS (2014), d'après senat.fr³⁰, réalisé sous Excel

²⁸ <http://clubassurance.blogspot.com/archive/2008/10/20/dossier-special-les-mutuelles-sans-intermediaire-et-l-assura.html> (16 avril 2014)

²⁹ <http://www.senat.fr/rap/r98-0452/r98-0452147.html> (16 avril 2014)

³⁰ <http://www.senat.fr/rap/r98-045/r98-0458.html> (16 avril 2014)

La commercialisation de produits d'assurance par les banques a de nombreux avantages, en effet la banque dispose d'un réseau d'agences extrêmement développé et il est donc facile d'utiliser ce réseau pour s'ouvrir à un autre marché car il ne pourra leur offrir qu'une meilleure rentabilité. Les agences bancaires ont tous les atouts dont doit disposer une agence d'assurance, à savoir notamment la proximité, la visibilité et le contact clients. Ainsi elles sont en mesure de proposer un plus grand nombre de produits à une clientèle déjà acquise, sachant que la gestion de produits d'assurance vie-capitalisation est très proche de celle des produits bancaires.

Une telle offre permet des augmentations de parts de marché aussi fulgurantes, car elle a aussi de nombreux avantages pour les clients. Pour commencer le client connaît son banquier puisqu'il le rencontre, pour la plupart, depuis de nombreuses années. Ainsi ce dernier connaît bien son client et ses habitudes, ce qui lui permet de lui proposer les offres les plus adaptées. De plus avec le développement de l'assurance aux guichets bancaires, les banques proposent un service plus simple (en regroupant les opérations) et un conseil complet. Enfin il faut garder à l'esprit que les produits d'assurance vie-capitalisation proposés par les bancassurances ont bien souvent des prix très avantageux puisqu'ils sont soumis à moins de charges de commercialisation qui s'amortissent plus rapidement.

Aujourd'hui, et avec la diminution des commissions³¹ liées aux crédits, les offres d'assurances chez les banques se diversifient. Celles-ci proposent désormais des assurances de biens et de personnes, en plus des assurances vie-capitalisation qu'elles proposaient jusqu'alors, pour encore plus de rentabilité.

Bien que 57% des français³² (entre 15 et 59ans) ont un contrat auto souscrit auprès d'un assureur contre seulement 6% chez un banquier, la concurrence s'intensifie entre les deux acteurs. En effet 40% des personnes interrogées envisageraient de souscrire un contrat d'assurance dommages chez leur banquier. De nos jours c'est le prix qui joue le rôle le plus important dans la décision de souscription, or il apparaît que si les assureurs savent proposer qualité et conseil, les banquiers eux proposent une approche plus tournée vers le produit en proposant notamment des offres « packagées » (la plus répandue étant le pack crédit et assurance auto).

³¹ Rémunération versée à l'occasion d'un contrat

³² http://www.empruntis.com/assurances-dommage/actualites/index.php?id_actu=1327 (22 avril 2014)

Entre 2012 et 2013 l'activité de la bancassurance montre une croissance³³ de 19% alors qu'elle n'avait évolué que de 8.2% l'année précédente, ce qui fait de la bancassurance un concurrent de taille pour les compagnies d'assurance, au même titre que les autres assureurs.

1.3.3. La concurrence digitale

Si le développement des mutuelles sans intermédiaires puis de la bancassurance ont été deux grandes ruptures dans la distribution des produits d'assurance, bien d'autres canaux font plus lentement leur apparition sur le marché.

Apparue au début du XXIème siècle avec l'éclatement de la bulle internet, il aura fallu quinze ans à l'e-assurance pour se développer en France. En effet dans les années 2000 on assiste à la création des premiers pure-players³⁴ tel qu'OK Assurance (filiale des AGF) qui stoppera finalement toutes activités en 2001. Ce n'est que depuis 2005 que les compagnies et mutuelles d'assurance s'intéressent réellement à ce nouveau mode de communication en transformant leurs sites basiques en véritables outils relationnels.

Depuis son apparition chez les particuliers, internet a facilité bien des démarches et 65% de la population avoue avoir un quotidien facilité par la révolution numérique. L'e-assurance est un concept qui attire majoritairement les jeunes familles urbaines appartenant aux catégories socio-professionnelles supérieures, déjà adeptes du e-commerce, et qui vise principalement à prendre contact avec de nouveaux clients afin de les inviter ensuite à se rendre en agence.

1.3.3.1. Les comparateurs d'assurance

Le principal avantage d'internet est l'accès facile et rapide à l'information. Aujourd'hui l'internaute peut trouver toutes sortes d'informations sur les sites dédiés aux assurances, il est ainsi de plus en plus exigeant et a même recours aux comparateurs en ligne afin d'être certain de son choix avant de se rendre en agence.

Les comparateurs sont des portails sur internet permettant aux utilisateurs de comparer différents contrats d'assurance en fonction de leurs tarifs. Cependant le premier inconvénient de ce type de comparaison est probablement le peu d'intérêt pour la qualité des informations

³³ <http://www.leseco.ma/finances/15655-bancassurance-l-activite-confirme-son-essor> (22 avril 2014)

³⁴ Entreprises exerçant uniquement sur internet

et la description des conditions du contrat. La rémunération de ces sites provient principalement des assureurs qui monnayent une mise en relation avec les clients potentiels. Les comparateurs d'assurance en ligne font donc d'internet un véritable outil de conquête.

Pour répondre à ce souci de qualité dans la comparaison des contrats d'assurance sur internet, AXA crée en 2012 le site Quialemeilleurservice.com, le premier comparateur de services (prestations proposées par l'assureur) en ligne, ne prenant pas en compte les tarifs proposés par les 24 assureurs mis en compétition. Ainsi l'internaute coche les services qui l'intéressent dans un moment de vie, et le comparateur détaillera les services proposés par chaque assureur. Une manière pour AXA d'être encore un peu plus visible sur internet puisque, dans la majorité des cas, l'assureur se retrouve sur le podium des meilleurs services proposés. Le PDG d'AXA France, Nicolas Moreau explique le but de la création d'un tel site : « *Ce comparateur doit nous permettre de mieux comprendre les attentes de nos clients, de nous challenger et de progresser, car nous ne sommes pas toujours sur le podium* ».

En 2010, 61% des internautes³⁵ comparaient les prix des différents assureurs via un comparateur en ligne et ces derniers sont, en 2011, à l'origine de 10% des souscriptions³⁶ de contrats d'assurance dommages.

1.3.3.2. Les assureurs de plus de plus visibles sur internet

Avec la nécessité grandissante pour chaque société d'être présente sur le net, les assureurs développent petit à petit leurs sites internet. Alors qu'ils n'étaient auparavant que de simples vitrines de compagnies d'assurance, aujourd'hui ils proposent de plus de plus de services. On y retrouve des conseils pratiques ainsi que des portails permettant d'accéder à son espace client, de gérer ses contrats, de demander des devis mais aussi de déclarer ou suivre un sinistre. Encore peu de compagnies proposent des souscriptions de contrats 100% en ligne et la majorité de ceux-ci sont des contrats auto qui nécessitent moins de conseils. AXA par exemple, développe depuis peu l'e-auto (assurance automobile sur internet). L'internaute fait une simulation du tarif sur le site axa.fr, si celui-ci lui convient, il peut choisir entre une souscription 100% en ligne ou être envoyé dans l'agence AXA participante la plus proche.

Toujours dans un souci de visibilité sur internet, et dans le but de s'imposer comme «*assureur leader du digital et du multi-accès* », AXA a annoncé le vendredi 11 avril 2014 son

³⁵ Baromètre des cyberconsommateurs 2010 Ifop pour Direct assurance

³⁶ Avis du Comité consultatif du secteur financier, banque-France.fr

partenariat « *stratégique* » d'un an avec Facebook, qui promet de belles perspectives quand on sait qu'aux Etats-Unis certains agents généraux AXA réalisent 60%³⁷ de leur activité grâce à ce réseau social. Ainsi AXA bénéficiera des équipes d'innovation et d'analyse Facebook pour l'aider dans l'augmentation de la visibilité de la marque sur le réseau social américain. Le but principal d'un tel partenariat est de toucher les jeunes actifs, en effet entre 2013 et 2015, AXA prévoit un budget de 800 millions d'euros pour sa conquête du monde digital.

Dans la même optique, en mai 2014, AXA a signé un partenariat mondial avec LinkedIn. Le but de celui-ci étant de soutenir les équipes ressources humaines, marketing, communication et distribution par la proposition de services premium et de prix négociés pour les salariés AXA. Véronique Weill, Directrice des Opérations du Groupe AXA : « *Par cet accord, nous espérons également continuer à améliorer la notoriété de notre marque auprès des clients mais également des talents, à travers la mise en place d'une stratégie de marque mondiale sur ce réseau. [...] Cet accord donnera, à nos équipes de distribution dans le monde, un accès à une plateforme de développement reconnue et permettra d'améliorer le service que ces équipes fournissent à nos clients grâce à davantage de proximité et de dialogue avec ceux-ci sur le plus grand réseau professionnel mondial* » (Site AXA, interne, 24 Juin 2014).

Pour se développer sur ce nouveau canal de distribution qu'est internet, certaines compagnies ont développé ce qu'on appelle la vente directe, en créant des filiales spécialisée dans la distribution d'e-assurance, avec des souscriptions exclusivement en ligne.

1.3.3.3. La vente directe

La vente directe en assurance correspond à la souscription de contrat sans aucun intermédiaire. Le consommateur souscrit une assurance (principalement auto, même si les autres tendent à se développer) 100% en ligne ou par téléphone, entrant ainsi en contact direct avec la compagnie qui va l'assurer et proposer des tarifs plus compétitifs.

La vente directe de contrats d'assurance a débuté dans les années 90 avec la création de filiales spécialisées par les compagnies d'assurance et se faisait exclusivement par téléphone. Avec l'arrivée d'internet celle-ci commence seulement à prendre ses parts sur le marché

³⁷ <http://www.lesechos.fr/entreprises-secteurs/finance-marches/actu/0203440670496-axa-s-allie-a-facebook-pour-accelerer-sa-mutation-digitale-664232.php> (14 avril 2014)

(3%³⁸ du marché de l'assurance en 2012, et des difficultés à progresser), les principales filiales spécialisées étant Amaguiz, Idmacif et DirectAssurance.

En 1992, AXA crée Direct Assurance, compagnie d'assurance à distance, distribuant principalement de l'assurance automobile (1,8 milliards d'euros de chiffre d'affaires sur 2,1 milliards), qui peine à générer des bénéfices. En effet les premiers ne seront réalisés que dix ans après son lancement. De son côté, Groupama lance en 2008 Amaguiz.com, un site internet qui propose des assurances auto adaptées au nombre de kilomètres parcourus. Enfin, la même année, la Macif ouvre IDMacif, un assureur en ligne qui se veut *low-cost* mais qui présente cependant 0.4 millions d'euros de perte en 2011 (sinistralité élevée, clients qui ne payent pas leurs cotisations, perte de clientèle...).

De nos jours, même si internet est bien ancré chez les consommateurs et qu'il prend de l'importance sur le marché de l'assurance, la souscription à 100% en ligne progresse encore peu et la vente directe sur internet n'est estimée qu'à 3% du marché des particuliers³⁹ dans le secteur de l'assurance.

Malgré ces difficultés, internet et la vente directe sont des acteurs entrant à surveiller sur le marché de l'assurance et restent une concurrence forte pour les agents généraux et leurs points de vente physiques. Il est donc capital pour le point de vente de travailler sur sa localisation et sa valeur ajoutée.

1.3.4. La grande distribution

La grande distribution dispose d'un très large potentiel de clientèle, et est désireuse de profiter de ce fichier clients important pour proposer des produits d'assurance dans le but de fidéliser un peu plus ses consommateurs.

Le premier à lancer cette tendance est Carrefour en 1989 avec « l'Épargne libre Carrefour » qui propose des assurances vie. Aujourd'hui l'enseigne n'est plus la seule intéressée par la distribution de produits d'assurance, Auchan et Casino annoncent de vraies intentions de développer leurs services financiers. Après « l'épargne libre Carrefour », Carrefour crée Carma en 1991 avec les Mutuelles du Mans. Il s'agit d'une société spécialisée

³⁸ <http://www.latribune.fr/entreprises-finance/banques-finance/industrie-financiere/20121109trib000730038/tops-et-flops-de-l-assurance-directe.html> (14 avril 2014)

³⁹ http://www.mba-enass-alumni.org/uploads/3/2/3/1/3231071/mba_enass_2012_becker_2012_distribution-assurance.pdf (15 avril 2014)

dans l'assurance IARD (incendie, accidents, risques divers) qui ne dégagera pas de bénéfices avant 1997. Aujourd'hui cette société (dont les produits sont distribués via Carrefour Banque) propose des assurances auto, multirisque habitation, famille ainsi que des complémentaires santé, et affiche des progressions impressionnantes (+ 16,1% en auto, + 21,2% en MRH, + 9,4% en protection famille et + 19% en santé en 2009⁴⁰).

De la même manière Casino propose ce type de produit en partenariat avec Amaline assurance, avec un mode de distribution bien différent puisque des cartes sont positionnées en tête de gondole. On trouve ainsi l'assurance chiens et chats au rayon croquette et l'assurance neige au rayon équipement pour les sports d'hiver.

D'autre comme IKEA propose depuis 1991 en collaboration avec la Société Suisse, des contrats MRH (multirisque habitation), assurance scolaire, hospitalisation et individuelle. A ce jour Ikea n'a vendu que quelques centaines de contrats et vise principalement les porteurs de la carte Ikea Family.

Si l'offre d'assurance est encore peu fournie et qu'elle ne représente pas encore un véritable concurrent pour les compagnies d'assurance, la grande distribution est un nouvel acteur sur le marché qui tend à se développer.

1.4. Conclusion

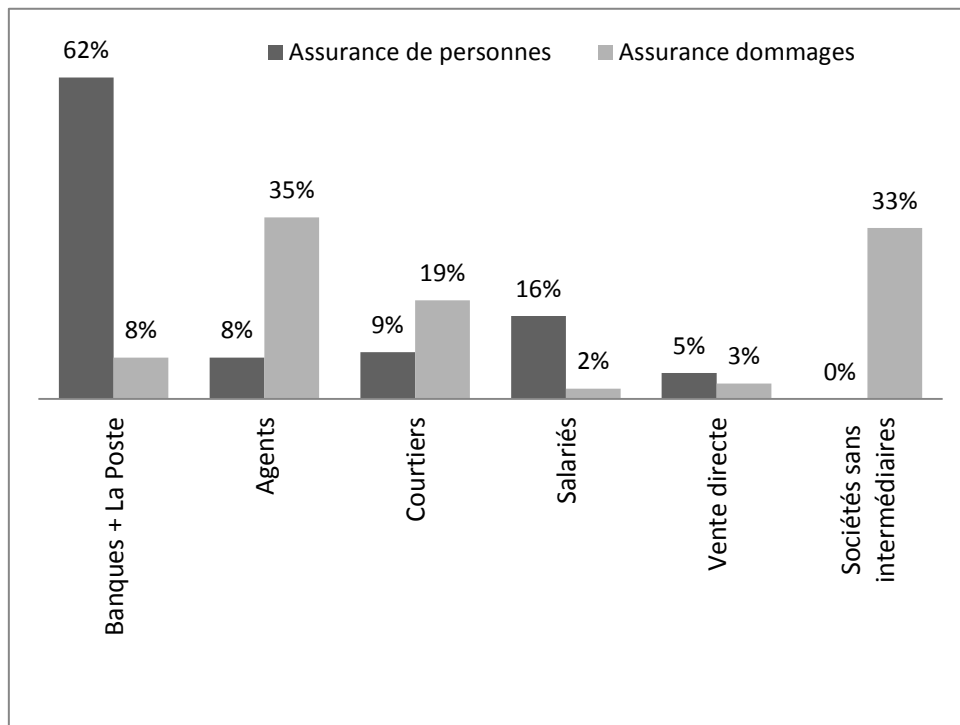
Comme il a été énoncé tout au long de cette première partie, les modes de distributions sont nombreux sur le marché et, par conséquent, les concurrents aussi. La figure 5, en plus de rappeler les différents types de concurrents présents dans la distribution d'assurance, montre les principaux modes de distribution pour les assurances dommages et les assurances de personnes.

En résumé, les assurances dommages sont principalement distribuées par les agents généraux (35%), qui gardent leur avance, et les sociétés sans intermédiaires (33%). Cependant au niveau des assurances de personnes, les agents généraux ne sont plus si bien représentés puisque ce sont les banques qui dominent avec 62% des contrats souscrits auprès d'eux. Les agents eux ne disposent que de 8%. Cette figure permet de mettre en avant le fait que chaque canal de distribution a une spécialité, même si la majorité d'entre eux distribue des assurances

⁴⁰ <http://www.argusdelassurance.com/tendances/l-assurance-nouvelle-source-de-revenus-pour-la-grande-distribution.49348> (17 avril 2014)

dommages comme des assurances de personnes (seules les sociétés sans intermédiaires ne proposent que des assurances dommages).

Figure 5 : Répartition des modes de distribution en France



Infographie Sophie STAËS (2014), d'après assemblée-nationale.fr⁴¹, réalisée sous Excel

L'offre proposée par chaque canal de distribution répond à des besoins et des clients différents. Il semblerait donc qu'il n'y ait pas encore de risque de cannibalisation du réseau de distribution « traditionnel », à savoir les compagnies d'assurance et leurs agents généraux. Cependant les concurrents sont de plus en plus nombreux sur un marché saturé, il est donc absolument nécessaire et indispensable de mettre en avant la valeur ajoutée d'un agent général mais aussi d'accroître celle-ci en commençant par étudier méticuleusement le lieu d'implantation d'une agence physique ainsi que son organisation, sa présentation mais aussi les produits et services qu'elle va proposer, tout en mettant en avant ses atouts : la proximité avec sa clientèle et les conseils qu'elle peut leur apporter.

Ainsi la seconde partie propose une analyse géographique du territoire de l'Ile-de-France afin de mieux comprendre celui-ci.

⁴¹ <http://www.assemblee-nationale.fr/12/pdf/rapports/r2217.pdf> (16 avril 2014)

2. LA PRISE EN COMPTE DES RÉALITÉS GÉOGRAPHIQUES : CONTEXTE D'UNE ÉTUDE DE TERRITOIRE EN RÉGION

Avec la saturation des marchés, apparaît une prise de conscience. Le territoire est un espace à partager avec ses concurrents et sur lequel les consommateurs ont des comportements différents. Les entreprises assistent alors au besoin de connaître leur territoire afin de s'y développer au mieux, mais aussi et surtout de s'implanter sur l'espace le plus approprié.

2.1. Introduction

L'intérêt de cette nouvelle partie est de mettre en avant les différences entre des zones géographiques proches (celles d'Ile-de-France) afin de mieux y comprendre par la suite le comportement des consommateurs. Pour cela un état des lieux de la région est effectué. En effet en marketing, connaître un territoire c'est avant tout connaître la population (et les clients), l'espace mais aussi les concurrents qui y sont installés, ceci dans le but de cibler ensuite les marchés à fort potentiel mais aussi de se faire une idée de sa position concurrentielle.

Pour cela il est nécessaire de disposer de certaines données géographiques mais aussi socio-économiques.

2.2. Les données disponibles

L'étude d'un territoire s'effectue à l'aide de différentes données, topographiques, socio-économiques, et clients. Les données disponibles sont regroupées en bases de données statistiques qui donnent des informations sur un pays, une région, un département, une commune ou encore un iris⁴². Les informations nécessaires à une étude de territoire varient selon le but de l'étude, ainsi elles peuvent être obtenues via divers acteurs tels que l'INSEE⁴³, l'IGN⁴⁴, des entreprises (EDF, SNCF, RATP) et donc être libres et gratuites ou bien au contraire payantes. D'autres données sont internes à l'entreprise qui effectue l'étude (données clients, points de vente, chiffre d'affaire...).

⁴² Ilots regroupés par des indicateurs statistiques, Découpage du territoire français réalisé par l'INSEE

⁴³ Institut national de la statistique et des études économiques

⁴⁴ Institut national de l'information géographique et forestière

2.2.1. Les données géographiques

Les données géographiques (ou topographiques) constituent la base indispensable à une carte. Il s'agit des informations permettant de visualiser un territoire, à savoir les reliefs, les limites de régions, départements et communes, l'hydrographie⁴⁵, le réseau routier mais aussi les infrastructures.

Figure 6 : Exemple de données géographiques sur le 1^{ER} arrondissement de Paris



L'information géographique est associée à une carte grâce à des coordonnées x et y qui permettent de représenter une multitude d'objets. La majorité de ces données sont payantes car généralement très précises, comme par exemple des altitudes ou des adresses. On en retrouve une grande partie sur le site de l'IGN mais aussi sur des sites tels que data.iledefrance.fr, opendata.paris.fr ou encore iau-idf.fr.

Infographie Sophie STAËS (2014), réalisée sous Geoconcept V7

De façon générale, les données géographiques de base regroupent les limites administratives, les routes, les espaces verts et les cours d'eau. Elles peuvent être de deux types : vecteur ou raster.

⁴⁵ Ensemble des lacs et cours d'eau

Les données vectorielles (type vecteur) sont des données représentables par des points, lignes, courbes ou surfaces. Chaque représentation est définie par des coordonnées x et y qui permettent de placer l'objet sur une carte. En plus des coordonnées géographiques, chaque donnée vecteur peut disposer d'attributs (données attributaires) qui donnent des informations sur l'objet (par exemple des données population, associées à un polygone « commune »).

Les données raster, elles, permettent une représentation de la réalité par des cellules sur une grille, comme le montre la figure 8. Il s'agit d'un ensemble de carrés (plus ou moins gros selon la résolution choisie) successifs qui permettent la couverture d'une aire géographique. Les données raster sont généralement des images satellites ou des photos aériennes, et contrairement aux données vectorielles elles ne disposent pas de base de données associée.

Figure 7 : Représentation vecteur

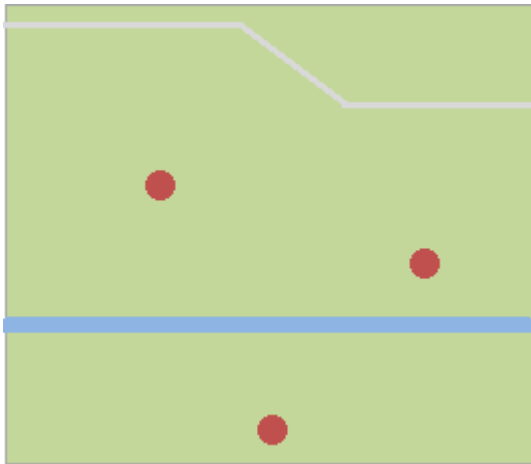
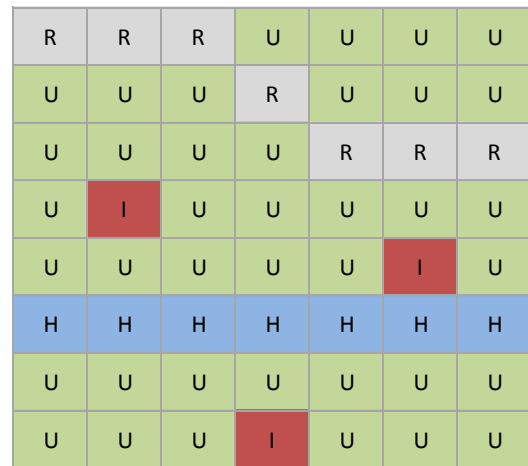


Figure 8 : Représentation raster



U	Zone urbaine
I	Infrastructures
H	Réseau hydrographique
R	Réseau routier

Infographie Sophie STAËS (2014), réalisée sous Excel et Paint

Dans une étude du territoire, les données topographiques font partie des données absolument nécessaires et ce sont principalement des données de types vecteur qui sont alors utilisées. Ce sont sur celles-ci que seront ajoutées toutes les autres informations, comme notamment les informations socio-économiques des communes ou iris.

2.2.2. Les données socio-économiques

Comme leur nom l'indique, les données socio-économiques renseignent à la fois des caractéristiques sociales et économiques d'une zone. Elles permettent de mieux comprendre un territoire ainsi que la population qui le compose grâce à l'étude des catégories socio-professionnelles. Ainsi il est possible de cerner le comportement social et économique d'une population.

Il existe différentes sources socio-économiques, cependant c'est certainement le site de l'INSEE qui produit le plus de données relatives à la population, l'économie et la démographie⁴⁶. L'INSEE est l'« institut national de la statistique et des études économiques » qui analyse les recensements de population. Ces données sont en grande majorité accessibles et mises à jour régulièrement.

Certains ministères produisent eux aussi de la donnée. C'est par exemple le cas des Affaires Sociales et de la Santé⁴⁷, qui dispose du répertoire FINESS (Fichier National des Etablissements Sanitaires et Sociaux) proposant des données sur les domaines médico-social, sanitaire et social. Toutefois, si ces données sont libres et accessibles, leur usage est en revanche règlementé⁴⁸. A l'échelle internationale, certaines données population sont disponibles sur le site de l'ONU (Organisation des Nations Unies)⁴⁹.

Cette courte liste n'est pas exhaustive. En effet de nombreuses institutions telles que les communautés de commune, métropoles, ou même des associations réalisent leurs propres bases de données. Dans ce cas il est fréquent que celles-ci ne soient pas libres.

Il est important de savoir que dans la plupart des situations, les données énoncées précédemment sont basées sur celles de l'INSEE et sont ensuite complétées. L'institut national de la statistique et des études économiques, est donc le premier fournisseur de données socio-économiques et le plus complet sur le sujet. C'est donc très souvent sur celles-ci que les études de territoire et de population sont effectuées, ce qui est le cas au sein d'AXA France, qui allie données INSEE et données internes à l'entreprise pour réaliser ses études.

⁴⁶ Données relatives à la population et à son dynamisme (évolution, natalité, mortalité, migration...)

⁴⁷ <http://www.drees.sante.gouv.fr/donnees,432.html> (22 mars 2014)

⁴⁸ « Réutilisation [...] autorisée dans le cadre d'une utilisation privée, ou dans celui de la réalisation de travaux. Dans ce dernier cadre, la réutilisation est autorisée à condition [...] que leur source soit clairement précisée aux destinataires des travaux ou des produits » <http://finess.sante.gouv.fr/jsp/recherche.jsp?mode=simple> (13 août 2014)

⁴⁹ <http://esa.un.org/unpd/wpp/index.htm> (22 mars 2014)

2.2.3. Les données internes

Les données entreprises sont particulièrement importantes aux études puisqu'elles constituent la connaissance clients. Les renseignements essentiels à une étude d'implantation d'agence chez AXA sont en premier lieu les données clients puis les données agences.

La base de données clients utilisée chez AXA pour les études est constituée des champs suivants : Nom, Adresse, Entité opérationnelle⁵⁰, Numéro de client, Numéro de contrat, Fractionnement⁵¹, Date début du contrat, Emmissibilité⁵², Commissions⁵³. Quant à la base de données agences, celle-ci dispose des informations suivantes : Nom de l'agent, Code Point de contact, Adresse, Nombre de portefeuille⁵⁴, Ancienneté, Date ouverture, Nombre de clients par entité opérationnelle, Nombre de clients par branche, Nombre de clients par tranche d'âge, Nombre de clients par catégorie socio-professionnelle, Nombre de contrats, Taux de pénétration, Cotisations⁵⁵.

La sauvegarde des données est primordiale dans l'activité d'une entreprise. Elles constituent l'historique de celle-ci. L'ensemble de ces données forment le système d'information de la société et sont utilisées dans de nombreux métiers. En plus des données clients et agence, AXA en produit chaque jour de nouvelles : partenaires, commerciaux, sinistres... ce qui fait de plus en plus de données entreprise à traiter. Dans les études d'implantation ces données sont associées à des territoires qu'il est tout aussi nécessaire de connaître.

2.3. L'Ile-de-France

Divisée en huit départements (Paris, Hauts-de-Seine, Seine-Saint-Denis, Val-de-Marne, Essonne, Seine-et-Marne, Val-d'Oise et Yvelines), l'Ile-de-France a une superficie de 12 011 km² et est particulièrement peuplée puisqu'elle ne représente que 3% du pays⁵⁶ mais regroupe 19% de la population française en 2013.

⁵⁰ Particulier ou Entreprise

⁵¹ Fréquence des paiements du contrat : Annuel, Semestriel, Trimestriel ou Mensuel

⁵² Primes théorique par an

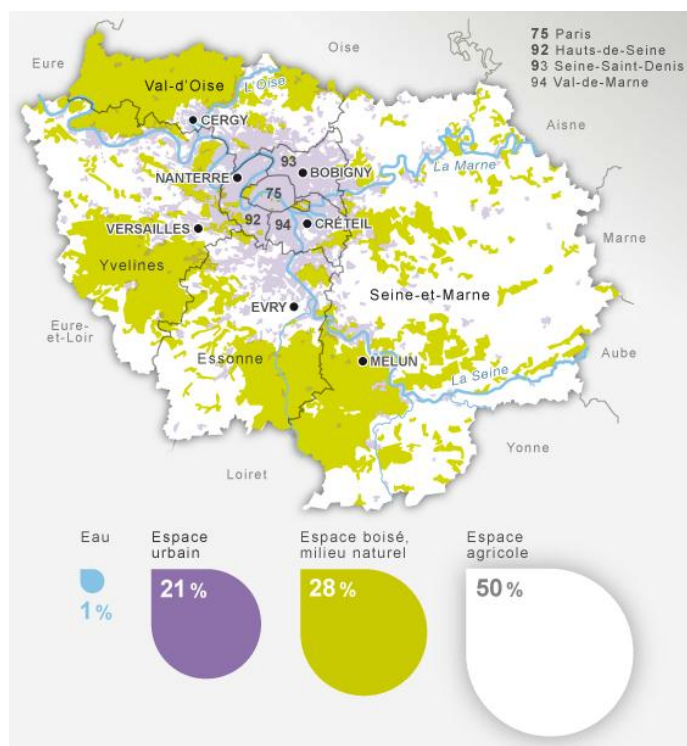
⁵³ Montant que touche l'agent sur un contrat

⁵⁴ Le portefeuille est l'ensemble des contrats souscrits auprès d'un agent

⁵⁵ Montant que paye l'assuré chaque mois, trimestre, semestre ou an pour son contrat d'assurance

⁵⁶ <http://ddc.arte.tv/nos-cartes/le-grand-paris-4-4> (6 mai 2014)

Illustration 1 : Départements de la région Ile-de-France



La région Ile-de-France, ou « région parisienne » constitue un territoire très varié puisqu'il est composé de 21% d'espace urbain, 28% d'espace naturel et 50% d'espace agricole. Son PIB par habitant de plus de 51 000 euros⁵⁷ en 2012, fait de l'Ile-de-France la région qui produit le plus de richesse du pays.

Carte d'identité de l'Ile-de-France, iledefrance.fr⁵⁸

2.3.1. Paris

Paris a la particularité d'être l'unique commune-département du pays. Située sur la Seine, la ville s'étend de chaque côté de celle-ci, la séparant ainsi en deux rives : la rive droite et la rive gauche. La commune est séparée de ses communes limitrophes par le boulevard périphérique et est entourée par les départements Hauts-de-Seine, Val-de-Marne et Seine-Saint-Denis.

Les arrondissements de Paris sont des divisions intra-communales créés en 1859. Ils disposent chacun de caractéristiques et d'un caractère propre, qui font d'eux des territoires différents les uns des autres.

Paris est le département le plus peuplé de l'Ile-de-France, avec une population de 2,24 millions d'habitants⁵⁹ en 2013, presque deux fois plus élevée que celle des autres départements de la région alors qu'il s'agit du moins étendu. Il s'agit de la capitale la plus

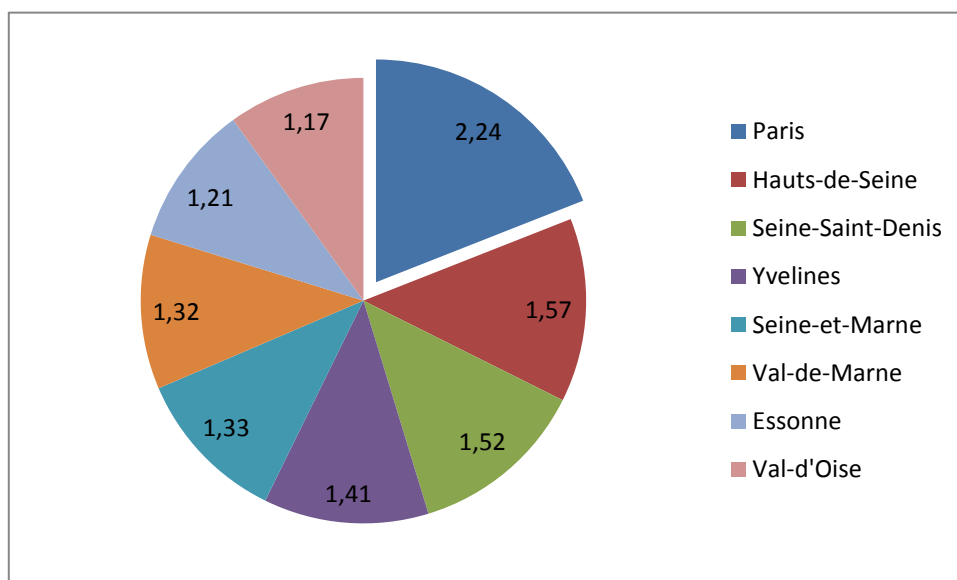
⁵⁷ http://www.insee.fr/fr/themes/tableau.asp?reg_id=20&ref_id=ecotc08101 (20août2014)

⁵⁸ <http://www.iledefrance.fr/territoire/carte-identite> (15 juillet 2014)

⁵⁹ <http://www.iledefrance.fr/territoire/carte-identite> (9 mai 2014)

dense d'Europe en termes de population avec une densité de 21 196 hab. /km²⁶⁰, alors qu'elle est loin d'être la plus étendue. En effet Rome est environ douze fois plus grande.

Figure 9 : Répartition de la population par département en millions d'habitants en 2013



Infographie Sophie STAËS (2014), d'après iledefrance.fr⁶¹, réalisé sous Excel

L'unité urbaine de Paris est composée de 412 communes⁶², dont celles de Paris et de sa petite couronne ainsi que 52 communes de Seine-et-Marne, 89 des Yvelines, 82 de l'Essonne et 65 du Val-d'Oise. Il s'agit d'une zone de plus de 10 millions d'habitants, où la population est nettement regroupée. En effet celle-ci ne représente que 23.7% de la surface de la région, mais abrite 88.7% de la population de l'Ile-de-France.

2.3.2. La petite couronne

La petite couronne de Paris est formée des départements Hauts-de-Seine (92), Val-de-Marne (94) et Seine-Saint-Denis (93), les départements limitrophes à la capitale. Elle regroupe 123 communes et dispose de 4,4 millions d'habitants⁶³ pour une superficie⁶⁴ de 657 km².

⁶⁰ <http://www.insee.fr/fr/bases-de-donnees/esl/comparateur.asp?codgeo=cv-7599> (9 mai 2014)

⁶¹ <http://www.iledefrance.fr/territoire/carte-identite> (9 mai 2014)

⁶² <http://www.insee.fr/fr/methodes/nomenclatures/zonages/zone.asp?zonage=UU2010&zone=00851> (9 mai 2014)

⁶³ <http://www.insee.fr/fr/regions/idf/reg-dep.asp?theme=1&suite=1> (12 mai 2014)

⁶⁴ <http://www.senat.fr/rap/r07-262/r07-26212.html> (source de l'ensemble des superficies, 20 août 2014)

Figure 10 : Département des Hauts-de-Seine



Le département des Hauts-de-Seine présente une superficie de 176 km² avec une densité de population de 8 987 hab. /km² et 1 581 628 habitants⁶⁵ en 2011 qui en font le département le plus petit de la petite couronne mais aussi le plus peuplé. Parce qu'il est entouré par les départements de la Seine-Saint-Denis (nord-est), des Yvelines (ouest), du Val-d'Oise (nord), de l'Essonne (sud), du Val-de-Marne (sud-est) et de Paris (est), et notamment grâce à sa forme, ce département est celui qui touche le plus de départements franciliens. Les Hauts-de-Seine sont composés de 36 communes⁶⁶.

Infographie Sophie STAËS (2014), réalisé sous Geoconcept V7

Figure 11 : Département Seine-Saint-Denis



Formé de 40 communes⁶⁷, le département de la Seine-Saint-Denis est étendu sur 236 km² au nord de Paris avec 1 529 928 habitants. Malgré un fort taux de chômage (12.6% en 2013⁶⁸), le département est dynamique grâce à de grands pôles économiques tels que la Plaine Saint-Denis, l'aéroport du Bourget et la proximité de la zone aéroportuaire Paris-Charles-de-Gaulle, les usines...

Infographie Sophie STAËS (2014), réalisé sous Geoconcept V7

⁶⁵ <http://www.insee.fr/fr/ppp/bases-de-donnees/recensement/populations-legales/france-departements.asp?annee=2011> (source de l'ensemble des populations, 20 août 2014)

⁶⁶ <http://www.insee.fr/fr/ppp/bases-de-donnees/recensement/populations-legales/departement.asp?dep=92>

⁶⁷ <http://www.insee.fr/fr/ppp/bases-de-donnees/recensement/populations-legales/departement.asp?dep=93>

⁶⁸ http://www.insee.fr/fr/themes/tableau.asp?reg_id=20&ref_id=tratc03301 (10 juin 2014)



Figure 12 : Département du Val-de-Marne

Le Val-de-Marne est un département très urbanisé avec cependant une faible partie d'espaces agricoles au sud-est. La population y est inégalement répartie, en effet alors qu'à Vincennes la densité de population est de 2 491 hab. /km² en 2009⁶⁹, elle est seulement de 370 hab. /km² à Santeny⁷⁰.

Infographie Sophie STAËS (2014), réalisé sous Geoconcept V7

2.3.3. La grande couronne

La grande couronne de Paris est composée de 4 départements, la Seine-et-Marne (77), l'Essonne (91), les Yvelines (78) et le Val-d'Oise (95). La grande couronne est bien plus étendue que la petite puisqu'elle a une superficie de 2 083 km². Elle est cependant à peine plus peuplée, avec 5,1 millions habitants⁷¹.



Figure 13 : Département Seine-et-Marne

La Seine-et-Marne est un département de l'Est parisien, particulièrement étendu puisque sa superficie de 5 915 km² représente à elle seule 49% de celle de l'Île-de-France. Très peuplé avec ses 1 338 427 habitants, le département est cependant principalement agricole. En effet 56% de son territoire est lié à l'agriculture. Le tourisme y est également développé grâce notamment à la présence du parc Disneyland Paris qui emploie 12 000 salariés⁷² (ce qui en fait le premier employeur du département) mais aussi à son parc hôtelier.

Infographie Sophie STAËS (2014), réalisé sous Geoconcept V7

⁶⁹ <http://www.insee.fr/fr/bases-de-donnees/esl/comparateur.asp?codgeo=com-94080> (8 juin 2014)

⁷⁰ <http://www.insee.fr/fr/bases-de-donnees/esl/comparateur.asp?codgeo=com-94070> (8 juin 2014)

⁷¹ <http://www.insee.fr/fr/regions/idf/reg-dep.asp?theme=1&suite=1> (14 juin 2014)

⁷² <http://www.seine-et-marne.fr/recrutements-disney> (14 juin 2014)

Figure 14 : Département de l'Essonne



Le département de l'Essonne est intégré à 40% dans l'agglomération parisienne. 77% du département est constitué d'espace rural, principalement au sud. Au nord se situe une zone urbaine sur laquelle se trouve une partie de l'aéroport d'Orly, deuxième aéroport de France avec 28,3 millions de passagers en 2013⁷³. Avec 196 communes, l'Essonne a une densité de peuplement hétérogène d'une commune à l'autre. En effet la population est particulièrement concentrée au nord-est.

Infographie Sophie STAËS (2014), réalisé sous Geoconcept V7

Figure 15 : Département des Yvelines



Le département des Yvelines est situé à l'ouest de Paris. 262 communes le composent et 1 408 765 habitants y sont regroupés sur 2 284 km². 47% de la surface du territoire est utilisé pour l'agriculture, quant à l'espace urbanisé, il est principalement concentré au nord du département et en représente 21,5%. Les principaux attraits touristiques des Yvelines sont le château de Versailles et le parc zoologique de Thoiry.

Infographie Sophie STAËS (2014), réalisé sous Geoconcept V7

⁷³ <http://www.aeroportsdeparis.fr/ADP/fr-FR/Groupe/Groupe-Strategie/LEssentiel/Trafic/> (20 août 2014)

Figure 16 : Département du Val-d’Oise



Le département du Val-d’Oise est peuplé de 1 180 365 habitants pour une superficie de 1 246 km². Il s’agit du département le plus jeune de l’Ile-de-France.

Infographie Sophie STAËS (2014), réalisée sous Geoconcept V7

2.3.4. La comparaison des départements

Certaines caractéristiques viennent d’être énoncées sur chacun des départements de l’Ile-de-France. Il sera à présent plus intéressant de comparer les différentes caractéristiques de ces huit territoires. Les données utilisées proviennent donc du site de l’INSEE et concernent les tranches d’âges et les catégories socio-professionnelles des populations de ces départements.

2.3.4.1. Les tranches d’âges

Le tableau 2 présente le nombre d’habitants par tranche d’âge et par département de la région Ile-de-France. Les données suivantes sont présentées telles qu’elles sont proposées par le site de l’INSEE.

Tableau 2 : Répartition de la population d’Ile-de-France par tranche d’âge

	Au 1er janvier 2013 (p)									
	Paris	Seine-et-Mame	Yvelines	Essonne	Hauts-de-Seine	Seine-Saint-Denis	Val-de-Marne	Val-d’Oise	Ile-de-France	France
moins de 20 ans	449 305	381 822	385 622	339 836	406 304	447 467	348 875	339 488	3 098 719	16 122 039
de 20 à 39 ans	783 770	360 920	357 220	328 012	472 623	453 645	387 021	324 897	3 468 108	16 158 670
de 40 à 59 ans	570 526	375 571	389 903	333 281	419 117	397 794	358 090	319 920	3 164 202	17 639 435
de 60 à 74 ans	298 787	163 167	185 187	154 754	186 564	161 680	160 027	137 617	1 447 783	9 698 383
de 75 ans ou plus	170 917	80 259	100 253	82 347	117 435	84 771	94 277	69 292	799 551	5 924 389
Total	2 273 305	1 361 739	1 418 185	1 238 230	1 602 043	1 545 357	1 348 290	1 191 214	11 978 363	65 542 916

Données INSEE au 1^{er} janvier 2013⁷⁴

⁷⁴ http://www.insee.fr/fr/themes/tableau.asp?reg_id=20&ref_id=poptc02104 (19 août 2014)

Celles-ci étant composées de nombres élevés, nous proposons le tableau 3, exprimé en pourcentage.

Tableau 3 : Répartition de la population d’Ile-de-France par tranche d’âge en pourcentage

	Paris	Seine-et-Marne	Yvelines	Essonne	Hauts-de-Seine	Seine-Saint-Denis	Val-de-Marne	Val-d'Oise	Ile-de-France	France
moins de 20 ans	20%	28%	27%	27%	25%	29%	26%	28%	26%	25%
de 20 à 39 ans	34%	27%	25%	26%	30%	29%	29%	27%	29%	25%
de 40 à 59 ans	25%	28%	27%	27%	26%	26%	27%	27%	26%	27%
de 60 à 74 ans	13%	12%	13%	12%	12%	10%	12%	12%	12%	15%
de 75 ans ou plus	8%	6%	7%	7%	7%	5%	7%	6%	7%	9%
Total	2 273 305	1 361 739	1 418 185	1 238 230	1 602 043	1 545 357	1 348 290	1 191 214	11 978 363	65 542 916

Infographie Sophie STAËS (2014), données INSEE, réalisé sous Excel

Ce tableau permet de mettre en avant le fait que la tranche d’âge majoritaire est différente d’un département à l’autre. En effet si à Paris, dans les Hauts-de-Seine, en Seine-Saint-Denis et dans le Val-de-Marne les 20 – 39 ans sont les plus représentés ; la Seine-et-Marne, l’Essonne et les Yvelines présente une population plus vieillissante avec une majorité de 40 – 59 ans. Enfin le Val-d’Oise se distingue puisque ce sont les moins de 20 ans qui y sont les plus nombreux, ce qui en fait le département le plus jeune de la région.

Il est tout de même remarquable que les proportions sont approximativement les mêmes d’un département à l’autre. Ainsi on retrouve une moyenne de 19% de plus de 60 ans sur l’ensemble de l’Ile-de-France.

2.3.4.2. L’emploi

Le tableau 4 propose la répartition du nombre d’actifs et d’inactifs par département de la région et se limite à la population de 15 à 64 ans. De même que précédemment il s’agit de données INSEE qui sont par la suite recalculées en pourcentage dans le tableau 5.

Tableau 4 : Population active ou inactive en Ile-de-France

	Au 1er janvier 2011									
	Paris	Seine-et-Marne	Yvelines	Essonne	Hauts-de-Seine	Seine-Saint-Denis	Val-de-Marne	Val-d'Oise	Ile-de-France	France
Actifs (effectif)	1 226 624	677 024	697 829	608 160	822 572	746 034	677 477	587 126	6 042 845	30 314 768
Inactifs (effectif)	370 475	215 364	228 851	200 182	235 091	275 444	216 471	202 038	1 943 916	11 519 669
Total (effectif)	1 597 099	892 388	926 680	808 342	1 057 663	1 021 478	893 947	789 164	7 986 761	41 834 437

Données INSEE au 1^{er} janvier 2011⁷⁵

Les actifs sont composés des hommes et des femmes en emploi ainsi que ceux au chômage, quant à la catégorie inactifs, elle est représentée par les retraités, les élèves et étudiants ainsi que les stagiaires non rémunérés.

Tableau 5 : Population active ou inactive en Ile-de-France

	Paris	Seine-et-Marne	Yvelines	Essonne	Hauts-de-Seine	Seine-Saint-Denis	Val-de-Marne	Val-d'Oise	Ile-de-France	France
Actifs (effectif)	77%	76%	75%	75%	78%	73%	76%	74%	76%	72%
Inactifs (effectif)	23%	24%	25%	25%	22%	27%	24%	26%	24%	28%
Total (effectif)	1 597 099	892 388	926 680	808 342	1 057 663	1 021 478	893 947	789 164	7 986 761	41 834 437

Infographie Sophie STAËS (2014), données INSEE, réalisé sous Excel

En règle générale, sur le tableau 5, la proportion d'actifs et d'inactifs par département est presque identique à celle retrouvée sur la région. Il ressort tout de même que les Hauts-de-Seine présentent un nombre d'actifs plus élevé que la moyenne. Cela peut être expliqué par le nombre d'emplois élevé sur ce département avec notamment la présence de la zone de La Défense. Le Val-d'Oise, quant à lui, dispose de plus d'inactifs que les autres départements, en effet il a été vu plus tôt qu'il s'agissait d'un département jeune, et donc probablement peuplé de nombreux étudiants (inactifs). Il en est de même pour la Seine-Saint-Denis.

Le nombre de chômeurs étant, dans les tableaux précédents, inclus dans le nombre d'actifs, le tableau 6 donne des indications supplémentaires sur le chômage dans chacun des départements.

⁷⁵ http://www.insee.fr/fr/themes/tableau.asp?reg_id=20&ref_id=tratc03101 (19 août 2014)

Tableau 6 : Taux de chômage en 2013

	Paris	Seine-et-Marne	Yvelines	Essonne	Hauts-de-Seine	Seine-Saint-Denis	Val-de-Marne	Val-d'Oise	Ile-de-France	France métropolitaine
Quatrième trimestre	8,1	7,9	7	7,3	7,6	12,6	8,6	9,8	8,6	9,7
Troisième trimestre	8,3	8	7,2	7,4	7,7	12,7	8,7	9,9	8,7	9,9
Deuxième trimestre	8,3	8	7,2	7,4	7,7	12,6	8,6	9,9	8,7	9,9
Premier trimestre	8,3	7,9	7,1	7,4	7,7	12,6	8,6	10	8,7	9,9

Données INSEE au 1^{er} janvier 2014⁷⁶

Ce nouveau tableau apporte de nouvelles informations sur les différents territoires d'Ile-de-France. Le taux de chômage en Seine-Saint-Denis par exemple, est particulièrement élevé par rapport à la région. Au contraire des Yvelines, de l'Essonne et des Hauts-de Seine qui sont des départements beaucoup moins touchés par le chômage que les autres d'Ile-de-France

Toutes ces informations peuvent s'avérer primordiales dans la couverture d'un territoire par une société.

2.4. Le maillage du territoire francilien

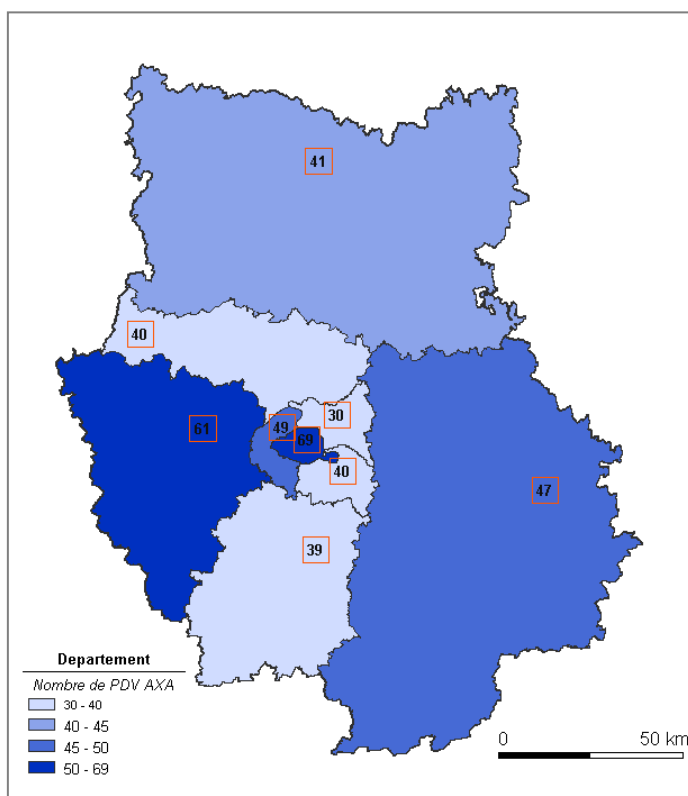
Le maillage territorial d'une société correspond à la couverture d'un territoire, par celle-ci, grâce à son réseau de distribution. L'optimisation du maillage peut se faire grâce au géomarketing et à l'analyse de données, et est primordial dans la recherche de meilleures visibilité et rentabilité des agences. Il est notamment important de rendre compte, pour une entreprise, de sa couverture réseau mais aussi de celle de ses concurrents.

2.4.1. Le réseau d'agents généraux AXA

L'annexe 1 présente les agences générales d'assurance AXA en Ile-de-France (l'Oise étant incluse dans la région, pour l'entreprise). On y remarque aisément que la concentration d'agence est bien plus élevée sur Paris et sa petite couronne que sur les cinq départements de la grande couronne. Les départements du centre semblent plus représentés en termes d'agences générales, cependant ceux de la grande couronne étant plus étendus il est judicieux de se demander si le nombre d'agence par département est réellement différent.

⁷⁶ http://www.insee.fr/fr/themes/tableau.asp?reg_id=20&ref_id=tratc03301 (19 août 2014)

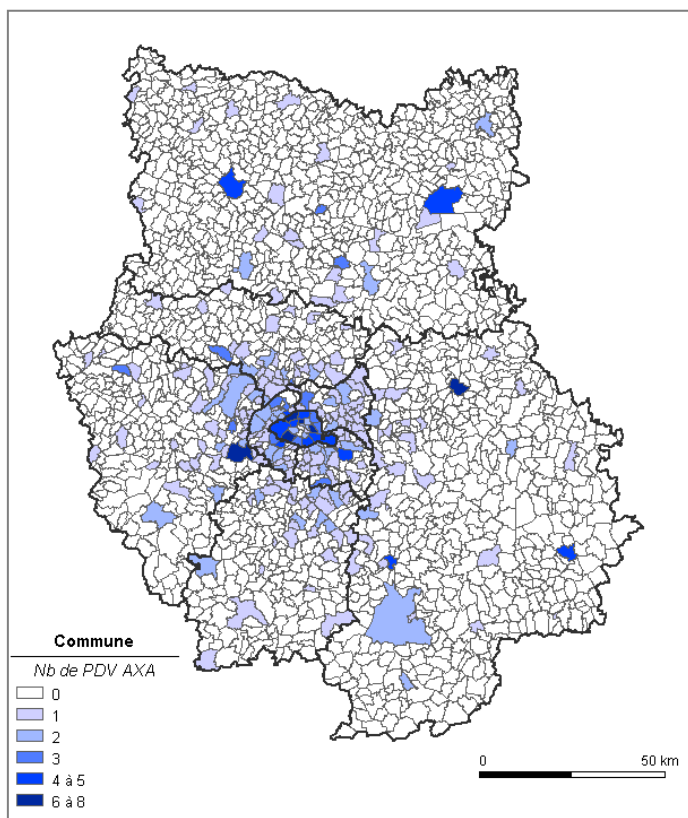
Figure 17 : Nombre d'agences générales AXA par département en Ile-de-France en 2014



Sur la figure 17, il est remarquable que des départements de taille très différentes tels que les Hauts-de-Seine et la Seine-et-Marne, présentent un nombre de points de ventes presque identique. Il en est de même pour les Yvelines et Paris ce qui ne devrait pas être le cas si le maillage territorial était parfait. Les explications possibles seraient donc : soit Paris intra-muros est surreprésenté en nombre d'agences, soit les Yvelines sont sous-représentées.

Infographie Sophie STAËS (2014), données internes AXA, réalisé sous Geoconcept V7

Figure 18 : Nombre d'agences générales AXA par commune en Ile-de-France en 2014



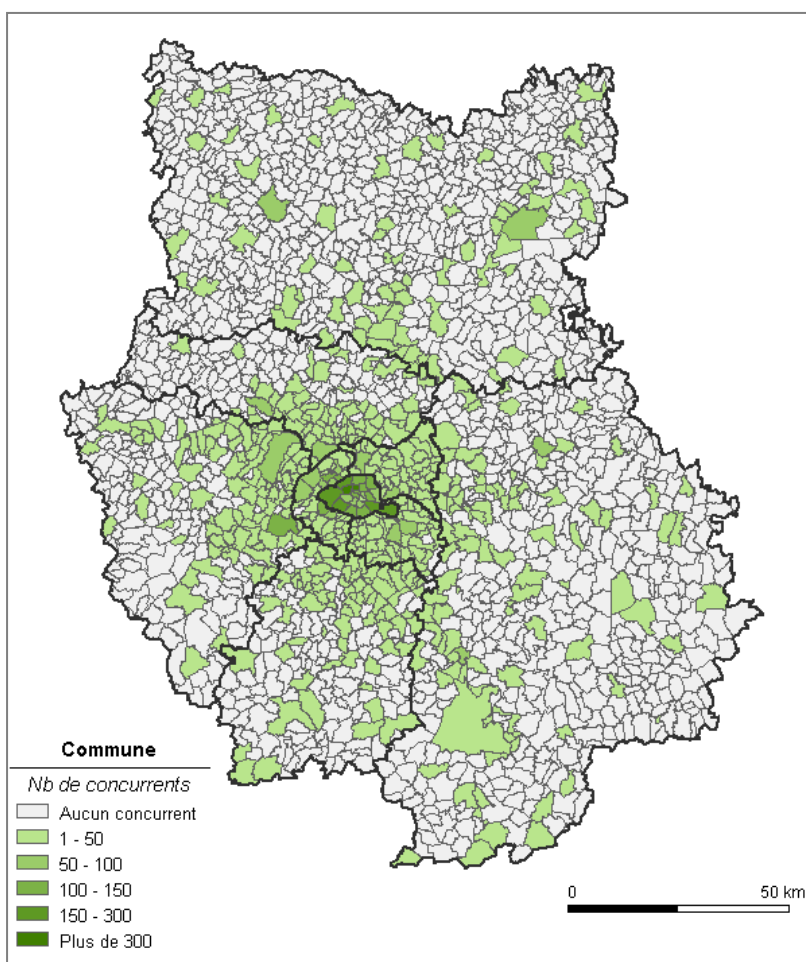
De plus on remarque que sur la grande couronne relativement peu de communes sont couvertes par une agence générale AXA, alors que sur celles qui le sont on retrouve parfois jusqu'à 8 points de ventes sur la même commune (c'est le cas de Versailles dans les Yvelines).

Infographie Sophie STAËS (2014), données internes AXA, réalisé sous Geoconcept V7

2.4.2. La localisation de la concurrence

Au sein du service Administration des réseaux chez AXA France, les concurrents qui sont pris en compte sont les agences bancaires (incluant La Poste) ainsi que les assureurs (agences générales, courtiers, mutuelles d'assurance...). On retrouve donc sur le territoire Ile-de-France, un nombre impressionnant de concurrents. De plus il est à noter, que les bases de données d'AXA ne disposent pas d'informations précises sur les enseignes d'assurance concurrentes, mais uniquement du type d'assurance dont il s'agit, alors que concernant les agences bancaires, la base de données renseigne de quelle société il est question.

Figure 19 : Nombre de concurrents (Banques et Assurances) par commune en 2014



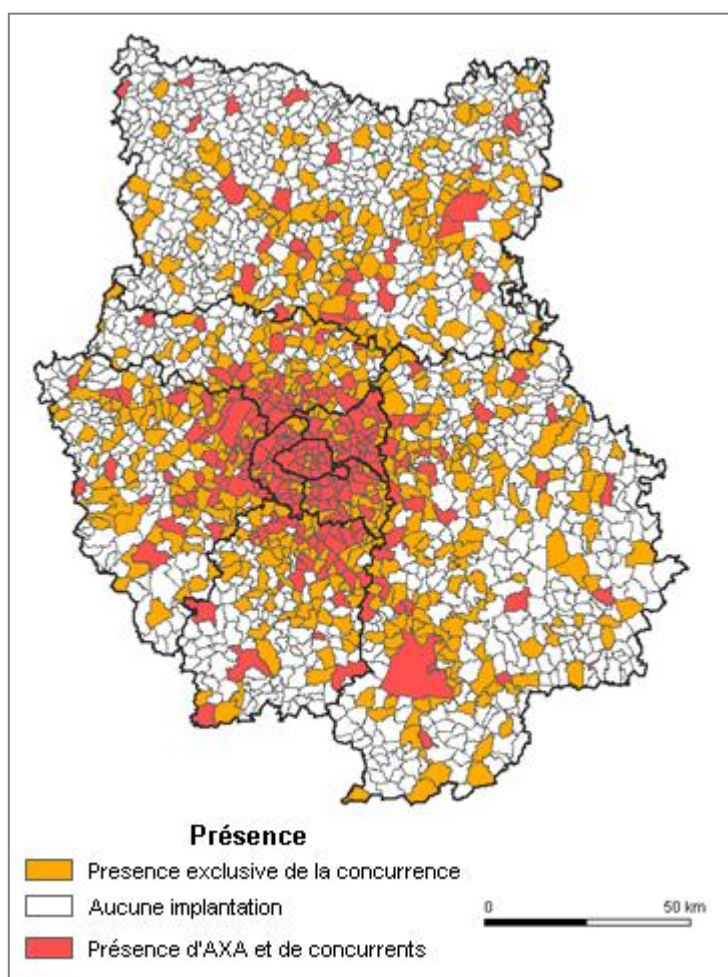
Compte tenu du nombre élevé de concurrents, une représentation cartographique à l'adresse serait peu lisible. C'est donc une représentation sous forme de thématique qui est choisie, et qui propose alors une analyse plus approfondie de la couverture de la région par les concurrents assureurs et bancaires confondus.

Infographie Sophie STAËS (2014), données internes AXA, réalisé sous Geoconcept V7

Comme pour la couverture du réseau AXA, les concurrents sont présents en plus grande densité sur Paris principalement, puis la petite couronne. La grande couronne étant

moins largement couverte, il est cependant remarquable que les enseignes concurrentes sont présentes dans de nombreuses communes. En revanche, les résultats sont très certainement faussés par le nombre très élevé de concurrents différents. Aussi une étude ou représentation individuelle par acteur ou enseigne semblerait plus appropriée.

Figure 20 : Présence d'AXA ou de la concurrence sur les communes d'Ile-de-France en 2014



L'exclusivité de la présence en revanche est intéressante à comparer. Si aucune commune d'Ile-de-France ne dispose d'une présence exclusive d'AXA, bon nombre de communes, notamment de la grande couronne, ne sont couvertes que par les concurrents ainsi que les premier et sixième arrondissements de Paris.

Infographie Sophie STAËS (2014), données internes AXA, réalisé sous Geoconcept V7

2.5. Conclusion

Grâce au croisement des données dont on dispose et d'une étude détaillée d'un territoire, il est aisément possible de bien connaître celui-ci. A partir de ces recherches il est plus facile de se faire une idée de la façon dont il faut s'implanter et se développer sur une zone.

Grâce aux rapides analyses et comparaisons qui ont été effectuées précédemment, il a été mis en avant que chaque département de la région Ile-de-France était différent des autres. La partie suivante va démontrer qu'ils sont aussi tout à fait différents sur le plan de la consommation en produits d'assurance.

3. L'OPTIMISATION DU TERRITOIRE GRÂCE À L'ANALYSE DU MARCHÉ

Comme il a été détaillé plus tôt, l'Ile-de-France est un territoire très étendu mais aussi très différent d'un département à l'autre, voire d'une commune à l'autre. Ainsi il est aisé de penser que les franciliens ont des profils socio-économiques mais aussi des consommations en produits d'assurance très différents en fonction de leur zone d'habitation.

L'analyse du marché a pour but de comprendre le consommateur ainsi que ses besoins afin de mieux y répondre. Elle nécessite des données précises et en nombre important, qui seront ensuite interprétées. Dans le cas de l'analyse du marché d'AXA, c'est le profil socio-économique des clients en Ile-de-France ainsi que leur consommation en produit d'assurance en fonction du département puis des arrondissements de Paris qui seront étudiés.

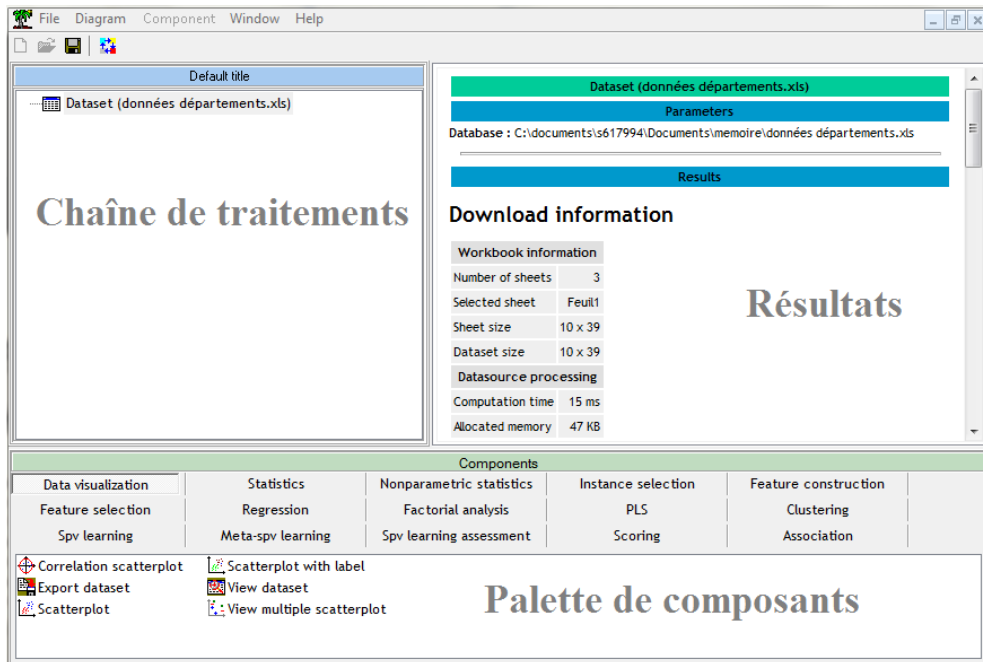
3.1. Introduction

Dans cette nouvelle partie, c'est l'utilisation du logiciel de Datamining Tanagra 1.4 qui va permettre d'analyser les différences entre les populations de la région et l'impact qu'ont celles-ci sur la consommation de produits d'assurance. L'extraction de connaissances à partir de données, ou Datamining, consiste en l'exploration des données dont dispose une entreprise via des logiciels automatiques ou semi-automatiques. Elle permet de mettre en évidence des corrélations entre certaines données, formuler des hypothèses ou encore rassembler des éléments ayant des caractéristiques statistiques proches.

Grâce à ces rapides analyses nous serons en mesure de déterminer si la recherche d'un modèle d'implantation optimal en quatrième partie s'effectuera sur l'ensemble de la région Ile-de-France, ou bien si au contraire il sera plus judicieux de se limiter à un territoire moins étendu.

La figure 21 montre la façon dont se présente le logiciel Tanagra. Celui-ci est décomposé en trois encadrés : la chaîne de traitements, la palette de composants et les résultats.

Figure 21 : Présentation du logiciel Tanagra



Capture d'écran Tanagra 1.4.41 après chargement des données départements

Pour commencer et avant chaque analyse, l'opérateur « Define status » est ajouté. Celui-ci permet de choisir les variables que l'on souhaite analyser. Une nouvelle fenêtre s'ouvre dans laquelle se trouve deux onglets : Target et Input. Les variables que l'on choisit d'étudier y sont ajoutées, les variables à décrire dans l'onglet Target et les variables descriptives (celles qui vont tenter d'expliquer celles en Target) dans l'onglet Input. Une fois cette opération réalisée, c'est un composant de la palette de composants qui est glissé dans la chaîne de traitements. Chaque composant propose une analyse et donc un résultat différent. Ce dernier s'affiche alors dans le cadre Résultats.

3.2. Le profil socio-économique de la région

Dans un premier temps ce sont les différents profils socio-économiques de la région d'Ile-de-France qui seront analysés, en commençant par les caractéristiques des départements puis celles des communes. Cette rapide analyse permettra de comprendre comment se répartie la population sur la région et quelles sont les tranches d'âges et les catégories socio-professionnelles qui caractérisent chaque département et chaque commune.

Pour cela les départements (comme les communes) vont être regroupés en plusieurs groupes aux caractéristiques proches grâce à la méthode des K-means (méthode des centres mobiles). Cette méthode, réalisée grâce à Tanagra, permet de rassembler des départements en

groupes de départements après l'analyse des données. Un certain nombre de clusters (groupes) est ainsi créé, regroupant des départements aux caractéristiques proches. Cette méthode présente cependant certains inconvénients, tels que la nécessité de fixer à l'avance le nombre de clusters sans indications préalables.

3.2.1. Les départements d'Ile-de-France

Le premier fichier utilisé est celui des départements (annexe 2), pour lesquels une analyse préalable et rapide de la population permettra par la suite de mieux comprendre la consommation en produits d'assurance de ces zones. Cette base de données est tirée de la base de données interne AXA, et est actualisée au 1^{er} janvier 2014. Ce fichier contient les informations population, nombre de ménages, âge, catégorie socio-professionnelles, nombre d'enfants, indice de revenu moyen ainsi que le nombre de clients AXA (particuliers et professionnels).

3.2.1.1. Analyse descriptive

On intègre alors les données départements au logiciel. Après avoir ajouté le composant « More Univariate cont stat 1 » qui permet une analyse succincte et univariée, les premiers résultats sont des présentations de variables. La figure 22, ci-dessous, permet par exemple de rendre compte des différences de revenus d'un département à l'autre. Un département a un indice de revenu moyen inférieur à 97,4 (100 étant la moyenne nationale), parallèlement trois départements de la région disposent d'un indice de revenu moyen supérieur à 148. De plus, huit départements sur neuf ont un indice de revenu moyen supérieur à la moyenne nationale.

Figure 22 : Analyse de la variable continue « Indice de revenu moyen »

Attribute	Stats		Histogram			
	Statistics		Values	Count	Percent	Histogram
Indice Rev moy men	Average	126,0000	x < 97,4000	1	11,11%	
	Median	119,0000	97,4000 =< x < 103,8000	0	0,00%	
	Std dev. [Coef of variation]	22,0227 [0,1748]	103,8000 =< x < 110,2000	1	11,11%	
	MAD [MAD/STDDEV]	17,5556 [0,7972]	110,2000 =< x < 116,6000	0	0,00%	
	Min * Max [Full range]	91,00 * 155,00 [64,00]	116,6000 =< x < 123,0000	3	33,33%	
	1st * 3rd quartile [Range]	117,00 * 149,00 [32,00]	123,0000 =< x < 129,4000	1	11,11%	
	Skewness (std-dev)	0,0740 (0,7171)	129,4000 =< x < 135,8000	0	0,00%	
	Kurtosis (std-dev)	-0,9701 (1,3997)	135,8000 =< x < 142,2000	0	0,00%	
			142,2000 =< x < 148,6000	0	0,00%	
			x >= 148,6000	3	33,33%	

Capture d'écran, résultat de l'analyse « More Univariate cont stat 1 » Tanagra 1.4.41

Cette analyse est bien entendu effectuée sur l'ensemble des variables quantitatives (continues) de la base de données. Cependant elles sont moins intéressantes.

On recherche ensuite une éventuelle corrélation entre chaque variable de la base de données département grâce à une corrélation linéaire. Pour cela un nouveau statut est défini avec l'ensemble des variables continues dans l'onglet Input, et ces mêmes variables dans l'onglet Target. Le composant « Linear correlation 1 » est ajouté, on obtient alors les résultats présentés en figure 23.

Figure 23 : Corrélation linéaire entre les variables continues du fichier Département

Y	X	r
Population	Pop femmes	0,9993
Nb menages	Pop 25-29 ans	0,9992
Population	Pop hommes	0,9990
Pop 60-74 ans	Pop Retraites	0,9970
Pop femmes	Pop hommes	0,9968
Pop 30-39 ans	Pop femmes	0,9968
Pop 40-49 ans	Pop hommes	0,9954
Pop 10-14 ans	Pop 15-17 ans	0,9942
Population	Pop 30-39 ans	0,9942
Pop 60-74 ans	Fam 0 enfant	0,9936
Nb menages	Pop Retraites	0,9928
Nb menages	Pop Artis commerc	0,9925
Pop Retraites	Fam 0 enfant	0,9925
Population	Pop 50-59 ans	0,9922
Population	Pop 40-49 ans	0,9918
Pop 30-39 ans	Pop 60-74 ans	0,9914
Pop 50-59 ans	Pop femmes	0,9914
Pop 50-59 ans	Pop hommes	0,9914
Pop 25-29 ans	Pop Artis commerc	0,9906
Nb menages	Pop 75 ans et plus	0,9903
Nb menages	Pop 60-74 ans	0,9900
Pop 30-39 ans	Pop Retraites	0,9899
Nb CII PartPro en cours	Pop 75 ans et plus	0,9894

Plus la valeur « r » est élevée (proche de 1), plus les variables sont corrélées entre elles. Si toutes les données population sont logiquement liées, on remarque que le nombre de clients particuliers et professionnels est fortement et positivement corrélé avec la population de plus de 75 ans (encadré vert). En extrapolant cette information, cela signifierait que plus le nombre de retraités est élevé sur un département, plus il y a de clients. Toutefois cette variable (nombre de clients) peut être liée à bon nombre d'autres.

Capture d'écran, résultat de l'analyse « Linear correlation 1 » Tanagra 1.4.41

A la suite de cela, le composant « Group Characterization » caractérise un par un les départements du fichier en calculant des statistiques comparatives avec les autres départements. Si cette étape est instructive sur la composition de chacun d'eux, elle est relativement limitée puisqu'elle ne propose que des comparaisons univariées (variable par variable). Ainsi sur la figure 24, la colonne « Group » renseigne la valeur de la variable sur le département étudié et est comparée à la colonne « Overall » qui donne la moyenne générale (ensemble des départements de la région Ile-de-France). Enfin le « Test value » est un indicateur qui permet de hiérarchiser les variables selon leur pertinence, proposant ainsi en première ligne celle qui caractérise le mieux le département par rapport aux autres.

Figure 24 : Description des départements par le composant « Group Characterization »

Nom=ESSONNE				Nom=VAL D'OISE				Nom=OISE				Nom=YVELINES			
Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1	
Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
Fam 2 enfants	0,10	75394,00 (-99999,00)	74227,89 (11712,85)	Fam 3+ enfants	0,27	42266,00 (-99999,00)	40296,56 (7338,38)	Pop Agriculteurs	1,88	4560,00 (-99999,00)	1430,22 (1667,22)	Indice Rev moy men	1,23	153,00 (-99999,00)	126,00 (22,02)
Pop Agriculteurs	-0,04	1367,00 (-99999,00)	1430,22 (1667,22)	Pop 10-14 ans	0,11	87049,00 (-99999,00)	85669,56 (12692,01)	Pop Ouvriers	-0,27	115055,00 (-99999,00)	122167,44 (25945,17)	Fam 2 enfants	1,11	87217,00 (-99999,00)	74227,89 (11712,85)
Indice Rev moy men	-0,05	125,00 (-99999,00)	126,00 (22,02)	Pop 15-17 ans	0,04	50079,00 (-99999,00)	49750,33 (7784,52)	Indice Rev moy men	-0,86	107,00 (-99999,00)	126,00 (22,02)	Pop 15-17 ans	1,09	58238,00 (-99999,00)	49750,33 (7784,52)
Pop Prof. Interm.	-0,13	153072,00 (-99999,00)	158962,44 (47030,72)	Pop 18-19 ans	-0,08	34085,00 (-99999,00)	34602,67 (6508,86)	Pop 75 ans et plus	-0,86	42327,00 (-99999,00)	75368,78 (38237,13)	Pop 10-14 ans	1,03	98763,00 (-99999,00)	85669,56 (12692,01)
Pop 50-59 ans	-0,31	139803,00 (-99999,00)	154772,56 (47621,50)	Pop Chomeurs	-0,16	51204,00 (-99999,00)	54560,00 (21281,96)	Pop PIS	-0,88	35374,00 (-99999,00)	132475,78 (109914,88)	Fam 3+ enfants	0,96	47353,00 (-99999,00)	40296,56 (7338,38)
Pop PIS	-0,33	96593,00 (-99999,00)	132475,78 (109914,88)	Fam 2 enfants	-0,19	72006,00 (-99999,00)	74227,89 (11712,85)	Pop Chomeurs	-0,95	34406,00 (-99999,00)	54560,00 (21281,96)	Pop 18-19 ans	0,74	39405,00 (-99999,00)	34602,67 (6508,86)
Pop 10-14 ans	-0,33	81517,00 (-99999,00)	85669,56 (12692,01)	Pop Ouvriers	-0,21	116717,00 (-99999,00)	122167,44 (25945,17)	Pop 25-29 ans	-1,04	55383,00 (-99999,00)	108531,56 (51151,18)	Pop 0-9 ans	0,59	178231,00 (-99999,00)	161033,56 (28989,50)
Fam 0 enfant	-0,36	126985,00 (-99999,00)	142301,00 (42337,54)	Pop Agriculteurs	-0,32	899,00 (-99999,00)	1430,22 (1667,22)	Nb menages	-1,06	280256,00 (-99999,00)	532291,67 (237249,28)	Pop 50-59 ans	0,34	171198,00 (-99999,00)	154772,56 (46261,50)
Pop 15-17 ans	-0,38	46759,00 (-99999,00)	49750,33 (7784,52)	Pop 20-24 ans	-0,33	76047,00 (-99999,00)	86580,78 (31879,91)	Pop Artis commerc	-1,09	18282,00 (-99999,00)	33703,00 (14202,37)	Pop Prof. Interm.	0,29	172496,00 (-99999,00)	158962,44 (47030,72)
Pop 20-24 ans	-0,40	73807,00 (-99999,00)	86580,78 (31879,91)	Pop 0-9 ans	-0,38	150033,00 (-99999,00)	161033,56 (28989,50)	Pop Retraites	-1,15	111553,00 (-99999,00)	183916,33 (62923,61)	Pop 40-49 ans	0,27	204374,00 (-99999,00)	192134,33 (46095,91)
Pop 60-74 ans	-0,42	121427,00 (-99999,00)	141009,33 (46826,18)	Pop Employes	-0,38	174908,00 (-99999,00)	193068,22 (47560,47)	Pop 60-74 ans	-1,20	84827,00 (-99999,00)	141009,33 (46826,18)	Pop hommes	0,20	664680,00 (-99999,00)	631308,33 (165554,05)
Fam 1 enfant	-0,43	76752,00 (-99999,00)	85042,33 (19503,29)	Indice Rev moy men	-0,41	117,00 (-99999,00)	126,00 (22,02)	Fam 0 enfant	-1,25	89457,00 (-99999,00)	142301,00 (42337,54)	Fam 0 enfant	0,18	149785,00 (-99999,00)	142301,00 (42337,54)

Nom=SEINE-ET-MARNE				Nom=HAUTS-DE-SEINE				Nom=SEINE-ST-DENIS				Nom=VAL-DE-MARNE				Nom=PARIS			
Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1	
Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
Pop Agriculteurs	1,48	3892,00 (-99999,00)	1430,22 (1667,22)	Indice Rev moy men	1,04	149,00 (-99999,00)	126,00 (22,02)	Pop Ouvriers	2,39	184189,00 (-99999,00)	122167,44 (25945,17)	Pop Employes	0,11	198522,00 (-99999,00)	193068,22 (47560,47)	Pop Artis commerc	2,47	68814,00 (-99999,00)	33703,00 (14202,37)
Pop Ouvriers	0,68	139860,00 (-99999,00)	122167,44 (25945,17)	Pop PIS	0,61	199671,00 (-99999,00)	132475,78 (109914,88)	Fam 3+ enfants	1,69	52729,00 (-99999,00)	40296,56 (7338,38)	Pop Retraites	0,06	187550,00 (-99999,00)	183916,33 (62923,61)	Pop 25-29 ans	2,45	234054,00 (-99999,00)	108531,56 (51151,18)
Pop 10-14 ans	0,65	93940,00 (-99999,00)	85669,56 (12692,01)	Fam 1 enfant	0,55	95733,00 (-99999,00)	85042,33 (19503,29)	Pop Chomeurs	1,16	79326,00 (-99999,00)	54560,00 (21281,96)	Fam 1 enfant	0,04	85750,00 (-99999,00)	85042,33 (19503,29)	Nb menages	2,44	1110912,00 (-99999,00)	532291,67 (237249,28)
Pop 15-17 ans	0,61	54491,00 (-99999,00)	49750,33 (7784,52)	Pop 0-9 ans	0,54	176780,00 (-99999,00)	161033,56 (28989,50)	Pop 10-14 ans	1,11	99791,00 (-99999,00)	85669,56 (12692,01)	Pop 75 ans et plus	0,01	75696,00 (-99999,00)	75368,78 (38237,13)	Pop 75 ans et plus	2,44	168544,00 (-99999,00)	75368,78 (38237,13)
Fam 2 enfants	0,46	79570,00 (-99999,00)	74227,89 (11712,85)	Pop Retraites	0,53	217539,00 (-99999,00)	183916,33 (62923,61)	Pop 15-17 ans	1,05	57955,00 (-99999,00)	49750,33 (7784,52)	Pop 60-74 ans	-0,02	140068,00 (-99999,00)	141009,33 (46826,18)	Pop 20-24 ans	2,40	163022,00 (-99999,00)	86580,78 (31879,91)
Fam 3+ enfants	0,44	43540,00 (-99999,00)	40296,56 (7338,38)	Pop 75 ans et plus	0,53	95564,00 (-99999,00)	75368,78 (38237,13)	Pop 0-9 ans	0,99	189622,00 (-99999,00)	161033,56 (28989,50)	Pop Prof. Interm.	-0,07	155571,00 (-99999,00)	158962,44 (47030,72)	Pop PIS	2,38	393883,00 (-99999,00)	132475,78 (109914,88)
Pop 18-19 ans	0,23	36108,00 (-99999,00)	34602,67 (6508,86)	Fam 0 enfant	0,51	163960,00 (-99999,00)	142301,00 (42337,54)	Pop Employes	0,90	230660,00 (-99999,00)	193068,22 (47560,47)	Fam 0 enfant	-0,09	138336,00 (-99999,00)	142301,00 (42337,54)	Pop Retraites	2,32	330151,00 (-99999,00)	183916,33 (62923,61)
Pop 0-9 ans	-0,07	159042,00 (-99999,00)	161033,56 (28989,50)	Pop 30-39 ans	0,50	247109,00 (-99999,00)	212485,33 (68570,70)	Pop 18-19 ans	0,76	39542,00 (-99999,00)	34602,67 (6508,86)	Pop 30-39 ans	-0,11	204781,00 (-99999,00)	212485,33 (68570,70)	Fam 0 enfant	2,30	239665,00 (-99999,00)	142301,00 (42337,54)
Pop 40-49 ans	-0,09	188095,00 (-99999,00)	192134,33 (46095,91)	Pop 60-74 ans	0,50	164540,00 (-99999,00)	141009,33 (46826,18)	Fam 1 enfant	0,43	93383,00 (-99999,00)	85042,33 (19503,29)	Pop Artis commerc	-0,12	31998,00 (-99999,00)	33703,00 (14202,37)	Pop 60-74 ans	2,30	248547,00 (-99999,00)	141009,33 (46826,18)
Pop Employes	-0,15	185843,00 (-99999,00)	193068,22 (47560,47)	Pop Prof. Interm.	0,41	178380,00 (-99999,00)	158962,44 (47030,72)	Pop 40-49 ans	0,29	205650,00 (-99999,00)	192134,33 (46095,91)	Nb menages	-0,14	499404,00 (-99999,00)	532291,67 (237249,28)	Pop femmes	2,28	1128929,00 (-99999,00)	670630,44 (201363,80)
Pop Artis commerc	-0,17	31325,00 (-99999,00)	33703,00 (14202,37)	Nb menages	0,39	624926,00 (-99999,00)	532291,67 (237249,28)	Fam 2 enfants	0,28	77554,00 (-99999,00)	74227,89 (11712,85)	Pop 25-29 ans	-0,16	100545,00 (-99999,00)	108531,56 (51151,18)	Pop 50-59 ans	2,28	263133,00 (-99999,00)	154772,56 (46261,50)
Fam 1 enfant	-0,21	80887,00 (-99999,00)	85042,33 (19503,29)	Pop femmes	0,37	746086,00 (-99999,00)	670630,44 (201363,80)	Pop hommes	0,28	671757,00 (-99999,00)	631308,33 (165554,05)	Pop femmes	-0,18	635255,00 (-99999,00)	670630,44 (201363,80)	Pop 30-39 ans	2,27	368002,00 (-99999,00)	212485,33 (68570,70)

Capture d'écran, résultat de l'analyse « Group Characterization » Tanagra 1.4.41

En sortent alors les informations descriptives de chaque département. Ainsi, et pour rester en surface, il est mis en évidence que :

- l'Essonne dispose de plus de familles de deux enfants que les autres départements.
- le Val-d'Oise présente une population plus jeune (10 – 17 ans) et les familles y sont plus grandes (plus de trois enfants).
- l'Oise est le département le plus peuplé en agriculteurs puisqu'il s'agit du département le plus ruralisé.

- les Yvelines ont une population plus riche (IRM⁷⁷ élevé), jeune et renferment des familles plutôt nombreuses.
- la Seine-et-Marne est aussi caractérisée par sa population d'agriculteurs mais aussi par son nombre d'ouvriers.
- les Hauts-de-Seine disposent d'une population relativement aisée, principalement de jeunes familles avec un enfant en bas âge et un nombre élevé de professions intellectuelles supérieures ou bien des retraités.
- la Seine-Saint-Denis est composée de nombreux ouvriers, chômeurs mais aussi de familles nombreuses (plus de trois enfants).
- le Val-de-Marne est très représenté en employés, retraités et familles avec peu d'enfants (un).
- enfin Paris est caractérisé par de nombreux artisans et commerçants, principalement des jeunes actifs de 20 à 30 ans ou des retraités (peu de familles avec des enfants en comparaison avec les autres départements) et un nombre de ménages élevé.

3.2.1.2. Création de clusters

A la suite de cette lecture rapide des données du fichier départements, vient l'étape de la création des clusters. Rappelons-le, la méthode des K-means permet de partitionner une population afin de créer des groupes aux caractéristiques proches.

Figure 25 : Répartition des départements dans trois clusters

Clusters		3	
Cluster	Description	Size	WSS
cluster n°1	c_kmeans_1	4	39,3643
cluster n°2	c_kmeans_2	4	31,9372
cluster n°3	c_kmeans_3	1	0,0000

La figure 25, ci-contre présente les premières informations obtenues après l'ajout du composant « K-Means 1 ». Dans cette opération de Datamining, il est donc choisi de structurer la base de données en trois classes de départements. Le premier groupe est

composé de quatre départements, le second de quatre également et enfin le troisième n'est formé que d'un seul département.

Capture d'écran, résultat de l'analyse « K-Means 1 » Tanagra 1.4.41

⁷⁷ Indice de revenu moyen

Afin de visualiser les trois groupes qui ont été créés, l'opération « Define status » est à nouveau effectuée. Dans l'onglet Input sont ajoutés l'ensemble des variables continues et dans l'onglet Target, la nouvelle variable discrète « Cluster_KMeans_1 » (qui s'est alors créée à l'étape précédente). Le composant « Group Characterization » est glissé dans la chaîne de traitements, on obtient alors trois groupes qui permettent de mieux comprendre la population d'Ile-de-France.

Figure 26 : Clusters auxquels appartiennent les départements d'Ile-de-France

	Nom	Cluster_Kmeans_1
1	ESSONNE	c_kmeans_1
2	VAL D'OISE	c_kmeans_1
3	OISE	c_kmeans_1
4	YVELINES	c_kmeans_2
5	SEINE-ET-MARNE	c_kmeans_1
6	HAUTS-DE-SEINE	c_kmeans_2
7	SEINE-ST-DENIS	c_kmeans_2
8	VAL-DE-MARNE	c_kmeans_2
9	PARIS	c_kmeans_3

En ajoutant l'opération « View Dataset 1 », il est possible de savoir quel département appartient à quel groupe (figure 26).

*Reproduction du résultat « View Dataset 1 »
Tanagra 1.4.41*

En effet cette étape permet de visualiser la table de données, sur laquelle a été ajoutée une dernière colonne contenant le numéro du cluster. Tanagra a ainsi choisi de regrouper l'Essonne, la Seine-et-Marne, le Val-d'Oise et l'Oise pour le groupe 1, puis les Yvelines, les Hauts-de-Seine, la Seine-Saint-Denis et le Val-de-Marne pour le deuxième groupe et enfin Paris forme le troisième groupe à lui seul.

Figure 27 : Caractéristiques des départements du cluster 1

Cluster_KMeans_1=c_kmeans_1			
Examples	[44,4 %] 4		
Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)			
Pop Agriculteurs	1,90	2679,50 (1816,52)	1430,22 (1667,22)
Pop Ouvriers	-0,23	119851,25 (13893,18)	122167,44 (25945,17)
Fam 3+ enfants	-0,97	37471,75 (6985,71)	40296,56 (7338,38)
Pop 10-14 ans	-1,04	80466,00 (14961,47)	85669,56 (12692,01)
Pop 15-17 ans	-1,07	46455,75 (8580,38)	49750,33 (7784,52)
Fam 2 enfants	-1,31	68148,50 (15331,68)	74227,89 (11712,85)
Pop 18-19 ans	-1,37	31073,50 (5978,33)	34602,67 (6508,86)
Pop PIS	-1,46	69146,75 (25265,71)	132475,78 (109914,88)
Pop Artis commerc	-1,49	25356,50 (5366,44)	33703,00 (14202,37)
Pop 20-24 ans	-1,51	67528,75 (15096,11)	86580,78 (31879,91)
Pop Prof. interm.	-1,54	130243,25 (34158,99)	158962,44 (47030,72)
Pop 25-29 ans	-1,57	76859,75 (14551,95)	108531,56 (51151,18)
Pop 75 ans et plus	-1,58	51445,75 (7132,77)	75368,78 (38237,13)
Pop Chomeurs	-1,60	41065,75 (7613,67)	54560,00 (21281,96)
Pop 40-49 ans	-1,67	161752,25 (30251,86)	192134,33 (46095,91)
Pop 50-59 ans	-1,68	123108,50 (24850,82)	154772,56 (47621,50)

Dans ce groupe, qui rappelons-le est composé de l'Essone, la Seine-et-Marne, le Val-d'Oise et l'Oise, la population est principalement caractérisée par un plus grand nombre d'agriculteurs que dans les autres départements. On remarque ensuite que l'ensemble des autres valeurs ont une valeur test négative, et la moyenne de ces valeurs est inférieure à la moyenne de la région. Ces départements présentent donc moins d'ouvriers, de familles nombreuses et d'adolescents (10 –

19 ans) qu'en moyenne en Ile-de-France.

Figure 28 : Caractéristiques des départements du cluster 2

Dans ce deuxième groupe, la caractéristique principale est la forte présence de jeunes enfants et de familles. En effet on remarque un nombre particulièrement élevé d'enfants de moins de neuf ans mais aussi de familles de deux à trois enfants et plus. Il s'agit d'une population plus ouvrière et employée que sur le reste de la région, on y retrouve aussi un nombre plus élevé de chômeurs. Ce groupe de département est composé de ceux de la petite couronne (Hauts-de-Seine, Seine-Saint-Denis, Val-de-Marne) et des Yvelines.

Cluster_KMeans_1=c_kmeans_2			
Examples		[44,4 %] 4	
Att - Desc	Test value	Group	Overrral
Continuous attributes : Mean (StdDev)			
Pop 0-9 ans	1,17	174439,25 (15326,58)	161033,56 (28989,50)
Fam 2 enfants	0,94	78583,00 (6352,72)	74227,89 (11712,85)
Fam 3+ enfants	0,90	42892,75 (8654,46)	40296,56 (7338,38)
Pop Employes	0,74	207001,00 (21317,19)	193068,22 (47560,47)
Fam 1 enfant	0,72	90619,25 (4709,84)	85042,33 (19503,29)
Pop 10-14 ans	0,70	89199,00 (11749,91)	85669,56 (12692,01)
Pop 15-17 ans	0,60	51599,75 (7602,79)	49750,33 (7784,52)
Pop 60-74 ans	0,39	148188,00 (11090,24)	141009,33 (46826,18)
Pop Ouvriers	0,36	125896,50 (39569,06)	122167,44 (25945,17)
Pop hommes	0,35	654033,75 (42223,02)	631308,33 (165554,05)
Pop 18-19 ans	0,34	35480,75 (4702,66)	34602,67 (6508,86)
Pop 30-39 ans	0,32	221028,25 (19232,08)	212485,33 (68570,70)
Pop Retraites	0,31	191684,25 (17526,85)	183916,33 (62923,61)
Pop femmes	0,29	694097,25 (45915,69)	670630,44 (201363,80)
Pop 40-49 ans	0,29	197506,25 (12073,25)	192134,33 (46095,91)
Pop Chomeurs	0,29	56984,75 (15674,11)	54560,00 (21281,96)

Capture d'écran, résultat de l'opération « Group Characterization » Tanagra 1.4.41

Figure 29 : Caractéristiques des départements du cluster 3

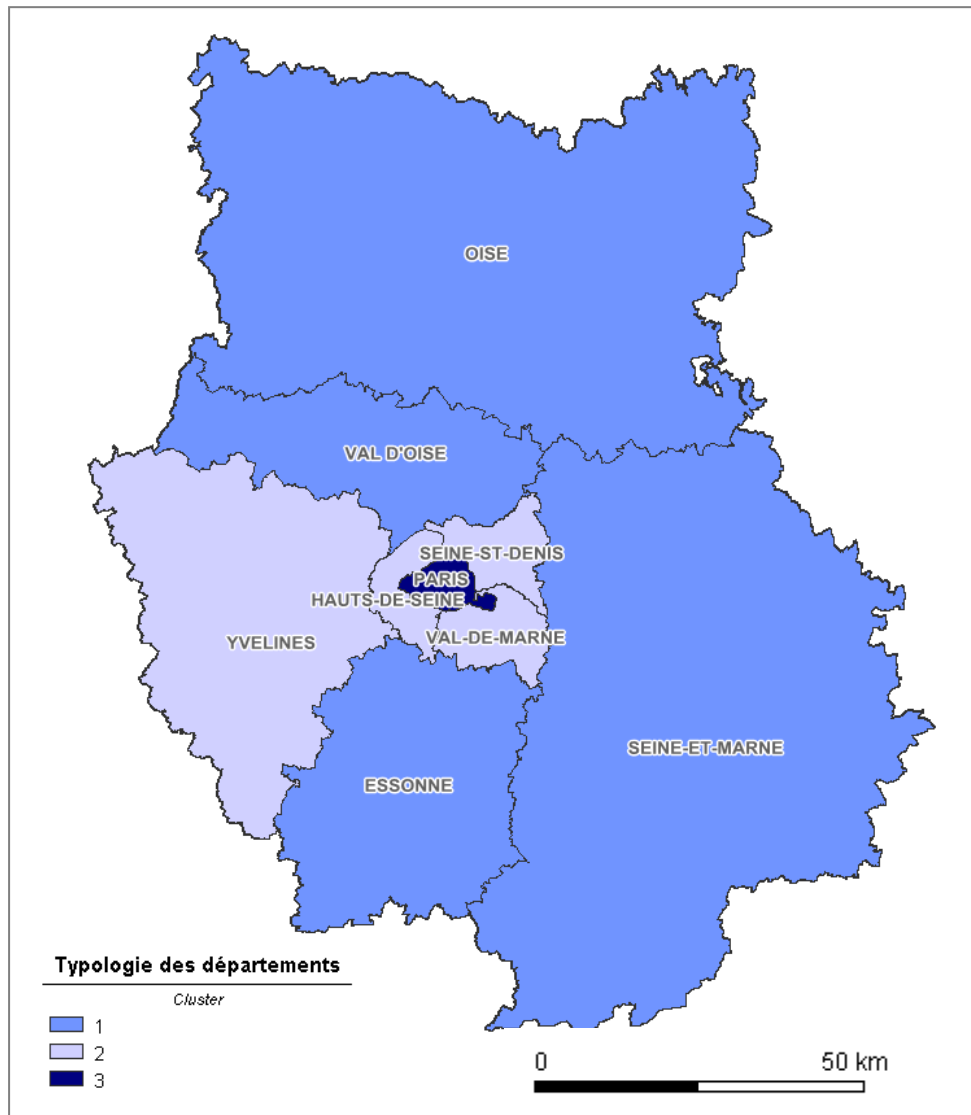
Cluster_KMeans_1=c_kmeans_3			
Examples		[11,1 %] 1	
Att - Desc	Test value	Group	Overrral
Continuous attributes : Mean (StdDev)			
Pop Artis commerc	2,47	68814,00 (-99999,00)	33703,00 (14202,37)
Pop 25-29 ans	2,45	234054,00 (-99999,00)	108531,56 (51151,18)
Pop 75 ans et plus	2,44	168544,00 (-99999,00)	75368,78 (38237,13)
Pop 20-24 ans	2,40	163022,00 (-99999,00)	86580,78 (31879,91)
Pop PIS	2,38	393883,00 (-99999,00)	132475,78 (109914,88)
Pop Retraites	2,32	330151,00 (-99999,00)	183916,33 (62923,61)
Fam 0 enfant	2,30	239665,00 (-99999,00)	142301,00 (42337,54)
Pop 60-74 ans	2,30	248547,00 (-99999,00)	141009,33 (46826,18)
Pop femmes	2,28	1128929,00 (-99999,00)	670630,44 (201363,80)
Pop 50-59 ans	2,28	263133,00 (-99999,00)	154772,56 (47621,50)
Pop 30-39 ans	2,27	368002,00 (-99999,00)	212485,33 (68570,70)
Pop hommes	2,21	996922,00 (-99999,00)	631308,33 (165554,05)
Pop 40-49 ans	2,17	292175,00 (-99999,00)	192134,33 (46095,91)
Pop Prof. interm.	2,14	259432,00 (-99999,00)	158962,44 (47030,72)
Pop Chomeurs	2,08	98838,00 (-99999,00)	54560,00 (21281,96)
Fam 1 enfant	1,90	122178,00 (-99999,00)	85042,33 (19503,29)

Enfin ce dernier groupe n'est formé que de Paris, commune département aux caractéristiques très spécifiques, elle se différencie par son nombre d'artisans et commerçants. On y retrouve une population plus jeune (20 à 29 ans) qui constitue les jeunes actifs, principalement cadres et professions intellectuelles supérieures, mais aussi des retraités de 75 ans et plus. Le nombre d'enfants y est faible puisque c'est le groupe dans lequel le nombre de famille sans enfant est le plus élevé.

Capture d'écran, résultat de l'opération « Group Characterization » Tanagra 1.4.41

Comme il a été vu précédemment, la table de données contenue dans Tanagra dispose à présent d'une colonne renseignant à quel cluster appartient un département. En exportant cette base (glisser « Export Dataset » dans la chaîne de traitements), puis en l'intégrant dans Geoconcept, il est possible de créer une thématique permettant de cartographier les trois clusters, comme sur la figure 30.

Figure 30 : Représentation cartographique des groupes de départements (clusters)



Infographie Sophie STAËS (2014), données exportées Tanagra, réalisé sous Geoconcept V7

A partir des groupes formés ci-dessus, une représentation cartographique permet une meilleure visualisation. Ainsi on distingue les départements aux caractéristiques proches et on remarque que si les départements de la petite couronne comme ceux de la grande se

regroupent majoritairement entre eux, celui des Yvelines est rattaché à ceux de la petite couronne.

Ce type d'analyse est le même sur n'importe quel territoire choisi. Ainsi, de la même façon, il va être possible d'analyser succinctement la population des communes d'Ile-de-France.

3.2.2. Le regroupement des communes

Dans cette nouvelle étape c'est un fichier contenant les données socio-économiques de l'ensemble des communes d'Ile-de-France qui est chargé dans Tanagra. Cette base de données de 1 993 communes contient exactement les mêmes informations que celles des départements.

Afin d'obtenir une segmentation par profils socio-économiques plus précise qu'en regroupant des départements, le même travail est donc réalisé avec l'ensemble des communes de la région.

Figure 31 : Répartition des communes dans huit clusters

Clusters 8			
Cluster	Description	Size	WSS
cluster n°1	c_kmeans_1	46	2060,1489
cluster n°2	c_kmeans_2	195	266,2655
cluster n°3	c_kmeans_3	885	506,4925
cluster n°4	c_kmeans_4	196	783,4468
cluster n°5	c_kmeans_5	15	229,9564
cluster n°6	c_kmeans_6	501	307,3501
cluster n°7	c_kmeans_7	10	1279,6245
cluster n°8	c_kmeans_8	145	1623,2023

Après avoir compris les données grâce aux comparaisons univariées, c'est l'opération de clustering qui a tout un intérêt. A ce niveau, huit groupes sont délimités, avec des groupes allant de 10 à 885 communes. La figure 31 ci-contre est, comme précédemment, le premier résultat. Celle-ci présente la répartition des communes dans chacun des huit groupes.

Capture d'écran, résultat de l'analyse « K-Means 1 » Tanagra 1.4.41

Le détail des huit groupes formés est expliqué ci-dessous grâce à la figure 32, et est limité aux premières variables, les plus explicatives. Enfin, une nouvelle représentation cartographique (figure 33) permet de visualiser les communes qui ont été regroupées entre elles.

Figure 32 : Caractéristiques des huit clusters de communes d’Ile-de-France

Cluster_KMeans_1=c_kmeans_1				Cluster_KMeans_1=c_kmeans_2				Cluster_KMeans_1=c_kmeans_3				Cluster_KMeans_1=c_kmeans_4			
Exemples				Exemples				Exemples				Exemples			
[2,3 %] 46				[9,8 %] 195				[44,4 %] 885				[9,8 %] 196			
Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
Fam 3+ enfants	27,30	2011,04 (729,87)	181,97 (459,67)	Population Agriculteurs	29,35	20,67 (5,89)	6,46 (7,12)	Indice Rev moy men	0,86	132,47 (32,95)	131,49 (45,53)	Fam 2 enfants	4,94	600,41 (229,86)	335,20 (791,63)
Population Ouvriers	26,93	6190,76 (2941,37)	551,68 (1436,49)	Population PIS	-3,02	74,13 (119,29)	598,23 (2551,24)	Revenu moyen par menage	0,85	38344,41 (9537,19)	38062,10 (13177,25)	Population 10-14 ans	4,51	671,39 (275,76)	386,87 (929,37)
Population 10-14 ans	26,64	3995,30 (1158,84)	386,87 (929,37)	Population 75 ans et plus	-3,03	94,67 (142,71)	340,35 (1191,58)	Population PIS	-8,45	58,03 (80,60)	598,23 (2551,24)	Population 15-17 ans	4,42	388,41 (161,49)	224,66 (545,90)
Population 0-9 ans	26,57	8004,61 (2262,58)	727,20 (1878,82)	Nb Cli PartProEnt en cours	-3,38	119,61 (173,06)	370,15 (1087,97)	Population 75 ans et plus	-10,02	40,90 (50,31)	340,35 (1191,58)	Population 18-19 ans	3,91	261,73 (112,96)	156,26 (397,22)
Population 15-17 ans	26,40	2325,46 (650,04)	224,66 (545,90)									Fam 3+ enfants	3,89	303,29 (152,80)	181,97 (459,67)

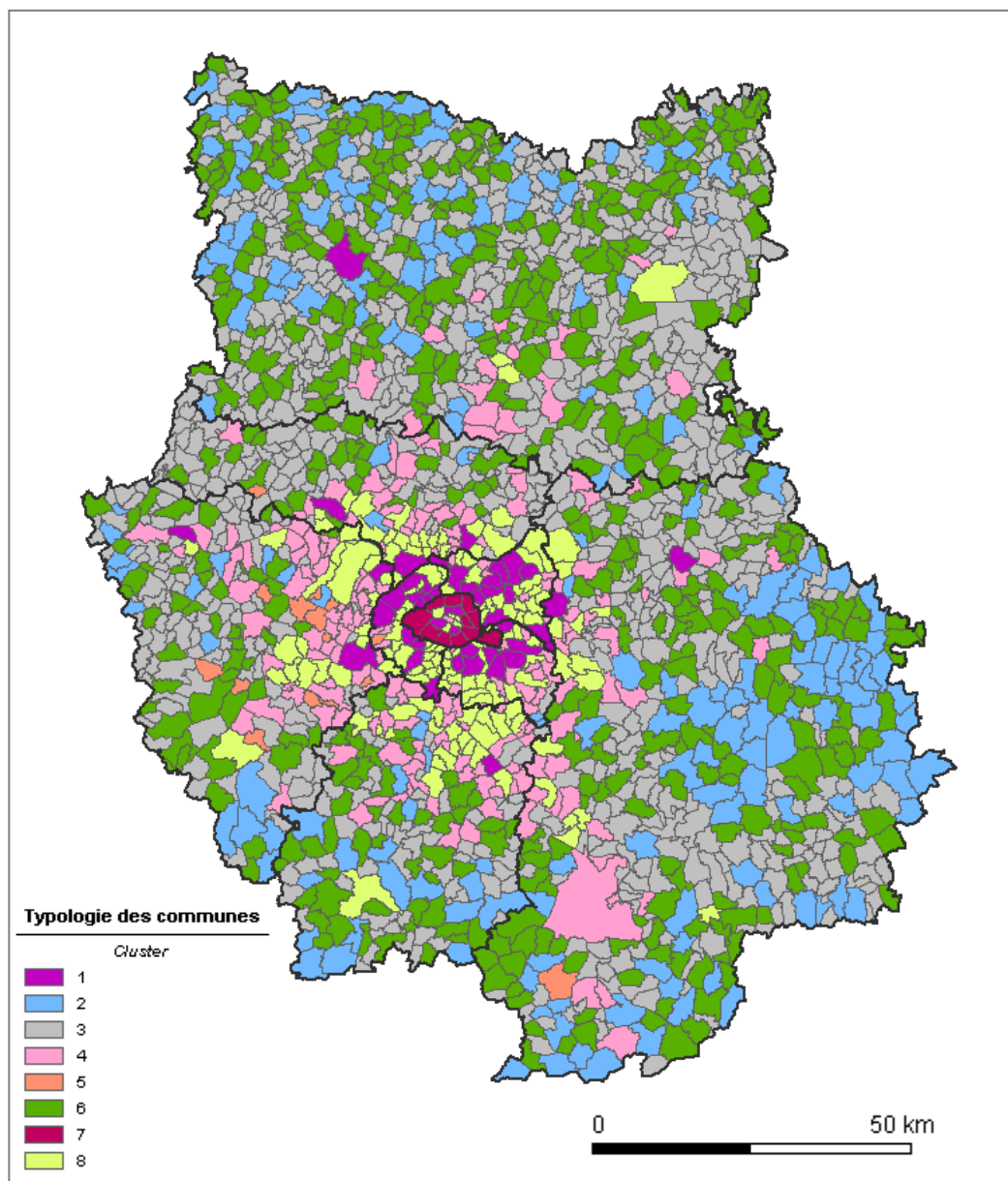
Cluster_KMeans_1=c_kmeans_5				Cluster_KMeans_1=c_kmeans_6				Cluster_KMeans_1=c_kmeans_7				Cluster_KMeans_1=c_kmeans_8			
Exemples				Exemples				Exemples				Exemples			
[0,8 %] 15				[25,1 %] 501				[0,5 %] 10				[7,3 %] 145			
Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
Revenu moyen par menage	24,77	122031,87 (52043,22)	38062,10 (13177,25)	Population Agriculteurs	10,87	9,45 (2,03)	6,46 (7,12)	Population PIS	35,87	29469,60 (8837,02)	598,23 (2551,24)	Fam 2 enfants	19,71	1583,33 (407,87)	335,20 (791,63)
Indice Rev moy men	24,76	421,53 (179,90)	131,49 (45,53)	Indice Rev moy men	-2,63	126,87 (31,20)	131,49 (45,53)	Population 75 ans et plus	34,51	13313,30 (3408,05)	340,35 (1191,58)	Population 10-14 ans	19,26	1818,97 (539,37)	386,87 (929,37)
Population PIS	-0,55	240,00 (302,96)	598,23 (2551,24)	Revenu moyen par menage	-2,63	36723,09 (9030,02)	38062,10 (13177,25)	Nb menages	34,24	86743,80 (15584,97)	2403,73 (7807,53)	Population 15-17 ans	19,00	1054,37 (319,70)	224,66 (545,90)
Population Artis commerc	-0,68	70,20 (71,58)	152,20 (470,60)	Population PIS	-5,49	56,54 (78,66)	598,23 (2551,24)	Population 25-29 ans	33,84	18138,70 (4093,81)	490,11 (1652,85)	Fam 3+ enfants	18,69	869,03 (366,14)	181,97 (459,67)
Nb Cli PartProEnt	-0,91	114,67 (132,33)	370,15 (1087,97)	Population 75 ans et plus	-6,38	46,43 (57,66)	340,35 (1191,58)	Population Artis commerc	33,12	5069,40 (1354,06)	152,20 (470,60)	Population 0-9 ans	18,23	3466,81 (1010,21)	727,20 (1878,82)
												Population Ouvriers	18,08	2629,41 (1294,80)	551,68 (1436,49)

Capture d’écran, résultat de l’opération « Group Characterization » Tanagra 1.4.41

- Le premier groupe est composé de communes peuplées de familles nombreuses, allant de trois enfants à plus. La population y est donc naturellement jeune et la catégorie socio-professionnelle principale est ouvrière. Représentées en violet sur la carte, ces communes sont principalement des communes de la petite couronne auxquelles s’ajoutent notamment Mantes-la-Jolie, Cergy, Beauvais, Meaux, Evry et Versailles.

- Le second groupe est représenté en bleu sur la carte. Ce sont des communes de la périphérie d’Ile-de-France, caractérisées par un taux d’agriculteurs plus fort que dans le reste de la région. C’est la principale caractéristique de ces communes cependant elles sont aussi fortement peuplées d’une population de 75 ans et plus.

Figure 33 : Représentation cartographique des groupes de communes (clusters)



Infographie Sophie STAËS (2014), données exportées Tanagra, réalisé sous Geoconcept V7

- Le troisième groupe, en gris, est le groupe le plus grand avec 885 communes. Il semble regrouper l'ensemble des communes qui n'entrent pas dans les autres groupes. Il est tout de même caractérisé par un indice de revenu moyen élevé (132 en moyenne) et une population aisée.
- Le quatrième groupe est représenté en rose sur la carte. Ce sont des communes très peuplées de la banlieue parisienne, principalement grande couronne. Elles sont caractérisées par un

grand nombre de familles de deux à trois enfants et plus et par conséquent par un nombre de jeunes très élevé (10 à 19 ans).

- Le cinquième groupe, lui, est composé de très peu de communes (15) regroupées majoritairement dans les Yvelines, en orange, et caractérisées par un indice de revenu moyen élevé et de nombreux cadres et professions intellectuelles supérieures. Le nombre d'artisans et commerçants y est aussi plus grand qu'ailleurs.

- Le sixième, en vert, est un second groupe représentatif des agriculteurs de la région parisienne, cependant le nombre y est bien moins élevé que sur le deuxième groupe.

- Le septième groupe, en rouge, n'est composé que de dix communes qui correspondent à dix arrondissements de Paris. Le nombre de ménages y est particulièrement élevé et la population y est plutôt constituée de cadres et professions intellectuelles supérieures. L'âge de la population, lui, est en majorité compris entre 25 et 29 ans ou atteint 75 ans et plus. Les caractéristiques sont donc les mêmes que celles retrouvées précédemment sur la commune - département de Paris.

- Enfin le huitième groupe est représenté sur la petite couronne de Paris, en jaune, et s'étend jusqu'à atteindre les bords de la grande couronne. La population y est ouvrière et jeune puisque composée de beaucoup de familles à deux, trois enfants ou plus.

3.3. La consommation de produits d'assurance

Mettre en avant les différences entre les habitants de chaque département, mais aussi de chaque commune de la région Ile-de-France, amène une autre question. Les produits d'assurance que privilégient les clients sont-ils les mêmes sur l'ensemble de la région ?

3.3.1. Une consommation différente d'un département à l'autre

Nous le verrons, la consommation de produits d'assurance n'est pas exactement la même d'un département à l'autre, et c'est toujours grâce au logiciel Tanagra, mais aussi avec les données clients d'AXA France que cela va être prouvé. Pour cette analyse, la base de données utilisée est le même fichier départements utilisé précédemment auquel on a ajouté des informations sur le nombre de clients (voir lignes ajoutées en annexe 3). Elle renseigne le nombre de clients total particuliers, professionnels et entreprises et le nombre de clients par branche de produits par département.

3.3.1.1. Analyse descriptive

Comme précédemment, le composant « Group Characterization » permet de dégager les informations les plus caractéristiques de chaque département, détaillées sur les deux figures suivantes.

Figure 34 : Description des départements par le composant « Group Characterization »

Nom=ESSONNE				Nom=VAL D'OISE				Nom=OISE				Nom=YVELINES			
Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1	
Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
Nb Cli Entreprise en cours	-0,30	3170,00 (-99999,00)	3956,22 (2591,89)	Nb Cli MRH	-0,62	36758,00 (-99999,00)	49005,89 (19741,95)	Nb Cli Automobile	0,31	27970,00 (-99999,00)	26695,56 (4049,06)	Nb Cli Automobile	1,29	31939,00 (-99999,00)	26695,56 (4049,06)
Nb Cli Epargne	-0,45	7241,00 (-99999,00)	8556,22 (2896,93)	Nb Cli Entreprise en cours	-0,62	2340,00 (-99999,00)	3956,22 (2591,89)	Nb Cli Banque	-0,05	4987,00 (-99999,00)	5030,11 (799,41)	Nb Cli Epargne	0,45	9849,00 (-99999,00)	8556,22 (2896,93)
Nb Cli Automobile	-0,47	24777,00 (-99999,00)	26695,56 (4049,06)	Nb Cli PartProEnt en cours	-0,66	62689,00 (-99999,00)	81967,33 (29390,31)	Nb Cli Prevoyance	-0,33	5177,00 (-99999,00)	5593,67 (1275,83)	Nb Cli Prevoyance	0,35	6042,00 (-99999,00)	5593,67 (1275,83)
Nb Cli PartProEnt en cours	-0,57	65123,00 (-99999,00)	81967,33 (29390,31)	Nb Cli PartPro en cours	-0,66	60349,00 (-99999,00)	78011,11 (26839,53)	Nb Cli Epargne	-0,33	7607,00 (-99999,00)	8556,22 (2896,93)	Nb Cli Sante	0,09	4648,00 (-99999,00)	4535,78 (1260,88)
Nb Cli MRH	-0,59	37417,00 (-99999,00)	49005,89 (19741,95)	Nb Cli Automobile	-0,71	23834,00 (-99999,00)	26695,56 (4049,06)	Nb Cli Sante	-0,42	4009,00 (-99999,00)	4535,78 (1260,88)	Nb Cli PartPro en cours	-0,04	77068,00 (-99999,00)	78011,11 (26839,53)
Nb Cli PartPro en cours	-0,60	61953,00 (-99999,00)	78011,11 (26839,53)	Nb Cli Sante	-0,81	3519,00 (-99999,00)	4535,78 (1260,88)	Nb Cli Entreprise en cours	-0,64	2287,00 (-99999,00)	3956,22 (2591,89)	Nb Cli PartProEnt en cours	-0,05	80414,00 (-99999,00)	81967,33 (29390,31)
Nb Cli Prevoyance	-0,64	4781,00 (-99999,00)	5593,67 (1275,83)	Nb Cli Banque	-0,92	4294,00 (-99999,00)	5030,11 (799,41)	Nb Cli PartProEnt en cours	-0,78	58983,00 (-99999,00)	81967,33 (29390,31)	Nb Cli MRH	-0,17	45712,00 (-99999,00)	49005,89 (19741,95)
Nb Cli Sante	-0,83	3492,00 (-99999,00)	4535,78 (1260,88)	Nb Cli Epargne	-1,00	5668,00 (-99999,00)	8556,22 (2896,93)	Nb Cli MRH	-0,79	33403,00 (-99999,00)	49005,89 (19741,95)	Nb Cli Banque	-0,19	4880,00 (-99999,00)	5030,11 (799,41)
Nb Cli Banque	-0,90	4312,00 (-99999,00)	5030,11 (799,41)	Nb Cli Prevoyance	-1,57	3596,00 (-99999,00)	5593,67 (1275,83)	Nb Cli PartPro en cours	-0,79	56696,00 (-99999,00)	78011,11 (26839,53)	Nb Cli Entreprise en cours	-0,24	3346,00 (-99999,00)	3956,22 (2591,89)
Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy			

Capture d'écran, résultat de l'analyse « Group Characterization » Tanagra 1.4.41

Il est alors mis en évidence que l'Essonne, comme le Val-d'Oise, avec des valeurs test négatives et des moyennes inférieures à celle de l'ensemble des départements, présentent nettement moins de clients que les autres départements, toutes branches confondues. L'Oise se démarque des autres départements par son nombre de clients assurés en automobile (tendance explicable par le fait que l'Oise est plus ruralisé que les autres départements et que la population ne s'y déplace pas en transport en commun comme sur la région Ile-de-France). Il en est de même pour les Yvelines qui est un département très étendu.

Figure 35 : Description des départements par le composant « Group Characterization » (suite)

Nom=SEINE-ET-MARNE				Nom=HAUTS-DE-SEINE				Nom=SEINE-ST-DENIS				Nom=VAL-DE-MARNE				Nom=PARIS			
Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1		Exemples		[11,1 %] 1	
Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral	Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
Nb Cli Banque	1,83	6490,00 (-99999,00)	5030,11 (799,41)	Nb Cli Epargne	0,67	10489,00 (-99999,00)	8556,22 (2896,93)	Nb Cli Prevoyance	0,54	6281,00 (-99999,00)	5593,67 (1275,83)	Nb Cli MRH	-0,21	44861,00 (-99999,00)	49005,89 (19741,95)	Nb Cli Entreprise en cours	2,59	10666,00 (-99999,00)	3956,22 (2591,89)
Nb Cli Automobile	1,22	31619,00 (-99999,00)	26695,56 (4049,06)	Nb Cli MRH	0,49	58640,00 (-99999,00)	49005,89 (19741,95)	Nb Cli MRH	-0,14	46152,00 (-99999,00)	49005,89 (19741,95)	Nb Cli PartPro en cours	-0,30	69988,00 (-99999,00)	78011,11 (26839,53)	Nb Cli PartProEnt en cours	2,49	155156,00 (-99999,00)	81967,33 (29390,31)
Nb Cli Sante	0,73	5453,00 (-99999,00)	4535,78 (1260,88)	Nb Cli PartPro en cours	0,45	90127,00 (-99999,00)	78011,11 (26839,53)	Nb Cli PartPro en cours	-0,29	70099,00 (-99999,00)	78011,11 (26839,53)	Nb Cli PartProEnt en cours	-0,30	73038,00 (-99999,00)	81967,33 (29390,31)	Nb Cli PartPro en cours	2,48	144490,00 (-99999,00)	78011,11 (26839,53)
Nb Cli Prevoyance	-0,02	5573,00 (-99999,00)	5593,67 (1275,83)	Nb Cli PartProEnt en cours	0,43	94475,00 (-99999,00)	81967,33 (29390,31)	Nb Cli PartProEnt en cours	-0,31	72972,00 (-99999,00)	81967,33 (29390,31)	Nb Cli Entreprise en cours	-0,35	3050,00 (-99999,00)	3956,22 (2591,89)	Nb Cli MRH	2,47	97841,00 (-99999,00)	49005,89 (19741,95)
Nb Cli Epargne	-0,06	8379,00 (-99999,00)	8556,22 (2896,93)	Nb Cli Automobile	0,31	27962,00 (-99999,00)	26695,56 (4049,06)	Nb Cli Entreprise en cours	-0,42	2873,00 (-99999,00)	3956,22 (2591,89)	Nb Cli Epargne	-0,38	7457,00 (-99999,00)	8556,22 (2896,93)	Nb Cli Sante	2,33	7476,00 (-99999,00)	4535,78 (1260,88)
Nb Cli Entreprise en cours	-0,17	3526,00 (-99999,00)	3956,22 (2591,89)	Nb Cli Banque	0,23	5215,00 (-99999,00)	5030,11 (799,41)	Nb Cli Sante	-0,44	3975,00 (-99999,00)	4535,78 (1260,88)	Nb Cli Prevoyance	-0,52	4935,00 (-99999,00)	5593,67 (1275,83)	Nb Cli Epargne	2,18	14869,00 (-99999,00)	8556,22 (2896,93)
Nb Cli PartProEnt en cours	-0,24	74856,00 (-99999,00)	81967,33 (29390,31)	Nb Cli Entreprise en cours	0,15	4348,00 (-99999,00)	3956,22 (2591,89)	Nb Cli Banque	-0,77	4416,00 (-99999,00)	5030,11 (799,41)	Nb Cli Sante	-0,55	3836,00 (-99999,00)	4535,78 (1260,88)	Nb Cli Prevoyance	2,09	8254,00 (-99999,00)	5593,67 (1275,83)
Nb Cli PartPro en cours	-0,25	71330,00 (-99999,00)	78011,11 (26839,53)	Nb Cli Prevoyance	0,09	5704,00 (-99999,00)	5593,67 (1275,83)	Nb Cli Epargne	-1,07	5447,00 (-99999,00)	8556,22 (2896,93)	Nb Cli Banque	-0,62	4533,00 (-99999,00)	5030,11 (799,41)	Nb Cli Banque	1,39	6144,00 (-99999,00)	5030,11 (799,41)
Nb Cli MRH	-0,44	40269,00 (-99999,00)	49005,89 (19741,95)	Nb Cli Sante	-0,10	4414,00 (-99999,00)	4535,78 (1260,88)	Nb Cli Automobile	-1,56	20396,00 (-99999,00)	26695,56 (4049,06)	Nb Cli Automobile	-1,04	22467,00 (-99999,00)	26695,56 (4049,06)	Nb Cli Automobile	0,64	29296,00 (-99999,00)	26695,56 (4049,06)
Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy			

Capture d'écran, résultat de l'analyse « Group Characterization » Tanagra 1.4.41

La Seine-et-Marne est un département sur lequel semble très bien se développer la banque, mais aussi l'assurance automobile (explicable par l'étendu du territoire et les communes rurales de l'est du département). Les quatre derniers départements, Hauts-de-Seine, Seine-Saint-Denis, Val-de-Marne et Paris, présentent un nombre de clients MRH⁷⁸ bien plus élevé que les autres départements. En effet il s'agit de la petite couronne de Paris et le nombre d'habitations y est plus élevé que sur la grande couronne. En plus de cela, les habitants des Hauts-de-Seine sont plus consommateurs de produits d'Epargne, ceux de Seine-Saint-Denis consomment plus de produits prévoyance. Quant au Val-de-Marne et à Paris, ils sont caractérisés par leur nombre très élevés de clients, entreprises pour le Val-de-Marne et toutes entités (particuliers, professionnels et entreprises) confondues pour Paris.

3.3.1.2. Analyse en composantes principales

Pour aller plus loin, l'analyse en composantes principales est une méthode d'analyse qui permet de représenter graphiquement les informations d'une base de données. Avec cette méthode seront comparés le nombre de clients par branche de produits d'assurance (Auto, MRH, Santé, Prévoyance, Epargne et Banque) et les variables quantitatives sur la population de chaque département. Ainsi on obtiendra des graphiques expliquant les liaisons entre ces variables.

⁷⁸ Multirisque Habitation

Les variables quantitatives sur la population sont analysées. Pour cela un nouveau statut est défini dans lequel sont ajoutées en Input toutes les variables relatives à la population et en Target, celles relatives au nombre de clients. Le composant « Principal Component Analysis 1 » est glissé dans la chaîne de traitements. Tanagra propose alors trois axes (qui correspondent à trois variables différentes) qui permettent d'expliquer plus ou moins la répartition de la consommation de produits d'assurance en Ile-de-France.

Figure 36 : Axes proposés par l'analyse en composantes principales

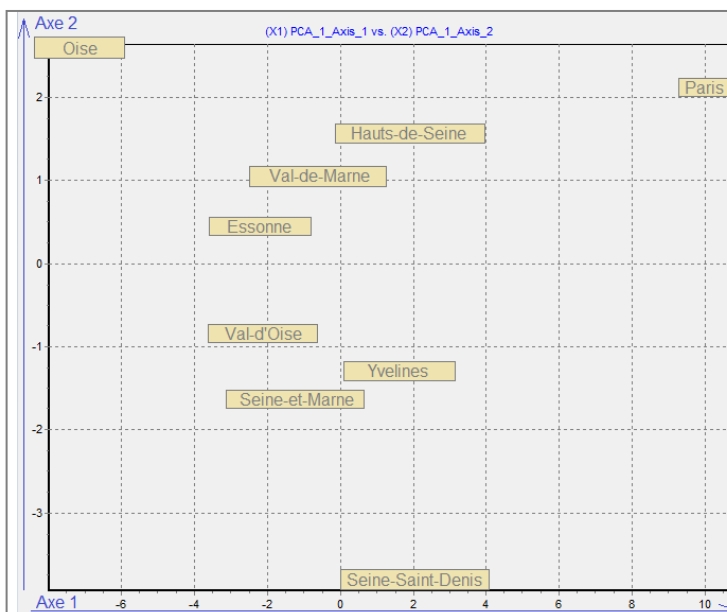
Eigen values				
Matrix trace = 28,00				
Axis	Eigen value	% explained	Histogram	% cumulated
1	21,459488	76,64%		76,64%
2	3,895610	13,91%		90,55%
3	1,554485	5,55%		96,11%

Cette première étape montre qu'un seul des trois axes, l'Axe 1, permet d'analyser plus de 70% des valeurs de consommation de produits d'assurance (76.64%).

Capture d'écran, résultat de l'analyse « Principal Component Analysis 1 » Tanagra 1.4.41

Avec cette étape, Tanagra a créé de nouvelles variables, les axes. Cependant il reste à rechercher à quelle variable correspond l'Axe 1 (qui est donc le plus intéressant). Pour cela, on ajoute le composant « Scatterplot 2 » qui nous permet d'obtenir un graphique. On renseigne alors « PCA_1_Axis_2 » en ordonnée, « PCA_1_Axis_1 » en abscisse et les départements en variable. Le résultat final est un graphique qui représente la répartition des départements de la région en fonction des Axes 1 (horizontal) et 2 (vertical).

Figure 37 : Répartition des départements d'Ile-de-France selon les axes 1 et 2



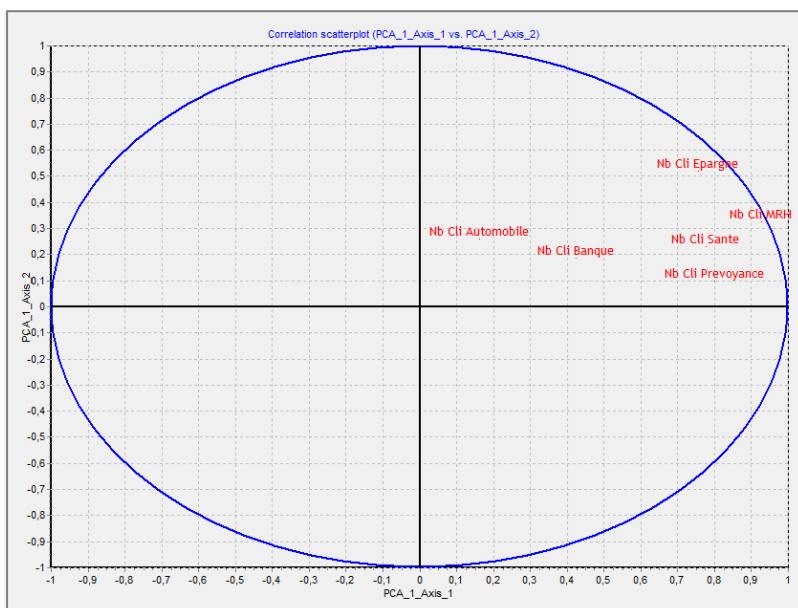
La variable correspondant à l'Axe 1 a une valeur très élevée sur Paris, en effet ce département se situe le plus à droite du graphique, en opposition à l'Oise qui se situe à la valeur la moins élevée de l'Axe 1. Cet Axe est donc une variable avec des valeurs de la plus élevée à la plus faible étant celles de Paris, suivi des Hauts-de-Seine, de la Seine-Saint-Denis, des Yvelines, du

Capture d'écran, résultat de l'opération « Scatterplot 2 » Tanagra 1.4.41

Val-de-Marne, de la Seine-et-Marne, de l'Essonne, du Val-d'Oise puis de l'Oise. Grâce à ce rapide classement, il semble apparaître que l'Axe 1 représente la population des départements. Il s'agit donc d'une possibilité d'explication de la répartition des clients par branche, ce qui signifierait que selon la densité de population d'un département, la consommation de produits d'assurance n'est pas la même.

A partir de cette composante principale qu'est la population, sont calculées les corrélations entre cette variable et les autres (à savoir le nombre de clients par branche). Pour cela on crée un nouveau statut : PCA_1_Axis_1 et PCA_1_Axis_2 en Target et les nombres de clients par branche en Input. On ajoute le composant « Correlation scatterplot 1 » à la chaîne de traitements. On obtient alors un cercle de corrélation sur lequel sont placées les variables : nombre de clients par branche. Plus celles-ci sont proches les unes des autres, plus elles sont corrélées, et il en est de même avec l'Axe 1 correspondant à la variable population.

Figure 38 : Cercle de corrélation entre l'Axe 1 et le nombre de clients par branche



Les branches Auto et Banque (au centre du cercle) sont proches l'une de l'autre mais également éloignées du cercle. Ceci peut vouloir dire qu'elles sont corrélées entre elles mais qu'elles ne caractérisent pas l'Axe 1 et donc ne sont pas expliquées par la population. Une explication logique en Ile-de-

France puisqu'une population élevée ne signifie pas toujours un grand nombre de voiture à assurer. En effet peu de franciliens utilisent une voiture (seulement 13,3% régulièrement)⁷⁹.

Capture d'écran, résultat de l'opération « Correlation scatterplot 1 » Tanagra 1.4.41

Le reste des variables sont relativement proches du cercle de corrélation, ce sont donc les variables les mieux représentées. Ce sont les branches Epargne, MRH et Santé qui

⁷⁹http://www.insee.fr/fr/themes/document.asp?ref_id=17224&page=alapage/alap353/alap353_graph.htm#graph3 (25 août 2014)

caractérisent le moins bien la variable correspondant à l'Axe 1 (qui rappelons-le, est la population des départements), car ce sont les plus éloignées de l'Axe.

Enfin la branche Prévoyance, quant à elle, est la plus proche de l'Axe 1 et se trouve seule dans son espace. Dans un premier temps cela signifie que c'est la branche la plus apte à caractériser l'Axe 1 et, dans un second temps, que les clients de cette branche ont un comportement unique, qui n'est lié à aucune autre branche.

3.3.1.3. Création de clusters

Dans le but de visualiser à présent sur une carte les départements aux consommations en produits d'assurance proches, on réalise à nouveau la méthode des K-means qui permet d'obtenir des clusters, comme il a été expliqué précédemment.

Figure 39 : Répartition des départements dans trois clusters

Clusters		3	
Cluster	Description	Size	WSS
cluster n°1	c_kmeans_1	4	3,6241
cluster n°2	c_kmeans_2	1	0,0000
cluster n°3	c_kmeans_3	4	5,9289

Dans cette analyse, trois clusters sont différenciés. Le premier et le troisième sont composés de quatre départements. Le second en revanche n'en comporte qu'un seul.

Capture d'écran, résultat de l'opération « K-Means 1 » Tanagra 1.4.41

Figure 40 : Caractéristiques des trois clusters de départements d'Ile-de-France

Cluster_KMeans_1=c_kmeans_1				Cluster_KMeans_1=c_kmeans_2				Cluster_KMeans_1=c_kmeans_3			
Examples [44,4 %] 4				Examples [11,1 %] 1				Examples [44,4 %] 4			
Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)			
Nb Cli MRH	-0,99	41297,00 (4896,60)	49005,89 (19741,95)	Nb Cli MRH	2,47	97841,00 (-99999,00)	49005,89 (19741,95)	Nb Cli Automobile	1,98	29872,50 (2205,31)	26695,56 (4049,06)
Nb Cli Prevoyance	-1,38	4898,25 (1098,93)	5593,67 (1275,83)	Nb Cli Sante	2,33	7476,00 (-99999,00)	4535,78 (1260,88)	Nb Cli Banque	1,15	5393,00 (744,56)	5030,11 (799,41)
Nb Cli Sante	-1,67	3705,50 (238,07)	4535,78 (1260,88)	Nb Cli Epargne	2,18	14869,00 (-99999,00)	8556,22 (2896,93)	Nb Cli Epargne	0,46	9081,00 (1321,33)	8556,22 (2896,93)
Nb Cli Epargne	-1,84	6453,25 (1041,99)	8556,22 (2896,93)	Nb Cli Prevoyance	2,09	8254,00 (-99999,00)	5593,67 (1275,83)	Nb Cli Sante	0,19	4631,00 (608,26)	4535,78 (1260,88)
Nb Cli Banque	-2,03	4388,75 (110,18)	5030,11 (799,41)	Nb Cli Banque	1,39	6144,00 (-99999,00)	5030,11 (799,41)	Nb Cli Prevoyance	0,06	5624,00 (357,55)	5593,67 (1275,83)
Nb Cli Automobile	-2,39	22868,50 (1901,67)	26695,56 (4049,06)	Nb Cli Automobile	0,64	29296,00 (-99999,00)	26695,56 (4049,06)	Nb Cli MRH	-0,58	44506,00 (10684,15)	49005,89 (19741,95)
Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy			

Capture d'écran, résultat de l'opération « Group Characterization » Tanagra 1.4.41

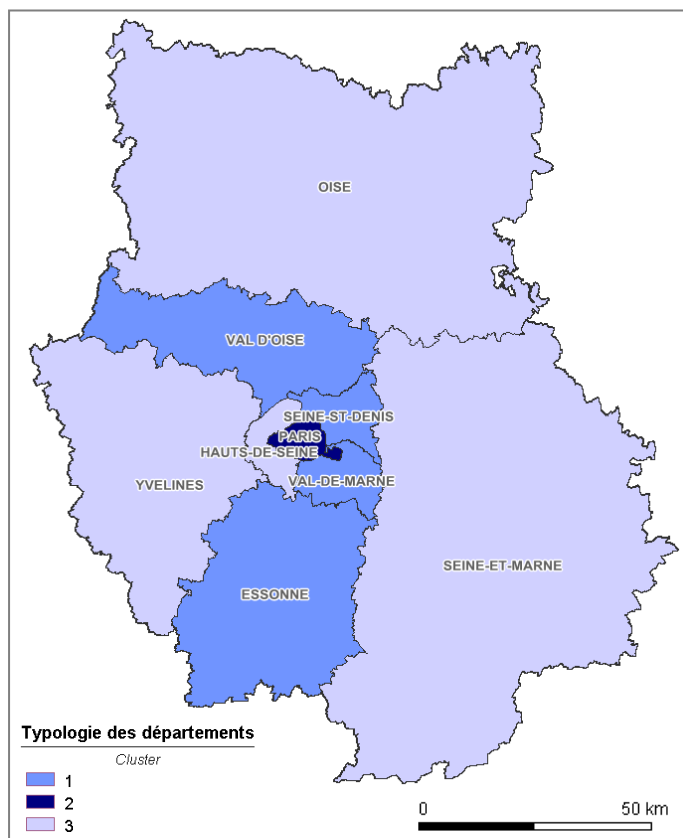
- Dans le premier cluster, quatre départements sont représentés. Il s'agit (comme il est visualisable sur la carte ci-dessous), du Val-d'Oise, de la Seine-Saint-Denis, du Val-de-Marne et de l'Essonne. La dominance de la branche de produits MRH est nettement remarquable, suivi des produits Prévoyance et Santé. En revanche les branches de produits Auto et Banque apparaissent en dernière position. Il en est concluable que pour ces départements l'habitation est primordiale devant la Banque et l'assurance automobile, certainement parce que bon nombre d'entre-eux ne dispose pas de voiture. Parallèlement à cela, au niveau de la quantité,

le nombre de clients sur l'ensemble des branches de ces départements est beaucoup plus faible qu'en moyenne sur la région.

- Au niveau du second cluster, un seul département est représenté, il s'agit de Paris qui présente donc des clients avec un type de consommation en produits d'assurance unique. On y observe une dominance des branches de produits MRH, Santé puis Epargne. En revanche, et comme précédemment, les branches de produits Banque et Auto viennent en dernier. Très peu de Parisiens ont une voiture, alors que le nombre d'habitations y est très élevé et que l'épargne et la prévoyance sont primordiales pour les clients de ce département.

- Enfin le troisième cluster, composé lui aussi de quatre départements, est formé de l'Oise, des Yvelines, des Hauts-de-Seine et de la Seine-et-Marne. Sur ces départements c'est l'assurance Auto qui domine. En effet ce sont des départements plus riches ou plus éloignés de Paris, au sein desquels les ménages sont plus équipés en automobiles. La branche Banque se situe en deuxième position ce qui signifie que les opérations bancaires y ont une place de choix. Au contraire, ce sont les branches de produits Prévoyance et MRH qui viennent en dernière position. Excepté pour la branche de produit MRH, le poids des clients de ces départements est plus fort que sur les autres départements.

Figure 41 : Représentation cartographique des groupes de départements (clusters)



A partir de ces données qu'on aura exportées, une carte réalisée sous Geoconcept permet la visualisation des départements regroupés en cluster.

Pour la suite, l'analyse de la consommation de produits d'assurance se limite aux arrondissements de Paris. En effet il a été étudié précédemment que les communes d'Ile-de-France sont très nombreuses et très différentes les unes des autres. Il serait donc difficile de ne les regrouper qu'en quelques clusters.

3.3.2. Une consommation particulière à Paris

La consommation de produits d'assurance sur Paris est particulière. Chaque arrondissement se comportant comme une commune, il en ressort des comportements de consommation différents. Mais ces comportements sont aussi très différents de ceux qui existent dans les autres communes de la région (celles des autres départements).

Pour cette nouvelle analyse le fichier utilisé est le même que celui des communes utilisé précédemment, il est cependant limité aux arrondissements de Paris et le nombre de clients par branche y a été ajouté.

3.3.2.1. Analyse descriptive

Figure 42 : Corrélation linéaire entre le nombre de clients AXA et les variables population

Y	X	r
Nb Cli PartProEnt en cours	Population 75 ans et plus	0,9763
Nb Cli PartProEnt en cours	Population Artis commerc	0,9643
Nb Cli PartProEnt en cours	Population Retraites	0,9477
Nb Cli PartProEnt en cours	Population PIS	0,9416
Nb Cli PartProEnt en cours	Fam 2 enfants	0,9232
Nb Cli PartProEnt en cours	Population 40-49 ans	0,8917

*Capture d'écran, résultat de l'opération
« Linear Correlation 1 » Tanagra 1.4.41*

Une simple corrélation linéaire permet de dégager une tendance sur Paris. Celle-ci met en avant le fait que le nombre de clients AXA est fortement corrélé avec le nombre de plus de 75 ans, ainsi que le nombre de retraités. Il est aussi lié au nombre d'artisans-commerçants (qui est plus élevé sur Paris que sur les autres

départements) et de professions intellectuelles supérieures.

Figure 43 : Description du 1^{ER} arrondissement par le composant « Group Characterization »

Nom=PARIS--1ER-ARRONDISSEMENT			
Examples		[5,0 %] 1	
Att - Desc	Test value	Group	Overral
Continuous attributes : Mean (StdDev)			
Nb Cli Automobile	-1,19 (-99999,00)	309,00	1464,80 (968,06)
Nb Cli Epargne	-1,21 (-99999,00)	189,00	743,45 (458,51)
Nb Cli Prevoyance	-1,22 (-99999,00)	132,00	412,70 (229,68)
Nb Cli Banque	-1,31 (-99999,00)	56,00	307,20 (191,59)
Nb Cli MRH	-1,32 (-99999,00)	1092,00	4892,05 (2869,57)
Nb Cli PartProEnt en cours	-1,34 (-99999,00)	2182,00	7757,80 (4168,71)
Nb Cli Sante	-1,38 (-99999,00)	89,00	373,80 (206,51)
Discrete attributes : [Recall] Accuracy			

Malgré cela, des comportements spécifiques par arrondissement existent. Chacun a ses particularités et, comme analysé précédemment (analyse des communes d'Ile-de-France), des populations différentes qui entraînent une consommation en produits d'assurance spécifique à chaque arrondissement.

Ces différences sont constatables grâce au composant « Group Characterization », qui décrit les arrondissements de Paris un par un.

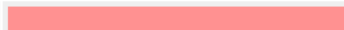


*Capture d'écran, résultat de l'opération
« Group Characterization » Tanagra 1.4.41*

La figure 43 est l'exemple du premier arrondissement, dans lequel les assurances automobiles dominent et la Santé vient en dernière position. Cependant en termes quantitatifs, cet arrondissement présente nettement moins de clients que les autres.

3.3.2.2. Analyse en composantes principales

De nouveau, et puisque la consommation en produits d'assurance sur les arrondissements de Paris y est spécifique, une analyse en composante principale est réalisée afin de mettre en avant les variables qui expliquent le plus la consommation de produit d'assurance sur ce département et les corrélations entre les branches.

Figure 44 : Axes proposés par l'analyse en composantes principales

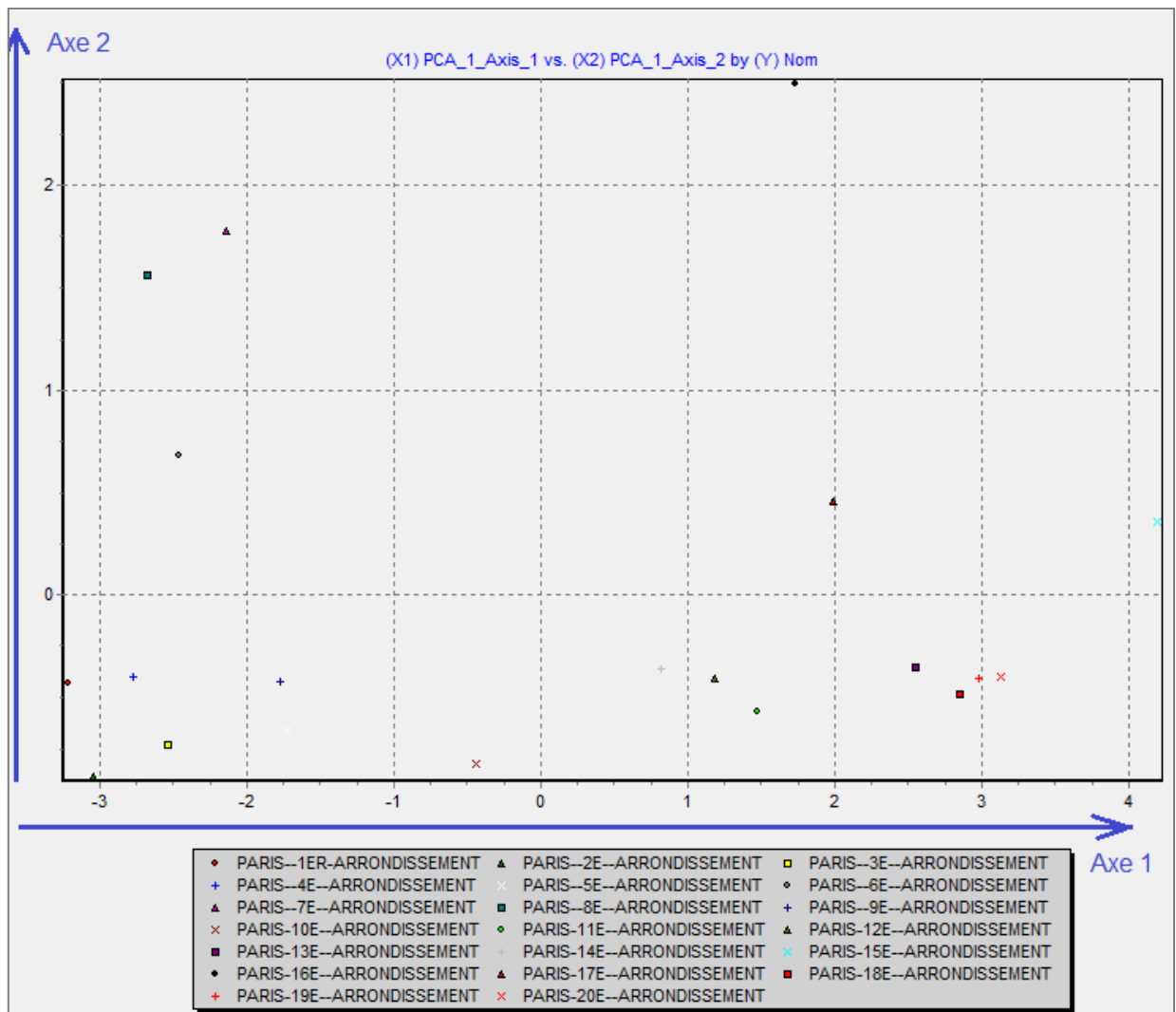
Eigen values				
Matrix trace = 7,00				
Axis	Eigen value	% explained	Histogram	% cumulated
1	5,994286	85,63%		85,63%
2	0,839071	11,99%		97,62%
3	0,149556	2,14%		99,76%
4	0,011651	0,17%		99,92%
5	0,004404	0,06%		99,99%
6	0,000739	0,01%		100,00%
7	0,000292	0,00%		100,00%
Tot.	7,000000	-	-	-

La première étape apparaît de la façon ci-contre. La figure 44 montre bien que l'Axe 1 permet à lui seul d'expliquer plus de 85% des variables. Combiné à l'Axe 2, la quasi-totalité des variables « Nb de clients par branche » (97.62%) sont expliquées.

Capture d'écran, résultat de l'analyse « Principal Component Analysis 1 » Tanagra 1.4.41

On réalise alors un graphique à partir de ces deux Axes, afin de déterminer à quelles variables ils correspondent. Pour cela, grâce au composant « Scatterplot 2 », les arrondissements sont placés en fonction des Axes 1 et 2, comme sur la figure 45.

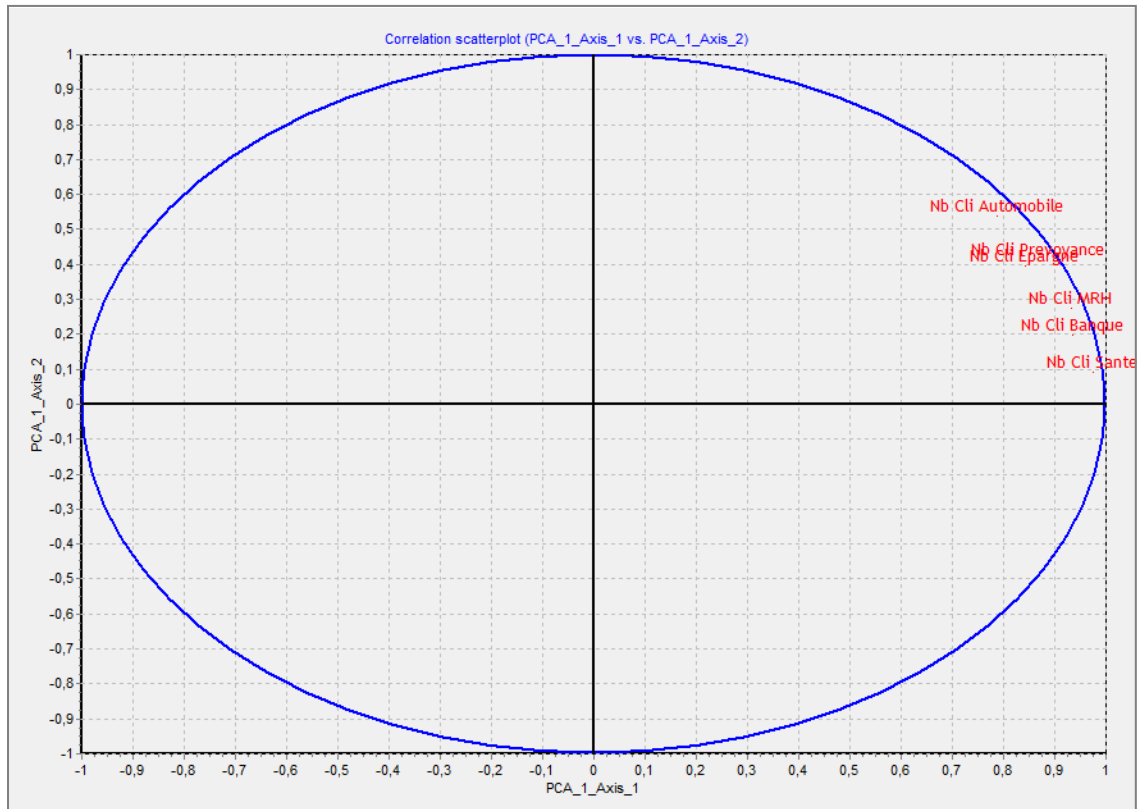
Figure 45 : Répartition des arrondissements de Paris selon les axes 1 et 2



Capture d'écran, résultat de l'opération « Scatterplot 2 » Tanagra 1.4.41

Le graphique obtenu place l'ensemble des arrondissements de Paris selon les deux variables Axe 1 et Axe 2. Le long de l'Axe 1 ce sont les arrondissements 15, 20, 19 et 18 qui ont les valeurs les plus élevées et les arrondissements 1 et 2 qui sont au plus bas. En triant la base de données afin de retrouver le classement évoqué ci-dessus, il apparaît que l'Axe 1 semble représenter la population de 40 à 49 ans ou les familles ayant un enfant. Pour l'Axe 2 ce sont les arrondissements 16, 7 et 8 qui se situent tout en haut de l'Axe. L'Axe 2 représente donc les revenus.

Figure 46 : Cercle de corrélation entre les Axes 1 et 2, et le nombre de clients par branche



Capture d'écran, résultat de l'opération « Correlation scatterplot 1 » Tanagra 1.4.41

Comme il est visualisable sur le cercle de corrélations, l'ensemble des branches de produits sont plus proches de l'Axe 1 que de l'Axe 2, de plus elles se situent toutes sur le cercle de corrélation. Cela signifie que la consommation de ces produits d'assurance est plus expliquée par le nombre de ménages de 40 – 49 ans ou de familles ayant un enfant (Axe 1), que par les revenus (Axe 2). On remarque que les branches Epargne et Prévoyance sont extrêmement proches l'une de l'autre (ce qui n'était pas le cas au niveau des départements), elles sont donc fortement corrélées. La branche Auto est isolée des autres, en effet elle n'est corrélée avec aucune autre branche et se trouve plus éloignée de l'Axe 1. Peu de parisiens ont des voitures et donc moins de clients disposent d'assurance automobile. Enfin les branches de produits MRH, Banque et Santé sont relativement rapprochées, elles sont probablement liées et ce sont les branches qui caractérisent le mieux l'Axe 1.

3.3.2.3. Création de clusters

Figure 47 : Répartition des arrondissements de Paris en trois clusters

Clusters	3		
Cluster	Description	Size	WSS
cluster n°1	c_kmeans_1	10	5,6714
cluster n°2	c_kmeans_2	7	2,0338
cluster n°3	c_kmeans_3	3	2,5226

Pour finir l'opération de Clustering regroupe les arrondissements ayant des habitudes de consommation en produits d'assurance proches. Trois groupes sont formés, le premier est composé de dix arrondissements, le second de sept et le troisième de trois.

Capture d'écran, résultat de l'opération « K-Means 1 » Tanagra 1.4.41

Figure 48 : Caractéristiques des trois clusters d'arrondissements de Paris

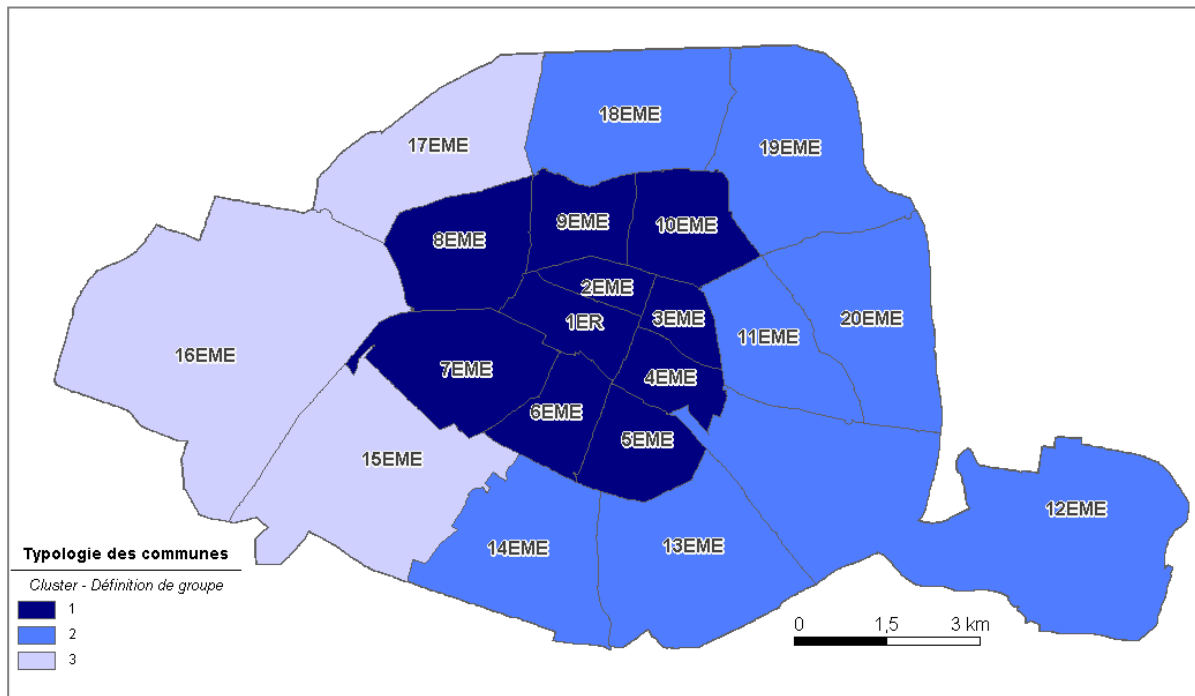
Description of "Cluster_KMeans_1"												
Cluster_KMeans_1=c_kmeans_1				Cluster_KMeans_1=c_kmeans_2				Cluster_KMeans_1=c_kmeans_3				
Examples [50,0 %] 10				Examples [35,0 %] 7				Examples [15,0 %] 3				
Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	Att - Desc	Test value	Group	Overall	
Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				Continuous attributes : Mean (StdDev)				
Nb Cli Automobile	-3,29	733,90 (360,45)	1464,80 (968,06)	Nb Cli Sante	2,48	533,71 (57,75)	373,80 (206,51)	Nb Cli Automobile	3,43	3279,00 (663,08)	1464,80 (968,06)	
Nb Cli Epargne	-3,50	375,60 (147,24)	743,45 (458,51)	Nb Cli Banque	2,16	436,86 (56,43)	307,20 (191,59)	Nb Cli Epargne	3,37	1586,33 (205,08)	743,45 (458,51)	
Nb Cli Prevoyance	-3,55	225,40 (79,74)	412,70 (229,68)	Nb Cli MRH	1,85	6552,43 (584,27)	4892,05 (2869,57)	Nb Cli Prevoyance	3,18	812,00 (104,29)	412,70 (229,68)	
Nb Cli MRH	-3,83	2369,50 (887,07)	4892,05 (2869,57)	Nb Cli Prevoyance	1,34	509,14 (69,57)	412,70 (229,68)	Nb Cli MRH	2,89	9426,33 (1165,01)	4892,05 (2869,57)	
Nb Cli Banque	-3,97	132,80 (48,58)	307,20 (191,59)	Nb Cli Epargne	1,15	907,71 (83,50)	743,45 (458,51)	Nb Cli Banque	2,66	586,00 (32,23)	307,20 (191,59)	
Nb Cli Sante	-3,99	185,00 (67,86)	373,80 (206,51)	Nb Cli Automobile	0,88	1731,43 (84,81)	1464,80 (968,06)	Nb Cli Sante	2,27	630,00 (90,67)	373,80 (206,51)	
Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				Discrete attributes : [Recall] Accuracy				

Capture d'écran, résultat de l'opération « Group Characterization » Tanagra 1.4.41

- Dans le premier cluster, c'est étonnamment l'assurance automobile qui domine, suivi de l'Epargne et de la Prévoyance. Il apparait donc qu'il s'agit d'arrondissements relativement riches, à savoir ceux du centre de Paris, du premier au dixième. Les branches de produits Banque et Santé viennent, elles, en dernière position. En revanche il est remarquable que sur l'ensemble de ces arrondissements, le nombre de clients par branche est nettement inférieur aux moyennes parisiennes. En effet ces arrondissements sont moins peuplés que ceux de la périphérie de Paris.
- Le second cluster regroupe les arrondissements de l'Est parisien, moins aisés, les branches de produits Auto et Epargne n'ont qu'une importance secondaire pour les clients de cette zone. Ce sont les branches Santé, Banque et MRH qui y ont un réel intérêt.
- C'est dans le troisième cluster que le nombre de clients est le plus élevé. Sur ces arrondissements les plus riches, 15^E, 16^E et 17^E, les branches de produits Auto, Epargne et

Prévoyance sont les plus développés. Au contraire la Santé et la Banque sont en dernière position.

Figure 49 : Représentation cartographique des groupes d'arrondissements de Paris (clusters), par type de consommation



Infographie Sophie STAËS (2014), données exportées Tanagra, réalisé sous Geoconcept V7

3.4. Conclusion

Grâce à ces différentes analyses il est mis en avant que les populations ont des profils et des consommations différentes d'un département à l'autre mais aussi d'une commune à l'autre. L'Ile-de-France est un territoire très diversifié et, notamment au niveau des communes de la région, avec des différences trop nombreuses et marquées pour qu'un seul modèle d'implantation convienne à l'ensemble d'entre elles. C'est pourquoi dans une quatrième partie, la recherche d'un modèle d'implantation optimal se limitera aux arrondissements de Paris.

De façon général, il a été observé que le nombre de clients AXA, quelles que soit les branches, est corrélé avec la population et plus particulièrement avec le nombre de ménage de 40 à 49 ans, les familles à un enfant, et, dans une moindre mesure, les revenus.

4. L'INTÉRÊT DU MODÈLE P-MÉDIAN DANS L'AMÉLIORATION DES ÉTUDES D'IMPLANTATION

Dans les études d'implantation d'AXA France, mais aussi de façon générale, la recherche d'une localisation commerciale est très souvent effectuée de façon intuitive, grâce à l'expérience, et non par des méthodes statistiques et rigoureuses. Cette dernière partie a donc pour but de comparer l'implantation d'une agence AXA de Paris intra-muros avec les résultats de la résolution du modèle P-médian.

4.1. Introduction

Il existe de nombreux modèles en géomarketing, et ceux-ci se regroupent en deux catégories : les modèles gravitaires (interaction spatiale) et les modèles de localisation-allocation.

Les modèles gravitaires sont très utiles dans la détermination de parts de marché ou de zones de chalandise, alors que les modèles de localisation-allocation sont plutôt utilisés dans la recherche d'une implantation optimale en fonction de la localisation des clients d'une enseigne. C'est donc sur l'un de ces derniers modèles que va se concentrer la recherche qui suit. On propose alors d'appliquer le problème p-médian afin de mettre en avant les avantages mais aussi les inconvénients de celui-ci pour le cas des agences AXA. Afin de faciliter le travail, et puisqu'il s'agit d'essais, l'application sera faite sur une seule agence AXA de Paris.

Les modèles de localisation-allocation ont pour objectif d'optimiser la distance entre l'offre et la demande. Par conséquent, c'est le nombre de points de vente ainsi que leurs implantations qui sont étudiés dans le but de minimiser les déplacements des clients. En effet, ces modèles ont pour hypothèse le fait que plus un point de vente est proche d'un consommateur plus il est probable que celui-ci le fréquente.

4.2. Le problème P-median

Le principe du problème p-médian est de calculer la somme des distances séparant l'implantation potentielle d'une agence et ses clients. Le point d'implantation potentiel présentant la somme des distances la plus faible sera alors le point le plus approprié à desservir la demande. Avec cette méthode il est possible de chercher la zone d'implantation d'une agence mais aussi d'un nombre infini.

Il est très fréquent que le nombre d'agences à implanter soit fixé au préalable, en effet une entreprise connaît son budget ainsi que le nombre de points de vente qu'elle peut ou a besoin d'ouvrir. Dans le cas qui suit, la recherche est concentrée sur l'implantation d'une seule agence. Puisque le « p » de « p-médian » correspond au nombre d'implantation à rechercher, il s'agit donc de la résolution du problème 1-médian.

La formulation mathématique, appliquée au cas présent est la suivante :

$$\text{Minimiser}$$

$$\sum_i \sum_j h_i d_{ij} x_{ij}$$

Où

i est le nœud de demande et j le nœud d'implantation potentiel de l'agence

h_i la demande au nœud i (=1 dans le cas non pondéré)

d_{ij} la distance entre la demande (clients) et le nœud j (implantation potentielle)

$x_{ij} = 1$ si la demande en i est satisfaite par j , 0 si non (dans le cas qui suit, ce paramètre est toujours égal à 1. En effet on cherche l'implantation possible d'une seule agence, or ce paramètre est utile dans le cas de l'ouverture de plus d'une. Le but étant de ne pas associer des clients de la première agence ouverte à la deuxième que l'on chercherait à implanter).

Dans tous les cas,

$$\sum_j x_{ij} = 1 \text{ (Tous les clients sont associés à une activité, et à une seule)}$$

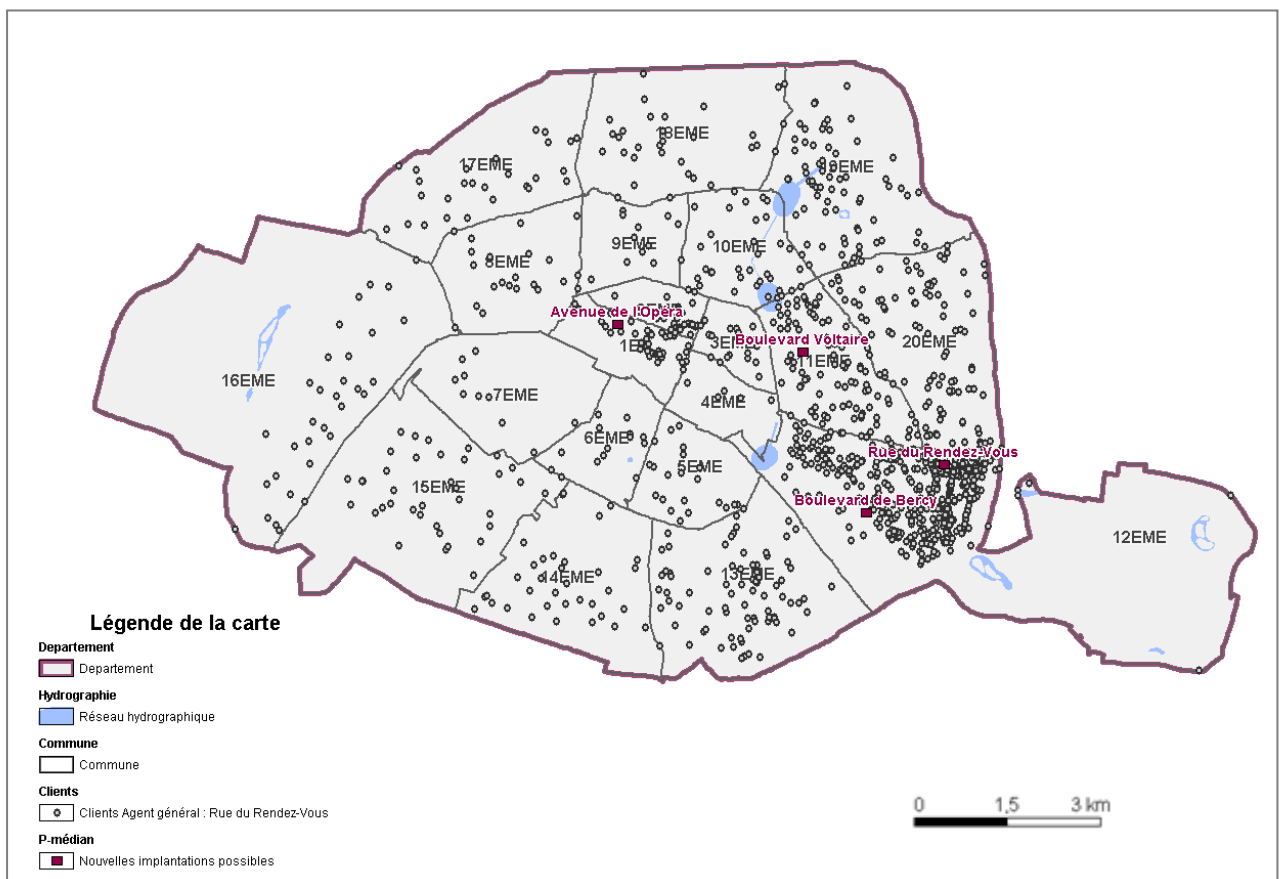
En résumé, pour tout i et pour tout j , on calcule la somme de : distances multipliées par demande. Le nœud j pour lequel la somme des distances sera la plus faible correspondra au meilleur site d'implantation.

L'agence AXA à partir de laquelle va être effectuée la résolution du problème est une agence du douzième arrondissement, située rue du Rendez-Vous. La question qui se pose est la suivante : Au garde de la localisation de ses clients, faut-il maintenir le point de vente sur son implantation actuelle ou bien déplacer celui-ci ?

4.2.1. Résolution du problème p-médian sans pondération

Ce modèle est réalisé en plusieurs étapes, la première étant la géolocalisation des clients de l'agence grâce au géocodage⁸⁰. La clientèle est représentée sur une carte à partir de la base de données d'adresses clients du point de vente, on observe ainsi, sur la figure 50, les zones où s'accumulent ceux-ci. C'est à partir de cette représentation que s'effectue la deuxième étape, la mise en avant des sites d'implantation potentiels.

Figure 50 : Représentation cartographique des clients d'une agence AXA du 12^E arrondissement et des sites d'implantation potentiels



Infographie Sophie STAËS (2014), données internes AXA France, réalisé sous Geoconcept V7

En effet, ce problème permet de choisir la meilleure implantation parmi des hypothèses déjà faites. Ici, quatre localisations potentielles ont été mises en avant et elles constitueront les implantations candidates :

- La rue du Rendez-vous, car c'est l'implantation actuelle.

⁸⁰ Attribution de coordonnées x et y à un point dans le but de le placer sur une carte

- L'avenue de l'Opéra, puisqu'aucune agence n'est implantée dans le premier arrondissement alors que celui-ci se situe au centre de Paris.
- Boulevard de Bercy, car peu d'agence AXA dans cette zone, de plus cela permettra d'éloigner l'agence d'une seconde actuellement implantée à moins d'un kilomètre de la rue du Rendez-vous.
- Le boulevard de Voltaire, car il s'agit d'une zone très passante et dynamique, peu éloignée de l'implantation actuelle.

Rappelons le, le principe est de calculer la somme des distances entre chaque clients et sites potentiels afin de déterminer lequel des sites dispose de la somme des distances la plus faible. Il s'agira alors du site le plus proche d'un maximum de clients et donc celui où il faut choisir de s'implanter. Étant donné que la demande est la même à chaque point client, à savoir 1 (puisque'un point ne représente qu'un client), le calcul n'a pas de nécessité d'être pondéré.

Tableau 7 : Distancier entre les clients et les quatres sites potentiels

Départ (numéro de client)	Arrivée (adresse du site)	Distance (mètres)
218146287	Rue du Rendez-Vous	4437
218146287	Boulevard Voltaire	2224
218146287	Boulevard de Bercy	3496
218146287	Avenue de l'Opéra	2233
1796805804	Rue du Rendez-Vous	6298
1796805804	Boulevard Voltaire	2916
1796805804	Boulevard de Bercy	5390
1796805804	Avenue de l'Opéra	1030
842807904	Rue du Rendez-Vous	5701
842807904	Boulevard Voltaire	2587
842807904	Boulevard de Bercy	4884
842807904	Avenue de l'Opéra	1272
⋮	⋮	⋮

Le calcul des distances est réalisé sous Geoconcept V7 grâce au module Distancier de l'onglet Itinéraire. Les points de départ choisis sont les clients de l'agence situés à Paris et les points d'arrivée sont donc les nouvelles implantations possibles. Les résultats obtenus sont exportés sous Excel, et on obtient alors un tableau dont les premières lignes correspondent au tableau 7.

Infographie Sophie STAËS (2014), données exportées de Geoconcept, réalisé sous Excel

Ce tableau va alors être retravaillé afin d'en obtenir un nouveau qui permettra de calculer pour chaque site, la somme des distances que parcourraient les clients pour s'y rendre. De plus les distances y sont converties en kilomètres et une partie du tableau a été masqué pour une meilleure lecture (en effet le tableau complet dispose en réalité de 1328 lignes).

Tableau 8 : Matrice de distance (km) entre les sites potentiels et les clients

Client	Site			
	Rue du Rendez-Vous	Boulevard Voltaire	Boulevard de Bercy	Avenue de l'Opéra
218146287	4,437	2,224	3,496	2,233
1796805804	6,298	2,916	5,39	1,03
842807904	5,701	2,587	4,884	1,272
1604998004	6,725	3,29	5,647	0,693
1074761204	7,088	3,945	5,831	0,688
3247646204	6,055	3,158	4,799	0,961
1744021004	6,002	3,28	4,745	0,907
1740509404	6,654	3,642	5,398	0,16
1722848504	5,789	2,354	5,04	1,391
⋮	⋮	⋮	⋮	⋮
1241982404	0,353	3,262	1,754	6,276
1789837604	0,727	3,635	1,988	6,65
1109390104	5,422	6,08	4,136	6,457
220608087	7,475	5,409	5,687	3,191
387137504	7,778	4,343	7,419	3,907
3211051004	6,429	4,139	6,905	5,058
1250392204	11,11	7,675	10,018	4,813
1241399204	1,597	4,146	3,241	7,16
2829607804	1,739	4,648	3,247	7,662
Somme des distances (km)	5015,821	4897,088	4791,023	6692,727

Infographie Sophie STAËS (2014), données exportées de Geoconcept, réalisé sous Excel

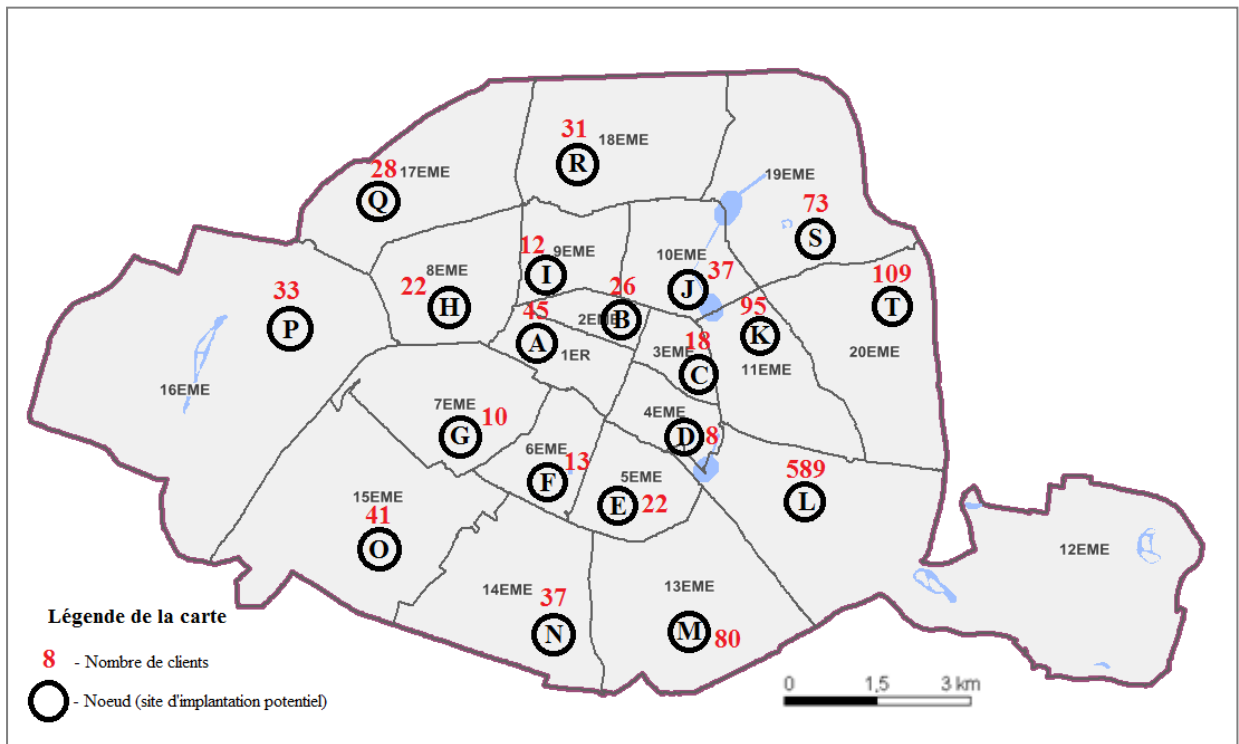
Sur le tableau 8, au bas de chaque colonne est calculée la somme des distances que parcourt chaque client. Il ressort qu'une implantation boulevard de Bercy, dans le douzième arrondissement est la plus judicieuse (il s'agit de la plus petite somme).

Cependant avec cette méthode la recherche d'une implantation optimale est biaisée par le fait que des sites potentiels sont déjà choisis. Pour compléter, confirmer ou non, cette proposition d'implantation, nous proposons d'effectuer le même travail mais en considérant que les sites d'implantation potentiels sont les vingt arrondissements de Paris.

4.2.2. Application du modèle p-médian pondéré

Pour cette nouvelle recherche, une notion de demande est ajoutée au calcul. En effet, nous avons choisi de créer un nœud par arrondissement (figure 51) au centre de celui-ci, représenté par une lettre. Les chiffres en rouge associés à ces lettres correspondent à la demande au niveau du nœud, à savoir ici, le nombre de clients sur l'arrondissement. On obtient alors une carte sur laquelle chaque arrondissement a un nombre de clients, chaque nœud présentera donc un poids différent. On applique toujours, comme précédemment le problème du 1-médian, cependant le calcul devra être pondéré.

Figure 51 : Représentation cartographique des nœuds du problème 1-médian



Infographie Sophie STAËS (2014), données internes AXA France, réalisé sous Geoconcept V7

Grâce à un calcul simple il est possible de connaître le nombre de solutions d'implantation à examiner, où n correspond au nombre de nœuds et p au nombre d'agences à placer. Ici nous disposons de vingt nœuds (les vingt arrondissements représentés par des lettres) et nous souhaitons implanter une agence. Le n est donc égal à vingt et p à un.

$$\frac{n!}{p!(n-p)!}$$

$$\text{Soit } \frac{20!}{1!(20-1)!} = 20$$

Dans le cas présent il y a donc vingt possibilités d'implantation différentes. Ce calcul est peu utile dans le cas présent, cependant il peut s'avérer très intéressant lorsque le nombre d'agences à planter est supérieur à un.

Sur la figure 51, les nœuds ne sont pas représentés au centre de chaque arrondissement afin que la carte soit plus lisible, cependant les distances sont calculées à partir du centre de chacun d'eux. De plus ces dernières n'ont pas été ajoutées à la carte mais sont précisées dans le tableau 9. Il est à noter qu'elles sont différentes d'un sens à l'autre car elles sont calculées par la route et prennent donc en compte les sens interdits.

De même que précédemment, on effectue sous Geoconcept un distancier entre l'ensemble des arrondissements, qu'on exporte ensuite vers Excel. A partir de ce premier tableau des distances entre chaque nœud, qu'est le tableau 9, il est possible d'effectuer une première somme des distances. On observe alors que c'est le deuxième arrondissement qui dispose de la somme la plus réduite.

Tableau 9 : Matrice de distance (km) non pondérée entre les arrondissements

Nœud	Site																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
A	0	1,088	2,708	2,15	3,045	2,425	3,347	2,678	2,441	2,991	3,776	6,449	5,268	4,516	5,076	6,217	3,674	4,268	5,151	5,31
B	1,36	0	2,743	3,074	3,969	3,201	3,721	2,441	1,911	2,793	4,067	7,234	6,192	5,058	5,578	6,44	3,418	3,588	4,954	5,331
C	2,437	2,282	0	1,578	3,232	3,318	5,157	4,699	4,17	2,696	2,215	5,374	5,053	4,99	6,601	8,09	5,677	4,617	4,644	3,844
D	2,138	2,731	1,963	0	2,493	2,665	4,503	4,588	4,352	3,481	2,461	5,036	4,212	4,337	5,948	7,437	5,584	5,402	5,575	4,181
E	3,194	3,787	3,441	1,837	0	1,601	4,258	5,36	5,408	4,537	3,88	5,856	3,002	2,745	4,884	7,918	6,423	6,458	6,698	5,6
F	2,611	3,204	3,708	2,548	1,583	0	2,761	4,144	4,824	4,585	4,59	6,771	3,737	2,29	3,387	6,477	5,207	6,125	6,745	6,31
G	3,626	4,301	5,449	4,409	3,751	2,334	0	3,314	4,69	6,207	6,372	8,696	5,574	3,232	1,965	4,349	4,435	6,163	8,164	8,083
H	3,432	3,388	5,088	5,355	6,089	4,743	2,985	0	3	4,91	6,184	9,57	8,286	6,131	4,921	4,241	1,964	4,089	6,82	7,448
I	2,385	1,552	3,244	4,06	4,955	4,14	4,335	2,361	0	2,387	4,337	7,723	7,178	5,997	6,271	6,395	2,911	2,501	4,257	5,601
J	2,943	2,232	1,891	3,447	4,457	4,28	5,59	3,836	2,717	0	2,776	6,162	6,59	5,952	7,526	7,932	4,702	2,485	2,435	4,04
K	4,115	3,404	1,924	2,308	3,65	4,32	6,464	5,822	5,149	2,802	0	3,504	4,923	5,874	7,603	9,672	6,799	4,916	4,142	1,958
L	7,186	7,008	5,528	5,121	6,377	7,048	9,191	9,426	8,805	6,503	3,853	0	4,477	7,253	10,01	12,48	10,4	8,572	6,434	3,372
M	5,524	6,117	5,257	3,69	2,98	3,806	6,299	7,647	7,737	6,739	4,674	4,437	0	3,135	5,723	9,122	8,71	8,66	8,674	6,505
N	4,671	5,264	5,412	4,067	2,876	2,35	4,348	6,053	6,884	6,351	5,587	7,006	3,447	0	3,589	6,988	7,116	8,185	8,512	7,418
O	5,207	5,851	6,524	5,483	4,704	3,264	2,458	4,642	6,239	7,364	7,446	9,469	6,07	3,274	0	3,594	5,763	7,628	9,524	9,158
P	6,172	6,782	8,622	8,094	7,918	6,524	4,585	4,613	6,705	8,56	9,72	12,39	9,366	6,756	3,795	0	4,649	7,558	10,38	11,1
Q	3,795	3,748	5,44	5,718	6,613	5,409	4,129	1,873	2,72	4,605	6,535	9,922	8,836	7,167	6,003	4,529	0	3,306	6,322	7,8
R	4,014	3,051	3,968	5,387	6,396	5,834	5,904	3,738	2,487	2,641	4,853	8,239	8,529	7,726	7,84	7,51	3,352	0	3,801	5,778
S	5,154	4,246	3,603	5,16	6,668	6,492	7,362	5,608	4,353	2,486	3,729	6,504	8,44	8,163	9,298	9,704	6,293	3,782	0	3,254
T	5,347	4,636	3,165	4,093	5,434	6,105	8,248	7,054	6,273	3,926	2,176	3,254	6,469	7,658	9,388	11,05	8,031	5,942	3,065	0
Somme des distances (m)	75,31	74,67	79,68	77,58	87,19	79,86	95,65	89,9	90,87	86,56	89,23	133,6	115,6	102,3	115,4	140,2	105,1	104,2	116,3	112,1

Infographie Sophie STAËS (2014), données exportées de Geoconcept, réalisé sous Excel

C'est donc le plus proche de l'ensemble des autres nœuds (arrondissements) mais cependant sans pondération. Cela signifie que la somme des distances suivante ne prend pas en compte le nombre de clients par nœud. Or sans ce paramètre le lieu d'implantation a peu d'intérêt, en effet il importe peu d'être situé le plus au centre possible de Paris. Il est donc absolument nécessaire de pondérer l'ensemble du tableau avec le poids (nombre de clients) de chaque nœud.

Les nœuds affichés en ligne correspondent à l'emplacement des clients. Afin de pondérer les distances parcourues pour atteindre chaque site, on multiplie donc chaque ligne par le nombre de clients présent sur le nœud correspondant. De cette façon le nombre de clients qui va se déplacer vers l'agence est pris en compte. C'est la résolution du problème 1-médian pondéré (distance multipliée par demande) et on obtient le tableau 10.

Tableau 10 : Matrice de coût (demande × distance)

Nœud	Site																				Pondération Nbr de clients par nœud
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
A	0	48,96	121,9	96,75	137	109,1	150,6	120,5	109,8	134,6	169,9	290,2	237,1	203,2	228,4	279,8	165,3	192,1	231,8	239	45
B	35,36	0	71,32	79,92	103,2	83,23	96,75	63,47	49,69	72,62	105,7	188,1	161	131,5	145	167,4	88,87	93,29	128,8	138,6	26
C	43,87	41,08	0	28,4	58,18	59,72	92,83	84,58	75,06	48,53	39,87	96,73	90,95	89,82	118,8	145,6	102,2	83,11	83,59	69,19	18
D	17,1	21,85	15,7	0	19,94	21,32	36,02	36,7	34,82	27,85	19,69	40,29	33,7	34,7	47,58	59,5	44,67	43,22	44,6	33,45	8
E	70,27	83,31	75,7	40,41	0	35,22	93,68	117,9	119	99,81	85,36	128,8	66,04	60,39	107,4	174,2	141,3	142,1	147,4	123,2	22
F	33,94	41,65	48,2	33,12	20,58	0	35,89	53,87	62,71	59,61	59,67	88,02	48,58	29,77	44,03	84,2	67,69	79,63	87,69	82,03	13
G	36,26	43,01	54,49	44,09	37,51	23,34	0	33,14	46,9	62,07	63,72	86,96	55,74	32,32	19,65	43,49	44,35	61,63	81,64	80,83	10
H	75,5	74,54	111,9	117,8	134	104,3	65,67	0	66	108	136	210,5	182,3	134,9	108,3	93,3	43,21	89,96	150	163,9	22
I	28,62	18,62	38,93	48,72	59,46	49,68	52,02	28,33	0	28,64	52,04	92,68	86,14	71,96	75,25	76,74	34,93	30,01	51,08	67,21	12
J	108,9	82,58	69,97	127,5	164,9	158,4	206,8	141,9	100,5	0	102,7	228	243,8	220,2	278,5	293,5	174	91,95	90,1	149,5	37
K	390,9	323,4	182,8	219,3	346,8	410,4	614,1	553,1	489,2	266,2	0	332,9	467,7	558	722,3	918,8	645,9	467	393,5	186	95
L	4233	4128	3256	3016	3756	4151	5413	5552	5186	3830	2269	0	2637	4272	5894	7353	6128	5049	3790	1986	589
M	441,9	489,4	420,6	295,2	238,4	304,5	503,9	611,8	619	539,1	373,9	355	0	250,8	457,8	729,8	696,8	692,8	693,9	520,4	80
N	172,8	194,8	200,2	150,5	106,4	86,95	160,9	224	254,7	235	206,7	259,2	127,5	0	132,8	258,6	263,3	302,8	314,9	274,5	37
O	213,5	239,9	267,5	224,8	192,9	133,8	100,8	190,3	255,8	301,9	305,3	388,2	248,9	134,2	0	147,4	236,3	312,7	390,5	375,5	41
P	203,7	223,8	284,5	267,1	261,3	215,3	151,3	152,2	221,3	282,5	320,8	409	309,1	222,9	125,2	0	153,4	249,4	342,4	366,2	33
Q	106,3	104,9	152,3	160,1	185,2	151,5	115,6	52,44	76,16	128,9	183	277,8	247,4	200,7	168,1	126,8	0	92,57	177	218,4	28
R	124,4	94,58	123	167	198,3	180,9	183	115,9	77,1	81,87	150,4	255,4	264,4	239,5	243	232,8	103,9	0	117,8	179,1	31
S	376,2	310	263	376,7	486,8	473,9	537,4	409,4	317,8	181,5	272,2	474,8	616,1	595,9	678,8	708,4	459,4	276,1	0	237,5	73
T	582,8	505,3	345	446,1	592,3	665,4	899	768,9	683,8	427,9	237,2	354,7	705,1	834,7	1023	1205	875,4	647,7	334,1	0	109
Somme des distances (m)	7295	7069	6103	5940	7099	7418	9510	9310	8845	6917	5154	4557	6828	8318	10618	13098	10469	8997	7651	5491	

Infographie Sophie STAËS (2014), données exportées de Geoconcept, réalisé sous Excel

Ce tableau 10 montre qu'en prenant en compte le nombre de clients, le résultat est différent et le nœud d'implantation le plus judicieux est le « L », soit le douzième arrondissement et non plus le deuxième. En effet la majorité des clients de l'agence se situe dans son arrondissement d'implantation actuel, soit le douzième.

4.3. Conclusion

Les recherches précédentes ont montré que l'agence AXA actuellement localisée rue du Rendez-Vous était très bien implantée au regard de l'emplacement de ses clients. En effet le douzième arrondissement est celui qui la garde le plus à proximité de sa clientèle. Cependant il serait probablement plus intéressant de se déplacer légèrement vers le sud, notamment vers le Boulevard de Bercy.

Ces propositions ne sont que les hypothèses d'une rapide recherche, mais elles permettent tout de même de rendre compte qu'un calcul simple est capable d'aider dans le choix d'une implantation.

Il est à noter que le modèle p-médian présente certains inconvénients. Il ne prend par exemple pas en compte le nombre de concurrents aux alentours, et est seulement axé sur la distance des clients. Il ne permet donc pas de choisir une implantation à lui seul mais plutôt de se faire une idée ou de compléter une étude déjà réalisée.

CONCLUSION

Si le nombre de points de vente des assureurs est élevé en France, il décroît depuis quelques années et la nécessité n'est plus d'être présent partout, n'importe comment, mais d'avoir les meilleures implantations avec moins de points de vente. En effet le nombre de concurrents sur le marché de l'assurance ne cesse d'augmenter. Il devient donc indispensable de redoubler d'efforts et de s'investir dans les études géomarketing afin d'être implanté au plus près des ses clients, acquis et potentiels.

Tout l'intérêt d'une étude géomarketing est de comprendre un territoire, sa population, ainsi que les clients dont l'entreprise dispose sur celui-ci. Le but étant ainsi de répondre au mieux à la demande qui s'y trouve.

Il a été montré que le logiciel Tanagra permettait de comprendre facilement la population d'Ile-de-France ainsi que ses habitudes de consommation en produits d'assurance. Celui-ci combiné à Geoconcept a permis d'expliquer les différences majeures entre les départements ainsi que les communes de la région. Il a donc été démontré que pour chaque zone, les comportements des consommateurs étaient différents et ce sont des paramètres à ne pas négliger dans une étude d'implantation. A la suite de cela, l'application du modèle p-médian sur Paris a permis de donner de précieuses indications dans l'étude de déménagement d'un point de vente AXA.

Cependant, si le modèle p-médian aide à choisir une localisation commerciale en fonction des clients d'un point de vente, celui-ci ne prend pas en compte la présence des concurrents qui sont, de plus, extrêmement nombreux sur Paris. Ce modèle a donc certaines limites puisqu'il peut tout à fait proposer d'implanter un point de vente dans une zone déserte, simplement si celle-ci se situe au centre de l'ensemble des clients de ce point de vente. Dans ce cas l'implantation serait mauvaise, puisque s'implanter dans une zone sans passage et sans concurrent ne serait sans doute pas rentable (pas de nouveaux clients). En effet, si les concurrents sont implantés et qu'ils y restent, cela signifie bien souvent qu'il s'agit d'une zone à fort potentiel.

Le choix d'un modèle d'implantation est difficile, et même si le modèle p-médian peut aider à conforter dans un choix d'implantation avant ou à la suite d'une étude géomarketing, il ne s'agit que d'un outil en plus et non une méthode à utiliser seule.

Il semble donc que pour l'heure, une bonne étude d'implantation ne puisse être menée que par la combinaison de plusieurs outils et modèles avec les connaissances et l'expérience du géomarketeur. Il existe cependant un grand nombre de modèles statistiques d'implantation différents, dont certains bien plus perfectionnés que celui du p-médian, qui permettrait peut être d'apporter plus de réponses.

BIBLIOGRAPHIE

WEBOGRAPHIE

BENATI Stefano et GARCIA Sergio (2012), *A p-median problem with distance selection*, Madrid, p. 2-15.

CLIQUET Gérard Cliquet et JOSSELIN Jean-Michel (2012), *Stratégies de localisation des entreprises commerciales et industrielles*, De Boeck Supérieur, p. 1-4.

GAY Jean-Christophe (2011), *Résolution du Problème du p-médian, application à la restructuration de bases de données semi-structurées*, p. 17-18.

MARTIN Pierre, *Petite histoire des agences d'assurance : au service du risque*

SCHINDLER Patrick (2013), *Analyse du secteur de l'assurance*, 3 p.

<http://www.journaldunet.com>

<http://www.infostat-marketing.com>

<http://eurogroupeconsulting.fr>

<http://www.marquesetreseaux.com>

<http://www.lalettredelassurance.com>

<http://www.argusdelassurance.com>

<http://www.alternatives-economiques.fr>

<http://www.senat.fr>

<http://clubassurance.blogspirit.com>

<http://www.empruntis.com>

<http://www.leseco.ma>

<http://www.lesechos.fr>

<http://www.latribune.fr>

<http://www.assemblee-nationale.fr>

<http://www.insee.fr>

<http://www.iledefrance.fr>

TABLE DES ILLUSTRATIONS, TABLEAUX ET FIGURES

Illustrations

N°	Titre	
1	Départements de la région Ile-de-France	28

Tableaux

N°	Titre	
1	Classement par nombre d'agents et de points de vente en France	11
2	Répartition de la population d'Ile-de-France par tranche d'âge	33
3	Répartition de la population d'Ile-de-France par tranche d'âge en pourcentage	34
4	Population active ou inactive en Ile-de-France	35
5	Population active ou inactive en Ile-de-France	35
6	Taux de chômage en 2013	36
7	Distancier entre les clients et les quatres sites potentiels	68
8	Matrice de distance (km) entre les sites potentiels et les clients	69
9	Matrice de distance (km) non pondérée entre les arrondissements	71
10	Matrice de coût (demande x distance)	72

Figures

N°	Titre	
1	Nombre de points de vente d'assurance en France	9
2	Raisons pour lesquelles un client effectue ses achats dans une boutique	10
3	Evolution des parts de marché dans l'assurance dommages	14
4	Evolution des parts de marché dans l'assurance vie-capitalisation	15
5	Répartition des modes de distribution en France	22
6	Exemple de données géographiques sur le 1 ^{ER} arrondissement de Paris	24
7	Représentation vecteur	25
8	Représentation raster	25
9	Répartition de la population par département en millions d'habitants en 2013	29
10	Département des Hauts-de-Seine	30
11	Département Seine-Saint-Denis	30
12	Département du Val-de-Marne	31
13	Département Seine-et-Marne	31
14	Département de l'Essonne	32
15	Département des Yvelines	32

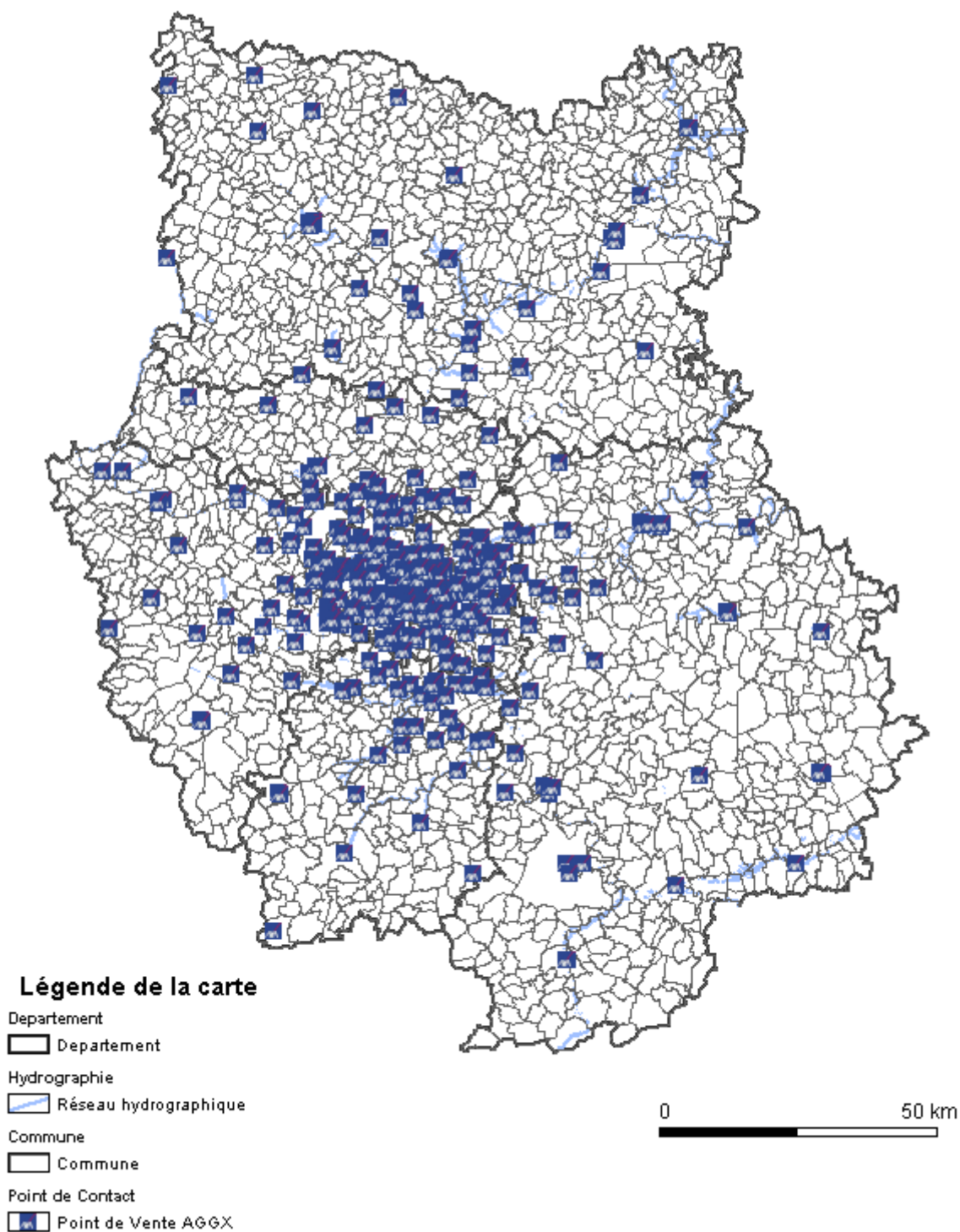
16	Département du Val-d'Oise	33
17	Nombre d'agences générales AXA par département en Ile-de-France en 2014	37
18	Nombre d'agences générales AXA par commune en Ile-de-France en 2014	37
19	Nombre de concurrents (Banques et Assurances) par commune en 2014	38
20	Présence d'AXA ou de la concurrence sur les communes d'Ile-de-France en 2014	39
21	Présentation du logiciel Tanagra	41
22	Analyse de la variable continue « Indice de revenu moyen »	42
23	Corrélation linéaire entre les variables continues du fichier Département	43
24	Description des départements par le composant « Group Characterization »	44
25	Répartition des départements dans trois clusters	45
26	Clusters auxquels appartiennent les départements d'Ile-de-France	46
27	Caractéristiques des départements du cluster 1	46
28	Caractéristiques des départements du cluster 2	47
29	Caractéristiques des départements du cluster 3	47
30	Représentation cartographique des groupes de départements (clusters)	48
31	Répartition des communes dans huit clusters	49
32	Caractéristiques des huit clusters de communes d'Ile-de-France	50
33	Représentation cartographique des groupes de communes (clusters)	51
34	Description des départements par le composant « Group Characterization »	53
35	Description des départements par le composant « Group Characterization » (suite)	54
36	Axes proposés par l'analyse en composantes principales	55
37	Répartition des départements d'Ile-de-France selon les axes 1 et 2	55
38	Cercle de corrélation entre l'Axe 1 et le nombre de clients par branche	56
39	Répartition des départements dans trois clusters	57
40	Caractéristiques des trois clusters de départements d'Ile-de-France	57
41	Représentation cartographique des groupes de départements (clusters)	58
42	Corrélation linéaire entre le nombre de clients AXA et les variables population	59
43	Description du 1 ^{ER} arrondissement par le composant « Group Characterization »	59
44	Axes proposés par l'analyse en composantes principales	60
45	Répartition des arrondissements de Paris selon les axes 1 et 2	61
46	Cercle de corrélation entre les Axes 1 et 2, et le nombre de clients par branche	62
47	Répartition des arrondissements de Paris en trois clusters	63
48	Caractéristiques des trois clusters d'arrondissements de Paris	63
49	Représentation cartographique des groupes d'arrondissements de Paris (clusters), par type de consommation	64
50	Représentation cartographique des clients d'une agence AXA du 12 ^E arrondissement et des sites d'implantation potentiels	67
51	Représentation cartographique des nœuds du problème 1-médian	70

TABLE DES ANNEXES

N°	Titre	
1	Localisation des points de vente AXA du réseau agents généraux en Ile-de-France	80
2	Données utilisées dans l'analyse du profil socio-économique par département	81
3	Données utilisées dans l'analyse de la consommation de produits d'assurance par département	81

ANNEXES

ANNEXE 1 : Localisation des points de vente AXA du réseau agents généraux en Ile-de-France



ANNEXE 2 : Données utilisées dans l'analyse du profil socio-économique par département

Departement	ESSONNE	VAL D'OISE	OISE	YVELINES	SEINE-ET-MARNE	HAUTS-DE-SEINE	SEINE-ST-DENIS	VAL-DE-MARNE	PARIS
Code Departement	91	95	60	78	77	92	93	94	75
Indice Rev moy men	125,00	117,00	107,00	153,00	119,00	149,00	91,00	118,00	155,00
Nb menages	420603	394690	280256	503096	432351	624926	524387	499404	1110912
Population	1134026	1105224	766313	1353957	1193511	1428678	1382928	1226961	2125851
Nb Cli PartPro en cours	61953	60349	56696	77068	71330	90127	70099	69988	144490
Nb Cli Entreprise en cours	3170	2340	2287	3346	3526	4348	2873	3050	10666
Nb Cli PartProEnt en cours	65123	62689	58983	80414	74856	94475	72972	73038	155156
Pop 0-9 ans	148094	150033	99106	178231	159042	176780	189622	153124	195270
Pop 10-14 ans	81517	87049	59358	98763	93940	81043	99791	77199	92366
Pop 15-17 ans	46759	50079	34494	58238	54491	46609	57955	43597	55531
Pop 18-19 ans	31547	34085	22554	39405	36108	32621	39542	30355	45207
Pop 20-24 ans	73807	76047	44928	84730	75333	89546	94233	77581	163022
Pop 25-29 ans	85643	80350	55383	97938	86063	127265	109543	100545	234054
Pop 30-39 ans	181643	171687	118496	208460	188427	247109	223763	204781	368002
Pop 40-49 ans	169378	171381	118155	204374	188095	200275	205650	179726	292175
Pop 50-59 ans	139803	127840	86685	171198	138106	167326	154573	144289	263133
Pop 60-74 ans	121427	106718	84827	145014	114813	164540	143130	140068	248547
Pop 75 ans et plus	54408	49955	42327	67606	59093	95564	65126	75696	168544
Pop Agriculteurs	1367	899	4560	1357	3892	132	127	202	336
Pop Artis commerc	25672	26147	18282	32429	31325	36622	32038	31998	68814
Pop Employes	171476	174908	101888	185875	185843	207541	236066	198522	275495
Pop Ouvriers	107773	116717	115055	107590	139860	96833	184189	114974	116516
Pop PIS	96593	73072	35374	151794	71548	199671	62860	107487	393883
Pop Prof. interm.	153072	139698	79686	172496	148517	178380	143810	155571	259432
Pop Retraites	155236	136826	111553	180578	154744	217539	181070	187550	330151
Pop Chomeurs	36108	51204	34406	44448	42545	56124	79326	48041	98838
Pop femmes	575354	563904	385916	689277	605182	746086	705771	635255	1128929
Pop hommes	558672	541320	380397	664680	588329	682592	677157	591706	996922
Fam 0 enfant	126985	110723	89457	149785	126722	163960	135076	138336	239665
Fam 1 enfant	76752	73431	49656	87611	80887	95733	93383	85750	122178
Fam 2 enfants	75394	72006	45624	87217	79570	77657	77554	71904	81125
Fam 3+ enfants	35755	42266	28326	47353	43540	37456	52729	34033	41211

ANNEXE 3 : Données utilisées dans l'analyse de la consommation de produits d'assurance par département

Departement	ESSONNE	VAL D'OISE	OISE	YVELINES	SEINE-ET-MARNE	HAUTS-DE-SEINE	SEINE-ST-DENIS	VAL-DE-MARNE	PARIS
Code Departement	91	95	60	78	77	92	93	94	75
Nb Cli PartPro en cours	61953	60349	56696	77068	71330	90127	70099	69988	144490
Nb Cli Automobile	24777	23834	27970	31939	31619	27962	20396	22467	29296
Nb Cli MRH	37417	36758	33403	45712	40269	58640	46152	44861	97841
Nb Cli Sante	3492	3519	4009	4648	5453	4414	3975	3836	7476
Nb Cli Prevoyance	4781	3596	5177	6042	5573	5704	6281	4935	8254
Nb Cli Epargne	7241	5668	7607	9849	8379	10489	5447	7457	14869
Nb Cli Banque	4312	4294	4987	4880	6490	5215	4416	4533	6144
Nb Cli Entreprise en cours	3170	2340	2287	3346	3526	4348	2873	3050	10666
Nb Cli PartProEnt en cours	65123	62689	58983	80414	74856	94475	72972	73038	155156

TABLE DES MATIÈRES

Remerciements	2
Clause déontologique.....	3
Résumé / Abstract.....	4
Sommaire.....	5
Introduction.....	6-7
1. La situation du marché de l'assurance : AXA face à ses concurrents.....	8-22
1.1 Introduction.....	8
1.2 Le réseau de points de vente comme vitrine historique des sociétés d'assurance...9-12	
1.2.1 L'avantage de la présence physique	
1.2.2 La visibilité	
1.3 La pression concurrentielle.....	12-21
1.3.1 Les mutuelles sans intermédiaires	
1.3.2 Le développement de la bancassurance	
1.3.3 La concurrence digitale	
1.3.3.1 Les comparateurs d'assurance	
1.3.3.2 Les assureurs de plus en plus visibles sur internet	
1.3.3.3 La vente directe	
1.3.4 La grande distribution	
1.4 Conclusion.....	21-22
2. La prise en compte des réalités géographiques : contexte d'une étude territoriale en région.....	23-39
2.1 Introduction.....	23
2.2 Les données disponibles.....	23-27
2.2.1 Les données géographiques	
2.2.2 Les données socio-économiques	
2.2.3 Les données internes	
2.3 L'Ile de France.....	27-36
2.3.1 Paris	
2.3.2 La petite couronne	
2.3.3 La grande couronne	
2.3.4 La comparaison des départements	
2.3.4.1 Les tranches d'âges	
2.3.4.2 L'emploi	
2.4 Le maillage du territoire francilien.....	36-39
2.4.1 Le réseau d'agents généraux AXA	
2.4.2 La localisation de la concurrence	
2.5 Conclusion.....	39
3. L'optimisation du territoire grâce aux études géomarketing et à l'analyse du marché.....	40-64
3.1 Introduction.....	40-41
3.2 Le profil socio-économique de la région.....	41-52

3.2.1	Les départements d'Ile-de-France	
3.2.1.1	Analyse descriptive	
3.2.1.2	Création de clusters	
3.2.2	Le regroupement des communes	
3.3	La consommation de produits d'assurance	52-64
3.3.1	Une consommation différente d'un département à l'autre	
3.3.1.1	Analyse descriptive	
3.3.1.2	Analyse en composantes principales	
3.3.1.3	Création de clusters	
3.3.2	Une consommation particulière à Paris	
3.3.2.1	Analyse descriptive	
3.3.2.2	Analyse en composantes principales	
3.3.2.3	Création de clusters	
3.4	Conclusion	64
<hr/>		
4.	L'intérêt du modèle p-médian dans l'amélioration des études d'implantation	65-73
<hr/>		
4.1	Introduction	65
4.2	Le problème p-médian	65-72
4.2.1	Résolution du problème p-médian sans pondération	
4.2.2	Application du modèle p-médian pondéré	
4.3	Conclusion	73
	Conclusion	74-75
	Bibliographie / Webographie	76
	Table des illustrations, tableaux et figures	77-78
	Table des annexes	79
	Annexes	80-81
	Table des matières	82-83