

Rapport final

L'implantation des banques coopératives en France
Les banques coopératives en France sont-elles implantées en fonctions de certaines caractéristiques de la population ?



Commanditaire

ARTIS Amélie

Tutrices

FAÏN Edwige

LETUÉ Frédérique



Table des matières

Remerciements	4
Résumés	5
Français	5
Anglais	5
Méthodologie	7
Les données géographiques des banques	7
Les données socio-économiques	10
Les analyses	11
Glossaire	14
Introduction	16
I. La diversité socio-économique du territoire français	17
Le salaire net horaire moyen	18
Les secteurs économiques	18
Le taux de chômage	20
L'âge de la population	21
II. La diversité des banques françaises	23
Présentation des banques	23
À l'échelle de la France Métropolitaine (hors Corse)	24
À l'échelle régionale	25
À l'échelle départementale	29
À l'échelle des zones d'emploi	31
III. Les différences entre les banques coopératives et commerciales	39
Statistiques descriptives	39
Analyse en Composante Principale	42
Régressions linéaires multiples	46
Conclusion	52
Annexes	53
Annexe 1 : Codes scraping	53
Annexe 2 : Visualisation des données	61
Annexe 3 : graphiques des résidus des régressions linéaire	62
Graphiques des résidus studentisés avant sélection de variable	62
Graphiques des résidus studentisés avant sélection de variable	62
Annexe 4 : Cahier des charges	63

Annexe 5 : Journal de bord.....76

Légende	
	Réunions avec les tuteurs
	Réunions avec les membres du projet
	Réunion avec le commanditaire

Annexe 6 : Compte-rendu des réunions.....83

Annexe 6: Schéma de nettoyage des données.....92

Exemple de flux Tableau Prep pour le nettoyage des variables sur les âges92

Remerciements

Ce rapport fait état du projet tutoré de deuxième année en formation du DUT Statistique et Informatique décisionnelle de l'académie de Grenoble.

Nous tenons à remercier notre commanditaire Amélie ARTIS, l'Institut de Science Politique de Grenoble et le département STID de Grenoble pour nous avoir permis de travailler sur ce projet.

Nous nous devons également de remercier Mme. LETUÉ et Mme. FAIN pour nous avoir donné la bonne marche à suivre. Nous les remercions aussi chaleureusement pour leur patience et leur pédagogie.

Pour finir, nous remercions l'ensemble de l'équipe éducative du département STID de Grenoble pour tous les cours sans lesquels la réalisation du projet aurait été impossible.

Résumés

Français

Dans le cadre de notre seconde année de DUT STID, nous réalisons un projet d'envergure par groupe afin de mettre en application les savoirs et savoir-faire acquis depuis le début de notre formation et de se familiariser avec la notion d'interdisciplinarité. Notre commanditaire, Amélie ARTIS, maître de conférence et enseignante-rechercheuse à Science Politique Grenoble, nous a sollicité pour que nous prenions part à ses recherches sur l'économie sociale et solidaire (ESS) en lui fournissant une étude sur l'implantation des banques coopératives en France métropolitaine. Nous allons donc essayer de mettre en évidence certaines caractéristiques de la population qui sembleraient avoir un impact sur l'implantation des banques coopératives en France.

Pour réaliser cela, nous avons suivi une méthodologie précise. Dans un premier temps, nous avons sélectionné des variables socio-économiques sur le site de l'Insee pour avoir diverses informations sur les différentes zones d'emploi en France métropolitaine. Ensuite, nous avons recueilli les adresses de quatre banques (coopératives et commerciales) grâce à une technique de fouille de données nommée scraping. Une fois l'ensemble de ces données récoltées et nettoyées, nous les avons analysées consciencieusement afin de pouvoir apporter des réponses à Amélie ARTIS. Après avoir effectué quelques statistiques descriptives sur nos données pour caractériser les zones d'emploi françaises, nous avons opéré une ACP ainsi qu'une régression linéaire afin de déterminer les variables ayant le plus d'impact sur l'implantation des banques en France.

Nos analyses nous ont permis d'apporter des éléments de réponse à notre problématique. Avec nos statistiques descriptives, nous avons remarqué des disparités dans l'implantation des différents types de banques. Ainsi les banques coopératives semblent implantées en plus grande quantité dans les zones d'emploi ayant une part d'agriculteurs importante et un salaire moins élevé. Cependant, l'étude des statistiques prédictives a révélé que les différences n'étaient pas aussi notables que ce que nous pensions. Nous pouvons donc en déduire que les banques coopératives restent encore relativement proches des ménages même si la diversité des différentes organisations n'est plus vraiment très marquée.

Anglais

In the context of our second year in a two-year technical degree in statistics and BI, we carry out a group project in order to apply the knowledge and the skills learnt since the beginning of our formation and to become more familiar with the notion of interdisciplinarity. Our sponsor, Amélie ARTIS, a university researcher and lecturer at Science Politique Grenoble, wants us to take part in her research about the social and solidarity economy by producing for her a study about the establishment of cooperative banks in metropolitan France. We are going to try to highlight some features of the population that seem to have an impact on the establishment of cooperative banks in France. In order to do so, we followed an accurate methodology. Firstly, we selected socio-economic variables on the Insee web site to have various information about the several "employment areas" in France. Then we collected addresses of four banks (cooperatives or commercials) thanks to a data

mining method called scraping. Once all of this data has been collected and cleaned, we analyzed it conscientiously to propose answers and results to Amélie ARTIS. After making some descriptive statistics on our data to characterize the french “employment areas”, we proceeded to a principal component analysis (PCA) and to a linear regression to determine the variables which have the biggest impact on the establishment of the cooperative banks in France.

Our analysis allowed us to answer to our problem. With the descriptive statistics we pointed some disparities on the establishment of the several banks types. In this way cooperatives banks seem to be established mostly in the employment areas with an important part of farmers and with a lower average income. However, the predictive statistics revealed that the differences were less important than we expected. We deduct that the cooperative banks are yet quite close to households even if the variety of the several organizations is not really pronounced.

Méthodologie

Les données géographiques des banques

Scraping

Pour récupérer les données géographiques des banques, nous sommes partis de plusieurs codes de scraping, implémentés sous le logiciel R, qui nous ont été fournis par des étudiants de l'année dernière qui travaillaient sur un projet similaire. Ce code permet littéralement de "gratter" les pages web pour en extraire l'information cherchée. Ici, nous avons utilisé cette méthode pour relever les adresses géographiques au sein des pages web des sites de ces banques. Ces programmes étant très peu commentés voire même obsolètes pour certains, nous avons dû les améliorer et les actualiser pour qu'ils nous permettent de collecter les bonnes adresses. Les quatre sites web associés aux quatre banques étudiées ayant des fonctionnements différents, nous avons donc quatre codes programmés en conséquence dont le fonctionnement va être détaillé dans cette partie.

Le Crédit Agricole

Pour obtenir les adresses des banques du Crédit Agricole en France grâce à R il nous a d'abord fallu télécharger et charger les packages R « rvest » et « stringr » qui nous seront utiles dans la suite du code.

Nous avons tout d'abord testé le code que nous allons utiliser sur le premier lien de la banque du Crédit Agricole pour voir si nous arrivions bien à extraire l'adresse présente sur le site. Une fois que cela a bien fonctionné pour une seule adresse, nous avons pu commencer à refaire cette démarche sur tous les liens grâce à une boucle. Nous avons donc commencé par créer une matrice vide de plusieurs colonnes et avec des milliers de lignes afin d'y insérer par la suite toutes les adresses obtenues grâce au scraping. Nous avons ensuite démarré notre boucle itérée à 2 car le premier lien qui contient l'adresse d'une banque est le lien <https://www.credit-agricole.fr/particulier/agence/sud-rhonealpes/vizille-0002.html>. Nous avons décidé de faire aller cette boucle jusqu'à 9999 pour être bien sûr de faire toutes les banques du Crédit Agricole (il y en a environ 7000 en France).

Pour pouvoir obtenir dans notre boucle tous les bons liens nous avons dû diviser le lien en 3 avec une première partie commune à tous les liens <https://www.credit-agricole.fr/particulier/agence/sud-rhonealpes/vizille-000>, une deuxième partie qui change avec la boucle (i) et une dernière partie également commune à tous les liens (.html). Ainsi à chaque avancement de la boucle nous avons obtenu un nouveau lien pour obtenir une nouvelle adresse de la banque Crédit Agricole. À partir de là nous avons pu obtenir l'adresse de la banque et la placer dans notre matrice vide de départ grâce au code disponible en annexes.

Enfin après avoir fait cela il ne restait plus qu'à remplacer toutes les virgules de la table par des espaces afin d'obtenir des données plus propres en sortie avant de réaliser cette sortie vers un emplacement choisi au préalable.

La Société Générale

Pour cette banque, il a été nécessaire d'installer les mêmes packages que pour le code précédemment présenté.

Le lien URL de la société générale a pour avantage d'être simple car sous le format <https://agences.societegenerale.fr/banque-assurance/agences-nomdedepartement-numeroedepartement>. En effet, une page de ce format nous renvoie toutes les adresses d'agences Société Générale présentes au sein d'un département. Il nous a donc suffi de créer deux listes à parcourir grâce à une boucle : une liste contenant les noms des départements et une liste contenant les numéros des départements. Une fois cette étape faite, nous avons créé une matrice vide « tablesortie » pour stocker à chaque fin de boucles les adresses pour chaque département. Nous avons initialisé notre boucle à 1 avant de la lancer. Cette boucle "tant que" a pour condition de ne pas dépasser la taille de la liste contenant les noms ou les numéros de département. Une fois encore nous avons décomposé notre lien en plusieurs parties. Ici nous l'avons décomposé en quatre parties. En première position une partie commune <https://agences.societegenerale.fr/banque-assurance/agences->, en deuxième position l'indice i parcourant la liste des noms de département, en troisième position le signe "-" et enfin l'indice i une nouvelle fois mais parcourant la liste des numéros de département. Nous avons ensuite concaténé tous ces liens pour en former un seul qui constitue notre page html pour le passage dans notre boucle. Nous avons ensuite essayé de lire cette page html. Après cela, nous avons stocké dans une matrice « table » les adresses du département en question en lisant cette page et en récupérant celles-ci dans le nœud html correspondant (que l'on trouve en regardant la structure et les balises du code html directement sur le code source de la page du site au préalable), ici « .agencyaddress ». Enfin nous avons concaténé notre table contenant les adresses pour le département présent au reste des adresses dans la matrice « tablesortie ». Pour terminer nous avons fait avancer notre indice de i à i+1 pour parcourir le département suivant lors du prochain passage dans la boucle. Nous avons par la suite converti notre matrice en fichier au format csv avant de la transférer sur vers emplacement choisi au préalable.

La Banque Populaire

Comme mentionné précédemment, il nous a fallu charger les packages R « rvest » et « stringr » pour faire fonctionner ce code. Dans un premier temps, il a été nécessaire que nous regardions attentivement la structure des liens des pages internet dans lesquelles nous voulions extraire des informations.

Pour la Banque Populaire, nous avons remarqué que les adresses des agences étaient regroupées par départements. En effet, il y a une page Internet par département dans laquelle sont regroupées toutes les agences bancaires ainsi que leurs adresses. Par exemple, pour obtenir des informations sur toutes les agences de la Banque Populaire en Isère, il faut se rendre sur ce lien :

<https://agences.banquepopulaire.fr/banque-assurance/agences-isere-38>. Notre but étant de collecter les adresses des agences de tous les départements français, nous avons élaboré une liste des départements et de leurs numéros pour pouvoir recréer les liens de chaque page Internet. Ainsi, avec une boucle "tant que", nous avons pu automatiser ce processus en initialisant une variable compteur (nommée i) qui a parcouru notre liste de départements. En effet, en gardant la partie des liens qui ne change pas et en faisant varier seulement la fin de ceux-ci, nous arrivons à retrouver tous les liens qui nous sont nécessaires. Ensuite, à l'aide de commandes spéciales, nous avons pu lire les pages html associées aux liens et extraire les adresses des agences bancaires contenues dans la balise html "div .em-poi-card_address". Chaque adresse collectée a été stockée dans une matrice "Tablesorties"

que nous avions créée au début du code et grâce à la commande “write.table”, elle a été exportée dans un fichier .csv à la fin de la boucle.

Le Crédit Mutuel

Afin de collecter les adresses des banques du Crédit Mutuel en France, un code sur le logiciel R nous a été fourni par les étudiants de l'année précédente. Nous avons donc d'abord analysé ce code et commenté ce que chaque ligne effectuait. Puis après avoir compris le code, nous l'avons lancé et nous nous sommes retrouvés avec les adresses mais sur une échelle plus petite que celle que nous voulions.

Dans un premier temps, nous avons dû installer et charger les packages R “rvest” et “stringr”. “rvest” nous a permis de parcourir et d'aller chercher le contenu d'une page web, pour le rendre exploitable par R et “stringr” nous a permis de manipuler les chaînes de caractères. Ensuite nous avons mis en place deux boucles.

La première nous a permis de parcourir les pages de chaque département du site web du Crédit Mutuel et de récupérer et stocker les noms des communes de chaque département dans une liste. Nous avons donc une boucle “tant que” avec un compteur j qui va de 1 à 95. Ce compteur j correspond au numéro de département et il nous a permis de parcourir les pages du site web et de modifier le début du lien selon le numéro de département j.

Pour obtenir le lien de la page nous avons donc concaténer le début du lien avec j à l'aide de fonctions provenant du paquet “stringr”. Nous avons ensuite utilisé des fonctions appartenant au paquet “rvest” afin de lire le code HTML de la page et récupérer les noms des communes qui sont présentes dans la balise afin de les stocker dans une liste. Enfin nous avons remplacé tous les espaces de la liste par des tirets.

Après cela nous avons procédé à la seconde boucle pour parcourir les pages de chaque ville où il y a une agence du Crédit Mutuel afin d'en extraire l'adresse et de la stocker dans une liste. La seconde boucle “tant que” a le même fonctionnement que la première, cependant le compteur i a parcouru chaque élément de la liste des villes en prenant le nom des villes comme valeur. Ainsi, nous avons obtenu une liste avec toutes les adresses des agences du Crédit Mutuel en France.

Suite à notre soutenance, nous nous sommes rendu compte qu'il existe un site web différent pour le Crédit Mutuel Bretagne. Par conséquent nous n'avons pu collecter les adresses des agences du Crédit Mutuel en Bretagne par manque de temps car les structures des deux sites web étant trop différentes nous n'avons pu adapter notre code.

Nettoyage et traitement

Pour nettoyer les données des banques nous avons dû séparer les adresses obtenues grâce au scraping en 3 variables différentes (voie, code postal et commune) grâce à une macro VBA.

Après avoir fait ça pour les quatre banques, nous avons supprimé les accents pour éviter tous les problèmes d'encodage que cela pouvait engendrer avant de supprimer les doublons sur Excel.

Une fois les doublons de chaque banque supprimée nous avons pu réunir toutes ces données sur un seul fichier Excel au format csv pour pouvoir l'importer dans une table sur l'entrepôt de données sur phpMyAdmin. C'est sur ce site que nous avons pu ajouter une

variable département créée grâce à la variable Code Postal (en prenant les deux premiers chiffres).

Comme nous voulons faire notre étude statistique en étudiant les zones d'emploi, nous avons dû ajouter cette donnée dans notre base de données des banques. Pour ce faire nous avons téléchargé sur le site de l'Insee un fichier contenant les communes et départements de chaque zone d'emploi afin de l'importer sur phpMyAdmin. Une fois cette nouvelle table importée nous avons pu obtenir une sortie contenant toutes les données des banques en plus de leur zone d'emploi grâce à une jointure contenant la commune et le département. Pour finir l'implantation de la base de données sur phpMyAdmin il nous a fallu rajouter une variable type de la banque pour pouvoir séparer la banque commerciale (Société Générale) des trois autres banques coopératives (Crédit Agricole, Crédit Mutuel et Banque Populaire) qui représente le centre de notre étude.

Les données socio-économiques

Le nettoyage et le traitement

Notre étude portant seulement sur la France métropolitaine, nous avons exclu les DOM-TOM des fichiers de données récupérés sur le site de l'Insee. De plus, nous avons préféré ne pas prendre en compte la Corse pour des raisons pratiques. En effet, certains fichiers de données ne nous apportent pas d'informations sur celle-ci.

Nous avons voulu travailler avec des données ayant comme échelle géographique les zones d'emploi. Or nous n'avons pas réussi à trouver toutes nos données à cette échelle, ce qui nous a menés à coder les zones d'emploi. Pour ce faire nous avons téléchargé le fichier sur le codage des zones d'emploi de 2020 sur l'Insee et utilisé le logiciel Tableau Prep Builder. Nous avons deux jeux de données qui ont nécessité un recodage par zone d'emploi qui sont le nombre d'établissements et la part de la population active. L'échelle géographique de ces deux jeux de données était les communes avec un code géographique pour chacune d'entre elles. Dans le fichier de codage de l'Insee, nous avions le code géographique, la commune, le code de la zone d'emploi et son nom. Nous avons donc créé une liaison entre ces deux tables avec le code géographique ce qui nous a donné une table avec le code géographique, la commune, le code de la zone d'emploi, son nom, le nombre d'établissements et la part de population active. Puis nous avons créé une étape d'agrégation et obtenu une table avec le nombre d'établissements et la part de la population active par zone d'emploi.

Nous avons remarqué une redéfinition des zones d'emploi en 2020, ainsi nous nous sommes retrouvés avec des données classées par zone d'emploi de 2020 et d'autres par zone d'emploi de 2010 pour des données de 2017. Nous avons fait le choix de garder les zones d'emploi de 2020 pour limiter la manipulation des données car nous avions plus de données avec les zones d'emploi de 2020 qu'avec celles de 2010. Afin de convertir les zones d'emploi de 2010 en zones d'emploi de 2020, nous avons utilisé le logiciel Tableau Prep Builder.

Dans un premier temps nous avons lié les tables des zones d'emploi de 2010 et des zones d'emplois de 2020 avec le code géographique des communes. Puis nous avons associé la table contenant la médiane du niveau de vie et le rapport interdécile avec la table des zones d'emploi en liant le code des zones d'emploi de 2010. Enfin avec une étape d'agrégation

nous avons obtenu les médianes de niveau de vie et le rapport interdécile par zones d'emploi de 2020.

Certaines données présentaient des données avec des effectifs, cependant avoir des données sous forme de proportions est plus intéressant pour nous. Nous avons donc calculé ces proportions sur Excel. Par exemple, pour la part de la population active dans le secteur primaire, nous avons divisé l'effectif d'individu dans le secteur primaire par l'effectif total de la population active.

Afin d'avoir un fichier unique avec toutes les variables socio-économiques par zone d'emploi, nous avons créé un fichier Excel constitué de 4 feuilles. La première feuille contient les variables socio-économiques par zone d'emploi avec des données en effectif. La deuxième feuille contient les mêmes données que la première mais nous avons cette fois converti ces données en proportions pour établir nos données finales. La troisième feuille contient les mêmes données que la deuxième mais sans unité. La dernière feuille contient le nom des variables et leur libellé. Enfin nous avons extrait la troisième feuille au format csv pour pouvoir réaliser son importation dans l'entrepôt de données phpMyAdmin.

Les analyses

Analyse descriptive

Les variables socio-économiques

Pour réaliser une analyse descriptive des variables socio-économiques nous avons réalisé des tris à plat et des graphiques élémentaires (tels que des diagrammes circulaires ou des diagrammes en barres) grâce à des tableaux croisés dynamiques sur Excel. Nous avons aussi calculé les coefficients de corrélation linéaire de Pearson pour certaines variables afin de mettre en évidence certaines ressemblances entre les banques.

Les banques

Pour réaliser l'analyse des différentes banques, nous avons dû créer un fichier contenant la proportion du nombre d'agences de chaque banque dans chaque zone d'emploi afin de pouvoir étudier leur répartition à l'échelle française. Ce fichier nous a également été utile pour pouvoir étudier l'impact de chaque donnée socio-économiques sur l'implantation des différentes banques. Ce fichier nous a aidé à présenter les banques et notamment à mettre en évidence leurs différentes implantations en France.

Analyse en composantes principales normée

Dans le but de déterminer si la présence d'une banque sur un territoire est liée avec certaines de nos variables socio-économiques, nous avons procédé à une Analyse en Composante Principales (ACP). Cette technique d'analyse de données est utile pour explorer des données multivariées, dégager des structures de liaisons entre des variables exclusivement quantitatives et regrouper les individus qui se ressemblent. En effet, on utilise l'ACP lorsqu'on dispose d'individus décrits par beaucoup de variables et qu'il est, de ce fait, difficile de les représenter graphiquement par exemple. Si nous voulions faire un graphique pour décrire nos individus, il faudrait faire une représentation en plusieurs dimensions et les résultats ne ressortiraient pas directement. L'ACP va donc nous fournir un résumé de

l'information contenue dans nos données en un nombre de dimensions le plus petit possible. Pour réaliser cette ACP, nous avons donc décidé de prendre en compte toutes nos variables socio-économiques ainsi que d'autres variables explicitant la répartition des quatre banques que nous étudions par zones d'emploi. Dans notre situation, la différence d'échelle de toutes nos variables nous laisse penser qu'une ACP normée serait plus significative. Effectivement, cela supprimera l'impact que la variance des variables pourrait avoir sur la représentation de nos données. Le logiciel SPAD nous a permis de faire ces analyses et d'en tirer certaines conclusions.

Pour paramétrier l'ACP sous SPAD, il nous a fallu dans un premier temps choisir le rôle de nos variables. Nous avons donc mis les variables "ZE 2020" et "Nom des ZE" en identifiant. Étant donné que nous voulions déterminer la liaison entre les variables socio-économiques et le nombre de banques par zones d'emploi, nous avons décidé de mettre ces quatre dernières en variables illustratives, pour voir où elles se situent par rapport aux autres sans que l'analyse ne les prenne en compte. Lorsqu'on fait tourner un ACP sur SPAD, le logiciel met à notre disposition deux types de sorties : des graphiques et des sorties numériques sur Excel. Dans un premier temps, nous avons regardé les graphiques pour s'assurer qu'il n'y avait pas d'individus atypiques qui biaisaient l'ACP. Nous avons donc remarqué que la zone d'emploi 1113 avait une coordonnée plus élevée que les autres sur les deux axes et que les zones d'emplois 1114 et 1115 avaient une coordonnée trop élevée sur le deuxième axe pour que l'on puisse les conserver dans notre analyse. Nous avons donc choisi de mettre ces individus en illustratif pour qu'ils ne soient pas pris en compte dans l'ACP afin que celle-ci ne soit pas biaisée par des individus trop atypiques.

Régression linéaire

Pour ajouter une notion prédictive à notre analyse, nous avons effectué une régression linéaire sur le nombre de banque par zone emploi en fonction des variables socio-économiques que nous avons sélectionné auparavant. Ces régressions permettent notamment d'identifier les variables influant sur la présence de chaque banque au sein d'une zone d'emploi. Ainsi, nous avons effectué quatre régressions linéaires pour les quatre banques étudiées pour analyser les stratégies commerciales des différentes banques.

Tout d'abord, nous avons remarqué que la répartition du nombre d'établissements pour banque données par zone d'emploi présentait une distribution très asymétrique. Nous avons donc décidé d'étudier le logarithme de cette qualité. De plus, ayant des zones emploi avec parfois aucun établissement associé pour certaines banques nous avons donc au final étudié le logarithme du nombre d'établissement pour une banque fixée plus un ($\ln(X+1)$). Nous avons supprimé les variables "population active travaillant dans le secteur tertiaire non marchand", "établissement de secteur tertiaire non marchand", la catégorie socio-professionnelle "Autres", "la part de la population âgée de 90 ans et plus" afin de bénéficier d'une matrice de rang plein. En effet, la régression serait impossible avec des variables liées par combinaison linéaire. C'est donc pour inhiber ce phénomène que nous avons supprimé ces variables.

Nous avons donc identifié le " $\log(X+1)$ " (où X représente le nombre d'établissements pour une banque fixée on le rappelle) comme variable Y à expliquer et le reste des variables (sauf citées précédemment) comme variables X explicatives. Nous avons donc lancé la commande "linear model" pour observer les résultats de cette régression. Les résultats associés nous ont entre autres permis d'apprécier la pertinence du modèle. Pour les quatre cas, la pertinence (via le test de Fisher) était vérifiée. Après avoir observé la table d'ANOVA

associée et le résumé des résultats obtenus, nous avons vérifié directement l'analyse des résidus pour pouvoir confirmer que l'utilisation de cette régression linéaire multiple a un sens. Pour cela, après avoir récupéré les résidus du modèle, nous les avons transformés en résidus standardisés puis studentisés. Nous avons, dans les quatre cas, des résidus qui semblent suivre une loi de Student. On observe pour les quatre banques un histogramme des résidus studentisés similaire à une loi de Student comme on peut l'observer sur la figure en annexe 3.

Une fois cette confirmation effectuée, nous avons effectué une sélection de variables avec la fonction "step" qui est une fonction pas-à-pas que nous avons programmé de manière descendante. Nous avons par la suite récupéré le modèle optimisé afin de l'étudier afin d'analyser les résultats finaux. Le résumé nous fournit la pertinence des variables de ce nouveau modèle (entre autres). Pour pouvoir utiliser ces informations, nous avons effectué un test de sous-modèle (test de Fisher) qui confronte notre modèle initial (avec toutes les variables) à notre "plus petit" modèle qui ne contient que les variables jugées les plus discriminantes par R pour ce modèle. Dans les quatre cas, la statistique de test n'appartient pas à la région de rejet. Nous avons donc conservé dans les quatre régressions effectuées le sous-modèle trouvé. Pour terminer notre analyse, nous avons étudié le nouveau modèle en observant une seconde fois la pertinence de celui-ci via le test de Fisher puis nous avons également effectué une nouvelle fois encore une étude des résidus(voir une fois encore l'annexe 3).

Pour les quatre banques encore, nos nouveaux modèles se sont avérés pertinents. Ainsi, nous avons pu identifier les variables intervenant dans la stratégie d'implantation des banques. Il faut noter que notre modèle aurait pu être affiné avec le calcul des VIF associé à chacune des variables pour atténuer les problèmes de colinéarités entre les variables. Nos tentatives de calculs ont été abandonnées par manque de cohérence et de justesse.

Glossaire

A

Analyse en composante principale (ACP)

Méthode d'analyse de données qui permet d'explorer des données multivariées et d'en extraire les informations les plus importantes.

B

Banque commerciale

Établissement financier dont les activités sont majoritairement tournées vers les particuliers (dépôts, placements, solutions d'épargne, crédit), les entreprises ou les collectivités publiques. Leur capital est détenu par des actionnaires n'étant pas nécessairement clients.

Banque coopérative

Association autonome de personnes volontairement réunies pour satisfaire leurs aspirations et besoins économiques, sociaux et culturels communs aux moyens d'une entreprise dont la propriété est collective et où le pouvoir est exercé démocratiquement.

E

Économie sociale et solidaire (ESS)

Ensemble d'entreprises organisées sous forme de coopératives, mutuelles, associations, ou fondations, dont le fonctionnement interne et les activités sont fondés sur un principe de solidarité et d'utilité sociale.

I

Institut national de la statistique et des études économiques (Insee)

Direction générale du ministère de l'Économie et des Finances. Il a pour mission de collecter, analyser et diffuser des informations sur l'économie et la société française sur l'ensemble de son territoire.

M

Macro Excel

Programme qui permet d'automatiser une séquence d'actions ou de calculs récurrents, et exécute des fonctions avancées que ne propose pas l'interface standard d'Excel. Elle peut être réalisée à l'aide de l'enregistreur de macros Excel ou dans l'éditeur VBA d'Excel.

P

PhpMyAdmin

Entrepôt de données qui permet de gérer des bases de données MySQL.

R

R

Langage de programmation utilisé pour le traitement de données et l'analyse statistique.

Rapport interdécile

Rapport entre le neuvième décile (niveau de vie qui sépare les 90 % qui touchent le moins des 10 % qui touchent le plus) et le premier décile (niveau de vie qui sépare les 10 % les plus pauvres des 90 % restants). Le rapport interdécile mesure combien de « fois » les plus riches reçoivent ce dont les pauvres disposent.

Régression Linéaire

Un modèle de régression linéaire est un modèle qui cherche à établir une relation linéaire entre une variable, dite expliquée, et une ou plusieurs variables, dites explicatives.

Rstudio

Environnement de développement gratuit, libre et multiplateforme pour R. un langage de programmation utilisé pour le traitement de données et l'analyse statistique.

S

Scraping

Technique de fouille de données qui permet d'extraire les informations d'un site web par l'intermédiaire d'un code.

SPAD

Logiciel destiné au Data Mining et à l'analyse prédictive des données.

T

Tableau Prep Builder

Logiciel de nettoyage et d'organisation de données.

Z

Zone d'emploi

Partition créée par l'Insee qui découpe le territoire français en espaces géographiques à l'intérieur desquels la plupart des actifs résident et travaillent.

Introduction

Au XIXe siècle, l'allemand Frédéric-Guillaume RAIFFEISEN a développé le concept de banque coopérative en créant des "caisses rurales" proches géographiquement de ses sociétaires et dans lesquelles ils pouvaient emprunter de l'argent relativement facilement. Ce concept a donné naissance à plusieurs banques en France, qui aujourd'hui encore fonctionnent en suivant le principe de l'entraide financière et qui s'inscrivent dans une logique d'économie sociale et solidaire.

En effet, les banques coopératives étant des établissements dirigés entièrement par leurs sociétaires et leur appartenant, indépendamment de la quantité de capital qu'ils apportent, elles font partie de l'économie sociale et solidaire. Cette dernière se définit comme étant un ensemble d'entreprises organisées sous forme de coopératives, mutuelles, associations, ou fondations, dont le fonctionnement interne et les activités sont fondées sur un principe de solidarité et d'utilité sociale. Ainsi, la finalité d'une banque coopérative n'est pas de permettre à ses actionnaires de faire du profit (ce qui est le cas pour les banques commerciales) mais plutôt de rendre l'accès au crédit plus facile pour chacun. On distingue donc, en France, les banques coopératives des banques commerciales. On peut alors se demander comment expliquer la diversité des organisations ou entreprises qui se développent dans notre système économique vu comme capitaliste. Notre commanditaire, Amélie ARTIS, organise ses recherches autour de cette problématique et a choisi de nous solliciter pour que nous contribuions à son étude sur l'économie sociale et solidaire. Plus particulièrement, elle a exprimé sa volonté d'en savoir plus sur l'implantation des banques coopératives en France métropolitaine.

C'est la raison pour laquelle nous nous sommes demandé s'il était possible d'identifier des caractéristiques de la population qui définissent l'implantation des banques coopératives en France. Pour répondre à cette problématique, nous avons mis en œuvre un certain nombre de techniques afin de produire une typologie des établissements que nous étudions.

À partir de données provenant de l'Insee ou directement du site Web de chaque banque étudiée, nous avons récolté les informations nécessaires à l'élaboration de statistiques descriptives et d'analyses factorielles. Après avoir démontré la diversité socio-économique des différents territoires français, nous nous efforcerons de mettre en évidence les disparités entre les banques et enfin étudierons le lien entre l'implantation des banques et les caractéristiques du territoire français.

I. La diversité socio-économique du territoire français

Notre étude a été basée sur les zones d'emploi. La zone d'emploi est l'espace géographique où la plupart des actifs vivent et travaillent. La division en zones d'emploi constitue une division territoriale propice à la recherche locale sur le marché du travail. C'est une référence pour la diffusion des taux de chômage locaux et des estimations de l'emploi. Cette zone est définie pour la France métropolitaine mais également pour les DOM. Un premier découpage par zone d'emploi avait été effectué en 2010 et a été actualisé en 2020. Sur l'Insee, la ventilation mise à jour est basée sur les mouvements de navetteurs domicile-travail observés lors du recensement de 2016.

Nous avons, à la demande de notre commanditaire Amélie ARTIS, utilisé le découpage de la France selon cette échelle. Ainsi, nous avons sélectionné différentes variables socio-économiques sur le site internet de l'Insee afin de dresser une caractérisation des territoires. Nous avons pour cela choisi des variables datant de 2017 pour nous assurer de l'existence des données. L'Insee nous fournissait ces informations par zone d'emploi actualisée ou non, ce qui a nécessité un nettoyage des données pour les uniformiser. Ces variables sont réparties selon trois catégories distinctes résumant chacune le profil de la population au sein de ces territoires.

Ainsi, nous distinguons les variables caractérisant la population avec:

- La répartition des classes d'âge
- La répartition des catégories socioprofessionnelles

Les variables décrivant le tissu économique du territoire avec :

- Le taux de chômage
- La part des emplois salariés
- La part de la population active
- Le nombre d'établissements

Les variables illustrant les revenus et les inégalités au sein du territoire avec :

- Le salaire net horaire
- La médiane du niveau de vie
- Le revenu moyen des différentes activités par heure
- Le rapport inter décile

Après avoir récupéré et nettoyé ces données afin de pouvoir les étudier par zone d'emploi, nous avons pu effectuer des analyses pour observer les disparités et les similitudes entre les différentes zones d'emploi sur le territoire français. Cela a constitué le point de départ de nos analyses pour traduire avant tout la disparité des types de territoires avant d'aborder la notion de banque par la suite.

Le salaire net horaire moyen

Commençons par étudier le salaire net horaire moyen par zone d'emploi pour voir s'il existe de grandes disparités entre les zones d'emplois.

Tableau 1 : Salaire net horaire moyen en fonction de la zone d'emploi

Salaire net horaire moyen	Nombre de Zone d'Emploi	Fréquence
Moins de 13€	149	53,21%
Entre 13€ et 16€	122	43,57%
Entre 16€ et 19€	5	1,79%
Plus de 19€	4	1,43%
Total	280	100,00%

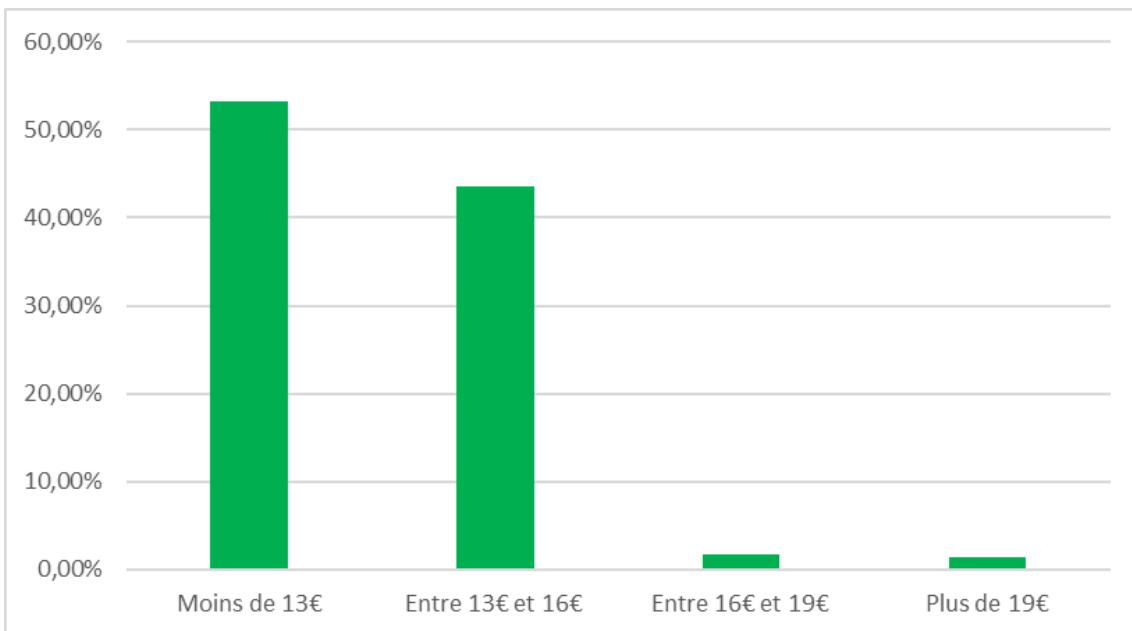


Figure 1 : Part des zones d'emploi en fonction du salaire net horaire moyen

Comme nous pouvons le voir sur le tableau 1 et la figure 1, la grande majorité (plus de 96 %) des zones d'emploi ont un salaire net horaire moyen inférieur à 16 €. Parmi celles-ci, plus de la moitié ont un salaire net horaire moyen inférieur à 13 €. Nous pouvons également observer le fait que 4 zones d'emploi ont un salaire net horaire moyen supérieur à 19 €. Ce sont les zones d'emplois de Versailles-Saint-Quentin, Seine-Yvelinoise, Paris et Rambouillet avec des salaires nets horaires moyens de respectivement 21,12 €, 20,42 €, 20,12 € et 20,05 €.

Les secteurs économiques

Présentons maintenant la part des différents secteurs économiques en France avant de les étudier par zone d'emploi pour voir s'il existe des zones d'emploi qui ne suivent pas la moyenne française.

Tableau 2 : Part des différents secteurs économiques en France

Secteur économique	Effectif	Fréquence
Primaire	243 186	1,00%
Secondaire	4 399 399	18,06%
Tertiaire Marchand	11 938 097	49,00%
Tertiaire Non Marchand	7 784 892	31,95%
Total	24 365 574	100,00%

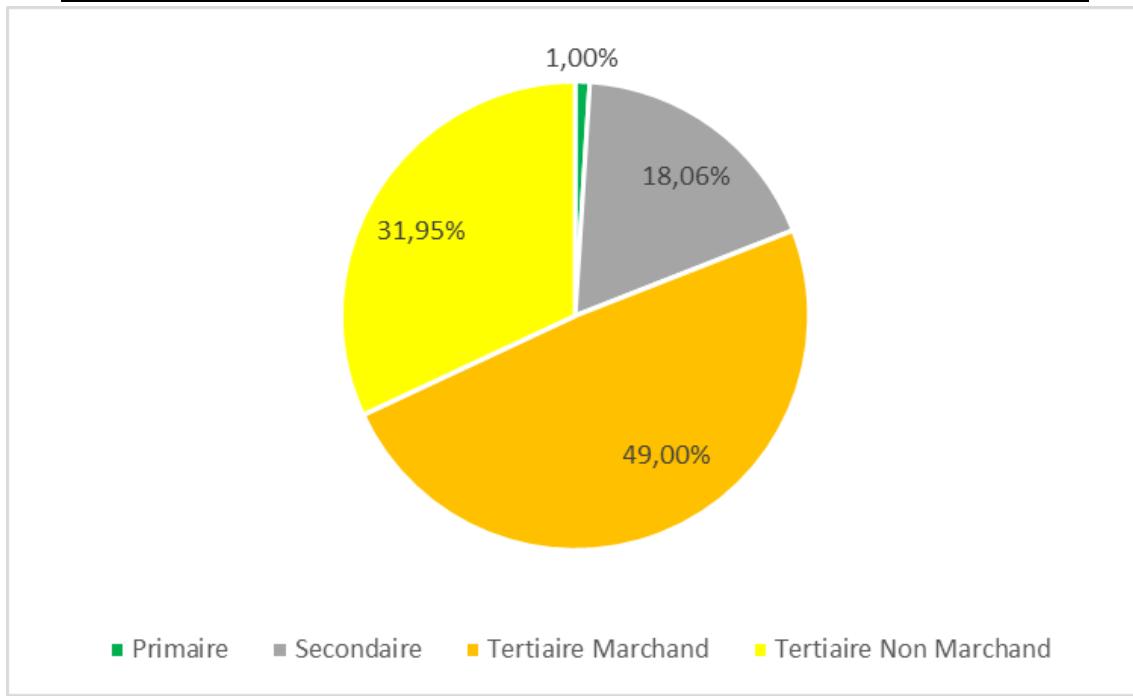


Figure 2 : Part de chaque secteur économique en France

Nous pouvons observer sur le tableau 2 et le graphique 2 que la répartition entre les 3 secteurs économiques est très inégale en France. En effet, le secteur primaire ne représente que 1 % de tous les emplois en France. Le secteur secondaire regroupe quant à lui quasiment un cinquième de la population française alors que le secteur tertiaire est le plus représenté en France avec plus de 80 % des emplois, notamment le secteur tertiaire marchand qui représente pratiquement un emploi sur 2 en France. On peut donc se demander si toutes les zones d'emploi suivent ce schéma ou s'il existe des disparités au sein de celles-ci.

Tableau 3 : Statistiques élémentaires sur les différents secteurs économiques

Secteur économique	Primaire	Secondaire	Tertiaire Marchand	Tertiaire Non Marchand
Minimum	0,02%	8,72%	27,00%	16,62%
Premier quartile	0,76%	17,53%	36,50%	29,70%
Médiane	1,36%	21,50%	40,40%	33,27%
Troisième quartile	2,30%	27,44%	45,62%	37,64%
Maximum	29,59%	47,98%	66,11%	51,51%
Moyenne	1,98%	23,04%	41,41%	33,57%
Variance	0,06%	0,58%	0,47%	0,40%
Écart-type	2,52%	7,66%	7,04%	6,38%
Étendue	29,57%	39,26%	39,11%	34,90%
Longueur interquartile	1,54%	9,91%	9,12%	7,94%

Nous observons sur le tableau 3 ci-dessus qu'il existe de très fortes disparités au sein des différentes zones d'emploi. En effet, nous pouvons voir qu'une zone d'emploi possède quasiment 30 % d'emplois dans le secteur primaire, ce qui est 30 fois plus que la moyenne française. Il s'agit de la zone d'emploi de Lesparre-Médoc, nous pouvons donc en déduire que cette zone d'emploi est une zone très agricole. La zone d'emploi de La Vallée de la Bresle-Vimeu compte quasiment 50 % de ses emplois dans le secteur secondaire, ce qui représente plus du double de la moyenne française, nous pouvons donc en déduire que cette zone d'emploi est très industrialisée. Enfin pour le secteur tertiaire il n'existe pas de très gros extrême même si on peut citer la zone d'emploi de Paris, qui possède les deux tiers de ses emplois dans le secteur tertiaire marchand, et celle de Châlons-en-Champagne, qui possède plus de la moitié de ses emplois dans le secteur tertiaire non marchand.

Le taux de chômage

Nous pouvons donc nous interroger maintenant sur le taux de chômage afin de voir s'il existe des disparités au sein des différentes zones d'emploi en France.

Tableau 4 : Taux de chômage en fonction de la zone d'emploi

Taux de chômage	Nombre de Zone d'Emploi	Fréquence
Moins de 6%	12	4,29%
Entre 6% et 9%	143	51,07%
Entre 9% et 12%	100	35,71%
Entre 12% et 15%	22	7,86%
Plus de 15%	3	1,07%
Total général	280	100,00%

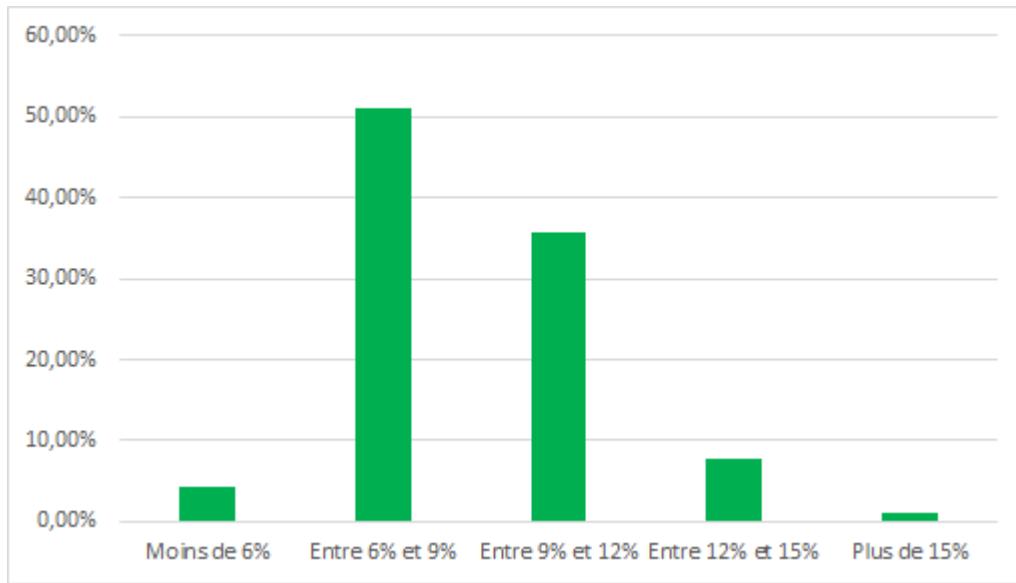


Figure 3 : Part des zones d'emploi en fonction du taux de chômage

Comme nous pouvons le voir dans le tableau 4 et la figure 3, plus de la moitié des zones d'emploi en France possèdent un taux de chômage inférieur à 9 % alors qu'environ 10 % des zones d'emploi ont un taux de chômage supérieur à 12 %. Les trois zones d'emploi avec les plus grands taux de chômage sont les zones d'Agde-Pézenas, Maubeuge et Calais avec des taux de chômage de respectivement 17,1 %, 15,7 % et 15,3 %.

L'âge de la population

Présentons maintenant la répartition en âge de la population au sein des différentes zones d'emploi en France pour voir s'il existe des zones d'emplois plus dynamiques que d'autres de ce point de vue-là.

Tableau 5 : Part moyenne des différentes catégories d'âge parmi toutes les zones d'emploi

Âge	Fréquence
Moins de 14 ans	17,78%
Entre 15 ans et 29 ans	15,62%
Entre 30 ans et 44 ans	17,61%
Entre 45 ans et 59 ans	20,59%
Entre 60 ans et 74 ans	17,87%
Entre 75 ans et 89 ans	9,14%
Plus de 90 ans	1,38%
Total	100,00%

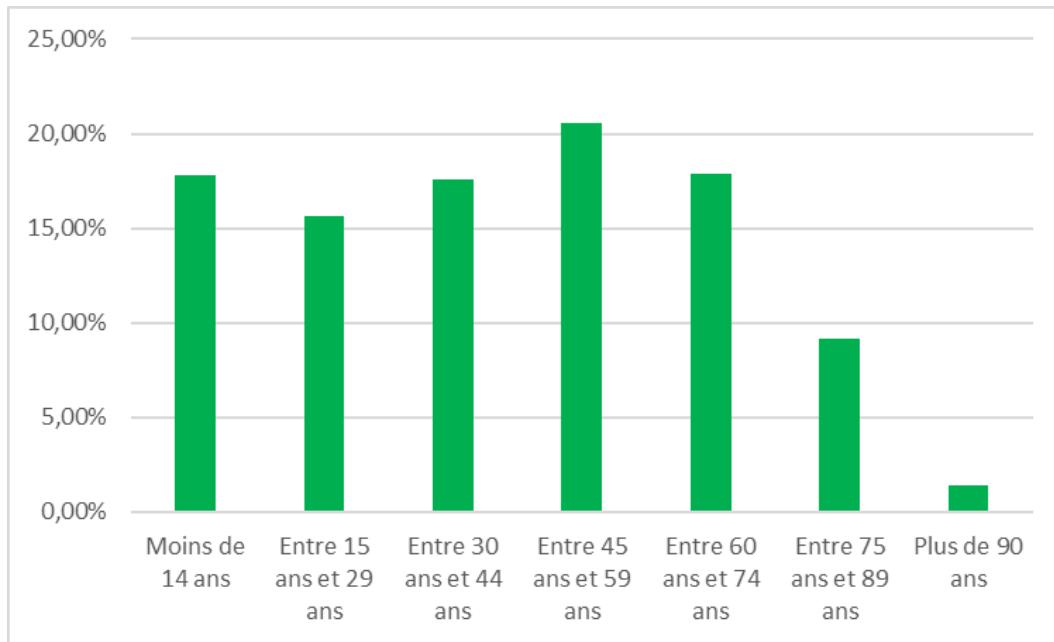


Figure 4 : Part moyenne des différentes catégories d'âge parmi toutes les zones d'emploi

Comme nous pouvons le voir dans le tableau 5 et la figure 4, en moyenne au sein d'une zone d'emploi plus de 30 % de la population à en dessous de 30 ans, un peu moins de 60 % de la population a entre 30 ans et 74 ans et seulement 10 % de la population a plus de 75 ans. Il est donc intéressant d'étudier cette classification au sein des zones d'emploi pour voir s'il existe de grandes disparités.

Tableau 6 : Statistiques élémentaires sur la répartition des âges de la population au sein des zones d'emploi

Âge	Moins de 14 ans	15-29 ans	30-44 ans	45-59 ans	60-74 ans	75-89 ans	Plus de 90 ans
Minimum	12,16%	10,65%	12,59%	10,33%	3,22%	0,79%	0,08%
Premier quartile	16,18%	13,94%	16,48%	19,93%	16,01%	7,74%	1,06%
Médiane	17,55%	14,99%	17,66%	20,52%	18,04%	9,19%	1,37%
Troisième quartile	18,96%	16,92%	18,76%	21,18%	19,88%	10,62%	1,68%
Maximum	40,58%	24,90%	23,97%	26,53%	27,56%	14,74%	2,51%
Moyenne	17,78%	15,62%	17,61%	20,59%	17,87%	9,14%	1,38%
Variance	0,07%	0,06%	0,03%	0,02%	0,09%	0,05%	0,00%
Écart-type	2,70%	2,50%	1,75%	1,41%	3,02%	2,20%	0,42%
Étendue	28,42%	14,25%	11,38%	16,20%	24,34%	13,95%	2,43%
Longueur interquartile	2,78%	2,98%	2,28%	1,25%	3,88%	2,88%	0,62%

Nous pouvons voir dans le tableau 6 ci-dessus qu'il existe de fortes disparités d'âge parmi les zones d'emploi. Nous pouvons prendre l'exemple de la zone d'emploi de Saclay qui est la zone d'emploi la plus jeune avec plus de 40 % de sa population qui a moins de 14 ans mais également moins de 1 % de sa population qui a plus de 75 ans. La zone d'emploi où la population est la plus âgée est quant à elle la zone de Foix-Pamiers avec quasiment 45 % de sa population qui a plus de 60 ans.

II. La diversité des banques françaises

Présentation des banques

Le Crédit Agricole

Le Crédit Agricole est connu sous le nom de «Banque verte» en raison de son activité initiale au service du monde agricole. C'est le plus grand réseau mondial de banques coopératives et de banques mutuelles. Le groupe Crédit Agricole se positionne aujourd'hui au deuxième rang des banques en France avec un PNB (Produit Net Bancaire) de 31,8 milliards en 2015 et à la troisième position des banques européennes. Le groupe Crédit Agricole intervient sur les trois grands métiers bancaires : la banque de détail, la gestion de l'épargne et la banque de financement et d'investissement. Le groupe Crédit Agricole est également présent sur les métiers du crédit-bail, de l'affacturage et de l'assurance.

Après l'acquisition de Crédit Lyonnais (rebaptisé LCL par la suite) en 2003, le groupe Crédit Agricole s'affirme aujourd'hui comme la première banque de détail avec plus de 8 900 implantations commerciales (7 000 agences Crédit Agricole et plus de 1 900 agences LCL). Le groupe, présent dans 50 pays, emploie plus de 138 000 collaborateurs en France et gère un portefeuille de 52 millions de clients.

La Banque Populaire

La Banque Populaire est un groupe mutualiste de services bancaires et financiers français. C'est une société anonyme à directoire et conseil de surveillance dont le siège social est situé à Paris, avenue Pierre Mendès France. C'est le deuxième groupe bancaire en France.

La Banque Populaire a été créée en 1878 pour permettre aux entrepreneurs de bénéficier de prêts bancaires pour réaliser des projets. En 1917, la loi Clémentel est adoptée dans le but d'organiser les prêts bancaires et de restaurer l'économie d'après-guerre. De nombreux changements sont intervenus au sein de la Banque Populaire, notamment en 2009, avec la fin de la Banque Populaire. En effet, après la crise de 2008, la Banque Populaire et la Caisse d'Épargne ont fusionné pour former BPCE, qui est le centre d'intérêt des deux groupes.

En 2018, l'organisation s'est implantée dans plusieurs régions pour se rapprocher des clients membres et conserver une certaine indépendance. Le groupe compte 12 Banques populaires régionales, CASDEN et Crédit Coopératif, et environ 3 330 agences en France, au service de 8 900 000 clients, dont 3 900 000 adhérents.

Le Crédit Mutuel

Banque coopérative et mutuelle fortement implantée en France à travers une organisation non centrale, le Crédit Mutuel a pour objectif principal la qualité des relations et des services qu'il apporte à ses adhérents et clients.

Ses valeurs de responsabilité, de solidarité et d'engagement social sont aussi stratégiques que la compétitivité des services, faisant de Crédit Mutuel une banque du secteur associatif français.

Le Crédit Mutuel compte 330 000 associations clientes et est un partenaire actif de toutes les associations, de la plus petite association basée sur le volontariat à la plus grande agence pour l'emploi. Il soutient les besoins du secteur et joue un rôle fondamental dans la mise en place de nouvelles formes de solidarité.

La Société Générale

Créée le 4 mai 1864, La Société Générale est l'une des plus importantes banques de France. Forte de ses 147 ans d'histoire, elle a formé l'actuelle banque LCL avec BNP Paribas et le Crédit Lyonnais, les trois plus anciennes banques de France.

D'un point de vue juridique, SG Bank est devenue une société privée depuis le 29 juillet 1987. Les salariés détiennent 7,13% des actions. Son siège social est situé au 29 Avenue de Paris et Avenue Haussmann 75009, Paris. L'entreprise compte 157 000 employés et exerce dans les domaines de la finance, de l'assurance, du conseil et de la banque. De plus, elle possède plusieurs filiales comme les entreprises ADL Automotive, Boursorama, Rosbank et Crédit du Nord.

La Société Générale est caractérisée par sa dimension mondiale avec une forte présence dans certains pays européens et africains. C'est un groupe également marqué par sa solidité financière avec une capitalisation qui était évaluée à 36,2 milliards d'euros au 31 décembre 2009. La même année, ses fonds propres étaient estimés à 42,2 milliards d'euros. La même année, le chiffre d'affaires de la banque atteignait les 21,73 milliards d'euros. Par ailleurs, elle est cotée à la bourse de Paris. Membre de l'indice CAC40, elle occupe le sixième rang en capitalisation boursière, en France.

La banque ne veut pas rester à la limite du développement durable prôné par tous les pays du monde. En tant que partenaire du développement durable, elle prend en compte ses responsabilités environnementales et sociales. Cependant, le salaire de son manager étant 520 fois supérieur au salaire moyen en France, il est difficile d'obtenir la certification ISO 14001 et l'agrément d'Entreprise Solidaire.

À l'échelle de la France Métropolitaine (hors Corse)

Tableau 7 : Tri à plat du nombre d'agences de chaque Banque en France

BANQUE	EFFECTIF	FREQUENCE
CREDIT AGRICOLE	5028	52,77%
BANQUE POPULAIRE	2010	21,09%
CREDIT MUTUEL	1333	13,99%
SOCIETE GENERALE	1158	12,15%
TOTAL	9529	100,00%

Nous pouvons observer dans le tableau 7 ci-dessus que suite au scraping, nous avons récolté au total 9529 adresses de banques.

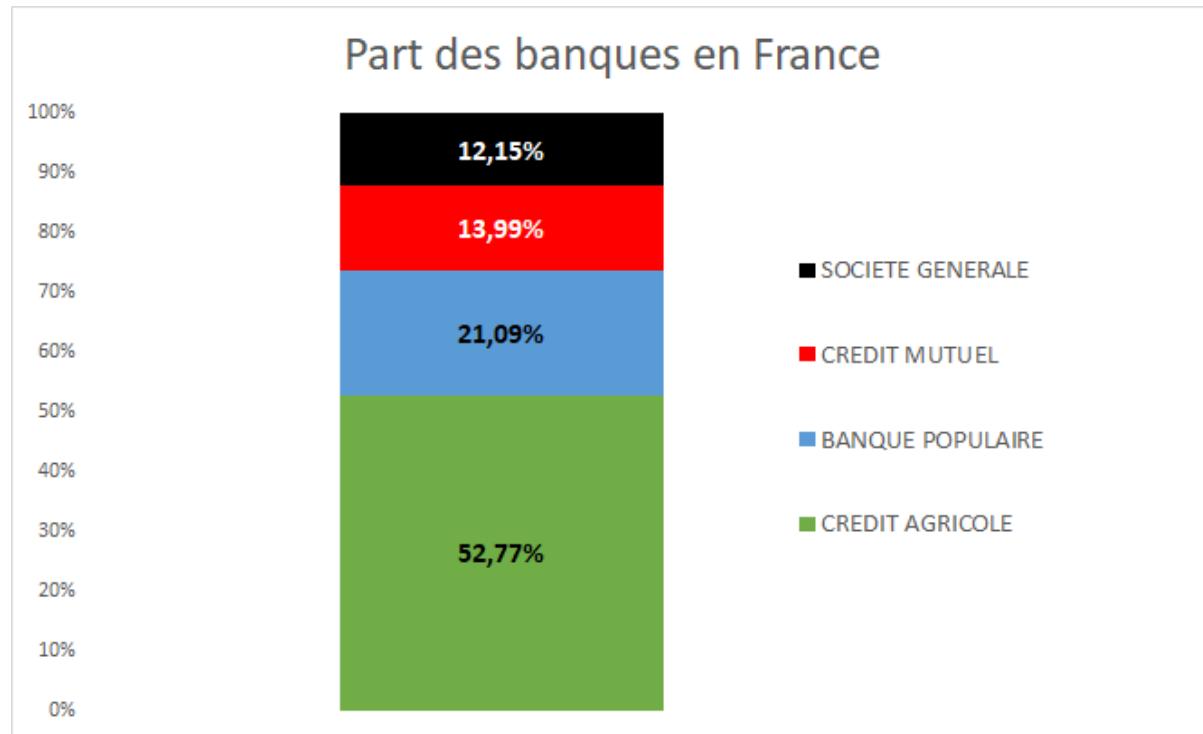


Figure 5 : Part des différentes banques en France

Sur la figure 5, nous pouvons voir que parmi les 9529 adresses collectées, un peu plus d'une adresse sur deux appartient est associée à une agence du Crédit Agricole, un cinquième des adresses viennent des agences de la Banque Populaire, environ 14 % sont des adresses des banques du Crédit Mutuel et environ 12 % appartiennent à la Société Générale.

À l'échelle régionale

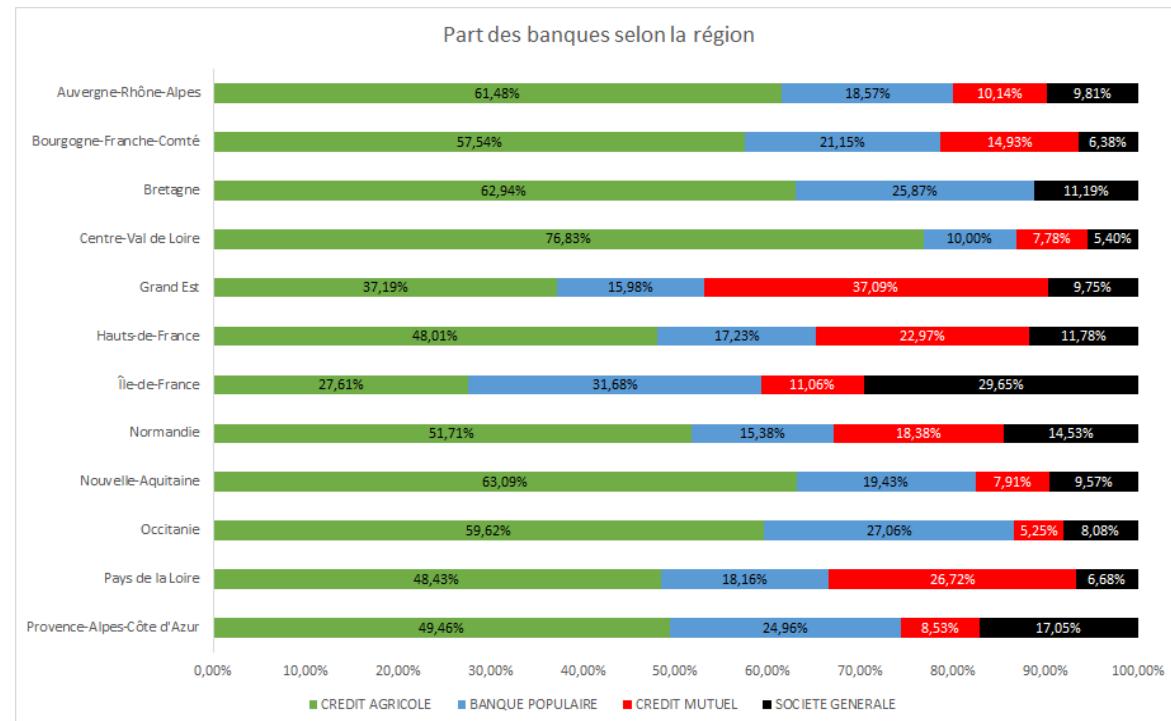


Figure 6 : Part des différentes banques dans chaque région

Sur la figure 6, nous pouvons remarquer que la banque la plus présente dans toutes les régions, excepté l'Île-de-France, est le Crédit Agricole. En effet, dans la région du Centre-Val de Loire, un peu plus de trois quarts des adresses collectées sont des agences du Crédit Agricole contre environ trois adresses sur dix en Île-de-France.

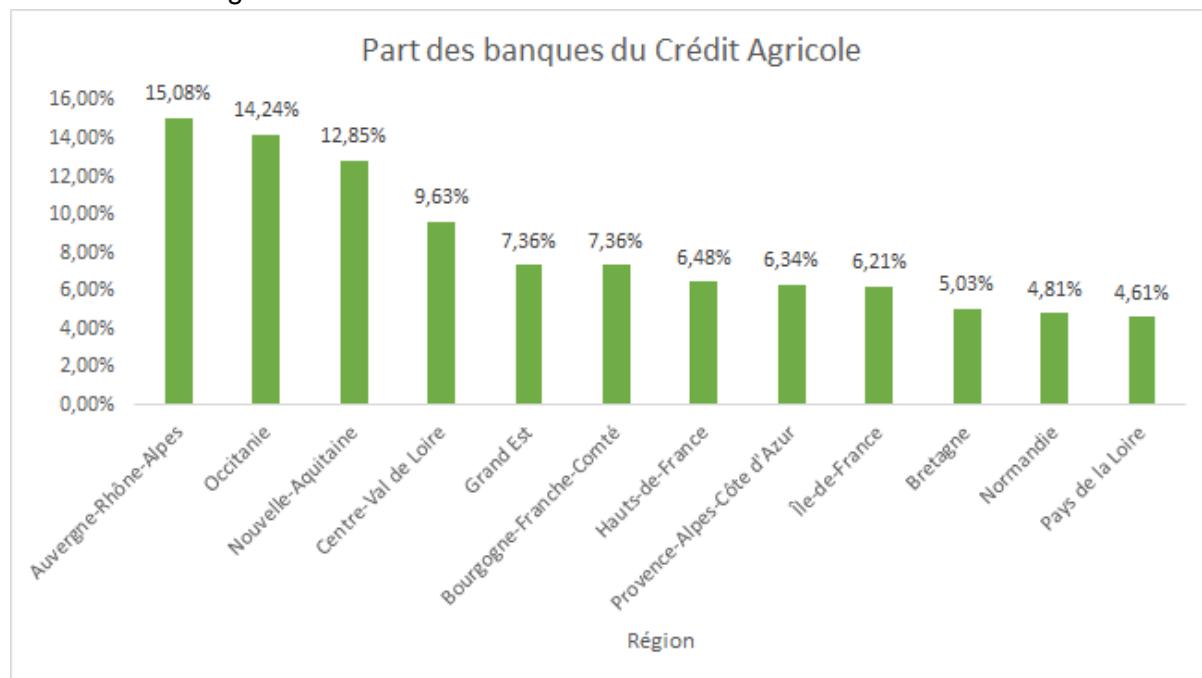
La Banque Populaire arrive en deuxième position dans huit régions avec une proportion d'agences allant d'une adresse sur dix à 3 adresses sur dix et quatre fois en troisième position avec une part allant de 16 % à environ 18 %.

Le Crédit Mutuel arrive quant à lui quatre fois en deuxième position avec 18% des adresses pour la Normandie, 23 % dans les Hauts-de-France, 27 % dans les Pays de la Loire et 37 % dans le Grand Est ; trois fois en troisième position avec une part allant de 8 % à 15 % ; quatre fois en dernière position avec une proportion des adresses allant de 5 % à 11 %.

Enfin, la Société Générale se situe aux deux dernières positions dans toutes les régions sauf l'Île-de-France avec une proportion d'établissements allant de 5 % à 17 %. La Société Générale arrive donc en deuxième position en Île-de-France avec 30 % des adresses.

Le Crédit Agricole

Nous allons maintenant nous concentrer sur le Crédit Agricole et sa présence dans les différentes régions.

*Figure 7 : Part des agences du Crédit Agricole selon les régions*

Sur la figure 7, nous pouvons observer que les régions où le Crédit Agricole est le plus implanté sont l'Auvergne-Rhône-Alpes avec 15 % des établissements et l'Occitanie avec 14 %. Quant aux régions les moins représentées, ce sont la Normandie et les Pays de la Loire avec respectivement 4,81 % et 4,61 %.

La Banque Populaire

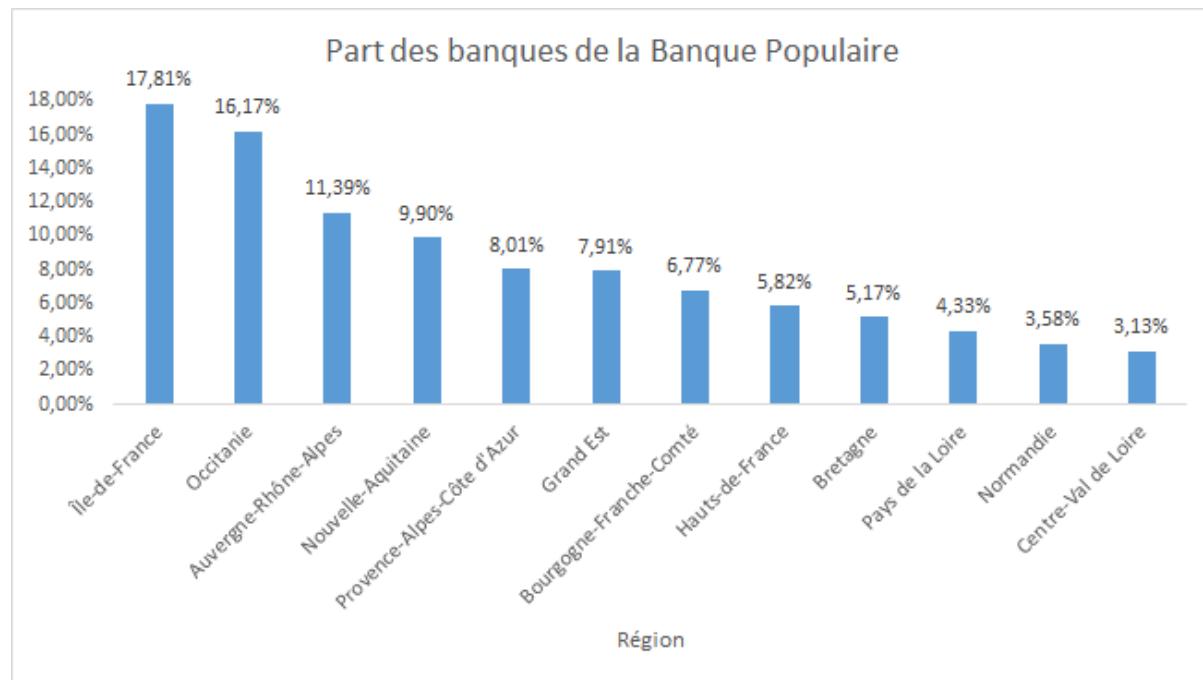


Figure 8 : Part des agences de la Banque Populaire selon les régions

Sur la figure 8, pour la Banque Populaire, nous pouvons voir que les régions les plus représentées en termes d'agences sont l'Île-de-France et l'Occitanie avec respectivement 17,81 % et 16,17 %. Au contraire la Normandie avec 3,58 % et le Centre-Val de Loire avec 1,13 % sont les régions avec le moins d'agences de la Banque Populaire.

Le Crédit Mutuel

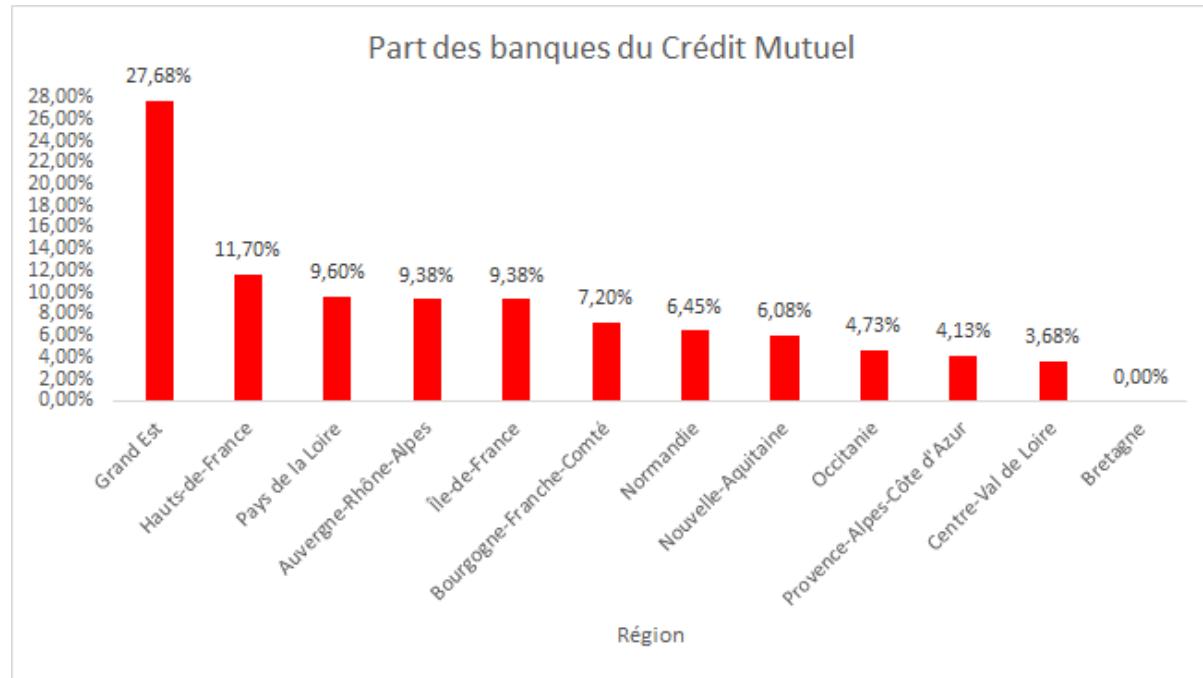


Figure 9 : Part des agences du Crédit Mutuel selon les régions

Sur la figure 9, nous pouvons observer que les régions où le Crédit Mutuel est le plus implanté sont le Grand Est avec 28 % des agences et les Hauts-de-France avec 12 %. Quant aux régions les moins représentées, nous avons la région Provence-Alpes-Côte d'Azur avec 4,13 % et le Centre-Val de Loire avec 3,68 %.

La Société Générale

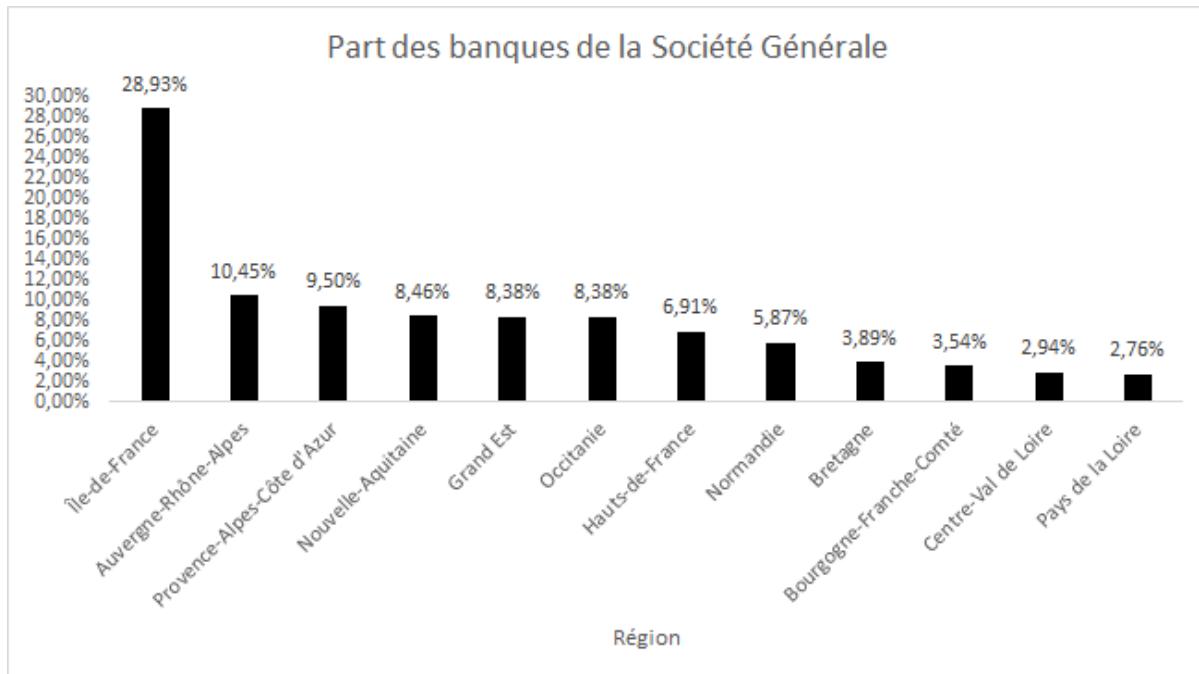


Figure 10 : Part des agences de la Société Générale selon les régions

Sur la figure 10, nous pouvons observer que les régions Île-de-France et Auvergne-Rhône-Alpes sont celles qui ont le plus d'agences de la Société générale avec respectivement 29 % et 10,5 %.

À l'échelle départementale

Le Crédit Agricole

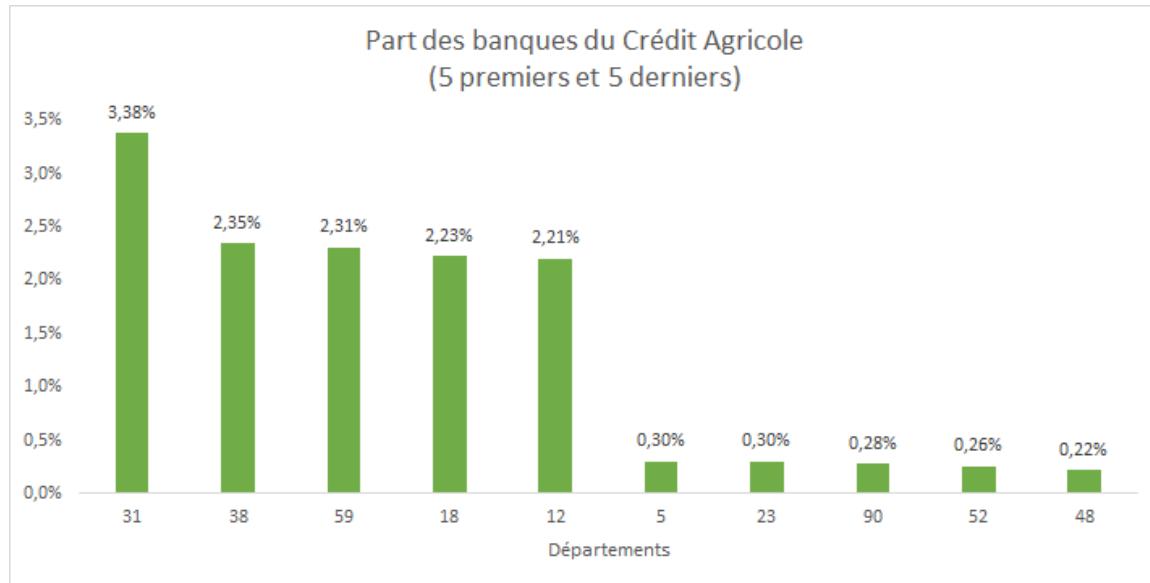


Figure 11 : Part des agences du Crédit Agricole selon les départements (5 premiers et 5 derniers)

Nous avons sur la figure 11 les cinq départements les plus représentés en termes d'établissements du Crédit Agricole et les cinq départements les moins représentés. Nous avons donc la Haute-Garonne en première position avec 3,38 % des établissements, l'Isère en deuxième position avec 2,35 %, la Haute-Marne en avant dernière position avec 0,26 % et la Lozère avec 0,22 %.

La Banque Populaire

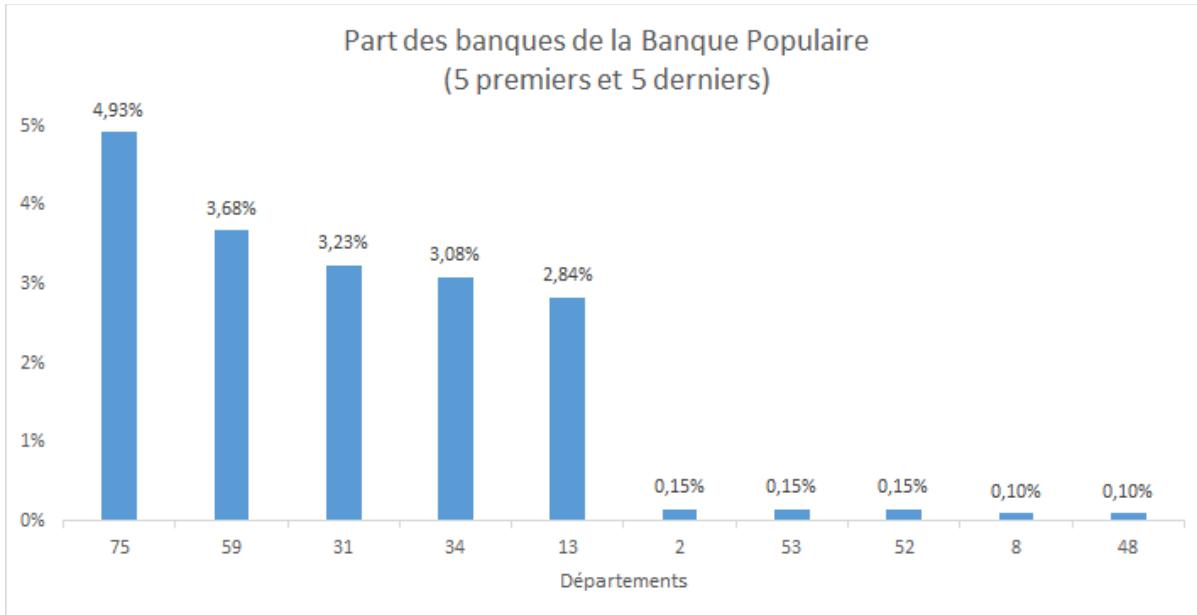


Figure 12 : Part des agences de la Banque Populaire selon les départements (5 premiers et 5 derniers)

Nous pouvons voir sur la figure 12 que Paris est le département où il y a le plus d'agences de la Banque Populaire avec 4,93 % des agences totales et le Nord est le

deuxième département avec le plus d'agences avec 3,68 %. Quant aux départements des Ardennes et la Lozère, ce sont les départements avec le moins d'agences de la Banque populaire.

Le Crédit Mutuel

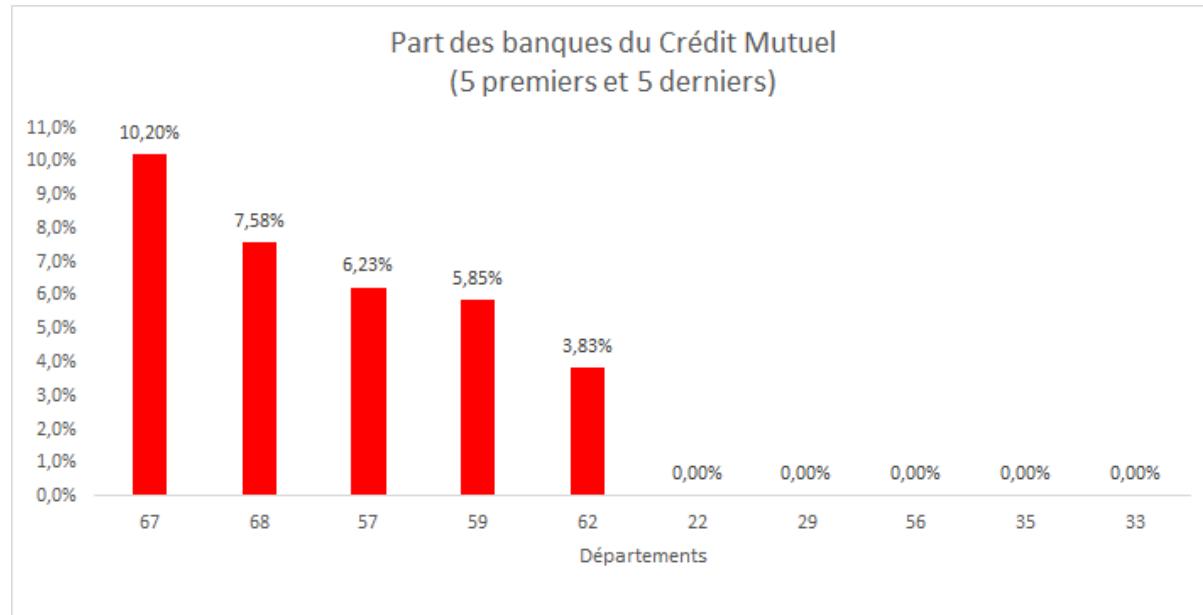


Figure 13 : Part des agences du Crédit Mutuel selon les départements (5 premiers et 5 derniers)

Nous pouvons observer qu'il existe des départements (hors Bretagne) où il n'y a pas d'agence du Crédit Mutuel comme par exemple le département de la Gironde (33). Néanmoins nous pouvons voir sur la figure 13 que le Bas-Rhin et le Haut-Rhin sont les départements avec le plus d'établissements du Crédit Mutuel avec respectivement 10,20 % et 7,58 % du nombre total d'agences Crédit Mutuel.

La Société Générale

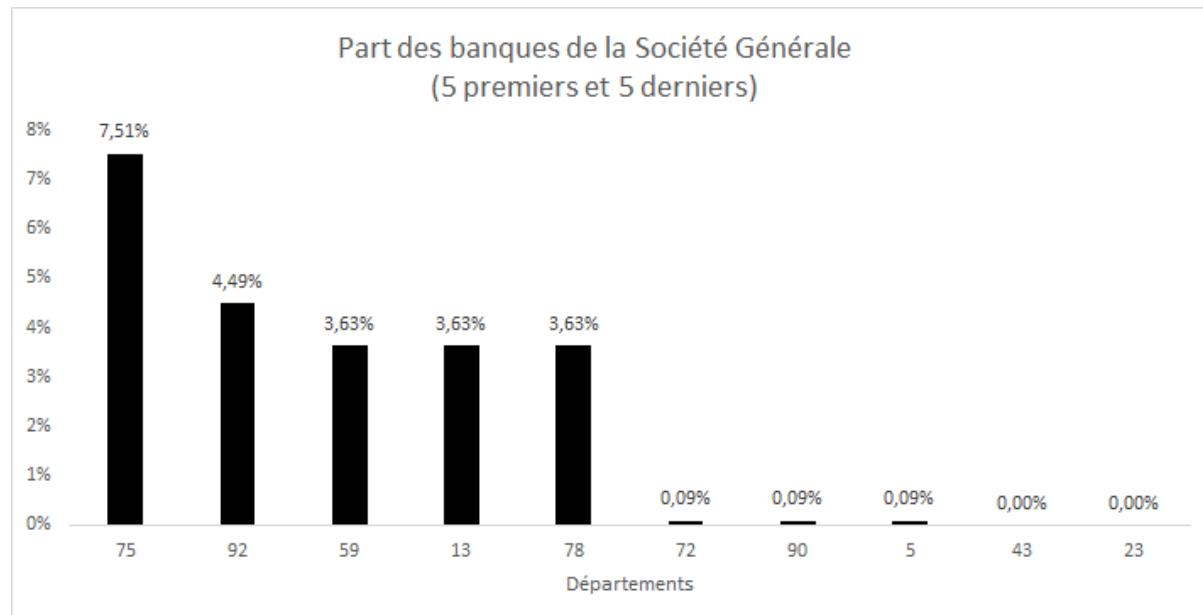


Figure 14 : Part des agences de la Société Générale selon les départements (5 premiers et 5 derniers)

Nous pouvons observer sur la figure 14 qu'il y a deux départements où il n'y a pas de banque de la Société Générale, ce sont ceux de la Creuse et la Haute-Loire. Quant aux départements de Paris et du Hauts-de-Seine, ce sont ceux qui ont le plus d'agences de la Société Générale avec 7,51 % pour Paris et 4,49 % pour le Hauts-de-Seine.

À l'échelle des zones d'emploi

Le Crédit Agricole

Tableau 8 : Répartition des différentes agences du Crédit Agricole par zone d'emploi en France

		Crédit Agricole																		
Zone d'Emploi	Effectifs	7625	8421	2405	2402	7621	7517	2413	2410	2706	8408	7505	5216	7524	9315	5315	7620	8409	4423	3216
Effectifs	147	144	105	83	73	68	65	64	57	57	57	53	51	51	50	49	49	48	47	43
Fréquence	2,9%	2,9%	2,1%	1,7%	1,5%	1,4%	1,3%	1,3%	1,1%	1,1%	1,1%	1,1%	1,0%	1,0%	1,0%	1,0%	1,0%	1,0%	0,9%	0,9%
Zone d'Emploi	7615	60	7602	2815	9312	7503	8431	2804	9304	5302	7611	7616	8428	3202	9318	8416	8433	4413	7610	2711
Effectifs	43	42	41	39	39	38	38	37	37	36	36	36	35	34	34	33	33	32	32	31
Fréquence	0,9%	0,8%	0,8%	0,8%	0,8%	0,8%	0,8%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,6%	0,6%
Zone d'Emploi	4412	7604	2713	7522	2401	2407	2718	7527	8401	1106	8419	5212	9301	5202	5313	2715	7512	7622	7624	4409
Effectifs	30	30	29	29	28	28	28	28	27	27	26	26	25	25	24	24	24	24	23	
Fréquence	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%
Zone d'Emploi	4425	7516	7609	8402	54	63	7501	7511	7519	7607	7618	56	4411	7510	8405	8406	53	1112	2403	2813
Effectifs	23	23	23	23	22	22	22	22	22	22	21	21	21	21	21	20	20	20	20	
Fréquence	0,5%	0,5%	0,5%	0,5%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%
Zone d'Emploi	5316	7502	7612	8420	2409	2714	3211	4410	5309	7520	8407	8435	1114	2408	2705	2805	4405	2415	2703	3218
Effectifs	20	20	20	20	19	19	19	19	19	19	19	19	18	18	18	18	18	17	17	
Fréquence	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,3%	0,3%
Zone d'Emploi	5210	7614	8432	3203	3213	3220	4407	5211	5318	7504	7509	8422	1104	1113	1115	2411	2412	2702	3205	4424
Effectifs	17	17	17	16	16	16	16	16	16	16	16	16	15	15	15	15	15	15	15	
Fréquence	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%
Zone d'Emploi	7513	7518	8403	51	64	2807	3212	3215	5219	5310	7506	7619	8413	8426	9311	1101	2712	2816	3217	7507
Effectifs	15	15	15	14	14	14	14	14	14	14	14	14	14	14	14	13	13	13	13	
Fréquence	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%
Zone d'Emploi	7523	8417	2708	3219	3222	5203	7528	7603	7608	7617	8415	9303	9310	52	57	59	1105	2716	4427	7514
Effectifs	13	13	12	12	12	12	12	12	12	12	12	12	11	11	11	11	11	11	11	
Fréquence	0,3%	0,3%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%
Zone d'Emploi	7526	7530	7531	8424	9308	1107	2809	2811	4404	4420	5206	5308	5317	7532	7613	8410	8423	9313	61	62
Effectifs	11	11	11	11	11	10	10	10	10	10	10	10	10	10	10	10	10	10	9	9
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%
Zone d'Emploi	2704	3206	3207	3210	4414	4421	4422	4426	5207	5213	5303	8414	8418	8429	58	1108	2707	2717	3208	3221
Effectifs	9	9	9	9	9	9	9	9	9	9	9	9	9	9	8	8	8	8	8	8
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%
Zone d'Emploi	4403	5304	7508	9305	9306	9307	9314	2406	2709	2802	2803	2817	3201	4406	5220	5301	7525	7606	8411	8434
Effectifs	8	8	8	8	8	8	8	7	7	7	7	7	7	7	7	7	7	7	7	
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%
Zone d'Emploi	55	2414	2710	2806	2810	2812	2814	2819	3214	4401	4402	4415	4417	4418	5214	5215	5306	7521	7601	8404
Effectifs	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%
Zone d'Emploi	8412	2701	2808	4416	4419	8427	8430	9309	9317	1103	4408	5208	5217	5311	5312	7515	7529	9316	1102	1110
Effectifs	6	5	5	5	5	5	5	5	5	5	4	4	4	4	4	4	4	4	3	3
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%
Zone d'Emploi	1111	2404	3204	5218	7623	8425	9302	2801	2818	3209	5201	5205	5209	5305	5307	5314	5319	5204	5221	7605
Effectifs	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2	1	1	
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%

Comme nous pouvons le remarquer dans le tableau 8 il existe treize zones d'emploi où il n'y a aucune agence du Crédit Agricole.

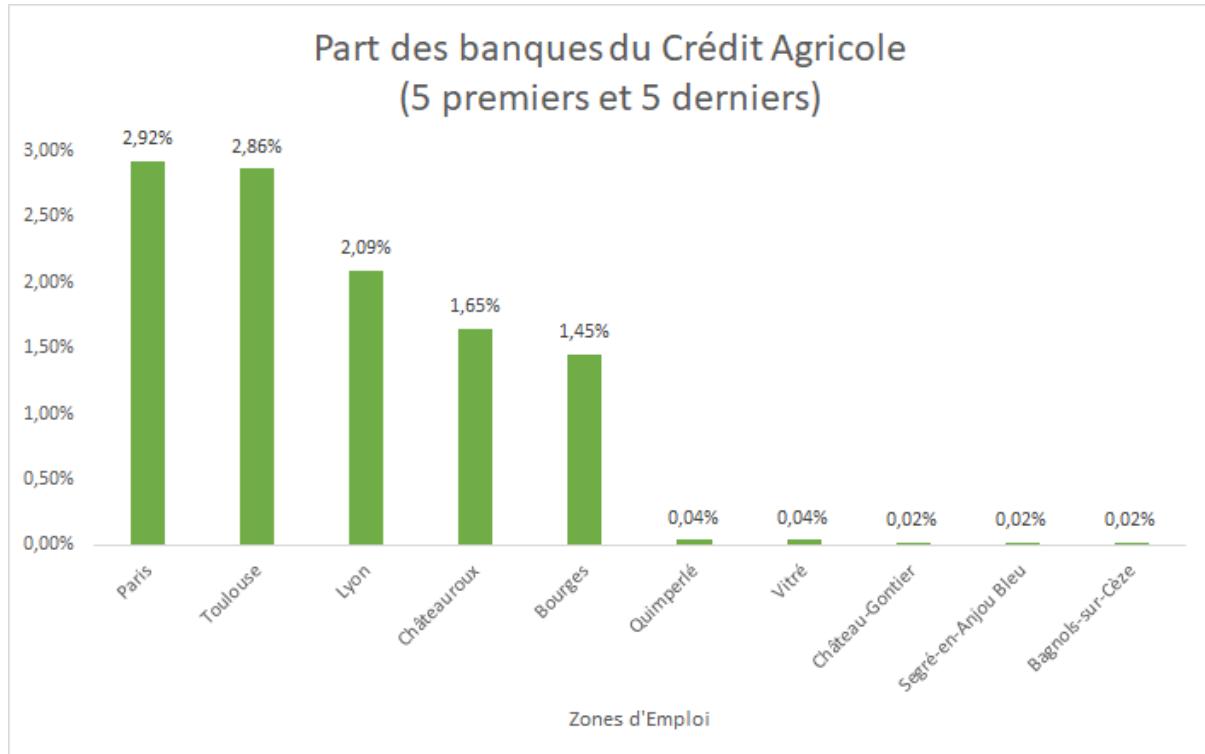


Figure 15 : Part des agences du Crédit Agricole selon les zones d'emploi (5 premiers et 5 derniers)

Sur la figure 15 ci-dessus, nous pouvons observer que les trois premières zones d'emploi avec le plus d'agences du Crédit Agricole sont Paris, Toulouse et Lyon avec respectivement 2,92 %, 2,86 % et 2,09 %. Remarquons qu'il s'agit de grands bassins de population avec plus de 400000 habitants.

La Banque Populaire

Tableau 9 : Répartition des différentes agences de la Banque Populaire par zone d'emploi en France

Banque Populaire																								
Zone d'Emploi	1109	7625	7616	8421	7620	7505	9312	3216	4423	5216	1115	9315	7618	8409	1114	5313	9318	1112	5315	2804				
Effectifs	223	64	48	46	45	34	34	30	30	30	29	24	23	21	20	18	18	17	17	16				
Fréquence	11,1%	3,2%	2,4%	2,3%	2,2%	1,7%	1,7%	1,5%	1,5%	1,5%	1,4%	1,2%	1,1%	1,0%	1,0%	0,9%	0,9%	0,9%	0,9%	0,8%				
Zone d'Emploi	4411	7624	8408	1104	2413	2706	5309	2711	3218	4410	5302	7503	7517	9304	1113	7604	53	2713	4412	7608				
Effectifs	16	16	16	15	15	15	15	14	14	14	14	14	14	14	13	13	12	12	12	12				
Fréquence	0,8%	0,8%	0,8%	0,8%	0,8%	0,8%	0,8%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,6%	0,6%	0,6%				
Zone d'Emploi	7610	8416	9301	1101	2815	7522	7524	7602	8401	8402	8431	1106	2702	4409	7617	2705	2715	3222	4405	4424				
Effectifs	12	12	12	11	11	11	11	11	11	11	11	11	10	10	10	10	9	9	9	9				
Fréquence	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%				
Zone d'Emploi	7512	7615	2718	4413	4425	5318	7507	7513	7601	7621	60	2712	2714	2813	5202	5212	5219	7607	7612	8405				
Effectifs	9	9	8	8	8	8	8	8	8	8	7	7	7	7	7	7	7	7	7					
Fréquence	0,5%	0,5%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%				
Zone d'Emploi	8426	9310	9313	1108	2405	2410	2708	3211	3213	3215	5210	5301	7509	7511	7518	7520	7623	8413	8414	8428				
Effectifs	7	7	7	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6					
Fréquence	0,4%	0,4%	0,4%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%				
Zone d'Emploi	1105	2402	2403	2707	2709	2716	2717	2817	3202	3203	3207	3212	3217	4407	4419	4420	5203	5214	5308	5317				
Effectifs	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5					
Fréquence	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%				
Zone d'Emploi	7504	7510	7516	7519	7523	7611	7619	8423	8432	8433	8435	51	52	2805	2811	4401	4414	5207	7501	7502				
Effectifs	5	5	5	5	5	5	5	5	5	5	5	5	4	4	4	4	4	4	4					
Fréquence	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	7514	7526	7531	7532	8406	8407	8415	8418	8419	9305	9306	9307	9308	9314	54	56	57	59	61	63				
Effectifs	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	3	3					
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	1111	2401	2406	2409	2704	2807	2816	3206	3208	3219	4408	4416	5217	5220	5310	5316	7506	7609	7614	8403				
Effectifs	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3					
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	8410	8420	8424	8434	9302	9317	64	1103	1110	2404	2408	2412	2703	2710	2803	2809	2814	2819	3201	3204				
Effectifs	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2					
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	3205	4404	4406	4415	4418	4421	4427	5206	5213	5303	5312	7521	7525	7527	7528	7530	7606	7613	8412	8417				
Effectifs	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2					
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	8422	9303	9311	9316	55	58	62	1102	1107	2407	2411	2414	2415	2701	2801	2802	2806	2808	2810	2812				
Effectifs	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1					
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	3210	3214	3220	3221	4422	4426	5201	5204	5211	5215	5221	5304	5305	5306	5307	5319	7515	7529	7603	7622				
Effectifs	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1					
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	8404	8411	8425	8430	2818	3209	4402	4403	4417	5205	5208	5209	5218	5311	5314	7508	7605	8427	8429	9309				
Effectifs	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0					
Fréquence	0,1%	0,1%	0,1%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%				

Comme nous pouvons le voir dans le tableau 9 lors de notre analyse, nous avons constaté seize zones d'emploi sans établissements de la Banque Populaire.

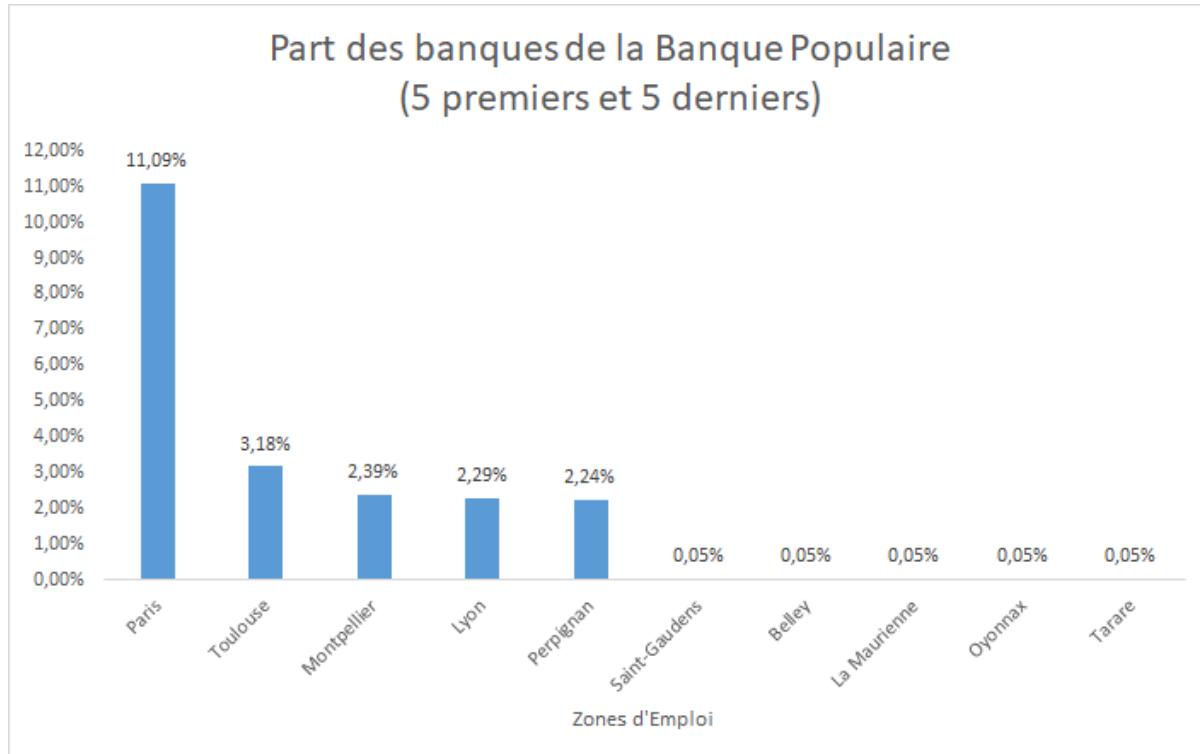


Figure 16 : Part des agences de la Banque Populaire selon les zones d'emploi (5 premiers et 5 derniers)

Nous pouvons observer sur la figure 16 ci-dessus les cinq premières zones d'emploi ayant le plus d'établissements de la Banque Populaire sont Paris, Toulouse, Montpellier, Lyon et Perpignan avec respectivement 11,09 %, 3,18 %, 2,39 %, 2,29 % et 2,24 %.

Le Crédit Mutuel

Tableau 10 : Répartition des différentes agences du Crédit Mutuel par zone d'emploi en France

Crédit Mutuel																								
Zone d'Emploi	4423	1109	4411	4409	4405	4424	3216	4420	7625	5216	8421	4418	4422	4419	2711	3213	2815	3218	5210	5212				
Effectifs	67	65	51	45	31	28	25	23	21	20	20	19	16	15	14	14	13	13	13	13				
Fréquence	5,0%	4,9%	3,8%	3,4%	2,3%	2,1%	1,9%	1,7%	1,6%	1,5%	1,5%	1,4%	1,2%	1,1%	1,1%	1,1%	1,0%	1,0%	1,0%	1,0%				
Zone d'Emploi	4408	5219	8416	4410	4412	5213	9312	2705	4416	8409	2714	2804	2805	3203	3205	3206	3217	7503	7520	7524				
Effectifs	12	12	12	11	11	11	10	10	10	9	9	9	9	9	9	9	9	9	9	9				
Fréquence	0,9%	0,9%	0,9%	0,8%	0,8%	0,8%	0,8%	0,8%	0,8%	0,8%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%	0,7%				
Zone d'Emploi	8401	1101	1112	2403	2706	2713	3202	3215	51	1104	1106	1114	1115	2410	2715	3212	3219	5203	5207	5211				
Effectifs	9	8	8	8	8	8	8	8	7	7	7	7	7	7	7	7	7	7	7	7				
Fréquence	0,7%	0,6%	0,6%	0,6%	0,6%	0,6%	0,6%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%				
Zone d'Emploi	7512	7517	9318	54	1113	2702	2709	2816	3208	5215	7616	9301	9304	9313	2413	2718	2811	3207	3211	3222				
Effectifs	7	7	7	6	6	6	6	6	6	6	6	6	6	6	5	5	5	5	5	5				
Fréquence	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,5%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%				
Zone d'Emploi	4413	4414	4425	5201	7506	7509	7526	8428	8431	8433	61	1105	2802	2806	2809	2810	2817	3204	5202	7602				
Effectifs	5	5	5	5	5	5	5	5	5	4	4	4	4	4	4	4	4	4	4	4				
Fréquence	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%				
Zone d'Emploi	8405	8407	8414	8415	8426	9315	57	60	2401	2404	2414	2704	2707	2708	2710	2819	4407	4421	5206	5208				
Effectifs	4	4	4	4	4	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3				
Fréquence	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	5214	7525	7528	7530	7608	7615	7624	8406	8408	8420	52	53	58	1108	2406	2407	2409	2411	2712	2717				
Effectifs	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	2	2	2				
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	2801	2803	2807	2808	2813	2814	2818	3220	3221	4401	4403	5205	5209	5217	5218	5220	7501	7511	7519	7522				
Effectifs	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2				
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	7604	7609	7610	7611	7612	7617	7618	7620	7621	8402	8410	8412	8413	8417	8418	8419	8423	8424	8430	8432				
Effectifs	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2				
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	8435	55	56	63	64	1102	1103	1110	1111	2402	2405	2408	2412	2415	2701	2703	2716	2812	3201	3210				
Effectifs	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1				
Fréquence	0,2%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	3214	4404	4415	4417	4426	4427	5204	7507	7510	7513	7518	7527	7531	7532	7601	7607	7613	7614	7623	8403				
Effectifs	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1				
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	8404	8422	8425	9302	9303	9305	9306	9307	9308	9310	9311	9314	9316	9317	59	62	1107	3209	4402	4406				
Effectifs	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	0	0	0				
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%				
Zone d'Emploi	5221	5301	5302	5303	5304	5305	5306	5307	5308	5309	5310	5311	5312	5313	5314	5315	5316	5317	5318	5319				
Effectifs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
Fréquence	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%				
Zone d'Emploi	7502	7504	7505	7508	7514	7515	7516	7521	7523	7529	7603	7605	7606	7619	7622	8411	8427	8429	8434	9309				
Effectifs	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
Fréquence	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%				

Nous pouvons observer sur le tableau 10 ci-dessus qu'il y a 46 zones d'emploi où il n'y a pas d'établissements du Crédit Mutuel.

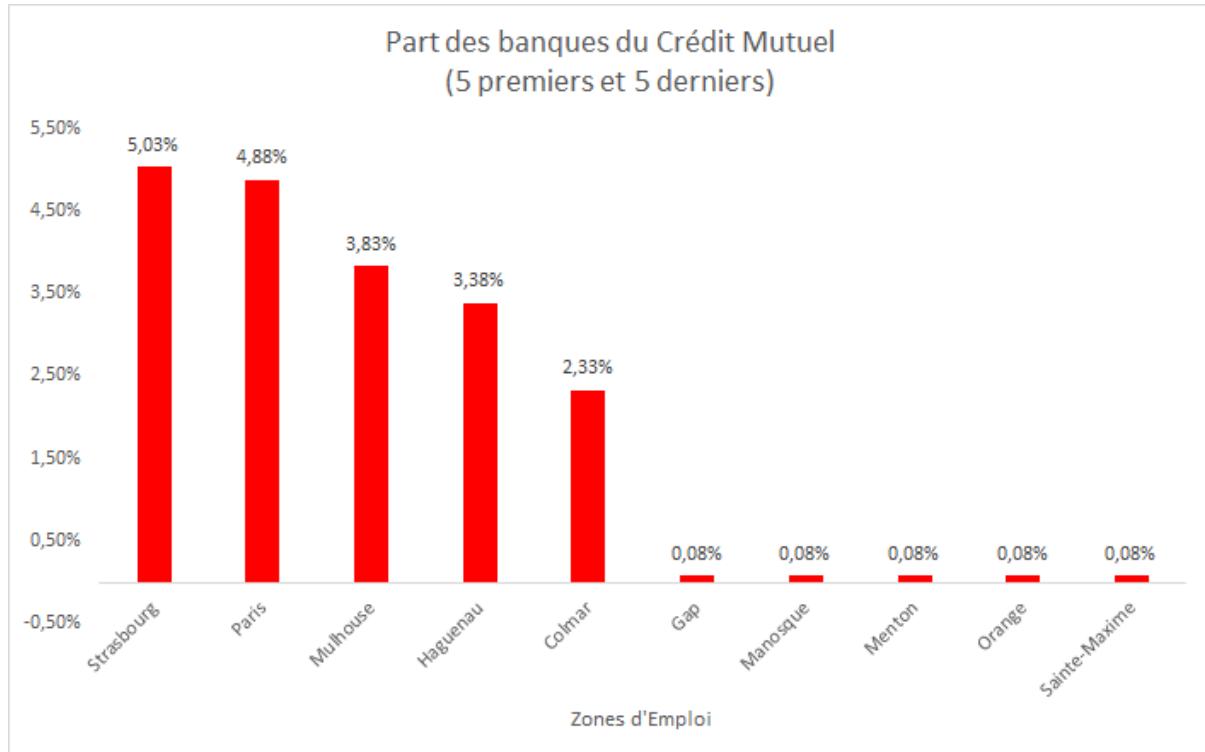


Figure 17 : Part des agences du Crédit Mutuel selon les zones d'emploi (5 premiers et 5 derniers)

Nous pouvons remarquer sur la figure 17 ci-dessus que les zones d'emploi de Strasbourg, Paris, Mulhouse, Haguenau et Colmar sont les zones d'emploi où le Crédit Mutuel est le plus implanté avec respectivement 5,03 %, 4,88 %, 3,83 %, 3,38 % et 2,33 % du nombre total d'agences du Crédit Mutuel.

La Société Générale

Tableau 11 : Répartition des différentes agences de la Société Générale par zone d'emploi en France

Société Générale																								
Zone d'Emploi	1109	8421	9312	7625	1114	7505	9315	1112	1115	4423	1101	1113	7616	2804	3216	5216	1104	2815	7503	5315				
Effectif	198	38	26	25	24	22	21	20	18	18	16	16	16	15	15	14	13	13	13	12				
Fréquence	17,1%	3,3%	2,3%	2,2%	2,1%	1,9%	1,8%	1,7%	1,6%	1,6%	1,4%	1,4%	1,4%	1,3%	1,3%	1,2%	1,1%	1,1%	1,1%	1,0%				
Zone d'Emploi	8409	9304	3218	4410	4411	7522	9301	9318	1106	3202	3211	4412	4413	8408	2711	4424	7524	7602	8416	1105				
Effectif	12	12	10	10	10	10	10	9	9	9	9	9	9	8	8	7	7	7	6					
Fréquence	1,0%	1,0%	0,9%	0,9%	0,9%	0,9%	0,9%	0,9%	0,8%	0,8%	0,8%	0,8%	0,8%	0,8%	0,7%	0,7%	0,6%	0,6%	0,6%	0,5%				
Zone d'Emploi	1108	2413	7620	53	3213	5202	5302	5309	5313	7618	8431	9313	52	57	2409	2706	2807	2813	2817	3217				
Effectif	6	6	6	5	5	5	5	5	5	5	5	5	5	4	4	4	4	4	4	4				
Fréquence	0,5%	0,5%	0,5%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%				
Zone d'Emploi	4407	4420	5318	7513	7517	7610	7612	7624	8401	8428	8433	9306	9308	54	60	1103	1107	2407	2410	2702				
Effectif	4	4	4	4	4	4	4	4	4	4	4	4	4	3	3	3	3	3	3	3				
Fréquence	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,4%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%				
Zone d'Emploi	2708	2713	2805	2811	2814	2819	3210	3212	3220	4405	4409	5308	7501	7509	7510	7516	7604	7608	7615	8405				
Effectif	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3	3				
Fréquence	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%	0,3%				
Zone d'Emploi	8432	8435	58	59	63	64	2401	2402	2403	2405	2705	2707	2715	2803	2808	3203	3204	3206	3208	3222				
Effectif	3	3	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2				
Fréquence	0,3%	0,3%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	4404	4408	4416	4419	4421	4425	5203	5207	5219	5301	7506	7507	7518	7520	7523	7526	7528	7601	7609	7613				
Effectif	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2				
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%				
Zone d'Emploi	7617	7621	8402	8412	8413	8414	8417	8423	8434	9314	51	55	56	61	62	1102	1110	1111	2404	2406				
Effectif	2	2	2	2	2	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1				
Fréquence	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,2%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	2408	2411	2412	2414	2415	2701	2703	2704	2710	2712	2714	2717	2718	2801	2802	2806	2809	2810	2812	3201				
Effectif	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1				
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	3205	3207	3214	3215	3219	3221	4401	4403	4414	4415	4417	4418	4426	5206	5210	5211	5212	5215	5217	5220				
Effectif	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1				
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	5304	5306	5307	5310	5311	5312	5316	5317	7502	7504	7512	7514	7515	7519	7521	7525	7530	7531	7532	7603				
Effectif	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1				
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	7607	7611	7614	7619	7623	8403	8404	8407	8410	8411	8415	8418	8422	8424	8425	8426	8430	9303	9305	9307				
Effectif	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1				
Fréquence	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%	0,1%				
Zone d'Emploi	9310	9311	9316	9317	2709	2716	2816	2818	3209	4402	4406	4422	4427	5201	5204	5205	5208	5209	5213	5214				
Effectif	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
Fréquence	0,1%	0,1%	0,1%	0,1%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%				
Zone d'Emploi	5218	5221	5303	5305	5314	5319	7508	7511	7527	7529	7605	7606	7622	8406	8419	8420	8427	8429	9302	9309				
Effectif	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0				
Fréquence	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%				

Dans le tableau 11 ci-dessus, nous pouvons observer 36 zones d'emploi sans aucune agence de la Société Générale.

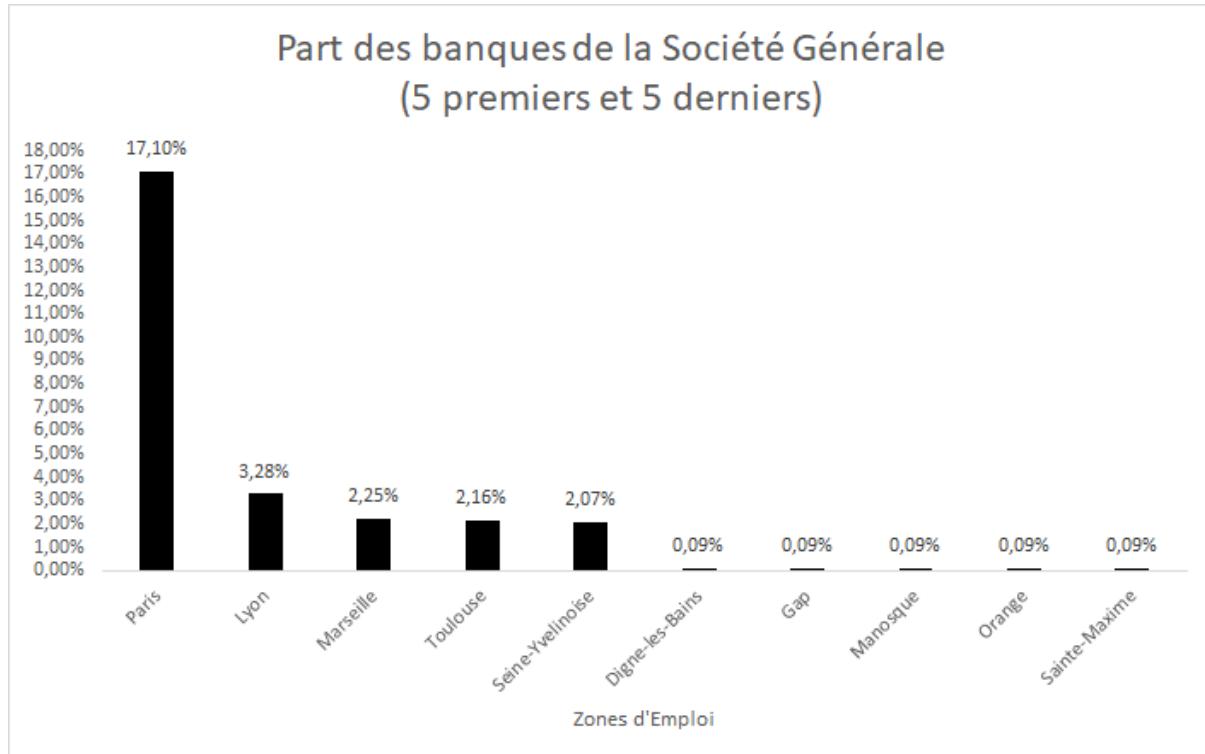


Figure 18 : Part des agences de la Société Générale selon les zones d'emploi (5 premiers et 5 derniers)

Dans la figure 18 ci-dessus, nous pouvons remarquer que presque deux banques sur dix de la Société Générale sont implantées dans la zone d'emploi de Paris. Les zones d'emploi de Lyon, Marseille et Toulouse arrivent en deuxième, troisième et quatrième position avec 3,28 %, 2,25 % et 2,16 % du nombre total d'agences.

III. Les différences entre les banques coopératives et commerciales

Nous pouvons désormais étudier les différents liens entre les variables socio-économiques et les différentes banques afin de voir s'il existe des corrélations entre certaines variables socio-économiques et le type de banque.

Statistiques descriptives

Le salaire net horaire moyen

De par nature, les banques coopératives ont été créées pour faciliter l'accès au crédit pour les personnes les plus défavorisées. Nous pouvons donc penser que les banques coopératives sont proportionnellement plus présentes dans les zones d'emploi où les salaires sont les plus faibles. Inversement nous pouvons penser que les banques commerciales sont plus présentes dans les zones d'emploi où les salaires sont les plus élevés.

Tableau 12 : Présence de chaque banque par zone d'emploi en fonction du salaire net horaire moyen dans cette zone

	Part de chaque banque			
	Banque Coopérative :			Banque Commerciale :
Salaire net horaire moyen	Crédit Agricole	Banque Populaire	Crédit Mutuel	Société Générale
Moins de 13€	38,93%	29,09%	31,41%	18,49%
Entre 13€ et 16€	55,94%	54,98%	60,90%	56,09%
Entre 16€ et 19€	1,84%	2,55%	2,33%	4,92%
Plus de 19€	3,64%	13,68%	6,02%	20,81%
Total	100%	100%	100%	100%

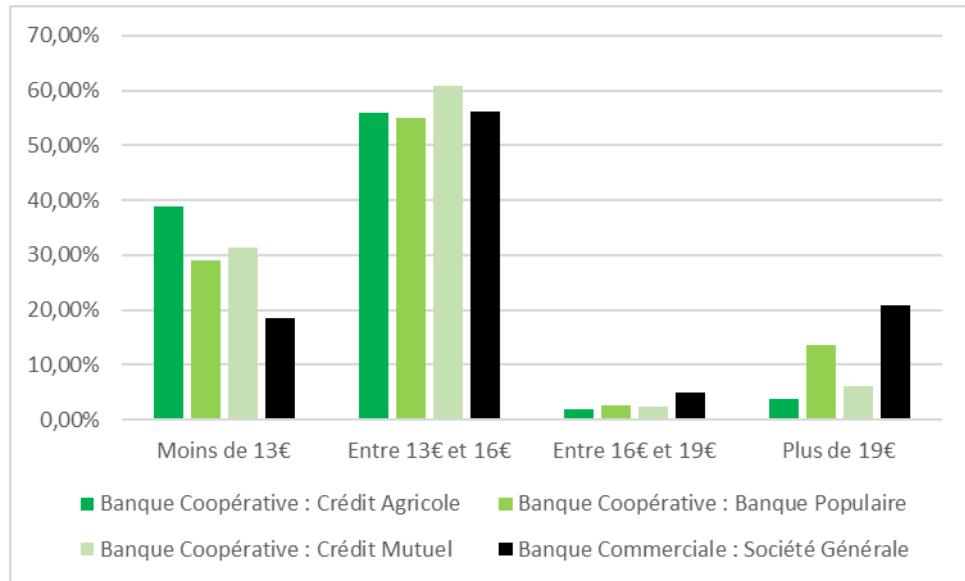


Figure 19 : Présence de chaque banque par zone d'emploi en fonction du salaire net horaire moyen dans cette zone

Tableau 13 : Coefficient de corrélation entre le salaire et le nombre de banques par zone d'emploi pour chaque banque

	Crédit Agricole	Banque Populaire	Crédit Mutuel	Société Générale
Coefficient de corrélation entre le salaire et le nombre de banques par zone d'emploi	0,3430	0,4769	0,3334	0,5097

Comme nous pouvons le voir sur le tableau 12 et la figure 19, les banques coopératives sont proportionnellement plus présentes que les banques commerciales dans les zones d'emploi où le salaire est inférieur à 13 € /h (en moyenne 33,1 % contre 18,5 %). De plus, les banques coopératives sont proportionnellement moins présentes que les banques commerciales dans les zones d'emploi où le salaire est supérieur à 19 €/h (en moyenne 7,8 % contre 20,1 %). Nos hypothèses de départ sont donc validées. Nous pouvons cependant voir que la banque coopérative Banque Populaire se rapproche un peu du modèle de la banque commerciale avec une plus forte présence que les autres banques coopératives dans les zones d'emploi où le salaire dépasse 19 €/h (13,7 % contre 3,6 % et 6 %) ainsi qu'une plus faible présence dans les zones d'emploi où le salaire est inférieur à 13€ /h (29,1 % contre 38,9 % et 31,4 %). Nous pouvons notamment observer ce rapprochement grâce aux données du tableau 13 car on observe que le coefficient de corrélation entre le salaire et le nombre de banques par zone d'emploi de la Banque Populaire se rapproche de celui de la banque commerciale Société Générale (0,48 contre 0,51) alors que celui des deux autres banques coopératives est beaucoup plus faible (0,34 pour le Crédit Agricole et 0,33 pour le Crédit Mutuel).

La part d'agriculteurs

De la même manière que pour les salaires nous pouvons donc penser que les banques coopératives sont proportionnellement plus présentes dans les zones d'emploi où la part d'agriculteurs est la plus forte. Inversement nous pouvons penser que les banques commerciales sont plus présentes dans les zones d'emploi où la part d'agriculteurs est la plus faible.

Tableau 14 : Présence de chaque banque par zone d'emploi en fonction de la part d'agriculteurs dans cette zone

Part d'agriculteurs	Part de chaque banque			
	Banque Coopérative :		Banque Commerciale :	
	Crédit Agricole	Banque Populaire	Crédit Mutuel	Société Générale
Moins de 1%	56,41%	71,60%	73,68%	78,80%
Entre 1% et 2%	27,89%	19,95%	17,16%	16,40%
Entre 2% et 3%	9,22%	5,30%	7,30%	3,28%
Entre 3% et 4%	4,22%	2,50%	2,21%	1,23%
Entre 4% et 5%	1,85%	0,70%	0,23%	0,43%
Plus de 5%	0,76%	0,25%	0,08%	0,17%
Total	100%	100%	100%	100%

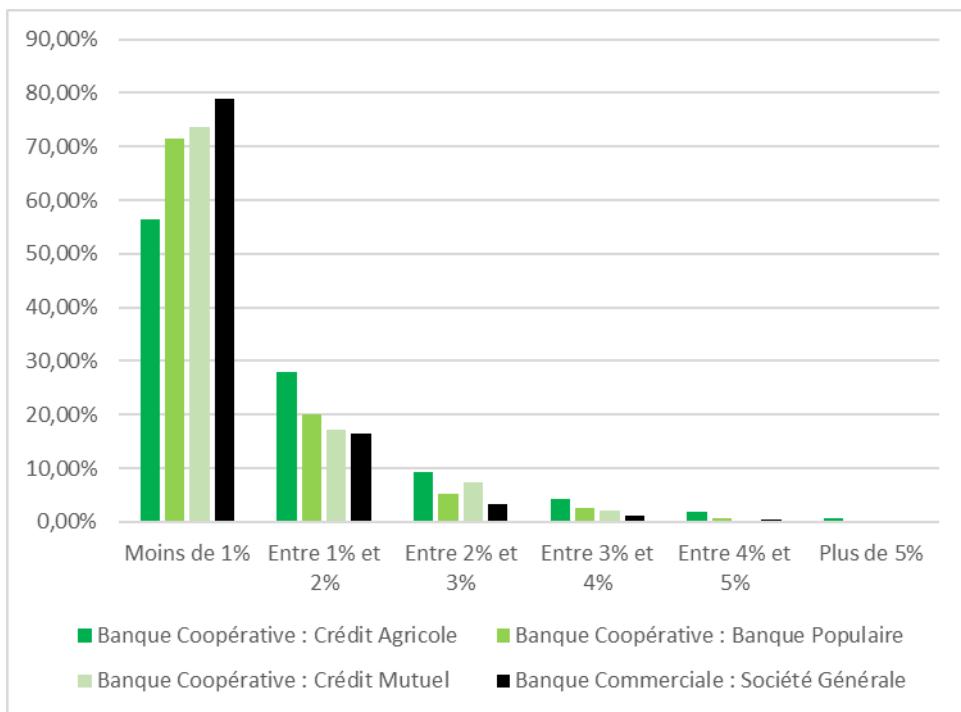


Figure 20 : Présence de chaque banque par zone d'emploi en fonction de la part d'agriculteurs dans cette zone

Comme nous pouvons le voir sur le tableau 14 et la figure 20, les banques commerciales sont proportionnellement plus présentes que les banques coopératives dans les zones d'emploi où la proportion d'agriculteurs est inférieure à 1 % (78,8 % contre en moyenne 67,2 %). De plus on peut voir que pour tous les intervalles au-dessus de la moyenne française d'agriculteurs (autour de 1 %), les banques coopératives sont proportionnellement plus présentes que les banques commerciales. Cela confirme donc nos hypothèses de départ.

Le taux de chômage

Nous pouvons donc maintenant nous interroger sur le lien entre la présence de banques coopératives et le taux de chômage car nous pouvons penser que les banques coopératives sont proportionnellement plus présentes dans les zones d'emploi où le taux de chômage est élevé. Inversement nous pouvons penser que les banques commerciales sont plus présentes dans les zones d'emploi où le taux de chômage est faible.

Tableau 15 : Présence de chaque banque par zone d'emploi en fonction du taux de chômage dans cette zone

Taux de chômage	Part de chaque banque			
	Banque Coopérative :			Banque Commerciale :
	Crédit Agricole	Banque Populaire	Crédit Mutuel	Société Générale
Moins de 6%	3,61%	3,29%	3,04%	2,85%
Entre 6% et 9%	51,85%	52,20%	57,53%	56,30%
Entre 9% et 12%	37,30%	32,15%	32,46%	32,50%
Entre 12% et 15%	7,03%	11,76%	6,49%	8,05%
Plus de 15%	0,56%	0,90%	1,14%	0,61%
Total	100%	100%	100%	100%

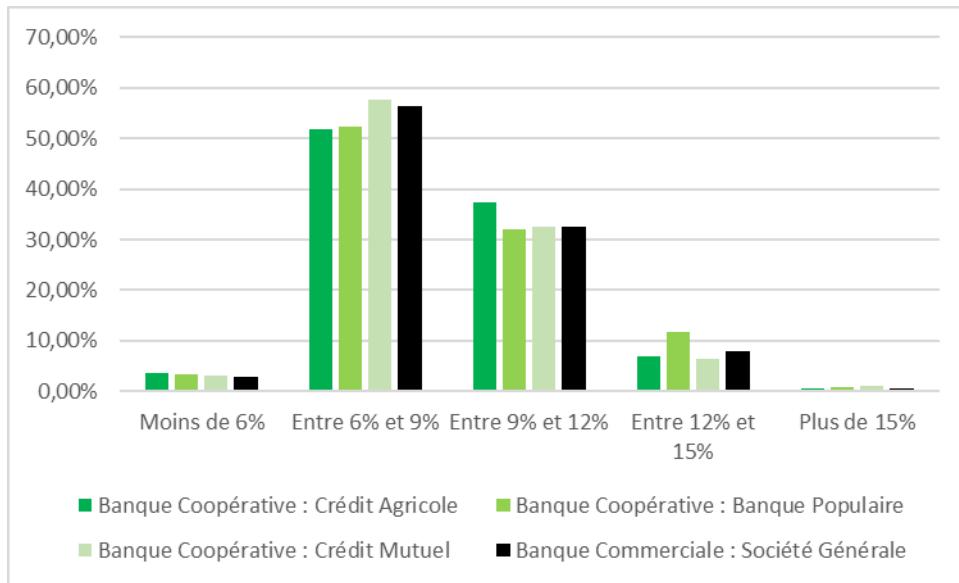


Figure 21 : Présence de chaque banque par zone d'emploi en fonction du taux de chômage dans cette zone

Comme nous pouvons le voir dans le tableau 15 et la figure 21, il ne semble y avoir aucune différence notable en ce qui concerne l'implantation des banques coopératives et commerciales par rapport au taux de chômage. En effet, la répartition des banques en fonction du taux de chômage semble être approximativement la même pour toutes les banques et tous les types de banques. Nos hypothèses de départ semblent donc erronées.

Analyse en Composante Principale

Dans le but de déterminer si certaines variables socio-économiques ont un impact sur le nombre de banques implantées dans chaque zone d'emploi, nous avons réalisé une Analyse en Composante Principale.

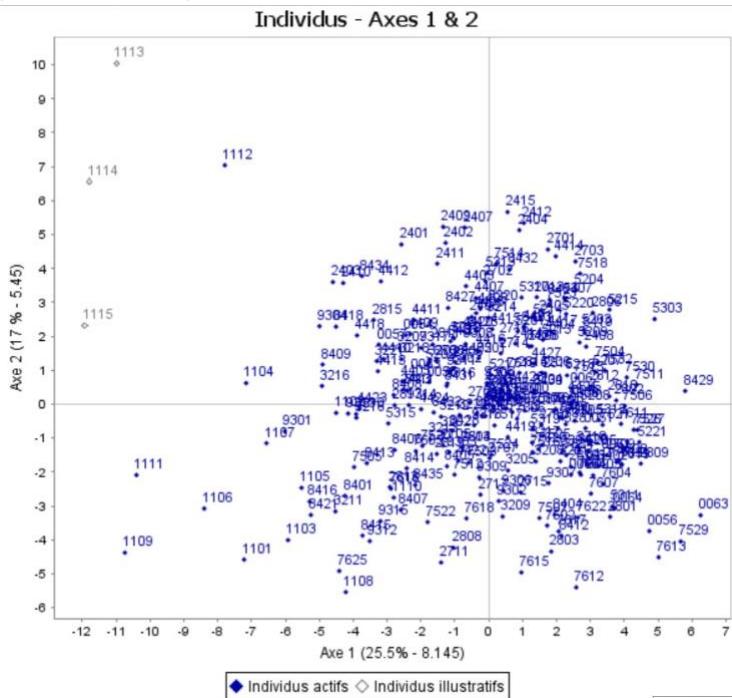


Figure 22 : Projection des individus sans les individus atypiques

La figure 22 nous montre la projection des individus actifs sur les deux premiers axes. Nous pouvons remarquer qu'ils ne sont pas distribués selon une structure particulière. Nous ne pouvons donc pas interpréter les axes par simple lecture graphique et nous avons dû pour cela regarder dans les sorties Excel.

Choix du nombre d'axes

La première étape, lorsque nous devons réaliser une ACP, est de sélectionner les axes que nous allons interpréter par la suite. Pour choisir le nombre d'axes que nous allons prendre en compte, nous nous basons sur plusieurs critères. Il faut, en effet, que la part d'information restituée par nos axes soit suffisamment importante et pour déterminer cela, nous pouvons essayer d'identifier un décrochement graphique notable sur la courbe de décroissance liée à la variance des axes. Il est également préférable que les valeurs propres associées aux axes retenus soient supérieures à 1.

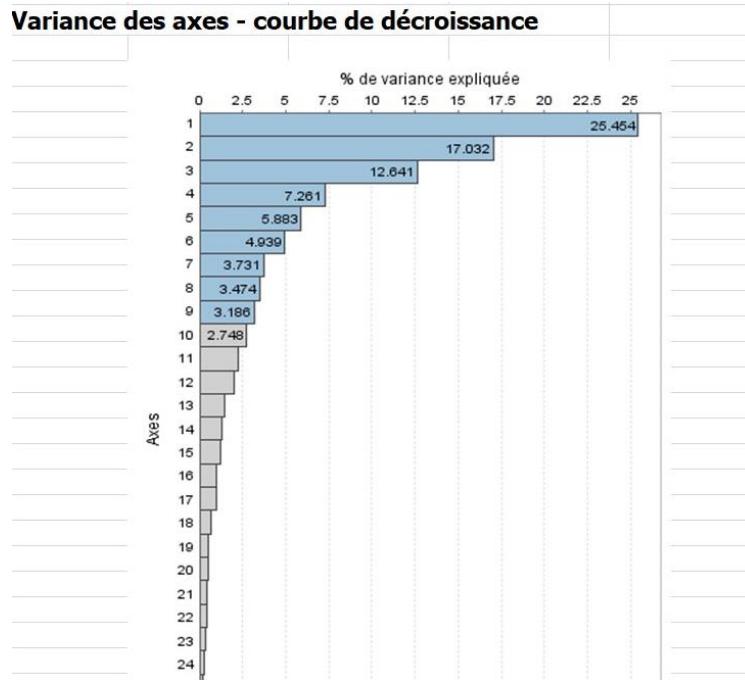


Figure 23 : Graphique du pourcentage de l'inertie portée par chaque axe

Nous remarquons donc, sur la figure 23, un décrochement entre la part d'inertie restituée par l'axe 3 et par l'axe 4. Nous avons donc choisi de retenir les trois premiers axes, qui, à eux seuls, restituent 55,1 % de l'information contenue dans nos données. Ce chiffre peut paraître insuffisant, mais en tenant compte du nombre important de variables que nous étudions, nous avons jugé que cette part était convenable.

Tableau 16 : Décomposition de la variance du nuage de point

Variance du nuage (trace de la matrice) : 32				
Axe	Variance de l'axe (valeur propre)	% de variance expliquée	% de variance expliquée cumulé	Retenu selon le critère de Kaiser
1	8,145	25,5	25,5	X
2	5,450	17,0	42,5	X
3	4,045	12,6	55,1	X

De plus, comme nous pouvons le constater dans le tableau 16, les valeurs propres associées aux axes 1, 2 et 3 sont supérieures à 1, ce qui nous conforte dans notre choix.

Interprétation des axes

Une fois les axes choisis, il nous a été nécessaire de les interpréter et d'identifier les variables responsables de leur construction.

L'axe 1 est donc à 50 % expliqué par les variables traitant de la part de la population dans les catégories socio-professionnelles des cadres et professions intellectuelles supérieures et des professions intermédiaires ainsi que par les variables des salaires horaires nets moyens pour ces CSP. Nous remarquons également que la variable qui retranscrit la part de personnes retraitées dans les zones d'emploi contribue assez significativement à la construction de l'axe 1. Celui-ci semble donc opposer les zones d'emploi comptant beaucoup de cadres et de professionnels intermédiaires, à celles qui ont une part importante de leur population qui est retraitée.

En ce qui concerne l'axe 2, les variables qui recensent la part de la population dans les différentes classes d'âges semblent être responsables à 84 % de sa construction. On peut donc dire que ce dernier oppose les zones d'emploi ayant une population majoritairement en dessous de la cinquantaine et celles qui ont une population plus vieillissante.

Environ 60 % de l'axe 3 est expliqué par les variables retranchant la part de la population active, le taux de chômage et le rapport interdécile.

En interprétant nos trois axes, nous nous sommes rendu compte que beaucoup de nos variables semblaient liées. En effet, le fait que les axes soient expliqués par beaucoup de variables dans une proportion presque égale nous permet de supposer que celles-ci ne sont pas indépendantes. Cela est expliqué en partie par le fait que beaucoup de nos variables sont des proportions. De ce fait, il est normal que les valeurs de certaines variables influencent les autres et nous allons pouvoir vérifier cela à l'aide du cercle des corrélations.

Corrélation entre les variables

Le cercle des corrélations nous offre une représentation graphique des variables actives et illustratives de notre analyse. Elles sont représentées par des flèches plus ou moins grandes et plus ou moins éloignées les unes des autres. Plus une flèche est proche du cercle de centre 0 et de rayon 1, plus la variable associée à cette flèche est bien représentée par l'analyse.

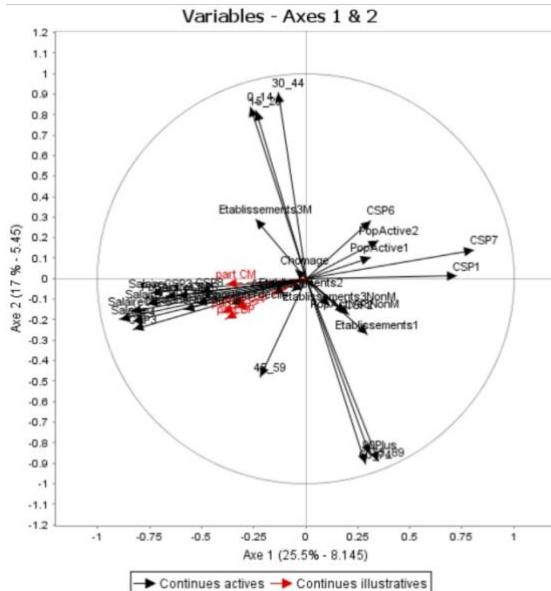


Figure 24 : Cercle des corrélations

Nous pouvons donc voir sur la figure 24 que les variables qui traitent de la part de la population appartenant aux différentes CSP et celles qui retrouvent le salaire horaire net moyen pour celles-ci, sont bien représentées par l'analyse. Il en est de même pour les variables qui présentent la répartition de la population en classes d'âges. Au contraire, les variables telles que les parts de la population active dans les secteurs primaire, secondaire et tertiaire ou encore la part du nombre d'établissements dans ces mêmes secteurs semblent assez mal représentés par notre ACP. Cela est cohérent avec l'interprétation des axes fournis précédemment, puisque nous pouvons remarquer que les variables les mieux représentées sont celles qui sont responsables de la construction des trois premiers axes. Grâce au cercle des corrélations, nous pouvons également identifier des structures de liaisons entre les variables. En effet, l'angle entre les projections de deux variables s'apparente au coefficient de corrélation qui existe entre celles-ci. Nous pouvons donc interpréter cela graphiquement, en disant que plus l'angle entre les projections de deux variables est aigu, plus le coefficient de corrélation entre ces deux variables sera élevé et donc plus elles seront liées positivement. À l'inverse, lorsque l'angle est obtus, nous pourrons dire que les variables sont corrélées négativement. Par exemple, nous pouvons voir que la part de la population étant cadre est corrélée positivement avec la variable du salaire net horaire moyen. Autrement dit, plus il y a de cadres (CSP3) dans une zone d'emploi, plus le salaire net horaire moyen de celle-ci sera élevé. Nous pouvons vérifier cela numériquement en regardant les sorties Excel dans lesquelles on peut retrouver les coefficients de corrélation entre nos variables. Le coefficient de corrélation entre les variables du salaire net horaire moyen et de la part de population travaillant dans la CSP3 est de 0,895, ce qui confirme qu'il existe une liaison très forte entre les deux. Pour identifier les variables qui semblent avoir un impact sur le nombre de banques implantées dans chaque zone d'emploi, nous pouvons regarder plus précisément l'endroit du graphique où sont projetées les variables illustratives (représentées en rouge sur le graphique).

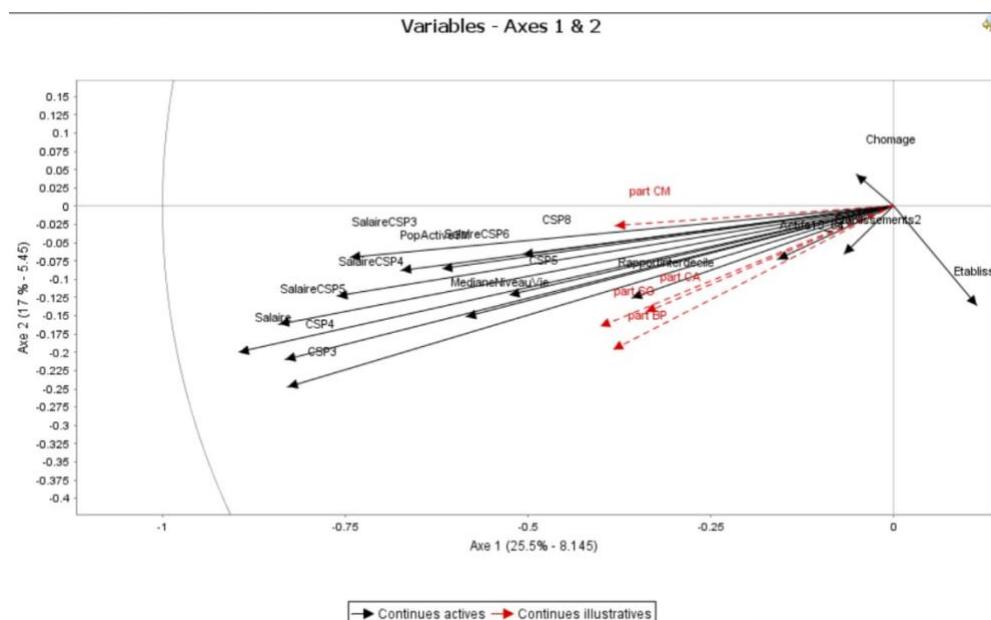


Figure 25 : Zoom sur une petite partie du cercle des corrélations

Nous pouvons voir clairement sur la figure 25 que les variables des parts des différentes banques par zones d'emploi sont très liées aux variables représentant les salaires et les CSP. Nous remarquons cependant que les variables qui présentent la répartition de la population en classes d'âge semblent très peu liées au nombre d'agences implantées dans une zone d'emploi.

Par ailleurs, il est utile de souligner que ces affirmations manquent de précision. En effet, nous pouvons voir que les extrémités des flèches qui représentent les variables de la répartition des différentes banques sont éloignées du cercle des corrélations. Cela signifie qu'elles ne sont pas représentées de manière optimale par notre ACP.
Nous allons donc préciser notre analyse en effectuant de multiples régressions linéaires.

Régressions linéaires multiples

Nous allons donc désormais étudier les différents liens entre les variables socio-économiques et le nombre d'agences par zone d'emploi pour chaque banque afin de voir si les banques coopératives ont toutes un profil similaire et si la banque commerciale correspond à un autre profil.

Le Crédit Agricole

Au vu de nos précédents résultats, comme le Crédit Agricole est une banque coopérative, nous pouvons penser que le nombre de ses agences va être plus grand dans les zones d'emploi où le salaire net horaire moyen est plus faible.

Tableau 17 : Tableau des variables explicatives du nombre d'agence Crédit Agricole par zone d'emploi

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	8.721e+01	2.894e+01	3.014	0.00283	**
Actifs15_64	2.254e+01	4.918e+00	4.584	7.09e-06	***
Chomage	-7.091e+00	3.912e+00	-1.812	0.07106	.
CSP1	-3.769e+01	7.795e+00	-4.835	2.28e-06	***
CSP2	-4.690e+01	7.160e+00	-6.550	3.04e-10	***
CSP4	-2.093e+01	6.757e+00	-3.097	0.00217	**
CSP5	-3.385e+01	6.063e+00	-5.583	5.91e-08	***
CSP6	-3.340e+01	4.246e+00	-7.867	9.67e-14	***
CSP7	-1.363e+01	3.035e+00	-4.490	1.07e-05	***
X0_14	-7.891e+01	2.893e+01	-2.728	0.00680	**
X15_29	-7.953e+01	2.903e+01	-2.740	0.00657	**
X30_44	-6.725e+01	2.849e+01	-2.360	0.01900	*
X45_59	-7.848e+01	2.866e+01	-2.738	0.00661	**
X60_74	-7.439e+01	2.839e+01	-2.620	0.00931	**
X75_89	-9.214e+01	3.502e+01	-2.631	0.00902	**
Salaire	-5.480e-01	1.187e-01	-4.619	6.07e-06	***
SalaireCSP4	2.120e-01	1.307e-01	1.623	0.10586	
MedianeNiveauVie	-6.836e-05	3.774e-05	-1.812	0.07121	.
RapportInterdecile	8.618e-02	6.099e-02	1.413	0.15881	
<hr/>					
Signif. codes:	0 ***	0.001 **	0.01 *	0.05 .	0.1 ' ' 1

Nous pouvons bien observer sur le tableau 17 que le salaire net horaire moyen est une variable explicative du nombre de banque Crédit Agricole par zone d'emploi car sa p-valeur est inférieure à 5 %. De plus sa t-valeur est négative, on peut donc en déduire que plus le salaire net horaire moyen est faible, plus il y aura en proportion d'agences du Crédit Agricole.

Nous pouvons également observer sur ce tableau que le taux de chômage n'est pas une variable explicative car sa p-valeur est supérieure à 5 %. Nous pouvons donc en déduire que le taux de chômage d'une zone d'emploi n'a pas d'incidence sur le nombre d'agences du Crédit Agricole dans cette zone. Enfin nous pouvons remarquer que la part de la population active de 15 ans à 64 ans est également une variable explicative du nombre d'agences Crédit Agricole par zone d'emploi et plus cette part est élevée plus le nombre d'agences est élevé. Nous remarquons également que les catégories socio-professionnelles influent toutes de la même manière, sans différence. Nous avons donc choisi d'écartier ce type de variable de notre analyse.

La Banque Populaire

Au vu de nos précédents résultats, nous pouvons penser que le nombre d'agences Banque Populaire ne va pas être plus grand dans les zones d'emploi où le salaire net horaire moyen est plus faible car nous nous sommes rendu compte que la Banque Populaire se rapprochait du profil des banques commerciales en ce qui concerne le taux d'implantation en fonction du salaire net horaire moyen.

Tableau 18 : Tableau des variables explicatives du nombre d'agences Banque Populaire par zone d'emploi

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	113.50042	27.94292	4.062	6.44e-05	***
PopActive3M	1.80860	0.79954	2.262	0.024517	*
Actifs15_64	10.18062	3.57902	2.845	0.004800	**
Chomage	-5.65255	3.51227	-1.609	0.108743	
CSP1	-28.61160	5.87642	-4.869	1.95e-06	***
CSP2	-26.69448	6.21429	-4.296	2.46e-05	***
CSP4	-13.76903	5.46175	-2.521	0.012299	*
CSP5	-15.74244	3.45977	-4.550	8.22e-06	***
CSP6	-20.33154	2.43250	-8.358	3.83e-15	***
CSP7	-10.06903	2.22220	-4.531	8.94e-06	***
X0_14	-100.10737	27.74463	-3.608	0.000370	***
X15_29	-101.69823	27.88875	-3.647	0.000321	***
X30_44	-92.64192	27.46141	-3.374	0.000855	***
X45_59	-100.64320	27.48548	-3.662	0.000303	***
X60_74	-101.42782	27.30064	-3.715	0.000248	***
X75_89	-110.43873	33.65913	-3.281	0.001175	**
SalairCSP4	-0.25204	0.08555	-2.946	0.003509	**
SalairCSP5	-0.55643	0.20349	-2.734	0.006676	**
RapportInterdecile	0.08655	0.05790	1.495	0.136202	

Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’	0.1 ‘ ’ 1

La variable Salaire n'étant pas présente dans le tableau 18, on peut donc en déduire que le salaire net horaire moyen n'a pas été conservé dans le modèle pour la Banque Populaire. On peut donc en déduire que ce n'est pas une variable qui influe sur la proportion du nombre d'agences Banque Populaire dans les zones d'emploi. Cependant nous pouvons voir que comme pour le Crédit Agricole la part de la population active de 15 ans à 64 ans est une variable explicative avec une t-valeur positive. Nous pouvons donc en déduire ici aussi

que plus cette part est élevée, plus le nombre d'agences est élevé. Cela est également le cas pour la part de la population active dans le secteur tertiaire marchand.

Le Crédit Mutuel

Au vu de nos précédents résultats, comme le Crédit Mutuel est une banque coopérative, nous pouvons penser que le nombre de ses agences va être plus grand dans les zones d'emploi où le salaire net horaire moyen est plus faible.

Tableau 19 : Tableau des variables explicatives du nombre d'agences Crédit Mutuel par zone d'emploi

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	18.59807	3.70703	5.017	9.67e-07	***
PopActive3M	2.86433	0.94283	3.038	0.002621	**
Actifs15_64	9.77028	4.21354	2.319	0.021174	*
Chomage	-12.01513	4.16103	-2.888	0.004206	**
Etablissements1	-4.35123	1.57208	-2.768	0.006044	**
Etablissements2	-4.37403	1.87558	-2.332	0.020451	*
Etablissements3M	-4.60597	1.19151	-3.866	0.000140	***
CSP1	-36.51921	6.86035	-5.323	2.19e-07	***
CSP2	-38.03707	7.17303	-5.303	2.42e-07	***
CSP4	-17.29269	6.37217	-2.714	0.007092	**
CSP5	-11.55535	4.06899	-2.840	0.004866	**
CSP6	-7.04911	2.86076	-2.464	0.014377	*
CSP7	-9.29970	2.62457	-3.543	0.000467	***
X0_14	-7.71089	2.88636	-2.671	0.008023	**
X75_89	-7.25780	3.64397	-1.992	0.047435	*
SalairCSP5	-0.92215	0.20435	-4.513	9.67e-06	***
RapportInterdecile	0.14012	0.06892	2.033	0.043062	*

Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’	0.1 ‘ ’ 1

Finalement nous observons sur le tableau 19 que le salaire horaire moyen net ne fait pas partie du modèle. Nous pouvons donc en déduire qu'il n'influe pas sur la proportion d'agences Crédit Mutuel dans les zones d'emploi. Cependant nous pouvons observer qu'une des variables explicatives est le salaire net horaire moyen des employés. Sa t-valeur étant négative nous pouvons donc en déduire que plus ce salaire est faible plus la proportion d'agences Crédit Mutuel est grande. De plus nous pouvons observer que la part d'agences du Crédit Mutuel par zone d'emploi est également plus grande lorsque la part de la population active de 15 ans à 64 ans est grande, lorsque la part de la population active dans le secteur tertiaire marchand est grande ainsi que lorsque le rapport interdécile des salaires est élevé. Enfin nous pouvons remarquer que contrairement aux deux autres

banques coopératives le nombre d'agences Crédit Mutuel par zone d'emploi est influencé par le taux de chômage. En effet, plus celui-ci est faible plus il y a d'agences Crédit Mutuel implantées.

La Société Générale

Maintenant que nous avons terminé l'étude des banques coopératives nous allons étudier la Société Générale qui est une banque commerciale afin de pouvoir comparer les deux types de banques. Au vu de nos précédents résultats, nous pouvons penser que la part d'agences de la Société Générale va être plus grande dans les zones d'emploi où le salaire net horaire moyen est plus élevé.

Tableau 20 : Tableau des variables explicatives du nombre d'agences Société Générale par zone d'emploi

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	6.325e+01	2.377e+01	2.661	0.008276 **
Chomage	-6.649e+00	3.179e+00	-2.091	0.037465 *
Etablissements1	1.004e+00	5.852e-01	1.716	0.087399 .
CSP1	-1.275e+01	5.055e+00	-2.521	0.012296 *
CSP2	-8.257e+00	4.427e+00	-1.865	0.063283 .
CSP3	3.203e+01	5.993e+00	5.344	2.00e-07 ***
CSP4	-1.266e+01	4.265e+00	-2.969	0.003267 **
CSP6	-8.718e+00	2.161e+00	-4.035	7.21e-05 ***
CSP7	-5.204e+00	2.090e+00	-2.490	0.013399 *
X0_14	-5.621e+01	2.346e+01	-2.396	0.017303 *
X15_29	-5.964e+01	2.354e+01	-2.534	0.011870 *
X30_44	-5.166e+01	2.324e+01	-2.223	0.027109 *
X45_59	-6.104e+01	2.319e+01	-2.632	0.008992 **
X60_74	-5.581e+01	2.310e+01	-2.416	0.016366 *
X75_89	-6.682e+01	2.843e+01	-2.350	0.019511 *
Salaire	-6.033e-01	1.470e-01	-4.104	5.45e-05 ***
SalaireCSP3	1.699e-01	4.878e-02	3.484	0.000581 ***
SalaireCSP4	3.366e-01	1.143e-01	2.945	0.003527 **
SalaireCSP5	-2.730e-01	1.852e-01	-1.474	0.141817
SalaireCSP6	1.831e-01	1.271e-01	1.440	0.151079
MedianeNiveauVie	-9.718e-05	3.200e-05	-3.037	0.002634 **
RapportInterdecile	8.669e-02	4.972e-02	1.744	0.082431 .

Signif. codes:	0 ‘***’	0.001 ‘**’	0.01 ‘*’	0.05 ‘.’
	1			

Étonnamment on observe dans le tableau 20 que plus le salaire horaire net moyen est faible plus la part d'agences Société Générale par zone d'emploi est grande ce qui ne correspond pas aux résultats précédents et à nos hypothèses. Cependant nous remarquons également que plus le salaire net horaire moyen des cadres ou des professions intermédiaires est élevé plus la part d'agences Société Générale dans les zones d'emploi est élevée ce qui correspond plus à nos hypothèses. De plus on peut remarquer que la part

de la population cadre et ouvrière sont les variables qui expliquent le plus la part d'agences Société Générale dans les zones d'emploi. En effet, on retrouve proportionnellement plus d'agences Société Générale dans des zones d'emploi où il y a plus de cadres et moins d'ouvriers, ce qui correspond beaucoup plus à l'image d'une banque commerciale.

Conclusion

Finalement après avoir effectué les statistiques descriptives nous aurions pu penser qu'il y avait deux profils bien distincts entre les banques coopératives et commerciales notamment axés autour du salaire et de la part des agriculteurs. Cependant nous avons pu remarquer après avoir fait l'étude des statistiques prédictives que les différences entre les banques coopératives et commerciales sont moins perceptibles que ce que nous aurions pu imaginer même si certaines différences initiales semblent persister.

Nous pouvons donc conclure à la suite de nos différentes analyses que les banques coopératives restent encore relativement proches des ménages même si la diversité des différentes organisations a été réduite.

Annexes

Annexe 1 : Codes scraping

Code scraping Crédit Agricole

```
### Installation des packages
install.packages("rvest")
install.packages("stringr")

### Chargement des packages
library(stringr)
library(rvest)

### Lecture de la première page pour tester le code
liendep<-"https://www.credit-agricole.fr/particulier/agence/sud-rhonealpes/vizille-0002.html"
#Lien de départ
Page<-read_html(x=liendep) #Lit le code html de la page web à partir du lien de départ
table<-html_text(html_nodes(Page, ".CardAgencyFunc-address")) #Dans le texte HTML,
chercher le noeud HTML comprenant ".CarAgencyFunc-address"

TableSorties<-matrix(NA,nrow=9999,ncol=20) #Création d'une matrice vide avec 9999
lignes et 20 colonnes.

i<-2 #on commence l'itération à 2 car le premier lien qui contient l'adresse d'une banque est
le lien "https://www.credit-agricole.fr/particulier/agence/sud-rhonealpes/vizille-0002.html"
while(i <9999){
    #on fait aller la boucle jusqu'à 9999 car il y a environ 7000 banques Crédit Agricole
    en France
    deblien<-"https://www.credit-agricole.fr/particulier/agence/sud-rhonealpes/vizille-000"
    #Début du lien html
    midlien<-i #Milieu du lien qui permet de faire une boucle
    finlien<=".html" #rajoute la terminaison du lien html
    lien1<-str_c(deblien, midlien) #Concatène la première partie du lien avec la fin du
    lien
    lien<-str_c(lien1,finlien) #Concatène la première partie du lien avec la fin du lien
    Page<-try(read_html(x=lien)) #Lit le code html associé à la page web du lien.
```

```

table3<-try(html_text(html_nodes(Page,".CardAgencyFunc-address")))
#Dans le texte HTML, chercher le noeud HTML comprenant ".CarAgencyFunc-
address" pour obtenir l'adresse de la banque
TableSorties<-rbind(table3,TableSorties) #rajoute cette adresse dans notre matrice
i<-i+1 #incrémentation
print(i) #afficher l'avancement de la boucle
print(lien) #affichage du lien
}
TableSorties2<-str_replace_all(TableSorties,pattern=",", replacement = " ") #remplace toutes
les virgules de la table par des espaces afin d'avoir des données plus propres en sortie csv
write.table(TableSorties,"C:/Users/33652/OneDrive/Bureau/STID2A/PROJETTUT/BD/CA/do
nneesCA.csv", sep="",row.names=FALSE) #permet de faire la sortie au format csv

```

Code scraping Crédit Mutuel

```

### Installation des packages
install.packages("rvest")
install.packages("stringr")

### Chargement des packages
library(stringr)
library(rvest)

### Déclaration des variables qui stockeront les adresses des banques
Sorties_corse<-matrix(NA,nrow=10,ncol=1)
Sorties<-matrix(NA,nrow=500,ncol=1)

### Pour le département de la Corse
##Récupération des villes situées en Corse
#Lien de la page qui donne accès aux villes situées en Corse
lienv<-"https://www.creditmutuel.fr/fr/banques/contact/trouver-une-
agence/BrowseLocality.aspx?SubdivisionId=FR-H"

#Lecture du code HTML de la page précédente
Pagev<-read_html(x=lienv)

#Récupération des chaînes de caractère contenues dans les balises <span> en une liste
listec<-html_text(html_nodes(Pagev,".lister span"))

#Remplacer tous les espaces (" ") présents dans éléments de la liste par un tiret ("")
listec<-str_sub(str_replace_all(listec,pattern=" ", replacement = "-"))
#La variable listec est donc une liste des villes de Corse

##Récupération des adresses des banques de chaque villes de Corse
#Déclaration du compteur qui nous permettra de parcourir listec dans la boucle while
i<-1

#Boucle qui permet la récupération des adresses de banques en Corse

```

```
while(i <=length(listec)){ #Tant que le compteur est inférieur ou égal à la longueur de la liste
  #Première partie de l'URL
  deblien<-https://www.creditmutuel.fr/fr/banques/contact/trouver-une-agence/SearchList.aspx?sub=true&type=branch&loca="
  #Nom de la ville positionnée à la ième place dans listec
  ville<-listec[i]
  #Concaténation de la première partie de l'URL avec la ville afin d'obtenir l'URL qui donne accès à l'adresse de la banque de la ville en question
  lien<-str_c(deblien, ville)
  #Lecture du code HTML de la page, le try permet à la boucle while de continuer jusqu'à la condition indiqué même s'il y a une erreur avec la lecture de la page
  Page<-try(read_html(x=lien))
  #Récupération du premier élément dans la balise <classe> lorsque class=invisible
  tablec<-try(html_text(html_node(Page,"em .invisible")))
  #Matrice qui stock les adresses
  Sorties_corse<-rbind(tablec,Sorties_corse)
  #Affiche la valeur de i
  print(i)
  #Affiche l'URL permettant d'accéder à l'adress de la banque de la ville
  print(lien)
  #Incrémantation de i
  i<-i+1
}

#Change les chaînes de caractère "\r\n" en chaîne de caractère vide dans la matrice qui répertorie les adresses
Sorties_corse<-str_sub(str_replace_all(Sorties_corse,pattern="\r\n",replacement = ""))
#Créer un fichier CSV avec les adresses collectées précédemment dans la bibliothèque du chemin précisé
write.table(Sorties_corse,"D:/Users/Nathalie/Documents/STID/2A/S3/Projet
tutoré/Données/Scraping/Crédit Mutuel/CM_corse.csv",sep="",row.names=FALSE)

### Pour les autres départements
##Récupération des villes des autres départements
#Déclaration du compteur et des listes
j<-1
liste<-c()
listev<-c()

#Boucle qui nous permet de parcourir les pages des 94 autres départements de la France métropolitaine et d'en extraire les villes
while(j<=95){
  #Le compteur ne permet pas de parcourir les département 01 à 09 car j est un compteur et ne peut prendre "01" en valeur
  #Si le compteur j est inférieur à 10, la première partie de l'URL va contenir un "0" à la fin pour parcourir les départements 01 à 09
```

```

if(j<10){
    debliend<-"https://www.creditmutuel.fr/fr/banques/contact/trouver-une-
    agence/BrowseLocality.aspx?SubdivisionId=FR-0"
    #Sinon, la première partie de l'URL ne va pas contenir de "0" à la fin
}
else{
    debliend<-"https://www.creditmutuel.fr/fr/banques/contact/trouver-une-
    agence/BrowseLocality.aspx?SubdivisionId=FR-"
}
#Concaténation de la première partie de l'URL avec le j afin d'obtenir l'URL du
département j
liend<-str_c(debliend,j)
#Lecture du code HTML de la page
Paged<-read_html(x=liend)
#Récupération du texte situé dans la balise <span> dans une liste
liste<-html_text(html_nodes(Paged,".lister span"))
#liste qui stock les villes à chaque fois que la boucle while est parcourue
listev<-c(listev,liste)
#Affiche la valeur de j
print(j)
#Affiche l'URL permettant d'accéder à la page qui contient les villes du département j
print(liend)
#Incrémantation du compteur
j<-j+1
}

#Remplace les espaces par des tirets de tous les éléments de la liste listev
listev<-str_sub(str_replace_all(listev,pattern=" ", replacement = "- "))

##Récupération des adresses des banques
#Déclaration du compteur i
i<-1

#Boucle qui va parcourir les différentes URL pour extraire les adresses des banques
while(i <=length(listev)){
    #première partie de l'URL
    deblien<-"https://www.creditmutuel.fr/fr/banques/contact/trouver-une-
    agence/SearchList.aspx?sub=true&type=branch&loca="
    #La ième ville de la liste
    ville<-listev[i]
    #Concaténation de deblien et ville afin d'obtenir URL permettant d'accéder à
    l'adresse de la banque de la ville i
    lien<-str_c(deblien, ville)
    #Lecture du code HTLM de la page, le try permet à la boucle while de continuer
    jusqu'à la condition indiqué même s'il y a une erreur avec la lecture de la page
    Page<-try(read_html(x=lien))
    #Récupération du premier élément dans la balise <classe> lorsque class=invisible
}

```

```

table<-try(html_text(html_node(Page,"em .invisible")))
#Matrice qui stock les adresses
Sorties<-rbind(table,Sorties)
#Affiche i
print(i)
#Affiche l'URL de la page avec l'adresse de la banque dans la ville i
print(lien)
#Incrémantation du compteur
i<-i+1
}

## Sortie des adresses
#Change les chaînes de caractère "\r\n" en chaîne de caractère vide dans la matrice qui
répertorie les adresses
Sorties<-str_sub(str_replace_all(Sorties,pattern="\r\n",replacement = " "))
#Créer un fichier CSV avec les adresses collectées précédemment dans la bibliothèque du
chemin précisé
write.table(Sorties,"D:/Users/Nathalie/Documents/STID/2A/S3/Projet
tutoré/Données/Scrapping/Crédit Mutuel/CM.csv",sep="",row.names=FALSE)

```

Code scraping Banque Populaire

```

### Installation des packages
install.packages("rvest")
install.packages("stringr")

### Chargement des packages
library(stringr)
library(rvest)

### Initialisation des variables nécessaires
# Liste des départements avec leurs numéros
dpt<-list("ain-01","aisne-02","allier-03","alpes-de-haute-provence-04","hautes-alpes-
05","alpes-maritimes-06","ardeche-07","ardennes-08","ariege-09","aube-10","aude-
11","aveyron-12","bouches-du-rhone-13","calvados-14","cantal-15","charente-16","charente-
maritime-17","cher-18","correze-19","corse-du-sud-2A","haute-corse-2B","cote-d-or-
21","cotes-d-armor-22","creuse-23","dordogne-24","doubs-25","drome-26","eure-27","eure-
et-loir-28","finistere-29","gard-30","haute-garonne-31","gers-32","gironde-33","herault-
34","ille-et-vilaine-35","indre-36","indre-et-loire-37","isere-38","jura-39","landes-40","loir-et-
cher-41","loire-42","haute-loire-43","loire-atlantique-44","loiret-45","lot-46","lot-et-garonne-
47","lozere-48","maine-et-loire-49","manche-50","marne-51","haute-marne-52","mayenne-
53","meurthe-et-moselle-54","meuse-55","morbihan-56","moselle-57","nievre-58","nord-
59","oise-60","orne-61","pas-de-calais-62","puy-de-dome-63","pyrenees-atlantiques-
64","hautes-pyrenees-65","pyrenees-orientales-66","bas-rhin-67","haut-rhin-68","rhone-
69","haute-saone-70","saone-et-loire-71","sarthe-72","savoie-73","haute-savoie-74","paris-
75","seine-maritime-76","seine-et-marne-77","yvelines-78","deux-sevres-79","somme-
80","tarn-81","tarn-et-garonne-82","var-83","vaucluse-84","vendee-85","vienne-86","haute-

```

```
vienne-87", "vosges-88", "yonne-89", "territoire-de-belfort-90", "essonne-91", "hauts-de-seine-92", "seine-saint-denis-93", "val-de-marne-94", "val-d-oise-95")
```

```
TableSorties<-matrix()
```

```
# Création d'une matrice vide pour stocker les résultats.
```

```
### Lecture des pages html de la Banque Populaire.
```

```
i<-1 # Initialisation d'un compteur i qui va permettre de parcourir notre liste dans la boucle.  
while(i <= length(dpt)){
```

```
# Création d'une boucle "while" pour pouvoir recréer tous les liens nécessaires.
```

```
deblien<-"https://agences.banquepopulaire.fr/banque-assurance/agences-"
```

```
#Début du lien de chaque page dans lesquelles nous voulons aller récupérer les adresses des banques.
```

```
midlien<-dpt[i]
```

```
# Utilisation du compteur i pour qu'il parcoure la liste des départements.
```

```
lien<-str_c(deblien, midlien)
```

```
# Concaténation du début de la première partie et de la deuxième partie du lien.
```

```
Page<-read_html(x=lien)
```

```
# Lecture de la page html associée au lien recréé ci-dessus grâce à la commande read_html.
```

```
table<-matrix(html_text(html_nodes(Page,"div .em-poi-card__address"))))
```

```
# La fonction html_nodes permet d'extraire des morceaux de code html. Cette commande prend en arguments l'URL de la page dans laquelle on cherche à extraire des informations et le nom de la balise qu'on veut extraire. Ici, par exemple, on veut extraire le texte qui est contenu dans les balises html se nommant ".em-poi-card__address". On utilise également la balise "html_text" pour afficher seulement le contenu textuel qui se trouve dans cette balise.
```

```
TableSorties<-rbind(table,TableSorties)
```

```
# Fusion des deux tableaux de données grâce à la commande "rbind".
```

```
i<-i+1 # Augmentation du compteur de 1.
```

```
print(i)
```

```
print(lien)
```

```
}
```

```
### Récupération des résultats dans un fichier csv.
```

```
TableSorties<-str_replace_all(TableSorties,pattern=",", replacement = " ")
```

```
# Remplacement des virgules par des espaces dans le fichier de sorties.
```

```
write.table(TableSorties,"C:/Users/baudj/Documents/STID2A/PROJET/Resultats/sortie_csv_3.csv",col=NA, sep="")
```

Code scraping Société Générale

```
### Installation des packages
```

```
install.packages("rvest")
```

```
install.packages("stringr")
```

```
### Chargement des packages
```

```
library(stringr)
```

```
library(rvest)
```

```
#Création listes départements & numéros
dept<-list("ain","aisne","allier","alpes-de-haute-provence","hautes-alpes","alpes-
maritimes","ardeche","ardennes","ariege","aube","aude","aveyron","bouches-du-
phone","calvados","cantal","charente","charente-maritime","cher","correze","corse-du-
sud","haute-corse","cote-d-or","cotes-d-
armor","creuse","dordogne","doubs","drome","eure","eure-et-loir","finistere","gard","haute-
garonne","gers","gironde","herault","ille-et-vilaine","indre","indre-et-
loire","isere","jura","landes","loir-et-cher","loire","haute-loire","loire-
atlantique","loiret","lot","lot-et-garonne","lozere","maine-et-loire","manche","marne","haute-
marne","mayenne","meurthe-et-
moselle","meuse","morbihan","moselle","nièvre","nord","oise","orne","pas-de-calais","puy-de-
dome","pyrenees-atlantiques","hautes-pyrenees","pyrenees-orientales","bas-rhin","haut-
rhin","rhone","haute-saone","saone-et-loire","sarthe","chambery","haute-
savoie","paris","seine-maritime","seine-et-marne","yvelines","deux-
sevres","somme","tarn","tarn-et-garonne","var","vaucluse","vendée","vienne","haute-
vienne","vosges","yonne","territoire-de-belfort","essonnes","hauts-de-seine","seine-saint-
denis","val-de-marne","val-d-oise")
numdpt<-
list("01","02","03","04","05","06","07","08","09","10","11","12","13","14","15","16","17","18","19
","2A","2B","21","22","23","24","25","26","27","28","29","30","31","32","33","34","35","36","37",
"38","39","40","41","42","43","44","45","46","47","48","49","50","51","52","53","54","55","56","5
7","58","59","60","61","62","63","64","65","66","67","68","69","70","71","72","73","74","75","76
","77","78","79","80","81","82","83","84","85","86","87","88","89","90","91","92","93","94","95")
```

```
#Initialisation boucle while
i<-1
tablesortie<-matrix()

while (i<=length(dept)){
  #####Construction du lien de la page web
  deblien<-"https://agences.societegenerale.fr/banque-assurance/agences-" #Début du
  lien
  mid1lien<-dept[i] #Suite du lien
  mid2lien<-" " #Suite du lien
  finlien<-numdpt[i] #Fin lien
  lien1<-str_c(deblie, mid1lien) #Concatène le début du lien et le milieu du lien
  lien2<-str_c(lien1, mid2lien) #Concatène le milieu du lien et la suite du lien
  lien<-str_c(lien2,finlien) #Concatène la première partie du lien avec la fin du lien
  ### Lecture de la page
  page<-try(read_html(x=lien))
  ### récupération du noeud HTML contenant les adresses des banques dans un
  département
  table<-matrix(html_text(html_nodes(page,".agencyaddress")))
  ### construction de la table
  tablesortie<-rbind(tablesortie,table)}
```

```
###Visualisation du résultat de l'itération
print(lien)
print(i)
###avancement de l'indice pour passer au département suivant
i<-i+1
}

#Convertir la matrice des données en fichier csv
TableSorties<-str_replace_all(tablesortie,pattern=",", replacement = "")
write.table(TableSorties,file="/Users/manonandre/Desktop/SG.csv",col=NA, sep="")
```

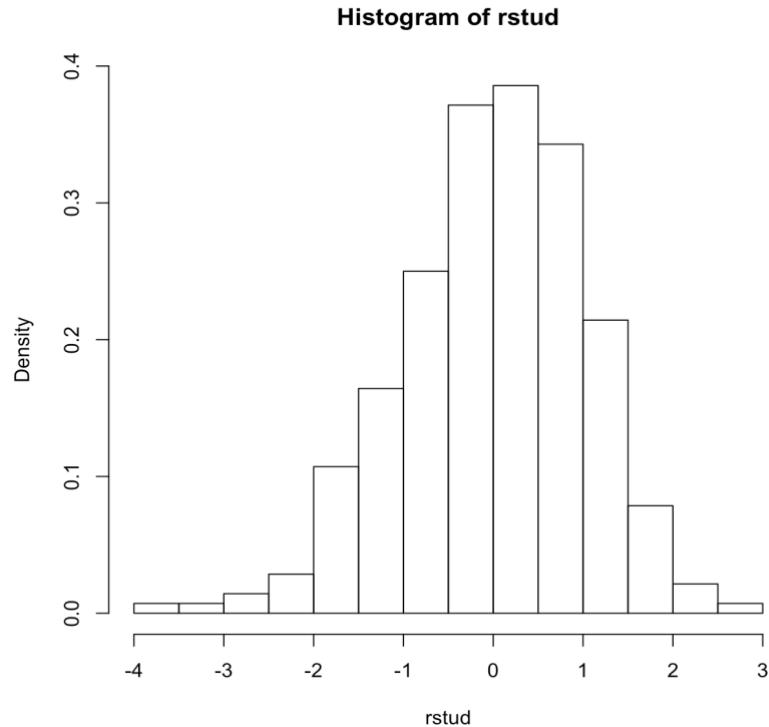
Annexe 2 : Visualisation des données

massonm8 BANQUEFINAL	
	VILLE : varchar(41)
#	DEPARTEMENT : int(2)
	BANQUE : varchar(16)
	TYPEBANQUE : varchar(30)
#	ZE : int(4)

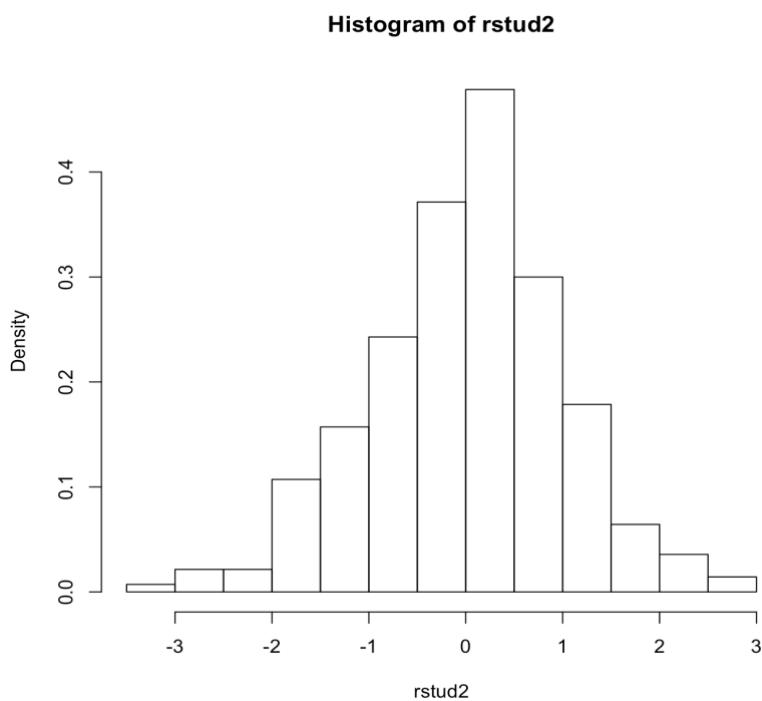
massonm8 SOCIOECO	
	ZE : int(4)
	NomZE : varchar(30)
#	PopActive1 : float
#	PopActive2 : float
#	PopActive3M : float
#	PopActive3NonM : float
#	PartActifs15_64 : float
#	Chomage : float
#	NbEtablissements1 : float
#	NbEtablissements2 : float
#	NbEtablissements3M : float
#	NbEtablissements3NonM : float
#	CSP1 : float
#	CSP2 : float
#	CSP3 : float
#	CSP4 : float
#	CSP5 : float
#	CSP6 : float
#	CSP7 : float
#	CSP8 : float
#	0_14 : float
#	15_29 : float
#	30_44 : float
#	45_59 : float
#	60_74 : float
#	75_89 : float
#	90Plus : float
#	Salaire : float
#	SalaireCSP3 : float
#	SalaireCSP4 : float
#	SalaireCSP5 : float
#	SalaireCSP6 : float
#	MedianeNiveauVie : float
#	RapportInterdecile : float

Annexe 3 : graphiques des résidus des régressions linéaire

Graphiques des résidus studentisés avant sélection de variable



Graphiques des résidus studentisés avant sélection de variable



Annexe 4 : Cahier des charges

IUT2-STID 2^e année

Implantation des banques coopératives en France



ANDRE Manon

BAUD Joséphine

MASSON Maxence

TANG Nathalie

Cahier des charges Projet Tutoré

Implantation des banques coopératives en France



Commanditaire

ARTIS Amélie

Tutrices

FAİN Edwige

LETUÉ Frédérique



ANDRE Manon, BAUD Joséphine, MASSON Maxence, TANG Nathalie

1

Table des matières

<i>Implantations des banques coopératives en France.....</i>	1
I) Présentation du projet.....	3
<i>Acteurs du Projet : les parties prenantes.....</i>	3
Maître d'ouvrage :.....	3
Maîtres d'œuvre :.....	3
Coordonnées téléphoniques :.....	3
Tutrices :.....	3
<i>Contexte du projet.....</i>	3
<i>Objectifs du projet.....</i>	5
<i>Contraintes liées au projet.....</i>	5
Crise sanitaire :.....	5
Délais :.....	5
Qualité :.....	6
<i>Ressources du projet.....</i>	7
II) Mise en œuvre du projet.....	9
<i>Tâches à réaliser.....</i>	9
<i>Organisation des RH.....</i>	10
<i>Risques associés au projet.....</i>	11
<i>Diagramme de GANTT.....</i>	14

I) Présentation du projet

Acteurs du Projet: les parties prenantes

Maître d'ouvrage :

- ARTIS Amélie : amelie.artis@iepq.fr

Maîtres d'œuvre :

Étudiants de deuxième année en DUT STID à l'IUT2 de Grenoble :

- ANDRE Manon : manon.andre1@etu.univ-grenoble-alpes.fr
- BAUD Joséphine : josephine.baud@etu.univ-grenoble-alpes.fr
- MASSON Maxence : maxence.masson@etu.univ-grenoble-alpes.fr
- TANG Nathalie : nathalie.tang@etu.univ-grenoble-alpes.fr

Coordinnées téléphoniques :

- ANDRE Manon : 06 76 60 46 72
- BAUD Joséphine : 07 82 38 73 34
- MASSON Maxence : 06 52 67 19 15
- TANG Nathalie : 06 03 50 27 22

Tutrices :

- FAÏN Edwige : edwige.fain@univ-grenoble-alpes.fr
- LETUË Frédérique : [frédérique.letue@univ-grenoble-alpes.fr](mailto:frederique.letue@univ-grenoble-alpes.fr)

Contexte du projet

Aujourd'hui, le spécialiste de la Statistique et de l'Informatique Décisionnelle intervient partout où il faut gérer des données (numériques ou autres), les organiser, en extraire l'information, l'analyser, en faire la synthèse pour éclairer celui qui prend les décisions. À l'heure actuelle, dans tous les secteurs d'activité, des quantités considérables de données, qui sont enregistrées, doivent être traitées.

Le DUT STID forme des techniciens de la Statistique et de l'Informatique Décisionnelle en 2 ans. C'est notamment un des DUT proposé par l'une des composantes de l'Université Grenoble Alpes : Institut Universitaire Technologique 2 (IUT2). L'Université Grenoble Alpes (UGA) est une université internationale, attractive et avec un rayonnement international marqué. L'UGA cherche à s'enraciner dans son territoire mais également d'être un acteur majeur dans le développement économique, culturel et social de ce territoire.

Pour répondre aux métamorphoses de la société actuelle, elle propose aux étudiants des formations riches, novatrices et variées.

Dans le cadre de notre formation, nous avons dû nous charger d'un projet dans le but de nous professionnaliser. Ce projet nous permettra de mettre en application nos connaissances et nos savoir-faire appris au cours de notre formation mais également de mettre en place une notion de transdisciplinarité.

La commanditaire de ce projet, Amélie ARTIS, est maître de conférence et enseignante-rechercheuse en économie sociale et solidaire à Science-Po-Grenoble. Le concept d'économie sociale et solidaire (ESS) désigne un ensemble d'entreprises organisées sous forme de coopératives, mutualistes, associations, ou fondations, dont le fonctionnement interne et les activités sont fondés sur un principe de solidarité et d'utilité sociale.

Dans le cadre de ses recherches, Amélie ARTIS est amenée à répondre à la problématique :

« Pourquoi et comment expliquer la diversité des organisations/entreprises qui se développent dans notre système économique vu comme capitaliste ? »

Le capitalisme est un système économique dont les fondements sont l'entreprise privée et la liberté du marché. Parmi ses recherches, on cherche ici à s'intéresser plus particulièrement au secteur bancaire coopératif. Selon l'Alliance Coopérative Internationale, une banque coopérative est « une association autonome de personnes volontairement réunies pour satisfaire leurs aspirations et besoins économiques, sociaux et culturels communs au moyen d'une entreprise dont la propriété est collective et où le pouvoir est exercé démocratiquement ». En d'autres termes, ces organismes sont gouvernés démocratiquement et chaque adhérent a le même pouvoir de décision indépendamment du montant de capitaux qu'il apporte à la banque.

En pratique, dans ce type de banques les décisions sont prises par les clients. Il n'y a pas d'actionnaire. Ces banques sont détenues par leurs clients qui peuvent être des personnes physiques ou morales. Ces individus détiennent les parts composant le capital social de la banque. Ce statut leur permet de toucher des dividendes annuels et de participer aux décisions prises lors des assemblées générales.

Historiquement, les banques coopératives ont joué un rôle important dans les systèmes financiers de presque tous les pays européens.

Ces réseaux ont été mis en place au cours du 19^e siècle pour faciliter l'accès aux crédits aux personnes issues de certaines catégories professionnelles qui éprouvaient des difficultés à emprunter. En les associant au pouvoir décisionnel des banques, leur accès à l'emprunt a été favorisé.

Objectifs du projet

Parmi ces travaux de recherche, nous sommes invités à étudier des questions plus précises :

- Est-ce qu'aujourd'hui l'essence historique des banques coopératives d'être proche des ménages est toujours observée ?
- Existe-t-il encore une diversité des organisations sur le territoire ?

Nos premières recherches se baseront sur de la recherche et de la sélection de données :

- Les variables socio-économiques sont à récupérer sur le site de L'Insee. Elles seront ensuite à trier selon leur pertinence. Il faudra enfin sélectionner une échelle territoriale pertinente : les zones d'emplois.
- Les données de localisation des banques coopératives seront à construire : mettre à jour les données du projet de l'an dernier et les comparer.

Il y a également la dimension base de données (où les données seront stockées) à prévoir en amont. Ces bases de données seront construites en fonction de ce que nous voudrons montrer du point de vue économique. En effet, notre objectif final, pour ce qui est des données, est de former une unique base de données en liant les données de localisation aux données socio-économiques. Cette base de données contiendra des données nettoyées et pertinentes entre elles.

Par la suite nous devrons mettre en place des procédures statistiques pour étudier la présence des banques coopératives. Cela nous mènera à déterminer une typologie de l'implantation des banques coopératives.

Contraintes liées au projet

Crise sanitaire :

- Enseignement à distance
- Fermeture du bâtiment
- Accès limité aux logiciels
- Équipement informatique personnel nécessaire
- Évaluations à distance

Délais:

- Les échéances connues :
 - Cahier des charges : 07/01/2021
 - Présentation orale du cahier des charges : 08/01/2021
 - Présentation orale en anglais durant la International Week : 21/01/2021
 - Rapport final : semaine du 22/03/2021

- o Présentation orale : semaine du 22/03/2021

Qualité :

- Niveau du travail attendu:
 - o Orthographe
 - o Constance et rigueur du travail
 - o Efficacité lors des missions
 - o Codage en différents langages informatiques
 - o Analyse fine des données
 - o Éveiller notre esprit d'initiative et notre esprit critique.
 - o Mise en place d'un cadre professionnalisant.
 - o Mise en application des savoirs et savoir-faire acquis pendant la formation.
 - o Développement du savoir-être.
- Contrôle par le maître d'ouvrage/les tutrices :
 - o Prévoir des réunions après chaque grande étape du projet terminée (avec comptes rendus).
 - o Informer régulièrement de nos avancements par mail, en visioconférence ou en face à face.

Ressources du projet

Ressources	Noms des propriétaires/ Organisme d'origine	Utilisation/Action	Budget
Données	Maîtres d'œuvre Insee	<ul style="list-style-type: none"> Utilisée pour répondre à la problématique 	Libre accès
Ressources matérielles			
Ordinateurs	UGA, IUT2, Maître d'œuvre	<ul style="list-style-type: none"> Communication et exploitation des données 	Effets personnels
Téléphones	Maîtres d'œuvre	<ul style="list-style-type: none"> Communication entre les maîtres d'œuvre 	Effets personnels
Disque durs/ Clé USB	Maîtres d'œuvre	<ul style="list-style-type: none"> Stockage des données et des documents 	Effets personnels
Ressources immatérielles			
Logiciel R Studio	UGA, IUT2, Maîtres d'œuvre	<ul style="list-style-type: none"> Collecte des données Manipulation de la base de données Élaboration de statistiques 	Libre accès
PhpMyadmin	Licence UGA, IUT2	<ul style="list-style-type: none"> Stockage des données Gestion des données 	Licence IUT2
Discord		<ul style="list-style-type: none"> Communication entre les maîtres d'œuvre Stockage des documents 	Libre accès
Google Drive		<ul style="list-style-type: none"> Rédaction des rapports écrits Partages de documents Stockage des documents 	Libre accès
Messagerie UGA	UGA	<ul style="list-style-type: none"> Communication entre les maîtres d'œuvre, le maître d'ouvrage et les 	Accès étudiant

		tutrices	
Plateforme Chamilo	UGA, IUT2	<ul style="list-style-type: none"> ▪ Partages de documents entre les maîtres d'œuvre et les tutrices 	Accès étudiant
Logiciel Microsoft	Licence UGA, IUT2, maître d'œuvre	<ul style="list-style-type: none"> ▪ Rédaction des rapports écrits 	Licence IUT
PhilCarto	Licence UGA, IUT2	<ul style="list-style-type: none"> ▪ Cartographier les données 	Libre accès
SPAD	Licence UGA, IUT2	<ul style="list-style-type: none"> ▪ Analyse des données ▪ Calculs statistiques 	Licence IUT
SAS	Licence UGA, IUT2	<ul style="list-style-type: none"> ▪ Calculs statistiques 	Licence IUT
Tableau Prep Et Tableau Builder	Licence UGA, IUT2, maître d'œuvre	<ul style="list-style-type: none"> ▪ Nettoyage des données ▪ Visualisation des données 	Licence IUT
Ressources humaines			
4 étudiants de 2 ^e année	UGA, IUT2	<ul style="list-style-type: none"> ▪ Réalisation du projet 	180h
Commanditaire (Amélie ARTIS)	UGA		
Tuteurs (Edwige FAİN et Frédérique LETUÉ)	UGA, IUT2	<ul style="list-style-type: none"> ▪ Supervision du projet ▪ Encadrement 	

II) Mise en œuvre du projet

Tâches à réaliser

- Récolter les données socio-économiques jugées pertinentes via l'Insee.
- Effectuer du ~~scrapping~~ de données pour créer les données géographiques des banques.
- Mener des requêtes en différents langages informatiques pour manipuler les données en fonction de notre problématique.
- Faire des analyses statistiques sur les résultats obtenus.

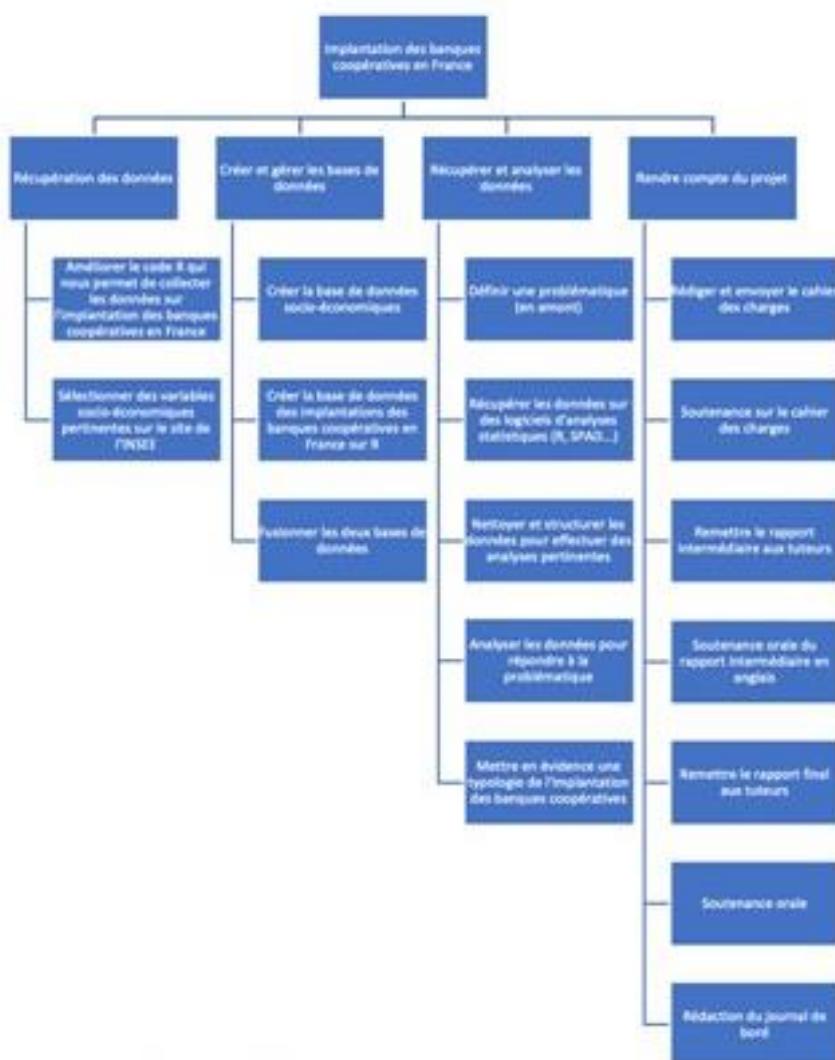


Figure 1 : Organigramme des tâches

Organisation des RH

T A C H E S		PARTIES PRENANTES						
		MAÎTRE D'OUVRAGE		TUTRICES		MAÎTRES D'ŒUVRE		
		ARTIS Amélie	FAIN Edwige	LETUE Frédérique	ANDRE Manon	BAUD Joséphine	MASSON Maxence	TANG Nathalie
	Redaction du journal de bord.		I	I	R	R	R	A/R/I
	Définir une problématique.	C/I	C/I	C/I	A/R	R	R	R
	QQQQCP		C/I	C/I	I	I	I	A/R
	Effectuer le cahier des charges.	I	C/I	C/I	A/R/I	R	R	R
	Soutenance du cahier des charges.		C/I	C/I	A/R	A/R	A/R	A/R
	Améliorer le code R (récolte des données sur l'implantation des banques coopératives en France).		I	I	R	R	R	A/R
	Sélectionner des variables socio-économiques qui semblent pertinentes sur le site de l'INSEE.	C/I	C/I	C/I	R	R	A/R	R
	Créer une base de données des variables socio-économiques sélectionnées précédemment.		I	I	R	A/R	R	R
	Remettre le rapport intermédiaire aux tuteurs.		C/I	C/I	A/R	R	R	R
	Soutenance du rapport intermédiaire en anglais.		C/I	C/I	A/R	A/R	A/R	A/R
	Créer une base de données des implantations des banques coopératives en France sur R.		C/I	C/I	R	A/R	R	R
	Fusionner les données des deux bases de données.		I	I	R	A/R	R	R
	Mettre en évidence une typologie de l'implantation des banques coopératives en France.		I	I	R	R	R	A/R
	Récupérer les données sur des logiciels d'analyses statistiques.		I	I	R	R	A/R	R
	Nettoyer et structurer les données pour effectuer des analyses pertinentes.		I	I	R	R	A/R	R
	Analyser les données pour répondre à la problématique.		I	I	R	R	R	A/R
	Remettre le rapport final aux tutrices et au commanditaire.	I	C/I	C/I	A/R	R	R	R
	Soutenance du rapport final.	I	C/I	C/I	A/R	A/R	A/R	A/R

Figure 2 : Matrice des ressources humaines

A	Responsable de la tâche
R	Réalise la tâche
C	Doit être consulté
I	Doit être informé

Risques associés au projet

	Catastrophique		Risque de perte de la clé USB sur laquelle l'avancement du projet est stocké			
Incidence	Très grave	Risque de ne pas terminer le projet	Risque de ralentissement dans notre progression suite à la maladie d'un des membres du groupe	Risque de produire une analyse biaisée de nos données		
	Majeure	Risque de ne pas respecter les échéances et les délais	Risque de problème de connexion de l'un des membres du groupe	Risque de mal sélectionner les variables parmi celles proposées par l'INSEE	Risque de ne pas réussir à fusionner nos données provenant de deux sources différentes	
	Sigificative		Risque de mauvaise compréhension du sujet du fait que nous n'avons eu que des réunions sur Zoom	Risque de ne pas pouvoir utiliser certains logiciels chez nous	Risque de ne pas retourner à l'université	
	Mineure			Risque de mal s'approprier le script. R élaboré par un groupe d'étudiants d'une année précédente		Risque de partir sur une mauvaise piste de réflexion
	Matrice des risques	Incrovable	Peu probable	Occasionnel	Probable	Presque certain
Probabilité						

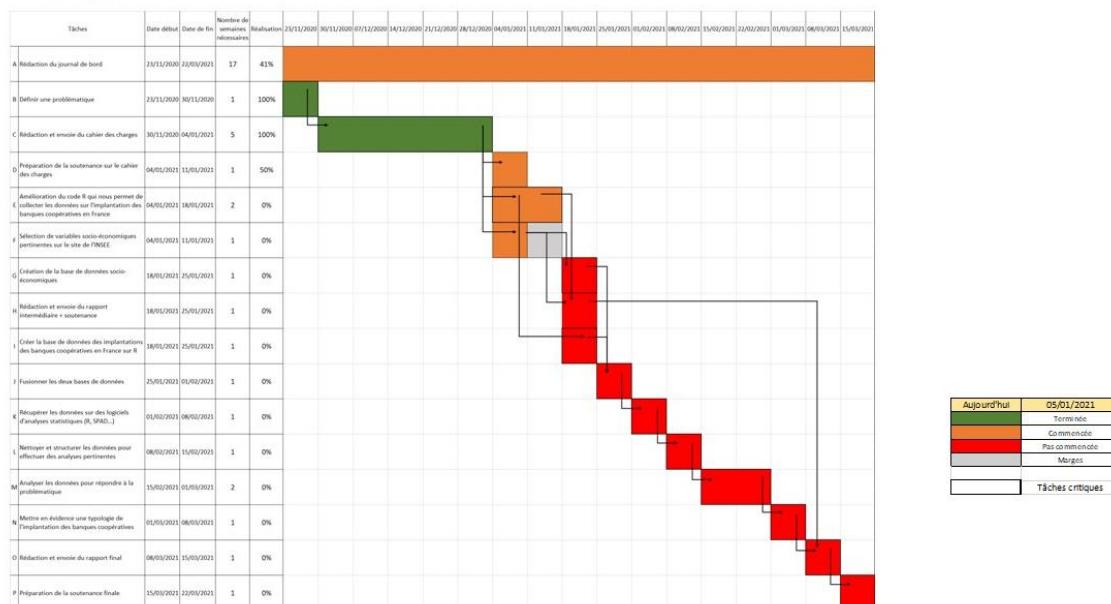
Figure 3 : Matrice des risques

Conséquences et solutions :

RISQUES	CONSEQUENCES	SOLUTIONS
Risque de ne pas respecter les échéances et les délais.	Prendre du retard sur la réalisation du projet.	- Convenir des délais à respecter avec nos tutrices et notre commanditaire, - Suivre le diagramme de GANTT.
Risque de mauvaise compréhension du sujet du fait que nous n'avons eu que des réunions sur Zoom.	Ralentissement du projet.	- Faire des réunions avec nos tutrices régulièrement afin de s'assurer que nous sommes toujours sur la bonne voie.
Risque de mal s'approprier le script R élaboré par un groupe d'étudiants d'une année précédente.	Exploitation seulement partielle de nos ressources.	- Lire le script attentivement et essayer de comprendre toutes les commandes qu'il contient. - Tester le script sur une petite quantité de données pour s'assurer que nous ayons bien compris son fonctionnement.
Risque de partir sur une mauvaise piste de réflexion.	Ralentissement du projet.	- Faire des réunions avec nos tutrices régulièrement afin de s'assurer que nous sommes toujours sur la bonne voie.
Risque de ne pas pouvoir utiliser certains logiciels chez nous.	Difficulté à réaliser certaines étapes du projet plus particulièrement celles qui impliquent notre base de données.	- S'assurer que les logiciels que nous prévoyons d'utiliser sont accessibles en ligne. - Demander des licences étudiantes si nécessaire.
Risque de problème de connexion de l'un des membres du groupe.	Augmentation de la charge de travail pour les autres membres du groupe.	- Prévoir des tâches réalisables hors connexion (réécriture d'un document, ...).
Risque de ne pas terminer le projet.	Insatisfaction de notre commanditaire.	- Faire preuve d'organisation. - Respecter le diagramme de GANTT.
Risque de perte de la clé USB sur laquelle l'avancement du projet est stocké.	Ralentissement du projet, devoir refaire une partie du projet.	- Sauvegarder notre avancement sur plusieurs supports. - Créer un dossier partagé en ligne.
Risque de ralentissement dans notre progression suite à la maladie d'un des membres du groupe.	Augmentation de la charge de travail pour les autres membres du groupe.	- Essayer de ne pas prendre de retard sur les tâches à faire pour pouvoir faire face à ce ralentissement.
Risque de ne pas retourner à l'université.	Ne faire que des réunions à distance et ne pas pouvoir travailler tous ensemble.	- Prévoir de réaliser l'intégralité du projet à distance.

Risque de produire une analyse biaisée de nos données.	Résultats faussés.	- Être minutieux et rigoureux lors de la réalisation de nos analyses. - Vérifier nos résultats entre les membres du groupe.
Risque de ne pas réussir à fusionner nos données provenant de deux sources différentes.	Avoir deux bases de données distinctes ce qui compliquerait les analyses.	- Nettoyer et organiser les deux jeux de données de façon à ce qu'on puisse les mettre en relation ensuite.
Risque de mal sélectionner les variables parmi celles proposées par l'INSEE.	Passer à côté de certains axes d'études.	- Étudier, dans un premier temps, toutes les variables pour être sûr de ne pas manquer d'informations importantes.

Diagramme de GANTT



Annexe 5 : Journal de bord

Date	Personnes concernées	Bilan du travail réalisé	Problèmes, questions
25/11/2020	La commanditaire, les tuteurs, les maîtres d'œuvres	Présentation du projet tutoré : implantations des banques coopératives en France. Cf. le compte rendu de la réunion.	
27/11/2020	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Rédaction du compte rendu. Recherche sur l'ESS et les banques coopératives. Mise en œuvre du cahier des charges.	
27/11/2020	Nathalie TANG Maxence MASSON	Réalisation du tableau QQOQCCP.	
27/11/2020	Manon ANDRE	Rédaction de l'introduction, du contexte et des informations du projet sur le cahier des charges. Élaboration du plan.	
27/11/2020	Joséphine BAUD	Réalisation du tableau des risques, réécriture des contraintes.	
27/11/2020	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Mise en commun des travaux effectués dans l'après-midi. Demande de réunion avec les tutrices.	
04/12/2020	Les tuteurs, les maîtres d'œuvres	Réunion pour la mise en route du projet. Cf. le compte rendu de la réunion	
04/12/2020	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Rédaction du compte rendu de la réunion. Recherche des variables pertinentes sur le site de l'INSEE Envoi d'un mail au commanditaire pour répondre à plusieurs questions formulées. Formulation de questions utiles pour la suite du projet.	

04/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Réunion pour départager les dernières tâches à effectuer pour le cahier des charges.	
04/01/2021	Nathalie TANG	Rédaction du compte rendu. Matrice des ressources humaines. Estimation du budget.	
04/01/2021	Manon ANDRE	Finalisation de la rédaction de la première partie du cahier des charges. Mise en page du cahier des charges.	
04/01/2021	Joséphine BAUD	Finalisation des risques, contraintes et solutions.	
04/01/2021	Maxence MASSON	Organigramme des tâches. Diagrammes de PERT et GANTT.	
05/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Mise en commun du travail effectué. Assemblage et finalisation du cahier des charges. Envoie du cahier des charges aux tuteurs.	
05/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Préparation du support pour la soutenance du cahier des charges. Chacun fait les slides correspondant à leur partie dans le cahier des charges.	
06/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Mise en commun des slides, finalisation du diaporama. Correction du cahier des charges. Dépôt du cahier des charges et du support pour la soutenance sur chamilo.	
07/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Finalisation du diaporama et entraînement oral.	

08/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Soutenance du cahier des charges.	
11/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Préparation du support pour la soutenance en Anglais. Chacun présélectionne des variables socio-économiques et commente le code R de l'année précédente. Mise en commun vendredi 15 janvier.	
12/01/2021	Maxence MASSON	Recherche des variables socio-économiques sur le site de l'INSEE et commentaire du code R	
13/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Préparation du support pour la soutenance en Anglais. Présentation de l'introduction de la soutenance de vendredi pendant le cours d'anglais.	
14/01/2021	Nathalie TANG	Recherche des variables socio-économiques sur le site de l'INSEE. Commenter les codes R.	
14/01/2021	Joséphine BAUD	Recherche des variables socio-économiques sur le site de l'INSEE.	
15/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Mise en commun de la sélection des variables socio-économiques sur le site de l'INSEE pour construire notre base de données.	
18/01/2021	Manon ANDRE Joséphine BAUD Nathalie TANG	Envoie de la liste des variables présélectionnées au commanditaire. Mise en commun du code R. Pour mercredi soir au plus tard : chacun aura fait les slides du diapo correspondants à leur partie.	On s'est aperçu que les url utilisés l'année précédente n'existent plus. De plus, la structure des coordonnées des banques a changé. Nous devons réfléchir à un autre code car le code du scraping était adapté à l'ancienne structure des sites web concernant les coordonnées.

20/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Finalisation du support pour la soutenance en anglais.	
21/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Soutenance en anglais durant l'International Week	
22/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Réunion suite aux questionnements rencontrés avec les scripts R avec Mme Letué.	
26/01/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Réunion avec la commanditaire pour la sélection des variables socio-éco.	
17/02/2021	Maxence MASSON	Amélioration du code crédit agricole	
12/02/2021	Joséphine BAUD Nathalie TANG	Vérification de l'année des données (2017)	
12/02/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Classification des données en trois catégories et répartition des codes R de chaque banque	
26/02/2021	Joséphine BAUD Maxence MASSON Nathalie TANG Manon ANDRE	Amélioration des codes de banques	
01/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Complément des données socio-éco Insee suite à la réponse d'Amélie ARTIS	
01/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Prise d'information autour du code R avec Mme Letué	
03/03/2021	Nathalie TANG	Lancement du code pour Crédit mutuelle	Deux échecs, problème dans le codage

05/03/2021	Maxence MASSON	Adresses des banques du Crédit Agricole collectées	
05/03/2021	Joséphine BAUD	Adresses des banques de la Banque Populaire collectées	
05/03/2021	Manon ANDRE	Adresses des banques de Société Générale collectées	
06/03/2021	Nathalie TANG	Adresses des banques du Crédit Mutuel collectées + nettoyage des données	
06/03/2021	Manon ANDRE	Nettoyage des données (adresses de banques de Société Générale)	
10/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Création d'une base de données avec les 4 banques Nettoyage des données socio-économiques Maquette d'un modèle en étoile pour l'entrepôt de données	
11/03/2021	Manon ANDRE	Réponse d'Amélie ARTIS, prise de rendez-vous pour une réunion	
13/03/2021	Maxence MASSON	Importation et nettoyage des données géographiques des banques sur phpMyAdmin	
14/03/2021	Joséphine BAUD Manon ANDRE	Définition de l'individu statistique et élaboration du plan du rapport final	
16/03/2021	Nathalie TANG	Nettoyage des données socio-économiques	reconversion ZE 2020-2010
17/03/2021	Nathalie TANG	Nettoyage des données socio-économiques	Création des ZE
17/03/2021	Joséphine BAUD	Rédaction du rapport (remerciements, introduction, résumés)	
18/03/2021	Nathalie TANG	Nettoyage des données socio-économiques	

18/03/2021	Manon ANDRE Amélie ARTIS Joséphine BAUD Nathalie TANG	réunion de mise au point sur les variables socio-économiques, les analyses et le plan du rapport	
19/03/2021	Nathalie TANG Joséphine BAUD Manon ANDRE	Nettoyage des données socio-économiques	
20/03/2021	Nathalie TANG Joséphine BAUD Manon ANDRE Maxence MASSON	Nettoyage des données socio-économiques	
20/03/2021	Nathalie TANG	Création d'un fichier Excel avec toutes les données + conversion en csv	
20/03/2021	Nathalie TANG Joséphine BAUD Maxence MASSON	Élaboration de statistiques descriptives sur les ZE et les banques et réalisation d'une première ACP	
20/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Rédaction du rapport Analyses	
21/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Rédaction du rapport Analyses Diaporama	
22/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Rédaction du rapport Analyses Diaporama	
23/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Rédaction du rapport Analyses Diaporama	
24/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Rédaction du rapport Diaporama Soutenance	

25/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Finalisation du rapport	
26/03/2021	Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Finalisation du rapport	

Légende	
	Réunions avec les tuteurs
	Réunions avec les membres du projet
	Réunion avec le commanditaire

Annexe 6 : Compte-rendu des réunions



Compte-rendu de la réunion n°1, 25/11/2020

Date et lieu : 25/11/2020 - visioconférence

Heure de début : 17h	Heure de fin : 18h
-----------------------------	---------------------------

Objet : Présentation du projet tutoré : implantations des banques coopératives en France

Présents :	Excusés :
<u>Tutrices :</u> Mme FAIN Mme LETUE	<u>Étudiants :</u> Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG
<u>Commanditaire :</u> Mme Amélie ARTIS	

La commanditaire, Amélie ARTIS, est maître de conférence en économie et enseignante-chercheuse à Science-Po-UGA. Dans le cadre de ses recherches, elle est amenée à répondre à la problématique :

« Pourquoi et comment expliquer la diversité des organisations/entreprises qui se développent dans notre système économique vu comme capitaliste ? »

Pour l'aider à répondre à cette question, elle fait appel à notre groupe de projet tutoré pour s'intéresser plus précisément au secteur bancaire coopératif. Ainsi nous sommes invités à étudier des questions plus précises :

- Est-ce qu'aujourd'hui l'essence historique des banques coopératives d'être proche des ménages est toujours observée ?
- Existe-t-il encore une diversité des organisations sur le territoire ?

Nos premières recherches se baseront sur de la recherche et de la sélection de données :

- Les variables socio-économiques sont à récupérer sur le site de L'Insee. Elles seront ensuite à trier selon leur pertinence. Il faudra enfin sélectionner une échelle territoriale pertinente : les zones d'emplois.
- Les données de localisation des banques seront à construire : mettre à jour les données du projet de l'an dernier et les comparer.

Par la suite, nous devrons mettre en place des procédures statistiques pour étudier la présence des banques coopératives. Cela nous mène à déterminer une typologie des banques coopératives.

Prochaines réunions : À définir – réunions avec les tuteurs pour	Objectifs : <ul style="list-style-type: none">- Essayer de s'imprégner du sujet (définitions, enjeux, tâches à prévoir, données)- Comment récolter les données ?- Faire une première ébauche du Cahier des charges
--	---

Compte-rendu de la réunion n°2, 04/12/2020

Date et lieu : 04/12/2020 - visioconférence

Heure de début : 14h

Heure de fin : 15h20

Objet : Réunion récapitulative et pistes de réflexion

Présents : <u>Tutrices :</u> Mme FAIN Mme LETUE	<u>Étudiants :</u> Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Ordre du jour : - Retour sur le compte-rendu de la première réunion - Aspect collecte, « scraping » et sélection des données - Gestion des données : où et comment les stocker - Utilisation des logiciels à distances - Attendus pour les prochaines échéances - Vision globale du projet
---	--	---

Retour sur le compte-rendu de la première réunion

Pour les futures réunions il sera souhaitable de plus structurer les compte-rendu et d'ajouter la notion d'ordre du jour.

Aspect collecte, « scraping » et sélection des données

Nous aurons accès au rapport des étudiants de l'an dernier pour la partie gestion des données, il contient notamment un code R en annexe que nous devrons améliorer et commenter pour que le fichier soit facile à lire pour les futurs étudiants.

Il faudra ensuite aller chercher sur le site de l'INSEE les données socio-économiques que nous trouvons intéressantes, il faudra trouver les indicateurs pertinents en fouillant sur le site. Nous pourrons nous diriger vers Amélie ARTIS pour savoir si nos choix sont pertinents.

Enfin nous devrons mettre en commun les données des banques avec les données socio-économiques pour pouvoir faire une étude statistique.

Gestion des données : où et comment les stocker

Il faudra créer une base de données et la stocker à plusieurs endroits différents. Il faudra faire des propositions à Amélie ARTIS pour voir sous quelle forme elle souhaite recevoir la base de données.

Utilisation des logiciels à distance

Nous devons lister les logiciels dont nous pensons avoir besoin pendant le projet et si nous n'y avons pas tous accès nous pouvons accéder aux salles 107 et 108 du bâtiment Michel Dubois via un VPN.

On peut également prévoir d'utiliser ces logiciels le plus tard possible au cours du projet car nous reviendrons sûrement en présentiel en février.

Il faudra donc intégrer cette contrainte au cahier des charges.

Attendus des prochaines échéances

Nous aurons une soutenance orale le 8 janvier 2021 qui portera autour du cahier des charges, nous devrons donc durant le mois de décembre nous concentrer sur sa rédaction et sur sa mise en forme au format PowerPoint. Ce fichier PowerPoint sera à déposer avant le jour de la soutenance.

Nous pourrons envoyer ce fichier PowerPoint à Amélie ARTIS avant la soutenance afin d'en discuter et d'avoir un premier retour sur notre cahier des charges.

Vision globale du projet

Lors de cette réunion nous avons fait émerger une ébauche de plan pour le projet :

- I) Récupération des données des banques + création base de données (« scrapping »).
- II) Recherche des données socio-économiques + lien avec les données des banques.
- III) Analyse des données.

Nous devrons nous approprier l'article d'Amélie ARTIS car notre étude sera similaire et cela nous permettra de trouver des pistes d'analyses et de mettre en lumière l'individu statistique.

L'avancée du projet devra se faire dans cet ordre :

- Savoir ce que nous voulons démontrer d'un point de vue économique
- Choisir en fonction les études statistiques que nous souhaitons réaliser
- En déduire les variables qui nous seront utiles
- Construire notre base de données

Après avoir fait ça nous pourrons enfin récolter les données sur les banques et les données socio-économiques dont nous avons besoin.

Prochaines réunions : <ul style="list-style-type: none">- Point sur le cahier des charges- Point sur les réponses de la commanditaire suite aux questions envoyées	Objectifs : <ul style="list-style-type: none">- Avancer le cahier des charges- Commencer à lire le code R de « scrapping »- Commencer à collecter des données pertinentes via le site de l'INSEE
--	---



Compte-rendu de la réunion n°3, 04/01/2021

Date et lieu : 04/01/2021 - visioconférence

Heure de début : 9h30

Heure de fin : 10h30

Objet : Réunion de mise en point du cahier des charges

Présents : <u>Étudiants</u> : Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG	Ordre du jour : - Mise en point du cahier des charges - Planification de la semaine du 4 janvier 2021
---	--

Mise en point du cahier des charges

Nous avons identifié les tâches à effectuer afin de finir le cahier des charges. Nous nous sommes ensuite partagé les tâches et chacun travaillera à sa convenance afin que les tâches soient finies pour le lendemain.

Planification de la semaine du 4 janvier 2021

- 4 janvier : Effectuer les tâches à faire pour finaliser le cahier des charges.
- 5 janvier : Mettre en commun du travail effectué la veille afin de finaliser et envoyer le cahier des charges aux tuteurs. Commencer le diaporama de la présentation orale du cahier des charges.
- 6 janvier : Continuer et finaliser le diaporama.
- 7 janvier : Envoyer le diaporama aux tuteurs et le déposer sur Chamilio avant midi. Commencer le projet ; lire, comprendre et commenter le code R de l'année précédente.
- 8 janvier : Soutenance du cahier des charges. Effectuer la liste des variables socio-économiques pertinentes et consulté le commanditaire.

Prochaines réunions : À définir.	Objectifs : - Finaliser le cahier des charges - Envoyer le cahier des charges aux tuteurs - Commencer le diaporama pour la présentation orale du cahier des charges
--	---



Compte-rendu de la réunion n°4, 22/01/2021

Date et lieu : 22 janvier 2021 en visioconférence

Heure de début : 9h **Heure de fin :** 9h55

Objet : Problème rencontré avec les scripts R

Présents :	Ordre du jour :
<u>Tutrices :</u> Mme LETUE	<u>Étudiants :</u> Manon ANDRE Joséphine BAUD Maxence MASSON Nathalie TANG

Les scripts R ne marchent pas, pour certains les liens URL de départ sont obsolètes.

Nous avons observé les codes sources des sites internet des banques afin de trouver l'accès aux adresses de ses dernières.

Nous devons créer une liste de département que nous pourrons parcourir avec une boucle dans les scripts.

Prochaines réunions :	Objectifs :
26/01/2021 avec la commanditaire	Créer la liste de département Modifier le code R pour inclure la liste



Compte-rendu de la réunion n°5, 19/03/2021

Date et lieu : 19 mars 2021 en visioconférence

Heure de début : 15h **Heure de fin :** 16h

Objet : Problème rencontré avec les scripts R

Présents :	Ordre du jour :
<u>Commanditaire :</u> Amélie ARTIS	<u>Étudiants :</u> Manon ANDRE Joséphine BAUD Nathalie TANG

Données socio-économiques

Les données socio-économiques liés à la caractérisation de la population sont trop lourdes et entraînent une perte des données conséquente. Mme ARTIS a suggéré de ne retenir seulement deux indicateurs : la pyramide des âges et les catégories socio-professionnelles. De plus elle nous fera parvenir le lendemain les fichiers en question pour gagner du temps de recherche.

Présentation des données géographiques

Amélie ARTIS est satisfaite des données récoltées.

Analyses statistiques

Après avoir demandé si une régression logistique conviendrait avec Mme ARTIS, nous avons convenu que le problème de granularité entre les données socio-économiques (à l'échelle de la zone emplois) et les données des banques (possible à l'échelle des communes) serait un trop gros frein à une étude statistiques efficace. Notre commanditaire nous a suggérée de mettre l'accent sur les statistiques descriptives et sur une ACP.

Plan du rapport

Après avoir présenté un plan trop descriptif de notre projet, Mme ARTIS nous a proposé un plan en trois parties avec une partie distincte sur les zones d'emploi, une partie distincte sur les banques, et enfin une partie où l'on croise ces deux informations pour répondre à notre problématique. Notre commanditaire nous a également suggéré de faire une partie méthodologie à part entière où l'on décrirait tous les aspects techniques, pour que notre développement ne laisse place qu'au résultats et aux analyses.

Prochaines réunions : Pas de prochaine réunion prévue	Objectifs : Terminer le rapport et les analyses statistiques
---	--

Annexe 6: Schéma de nettoyage des données

Exemple de flux Tableau Prep pour le nettoyage des variables sur les âges

