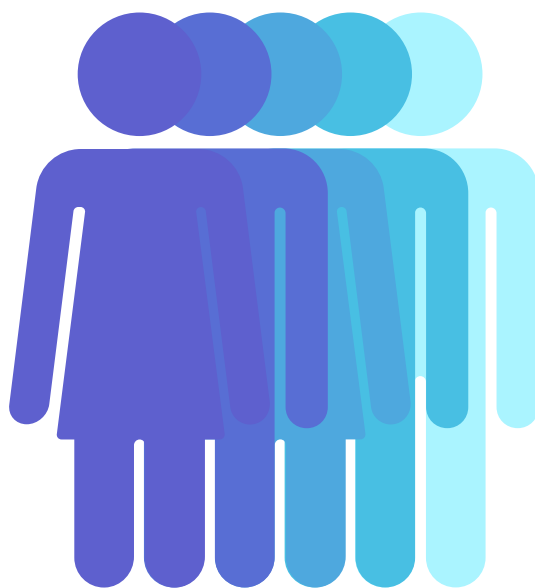


ÉTUDE DES PROFILS CLIENTS

CONNAISSANCE ET SEGMENTATION DE LA CLIENTÈLE



Réalisée par :

Carmen Andreea NITA – Chargée d'études marketing
Angéline BAERT – Chargée d'études marketing

A destination de :

Mme. Maureen DHONDT – Directrice marketing

Sommaire

Introduction.....	02
-------------------	----

Partie 1. Exploration Marketing

Les informations à notre disposition.....	06
---	----

Les caractéristiques de nos clients.....	07
--	----

Notre activité en quelques chiffres.....	09
--	----

Analyse croisée.....	12
----------------------	----

Retraitement des données.....	14
-------------------------------	----

Classification des clients.....	15
---------------------------------	----

Préconisations.....	18
---------------------	----

Partie 2. Focus Technique

Modélisation Statistique.....	21
-------------------------------	----

Explication des Décisions.....	23
--------------------------------	----

Introduction

La connaissance client est essentielle pour la personnalisation des offres et l'optimisation des campagnes marketing en fonction des besoins et des préférences des clients. Elle favorise la fidélisation en créant des relations solides et en anticipant la demande, ce qui réduit les coûts et renforce l'avantage concurrentiel. De plus, elle guide l'innovation en identifiant les besoins émergents du marché. Enfin, elle facilite la gestion des plaintes et des retours, contribuant ainsi à maintenir la satisfaction client et la rentabilité sur le long terme.

Contexte

Vous nous avez exprimé votre souhait d'établir un profil des clients se connectant au site web de BanqTout. Votre demande traduit un besoin évident d'identifier les différents types d'individus composant votre clientèle afin de pouvoir adapter, voire entreprendre des campagnes marketing correspondant aux clients.

Afin de répondre à votre demande, nous disposons d'une base de données que vous nous avez fourni. Cette base répertorie la liste **des**

internauts ayant visités le site web, ainsi que leurs caractéristiques.

Certains de ces internautes sont déjà clients et d'autres ne le sont pas. Cette dernière catégorie peut être considérée comme des clients potentiels qui sont intéressés par vos

offres de produits. Grâce à ces informations nous pourrons

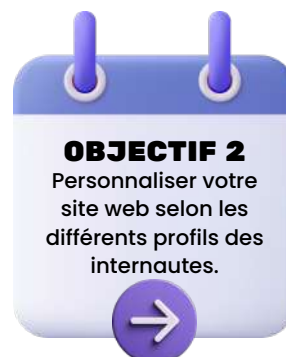
faire le nécessaire afin de satisfaire votre besoin.

Objectifs

Vous nous avez exprimé ce besoin afin de répondre à un double objectif :



Il est devenu indispensable de comprendre les besoins des clients mais surtout de savoir qui compose finalement cette clientèle. Identifier les clients et leurs caractéristiques rendra les actions côté marketing bien plus efficaces. C'est ainsi que l'on en vient au second objectif.



Comme précisé précédemment, l'identification des visiteurs sur le site web est une étape indispensable pour avoir une campagne marketing adaptée au type d'internaute.

La personnalisation du site internet est une excellente idée pour amener à concrétiser à l'achat, les internautes qui sont pour le moment simples visiteurs. En revanche, pour les internautes déjà clients, l'objectif sera de mettre en avant les autres produits pour lesquels ils pourraient avoir un intérêt particulier et donc inciter à l'achat.

Méthodologie

Pour ce faire, vous nous avez fourni une base de données comprenant une multitude de renseignements sur les individus (clients ou non) visitant votre site internet mais également sur leur crédit s'ils en ont un, le nombre de connexion au site ou encore la façon avec laquelle ils ont eu accès au site. Afin de répondre à votre besoin, nous avons choisi de procéder de la manière suivante.

Tout d'abord, nous devons bien évidemment prendre en main la base de données, et donc les informations qui nous seront utiles.

Ensuite, après avoir retraité la base, nous avons opté d'entreprendre une analyse des correspondances multiples (ACM) pour ensuite procéder à une classification ascendante hiérarchique (CAH) qui nous donnera les réponses à vos questions via une méthode de clustering. Cette méthode consiste à regrouper les clients selon certains critères déterminants. Vous pourrez retrouver les détails de nos choix dans une partie dédiée en fin de rapport (cf Explication des décisions).

Le rapport se compose en **deux temps**. Dans une première partie, nous vous donnerons une interprétation des résultats obtenus, accompagnée bien évidemment des préconisations associées que nous prodiguons en conséquence. Dans un second temps, nous mettrons en évidence notre méthodologie et nos choix relatifs à l'élaboration de ce projet.

PARTIE 01

Connaissance et Segmentation de la Clientèle : Exploration Marketing















Les informations à notre disposition

Avant toute analyse, une étape d'exploration est nécessaire pour une meilleure compréhension de nos données, ce qui nous permettra d'adopter la meilleure approche possible pour les traiter.

Exploration de données. Description et analyse du fichier source

Pour la réalisation de notre étude, nous disposons d'une base de données composée initialement de 50000 observations et 14 variables, y compris l'identifiant des enregistrements. L'unité d'observation pour laquelle les données sont recueillies est l'individu qui accède au site web de Banq Tout.

Les variables décrivent les individus en fonction de leurs caractéristiques (âge, revenu, situation familiale, etc.), mais aussi en fonction de leur activité relative aux services et produits proposés par Banq Tout (activité, montant du crédit, taux de remboursement, etc.). Ces variables sont détaillées ci-dessous.

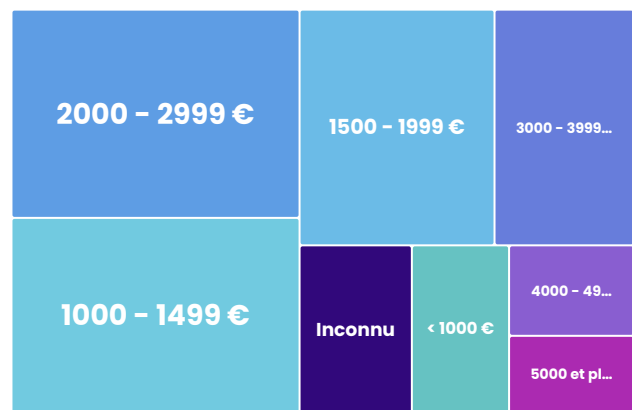
NOS CLIENTS	 Numéro Client Identifiant de l'internaute	 Situation familiale Situation familiale de l'internaute	 Logement Type du logement de l'internaute
	 Revenu financier Montant des revenus	 Age Tranche d'âge	 Département Département de l'internaute
BANQ TOUT	 Ancienneté Ancienneté du crédit principal détenu par le client	 Utilisation crédit Agrégat donnant le taux de remboursement du crédit principal	 Montant crédit Montant du crédit principal (donnée vide si l'internaute n'est plus client)
	 Assurance Crédit principal assuré ou non	 Connexions/mois Nombre de connexions web sur un mois	 Source web L'origine de la connexion web
	 Activité Activité du crédit principal (actif si la dette est positive, inactif si crédit remboursé)	 Type client Type de crédit détenu par l'internaute	

Les caractéristiques de nos clients

La compréhension approfondie des caractéristiques de notre clientèle est essentielle pour façonner efficacement nos stratégies commerciales. Cette section explore en détail les comportements et les préférences de nos clients, offrant ainsi un aperçu précieux qui guidera nos décisions futures et renforcera notre engagement envers leur satisfaction.

Revenu

L'évaluation des revenus des clients revêt une importance cruciale pour les banques, et une analyse approfondie de cette variable est essentielle pour notre étude.



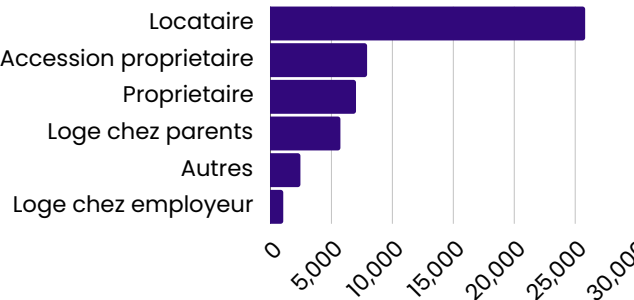
À la suite de cette analyse, une observation clé émerge : **une partie significative des visiteurs de notre site (24%) affichent un revenu relativement élevé**, situé entre 2000 et 2999€. Ensuite, une part importante, soit plus de 22% de la population, présente des revenus plutôt faibles, compris entre 1000 et 1499 €. Les revenus de 1500 à 1999 € représentent quant à eux plus de 18% de notre échantillon. Les revenus très élevés (au-delà de 3000 €) sont également bien représentés, constituant 21.5% de notre base (avec 13% ayant un revenu de 3000 à 3999 €, 4.5% pour l'intervalle 4000 à 4999 € et 4% pour les revenus de 5000 € et plus). Notons également qu'il existe une petite proportion de 7,5% pour les personnes dont le revenu est inconnu.

Observation

Il est notable que les individus bénéficiant de revenus relativement élevés sont ceux qui optent le plus fréquemment pour des crédits. Cette corrélation s'aligne de manière cohérente avec les critères traditionnels d'octroi de crédit, généralement liés au niveau de revenu.

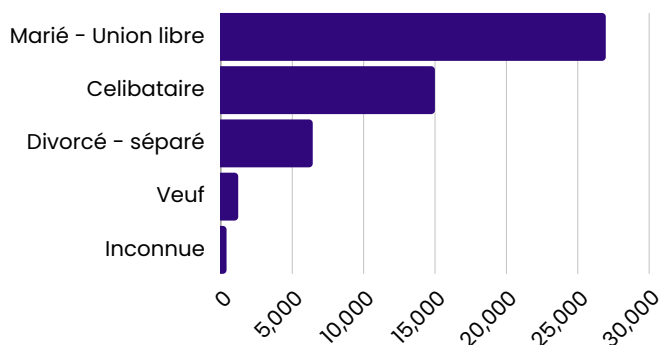
Logement

En ce qui concerne le logement de nos clients, nous avons identifié six catégories principales. Selon le diagramme à barres ci-dessous, il apparaît que plus de la moitié de nos visiteurs sont actuellement en location.



Les personnes **possédant leur propre logement** constituent **14%** de notre population, et les personnes qui sont en situation d'accession propriétaire constituent 16%. Les 18% restant se trouvent dans une situation de logement partagé, que ce soit chez leurs parents, leur employeur ou autre. Les locataires semblent avoir **une propension plus élevée à souscrire à des crédits** ou à d'autres services proposés par la banque.

Situation familiale



Concernant la situation familiale, les données révèlent que **plus de la moitié de nos visiteurs sont mariés ou en union libre**. Les célibataires représentent, quant à eux, 30% de la population, tandis que nous observons un pourcentage de 13% pour les personnes divorcées ou séparées. Les veuf(ve)s représentent 2% de la population, et on compte moins de 1% des individus pour lesquels la situation familiale est inconnue.

Observation

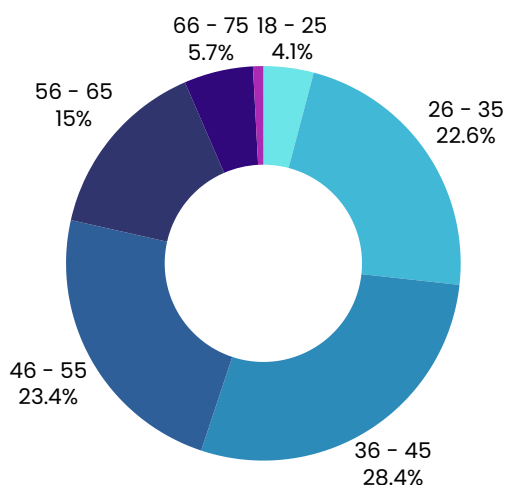
Cette distribution offre une information cruciale, mettant en lumière le fait que les individus en couple sont davantage enclins à recourir à des services de crédit. Cette constatation peut orienter nos stratégies marketing en ciblant spécifiquement cette catégorie de clients pour des offres de crédit.

de **36 à 45 ans**, ce qui en fait les catégories les plus prédominantes. En second lieu, les 46-55 ans représentent un peu plus de 20% de notre base, suivis par les visiteurs plus seniors, âgés de plus de 56 ans, qui constituent 15% de la population. Les jeunes de moins de 25 ans sont sous-représentés dans notre échantillon, représentant moins de 5%. De cette observation, nous pouvons déduire que **les individus âgés de plus de 26 ans et de moins de 45 ans sont les plus enclins à recourir à des crédits**. Cette tranche d'âge correspond généralement à des individus actifs sur le plan professionnel, disposant de revenus et aptes à souscrire à des crédits.

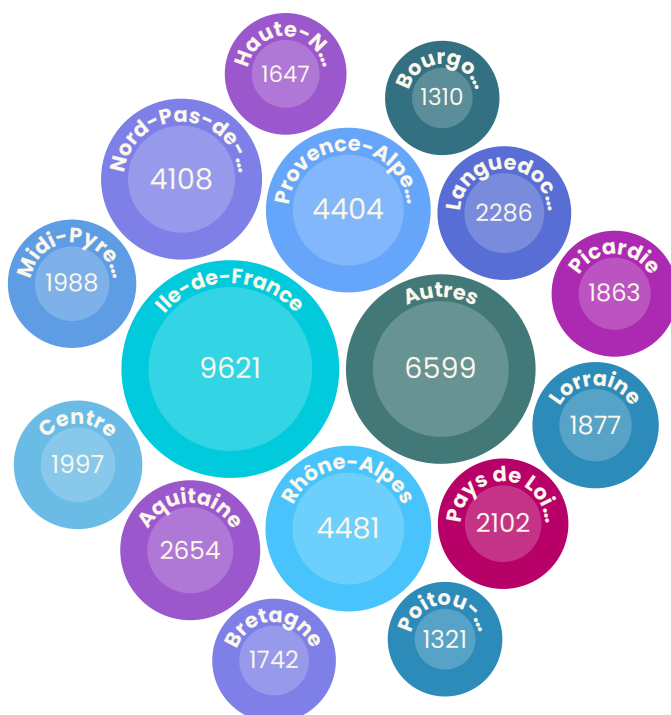
Départements

En ce qui concerne la localisation de nos clients par département, il est notable que les départements d'**Île-de-France, Rhône-Alpes, Provence - Alpes - Côte d'Azur, et Nord - Pas - de - Calais occupent les premières positions en termes de représentation** dans notre base. Ils contribuent respectivement à hauteur de 19%, 9%, 8,8%, et 8,2% de notre clientèle. Cependant, compte tenu de notre focalisation sur l'activité des clients sur le site web, cette information revêt une importance moindre dans le cadre de notre étude.

Age



D'après le graphique ci-dessus, il est apparent que **plus de la moitié des visiteurs** de notre site se situent dans les tranches d'âge de **26 à 35 ans** et

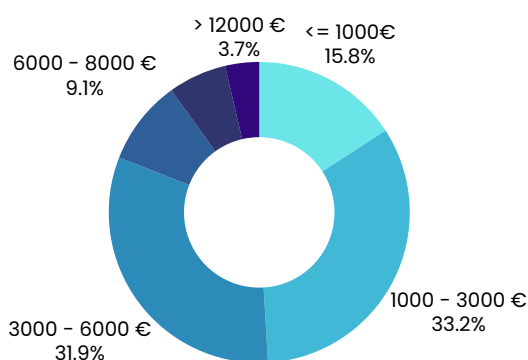


Notre activité en quelques chiffres

Cette section offre une perspective chiffrée de notre activité, fournissant une analyse concise mais détaillée de nos performances clés. En examinant les chiffres significatifs, nous visons à mettre en lumière les tendances, les réussites et les domaines d'amélioration, contribuant ainsi à une vision globale de la santé et de la croissance de BanqTout.

Montant du crédit

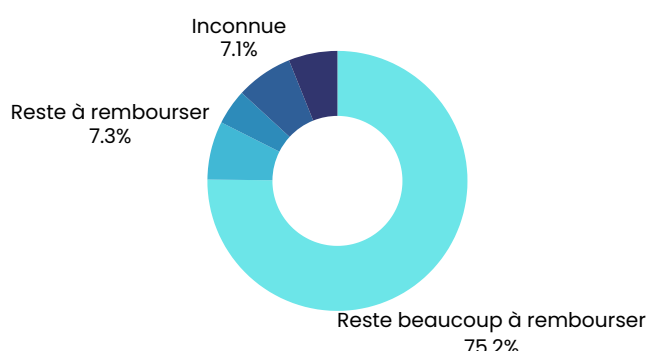
L'analyse des montants de crédit souscrits par nos clients révèle que, avec 31% de la clientèle préférant des crédits de 1000 à 3000 €, une demande notable pour des prêts de taille modérée se dessine, pouvant être associée à **des dépenses courantes, des projets spécifiques ou des besoins financiers ponctuels**.



Similairement, une proportion équivalente de 30% a opté pour des crédits plus élevés, compris entre 3000 et 6000 €, regroupant des individus avec des projets plus importants ou des besoins financiers spécifiques nécessitant des montants plus élevés.

Les clients favorisant des montants de crédit inférieurs à 1000 € représentent 15% de notre clientèle, possiblement caractérisés par des besoins financiers modestes mais trouvant avantageux de recourir à un crédit.

Remboursement du crédit

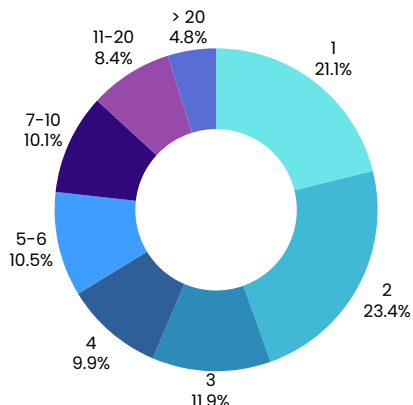


La prévalence marquée de la catégorie 'Reste beaucoup à rembourser' (75%) indique que la majorité de nos clients est encore engagée dans le remboursement substantiel de leur crédit principal, suggérant la nécessité de programmes de gestion financière et d'offres de refinancement. La catégorie 'Reste à rembourser' (7%) représente une phase intermédiaire, nécessitant des communications ciblées pour encourager la continuité des remboursements.

La réussite financière des clients dans la catégorie 'Crédit remboursé' (6%) offre une opportunité de **fidélisation**, tandis que les clients 'Reste peu à rembourser' (4%) pourraient être **incités à finaliser rapidement leur remboursement avec des offres spéciales**. Enfin, la catégorie où la situation de remboursement est inconnue (6%) requiert une enquête approfondie pour mieux comprendre la dynamique financière de ces clients, suggérant des initiatives pour collecter des informations manquantes.

Connexion mensuelle

L'analyse de la fréquence de connexion mensuelle de nos clients révèle une diversité significative dans leur **engagement avec notre plateforme en ligne**, ce qui est d'une importance forte dans le contexte de notre étude.



Une proportion notable de 35% se connecte régulièrement, entre 2 et 3 fois par mois, indiquant une utilisation fréquente de nos services. Un groupe de 30% se connecte de 4 à 10 fois, témoignant d'un engagement élevé, tandis que 21% se connectent seulement une fois par mois, reflétant un usage plus sporadique. Les clients **particulièrement actifs**, se connectant plus de 20 fois mensuellement, représentent 5% de notre base.

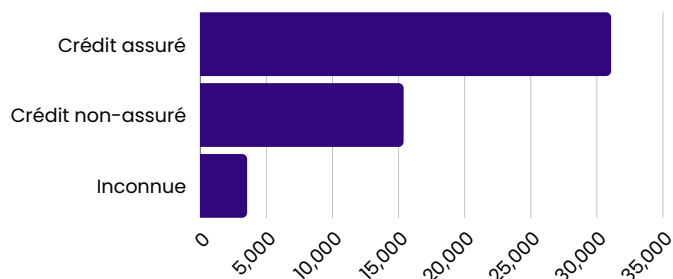
Observation

Ces résultats soulignent l'importance d'adapter nos stratégies de communication et d'offres en fonction des différents niveaux d'engagement, cherchant à maximiser la satisfaction client tout en encourageant une utilisation plus régulière de nos services en ligne.

Assurance du crédit

L'analyse révèle une distribution significative parmi nos clients en matière de souscription à des crédits avec ou sans assurance. La majorité, représentant **62%**, a choisi un crédit assorti d'une assurance, soulignant une sensibilité accrue à la protection financière liée à leur emprunt. Cette préférence suggère une **préoccupation pour la sécurité financière en cas d'imprévu**.

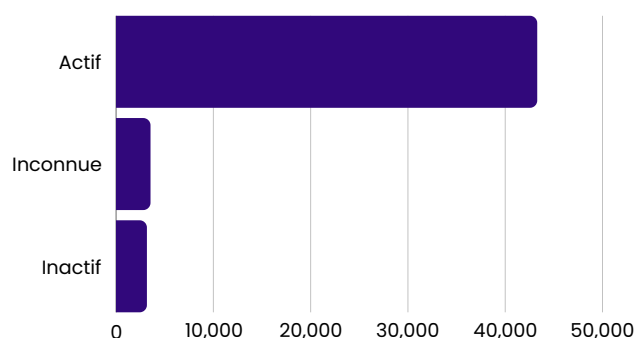
À l'inverse, 30% ont opté pour un crédit non-assuré, reflétant une confiance particulière dans leur situation financière ou une aversion personnelle à l'ajout de couvertures supplémentaires. Cette diversité souligne l'importance d'offrir des options flexibles pour répondre aux besoins individuels.



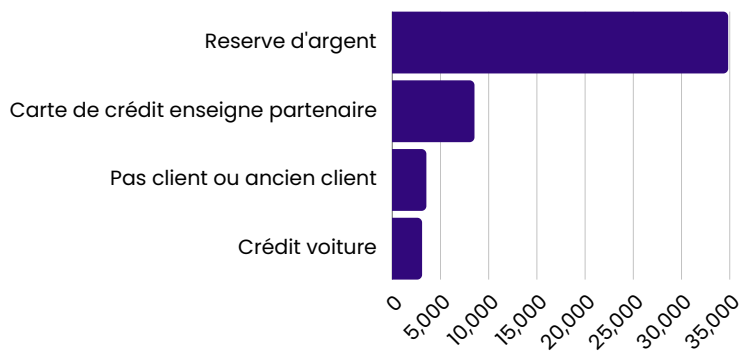
Activité

L'analyse de la répartition des visiteurs de notre site web offre des insights significatifs sur l'engagement de notre audience. Avec une prédominance écrasante de 86% de clients actifs, la fidélité envers BanqTout est apparente, soulignant l'importance cruciale de cette base pour la stabilité.

La présence de 6% de clients inactifs ou d'anciens clients indique une catégorie nécessitant une attention particulière, suggérant des opportunités pour des stratégies de réactivation ou des offres personnalisées. La portion de 7% avec une activité inconnue met en avant l'importance de collecter des données supplémentaires pour mieux comprendre leur comportement, offrant ainsi la possibilité d'une personnalisation accrue de nos interactions pour répondre à leurs besoins spécifiques.

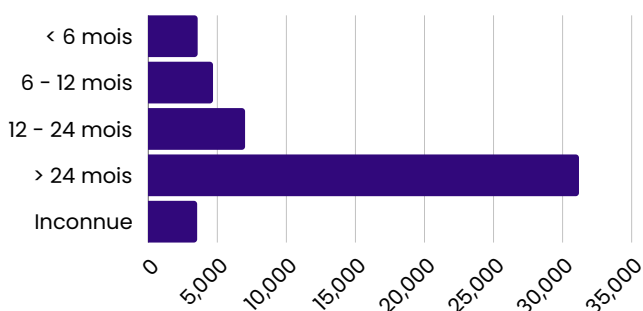


Type de crédit



Le diagramme ci-dessus offre une représentation détaillée de la composition des visiteurs de notre site, les classant selon leurs interactions avec les services bancaires. Une observation majeure révèle que **70%** de nos visiteurs sont des clients utilisant principalement une **réserve d'argent**, soulignant la **fidélité à ce service particulier**. Par ailleurs, 17% de notre audience détient une carte de crédit provenant d'une enseigne partenaire, mettant en lumière l'impact positif des collaborations sur notre portefeuille client. Avec 6% souscrivant au crédit voiture, une niche spécifique de besoins financiers est identifiée au sein de notre base d'utilisateurs. Enfin, les anciens clients et les prospects, représentant conjointement 7% des visiteurs, suscitent l'intérêt pour une exploration approfondie afin de comprendre les motivations des anciens clients et les attentes des prospects.

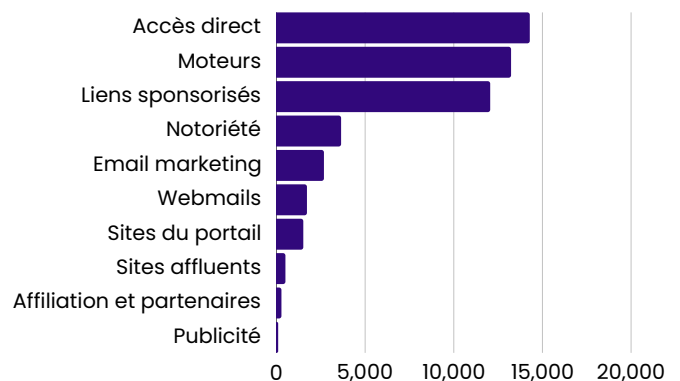
Ancienneté



La catégorie de plus de 24 mois, majoritaire à 62%, reflète une loyauté significative des clients avec une **relation établie de plus de deux ans**, les positionnant comme des **ambassadeurs potentiels** bien versés dans nos services. La tranche de 12 à 24 mois (14%) représente une

phase de transition offrant des opportunités pour renforcer l'engagement à long terme. Avec une ancienneté de 6 à 12 mois (9%), ces clients sont encore dans une phase récente, justifiant la mise en place de stratégies d'engagements spécifiques. Les clients de moins de 6 mois (7%) offrent une opportunité stratégique pour des initiatives de bienvenue et des incitations visant à favoriser leur fidélité. Pour la portion de clients dont l'ancienneté est inconnue (7%), la collecte de données supplémentaires est essentielle pour mieux comprendre leur comportement et adapter nos approches en conséquence.

Source de la connexion



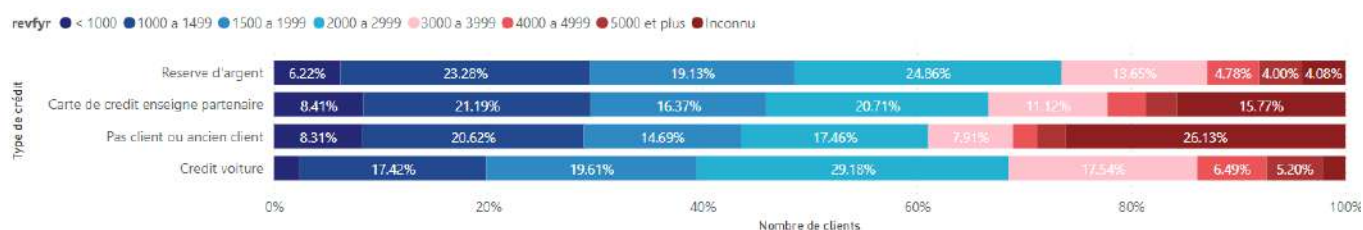
La catégorie "Accès direct" à 29% indique qu'environ un tiers des clients connaissent préalablement notre site et préfèrent y accéder directement, soulignant **l'importance de maintenir une présence en ligne mémorable**. Les 26% arrivant via les moteurs de recherche soulignent l'importance de l'optimisation pour un positionnement optimal dans les résultats de recherche. Les "Liens Sponsorisés" à 24% mettent en lumière l'impact significatif des campagnes publicitaires payantes dans la génération de trafic. La connexion via la notoriété à 7,29% souligne le rôle crucial de la réputation de notre marque, résultant d'efforts de relations publiques, de recommandations positives ou de la reconnaissance de notre marque dans l'industrie. Les 5% provenant de campagnes d'email marketing mettent en évidence l'efficacité de nos communications par email pour stimuler l'engagement en ligne, suggérant la nécessité de continuer à optimiser ces campagnes pour maximiser leur impact.

Analyse croisée

Dans la section suivante, nous procéderons à une analyse croisée entre nos chiffres et les chiffres de nos clients. Cette approche vise à identifier des tendances et des opportunités stratégiques pour orienter nos futures décisions commerciales.

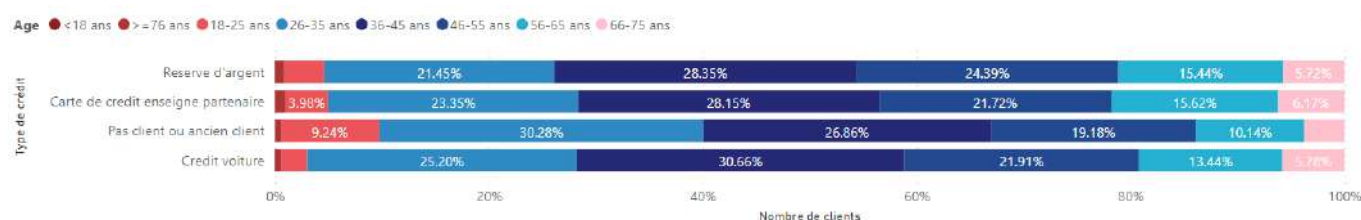
Type de crédit en fonction du revenu

L'analyse des statistiques descriptives révèle des tendances significatives parmi les détenteurs de différents types de crédit en fonction de leurs revenus.



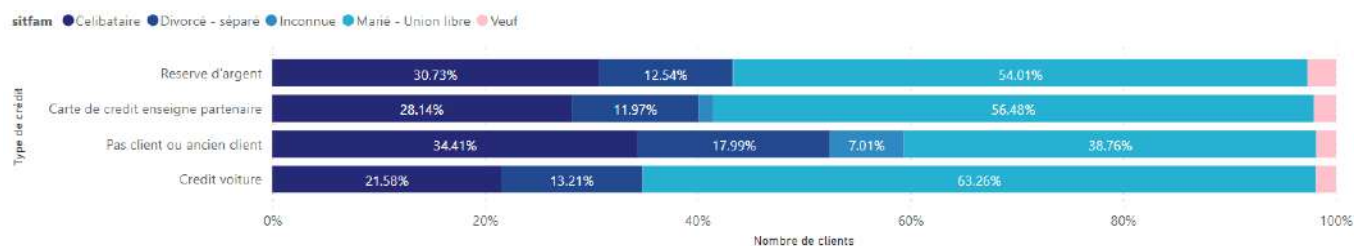
Les clients ayant des **revenus** compris entre **2000 et 2999 €** (24.86%) montrent une préférence marquée pour la **réserve d'argent**. une concentration significative est également observée dans la tranche de revenus de 1000 à 1499 € (23.38%). La **carte de crédit enseigne partenaire** affiche une répartition équilibrée entre les tranches de **1000 à 1499 € et de 2000 à 2999 €** (21% chacune), ainsi que 16% dans la tranche de 1500 à 1999 €, indiquant une accessibilité élargie pour ce type de crédit. Pour le crédit voiture, une nette préférence émerge pour les revenus de 2000 à 2999 € (29%), suivis par les tranches de 1500 à 1999 €, 1000 à 1499 € et 3000 à 3999 € avec des pourcentages significatifs (19%, 17%, et 17% respectivement).

Type de crédit en fonction de l'âge



Le crédit voiture est principalement souscrit par la **catégorie d'âge 26-35 ans**, affichant un **comportement similaire à la tranche d'âge 36-45 ans**. Les plus jeunes, âgés de 18-25 ans, sont davantage enclins à souscrire à la carte de crédit enseigne partenaire, tout comme les personnes âgées de plus de 56 ans, probablement pour l'accès à des offres promotionnelles, des réductions ou des récompenses attrayantes. Quant à la catégorie d'âge 46-55 ans, elle semble présenter une préférence marquée pour le type de crédit réserve d'argent.

Type de crédit en fonction de la situation familiale



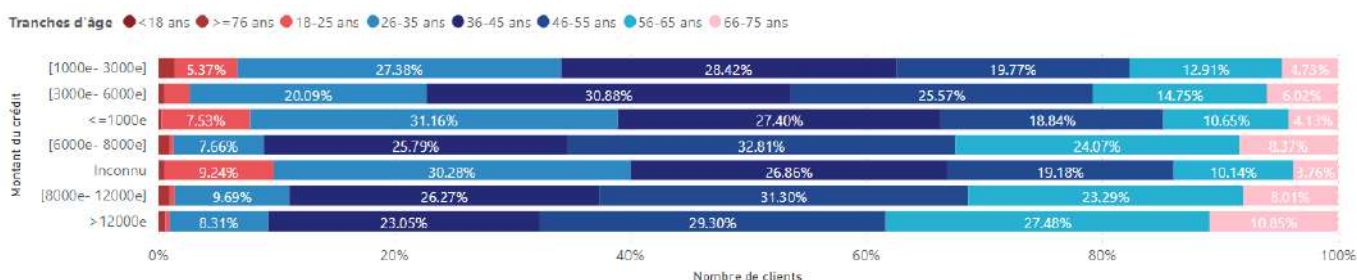
La souscription aux différents types de crédit varie en fonction de la situation familiale des individus. Les personnes mariées/remariées sont les principaux souscripteurs du crédit voiture, représentant 63% de ceux qui optent pour ce type de crédit. Ces personnes pourraient être motivées par des engagements familiaux tels que **l'achat d'une voiture familiale** ou **des besoins de mobilité accrus** en raison des responsabilités familiales. Ils affichent également une préférence marquée pour la carte de crédit enseigne partenaire, représentant plus de la moitié de cette catégorie avec un pourcentage de 56%. En revanche, **les célibataires privilégient la réserve d'argent**, constituant 31% de cette catégorie. Les personnes divorcées/séparées montrent une propension plus élevée à souscrire au crédit voiture, représentant 22% de cette catégorie.

Montant du crédit en fonction de la situation familiale



Les individus **mariés/remariés** affichent une propension significative à souscrire à **des crédits dépassant les 12 000 €**, représentant 74% de cette catégorie. Cette préférence suggère des engagements financiers importants liés à des projets majeurs. Dans la tranche de crédits de 1000 à 3000 €, les mariés/remariés restent majoritaires à 52%, tandis que les célibataires représentent 33%. En revanche, les divorcés montrent une préférence pour les crédits compris entre 3000 et 6000 €.

Montant du crédit en fonction de l'âge



Les personnes **les plus jeunes de notre clientèle**, regroupant ceux de moins de 18 ans, de 18 à 25 ans, et de 26 à 35 ans, optent principalement pour des **crédits de valeur modeste**, soit inférieurs à 1000 € ou compris entre 1000 et 3000 €. En revanche, **les individus plus âgés**, de 56 ans et plus, montrent une **préférence pour des crédits dépassant les 6000 €**. Les clients d'âge moyen, situés entre 36 et 55 ans, sont plus enclins à souscrire des crédits dont la valeur se situe entre 3000 et 8000 €. Ces tendances d'emprunt reflètent une diversité de préférences et de besoins financiers au sein de différentes tranches d'âge au sein de notre clientèle.

Retraitement des données

Données gardées après l'analyse préliminaire

Après analyse, il n'y a aucune donnée manquante. La base de données ne nécessite donc aucun traitement à ce sujet. Il s'agit désormais d'étudier les valeurs étant potentiellement aberrantes. Nous avons pu remarquer un fait qui nous semblait être aberrant : des clients âgés de moins de 18 ans. Or, cette situation n'est pas possible car il est légalement interdit de souscrire à un prêt à son nom en étant mineur. Par conséquent, toutes les personnes dans cette situation sont retirées de la base de données pour inexactitude de l'information.

Concernant **le choix des variables**, nous avons décidé de ne pas conserver certaines d'entre elles jugées **non discriminantes** : le département de l'internaute, le nombre de connexions par mois et l'origine de la connexion web. D'autres variables ont été retirées du modèle en raison des informations similaires à celles que dévoilent d'autres variables du modèle. Vous retrouverez les justifications de ces choix dans la partie technique dédiée.

Finalement, après nos retraitements, nous nous retrouvons avec les variables suivantes dans notre modèle pour procéder au clustering.

Variable explicatives	Libellé
num_cli	Identifiant client
typ_client	Type de crédit détenus
montant_credit	Montant du crédit principal
anciennete	Ancienneté du crédit principal
utilisation_credit	Agrégat donnant le taux de remboursement du crédit principal
assurance	Crédit principal assuré ou non
logmt	Type de logement
revfyr	Montant des revenus
tr_age	Âge
sitfam	Situation familiale

Certaines modalités ont fait l'objet de regroupements :

- Montant des revenus : >3000 : "3000 à 3999", "4000 à 4999" et "5000 et plus"
- Montant du crédit principal :]6000e- 12000] : "]6000e- 8000e]" et "]8000e- 12000e]"
- Ancienneté : <12 : "- de 6m" et "6-12m"
- Âge : >=66 ans : "66-75 ans" et ">=76 ans"

Nous vous invitons à vous référer à la partie technique pour avoir les explications relatives à nos décisions.

Classification des clients

Afin de pouvoir répondre à votre demande, nous avons donc initié une Analyse des Correspondances Multiples (ACM) pour ensuite procéder à une Classification Ascendante Hiérarchique (CAH). Notre analyse se base sur un échantillon de la population globale de 49 997 internautes. Cet échantillon correspond à 20% de cette base. Des détails sur la performance de notre modèle vous sont donnés dans la partie technique.

Ainsi, nous pouvons immédiatement passer à la présentation des résultats. Notre modèle montrait clairement qu'une division en quatre groupes était optimale. On retrouve donc 4 clusters d'internautes distincts. Le premier est composé de 3242 internautes. La seconde classe contient 3835 internautes et est celle regroupant le plus de personnes. Pour le troisième cluster, nous avons 723 individus et enfin 2200 pour l'ultime cluster. Nous allons désormais voir les caractéristiques de ces classes. Des modalités ressortent-elles davantage pour une certaine catégorie d'internautes ? C'est ce que nous allons voir ici.

Cluster 1

La première classe créée est **principalement composée de clients détenant une réserve d'argent**. Ces clients représentent en effet 85.38% de la classe. Pour la plupart des internautes, le montant de leur crédit principal est supérieur à 3 000€. On retrouve 70% des internautes de ce groupe avec un crédit entre 3 000 et 12 000€ (38.96% pour '3000e-6000' et 32.26% pour '6000e-12000').

Il est important de noter également que 88.89% des 10 000 internautes de l'échantillon ayant un crédit supérieur à 12 000€ se trouve dans ce cluster 1. Ces crédits sont d'ailleurs majoritairement assurés (74.09%). Ensuite, plus de 75% des individus du cluster sont clients depuis plus de 2 ans et ont encore la quasi-totalité de leur crédit à rembourser. Concernant leur logement, on y retrouve principalement des locataires mais également des propriétaires (accession propriété et propriétaire). Nous recensons dans cette première classe **54.6% des propriétaires de l'échantillon et 66.19% des personnes en accession à la propriété**. En outre, on retrouve donc plus de 55% de propriétaires dans notre cluster et 40.8% de locataires.

Ces internautes sont d'ailleurs une grande majorité mariés (87%) et ont entre 36 et 55 ans (62% au cumul). Enfin, le montant de leur revenu est bien souvent supérieur à 2 000 €. Ce cluster contient 71% des internautes ayant un revenu supérieur à 3000€ et 51% de ceux ayant un revenu entre 2 000 et 2 999€.

Cluster 1

En bref, il s'agit donc ici visiblement de personnes assez aisées, étant mariés et ayant déjà investi dans un bien immobilier. Leur situation leur permet de souscrire à un emprunt bancaire d'un montant faisant partie des plus élevés. Ils sont détenteurs pour la plupart d'une réserve d'argent et clients depuis plus de 2 ans désormais.

Cluster 2

Tout comme la première classe, les clients de ce groupe **sont 93.59% à posséder une réserve d'argent**. Contrairement au précédent, les montants des crédits sont plus faibles. Globalement, le montant des crédits se situe entre 1 000 et 6 000€ (35.38% pour '1000e-3000' et 39.95% pour '3000e-6000'). À nouveau, le montant à rembourser de ce crédit est encore élevé (93.9% pour '+++') et ce dernier est la plupart du temps assuré (80.76%).

La distinction d'ancienneté des clients est plus fine qu'auparavant avec 58.36% des personnes étant clients depuis plus de 2 ans. Du côté des caractéristiques personnelles, on retrouve dans cette classe 63% des internautes de l'échantillon ayant entre 18 et 25 ans et 56% de ceux ayant 26 et 35 ans. 68% des personnes de cette classe ont moins de 45 ans. **Leurs revenus sont parmi les plus faibles** avec plus de 70% gagnant moins de 2000€. La moitié des personnes gagnant moins de 1000€ sont affectées à ce cluster.

Enfin, ils sont une majorité de célibataires (51.6%) et les personnes divorcées, séparées sont souvent mis dans ce groupe également (50.8% des divorcés sont dans ce cluster). Le client est soit locataire (66%), soit il loge encore chez ses parents (21% du cluster et 71.4% des personnes dans ce cas dans l'échantillon total sont dans ce cluster dans cette situation).

Cluster 2

Ce cluster est un regroupement d'internautes relativement jeunes et célibataires, ayant plutôt un revenu modeste et n'ayant pas encore d'investissement immobilier concret. Leurs crédits sont plutôt faibles et ils possèdent une réserve d'argent pour la plupart.

Cluster 3

Cette classe est composée, sans aucun doute, des internautes n'étant pas clients de BanqTout. On y retrouve en effet 100% des internautes **non clients** de notre échantillon. Quant aux aspects personnels, ils sont 51% à vivre en location et 40% à être mariés. **Les distinctions sont plus vastes** car il s'agit d'un regroupement prenant en compte comme **seule caractéristique le fait d'être non client**. On y retrouve donc des profils différents avec une proportion quasiment identique.

Cluster 3

Dans ce troisième cluster, on retrouve des profils différents avec une proportion quasiment identique.

Cluster 4

Enfin, pour l'ultime groupe de notre classification, nous retrouvons 91.3% des 10 000 internautes possédant une **carte de crédit**. Les montants des crédits sont les plus faibles avec **plus de 90% du cluster avec un crédit inférieur à 3 000€**. Ces crédits sont pour la plupart non assurés (69.14%).

À nouveau, ces internautes sont clients depuis plus de 2 ans pour la plupart (69%). La somme qu'ils leur restent à rembourser est plutôt dérisoire : 91% des '+' et 65% de '++' sont envoyés dans cette classe. Les informations à caractère personnel nous renseignent que ces clients sont assez différents les uns des autres. On retrouve aussi bien des personnes mariées (48%), des célibataires (30%) et 76.8% des veufs sont affectés à ce cluster.

Au niveau de l'âge, la distinction n'est pas plus évidente non plus. On retrouve une répartition quasiment parfaite, avec cependant 50% des personnes ayant entre 26 et 45 ans. On y retrouve également des locataires, des propriétaires, et **le montant des revenus est plutôt dans la moyenne basse de notre échantillon** : 40% gagnent moins de 1 500€, 38.8% entre 1 500 et 3 000€ et 16% pour un revenu supérieur à 3 000€. À noter que 36% d'entre eux ne désiraient pas divulguer leur revenu.

Cluster 4

Ce dernier cluster distingue davantage les internautes par rapport au type de crédit détenu (ici, la carte de crédit) plutôt qu'aux caractéristiques personnelles. Il s'agit davantage d'internautes avec donc une carte de crédit, un montant de crédit faible et ce dernier n'étant pas assuré.

Préconisations

Si nous devons résumer les caractéristiques de chacun d'entre eux selon les 9 variables conservées, nous avons ceci :

Profil 1. Profils Aisés et Bien Établis

Le cluster 1 regroupant plutôt des personnes avec un revenu assez riche et ayant déjà souscrit aux prêts les plus élevés, nous vous recommandons de **cibler ces clients par le biais d'une campagne de communication personnalisée** mettant en avant des offres de prêts importants ainsi que des avantages exclusifs pour les propriétaires et ceux mariés par exemple.

Même si presque 75% des clients de cette classe ont un crédit assuré, il est important de continuer à mettre l'accent sur la **protection financière** qui est garantie par nos produits d'assurance adaptés à leur niveau de crédit élevé.



Profil 2. Jeunes Actifs avec Revenus Modestes



Pour accompagner ces jeunes dans le développement de leurs projets personnels, il serait intéressant de **leur proposer des produits de crédit adaptés à leurs revenus**, moyennant pourquoi pas **des modalités de remboursement plus flexibles** par exemple.

Parce que le style de vie des internautes ne devrait pas impacter leur développement personnel, proposer des offres spéciales pour les personnes seules, célibataires pourrait être une solution pour ces clients qui ont peut être du mal à trouver des solutions adaptées à leur situation.

Profil 3. Non-clients Potentiels

Le cluster des non-clients est très clairement un territoire à conquérir. Ce sont **des clients potentiels** qu'il faut convaincre. Continuez à lancer des campagnes **marketing d'acquisition**.

Essayez de les personnaliser **en fonction des caractéristiques des clients** (majorité de mariés et célibataires, entre 26 et 45 ans...) présents dans cette classe pour maximiser la conversion de ces internautes. Procédez par ordre d'importance en visant dans un premier temps les internautes ayant les caractéristiques les plus fréquentes pour attirer un maximum de nouveaux clients se reconnaissant dans votre annonce.



Profil 4. Détenteurs de Cartes de Crédit avec Profils Variés



Pour cette dernière classe axée sur les clients ayant une carte de crédit non assurée, il faudrait inciter ces derniers à **souscrire à une assurance** en mettant en avant la **protection financière** qu'elle procure. Cette démarche se ferait idéalement via une **communication multicanale** car ce groupe présente une grande diversité de profils.

En suivant ces recommandations, BanqTout peut maximiser la satisfaction client, améliorer la rétention, et attirer de nouveaux clients en répondant de manière précise aux besoins de chaque segment. Vous pouvez par ailleurs lancer une campagne de récompenses pour les clients les plus fidèles. Ce dernier pourrait attirer des nouveaux clients et aider à la conservation des clients déjà conquis.

Afin d'analyser les impacts de ces démarches, n'hésitez pas à créer des métriques pour suivre le nombre de nouveaux clients, le nombre de souscriptions à une assurance etc.

Suivez de façon régulière les évolutions de ces différents clusters en surveillant les changements de comportements financiers des clients et non-clients de chaque segment afin d'ajuster les stratégies en conséquence.

PARTIE 02

Connaissance et Segmentation de la Clientèle : **Focus Technique**

Modélisation Statistique

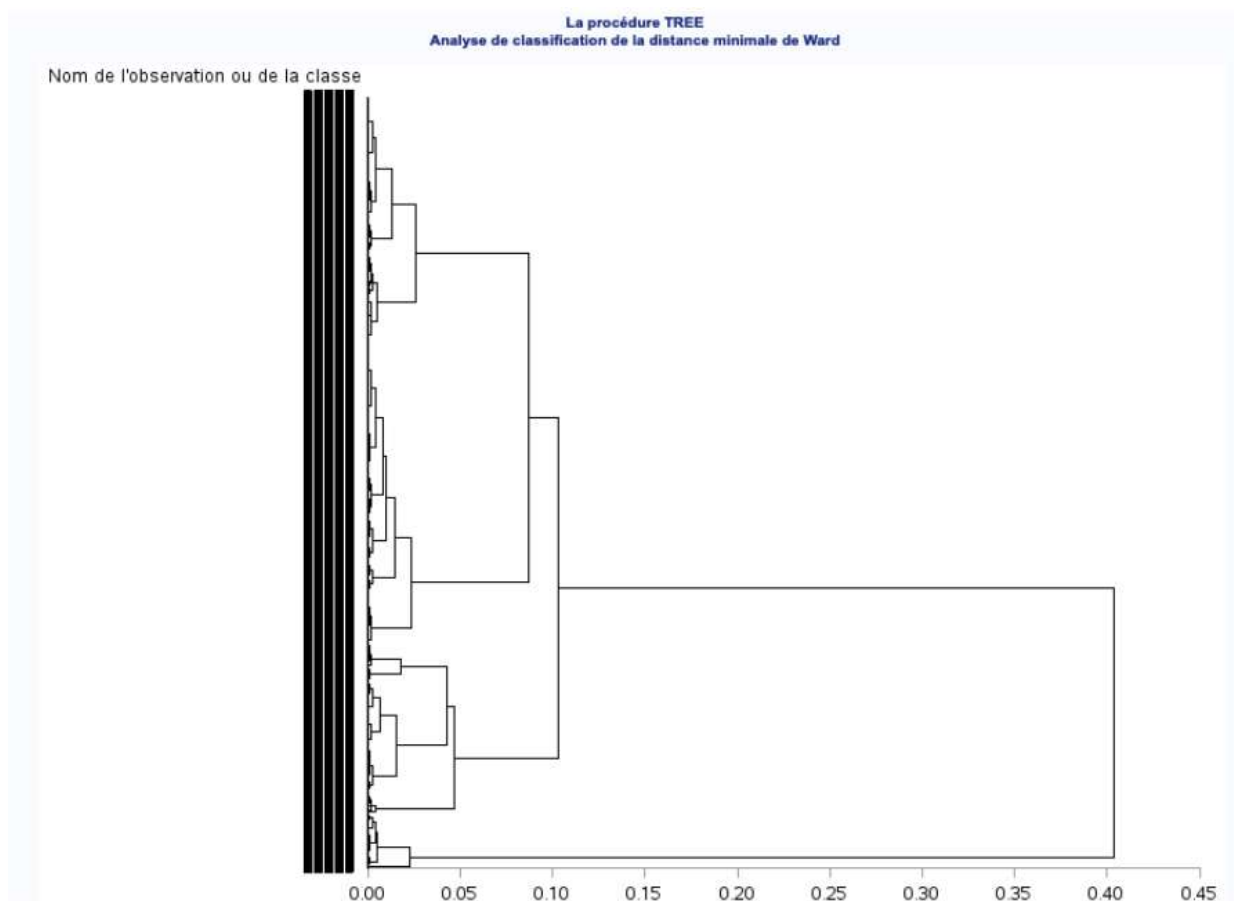
Données gardées après l'analyse préliminaire

Pour mener à bien notre étude, nous avons effectué une ACM puis une CAH. Notre population observée étant importante (50 000 internautes), nous avons pris un échantillon aléatoire de 20% de la base soit 10 000 internautes.

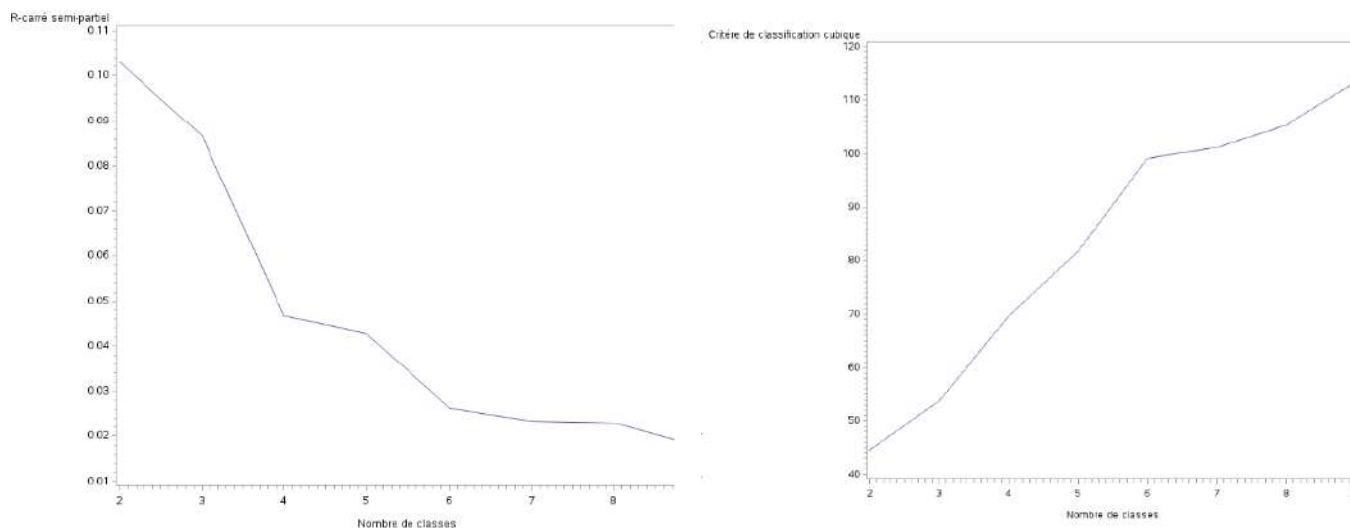
Notre base aléatoire est composée de plus de 50% d'individus mariés, 30% de célibataires et 13% de personnes divorcées. Le reste des internautes n'ont pas référencé leur situation maritale ou alors ils sont veufs. 75% de notre échantillon a entre 26 et 55 ans. Il y a 20% de personnes de plus de 55 ans et seulement 407 jeunes de moins de 26 ans. La plupart des 10 000 internautes ont une réserve d'argent (69%), contre 17% de carte de crédit et 6% environ un crédit de voiture. 723 d'entre eux sont non clients. L'ACM était indispensable dans la construction de la CAH.

De plus, cette analyse des correspondances multiples nous a permis de connaître les variables n'ayant aucun impact sur notre analyse et qui sont donc à retirer du modèle. Nous avons également pu identifier les regroupements potentiels entre certaines modalités de variable. Le détail de ces derniers se trouve dans la prochaine section.

Enfin, la classification ressemble à ceci :



On identifie un cluster assez petit, représentant probablement les internautes non clients. Ensuite, on a pu distinguer plusieurs autres clusters se divisant de manière plus ou moins égale. Afin de choisir le nombre de clusters dans lesquels notre population sera répartie, nous avons décidé d'étudier deux indicateurs mesurant la qualité de la classification : le R2 semi-partiel qui calcule la perte d'inertie interclasse provoquée en regroupant 2 clusters, et le critère de classification cubique qui permet de bien sélectionner le nombre de groupes.



Au vu de ces deux graphiques, nous avons décidé de sélectionner une division en 4 clusters pour notre classification. Les deux graphiques poussent à penser qu'il s'agit de la meilleure option pour la CAH de notre modèle. On aurait également pu penser qu'il aurait fallu prendre 6 classes, cependant, après analyse, les 4 clusters restent la meilleure option.

Explication des Décisions

Comme exprimé dans la partie plutôt métier, nous avons supprimé les observations où l'internaute avait moins de 18 ans car il est légalement impossible qu'une personne mineure souscrive à un prêt à son nom. Ensuite la suppression des variables relatives au département de l'internaute, au nombre de connexions par mois et à l'origine de la connexion web. s'explique par le fait qu'elles n'apportent aucune information sur les caractéristiques des internautes dans le cadre d'un clustering pour mieux connaître la clientèle.

Ces variables ne sont pas discriminantes et ne sont ainsi pas conservées dans notre modèle. Ces variables auraient pu être intéressantes dans le cadre d'une analyse des canaux de communication les plus efficaces. D'autres variables n'ont pas été conservées dans le modèle en raison de leur similitude. En effet, *active* et *utilisation_credit* indiquent tous deux des informations sur l'état de leur crédit. Cependant, *utilisation_credit* apporte plus de granularités et donc d'informations donc on conserve celle-ci.

La méthodologie pour l'ensemble des regroupements de modalités est la suivante : on observe à l'aide de l'ACM les modalités se trouvant à proximité. Suite à cela, il s'agit de les regrouper uniquement si ces modalités peuvent être regroupées selon nous, d'un point de vue métier. Il faut donc réfléchir d'un point de vue technique avec les positions sur les axes des différentes dimensions de l'ACM mais aussi d'un point de vue métier pour ne pas raisonner de manière contre-productive et avoir un modèle proche de la réalité.

Par exemple, si les 18-25 ans sont proches des plus de 75 ans, nous n'allons pas les regrouper car d'un point de vue métier, il semble illogique de le faire. Ces deux catégories ne représentent pas les mêmes besoins à ces âges et ne sont pas aux mêmes étapes en termes de projets de vie.

Pour rappel, les regroupements faits à l'aide de cette méthode sont les suivants :

- Montant des revenus : >3000 : "3000 a 3999", "4000 a 4999" et "5000 et plus"
- Montant du crédit principal :]6000e- 12000] : "[6000e- 8000e]" et "]8000e- 12000e]"
- Ancienneté : <12 : "- de 6m" et "6-12m"
- Âge : >=66 ans : "66-75 ans" et ">=76 ans"

À noter qu'un autre regroupement aurait pu être fait mais nous avons décidé de ne pas le faire. Selon la contribution et la position sur la dimension 2, les deux modalités "<1000" et "1000-1499" de la variable du revenu du foyer (*revfyr*) auraient pu être regroupées. Cependant, d'un point de vue métier, ces critères peuvent décrire deux types de personnes bien distinctes. Les personnes ayant un revenu <1000 sont des individus en contrat d'intérim ou à temps-partiel. L'intérim ne garantissant pas la sûreté de l'emploi, les chances d'obtenir un prêt sont plus minces.



***Merci pour
votre attention !***

CONTACT :

Andreea NITA - anita@banqtout.com
Angéline BAERT - abaert@banqtout.com

*hello@banqtout.com
www.banqtout.fr*