

Université Cheikh Anta Diop
Faculté des Sciences et Techniques
Département de Mathématiques et Informatiques



Master 1 Modélisation Statistique et Informatique

Analyse des Correspondances Multiples

Maigbibé

Présenté par :

MAYENE Bienvenue Schékina

Sommaire :

Introduction.....	2
I-/ PRESENTATION DES DONNEES.....	3
a) - Description de la base.....	3
b) - Statistiques descriptives sur la base de données.....	4
II-/ ANALYSE DES CORRESPONDANCES MULTIPLES.....	8
a) Choix des axes à retenir.....	8
b) Description du nuage des individus.....	9
c) Description du nuage des variables.....	12
d) Variables supplémentaires.....	15
III-/ INTERPRETATION DES RESULTATS.....	17
a) – Description du nuage des individu-variables.....	17
b) - Résumé des informations obtenues et Identification des groupes.....	17
Conclusion.....	19

Introduction :

L'Analyse des Correspondances Multiples (ACM ou MCA pour *multiple correspondence analysis*) est une extension de l'analyse factorielle des correspondances pour résumer et visualiser un tableau de données contenant plus de deux variables catégorielles.

Elle est aussi considérée comme une généralisation de l'analyse en composantes principales lorsque les variables à analyser sont catégorielles plutôt que quantitatives

C'est une méthode d'analyse factorielle par excellence car elle se prête bien à l'analyse des bases de données issues d'enquêtes ou de sondages avec pour objectif l'identification des groupes de personnes ayant un profil similaire dans leurs réponses aux questions

L'Analyse des correspondances multiples d'un tableau de données qualitatives correspond une AFC appliquée sur un tableau déduit du tableau de données initial, ce tableau est appelé **tableau disjonctif complet**. Dans ce tableau, une variable qualitative à k modalités est remplacée par k variables binaires, chacune correspondant à une des modalités.

L'ACM permettra donc d'étudier les ressemblances entre individus du point de vue de l'ensemble des variables, de faire un bilan des liaisons entre variables et également d'étudier les associations entre les catégories des variables

Aussi, à l'instar de l'Analyse des composantes principales et de l'Analyse factorielle des correspondances, les individus ou groupes d'individus (lignes) peuvent être caractérisés par les modalités des variables (colonnes). Si certaines variables du tableau de données sont quantitatives, il est possible de les intégrer dans l'analyse en les découpant en classes.

Notre travail portera sur la base de données nommée « **Credit** ». Il consistera tout d'abord à décrire cette base de données, ensuite réaliser l'Analyse des correspondances multiples (ACM) sur celle-ci avec le logiciel **R** en utilisant les packages *FactoMineR* (pour l'analyse) et *factoextra* (pour la visualisation des données) et enfin interpréter les résultats obtenus.

I-/ Présentation des données

a) Description de la base :

La base de données sur laquelle reposera notre travail est nommée « **credit** ». Cet ensemble de données représente 66 clients ayant souscrit un crédit à la consommation dans un organisme de crédit. Il y a un total de 66 observations sur 11 variables.

Le but de cette étude est de caractériser la clientèle de l'organisme de crédit. Nous voulons dans un premier temps mettre en évidence différents profils de comportements bancaires, c'est-à-dire effectuer une typologie des individus.

Nous voulons ensuite étudier la liaison entre le signalétique (CSP, âge, etc.) et les principaux facteurs de variabilité des profils de comportements bancaires (i.e. caractériser les clients aux comportements particuliers).

Signification des variables :

Les 11 variables qualitatives de notre jeu de données sont les suivantes :

[1] "Marche" : Motif de l'emprunt

[2] "Apport" : Apport personnel du client avant de réaliser l'emprunt

[3] "Impaye" : Nombre d'échéances impayées par le client

[4] "Assurance" : Type d'assurance contracté par le client (AID : assurance invalidité et décès ou Senior : pour les plus de 60 ans)

[5] "Endettement" :

[6] "Famille" : Situation Matrimoniale

[7] "Enfants" : Nombre d'enfants du client

[8] "Logement" : Conditions de logement du client

[9] "Profession" : Profession du client

[10] "Intitule" : Sexe du client

[11] "Age"

b) - Statistiques descriptives sur la base de données

L'objectif dans cette section sera d'inspecter et d'interpréter visuellement les profils des individus et des colonnes.

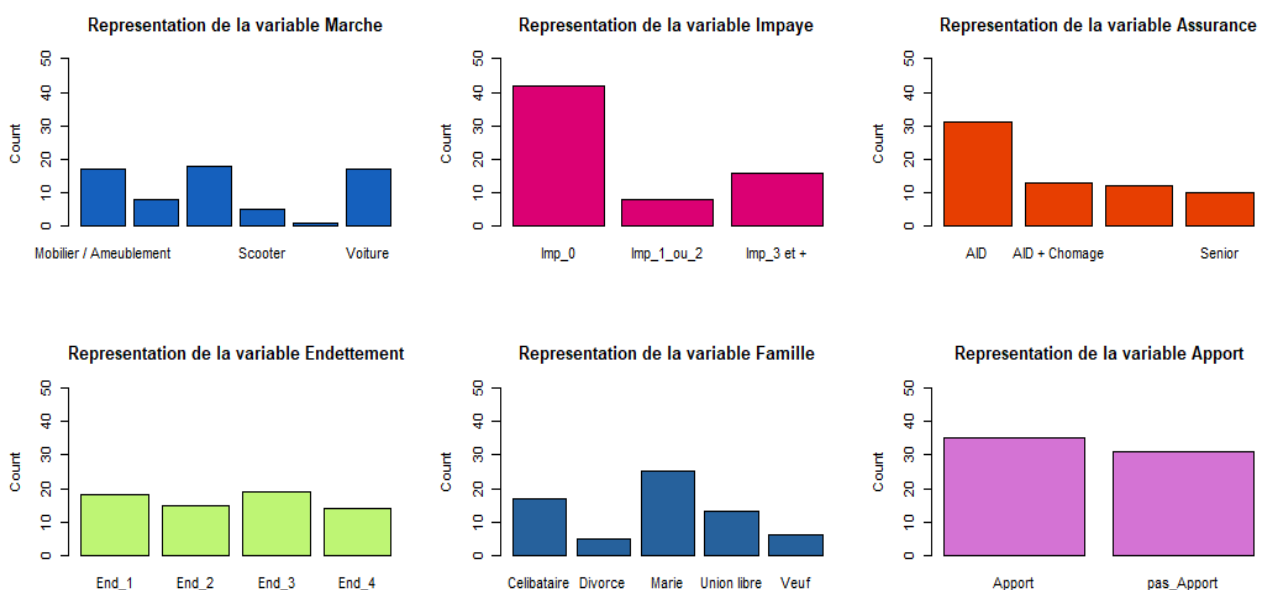
Avant de pouvoir faire des statistiques descriptives sur l'ensemble de la base de données, nous allons procéder au recodage de toutes nos variables qualitatives. En effet, en visualisant nos variables à travers la commande `str(credit)`, nous remarquons que nos variables qualitatives sont de type *character*.

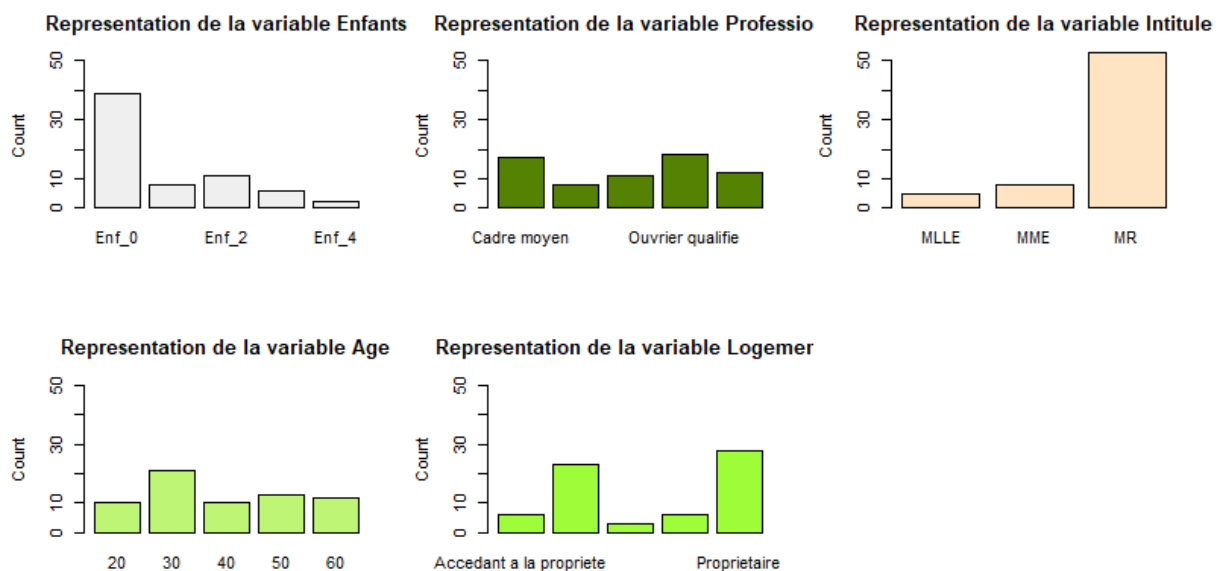
Si nous laissons ce type inchangé, nous ne pourrions pas mener correctement notre travail. Nous allons donc modifier le type de ces variables en type *factor*.

D'après les résultats de la fonction '*summary*', nous pouvons voir que la variable '*Age*' n'est pas considérée comme une variable qualitative. Il est donc nécessaire de la transformer comme telle.

En ACM, il est important de vérifier qu'il n'y a pas de modalités rares c'est-à-dire avec un faible effectif. En effet, les modalités rares ont une forte contribution dans la construction des axes et il est donc souhaitable d'en avoir peu au risque de biaiser nos interprétations.

Pour cela, nous préconisons de visualiser la fréquence des catégories des variables, nous obtenons les graphes suivants :





D'après les résultats obtenus et d'après l'observation des graphes ci-dessus, nous constatons que dans notre base de données un seul individu prend la modalité '*Side-car*' de la variable « *Marche* ». Nous constatons également qu'il n'y a que deux individus qui prennent la modalité '*Enf_4*' de la variables « *Enfants* ».

Lorsque certaines variables admettent des modalités à faibles effectifs, plusieurs solutions sont envisageables pour éviter que ces modalités influencent trop l'analyse :

- Regroupement naturel de certaines modalités : solution préconisée dans le cas de modalités ordonnées
- Ventilation (répartition aléatoire) des individus associés aux modalités rares dans les autres modalités : certains programmes permettent une ventilation automatique.
- Suppression des individus qui prennent des modalités rares.

Un *Side-car* étant physiquement plus proche d'une moto, Il est donc naturel de regrouper cette modalité avec la modalité '*Moto*' de la variable « *Marche* ».

Notons que nous ne ferons pas la même chose pour la modalité '*Enf_4*' de la variables « *Enfants* » car cette variable sera considérée comme une variable supplémentaire et donc la modalité '*Enf_4*' n'aura aucune influence dans la construction des axes factorielles.

Analyse Bivariée :

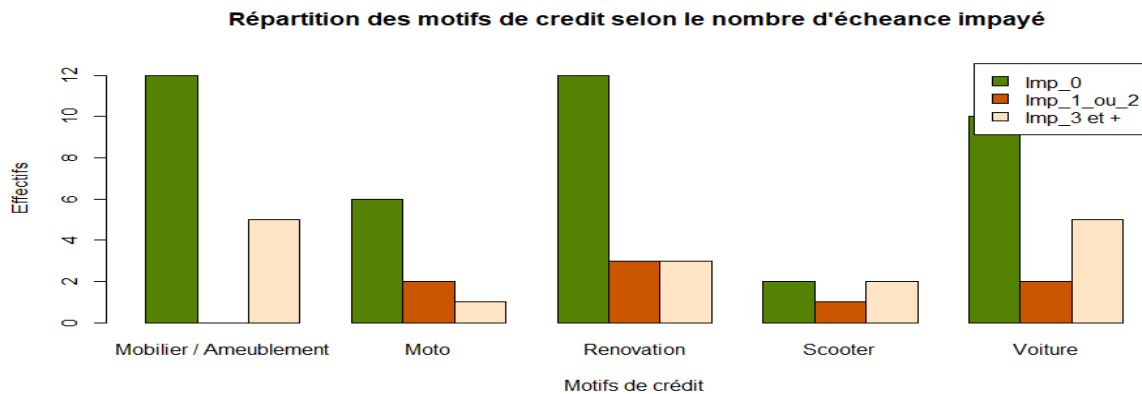
Dans l'analyse bivariée des variables de notre base, nous allons étudier la dépendance de certaines variables par rapport à la variable « *Impaye* »

- Première étude

Dans cette première étude, nous cherchons à voir si les différents motifs d'emprunts ont un effet sur le nombre d'échéances impayées par le client. Autrement dit nous cherchons à comparer les différentes modalités de la variables « *Marche* » selon une modalité précise de la variable « *Impaye* »

Nous aurons alors comme variable indépendante (VI) la variable « *Marche* » et comme variable dépendante (VD) = « *Impayé* ».

En construisant le tableau de contingence des variables « *Impaye* » (en ligne) et « *Marche* » (en colonne) et en réalisant le diagramme en barre de ce tableau, nous obtenons le graphique ci-dessous



D'après le graphe ci-dessus, nous pouvons dire que les clients qui ont contracté un crédit pour l'achat d'une voiture et les clients ayant contracté un crédit pour éventuellement s'acheter une maison ou pour la meubler sont ceux ayant le plus d'échéances impayées parmi les autres clients ayant eux aussi d'autres motif de crédit.

Nous voyons aussi que les clients ayant pour motif d'emprunt la rénovation et les clients ayant pour motif d'emprunt le mobilier ou l'ameublement sont ceux ayant le plus grand nombre d'échéances payées.

Les clients ayant pour motif d'emprunt l'achat éventuel de scooter sont ceux ayant le moins d'échéances payées parmi les autres clients ayant eux aussi d'autres motif de crédit.

Dans la pratique pour s'assurer qu'il existe réellement une association 'significative' entre deux variables qualitatives on utilise un test statistique. Ainsi, dans le cas nous pouvons utiliser le test de chi-2 pour évaluer s'il existe une dépendance significative entre les variables « *Impaye* » et « *Marche* »

Hypothèses de test :

Ho : Les deux variables sont indépendantes vs **H1** : Les deux variables sont dépendantes.

Si la p-value est inférieur au seuil de 5%, alors on rejette l'hypothèse nulle d'indépendance

Dans notre cas nous avons : **p-value = 0,6467** donc supérieure au seuil.

Alors nous ne rejetons pas Ho.

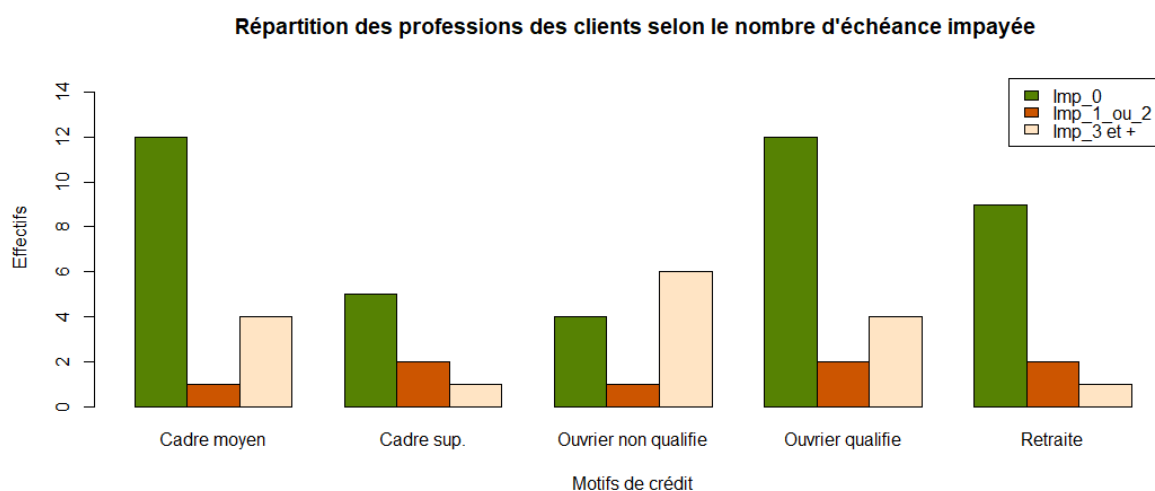
Nous pouvons donc affirmer qu'il n'existe pas **une association statistique significative** entre les variables « *Marche* » et « *Impaye* ».

○ Deuxième étude

Dans cette deuxième étude, nous cherchons à voir si les différentes professions des clients ont un effet sur le nombre d'échéances impayées par le client. Autrement dit nous allons chercher à comparer les différentes modalités de la variables « *Profession* » selon une modalité précise de la variable « *Impaye* »

Nous aurons alors comme variable indépendante (VI) la variable « *Profession* » et comme variable dépendante (VD) = « *Impayé* ».

En construisant le tableau de contingence des variables « *Impaye* » (en ligne) et « *Profession* » (en colonne) et en réalisant le diagramme en barre de ce tableau, nous obtenons le graphique ci-dessous :



D'après le graphe ci-dessus, nous pouvons dire que les ouvriers non qualifiés sont les clients ayant le plus d'échéances impayées par rapport aux autres clients.

Nous voyons aussi que les ouvrier qualifiés et les cadres moyens sont les clients ayant le plus d'échéances payées

Nous allons également utiliser le test de chi-2 pour évaluer s'il existe une dépendance significative entre les variables « *Impaye* » et « *Profession* ».

Dans notre cas nous avons : **p-value = 0,299** donc supérieure au seuil.

Alors nous ne rejetons pas H_0 .

Nous pouvons donc affirmer qu'il n'existe pas **une association statistique significative** entre les variables « *Profession* » et « *Impaye* ».

II -/ Analyse des correspondances multiples

L'analyse factorielle des correspondances est utilisée lorsque **les variables dites d'intérêts sont qualitatives**. On les appelle encore variables actives

Ces variables actives seront les seuls à participer à la construction des axes et elles seront utilisées pour comparer et décrire les individus. Les variables supplémentaires quant à elles caractériseront les groupes d'individus et les relations entre les variables actives.

Etant donné que notre base porte sur 66 clients ayant souscrit un crédit à la consommation dans un organisme de crédit, nous aurons comme variables actives les variables directement liées à ce sujet-là c'est-à-dire les variables indexées de la position 1 à la position 5. Les variables supplémentaires seront les variables indexées de la position 6 à 12.

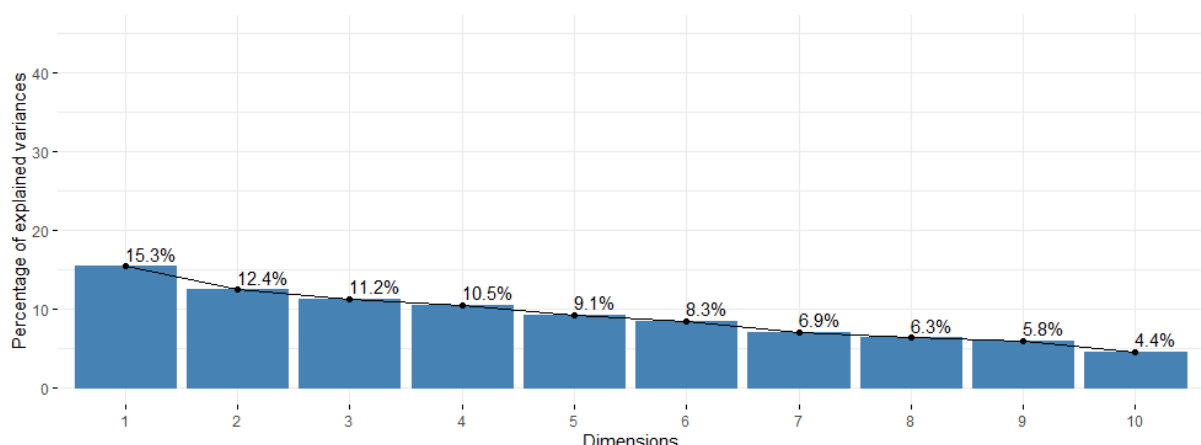
a-) Choix des axes à retenir :

Pour déterminer le nombre d'axe principaux à retenir, nous devons examiner les valeurs propres. Celles-ci correspondent à la quantité d'information retenue par chaque axe. Elles sont grandes pour le premier axe et petites pour l'axe suivant.

La proportion des variances retenues par les différentes dimensions (axes) peut être extraite à l'aide de la fonction ***get_eigenvalue()*** du package *factoextra*.

Plusieurs solutions existent pour déterminer le nombre d'axes à analyser en ACM.

Celle que nous allons utiliser, consiste à regarder le diagramme en barre des valeurs propres. Le nombre d'axes est déterminé par le point, au-delà duquel les valeurs propres restantes sont toutes relativement petites et de tailles comparables

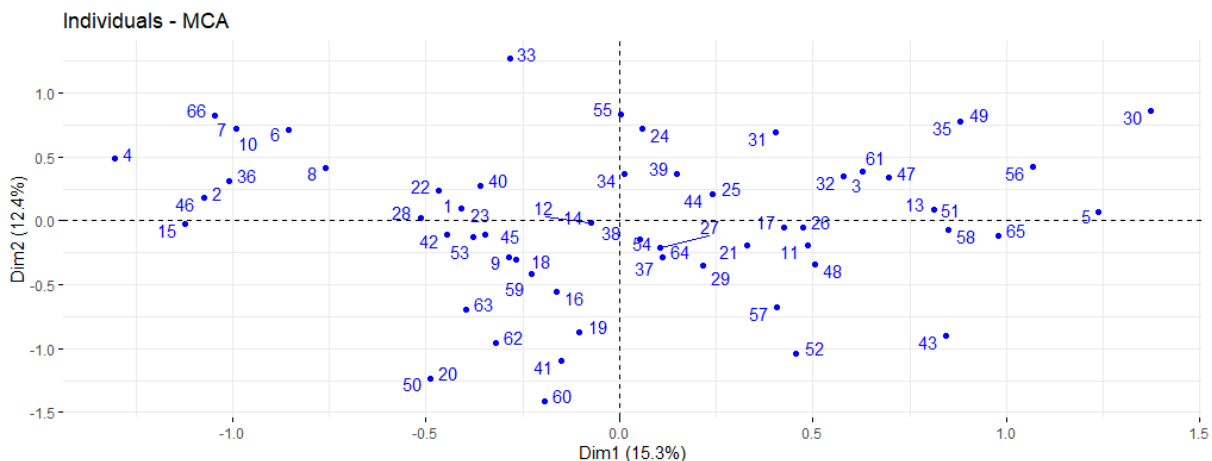


Nous constatons d'après le graphique ci-dessus que les proportions de variances retenues par les axes sont décroissantes suivant les axes et très faible, cela est classique en ACM.

Les deux premiers axes expriment environ **28 %** de l'inertie totale. Autrement dit, 28 % de l'information du tableau de données est résumée par les deux premières dimensions, ce qui est relativement important dans le cadre d'une ACM.

Nous pouvons donc nous contenter de décrire ces deux premiers axes.

b-) Description du nuage des individus



Pour décrire le nuage des individus, nous allons dans un premier temps nous focaliser sur les individus les plus originaux c'est-à-dire qui ont une grande disto et qui dans un deuxième temps contribuent fortement à la construction de l'axe.

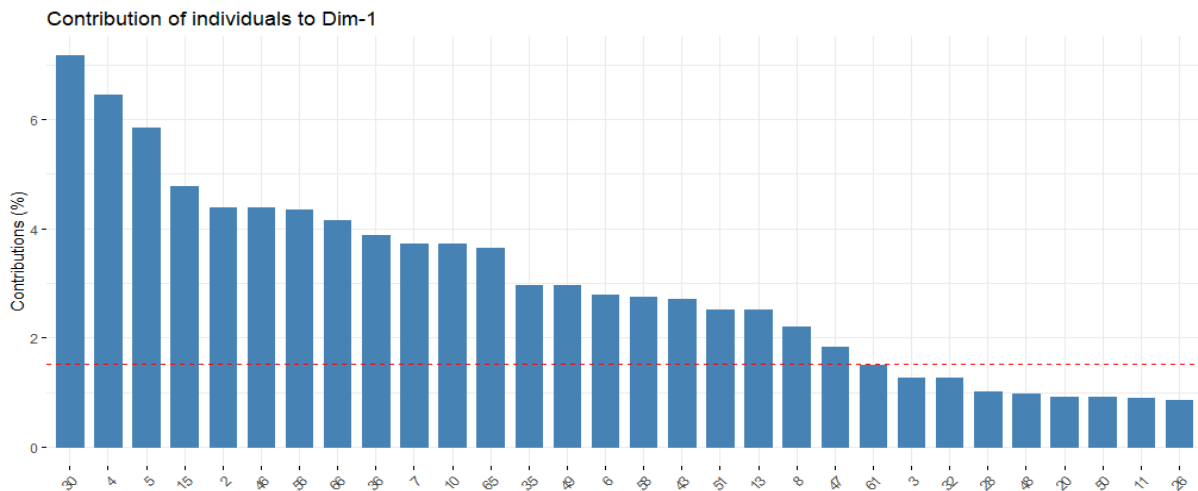
➤ Commentaires pour l'axe 1

Sur le graphique, nous pouvons voir qu'il y'a des individus qui se démarquent des autres individus et qui sont assez éloignés du centre. Nous remarquons que les individus 6, 7, 10 et 66 sont proches, c'est-à-dire que ces individus prennent à peu près les mêmes valeurs par rapport aux variables. Ils ont chacun une grande disto et sont tous de coordonnées négatives. Nous dirons que ces individus prennent des fortes valeurs sur l'axe 1.

Nous remarquons également que les individus 36, 46, 2 et 15 sont également proches, ont une grande disto et prennent aussi des fortes valeurs sur l'axe 1.

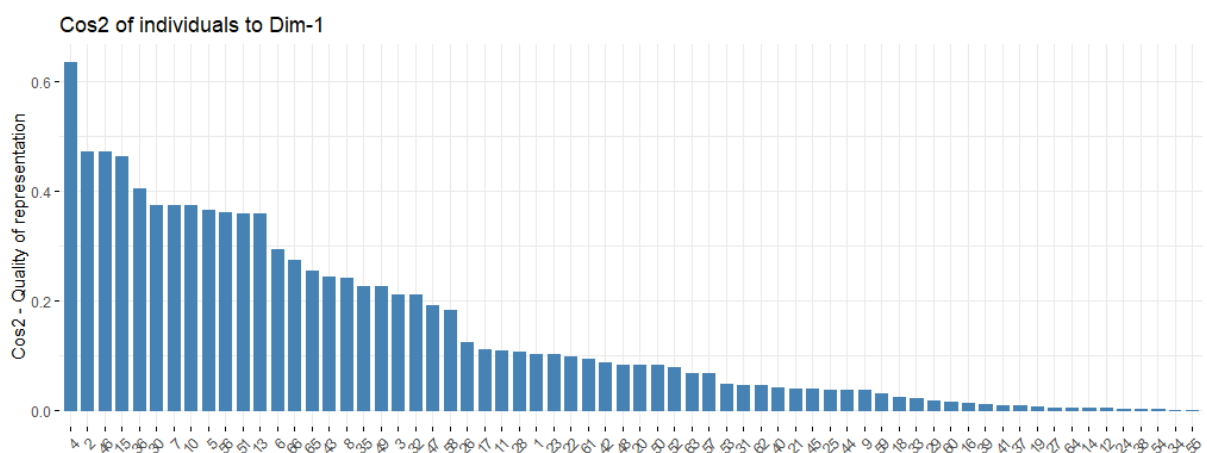
Comme autres individus atypiques nous remarquons de suite les individus 4 et 30 qui sont aussi très éloignés du centre et par conséquent ont une grande disto. L'individu 30 est de coordonnée positive et l'individu 4 de coordonnée négative. Nous dirons également que ces individus prennent des fortes valeurs sur l'axe 1.

La droite en pointillée rouge sur le graphique ci-dessous, indique la valeur moyenne attendue, si les contributions étaient uniformes.



Selon le graph ci-dessus et d'après les résultats obtenus, nous pouvons voir qu'il s'agit de deux groupes qui contribuent très fortement à la construction de l'axe 1.

Nous avons d'une part les individus 4, 15, 2, 46, 66, 36, 7, 10, 6 et d'autres part les individus 30, 5 et 56. Le fait que ces individus soient opposés signifie que ce sont des voitures très différentes puisque le 1^{er} axe est celui qui sépare au mieux les points. Donc ces deux groupes d'individus ont des comportements très différents et ce sur l'ensemble de leurs caractéristique observables



Nous pouvons aussi voir sur le graphique ci-dessus que seul l'individu 4 a une qualité de représentation sur l'axe 1 à peu près considérable. Tous les autres individus qui ont une grande contribution dans la construction de l'axe 1 ne sont pas bien représentés. Nous allons donc essayer de voir leur qualité de représentation sur le plan (Voir figure M).

Sur cette figure, nous pouvons voir que ce sont les individus 4, 7, 10 et 30 qui sont à peu près bien représentés sur le plan factoriel avec un cosinus carré supérieur à 0,5.

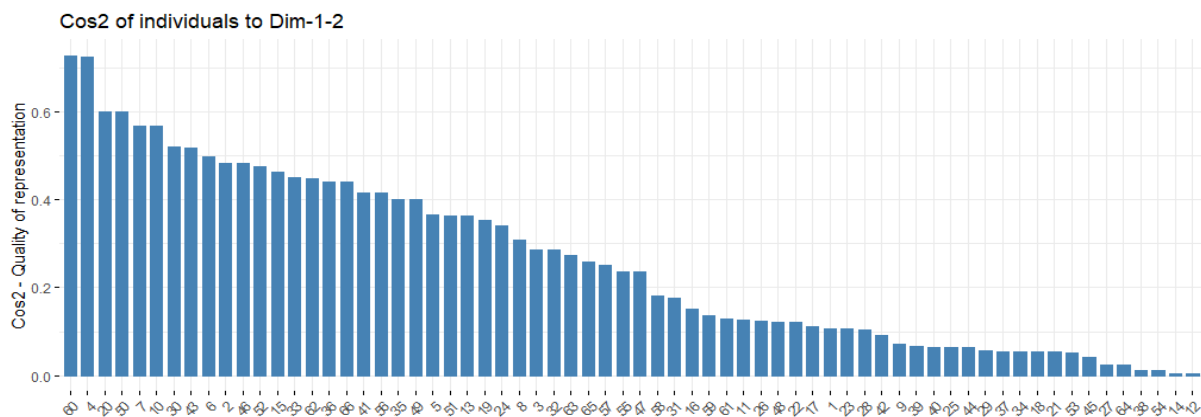
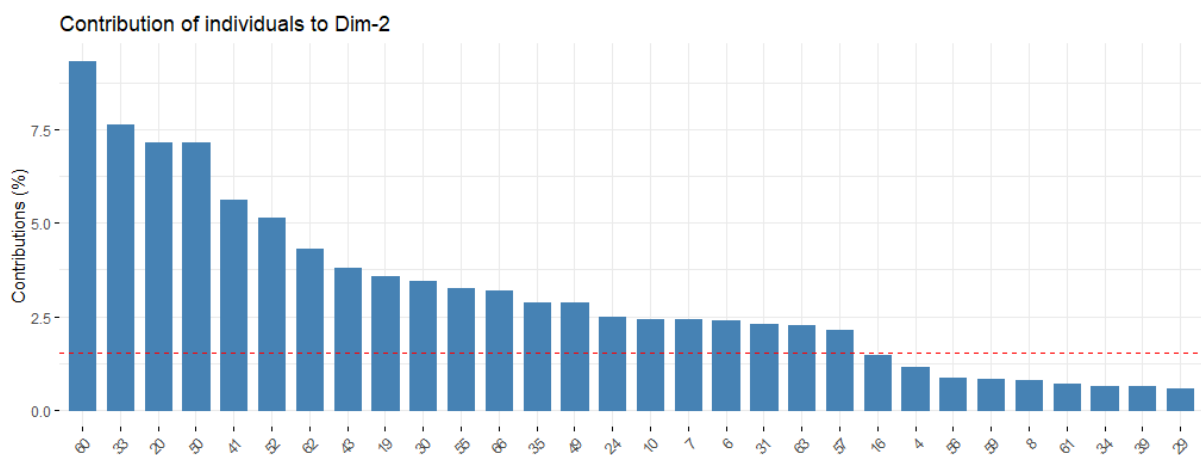


Figure M

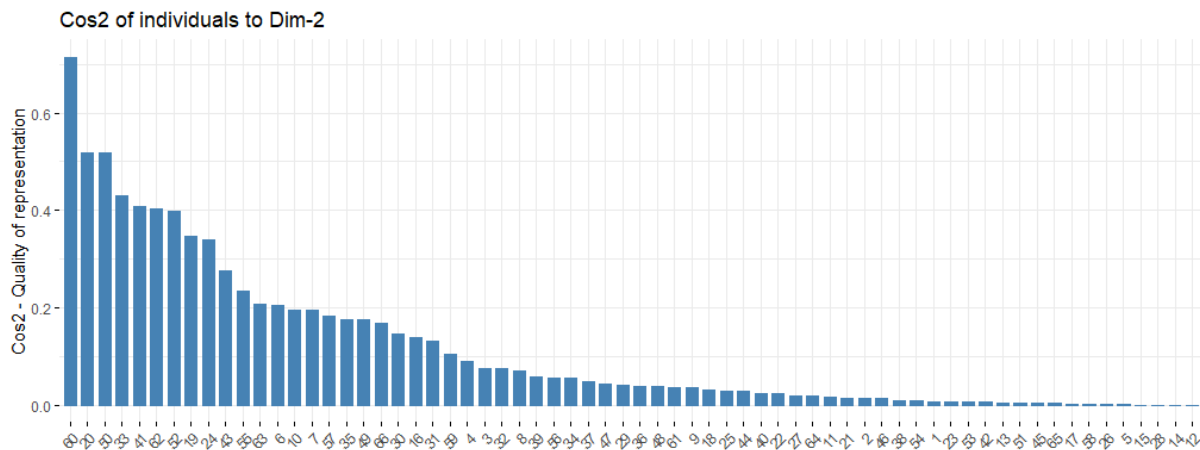
➤ Commentaires pour l'axe 2

Sur le graphique, nous pouvons voir qu'il y'a des individus qui se démarquent des autres et qui sont assez éloignés du centre. Nous remarquons que les individus 20, 50, 60 et 33 sont assez éloignés du centre c'est-à-dire qu'ils ont chacun une grande disto. Les individus 20, 50 et 60 sont de coordonnées négatives tandis que l'individu 33 est de coordonnée positive. Nous dirons que ces individus prennent des fortes valeurs sur l'axe 2.



Selon le graph ci-dessus et d'après les résultats obtenus, nous constatons que les individus atypiques que nous avons remarqués contribuent le plus à la construction de l'axe 2.

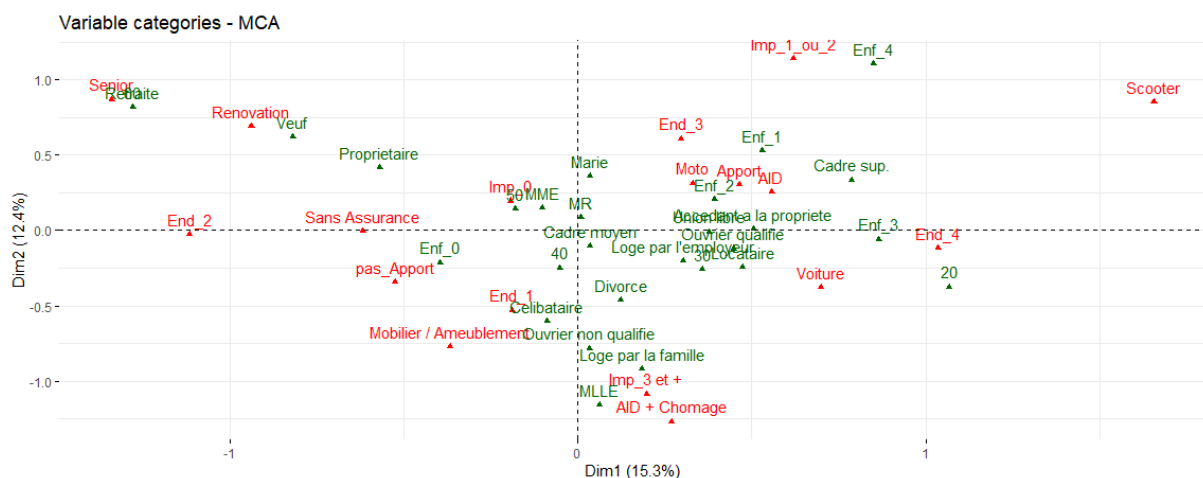
Il s'agit des individus 60, 20, 50 et 33. Le fait que ces individus n'ont pas les mêmes coordonnées signifie que nous sommes en présence de deux groupes. Le premier constitué des individus 60, 20, 50 de coordonnées négatives et le 2^e constitué uniquement de l'individu 33 de coordonnée positive. Ces groupes d'individus ont des comportements très différents sur l'ensemble de leurs caractéristique observables



Nous pouvons aussi voir sur le graphique ci-dessus que seul les individus 60, 20 et 50 ont une qualité de représentation sur l'axe 2 à peu près considérable. L'individu 33 qui a une grande contribution dans la construction de l'axe 2 n'est pas bien représentés. Nous allons essayer de voir sa qualité de représentation sur le plan (Voir figure M, page 11).

Sur la figure M, nous pouvons voir que l'individu 33 est à peu près bien représenté sur le plan factoriel avec un cosinus carré supérieur à 0,5.

c-) Description du nuage des variables

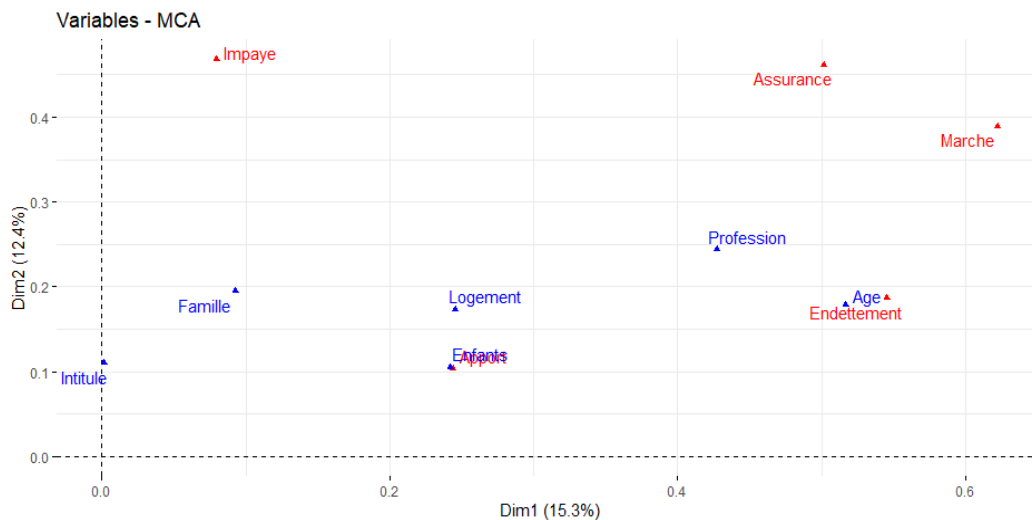


NB : Ce sont les catégories des variables qualitatives de notre base qui sont représentées sur le graphique ci-dessus.

Les catégories des variables actives sont en rouge et ceux des variables supplémentaires en verts.

Corrélation entre les variables et les axes principaux :

Le graphique ci-dessous nous permet d'identifier les variables les plus corrélées avec chaque axe. Les corrélations au carré entre les variables et les axes sont utilisés comme coordonnées.

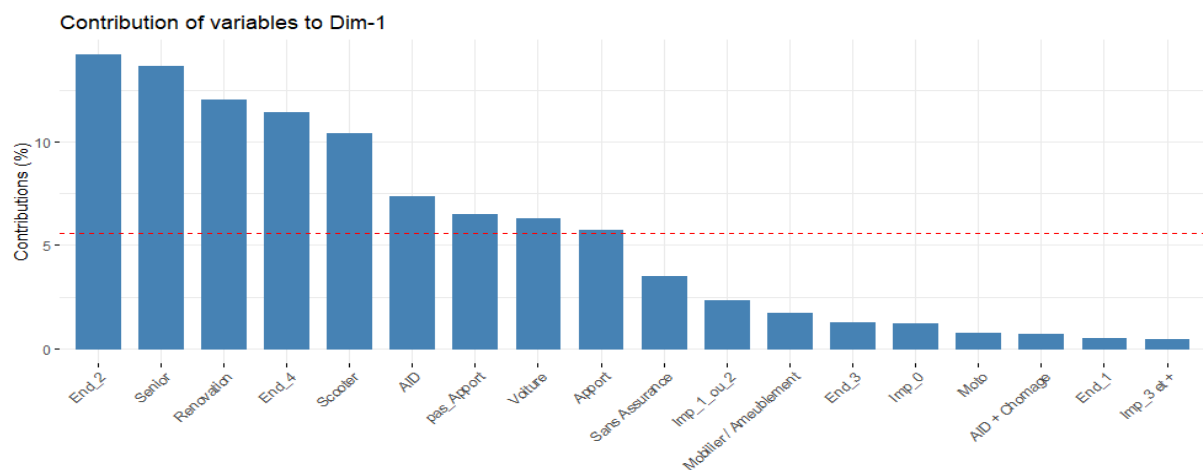


Nous constatons que les variables « Marche » et « Endettement » sont les plus corrélées avec la dimension 1. De même, la variable « Impaye » est la plus corrélée avec la dimension 2.

Pour décrire le graphe des variables, nous allons dans un premier temps observer les catégories de variables regroupées et éloignées, ensuite regarder la contribution de ces catégories dans la construction des axes et enfin leurs qualités de représentation.

Commentaires pour l'axe 1 :

Sur le graphique, nous pouvons voir qu'il y'a des catégories qui se démarquent des autres et qui sont assez éloignées du centre. Nous remarquons que les catégories « End_2 », « Senior », « Renovation » et « Scooter » sont assez éloignées du centre c'est-à-dire qu'ils ont chacune une grande disto. Les catégories « End_2 », « Senior », « Renovation » et « Scooter » sont de coordonnées négatives tandis que la catégorie « Scooter » est de coordonnée positive. Nous dirons que ces modalités prennent des fortes valeurs sur l'axe 1.



Selon le graph ci-dessus et d'après les résultats obtenus, nous constatons que les modalités atypiques que nous avons remarquées contribuent le plus à la construction de l'axe 1.

Il s'agit des catégories « End_2 », « Senior », « Renovation » et « Scooter ». Le fait que ces catégories n'ont pas les mêmes coordonnées signifie que nous sommes en présence de deux groupes. Le premier constitué des catégories « End_2 », « Senior », « Renovation » de

coordonnées négatives et le 2^e constitué uniquement de la modalité « *Scooter* » de coordonnée positive. Ces groupes de modalités ont des comportements très différents sur l'ensemble de leurs caractéristique observables

En faisant la commande `head(var$cos2, 10)`, nous nous sommes rendu compte que les cosinus carré de ces catégories était très faible sur l'axe 1, nous avons donc décidé de regarder la qualité de représentation sur le plan (voir graphique juste en bas).

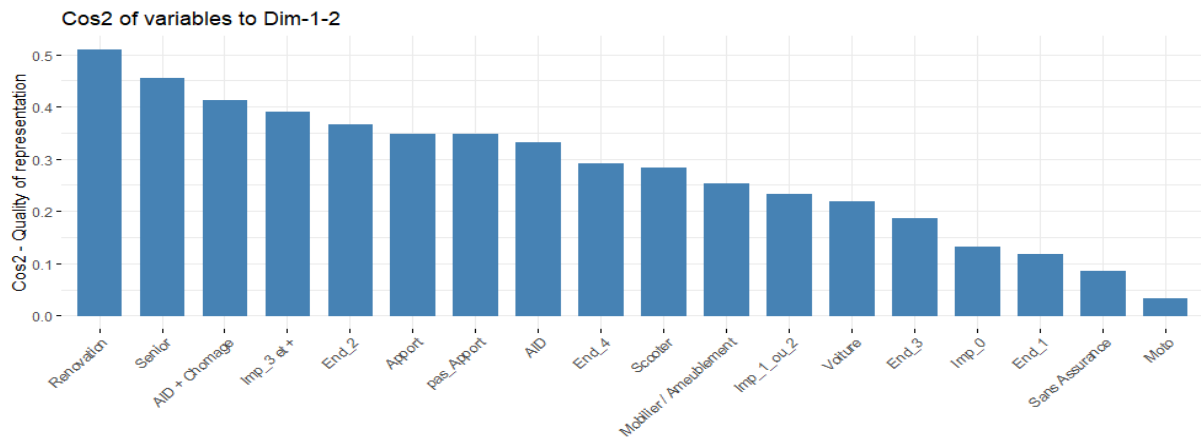
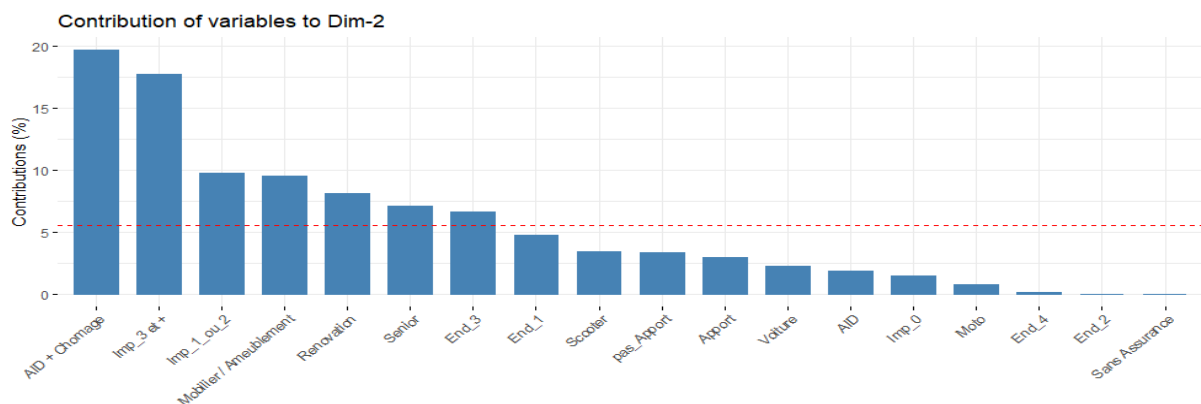


Figure L

En regardant le graphique, nous pouvons voir que seules les catégories « *Renovation* » et « *Senior* » sont à peu près bien représentés sur le plan avec un cosinus carré supérieur à 0,45.

➤ Commentaires pour l'axe 2 :

Sur le graphique des catégories, nous pouvons voir qu'il y'a des catégories qui se démarquent des autres et qui sont assez éloignés du centre. Nous remarquons que les catégories « *AID + Chomage* », « *Imp_3 et +* » et « *Imp_1_ou_2* » sont assez éloignés du centre c'est-à-dire qu'ils ont chacune une grande disto. Les catégories « *AID + Chomage* » et « *Imp_3 et +* » sont de coordonnées négatives tandis que la catégorie « *Imp_1_ou_2* » est de coordonnée positive. Nous dirons que ces modalités prennent des fortes valeurs sur l'axe 2.



Selon le graph ci-dessus et d'après les résultats obtenus, nous constatons que les modalités atypiques que nous avons remarquées contribuent effectivement le plus à la construction de l'axe 2.

Il s'agit des catégories « *AID + Chomage* », « *Imp_3 et +* » et « *Imp_1_ou_2* ». Le fait que ces catégories n'aient pas les mêmes coordonnées signifie que nous sommes en présence de deux groupes. Le premier constitué des catégories « *AID + Chomage* », « *Imp_3 et +* » de coordonnées négatives et le 2^e constitué uniquement de la modalité « *Imp_1_ou_2* » de coordonnée positive. Ces groupes de modalités ont des comportements très différents sur leurs caractéristiques observables

En faisant la commande **head(var\$cos2, 10)**, nous nous sommes rendu compte également que les cosinus carré de ces catégories était beaucoup trop faible sur l'axe 2, nous avons donc décidé de regarder la qualité de représentation sur le plan (voir graphique page ...)

En regardant la figure L de la page 14, nous pouvons voir que seules les catégories « *AID + Chomage* », « *Imp_3 et +* » sont les mieux représentées parmi toutes les catégories contribuant à la construction de l'axe 2.

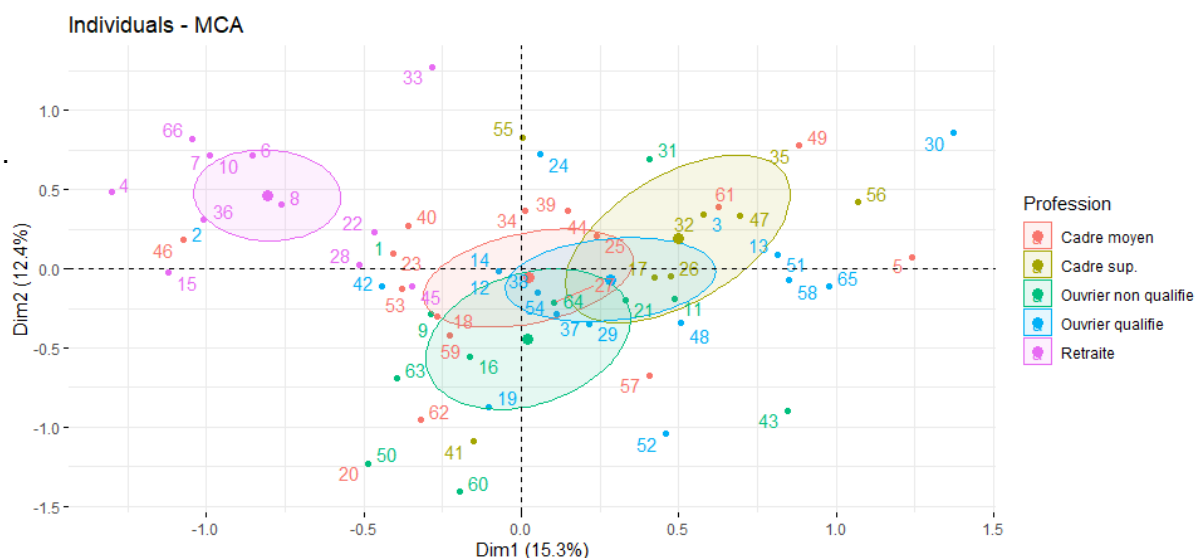
d-) Variables supplémentaires

Dans cette partie, nous allons colorons les individus selon les catégories des variables supplémentaires.

Rappelons que les variables supplémentaires sont utilisées pour caractériser ou illustrer les différents groupes d'individus existant dans notre base.

Les individus sur lesquels nous allons plus nous focaliser sont ceux que nous avons retenu lors de la description du nuage des individus.

➤ Caractérisation des individus avec la variable « Profession »

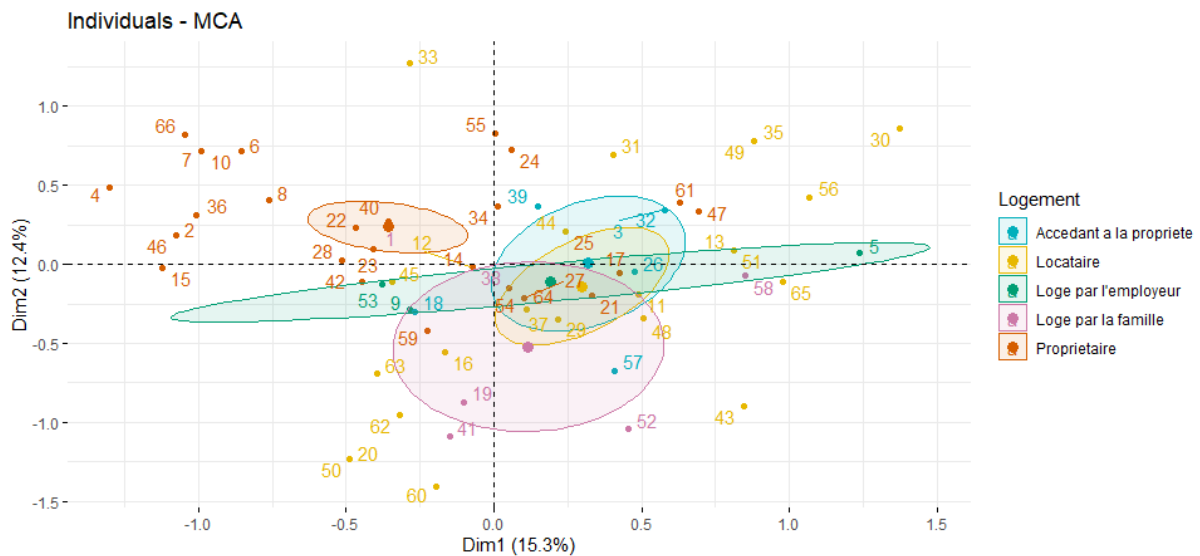


Nous remarquons sur le graphique que l'axe 1 oppose clairement deux groupes d'individus. Nous avons d'une part les ouvriers qualifiés, les cadres supérieurs, les cadres moyen et d'autre part nous avons majoritairement les retraités.

Les individus que nous avons retenu lors de l'analyse de l'axe 1 à savoir les individus 4, 7, 10 sont des retraités et ils s'opposent à l'individu 30 qui est un ouvrier qualifié

Les individus que nous avons retenu lors de l'analyse de l'axe 2 à savoir les individus 60, 20 et 50 sont des ouvriers non qualifiés et cadre moyen et ils s'opposent à l'individu 33 qui est un retraité.

➤ Caractérisation des individus avec la variable « Logement »



Nous pouvons voir sur le graphique que l'axe 1 oppose clairement deux groupes d'individus. Nous avons d'une part les propriétaires et d'autre part les locataires.

Les individus que nous avons retenu lors de l'analyse de l'axe 1 à savoir les individus 4, 7, 10 sont des propriétaires et ils s'opposent à l'individu 30 qui est un locataire

-	+
Catégories	Catégories
« AID + Chomage » « Imp_3 et + »	
Individus	Individus
60, 50, 20	33

➤ Identification des groupes d'individus à l'aide des variables

Nous pouvons dire que l'axe 1 oppose deux groupes d'individus :

Premièrement nous avons les individus 4, 7 et 10 qui d'une part ont souscrit à un crédit avec comme motif la rénovation et d'autre part ont une assurance de type Senior. Ces individus sont généralement des retraités et ils sont également propriétaires.

Deuxièmement nous avons l'individu 30 qui est un simple ouvrier qualifié et qui est locataire

En d'autres termes il n'y'a pas beaucoup de retraités qui sont des locataires parce que la plupart sont propriétaires et il n'y a pas beaucoup d'ouvriers propriétaires parce que la plupart sont locataires.

Les retraités souscrivent plus à des contrats d'assurances de type senior et ceci n'est pas le cas chez les ouvriers.

L'information que nous pouvons en tirer est que les retraités propriétaires et les ouvriers locataires n'ont pas les mêmes motifs lorsqu'ils viennent souscrire à un crédit. Cela est certainement dû au fait qu'ils ne sont pas dans les mêmes conditions de vies.

Nous pouvons dire que l'axe 2 oppose deux groupes d'individus :

D'abord nous avons les individus 60, 20 et 50 qui d'une part ont souscrit à une assurance de type AID + Chomage et d'autre part, ont plus de 3 échéances impayées. Ces individus sont généralement des ouvriers non qualifiés ou des cadres moyens.

Ensuite nous avons l'individu 33 qui est un retraité.

En d'autres termes, la plupart du temps les ouvriers ont beaucoup plus d'échéances à payer comparer aux retraités qui ont moins d'échéances à payer dans le cadre du remboursement d'un crédit.

Les ouvriers souscrivent plus à des contrats d'assurances garantissant des indemnités en situation de chômage. Ceci n'est pas le cas des retraités

L'information que nous pouvons en tirer est que les retraités et les ouvriers ne rencontrent pas les mêmes difficultés dans le remboursement de leurs emprunts. Ceci peut être expliquer par le fait que les retraités ont certainement déjà passé plusieurs années à rembourser leurs crédits alors que les ouvriers actuels sont encore au début de leurs période de remboursement.

Conclusion

L'analyse des correspondances multiple (ACM) est un outil de synthèse de l'information, très utile lorsque l'on est en présence d'une somme importante de données qualitatives à traiter et interpréter. Le principe de cette analyse est de chercher le plan qui résume le mieux l'information contenue dans le tableau disjonctif complet. On obtient ainsi une représentation approchée du nuage dans un espace de faible dimension.

La limite avec les méthodes d'analyse de données telle que l'ACM par exemple est qu'elles sont juste descriptives et non explicatives. Elles ne peuvent pas être utilisées pour montrer des liens de causalités. Elles nous aident cependant à formuler des hypothèses.