Segmentation multicanale

Clusterings

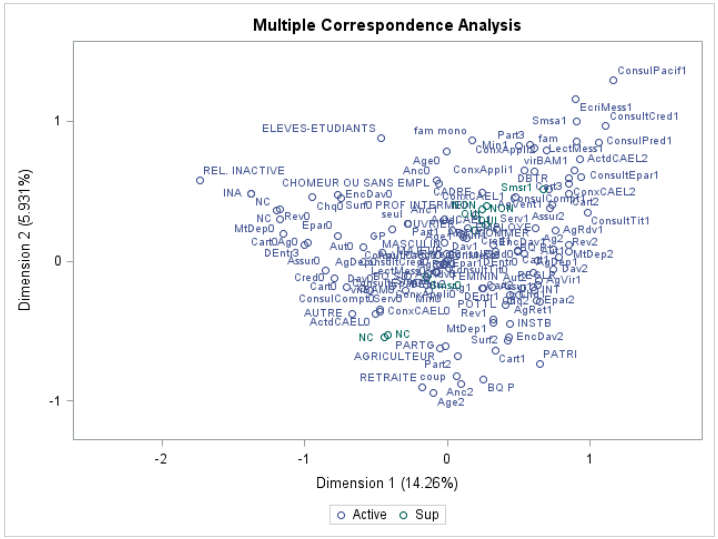
3 approches différentes

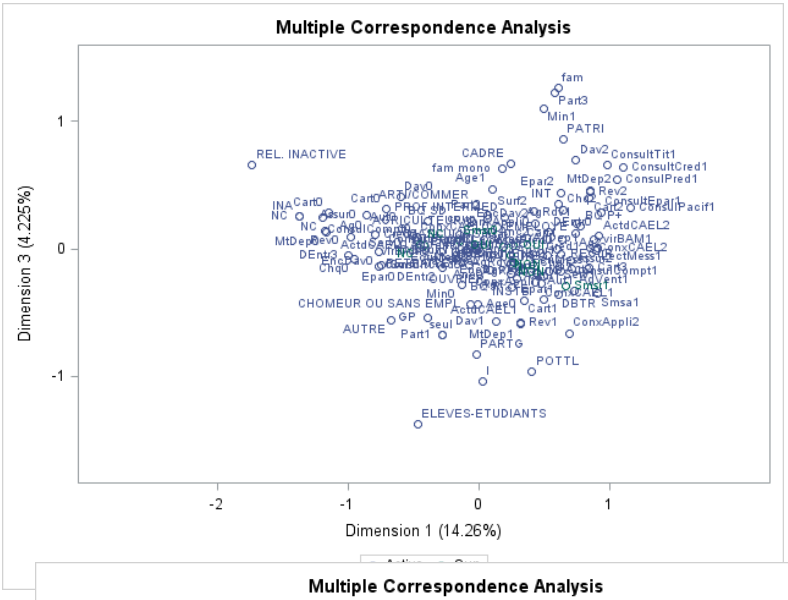
- Clustering sur l'AFCM de toutes les variables qualitatives (y compris les quantitatives discrétisées)
- Clustering sur l'ACP de toutes les variables quantitatives
- Clustering sur l'AFCM + ACP des variables quantitatives et qualitatives

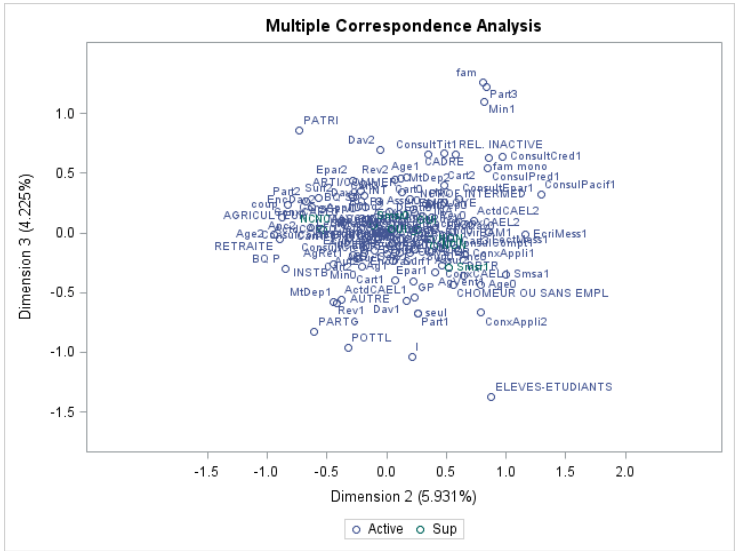
1ere approche : Clustering sur l'AFCM des variables qualitatives - Méthode

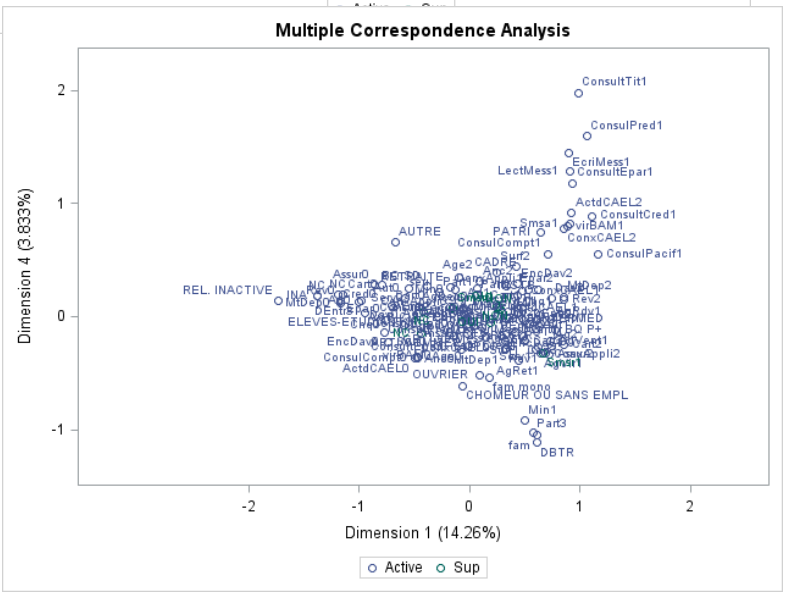
- Recoder toutes les variables quantitatives en qualitatives en choisissant des classes en fonction de la distribution des variables
- Y ajouter toutes les variables qualitatives initiales
- Réaliser une AFCM sur toutes ces variables dans le but de réduire la dimension et optimiser la méthode
- Faire un k-means avec un nombre de classes égal à 10% de l'échantillon sur les résultats de l'AFCM
- Réaliser un clustering hiérarchique en initialisant les barycentres obtenus par le k-means
- Déterminer le nombre de classes optimal
- Lancer un k-means avec le nombre de classes optimal

Résultats de l'AFCM

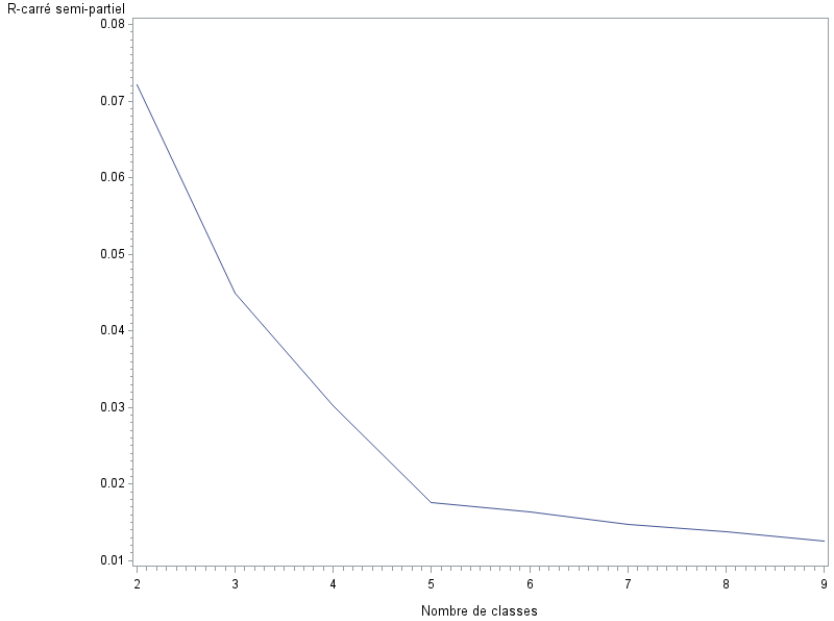






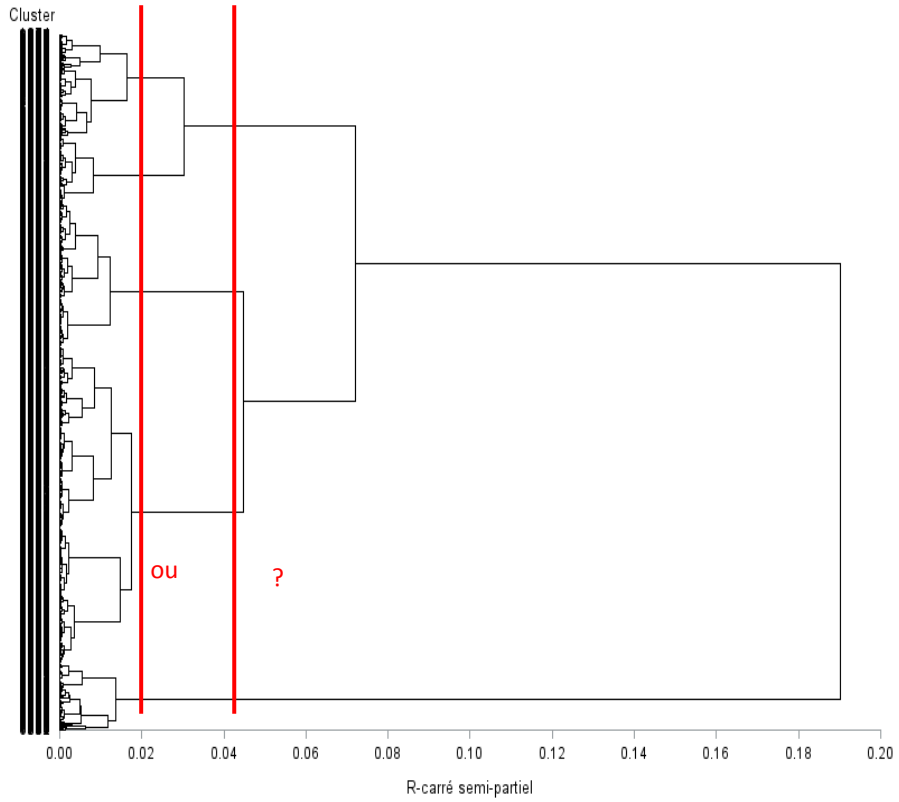


Résultats du clustering

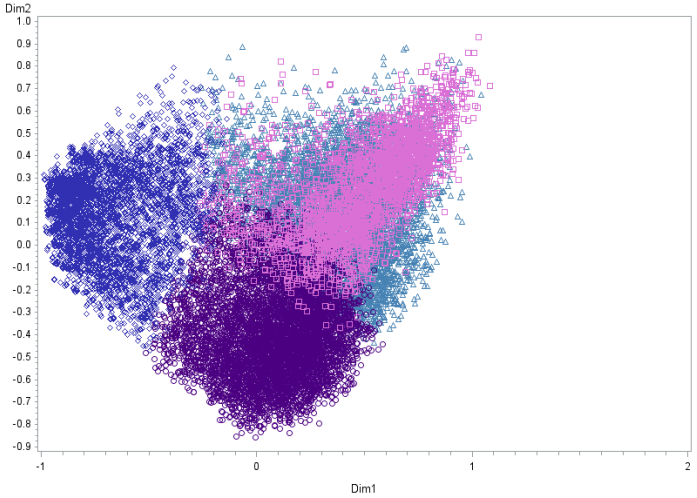


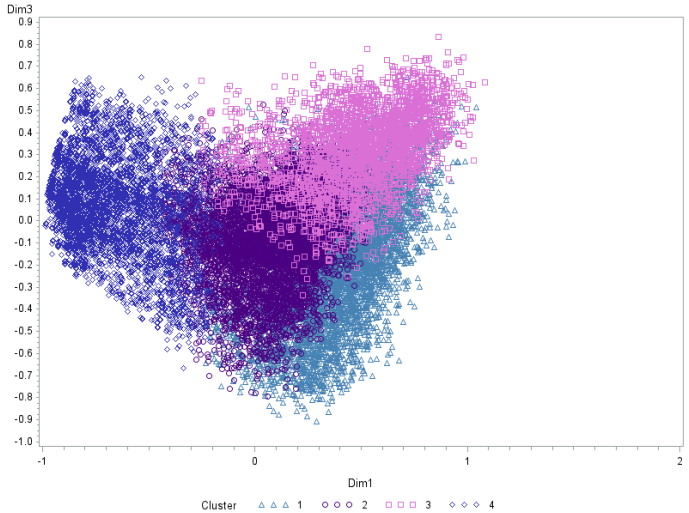
- 5/4 classes suggérées par les graphs. CCC = 24.7 pour 5 classes et 18.1 pour 4 classes

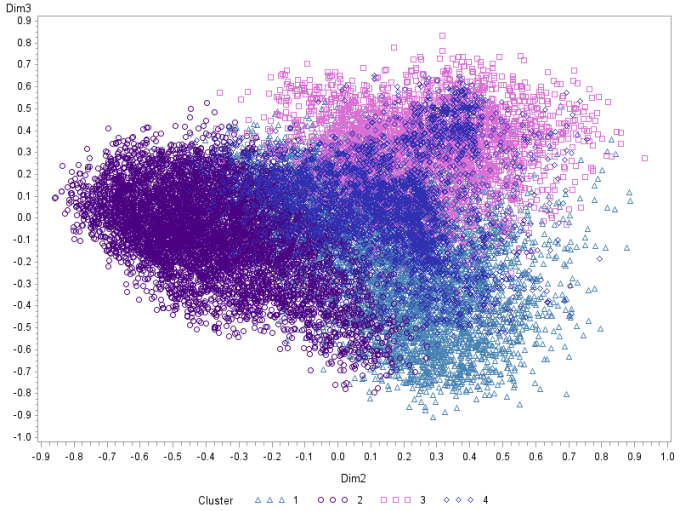
Clustering hiérarchique - dendogramme



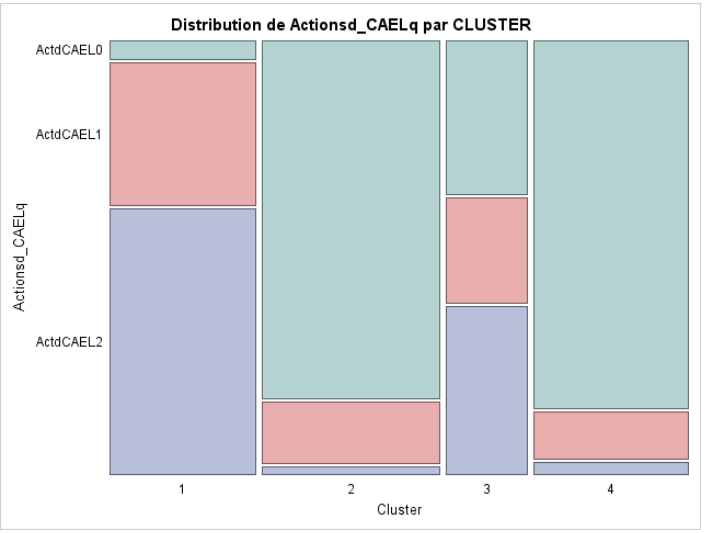
Représentation des clusters dans les plans de l'AFCM

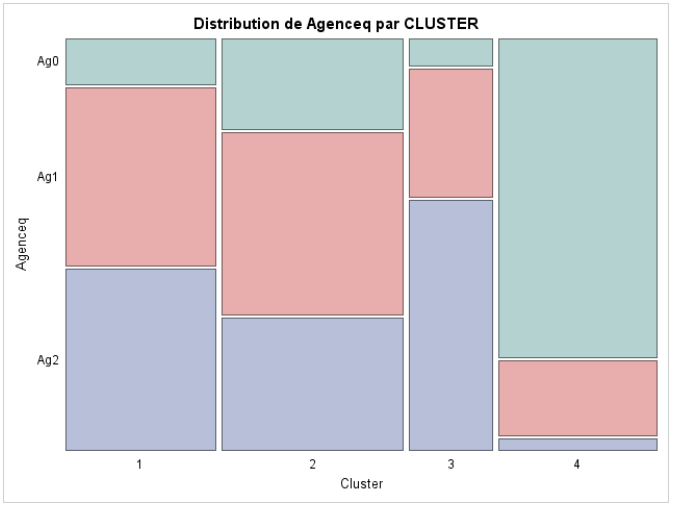


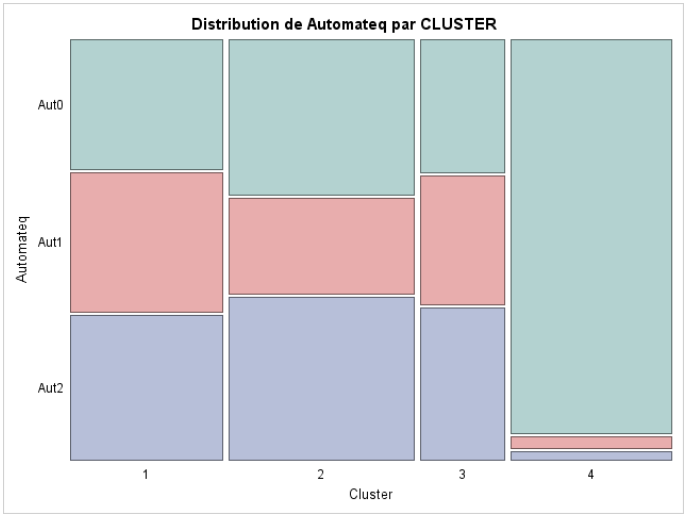


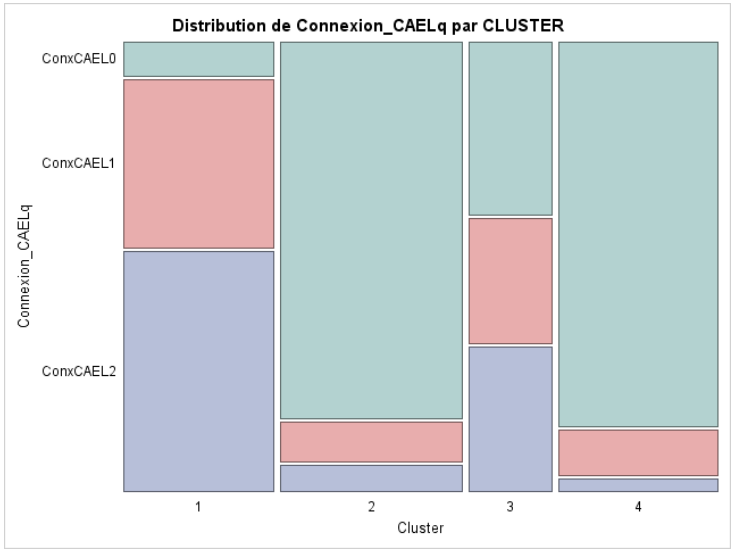


Résultat du clustering

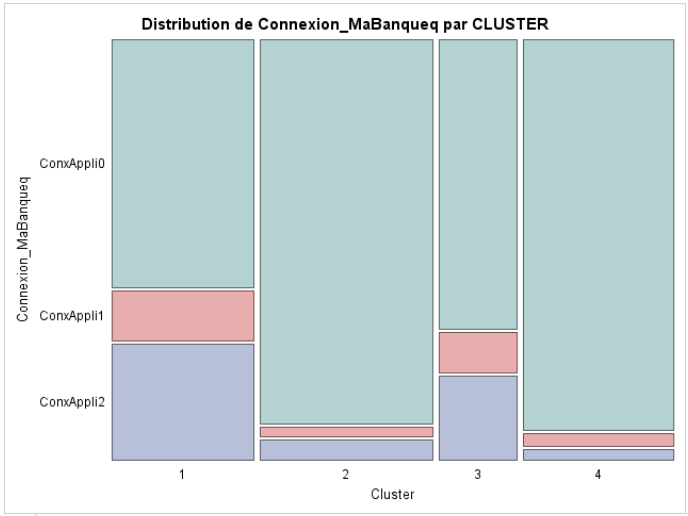


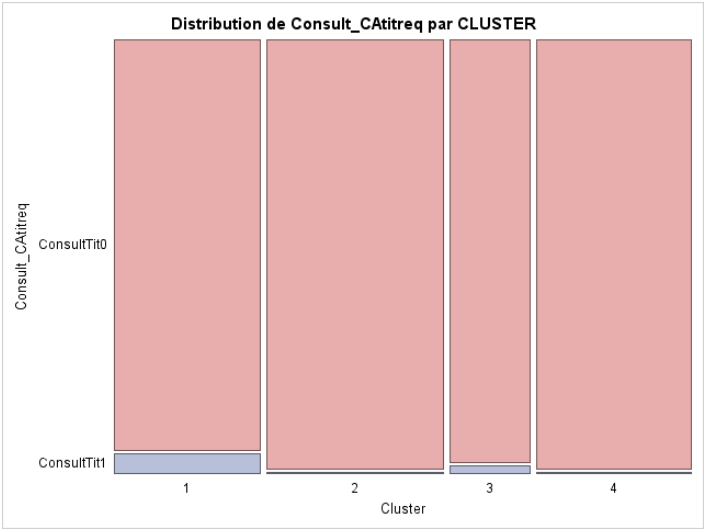


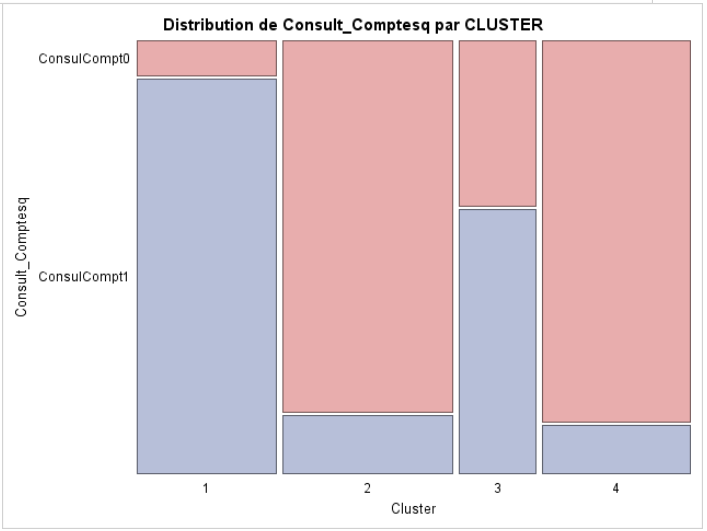


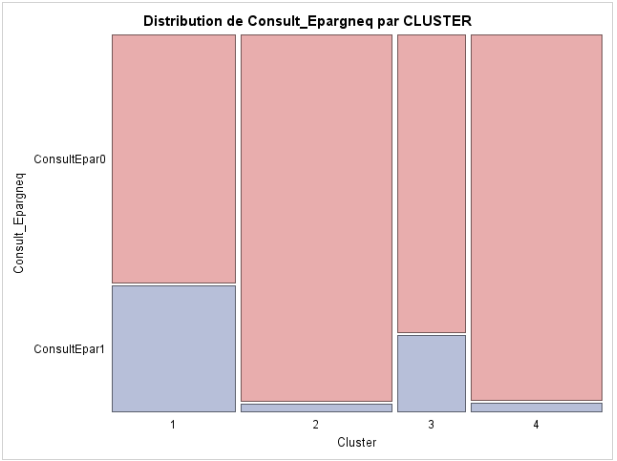


Résultat du clustering (2)

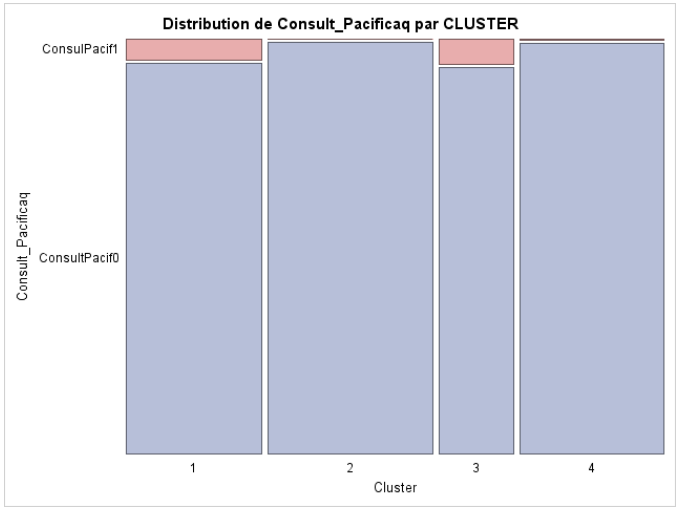


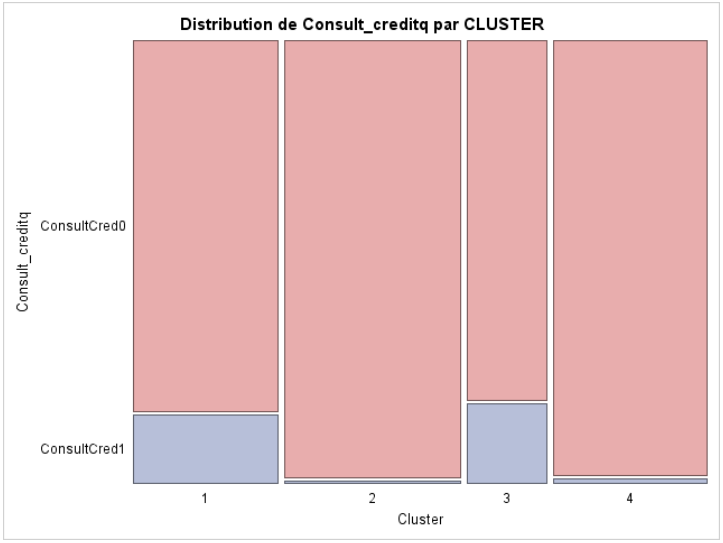


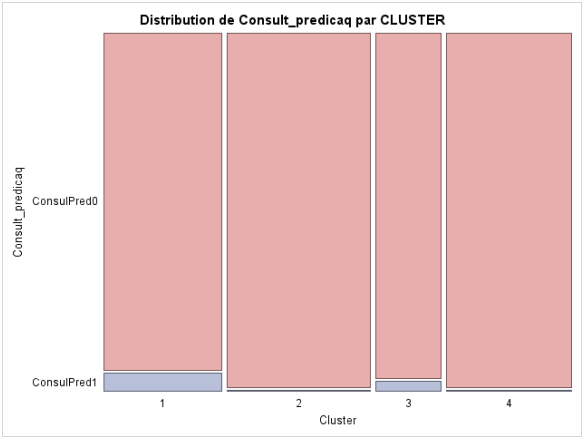


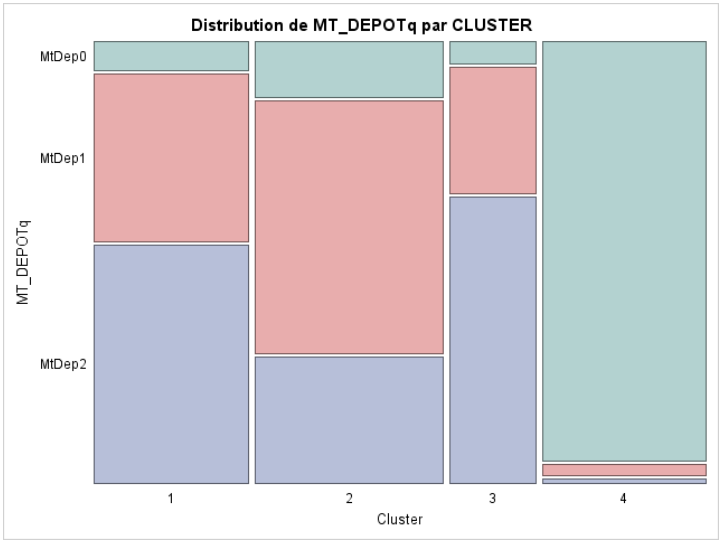


Résultat du clustering (3)

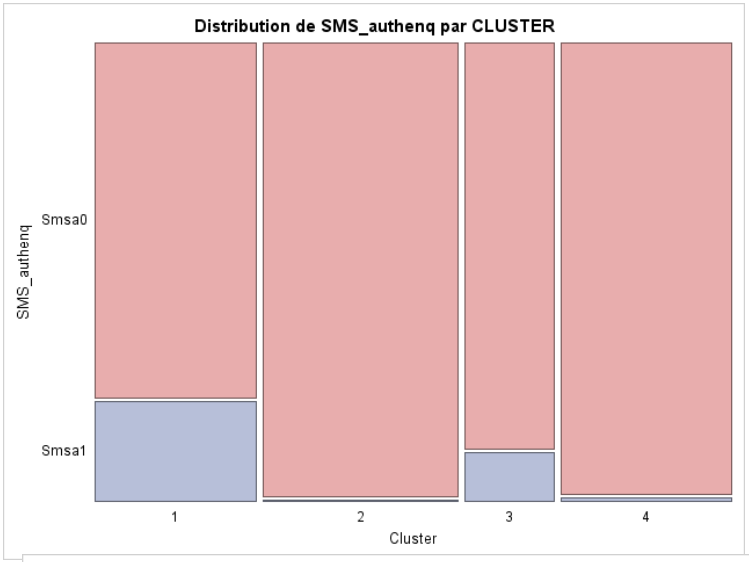


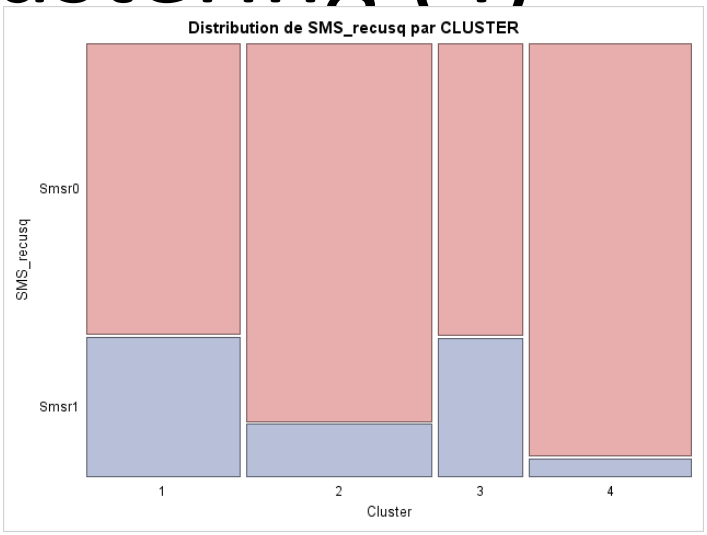


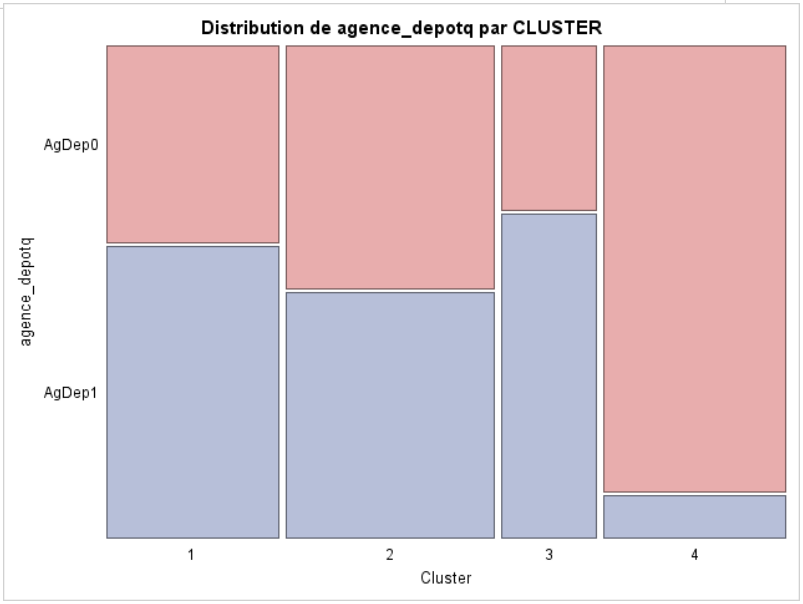


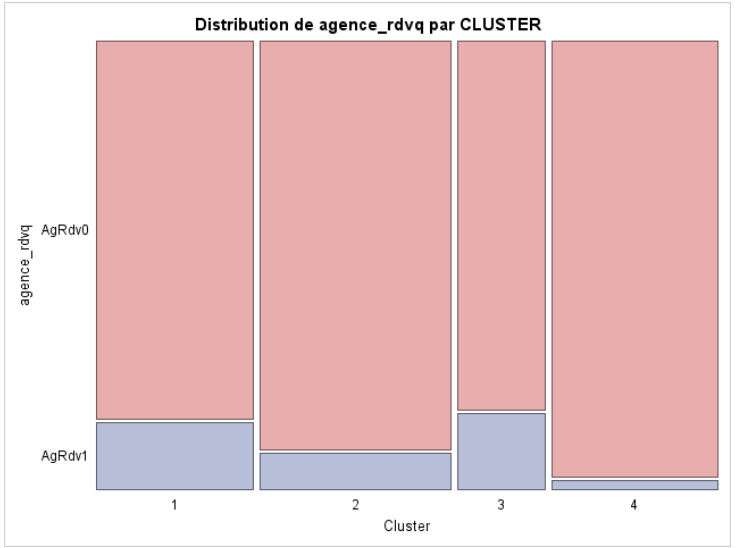


Résultat du clustering (4)

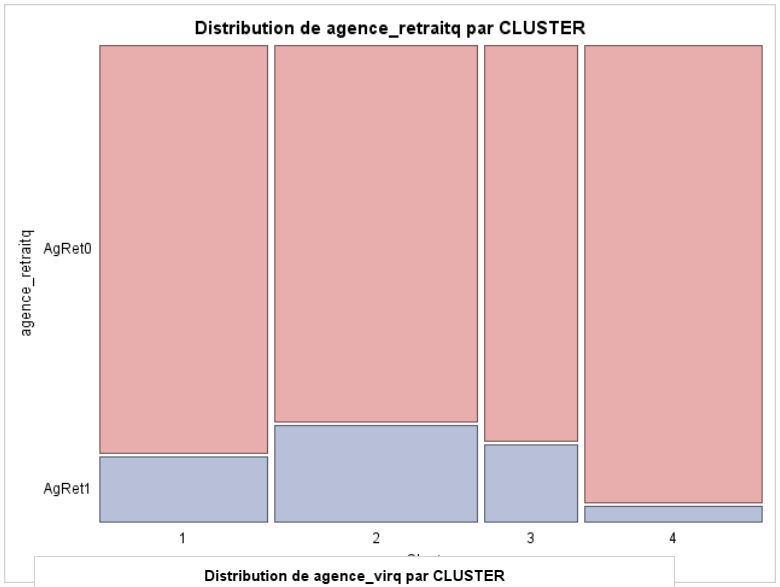


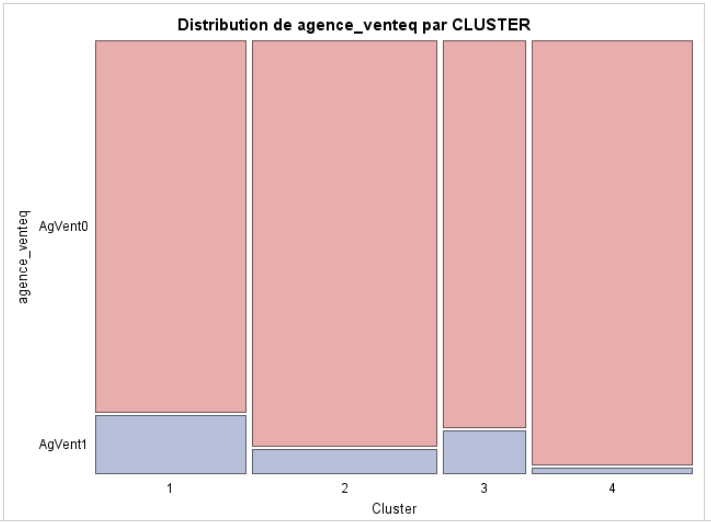


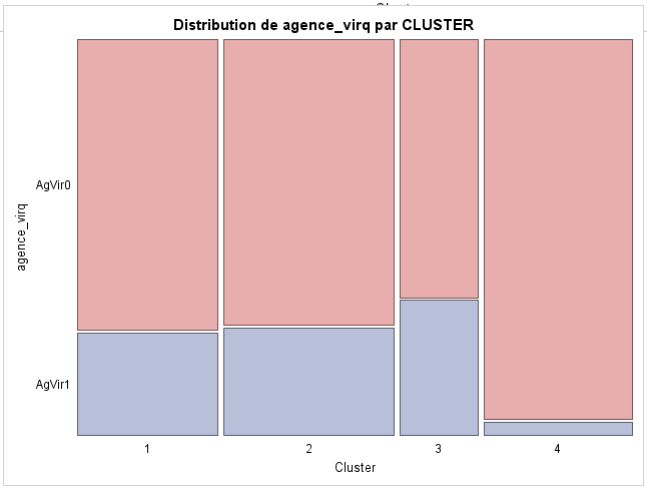


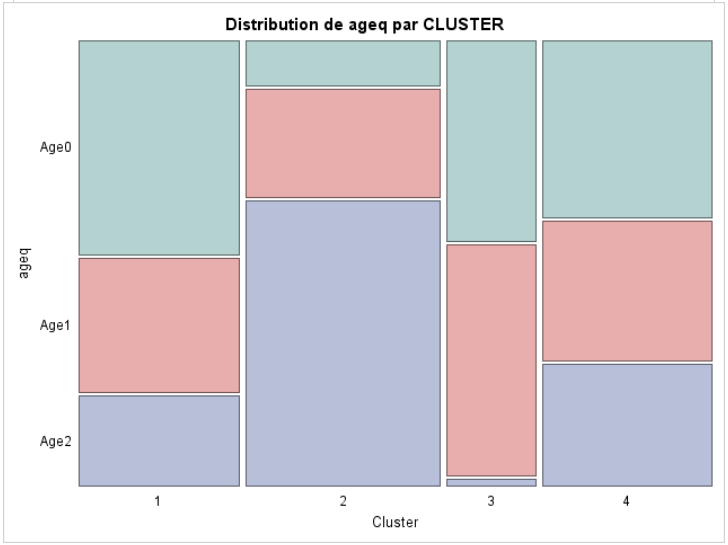


Résultat du clustering (5)

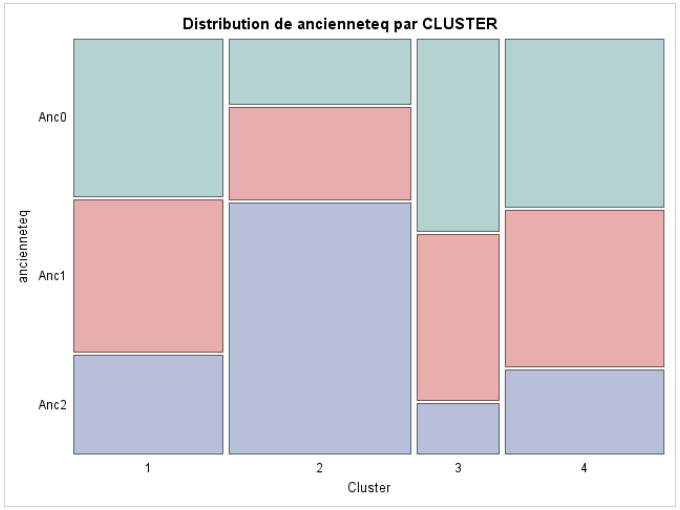


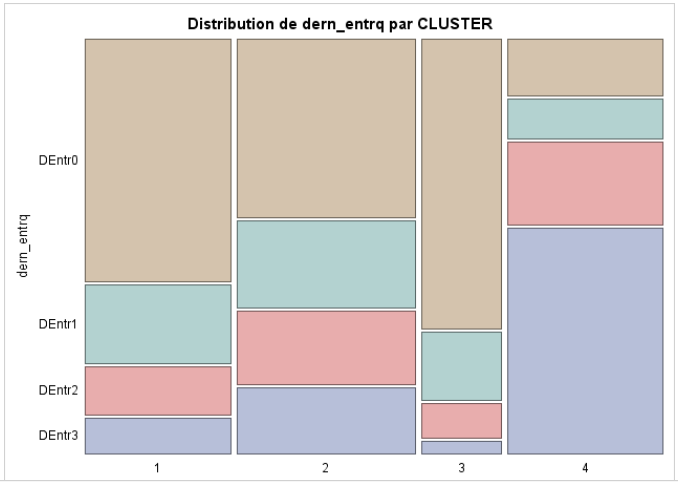


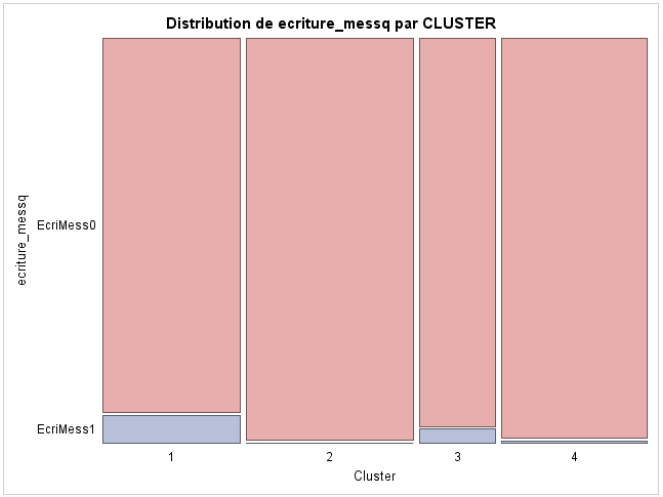


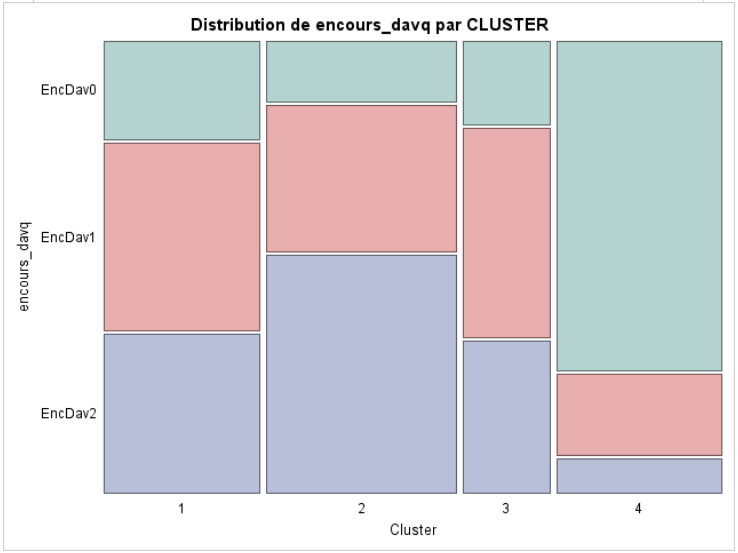


Résultat du clustering (6)

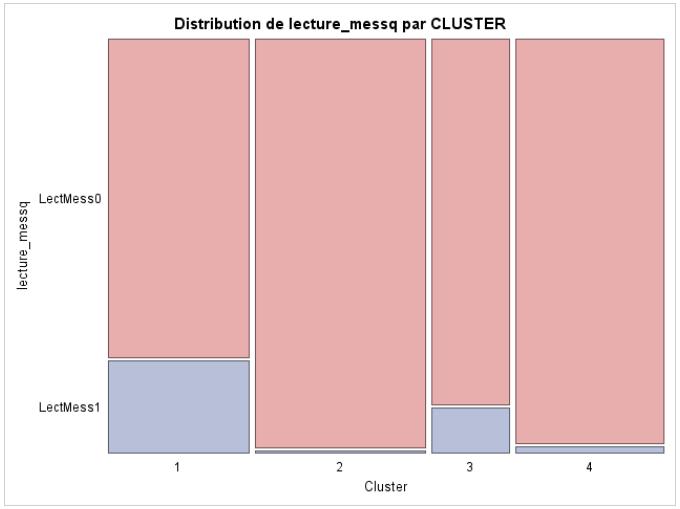


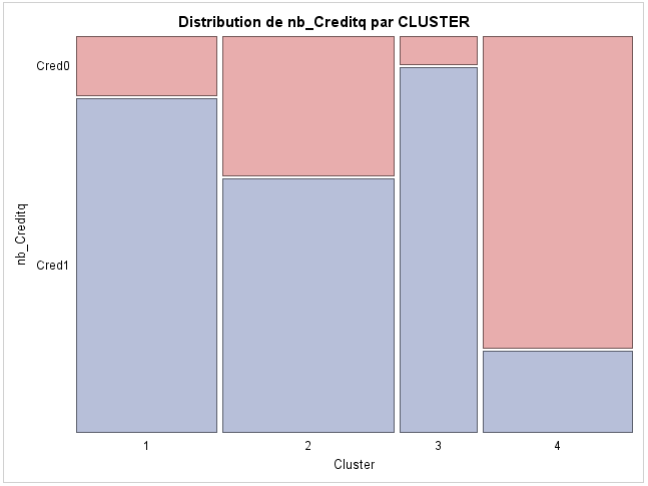


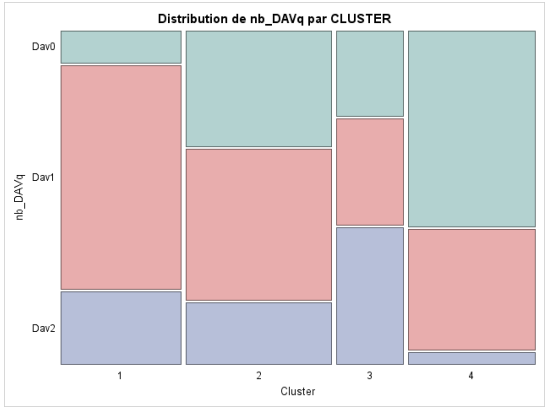


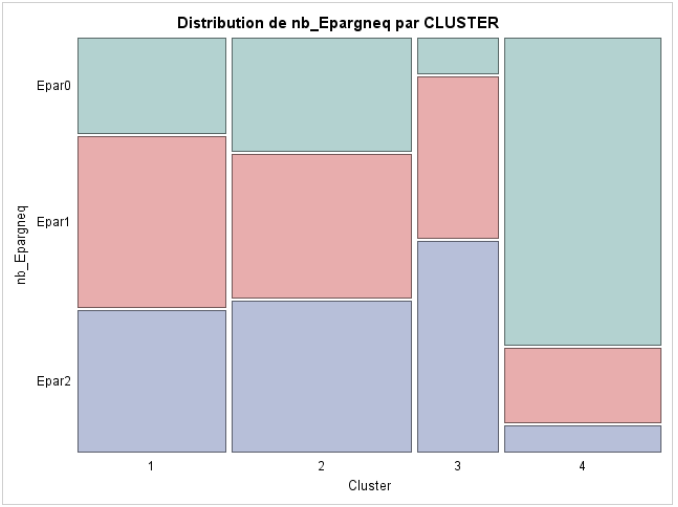


Résultat du clustering (7)

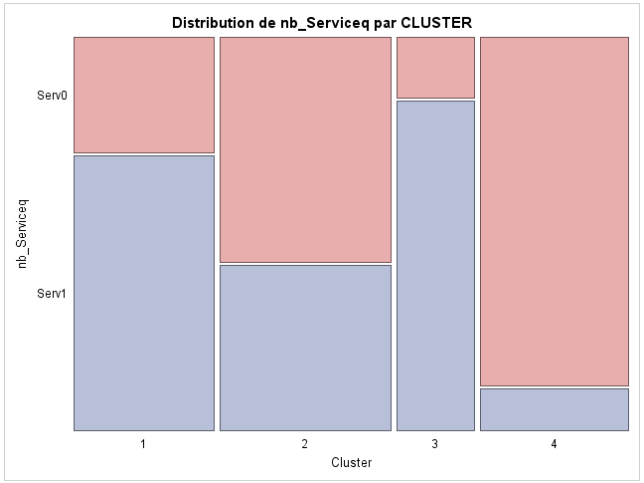


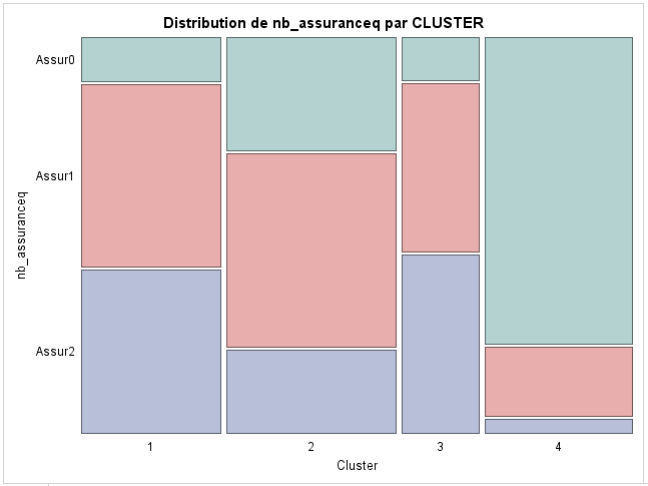


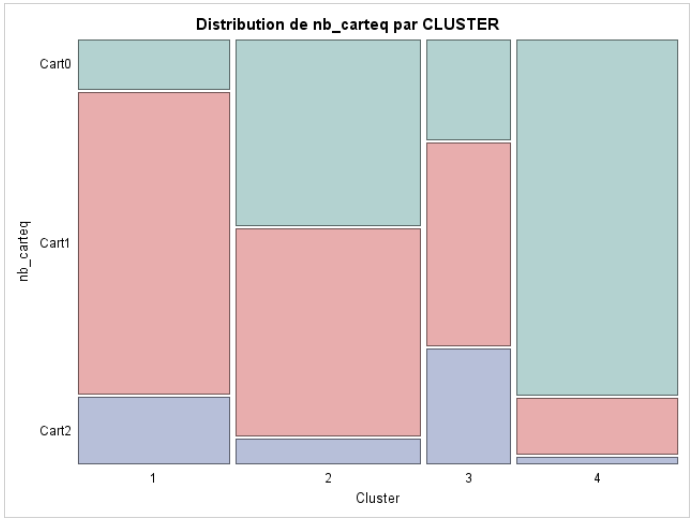


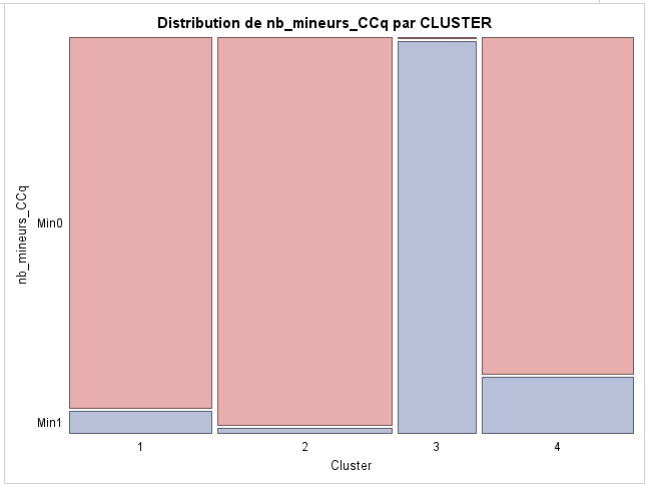


Résultat du clustering (8)

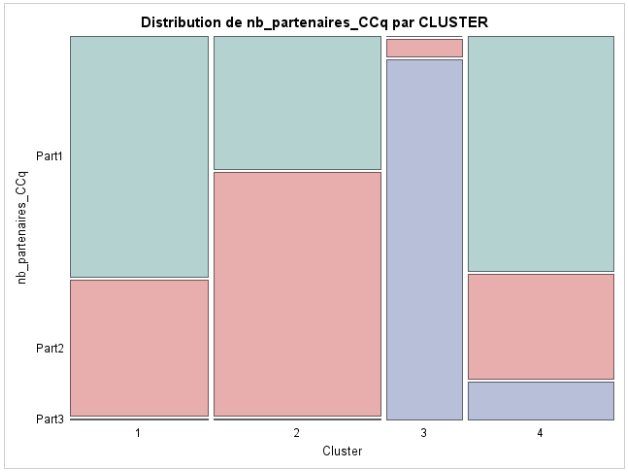


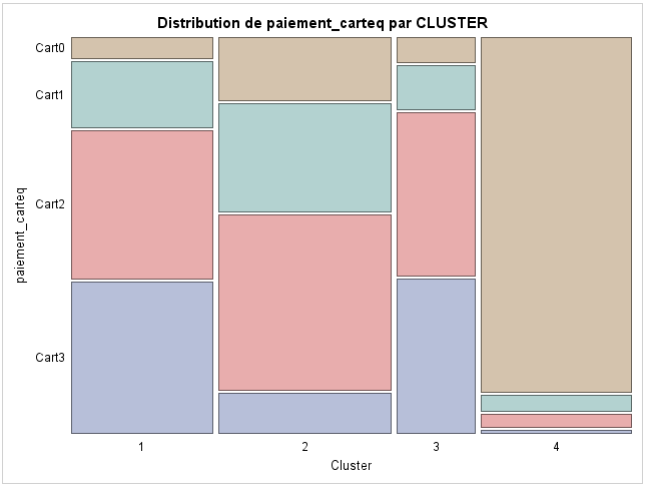


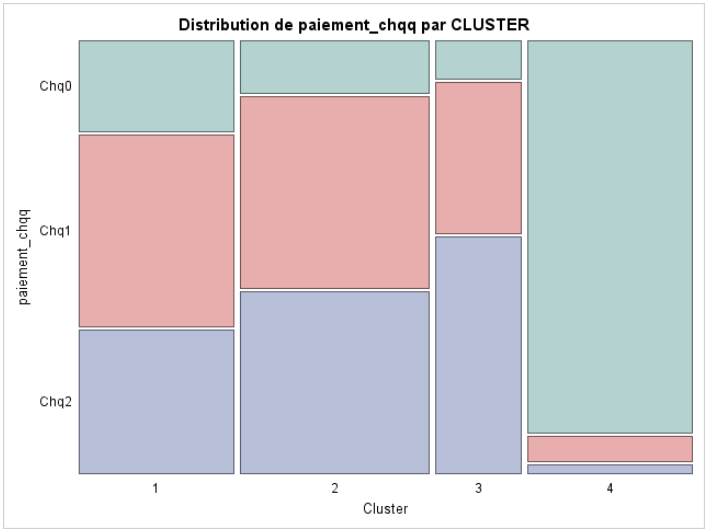


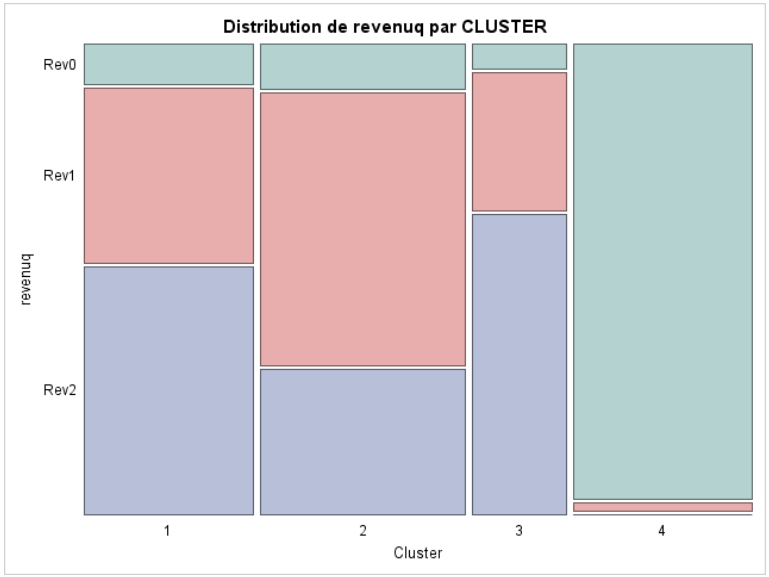


Résultat du clustering (9)

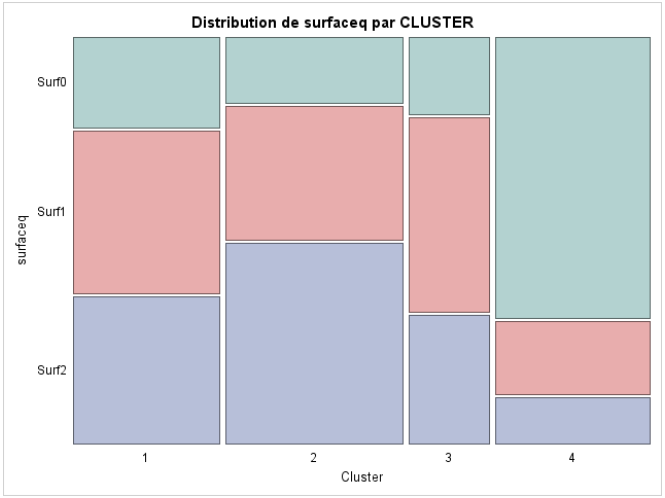


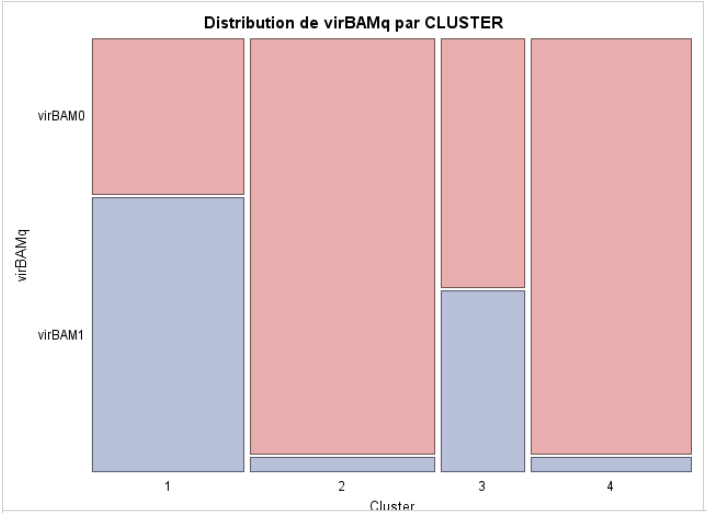


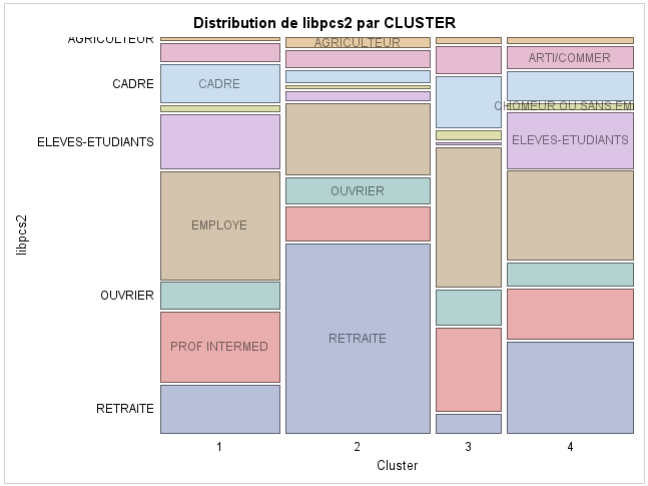


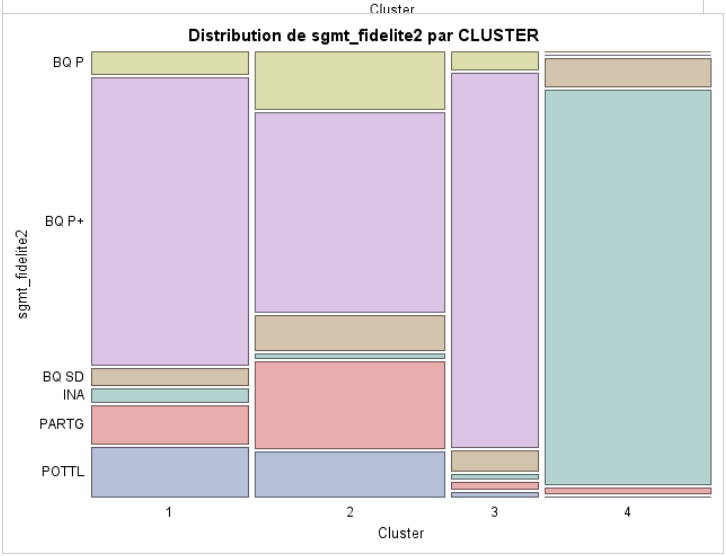


Résultat du clustering (10)

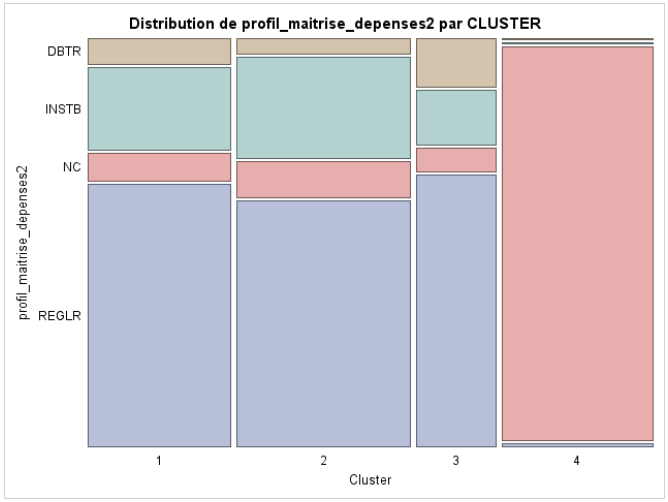


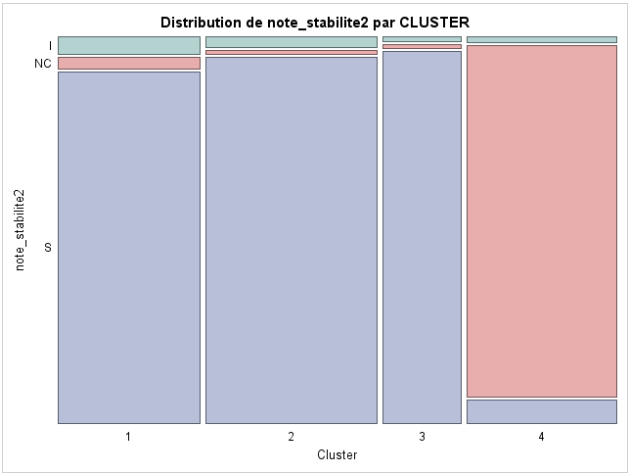


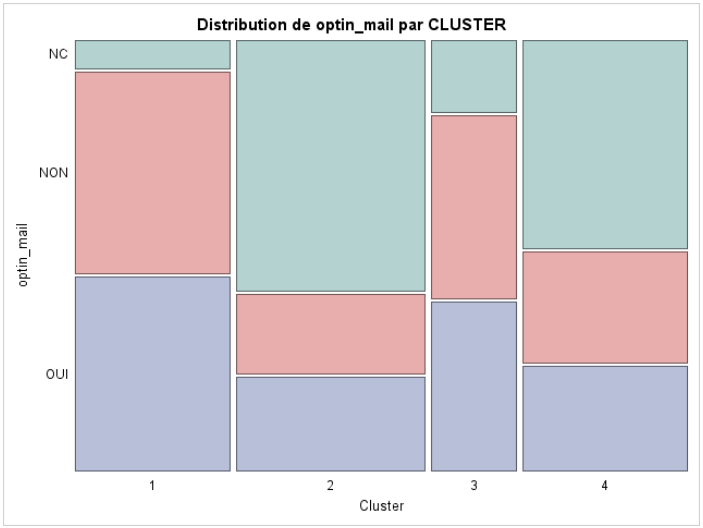


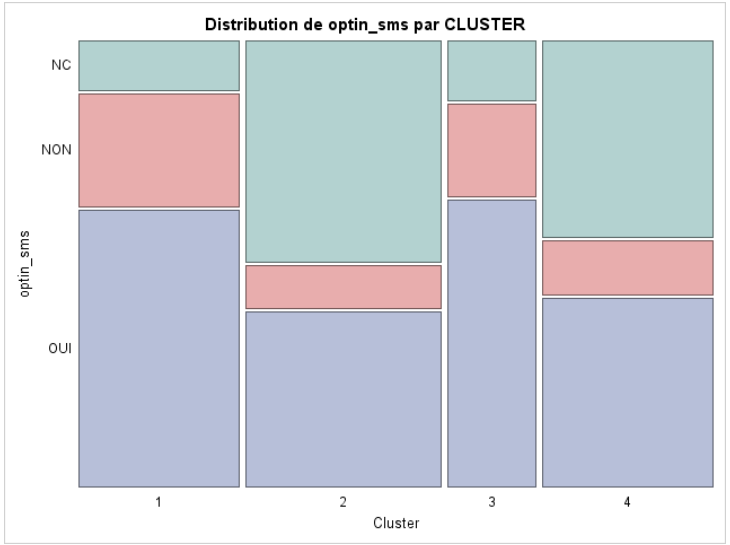


Résultat du clustering (11)

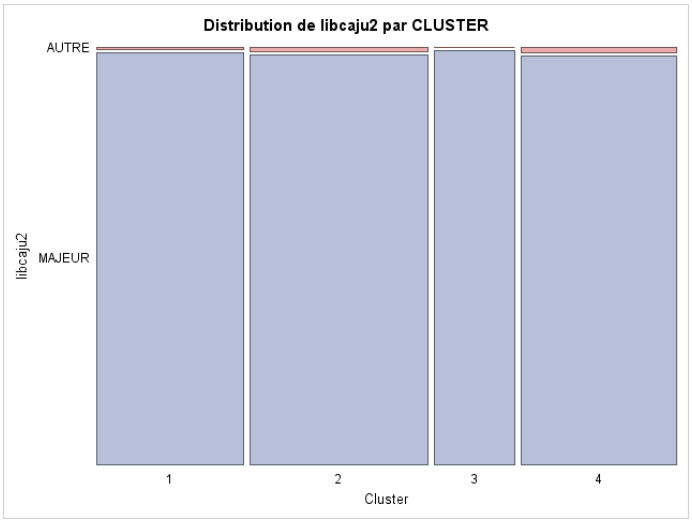


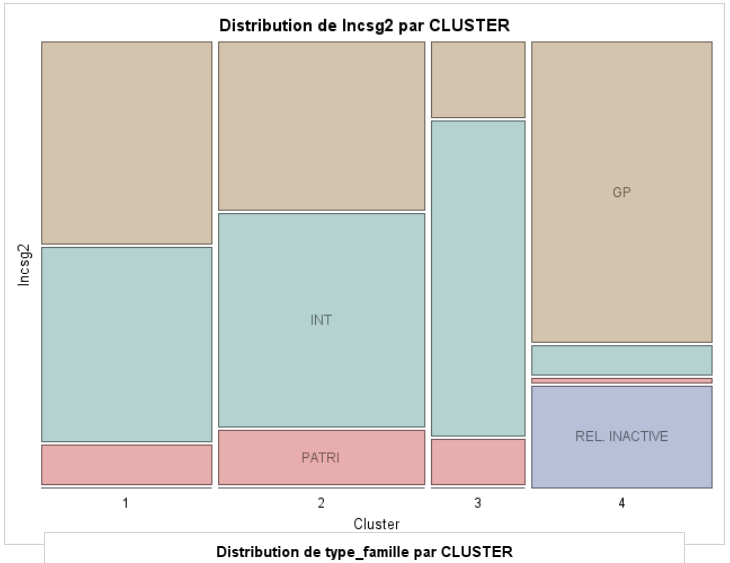


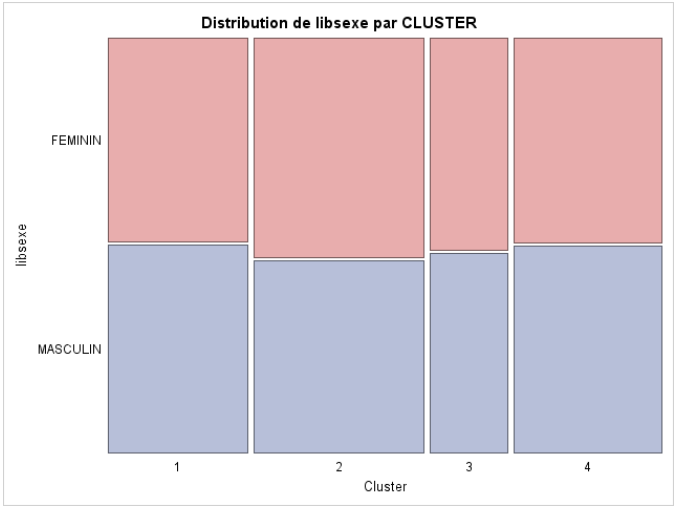


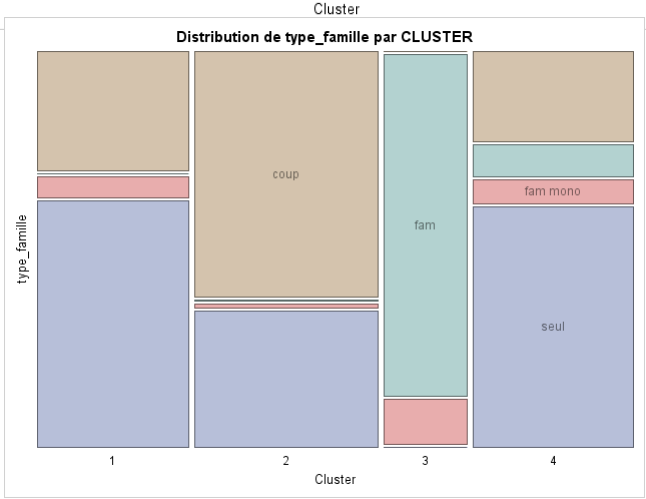


Résultat du clustering (12)

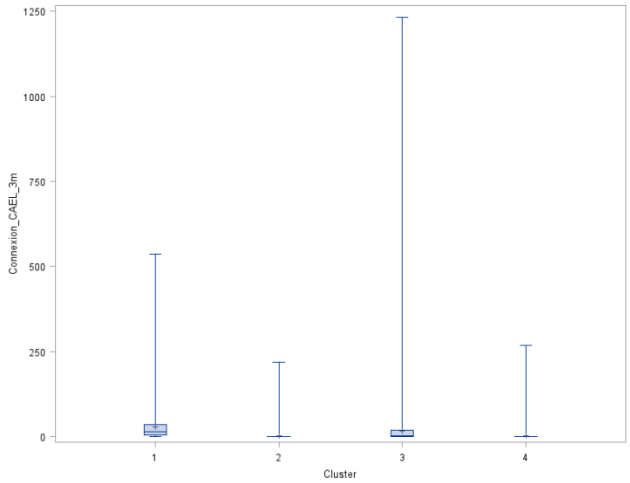


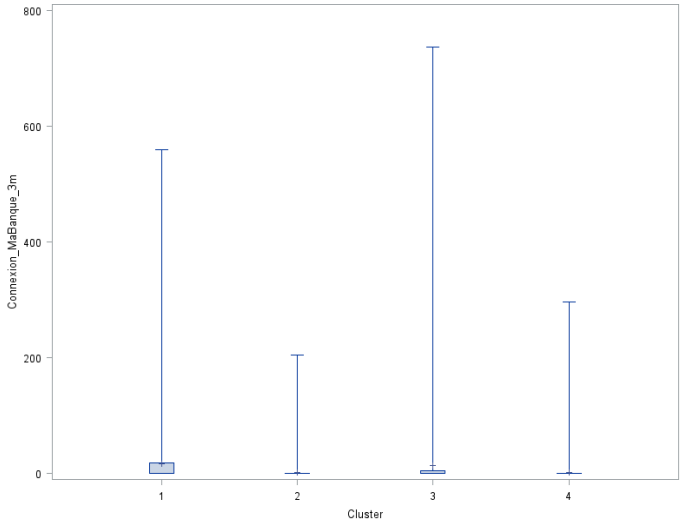


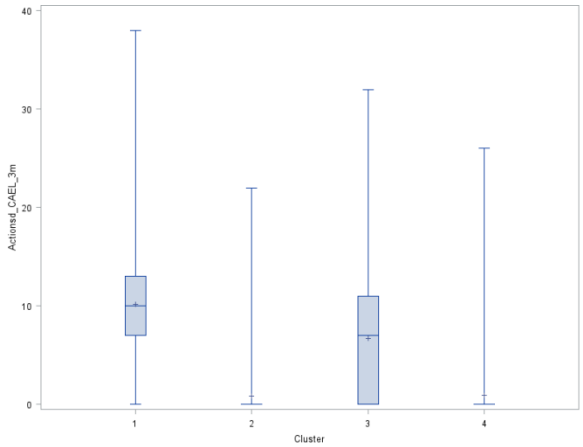


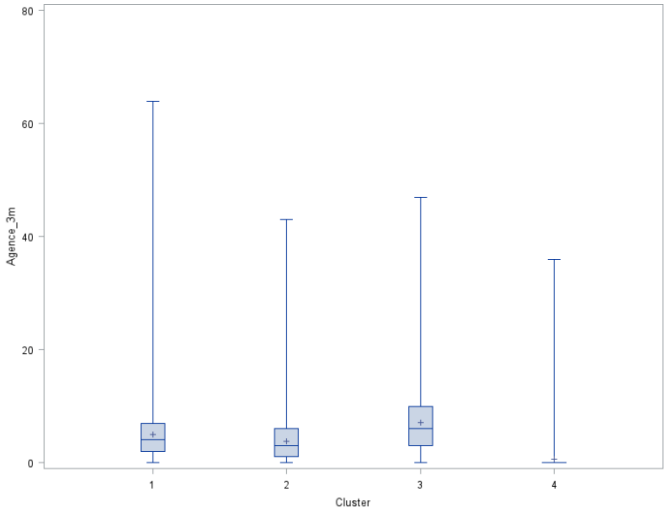


Distribution des variables de contact par cluster

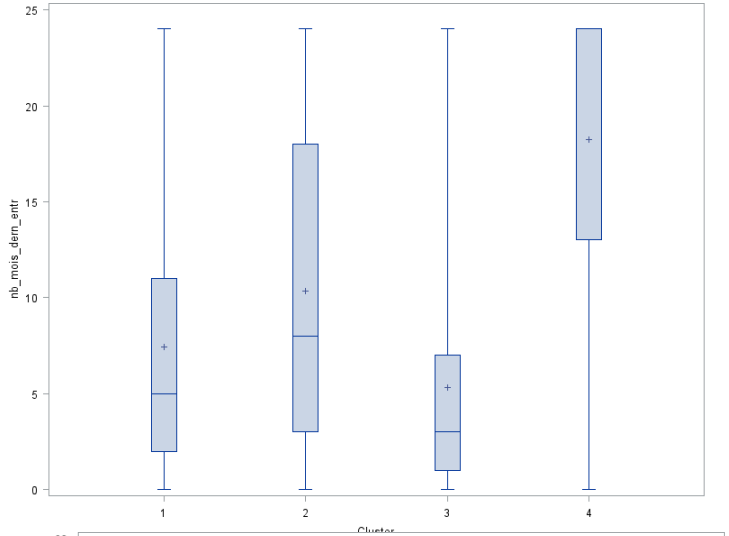


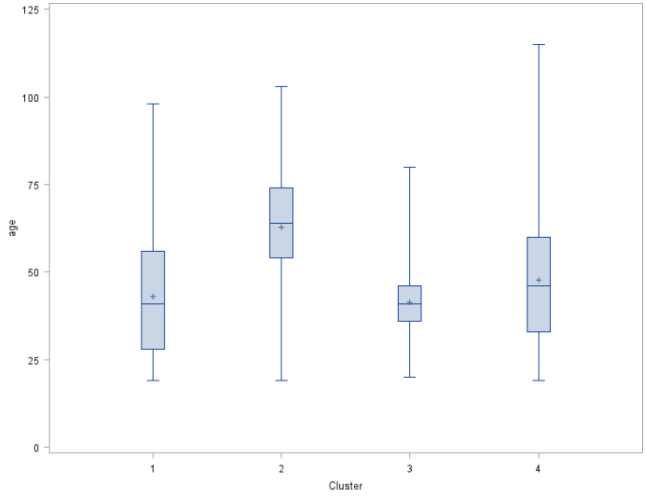


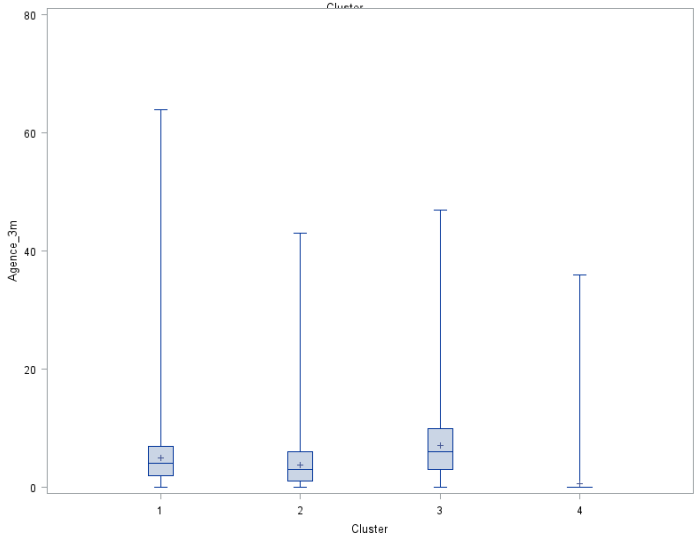


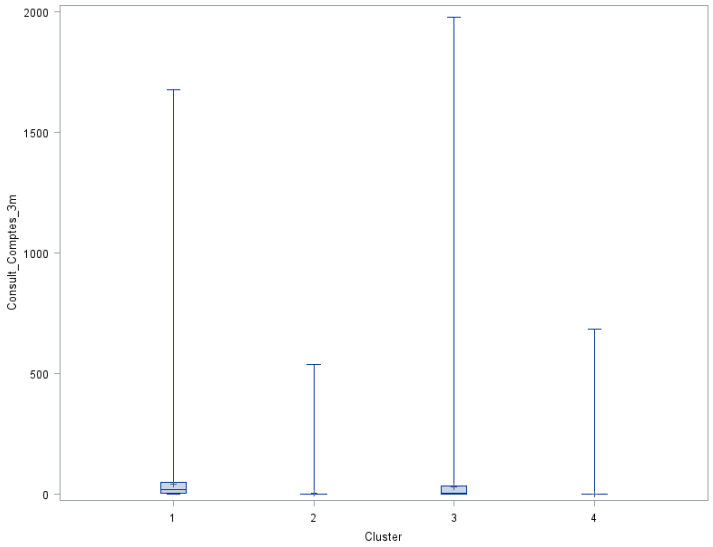


Distribution des variables de contact par cluster (2)









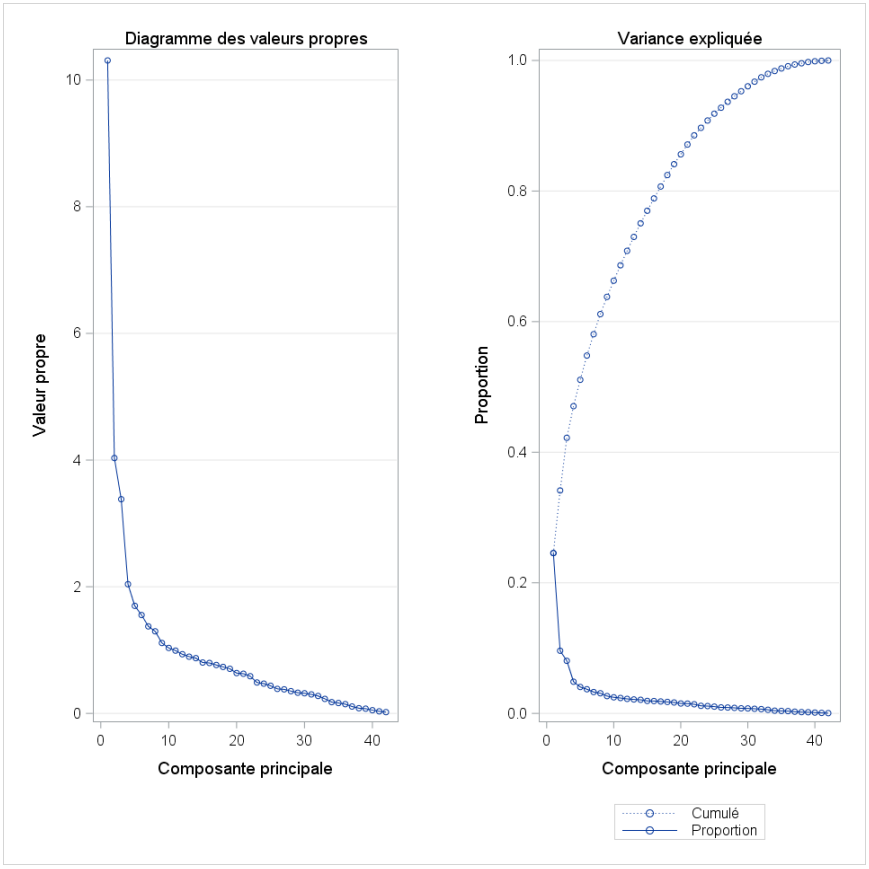
Résumé de la 1ere approche

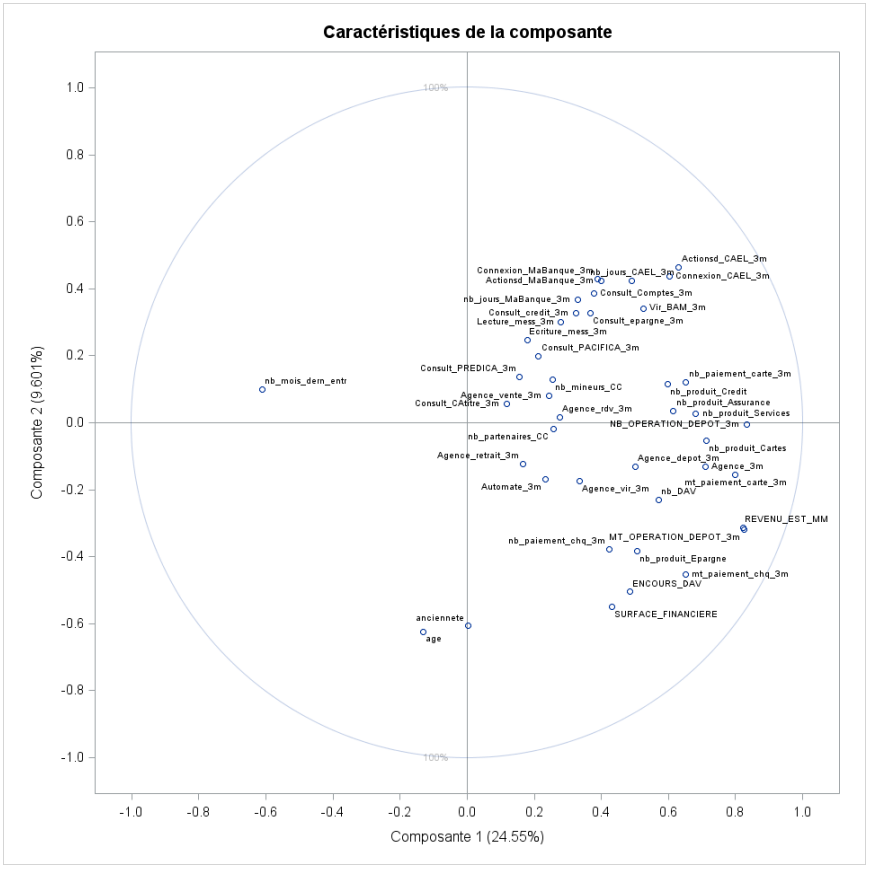
- Le cluster 1 est le cluster des clients **les plus actifs**, tous canaux confondus. 50 % d'entre eux sont optin mail et 63% sont optin sms. Ces clients font partie des plus « jeunes » (médiane et moyenne presque égales à celles du cluster 3, plus de variance dans ce cluster). **Canal privilégié** : mail ou sms.
- Le cluster 2 est le cluster des clients les **plus âgés** : l'essentiel des contacts se font en agence (automates compris). **Canal privilégié** : courrier ou contact conseiller
- Le cluster 3 est celui des clients **fidèles**. Les canaux les plus consommés par ces clients sont l'agence et CAEL. **Canal privilégié** : mail + contact direct?
- Le cluster 4 est celui des « inactifs »

2eme approche : ACP sur les variables quanti - Méthode

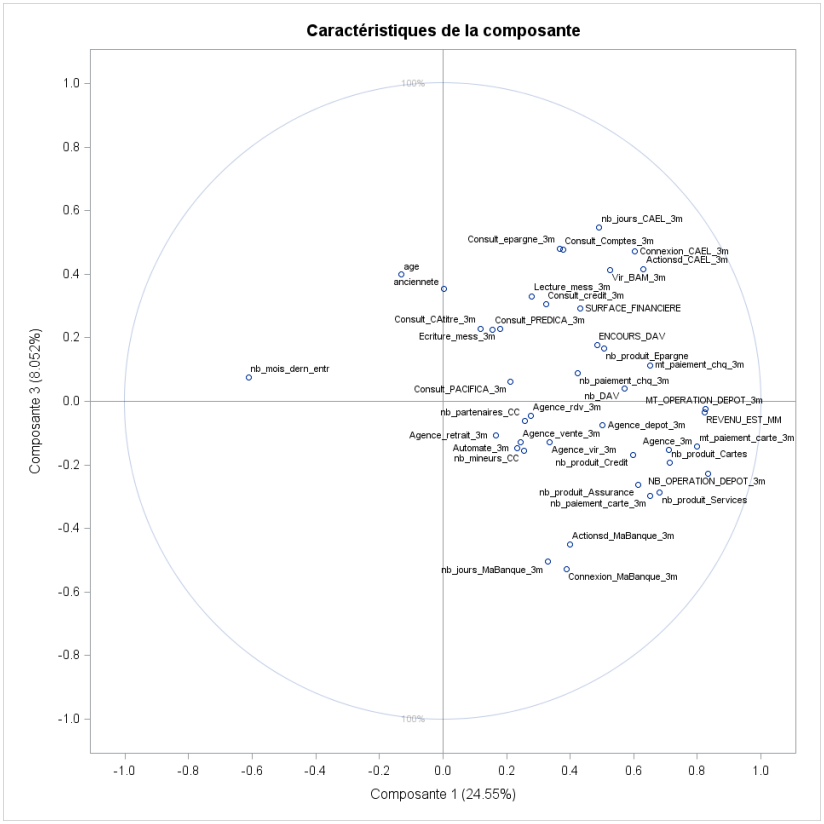
- Sélectionner les variables quantitatives intéressantes et effectuer des transformations si nécessaire (ex : transformation en log…)
- Réaliser une ACP sur ces variables pour réduire la dimension
- Faire un k-means avec 2000 classes (10% de n) sur les résultats de l'ACP
- Faire un clustering hiérarchique en utilisant les barycentres du k-means
- Choisir le nombre de classes optimales
- Faire un k-means en précisant le nombre de classes optimales
- (Préciser méthode des outliers SAS)

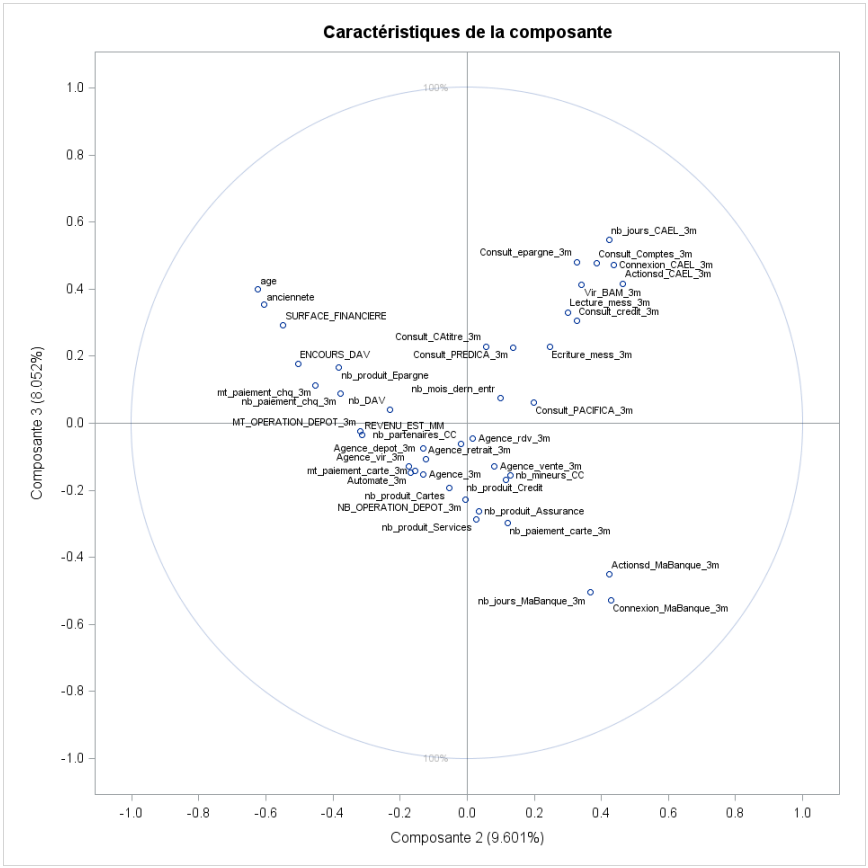
Résultats de l'ACP





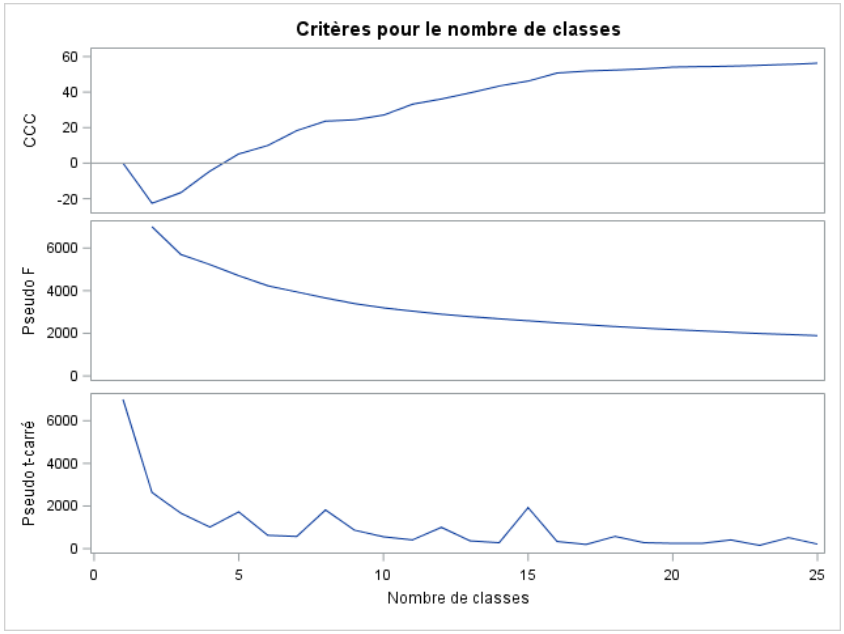
Résultats de l'ACP (2)

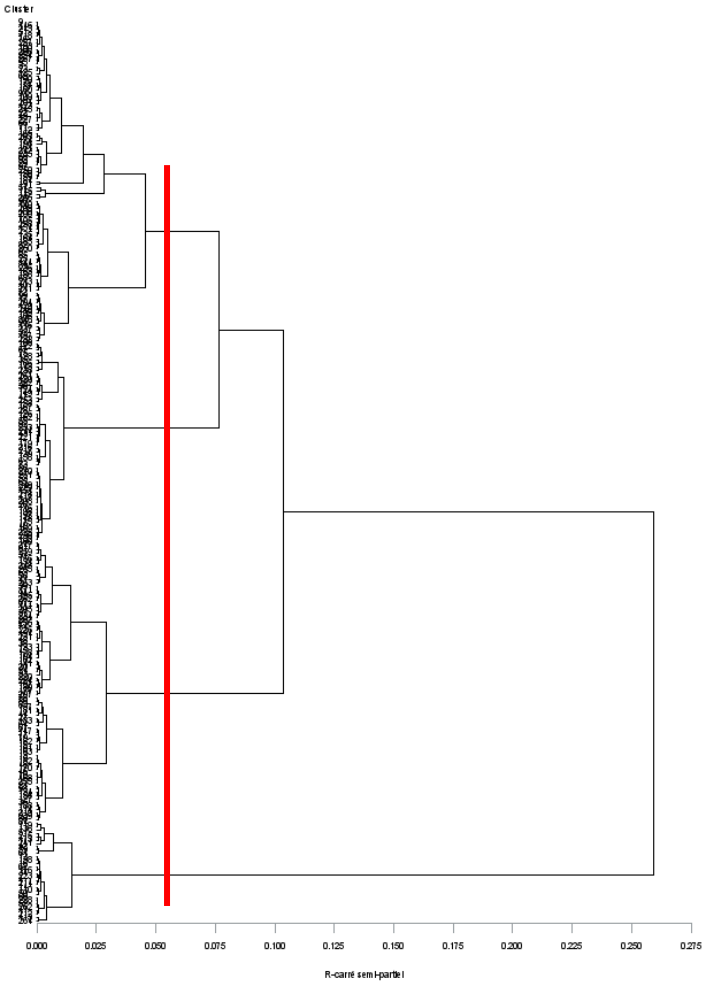




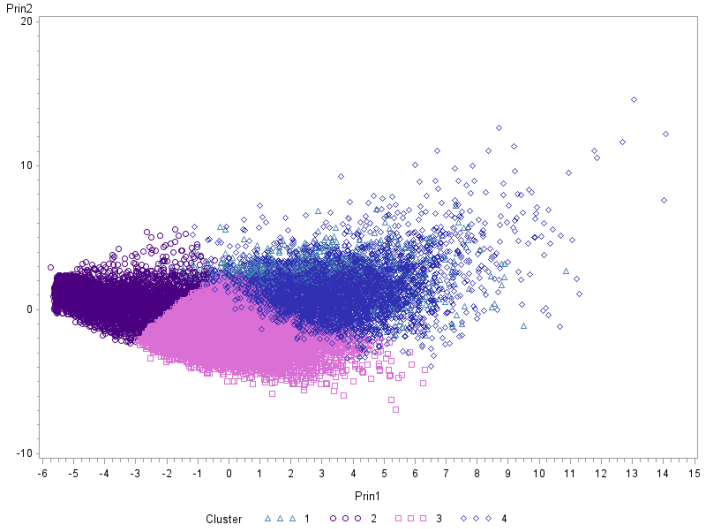
Résultats de la CAH

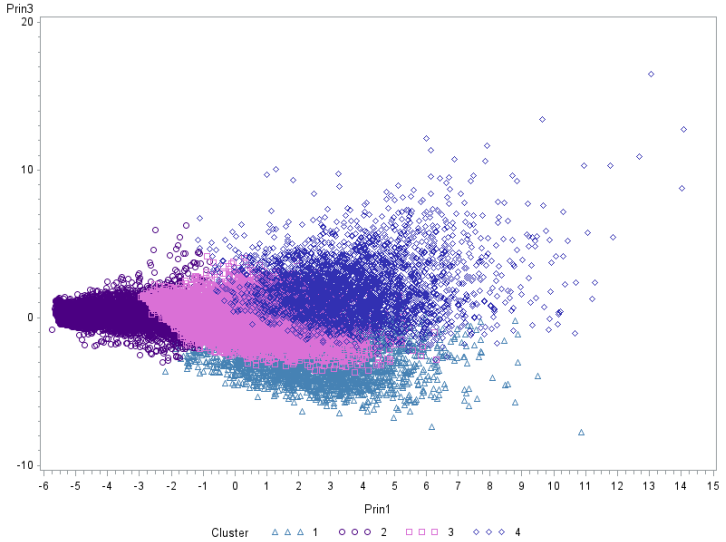
- 4 classes suggérées

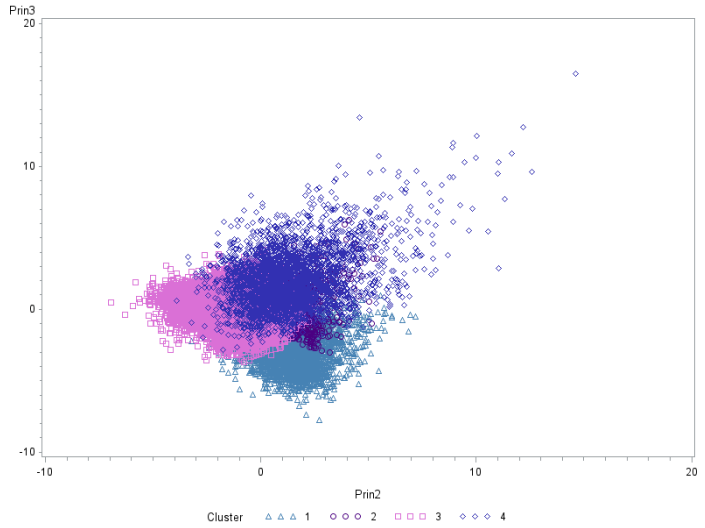




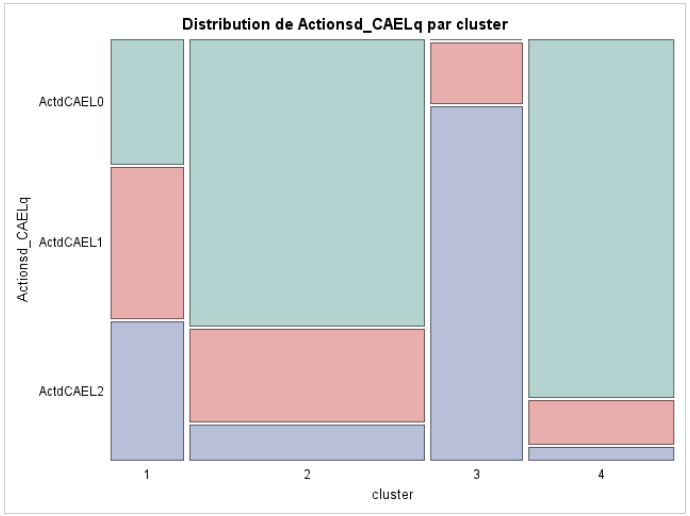
Représentation des clusters dans le plan de l'ACP

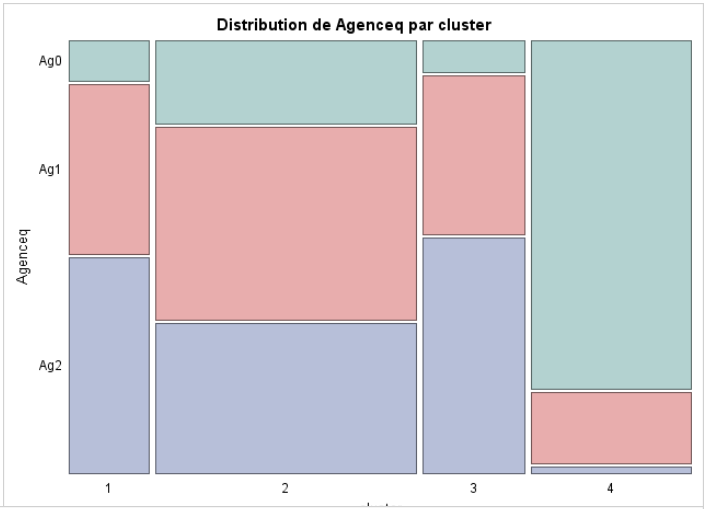


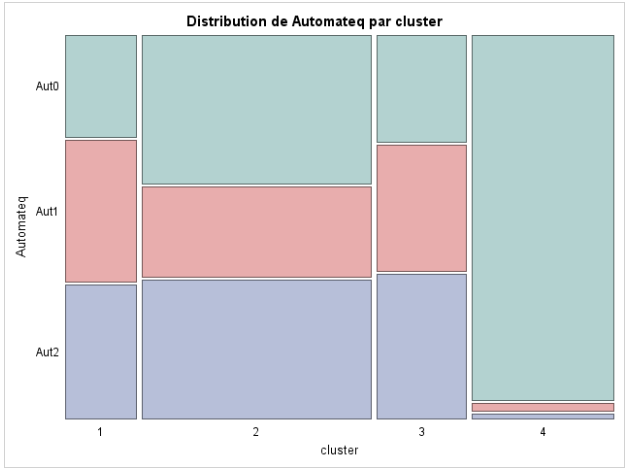


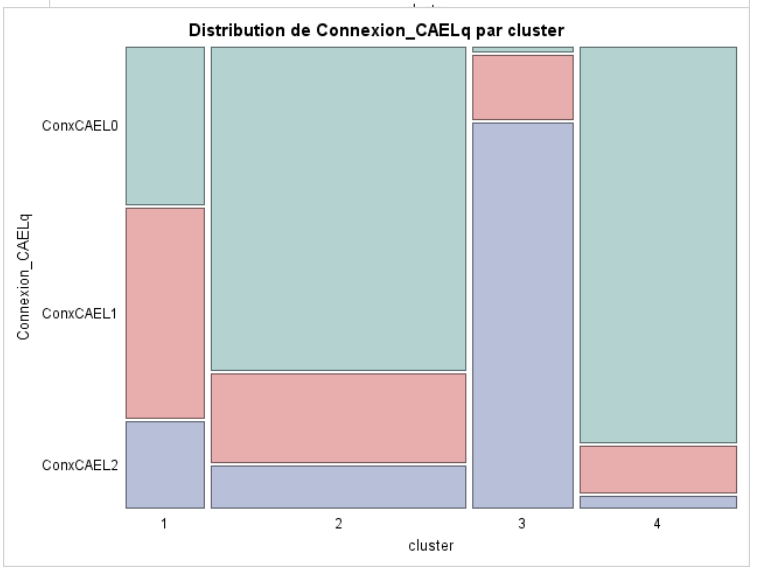


Résultats du clustering

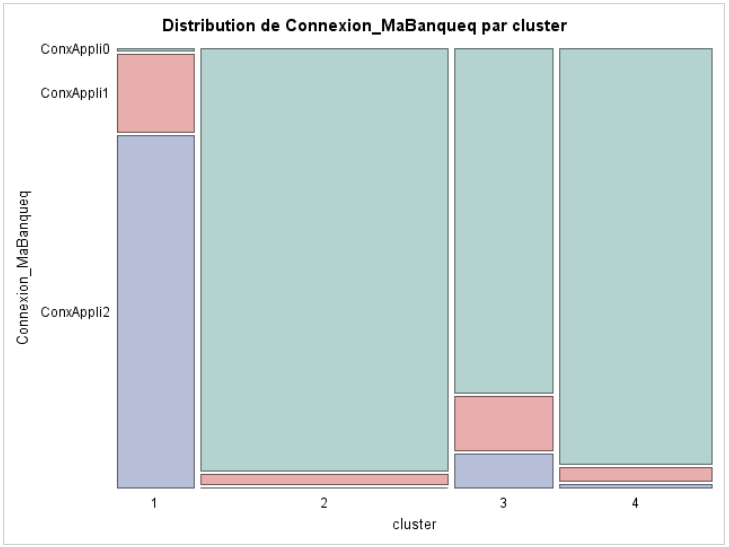


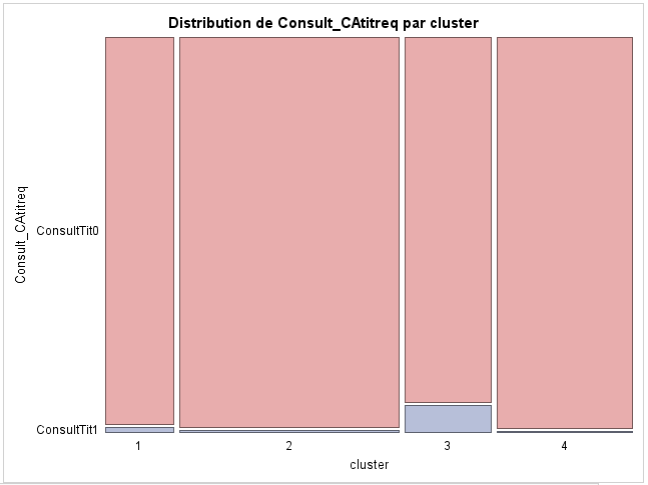


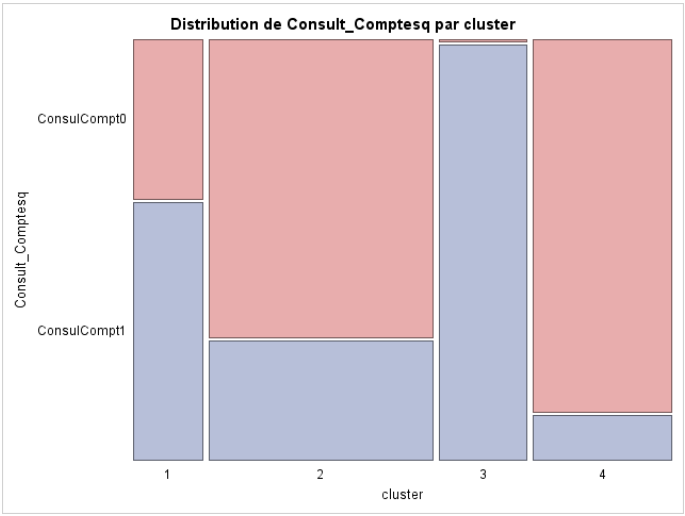


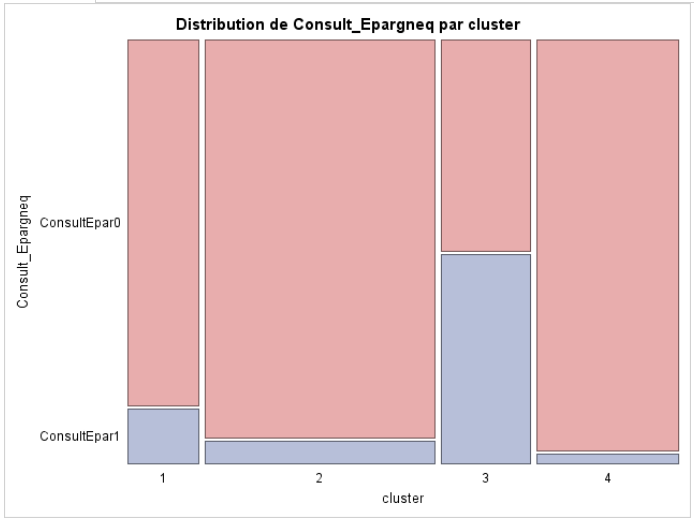


Résultats du clustering (2)

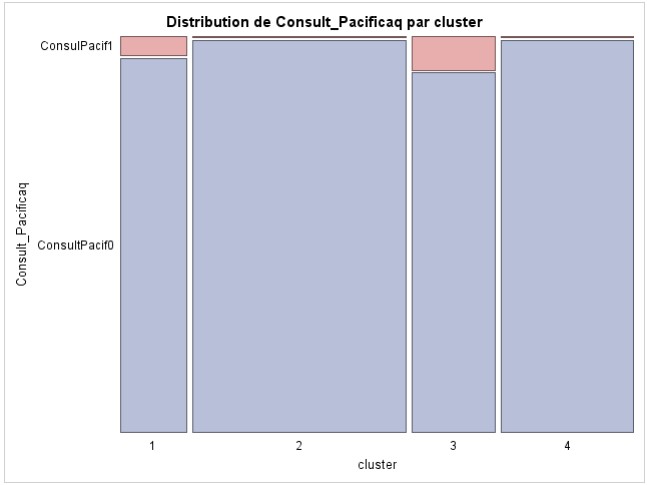


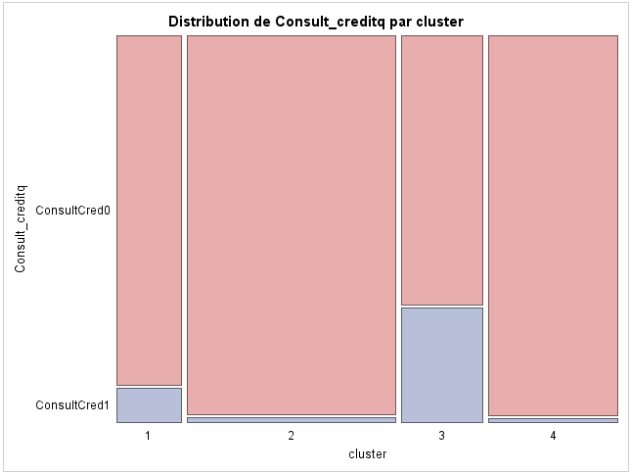


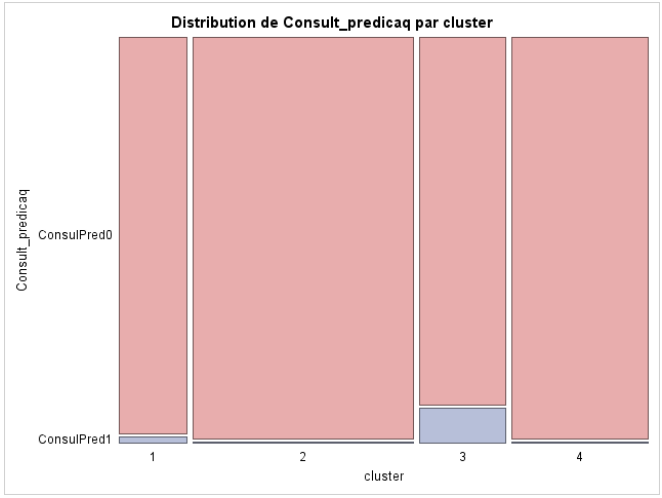


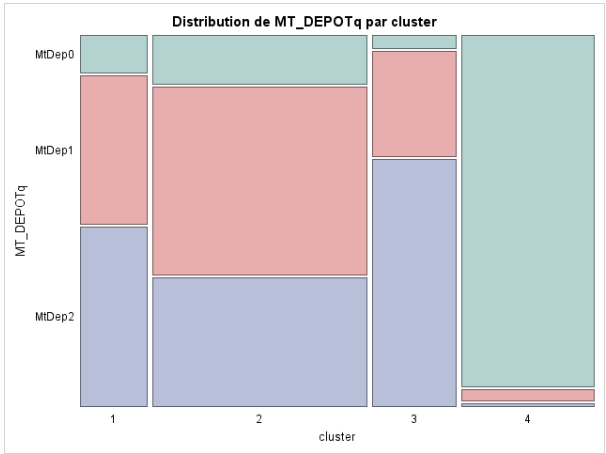


Résultats du clustering (3)

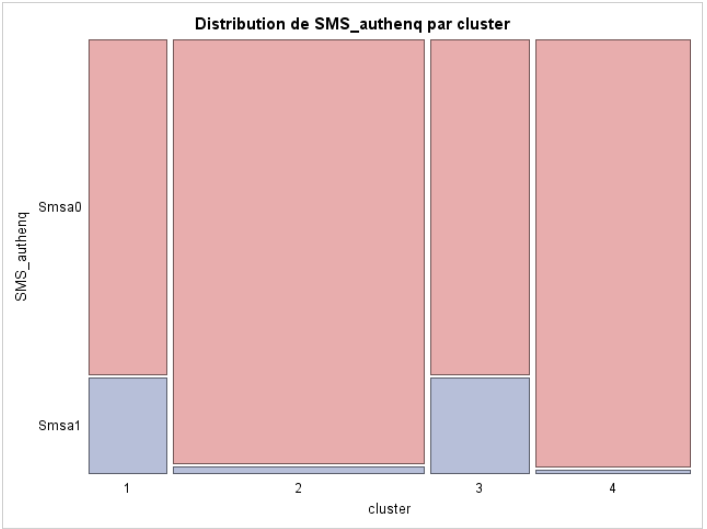


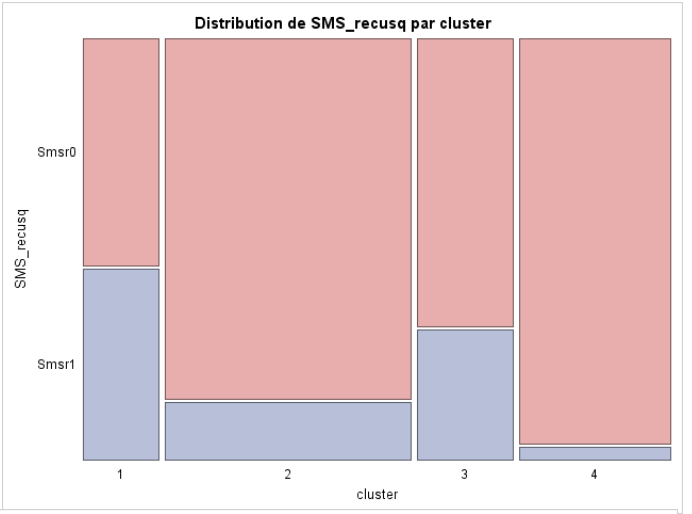


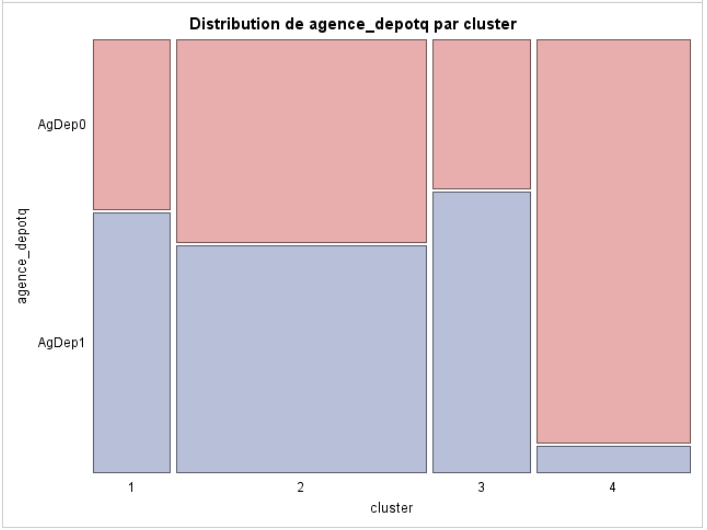


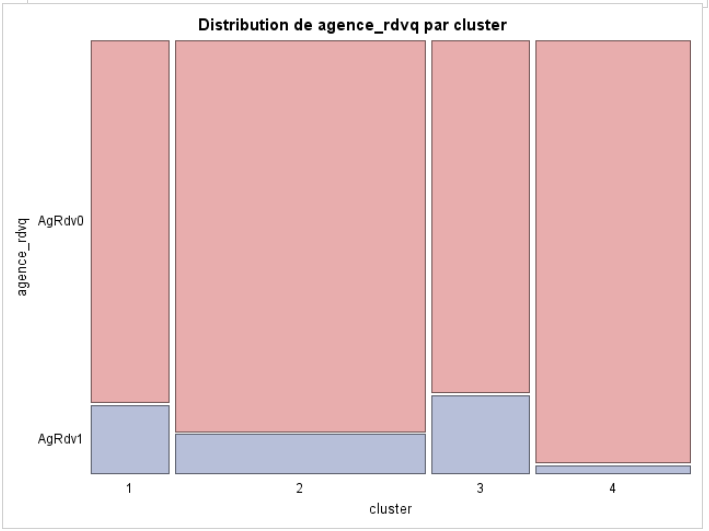


Résultats du clustering (4)

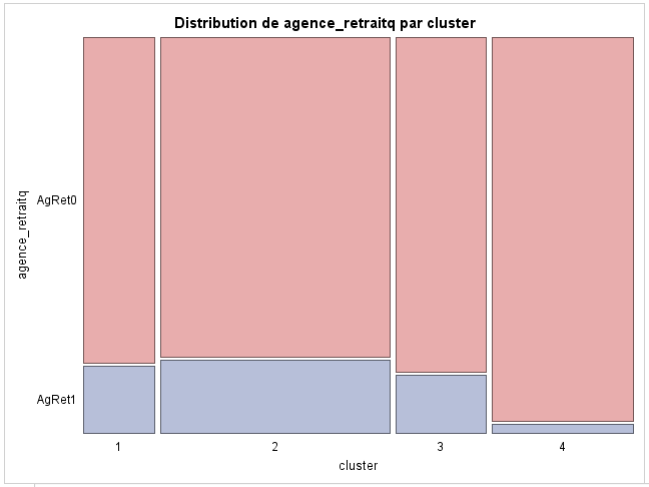


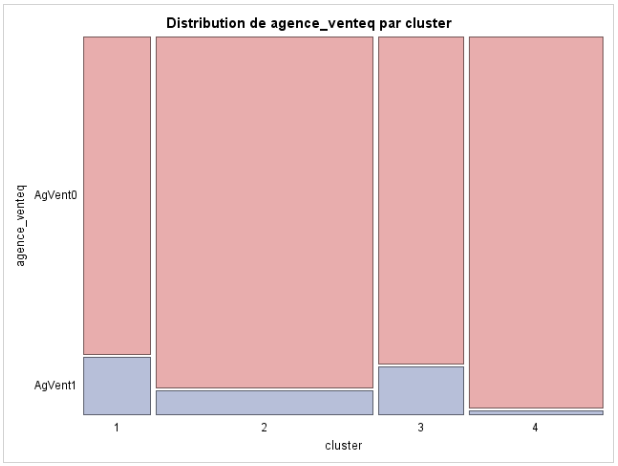


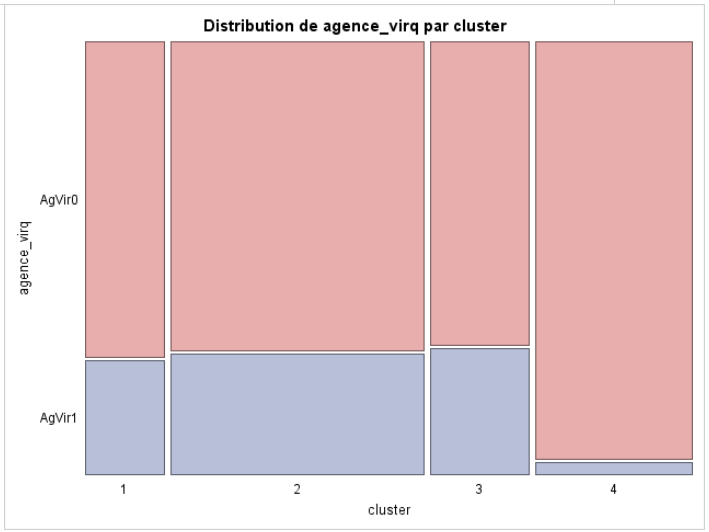


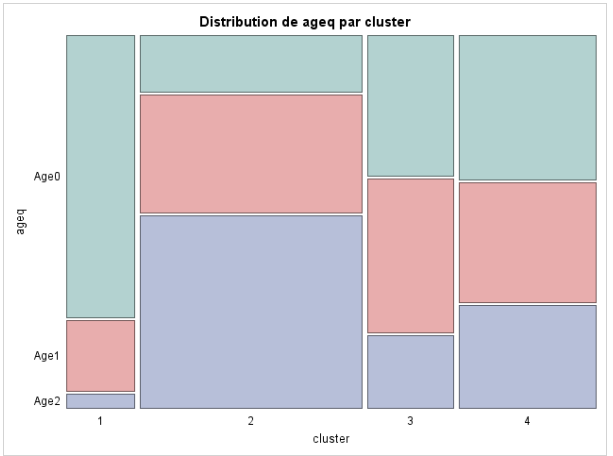


Résultats du clustering (5)

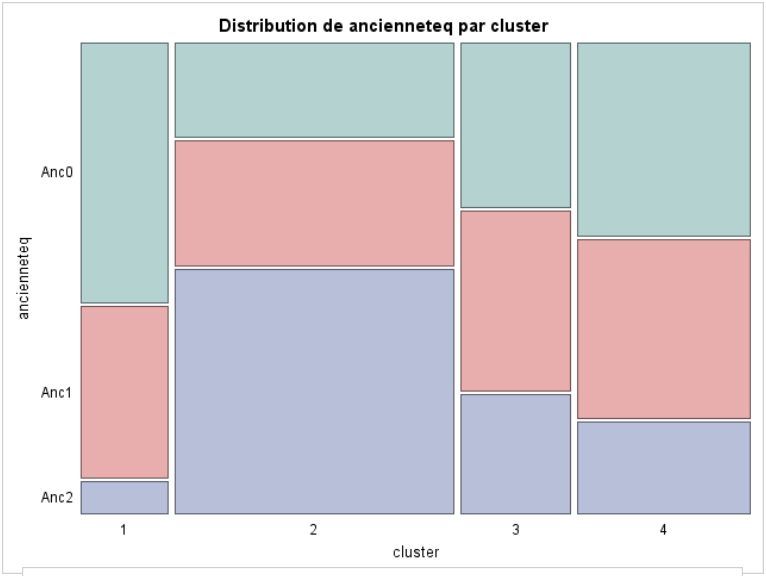


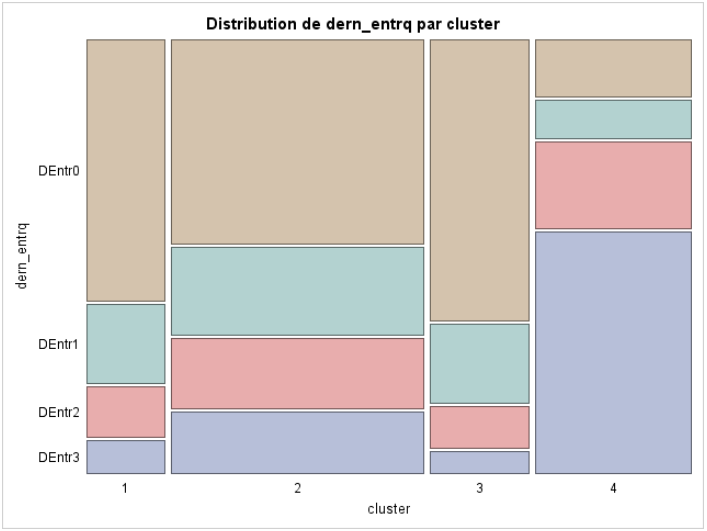


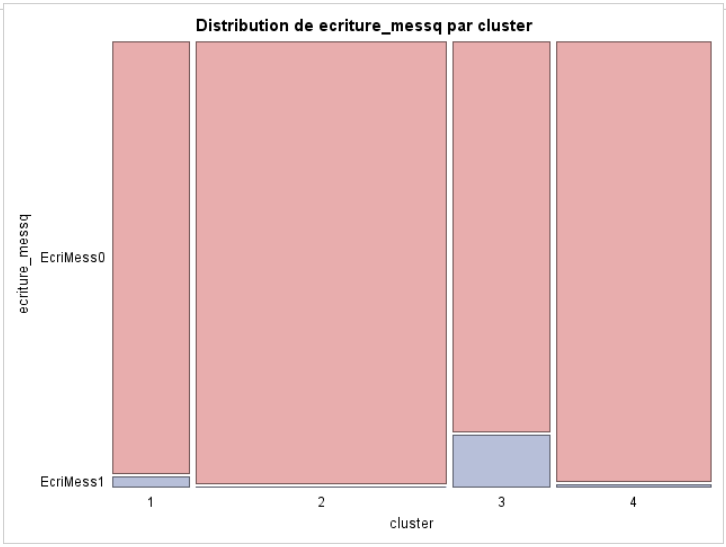


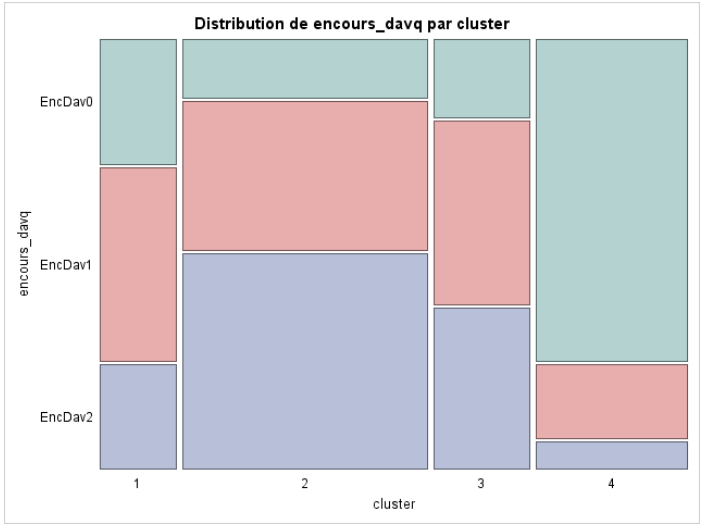


Résultats du clustering (6)

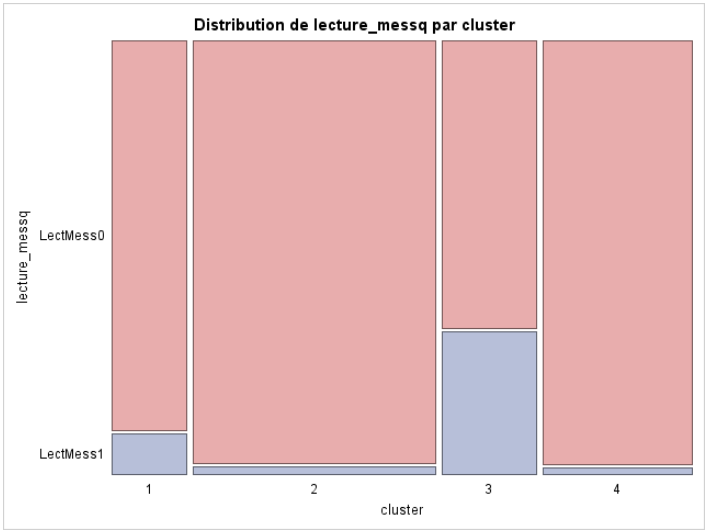


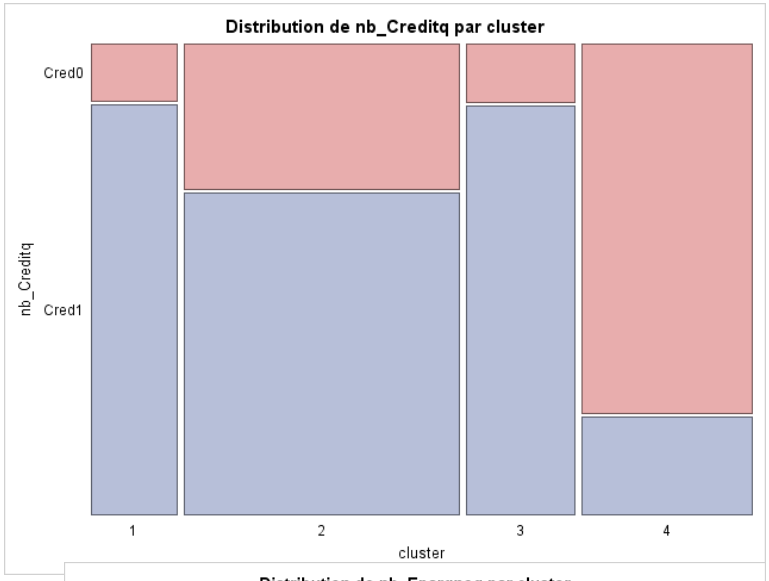


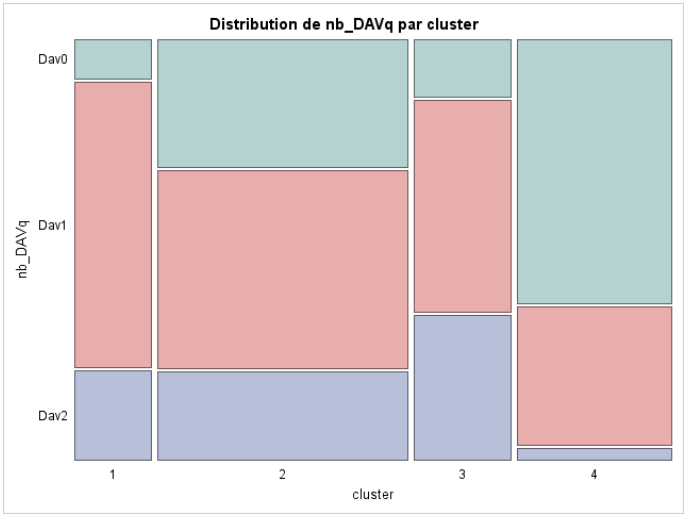


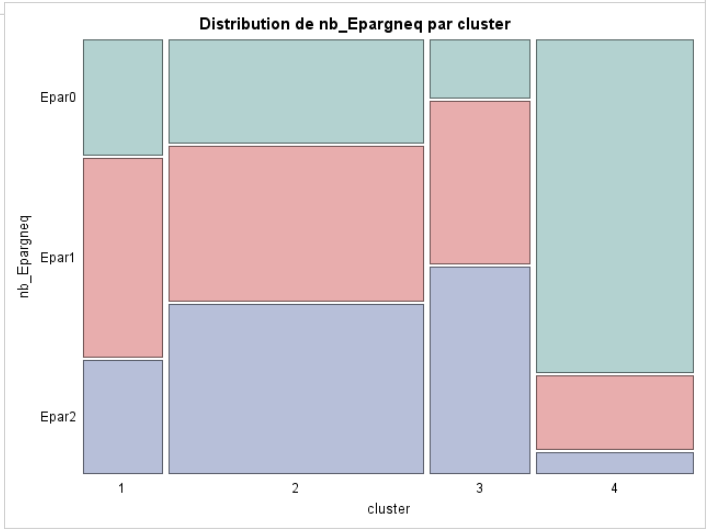


Résultats du clustering (7)

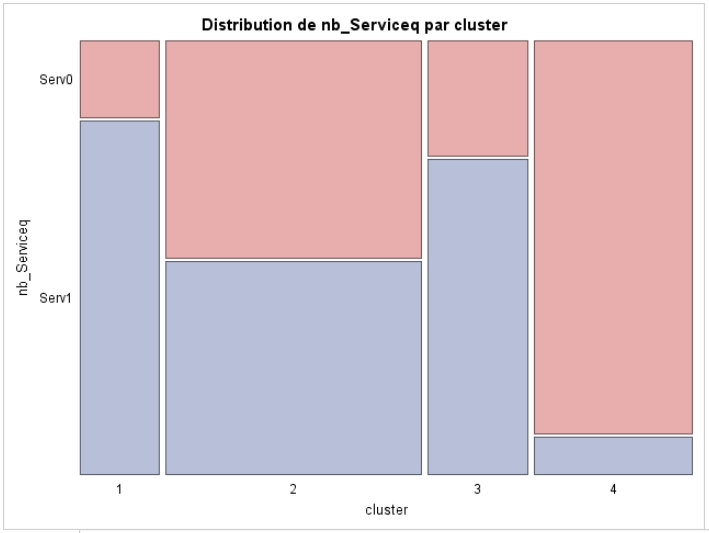


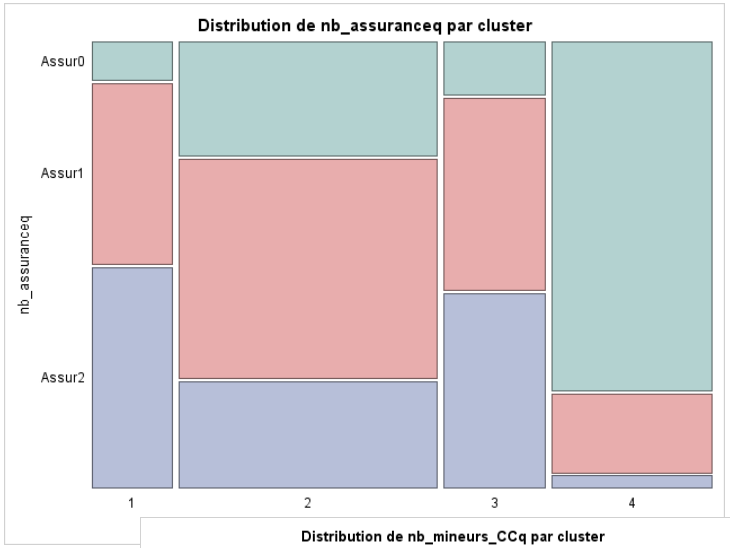


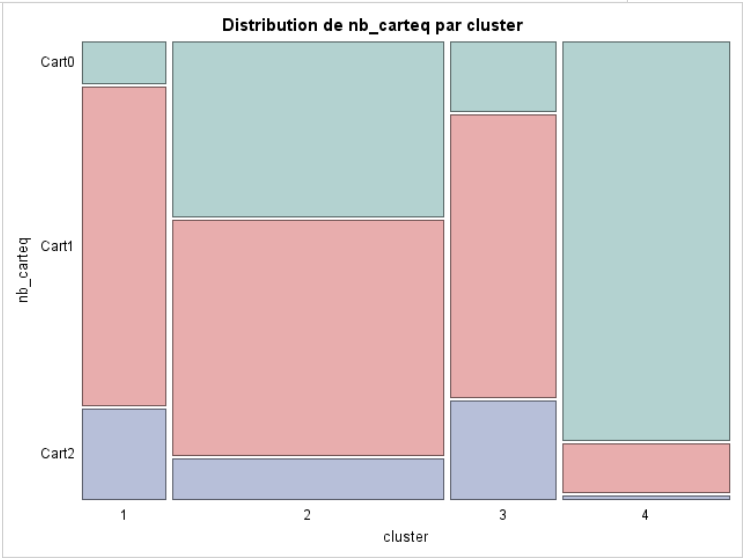


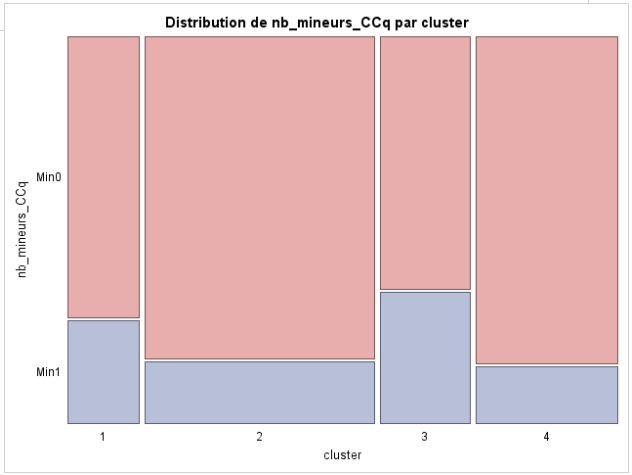


Résultats du clustering (8)

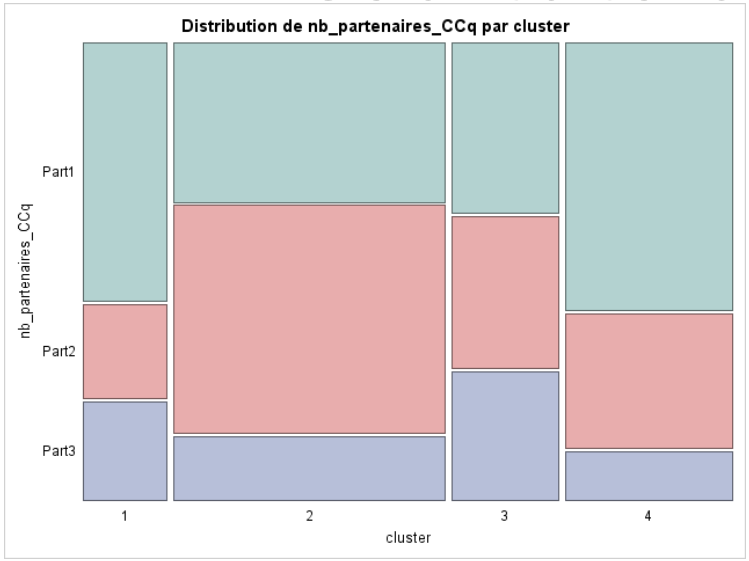


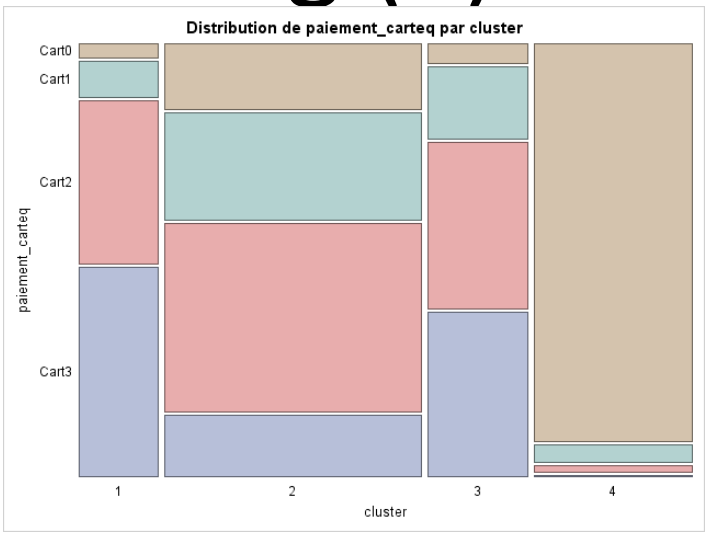


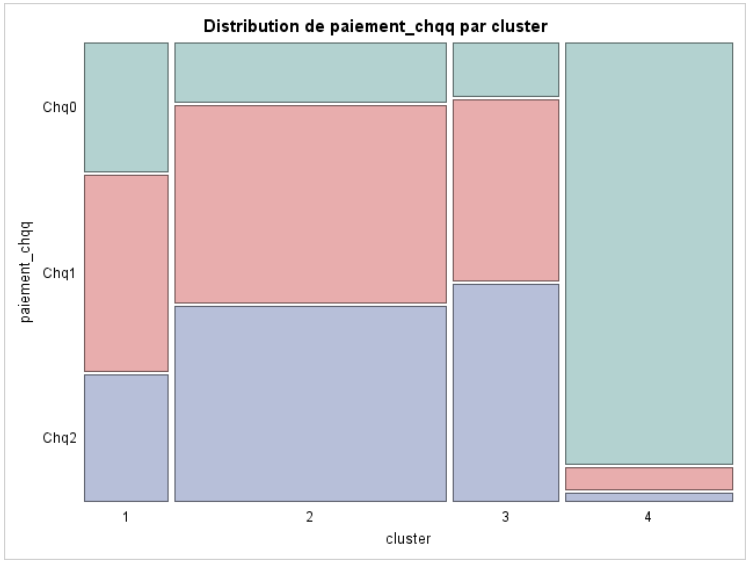


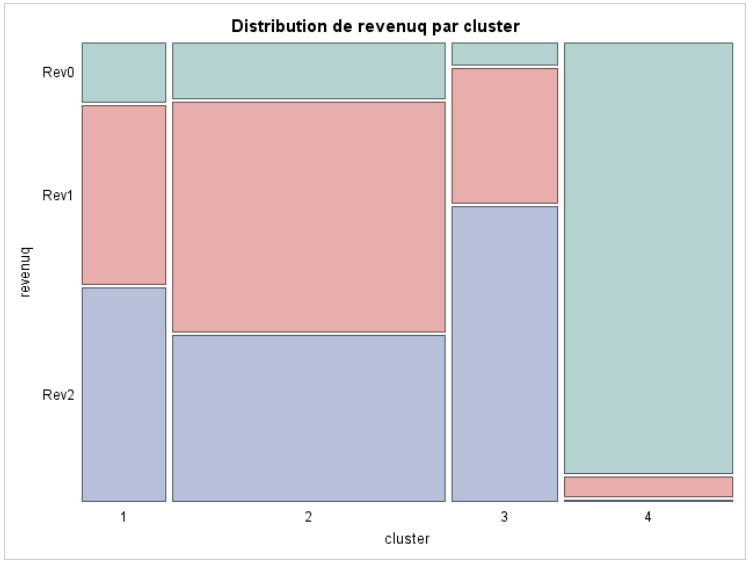


Résultats du clustering (9)

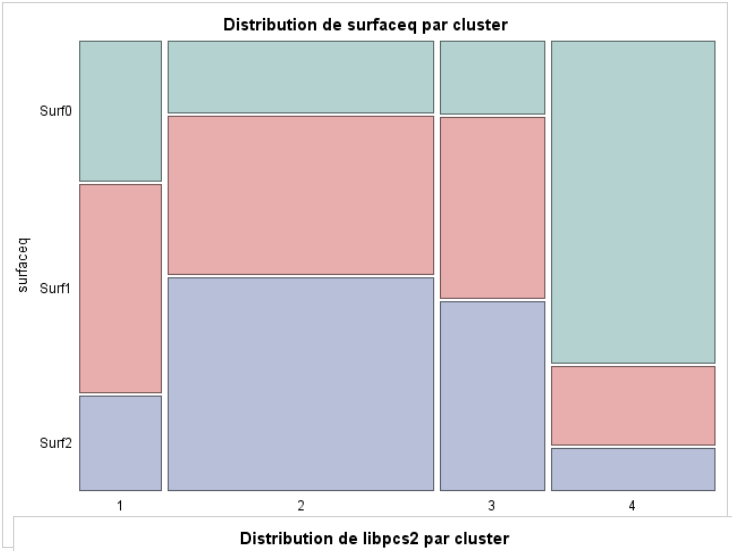


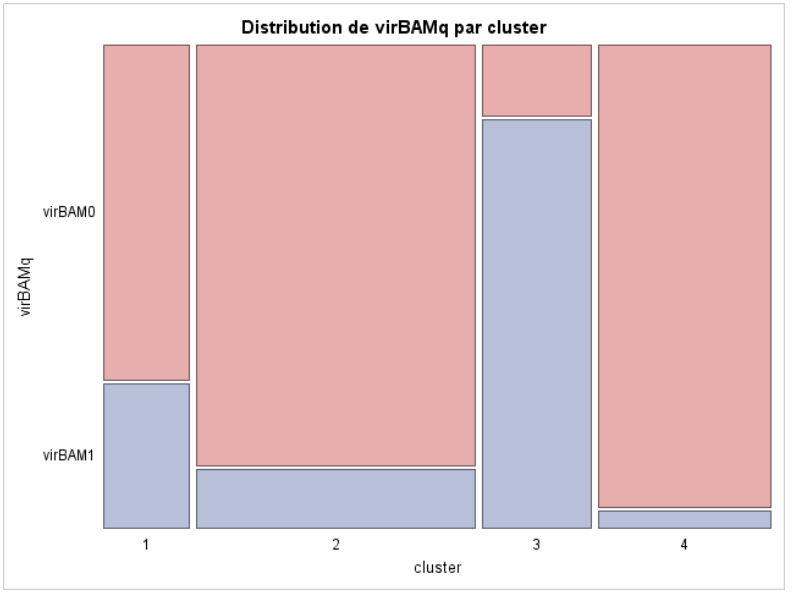


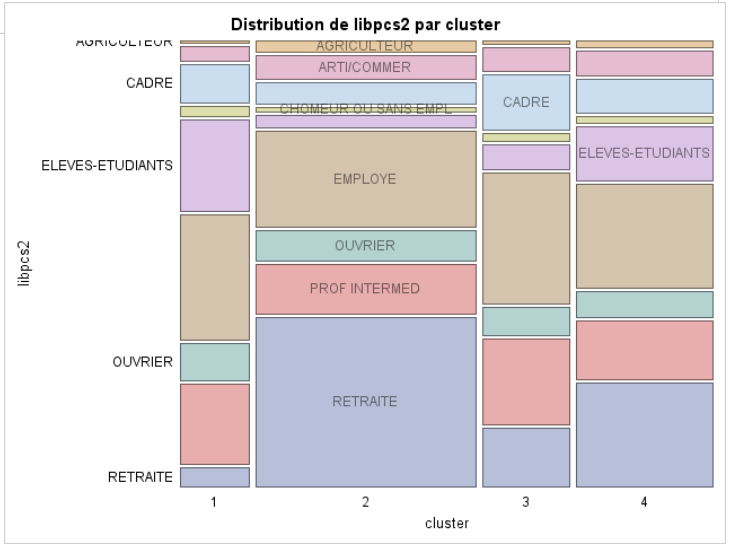


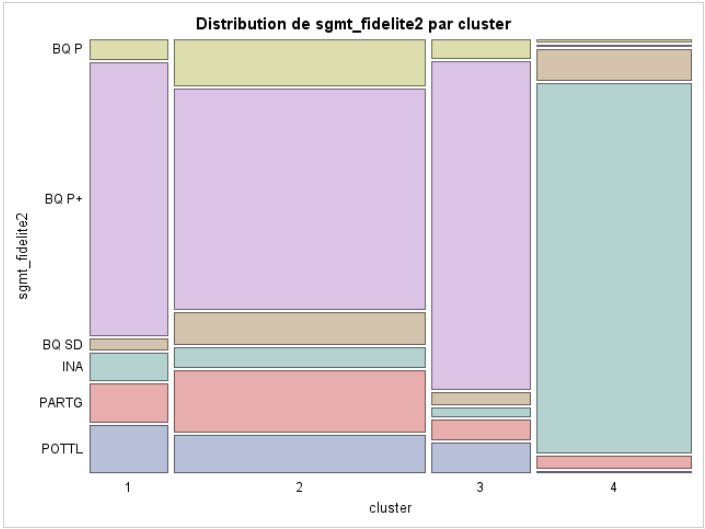


Résultats du clustering (10)

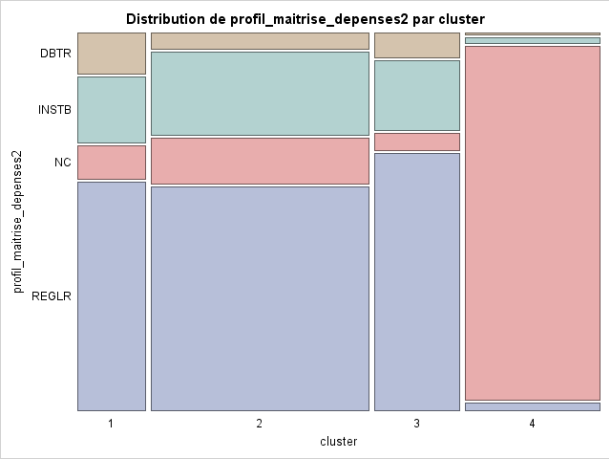


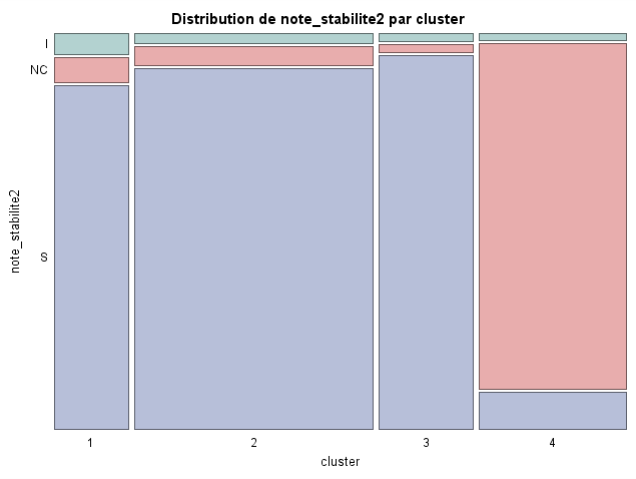


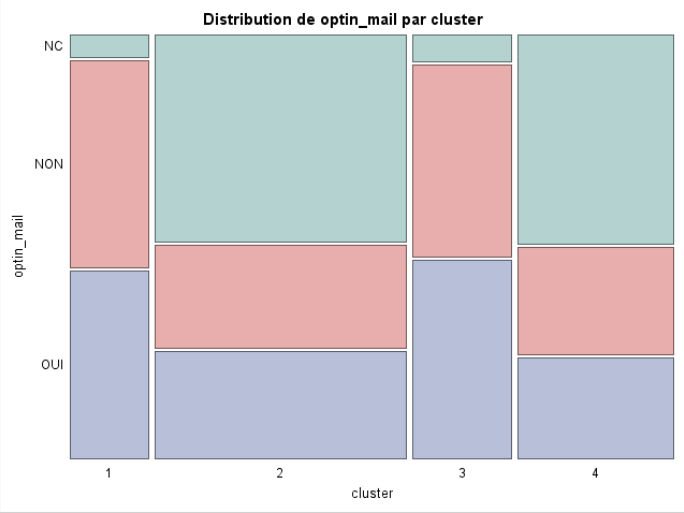


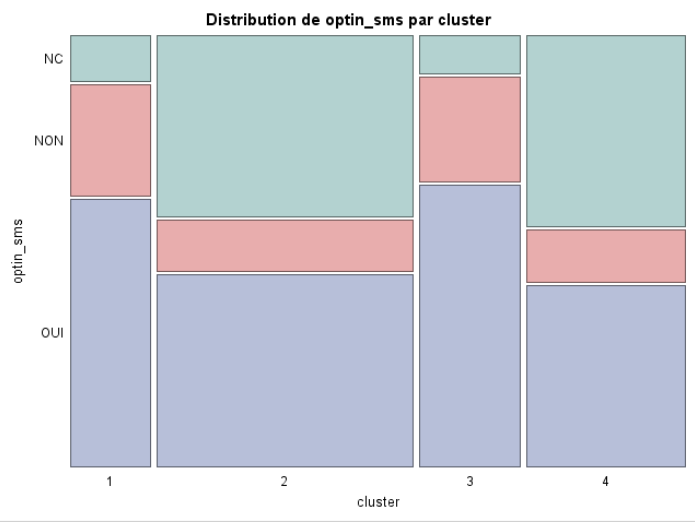


Résultats du clustering (11)



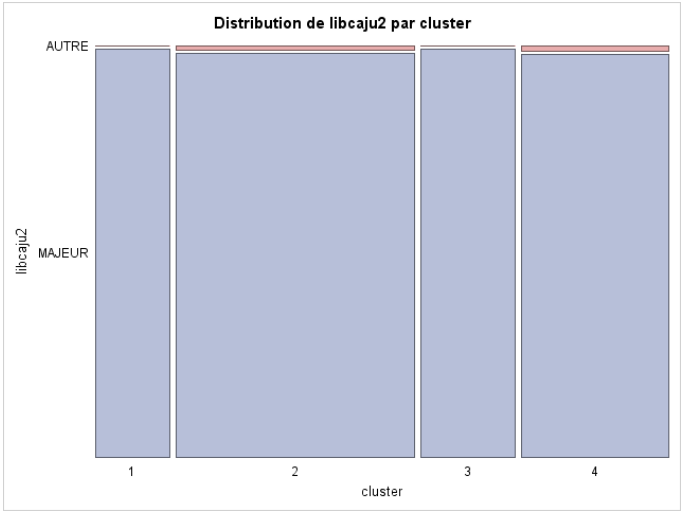


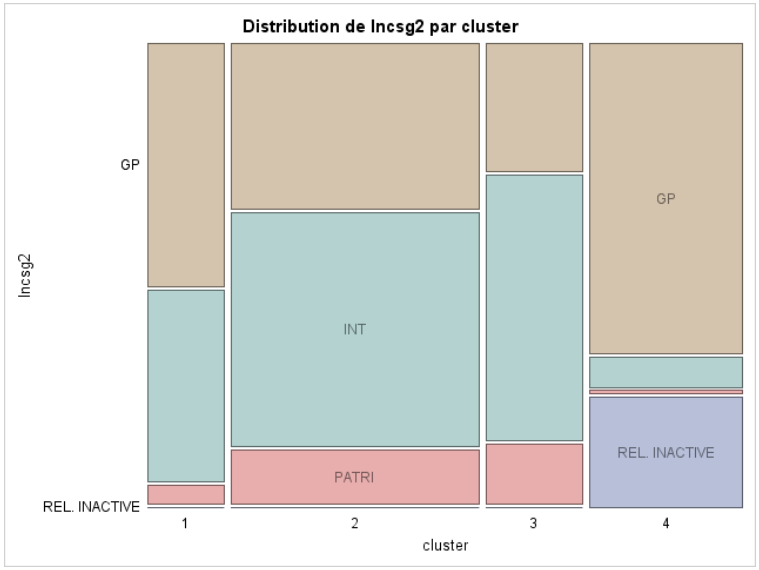


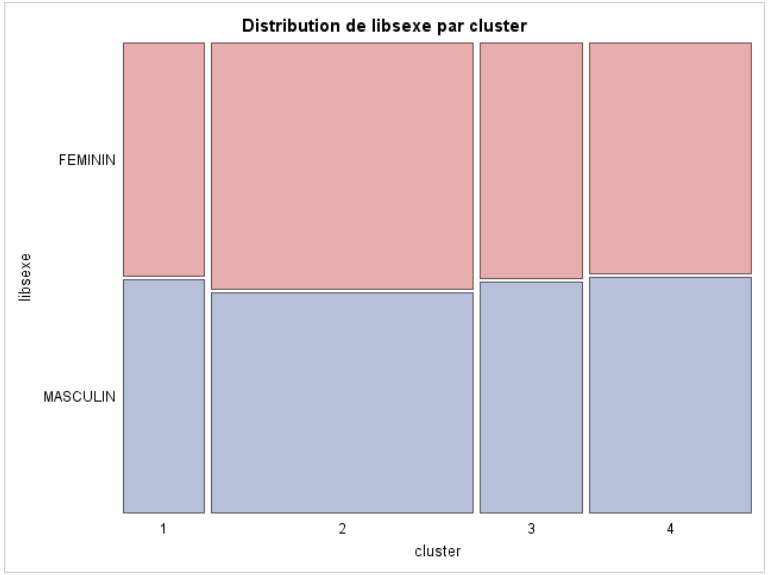


Résultats du clustering (12)

- Voir fichier « Box plots - 2eme approche (tout acp) » pour les box plots







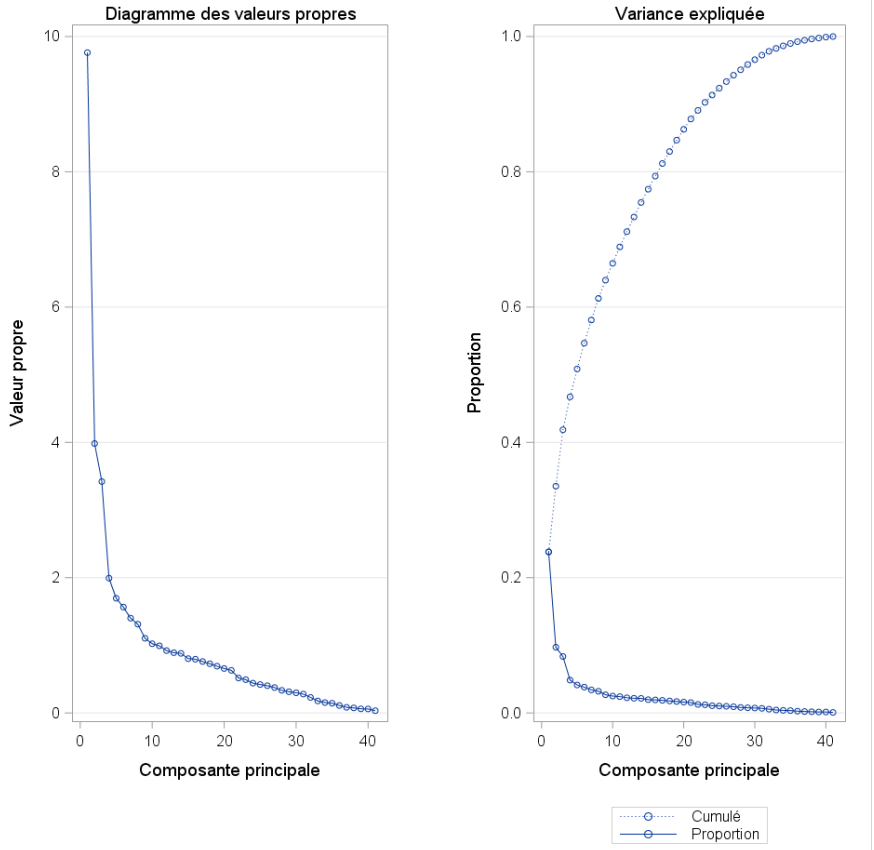
Résumé de la 2eme approche

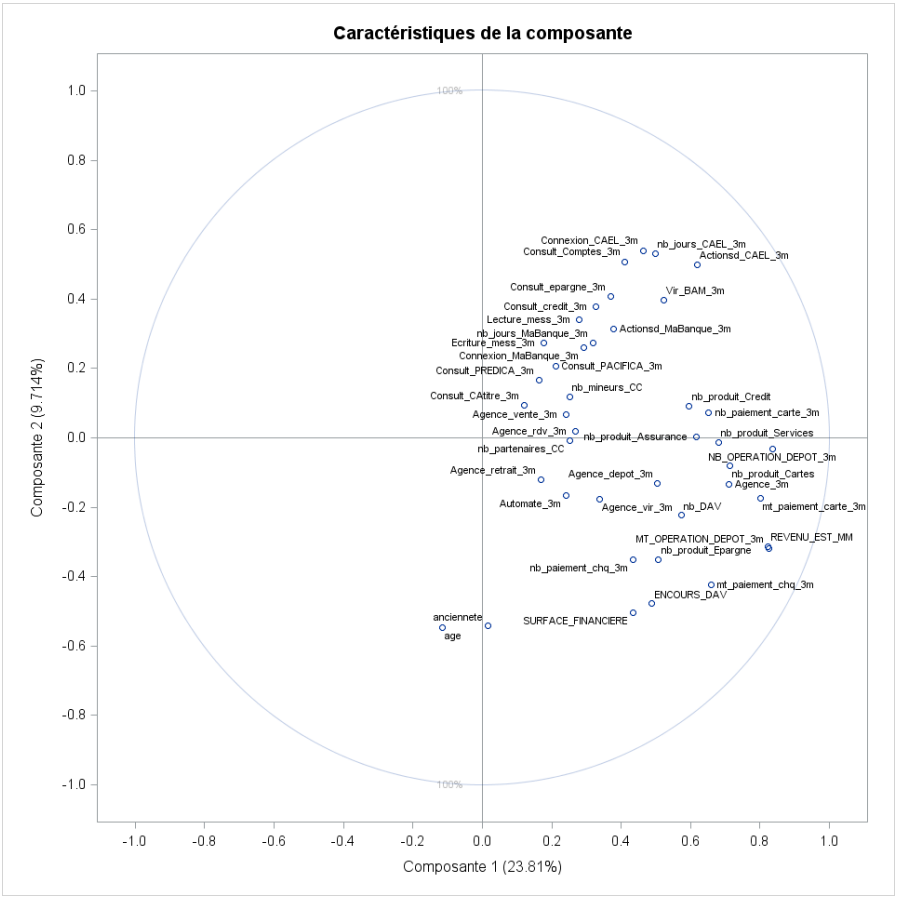
- Le cluster 1 se démarque par une forte utilisation de **ma banque**. Ces clients consomment également les autres canaux (agence et CAEL) et ils sont les plus jeunes. **Canaux privilégiés** : mail ou sms ou contact direct en agence
- Le cluster 2 a comme canal privilégié l'**agence**. Attention, ces clients ne se rendent pas plus en agence que les clients des clusters 1 et 3, mais ils se démarquent par le fait qu'ils consomment (presque) uniquement de l'agence (automates compris). Sans surprise, ce cluster regroupe le plus de retraités et l'âge moyen y est considérablement plus élevé que dans les autres clusters. **Canal privilégié** : courrier ou contact direct en agence.
- Le cluster 3 se démarque par une forte utilisation de **CAEL**. Ces clients se rendent aussi beaucoup en agence et utilisent ma banque (moins que le cluster 1). L'âge moyen est « intermédiaire » : plus vieux que le cluster 1 mais plus jeune que le 3**. Canal privilégié**: MS ou mail ou contact direct?
- Le cluster 4 est celui des clients inactifs.

3eme approche : clustering sur AFCM + ACP – Méthode

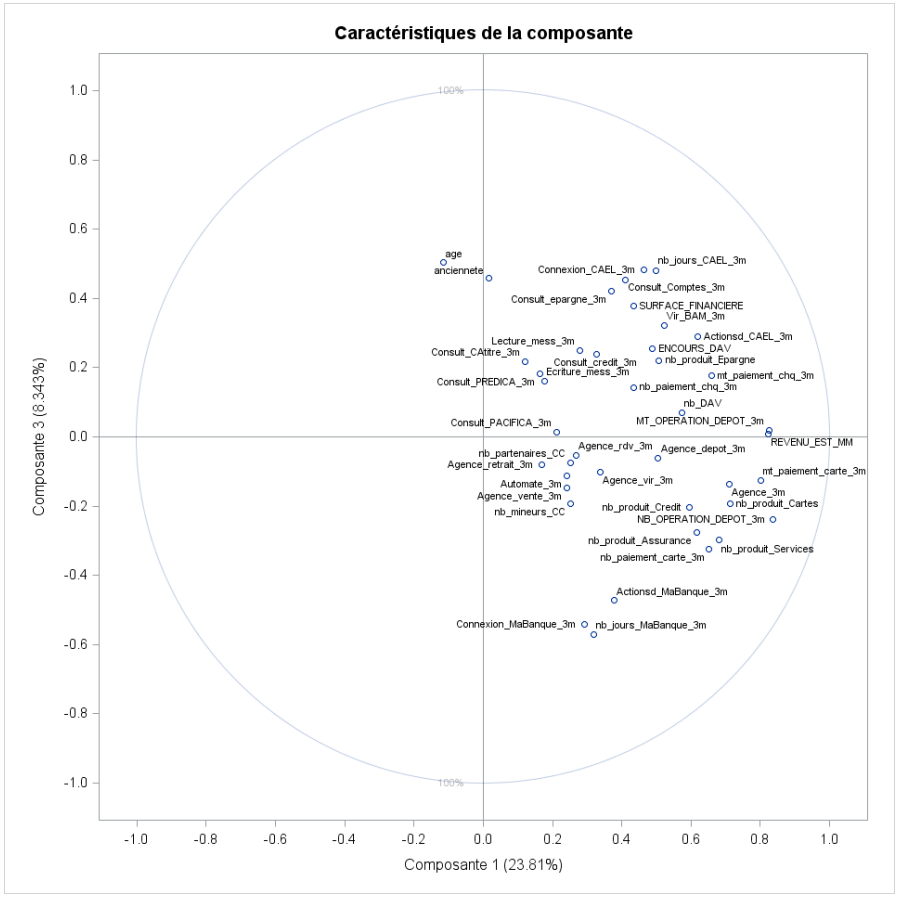
- Supprimer les outliers de la base d'étude contenant les variables quantitatives en étudiant la distribution de chaque variables
- Transformer les variables quantitatives (même méthode que pour la 2eme approche)
- Réaliser une ACP sur ces variables
- Réaliser une AFCM sur les variables qualitatives
- Stocker les résultats de l'ACP + AFCM dans une même base
- Lancer un k-means sur ces résultats avec 2000 classes (10% de 20000)
- Faire un clustering hiérarchique en initialisant les barycentres obtenus par le k-means
- Déterminer le nombre de classes optimales
- Lancer un k-means avec le nombre de classes optimales

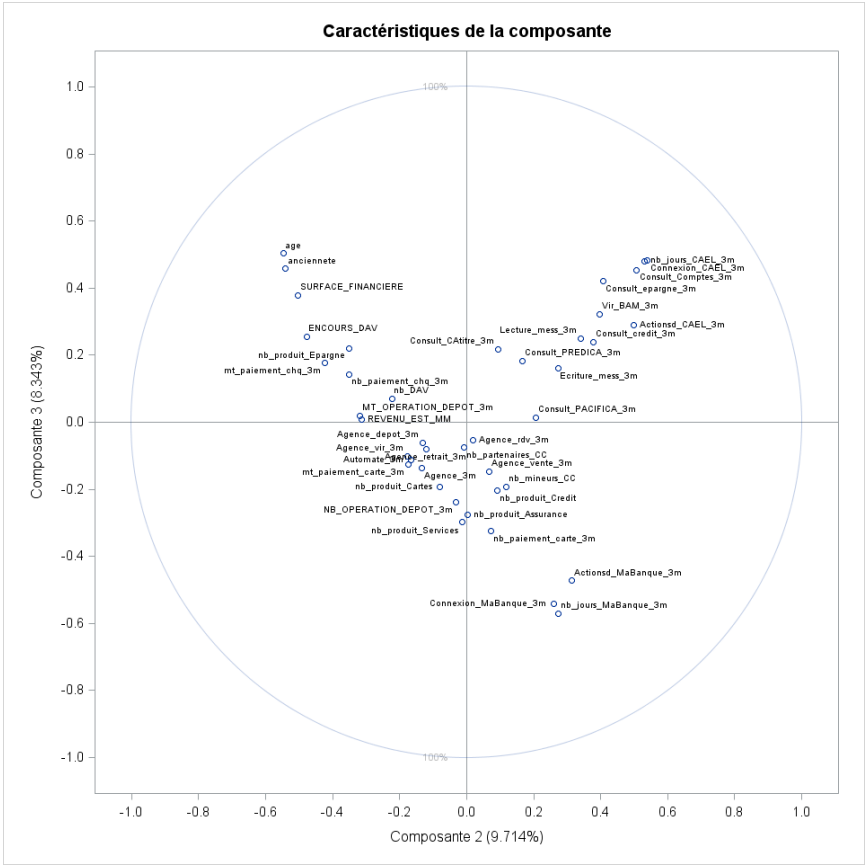
Résultats de l'ACP



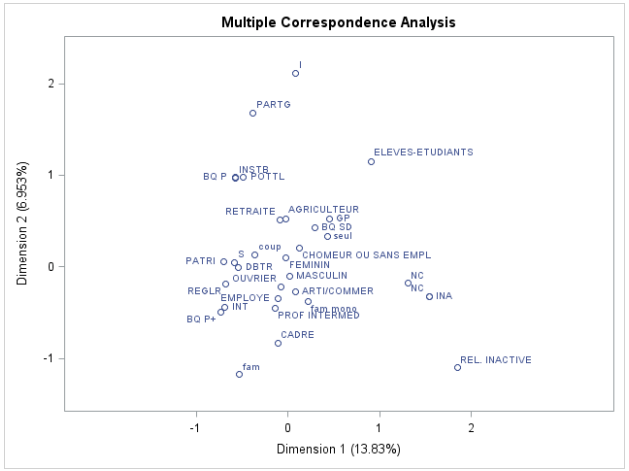


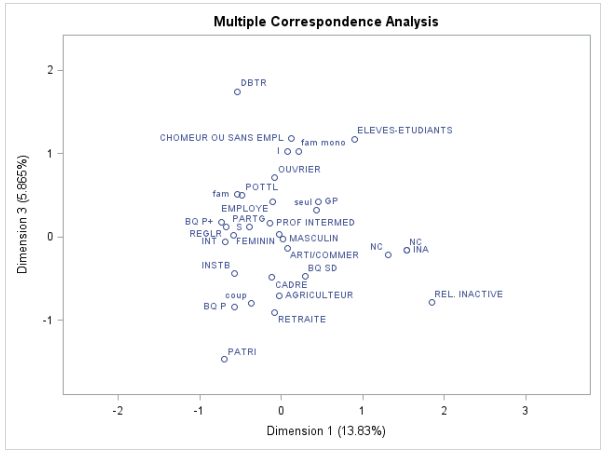
Résultats de l'ACP (2)

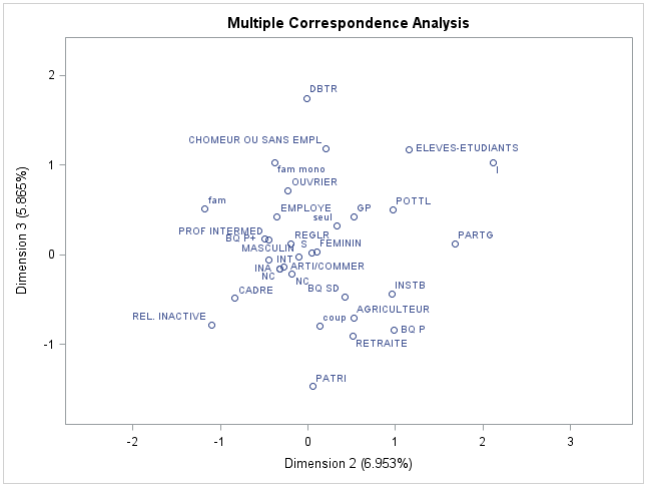


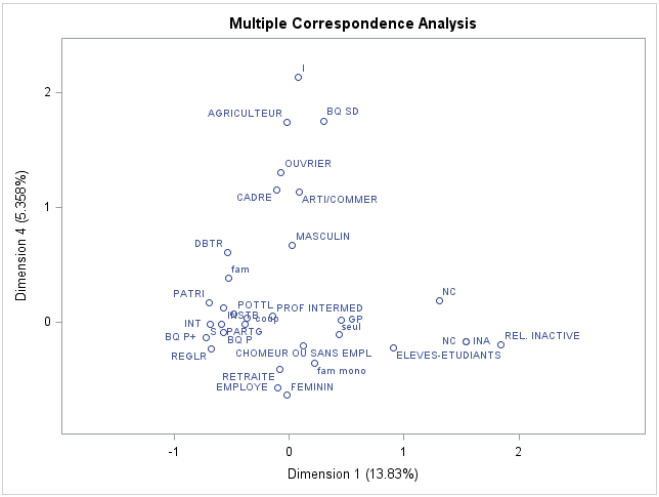


Résultats de l'AFCM

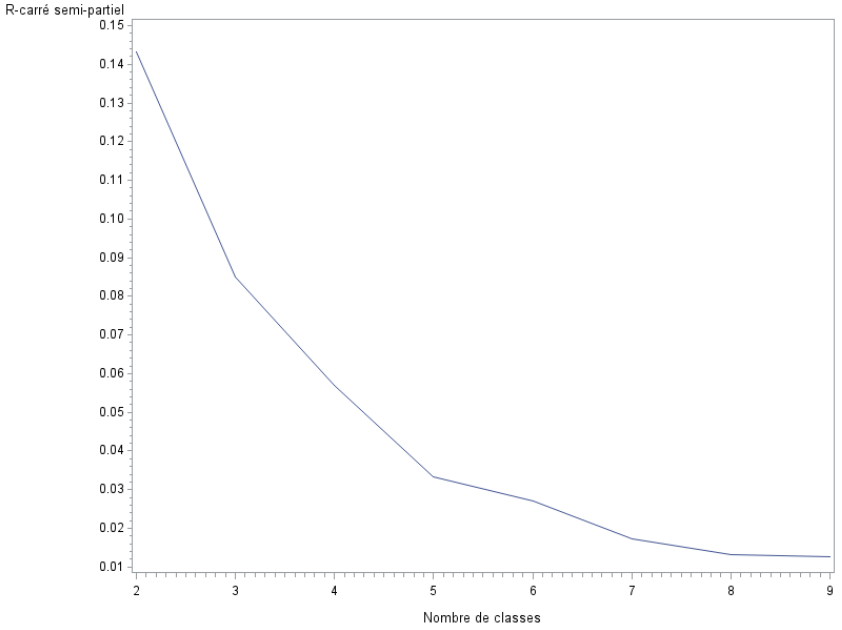




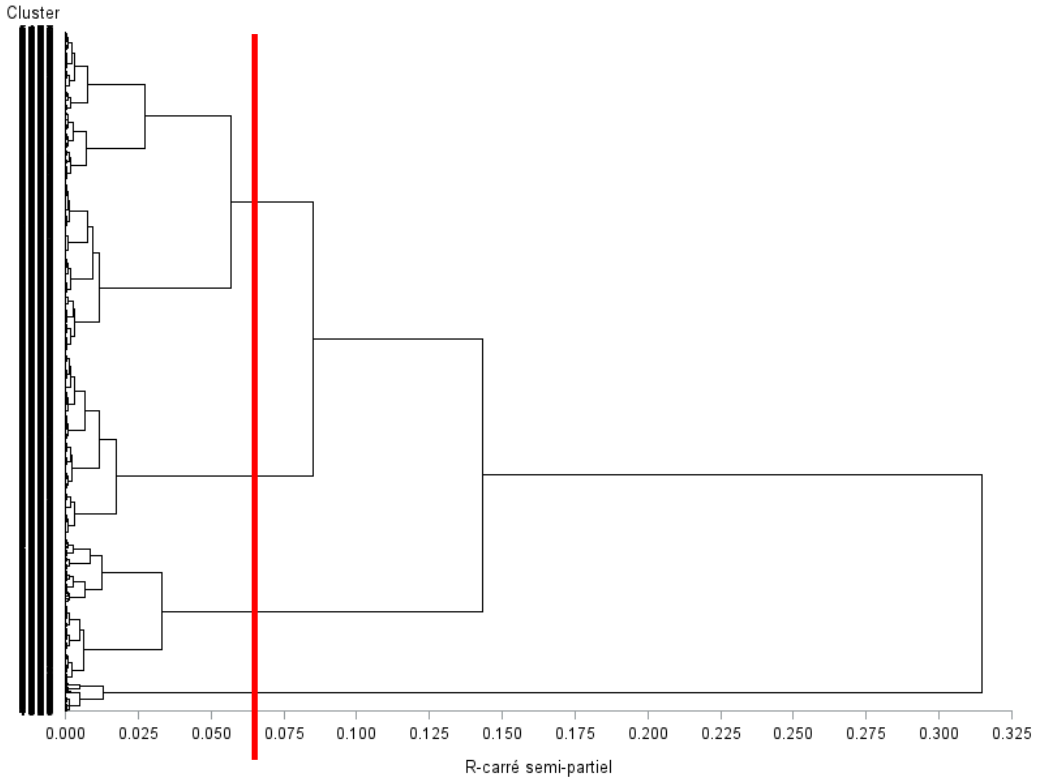




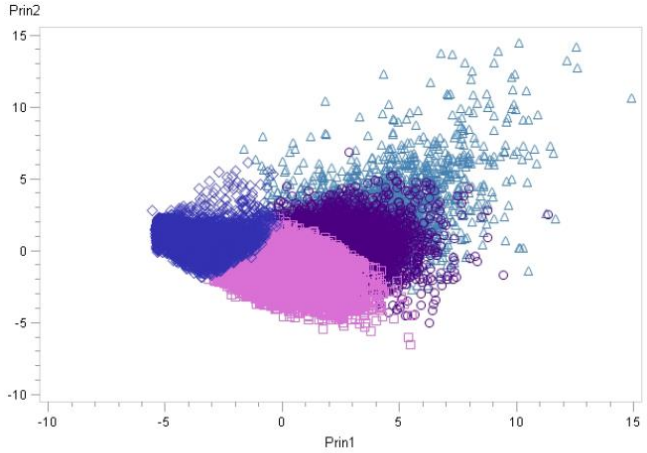
Résultats de la CAH

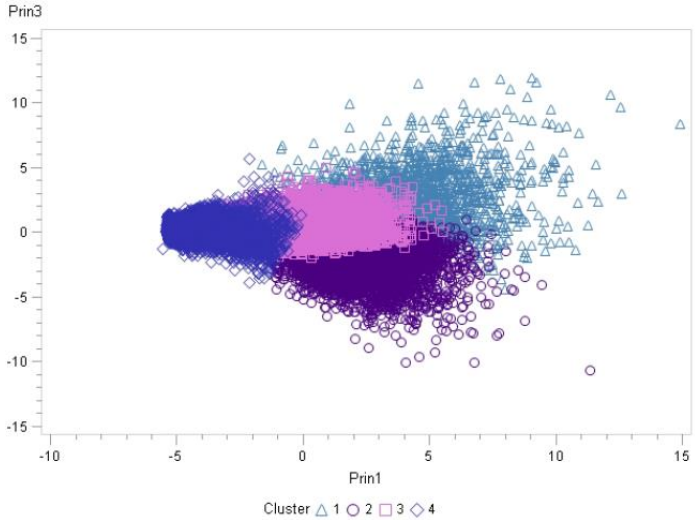


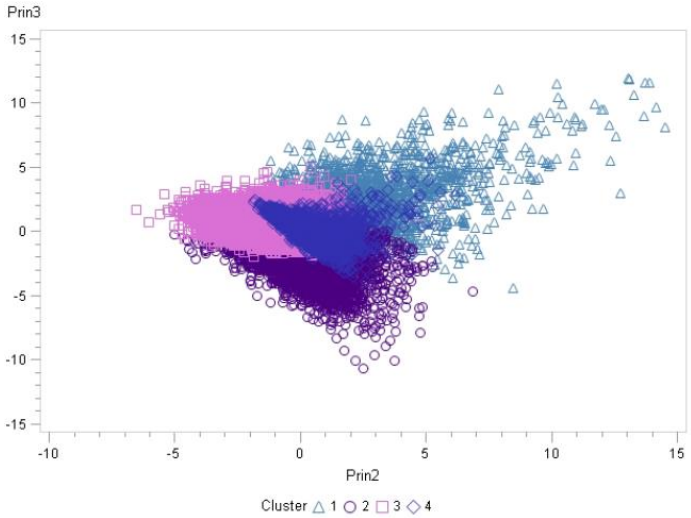
Résultats de la CAH (2)



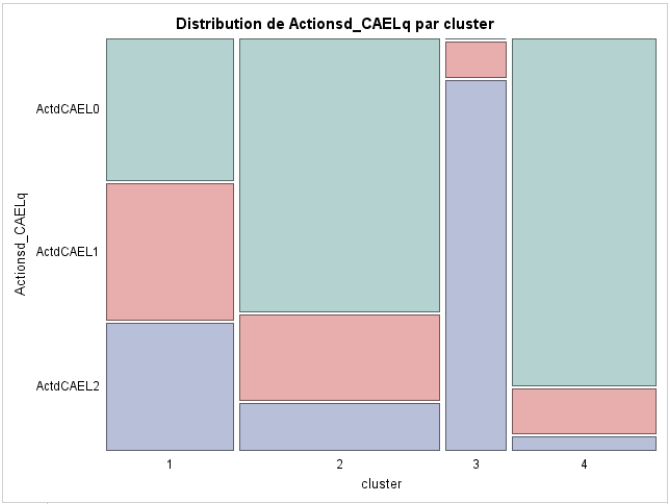
Représentation des clusters dans les axes de l'ACP

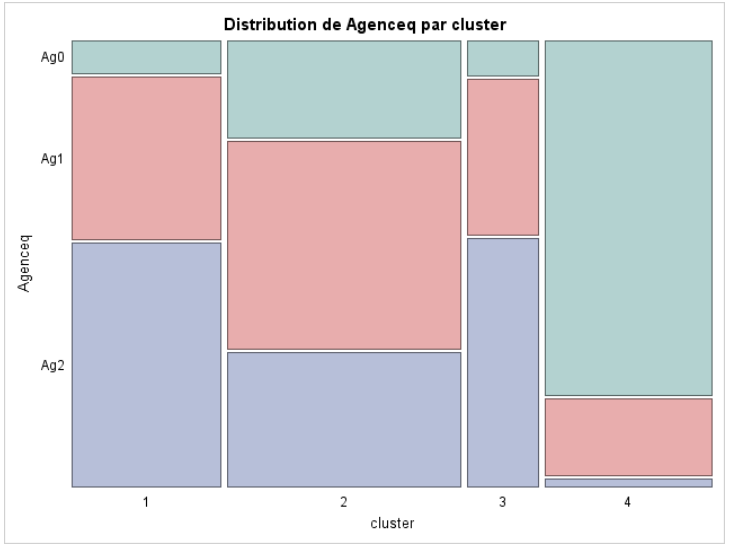


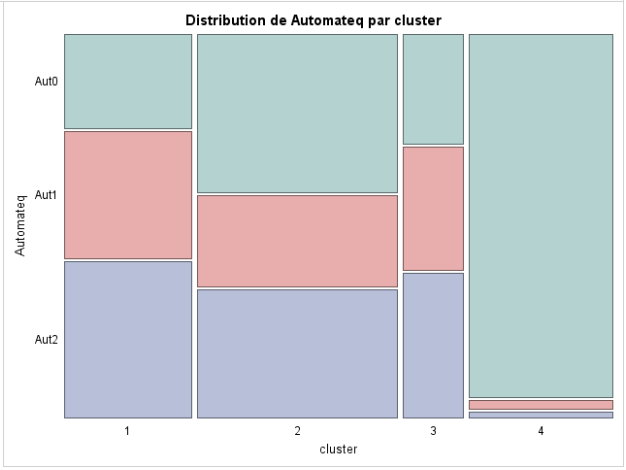


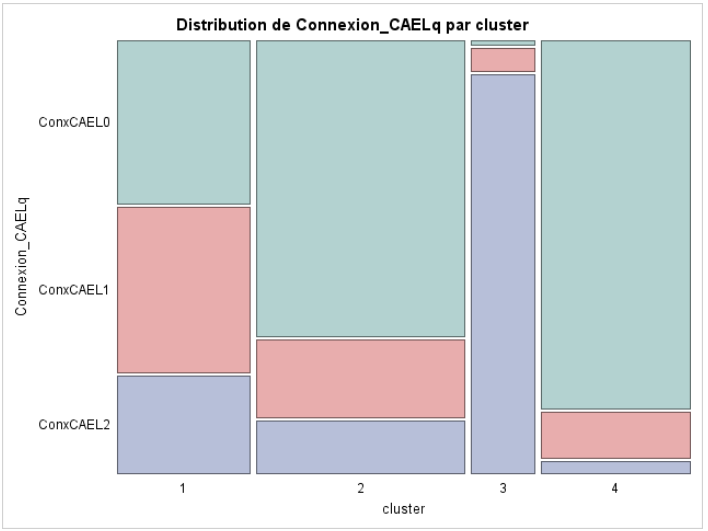


Résultats du clustering

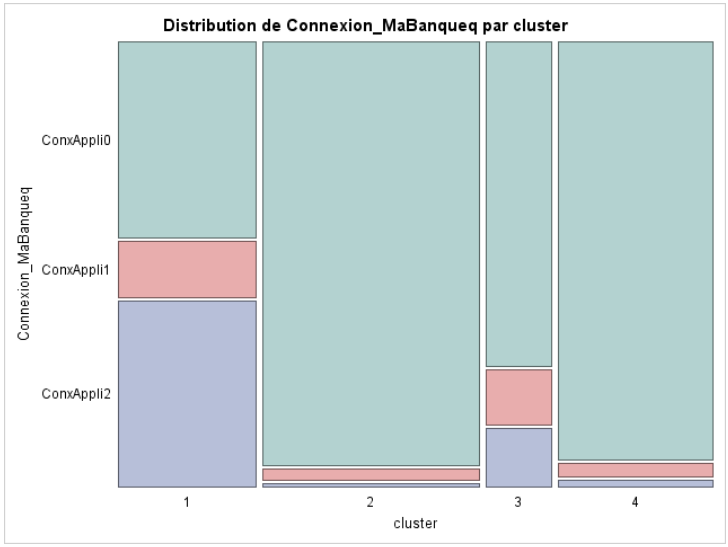


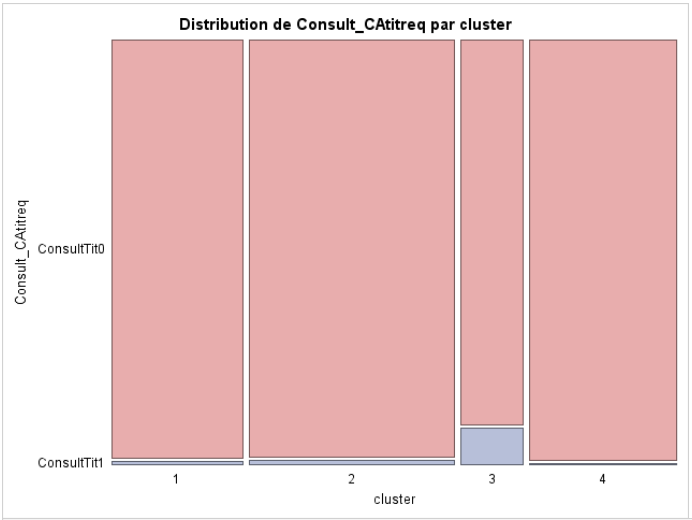


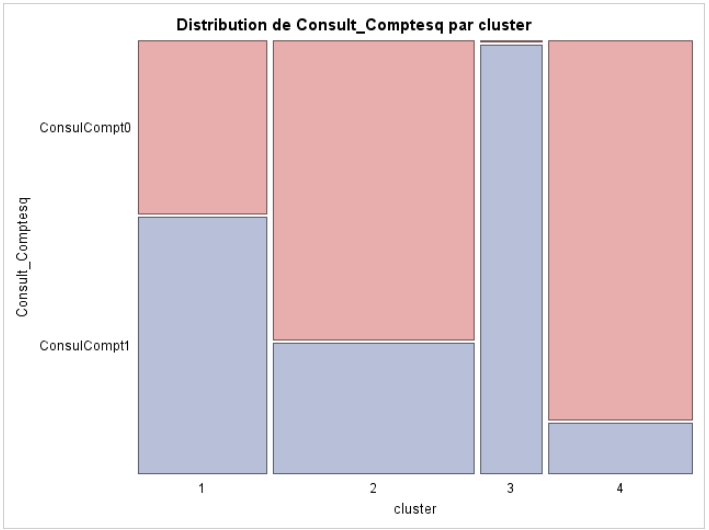


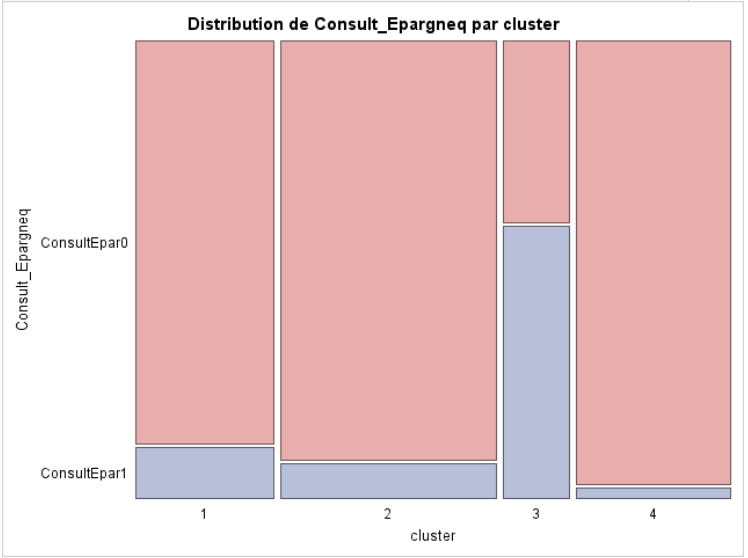


Résultats du clustering (2)

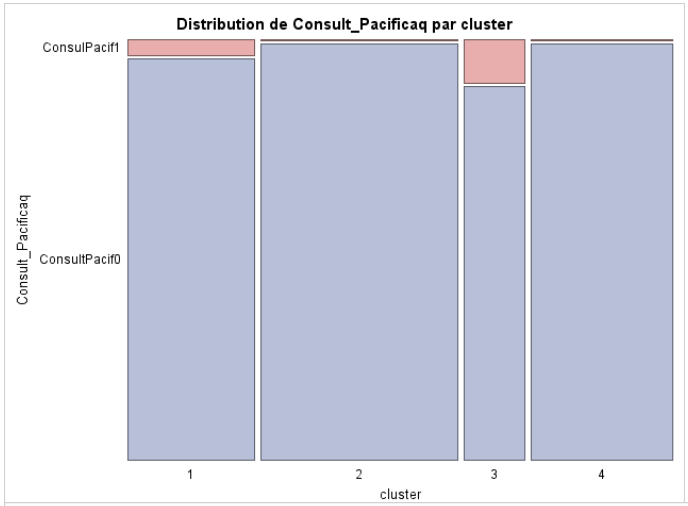


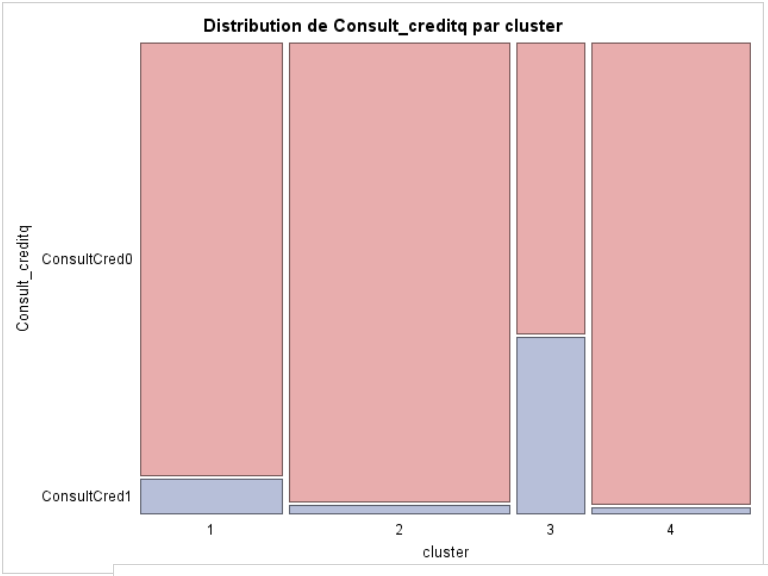


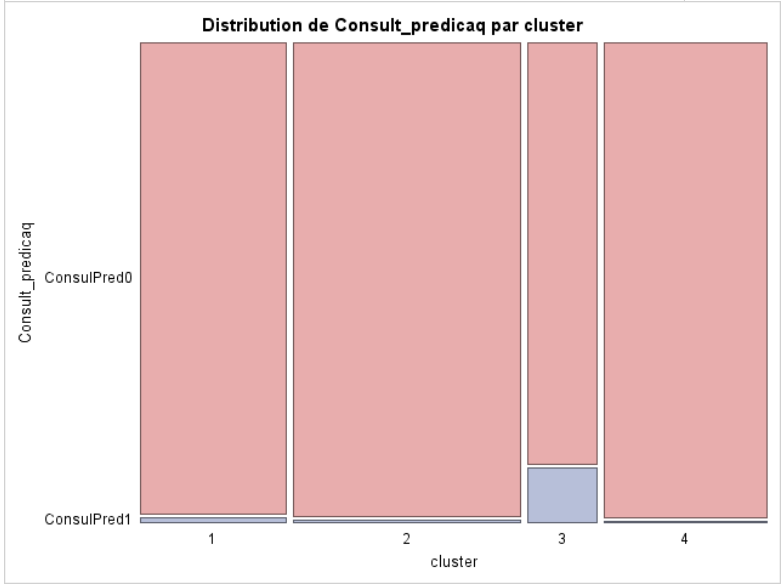


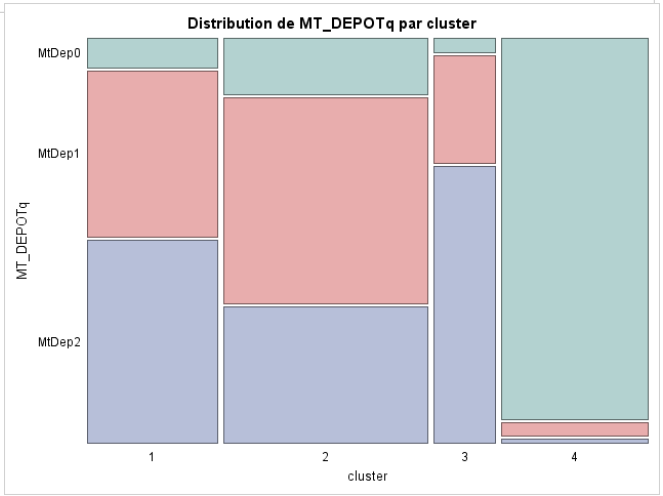


Résultats du clustering (3)

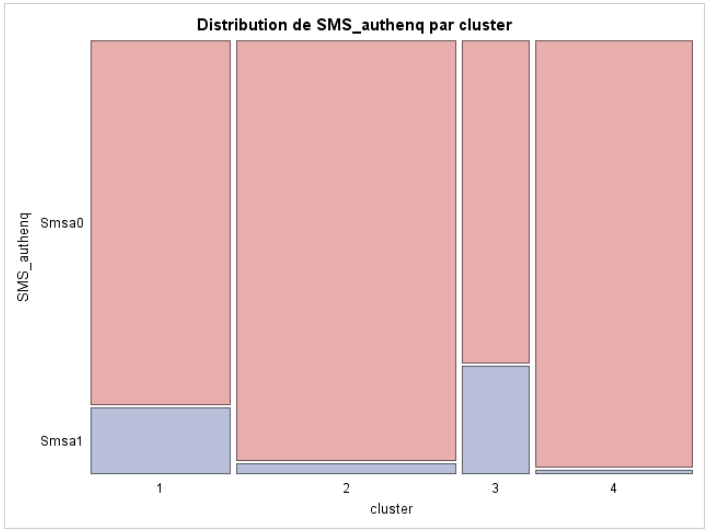


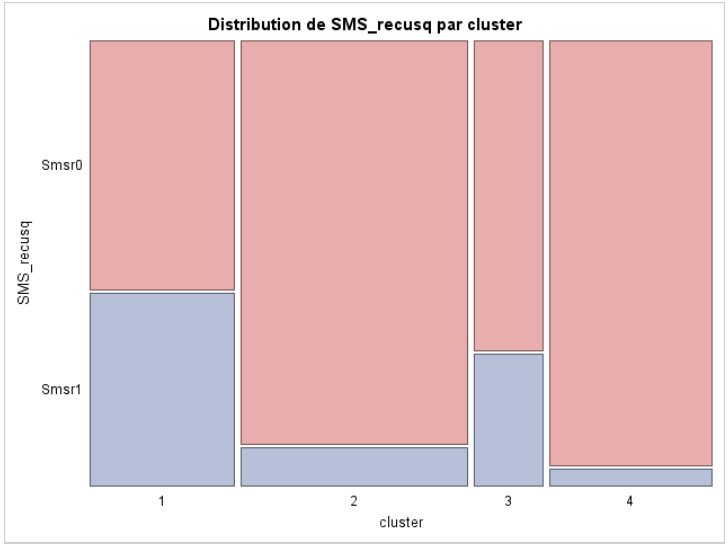


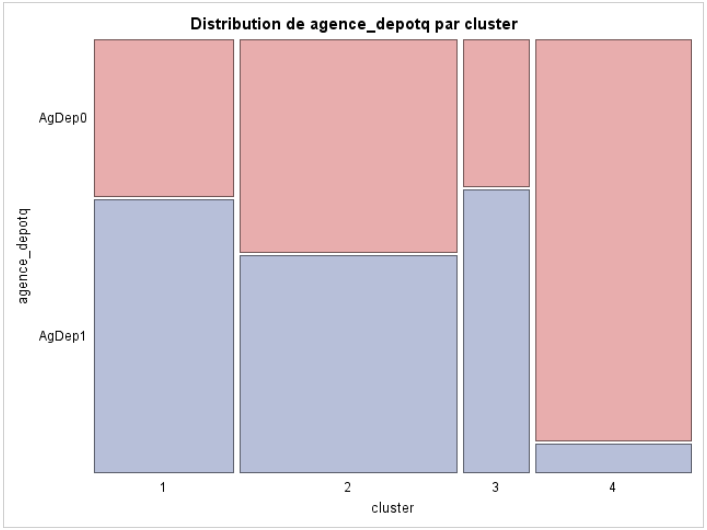


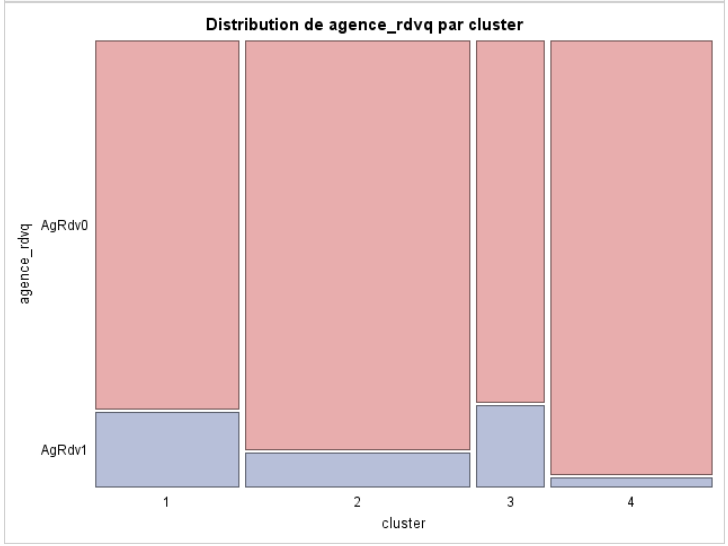


Résultats du clustering (4)

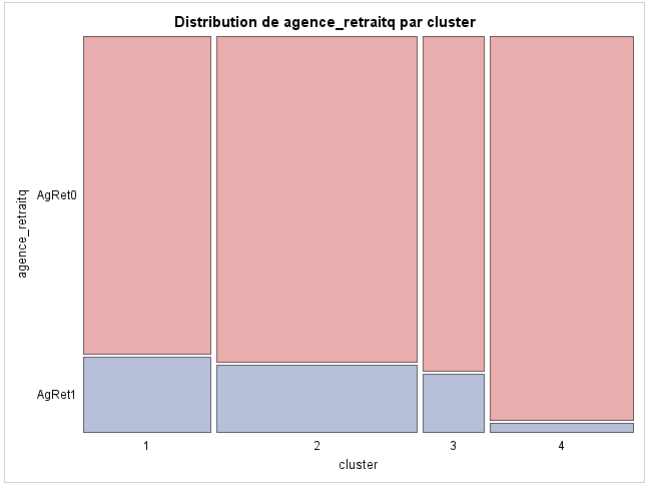


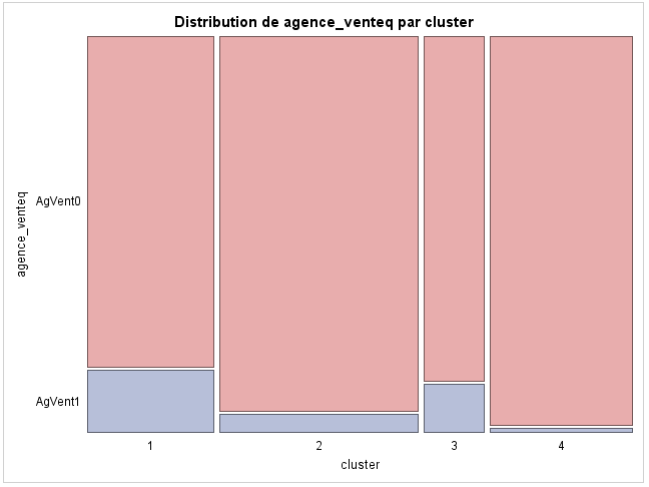


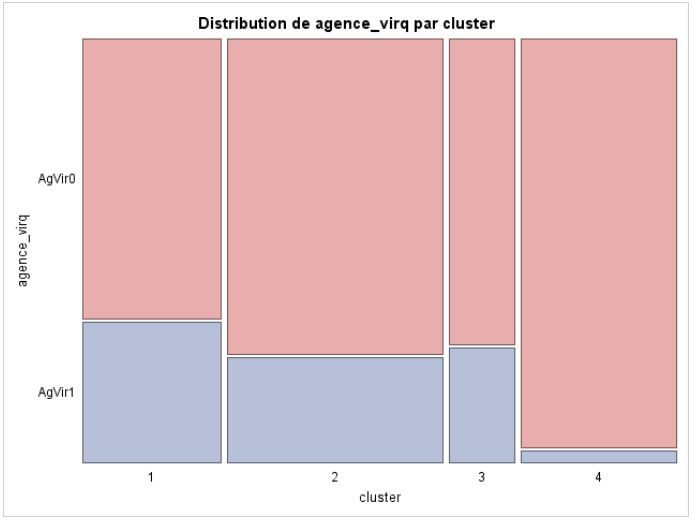


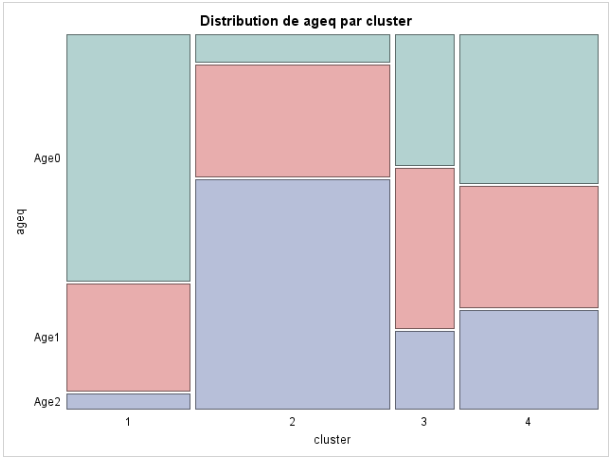


Résultats du clustering (5)

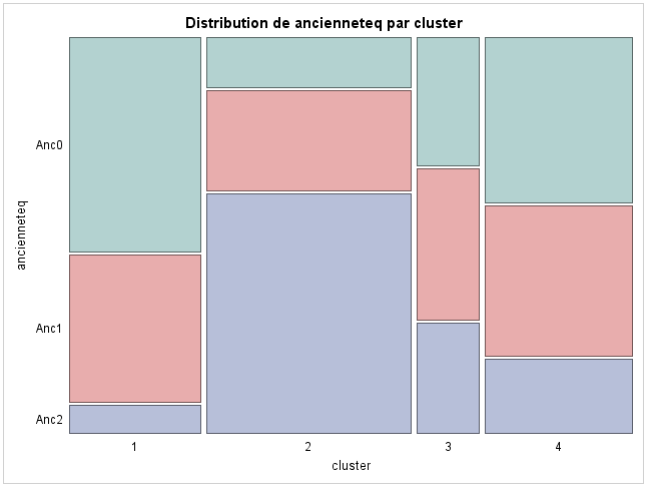


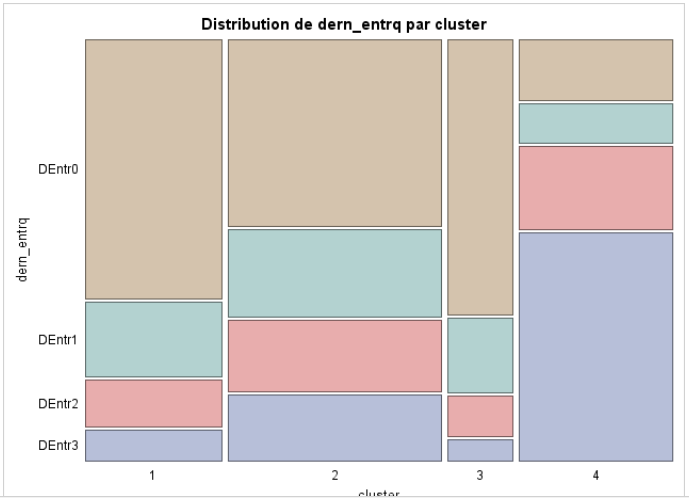


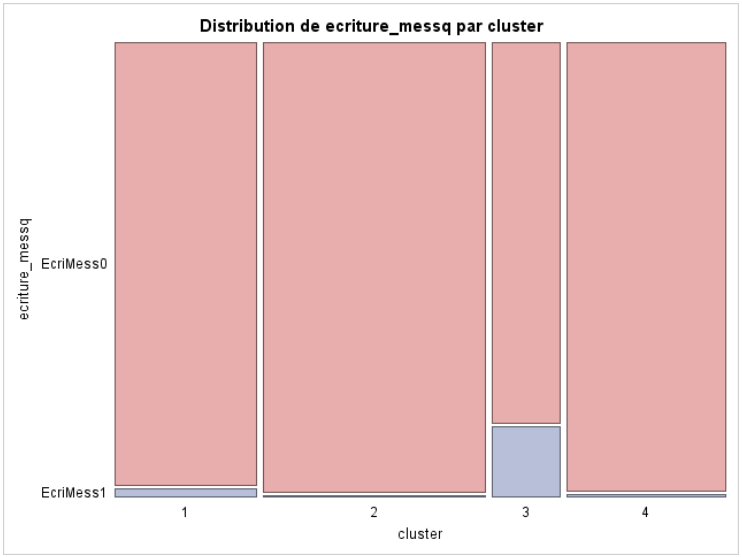


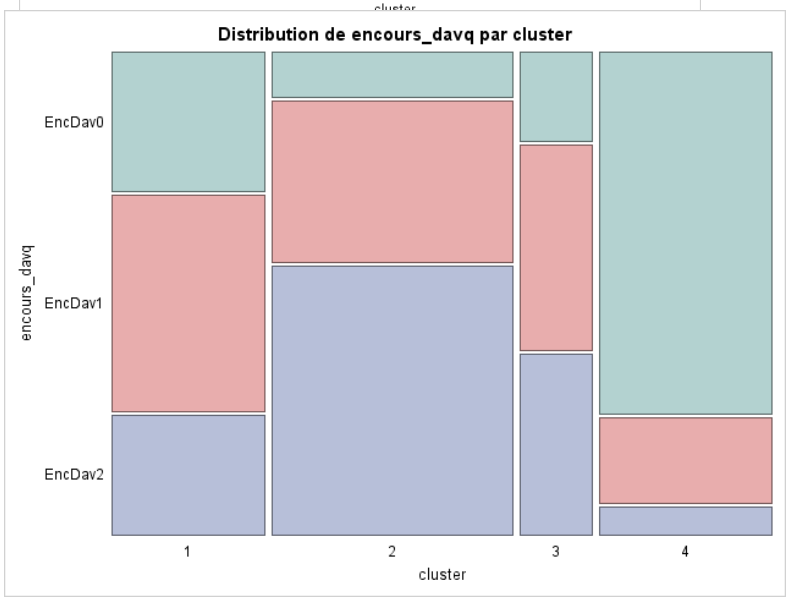


Résultats du clustering (6)

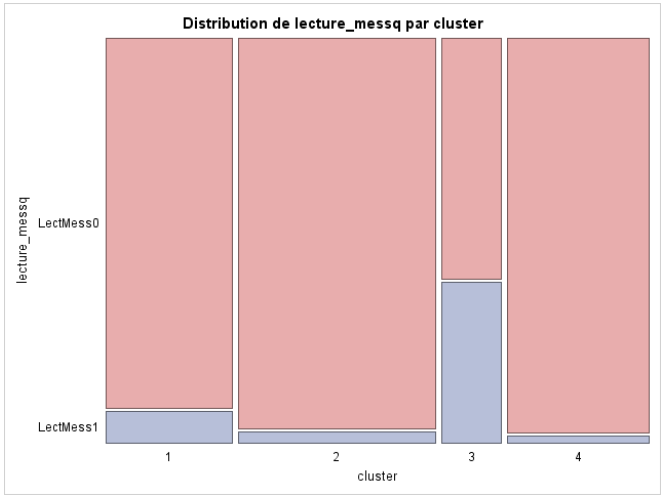


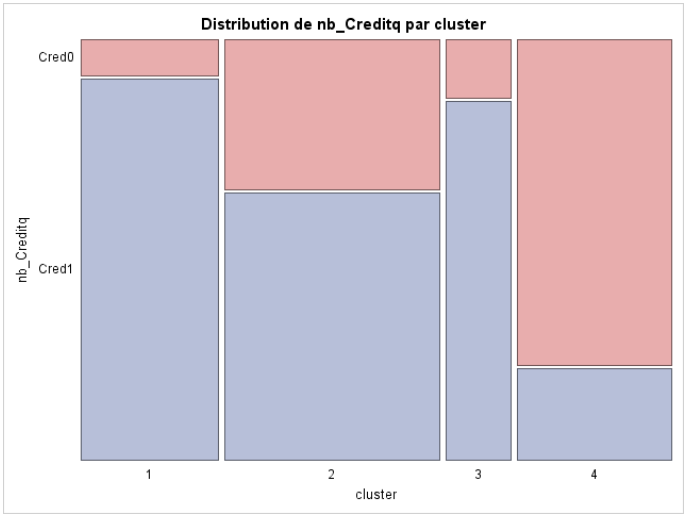


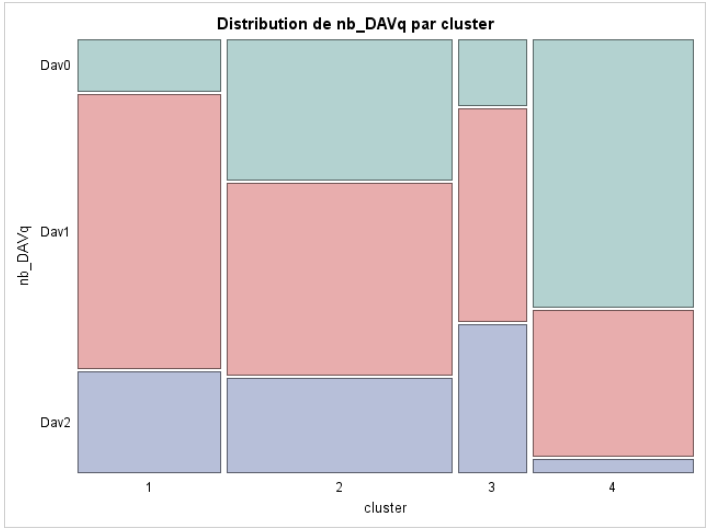


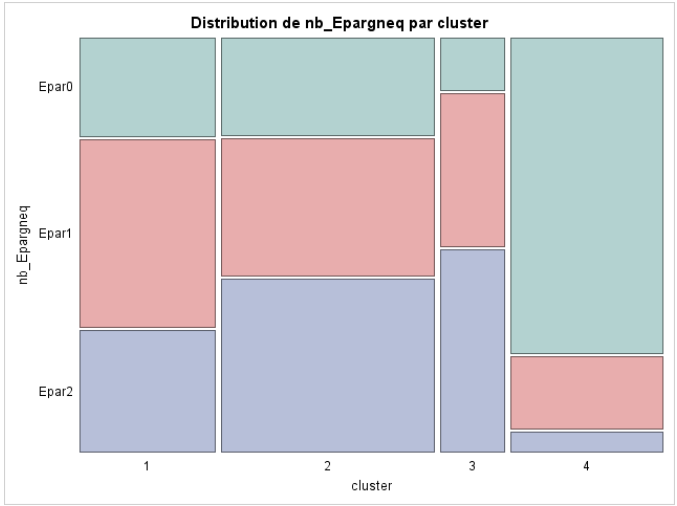


Résultats du clustering (7)

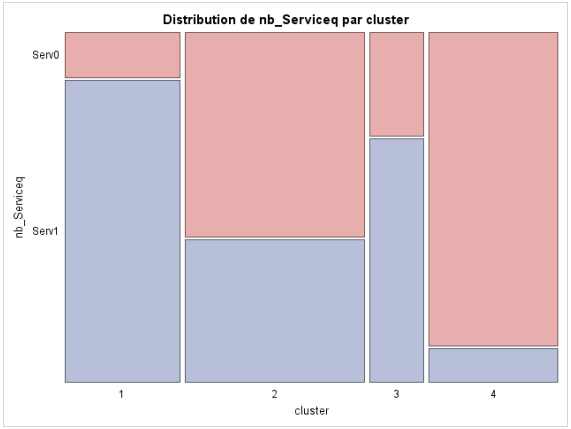


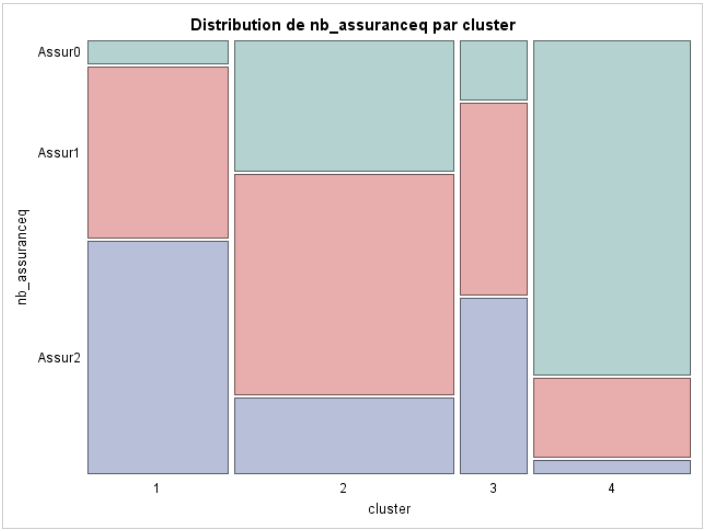


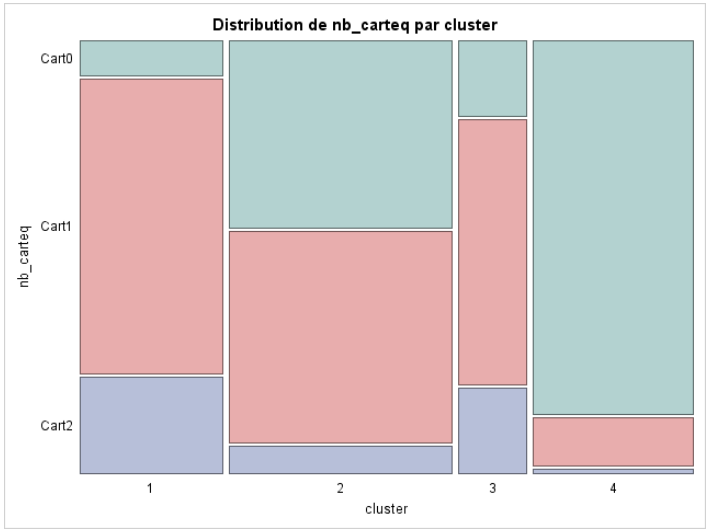


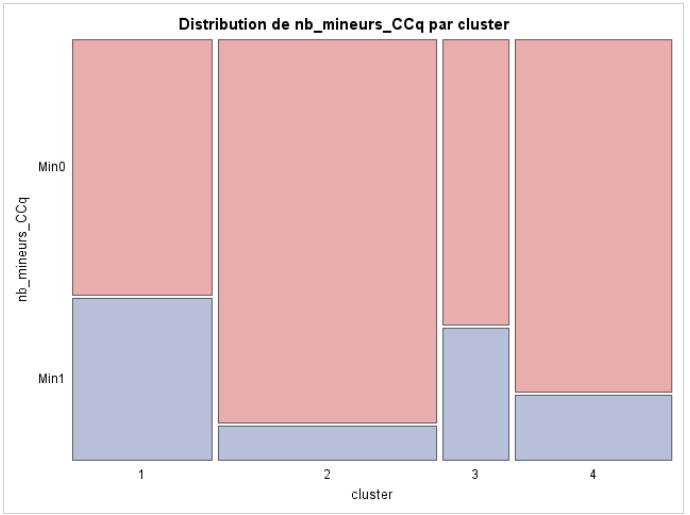


Résultats du clustering (8)

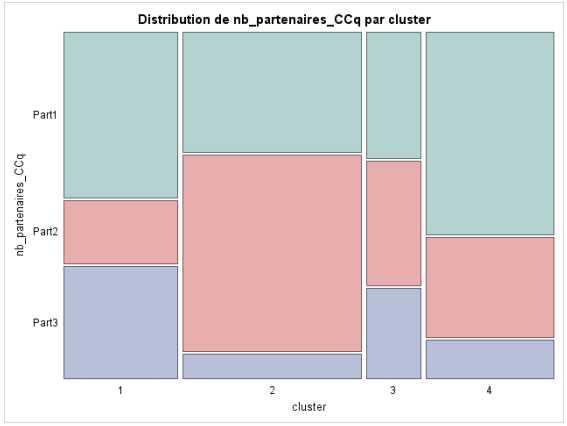


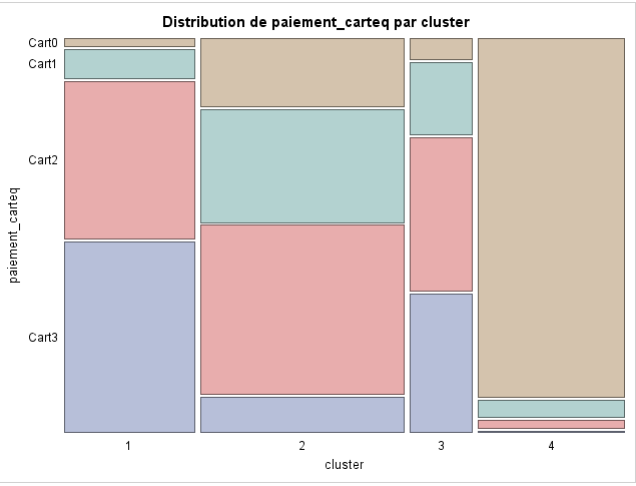


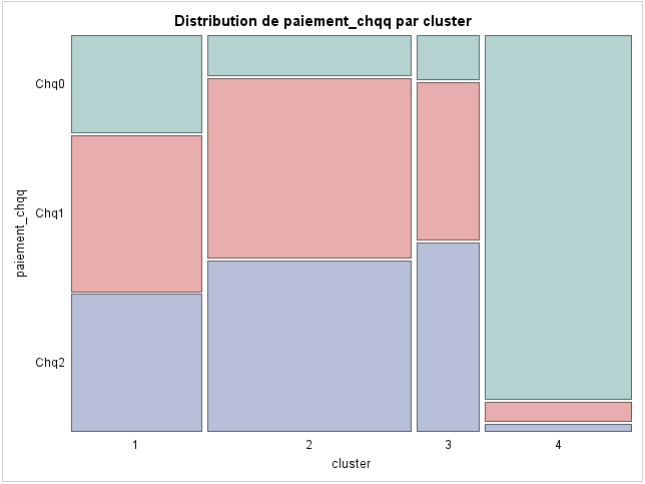


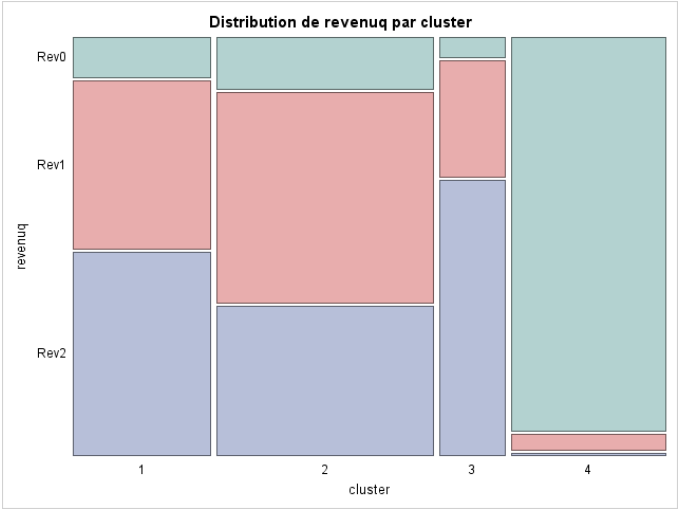


Résultats du clustering (9)

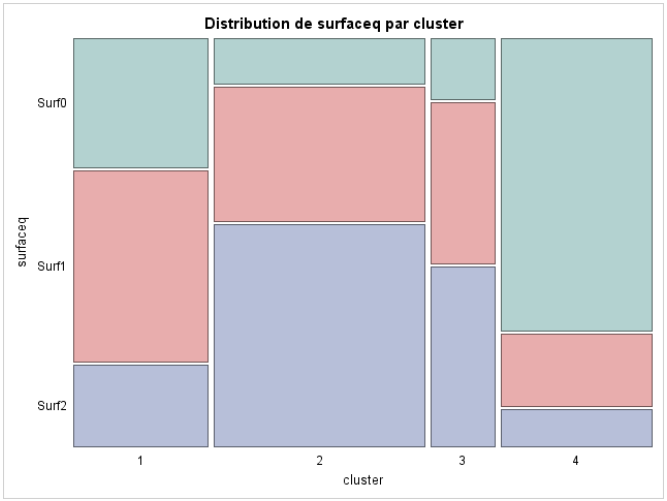


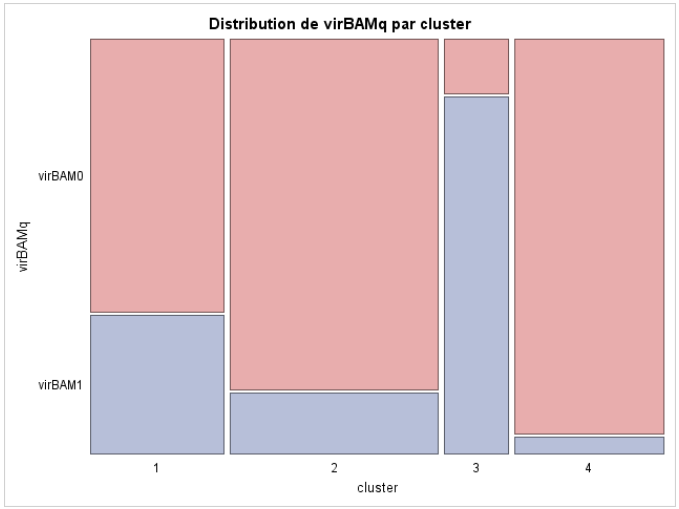


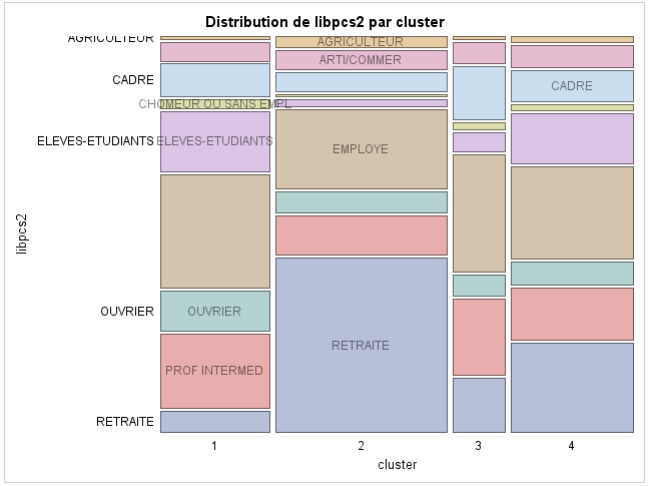


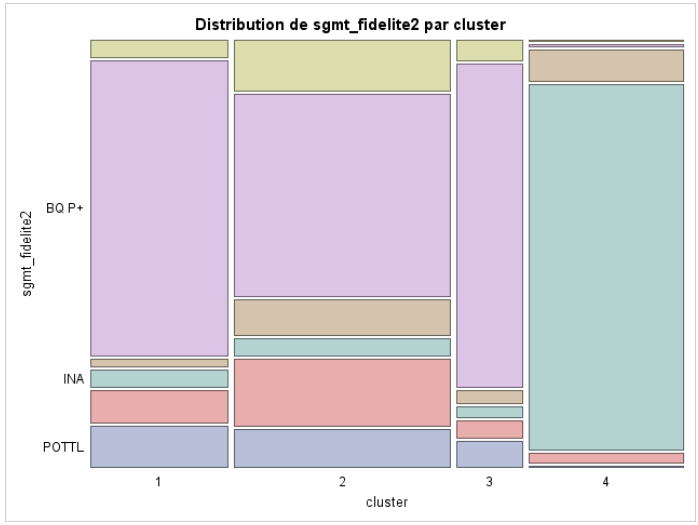


Résultats du clustering (10)

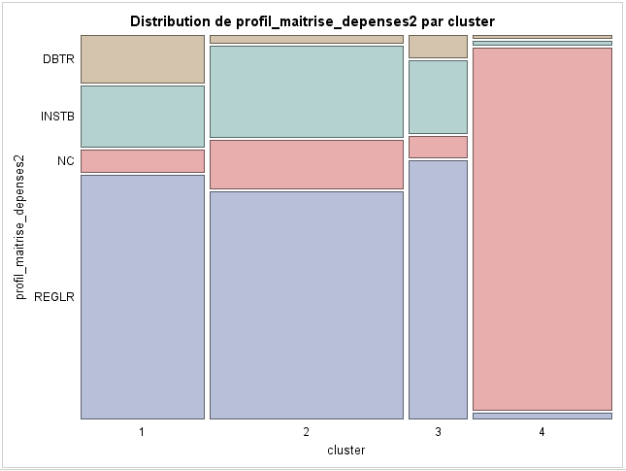


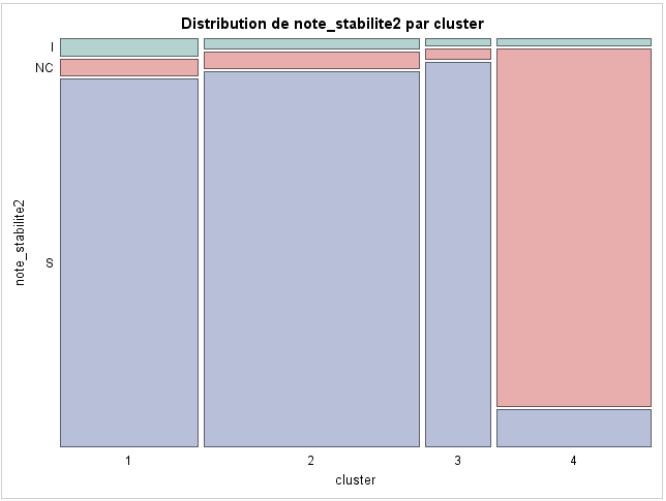


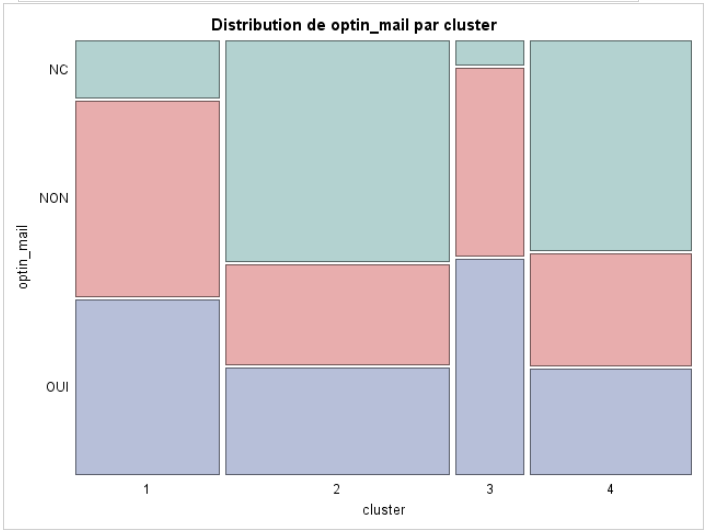


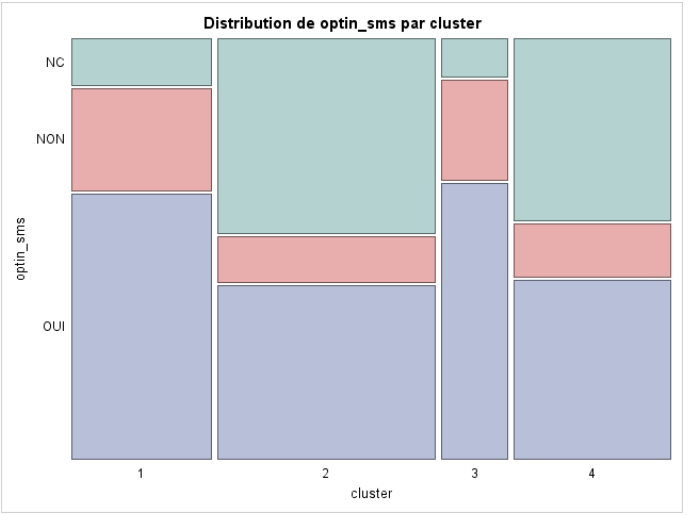


Résultats du clustering (11)

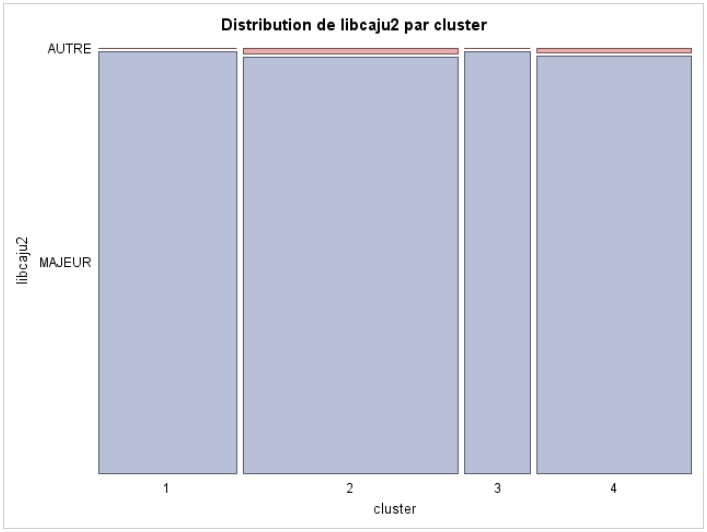


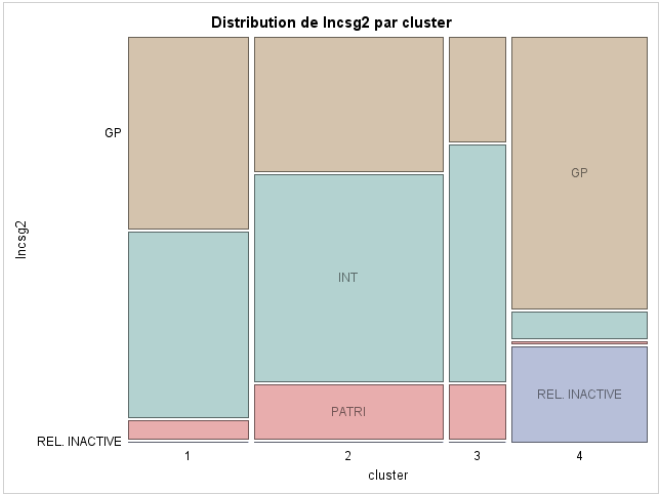


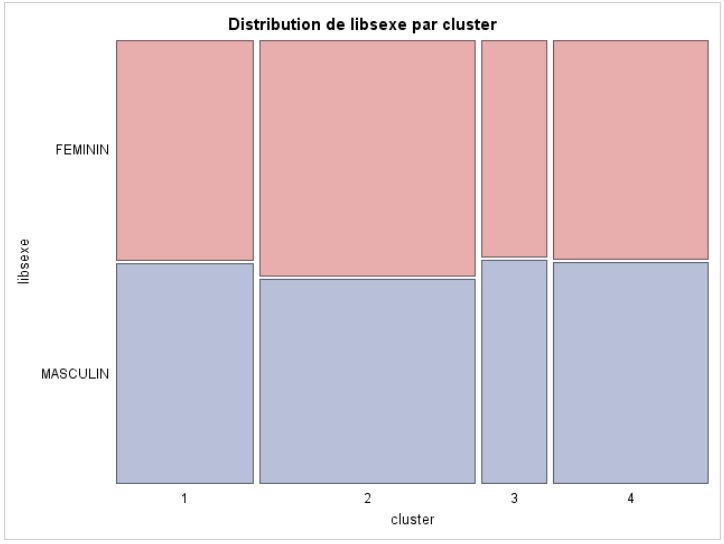


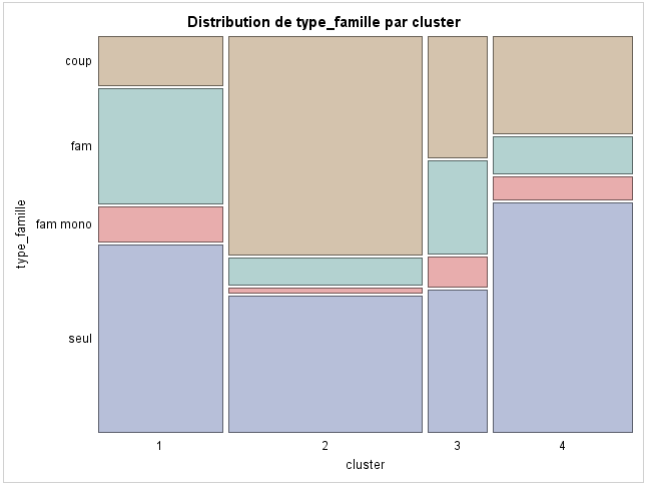


Résultats du clustering (12)









Résumé de la 3eme approche

- Le cluster 1 est le cluster des clients les **plus jeunes**. Ceux-ci utilisent **Ma Banque** plus que les autres clusters. Ils consomment également beaucoup d'agence et CAEL. **Canal privilégié** : mail ou sms ou contact direct en agence
- Le cluster 2 regroupe les clients les **plus âgés**. Ils consomment presque exclusivement de l'**agence** (utilisation de CAEL également mais beaucoup moins que les clusters 1 et 3). **Canal privilégié** : courrier ou contact direct en agence
- Le cluster 3 utilise davantage **CAEL** que les autres clusters. L'âge de ces clients est **intermédiaire**: ils sont plus âgés que les clients du cluster 1 mais plus jeunes que ceux du cluster 3. Ces clients sont aussi utilisateurs de ma banque et se rendent beaucoup en agence. **Canal privilégié** : MS ou mail ou contact direct en agence
- Le cluster 4 est celui des clients « **inactifs** ».

Quelle est la meilleure approche ?

- La **deuxième** (uniquement ACP) approche semble être plus tranchée sur les **variables de contacts** : on distingue bien les utilisateurs majoritaires sur ma banque et CAEL. On distingue également le cluster des retraités qui effectuent l'essentiel des opérations en agence ou sur des automates (situés en agence?)