

Typologie des consommateurs et Mesure de la loyauté/fidélité

Stéphanie LEDAUPHIN-MENARD¹, Sébastien LÊ²

¹ Équipe de Statistique - Laboratoire IRMAR Université Rennes 2-Haute Bretagne- Place du recteur H. Le Moal, CS 24307, 35043 RENNES - Email : stephanie.ledauphin@uhb.fr

² AGROCAMPUS OUEST Centre de Rennes Laboratoire de Mathématiques Appliquées 65, rue de St-Brieuc CS 84215 35042 RENNES Email : sebastien.le@agrocampus-ouest.fr

Résumé Face aux problèmes de pouvoir d'achat, les consommateurs changent leur manière d'acheter : certains se dirigent vers les marques premiers prix d'autres restent fidèles à leurs marques habituelles. Les industriels souhaitent définir des typologies de consommateurs concernant leur fidélité à une marque au cours du temps. Après une description du jeu de données concernant le suivi de consommation d'un panel de consommateurs sur deux marchés, nous avons défini une mesure de la loyauté/fidélité envers une marque sur la période considérée. Nous avons ainsi pu mettre en avant quatre groupes de consommateurs et visualiser la position des consommateurs vis-à-vis des différentes marques présentes sur le marché considéré. Dans un second temps, nous nous sommes attachés à caractériser la régularité de cette fidélité au cours du temps et avons obtenu une typologie des consommateurs. Enfin, nous avons cherché le lien éventuel entre les marques et les typologies considérées.

Mots Clés : Indicateur de loyauté, Indicateur de stabilité, Profils, Analyse Factorielle des Correspondances, Analyse Factorielle Multiple.

1 Description des données

Le jeu de données comprend 3745296 lignes correspondant aux informations recueillies auprès de 10068 consommateurs concernant 6 marques (les marques A,B,C (marché 1) et les marques D,E,F (marché 2)) pendant 62 semaines. Le fichier est composé de nombreuses valeurs nulles (plus de 88%), la répartition des valeurs non nulles est résumée dans la figure 1. La somme dépensée par les différents consommateurs varie énormément d'un individu à l'autre (cf figure 2). Certains n'achètent qu'une seule marque d'autres plusieurs simultanément. Les dépenses sont plus ou moins importantes, ponctuelles ou régulières. . .

1.1 Le tableau par marché : notations

Nous avons étudié dans un premier temps le premier marché (produits A , B et C), puis le second marché (produits D , E et F). Nous avons créé un tableau à 3 entrées M^1 de dimension I (nombre d'individus, ici 10068)* J (nombre de produits, ici 3)* T (nombre de dates, ici 62). La valeur M_{ijt}^1 correspond à la somme dépensée par l'individu i pour le produit j au cours de la semaine t sur le premier marché.

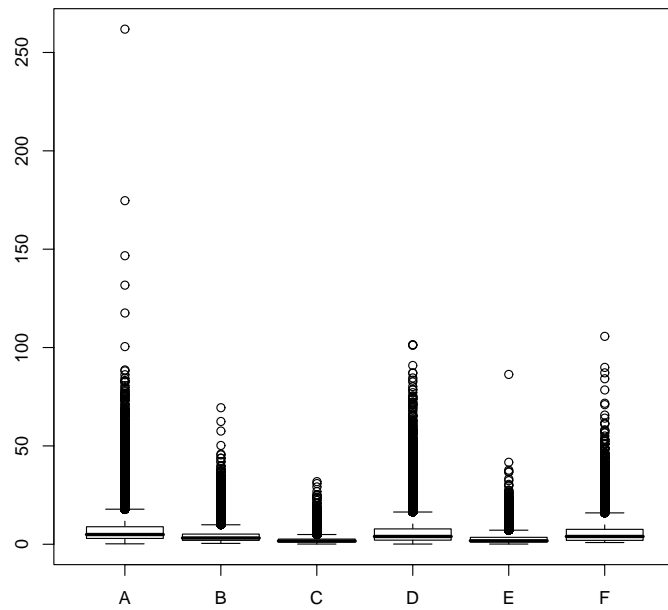


FIG. 1 – Box-plot des dépenses hebdomadaires non nulles

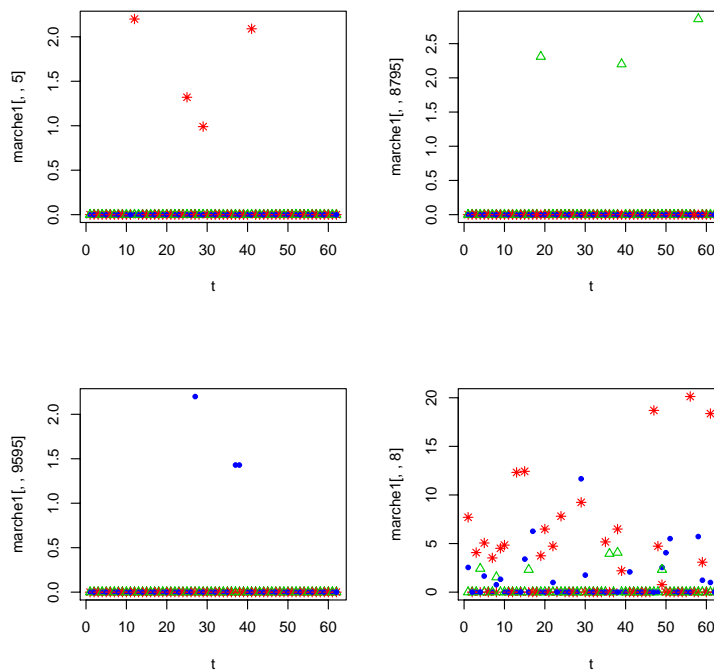


FIG. 2 – Valeurs dépensées pendant 62 semaines sur le premier marché pour quatre individus (5;8795;9595 et 8); (* produit A; Δ produit B; \bullet produit C)

1.2 Les données de profils par marché

Nous nous intéressons à la loyauté d'un individu envers une marque et nous n'allons plus considérer les sommes dépensées pour chaque marque mais la proportion dépensée sur le marché considéré. En effet, l'intérêt n'est pas dans les différences de budget parmi les consommateurs mais bien dans le fait qu'un consommateur dépense plus pour une marque que pour une autre. Deux individus achetant toujours une même marque sont des consommateurs très fidèles quelque soit leur budget.

Pour un marché (par exemple les produits A , B et C du marché 1), nous allons ainsi calculer ce qui pourrait être appelé 'profil' de chaque individu (la somme pour toutes les dates et pour tous les produits d'un individu vaut 1) :

$$\forall i \in [1, I] \quad j \in [1, J] \quad t \in [1, T] \quad P_{ijt}^1 = \frac{M_{ijt}^1}{\sum_{j=1}^J \sum_{t=1}^T M_{ijt}^1}$$

1.3 Les données de profils (agrégation suivant le temps)

Pour simplifier les calculs et les visualisations de résultats, il est possible d'agréger suivant une autre unité de temps que les 62 semaines, par exemple par mois ou par trimestre. Dans l'exemple, nous obtenons alors deux tableaux par marché, de dimensions $10068*3*15$ ou $10068*3*5$, selon l'unité choisie. Cela permet aussi d'éliminer des lignes nulles (semaines sans achat sur le marché) qui peuvent être nombreuses selon le marché considéré.

1.4 Les marges

Nous allons faire la somme pour chaque individu et chaque produit des valeurs P_{ijt}^1 au cours du temps et ainsi n'obtenir qu'un tableau *margét* de dimension $I * J$ (ici $10068*3$). L'étude de cette marge nous permet de déterminer pour chaque individu la répartition de ses achats dans les différentes marques sur toute la période considérée. Pour les quatre individus (5, 8, 8795 et 9595) vus précédemment (cf figure 2), nous obtenons par exemple :

- individu 5 : A : 1, B : 0, C : 0
- individu 8795 : A : 0, B : 1, C : 0
- individu 9595 : A : 0, B : 0, C : 1
- individu 8 : A : 0.71, B : 0.07, C : 0.22.

Il est aussi possible de faire la somme pour chaque date et chaque produit des valeurs P_{ijt}^1 des différents individus et obtenir le tableau *margei* de dimension $J * T$ (ici $3*62$, ou $3*15$ ou $3*5$ selon l'unité de temps choisie). L'étude cette marge permet de déterminer pour chaque produit l'évolution de la répartition des achats dans les différentes marques au cours du temps.

L'étude du tableau *margej*, somme pour chaque individus et pour chaque date des valeurs P_{ijt}^1 , permet de déterminer pour chaque individu l'évolution de la répartition de ses achats au cours du temps.

2 Méthodologie

2.1 Mesure de la loyauté

Le tableau *marget* indique, pour chaque individu, la part du budget attribuée à chacune des marques. La valeur maximale peut être considérée comme un indicateur de loyauté. Ainsi les individus 5, 8795 et 9595 sont très fidèles ; Ils n'ont acheté qu'une seule marque pendant les 62 semaines. L'individu 8, quant à lui, a acheté majoritairement la marque *A* (71% du budget) mais sa loyauté est moindre. Il est, en fait, plus judicieux de considérer l'écart entre la valeur maximale et la suivante. Posons :

$$IL_i = marget_{ij_1} - marget_{ij_2}$$

$$\text{avec } j_1 = \{j \in [1, J], marget_{ij} = \sup_{j \in [1, J]} marget_{ij}\}$$

$$\text{et } j_2 = \{j \in [1, J], marget_{ij} = \sup_{j \in [1, J] \setminus j_1} marget_{ij}\}$$

L'indicateur *IL* est compris entre 0 et 1. Si une seule marque est achetée (profil du type $\{1, 0, \dots\}$), alors l'indicateur de loyauté vaut 1. A l'opposé, si toutes les marques sont achetées dans les mêmes proportions (profil du type $\{\frac{1}{J}, \frac{1}{J}, \dots\}$) ou si les deux marques principales sont achetées dans les mêmes proportions, alors *IL* vaut 0. L'individu n'est en effet pas loyal envers une marque en particulier.

Nous avons à partir de cet indicateur *IL* défini quatre groupes de consommateurs :

- Les consommateurs très fidèles ($IL = 1$) n'ont acheté qu'une seule marque pendant les 62 semaines.
- Les fidèles ($0.5 \leq IL < 1$) ont tenté, à un moment ou à un autre, une autre marque (promotion, produit absent, bon de réduction), mais ils gardent une grande majorité de leur achat pour une même marque.
- Les indécis ($0.25 \leq IL < 0.5$) ne sont pas aisément caractérisables. Ils ont globalement une préférence pour une marque (comme l'individu 8 précédemment décrit).
- Les très indécis ($IL < 0.25$) achètent aussi bien une marque qu'une autre.

Pour les individus vus précédemment, les individus 5, 8795 et 9595 sont très fidèles ($IL = 1$). L'individu 8 quant à lui a un indicateur de loyauté de 0.48. D'après la figure 2, l'individu 8 (indécis) a toujours acheté en majorité la marque *A*, mais nous voyons bien ici que la notion de loyauté au cours du temps prend toute son importance. Certains individus pour des mêmes valeurs d'*IL* peuvent être stables, modifier leur préférence d'une marque vers une autre (glisseurs), d'autres encore essaient une autre marque pour finalement revenir vers la marque de préférence (zappeurs). L'étude du temps est peu judicieuse pour le dernier groupe (très indécis), puisque ces derniers sont à priori dans une période d'incertitude et n'ont pas vraiment de marque de préférence. En ce qui concerne les deux groupes intermédiaires (fidèles et indécis), l'étude du temps permet de distinguer les consommateurs stables, des zappeurs ou des glisseurs. L'étude du temps pour le premier groupe peut s'avérer utile pour mettre en avant la régularité des achats. Un individu très fidèle ($IL = 1$) peu stable achète rarement sur le marché alors qu'un individu très fidèle et très stable achète fréquemment et toujours la même marque.

2.2 Analyse Factorielle des Correspondances

Il est possible de visualiser l'ensemble des consommateurs sur un diagramme triangulaire du tableau *marget*. Les individus ayant des profils similaires (même répartition des achats dans les différentes marques) sont proches les uns des autres. Plus un individu est proche d'un sommet du triangle plus la proportion des sommes dépensées pour la marque correspondante est importante. Le sommet correspond à une valeur de 1 c'est à dire 100% du budget. Cette représentation graphique n'est possible que dans le cas où le marché est constitué de trois produits ($J = 3$). Si le nombre de produits est supérieur à 3, il est possible de réaliser une Analyse Factorielle des Correspondances ([1], [2]). Le profil d'un consommateur i (i^{eme} ligne du tableau *marget*) est appelé profil-ligne. Il a été prouvé que le nuage des profils lignes i appartient à un hyperplan. Dans ce nuage de points, la ressemblance entre les profils est mesurée au travers de la distance du χ^2 . Dans le cas d'un marché possédant trois marques comme dans notre exemple, le nuage des profils lignes appartient à un plan et la représentation graphique est très proche du diagramme triangulaire. Le nuage des profils-colonnes (profil des marques) peut se superposer à la représentation du nuage des profils-lignes (relations de transition) et ainsi il est possible de situer les différentes marques auprès des profils d'individus.

2.3 Analyse Factorielle Multiple et mesure de la stabilité

L'Analyse Factorielle Multiple (AFM) peut permettre de quantifier la stabilité d'un profil au cours du temps ([3]). Il suffit de réaliser une AFM sur les T tableaux de profils (au cours des trimestres par exemple). Cette analyse permet de comparer les T tableaux, c'est à dire l'évolution ou la stabilité du profil au cours du temps. En effet, l'inertie-intra associée à chaque individu diminue avec la proximité des profils au cours du temps. Nous posons alors IS indicateur de stabilité au cours du temps qui vaut 1 si l'inertie-intra est minimum et qui décroît vers 0 pour une inertie-intra maximum. La variable ST prend deux valeurs 1 pour une inertie-intra inférieure à la moyenne de toutes les valeurs et 0 sinon. Cette variable sépare les consommateurs en deux groupes : les consommateurs stables ($ST = 1$) et les consommateurs non réguliers ($ST = 0$).

2.4 Analyses supplémentaires

Analyse de la variance

Une fois les groupes de consommateurs définis, reste la question de l'influence de la marque sur le comportement des consommateurs. Nous proposons pour finir de réaliser des analyses de la variance à un facteur :

- l'ANOVA de la variable IL (quantitative) et du facteur $pref$ (à J modalités) indique si la loyauté dépend des préférences (marque achetée en majorité)
- l'ANOVA de la variable IS (quantitative) et du facteur $pref$ indique si la stabilité de la loyauté au cours du temps dépend de la marque achetée en majorité.

Tests d'égalité des moyennes

Un test d'égalité des moyennes des valeurs IL selon la marque achetée en majorité (ou préférence) permet alors de noter l'influence éventuelle d'une marque sur l'indicateur

loyauté. Le test, pour la marque A par exemple, est le suivant :

$$H_0 : \overline{IL}_A = \overline{IL}_{\bar{A}}$$

$$H_1 : \overline{IL}_A \neq \overline{IL}_{\bar{A}}$$

avec \overline{IL}_A l'indicateur de loyauté moyen pour les individus ayant une préférence pour A et $\overline{IL}_{\bar{A}}$ l'indicateur de loyauté moyen pour les individus ayant une préférence pour les autres marques.

De la même manière, il est possible de mettre en évidence l'influence d'une marque sur la stabilité des consommateurs.

3 Résultats

3.1 Etude du premier marché

En ce qui concerne le premier marché, 63 individus n'ont rien dépensé pendant les 62 semaines considérées. Nous ne conservons donc que 10005 consommateurs. La figure 3 représente le résultat de l'Analyse Factorielle des Correspondances du tableau *marget*. La couleur attribuée aux individus correspond à l'indicateur de loyauté IL (voir tableau 1 : dégradé de gris : du plus loyal au moins loyal). Sur ce marché, on note une large préférence pour la marque A .

3.1.1 La loyauté/fidélité

Le tableau 1 représente la répartition des consommateurs selon leur préférence (marque achetée en majorité) et leur indicateur de loyauté (IL). Un nombre important d'individus est très fidèle à une marque sur ce marché (781) et plus des deux tiers des individus ont un indicateur de loyauté supérieur à 0.5.

préférence pour :	très fidèles	fidèles		indécis	très indécis	total
	$IL = 1$ noir	$IL \in [0.75; 1[$ gris très foncé	$IL \in [0.5; 0.75[$ gris foncé	$IL \in [0.25; 0.5[$ gris	$IL < 0.25$ gris clair	
A	708	3646	2889	1485	632	9360
B	12	2	9	40	116	179
C	61	13	51	105	236	466
total	781	3661	2949	1630	984	10005

TAB. 1 – Répartition des consommateurs selon leur préférence et leur indicateur de loyauté

3.1.2 La stabilité

La figure 4 représente le résultat de l'Analyse Factorielle des Correspondances (AFC) du tableau *marget*. La couleur attribuée aux individus correspond à la valeur de ST : les individus stables ($ST = 1$) sont en foncé et les autres en clair. Nous pouvons noter que certains individus très proches du produit A sont jugés comme non stables, néanmoins une grande majorité d'entre eux sont jugés comme stables. Dans le tableau 2, la répartition des consommateurs stables selon leur préférence et leur indicateur de loyauté, montre clairement que les individus stables sont plus fidèles à la marque A qu'aux autres marques.

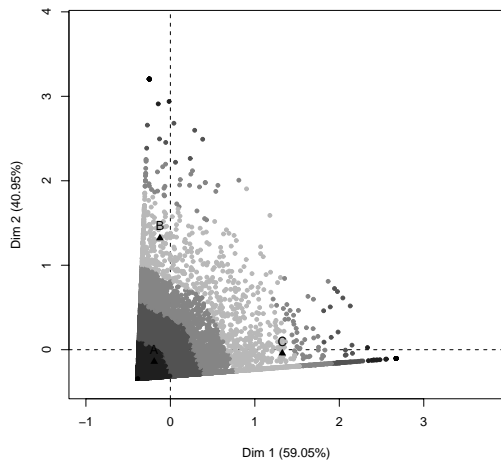


FIG. 3 – AFC de *marget* (dégradé de gris : du plus loyal au moins loyal)

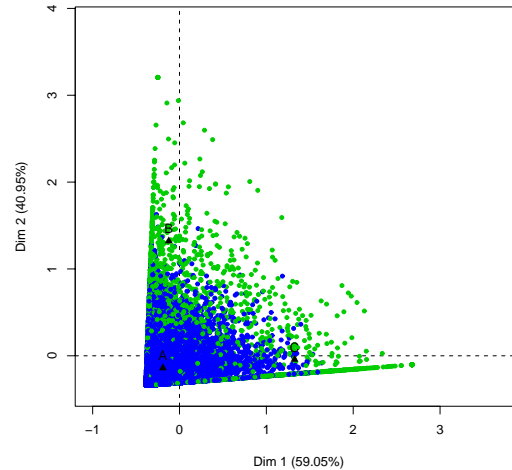


FIG. 4 – AFC de *marget* (foncé : très stable, clair : peu stable)

Même les individus très fidèles aux marques *B* et *C* ne sont pas réguliers dans leur achat sur ce marché.

	$IL = 1$	$IL \in [0.75; 1[$	$IL \in [0.5; 0.75[$	$IL \in [0.25; 0.5[$
A	708 (<i>493</i>)	3646 (<i>3447</i>)	2889 (<i>2532</i>)	1485 (<i>990</i>)
B	12 (<i>0</i>)	2 (<i>0</i>)	9 (<i>0</i>)	40 (<i>0</i>)
C	61 (<i>0</i>)	13 (<i>0</i>)	51 (<i>0</i>)	105 (<i>11</i>)
total	781	3661	2949	1630

TAB. 2 – Répartition des consommateurs selon leur préférence et leur indicateur de loyauté; en italique le nombre d'entre eux jugés stables

3.1.3 Influence de la marque

Les analyses de la variance réalisées sur les variables IL et IS indiquent une nette influence de la marque à la fois sur la loyauté et sur la stabilité des consommateurs. Le tableau 3 indique que les consommateurs achetant en priorité la marque *A* ont un indicateur de loyauté moyen de 0.68 alors que les individus achetant une autre marque que *A* ont un indicateur de loyauté de 0.33. Les trois tests sont significatifs et permettent d'affirmer que la loyauté est plus grande pour le produit *A* que pour les deux autres produits.

	\overline{IL}_{prod}	$\overline{IL}_{\overline{prod}}$	p-value
A	$\overline{IL}_A = 0.68$	$\overline{IL}_{\bar{A}} = 0.33$	$< 2.2 * 10^{-16}$
B	$\overline{IL}_B = 0.25$	$\overline{IL}_{\bar{B}} = 0.67$	$< 2.2 * 10^{-16}$
C	$\overline{IL}_C = 0.36$	$\overline{IL}_{\bar{C}} = 0.67$	$< 2.2 * 10^{-16}$

TAB. 3 – Test d'égalité des moyennes de IL selon la marque achetée en priorité

Par un procédé analogue, l'indicateur de stabilité IS vaut en moyenne 0.992, respectivement 0.898 et 0.951, pour les consommateurs achetant en majorité la marque *A*,

respectivement B et C . Les consommateurs présents sur le premier marché et achetant en majorité la marque A sont donc plus fidèles et plus stables que ceux préférant les marques B ou C .

3.2 Etude du second marché

En ce qui concerne le second marché, 30 individus n'ont rien dépensé pendant les 62 semaines considérées. Nous ne conservons donc que 10038 consommateurs. La figure 5 représente le résultat de l'Analyse Factorielle des Correspondances du tableau *marget*. La couleur attribuée aux individus correspond à l'indicateur de loyauté IL (voir tableau 4 : dégradé de gris : du plus loyal au moins loyal). Pour ce second marché, les individus sont plus répartis quant à leur préférence.

3.2.1 La loyauté/fidélité

Le tableau 4 représente la répartition des consommateurs selon leur préférence et leur indicateur de loyauté (IL). Pour ce second marché, les indicateurs de loyauté sont globalement plus faibles : plus de la moitié des consommateurs ont un indicateur inférieur à 0.5 et presque le tiers des individus ont un indicateur inférieur à 0.25.

préférence pour :	très fidèles	fidèles		indécis	très indécis	total
	$IL = 1$ noir	$IL \in [0.75; 1[$ gris très foncé	$IL \in [0.5; 0.75[$ gris foncé	$IL \in [0.25; 0.5[$ gris	$IL < 0.25$ gris clair	
D	228	1428	1941	1778	1516	6891
E	198	300	439	551	782	2270
F	5	15	113	246	498	877
total	431	1743	2493	2575	2796	10038

TAB. 4 – Répartition des consommateurs selon leur préférence et leur indicateur de loyauté

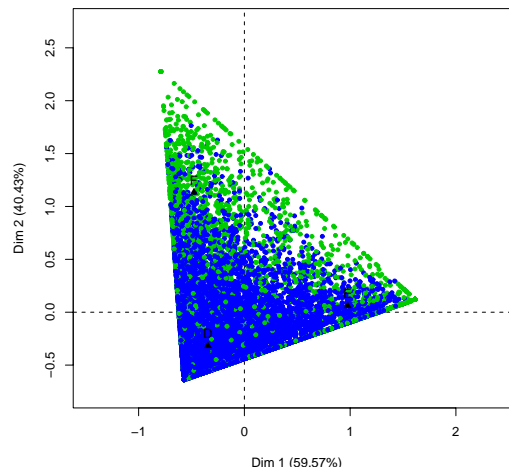
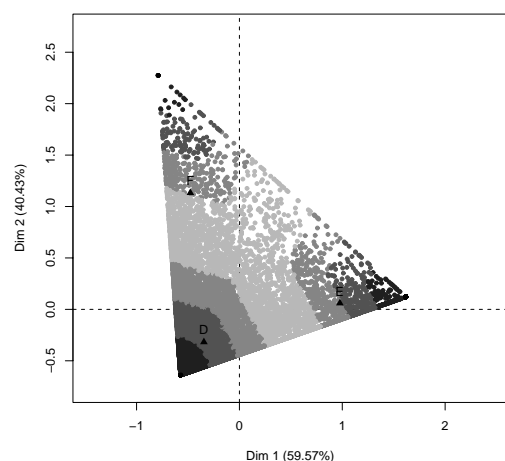


FIG. 5 – AFC de *marget* (dégradé de gris : du plus loyal au moins loyal)

FIG. 6 – AFC de *marget* (foncé : très stable, clair : peu stable)

3.2.2 La stabilité

La figure 6 représente le résultat de l'Analyse Factorielle des Correspondances (AFC) du tableau *marget*. La couleur attribuée aux individus correspond à la valeur de ST : les individus stables ($ST = 1$) sont en foncé et les autres en clair. Nous pouvons noter que certains individus très proches du produit D sont jugés comme non stables, néanmoins une grande majorité d'entre eux sont jugés comme stables.

Dans le tableau 5, la répartition des consommateurs stables selon leur préférence et leur indicateur de loyauté, montre clairement que les individus stables sont plus fidèles à la marque D qu'à la marque F . Même les individus très fidèles à la marque F ne sont pas réguliers dans leur achat sur ce marché. La marque E attire, elle aussi, des individus réguliers mais dans une moindre mesure.

	$IL = 1$	$IL \in [0.75; 1[$	$IL \in [0.5; 0.75[$	$IL \in [0.25; 0.5[$
D	228 (110)	1428 (1239)	1941 (1728)	1778 (1508)
E	198 (2)	300 (88)	439 (232)	551 (358)
F	5 (0)	15 (0)	113 (11)	246 (67)
total	431	781	3661	2949

TAB. 5 – Répartition des consommateurs selon leur préférence et leur indicateur de loyauté ; en italique le nombre d'entre eux jugés stables

3.2.3 Influence de la marque

Comme pour le premier marché, les analyses de la variance réalisées sur les variables IL et IS indiquent une nette influence de la marque à la fois sur la loyauté et sur la stabilité des consommateurs. Le tableau 6 indique que les consommateurs achetant en priorité la marque D ont un indicateur de loyauté moyen de 0.51 alors que les individus achetant une autre marque que D ont un indicateur de loyauté de 0.40. Les trois tests sont significatifs et permettent d'affirmer que la loyauté est plus grande pour le produit D que pour les deux autres produits. Néanmoins, la différence entre les marques est moins importante que sur le premier marché. Par un procédé analogue, l'indicateur de stabilité IS vaut

	\overline{IL}_{prod}	$\overline{IL}_{\overline{prod}}$	p-value
D	$\overline{IL}_D = 0.51$	$\overline{IL}_{\overline{D}} = 0.40$	$< 2.2 * 10^{-16}$
E	$\overline{IL}_E = 0.45$	$\overline{IL}_{\overline{E}} = 0.48$	$< 7.8 * 10^{-6}$
F	$\overline{IL}_F = 0.26$	$\overline{IL}_{\overline{F}} = 0.50$	$< 2.2 * 10^{-16}$

TAB. 6 – Test d'égalité des moyennes de IL selon la marque achetée en priorité

en moyenne 0.982, respectivement 0.960 et 0.923, pour les consommateurs achetant en majorité la marque D , respectivement E et F . Les consommateurs présents sur le second marché et préférant la marque D sont donc plus fidèles et plus stables que ceux achetant en majorité la marque F . La marque E est tant au niveau de la loyauté que de la stabilité intermédiaire entre les marques D et F .

4 Conclusion

Dans l'étude d'un suivi de consommation, nous avons défini une mesure de la loyauté pour un consommateur envers une marque sur la période considérée. Nous avons ainsi pu définir des groupes de consommateurs (fidèles ou non) et visualiser la position des consommateurs vis-à-vis des différentes marques. Sur le premier marché, les consommateurs se situent globalement à proximité de la marque *A*. Cette marque semble détenir une part très importante du marché. Sur le second marché, la prépondérance de la marque *D* est notable mais beaucoup moins écrasante que la marque *A* du premier marché. Les consommateurs sont mieux répartis sur ce second marché.

Dans un second temps, nous avons cherché à caractériser la régularité de cette fidélité au cours du temps et avons obtenu une typologie des consommateurs : les très fidèles, les fidèles stables, les fidèles non stables (zappeurs ou glisseurs), les indécis stables, les indécis non stables et les très indécis.

Enfin, nous avons cherché le lien éventuel entre les marques et les typologies considérées. La marque joue un rôle dans la loyauté mais aussi dans la stabilité des profils de consommation. Ainsi la marque *A* (respectivement *D*) attire des consommateurs fidèles et stables sur le premier marché (respectivement second marché). A l'opposé les marques *B* et *C* du premier marché et la marque *F* du second marché attirent des consommateurs moins fidèles et moins stables. La marque *E* est une marque intermédiaire qui prend une part de marché assez conséquente et attire des consommateurs relativement réguliers dans leur consommation.

Une étude plus approfondie peut être menée sur les consommateurs non stables à savoir vers quelle marque se dirigent les consommateurs "glisseurs". Ces derniers peuvent en effet quitter leur préférence dominante pour une autre ou atteindre cette préférence. De plus, l'étude simultanée des deux marchés est possible, en particulier l'Analyse Factorielle des Correspondances permet de visualiser si certaines marques (sur des marchés différents) correspondent aux mêmes profils de consommateurs. Une analyse de la variance avec pour facteur le marché permettra ainsi de comparer l'éventuel effet marché vis-à-vis de la loyauté ou de la stabilité.

Références

- [1] Benzecri, J.-P. : L'analyse des données tome 2 : l'analyse des correspondances, Paris : Bordas (1980)
- [2] Greenacre, M.J. : Correspondence Analysis in Practice, London : Academic Press (1993)
- [3] Escofier, B., Pagès, J. : Multiple Factor Analysis (AFMULT package), Computational Statistics and Data Analysis, **18** (1994) 121-140