

Recherche Marketing: Collecte et Analyse des données

Michel Calciu

Cours à l'Université de Lille 1 -
2006/2007

Introduction

Le concept du marketing

Dans l'époque actuelle où le savoir produire ne suffit plus pour assurer la place d'une entreprise ou d'un produit sur le marché à cause d'un environnement très concurrentiel, l'approche marketing semble offrir la clé du succès.

L'idée centrale du marketing est la concentration sur le consommateur (ou le client) et l'adaptation à ses attentes. Le marketing c'est un état d'esprit, c'est un art, mais c'est aussi une science.

Traditionnellement on fait la distinction entre deux phases du marketing, la *phase analytique* ou la recherche marketing (Marketing Research) et la *phase opérationnelle* qui regroupe le marketing opérationnel et stratégique.

Le rôle de la recherche marketing

La *recherche marketing* développe un ensemble de techniques pour l'étude du consommateur, de la clientèle et du marché. Elle détecte et enregistre les besoins et motivations des clients, qu'elle déduit souvent à partir des attitudes exprimées ou des comportements observés. Elle dispose de tout un arsenal cohérent de méthodes pour arriver à évaluer les marchés et leurs attentes, pour identifier la perception (positionnement) des produits et des services par rapport aux attentes et aux concurrents, pour segmenter la clientèle afin d'isoler et cibler les groupes que l'entreprise avec ses moyens et ses compétences peut le mieux servir.

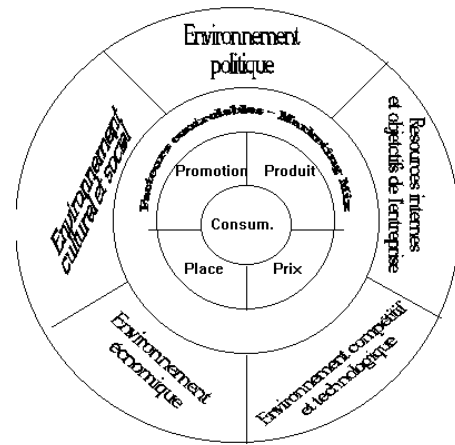


Figure 1. Le marketing et le consommateur.

Le marketing opérationnel et le marketing stratégique utilisent les informations fournies par la recherche marketing pour développer et appliquer une politique de produit, de prix, de communication et de distribution adaptée aux attentes de la clientèle ciblée.

Objectifs du cours

L'objectif de ce cours est de présenter les *techniques de recueil, traitement et analyse des données* intégrées au *cadre conceptuel de la recherche marketing*. Ces techniques sont illustrées par de nombreux *exemples et études de cas*. La solution de la plupart des traitements des données qu'elles impliquent est présentée sur *tableur* (Excel), qui est l'instrument microinformatique le plus utilisé dans les entreprises. Un *logiciel d'analyse des données* multidimensionnelles écrit par l'auteur est mis à disposition. Il intègre sur le *tableur* les méthodes d'analyse multivariée les plus importantes pour la pratique de la recherche marketing.

La typologie des études de marché

Toute recherche commence par une phase **exploratoire**. Les entretiens individuels et/ou de groupe sont des moyens de découverte de la problématique du marché (ou domaine) étudié. Avec la problématique éclairée on passe à la phase **descriptive**, à l'aide de techniques *d'observation* et de *communication*. Les techniques d'observation permettent d'enregistrer le comportement d'une manière inobtrusive. Mais dans la plupart des cas l'information est obtenue par enquête sur échantillon à l'aide d'un questionnaire. Dans cette catégorie on trouve les enquêtes ponctuelles, les enquêtes

omnibus les panels. Par fois l'*expérimentation* à travers l'étude **causale** est aussi exigée. Les tests de concept, les copy tests, tests de produit, les marchés test servent à accroître la qualité de la décision managériale.

La collecte des données et les sources d'informations

Information primaire et secondaire

Pour obtenir des informations il y a le choix entre des données préexistantes (**information secondaire**) publiées ou non publiées, externes ou internes à l'entreprise qui sont souvent peu coûteuses mais aussi peu adaptées aux objectifs de l'étude et des données recueillies directement en observant ou en interrogeant les clients (**information primaire**).

Catégories de données secondaires

Les *données internes* sur les ventes, les clients, les actions marketing passées, les coûts, sur la distribution et les acheteurs organisés souvent en base de données, sont des outils redoutables (voir essentiels) pour cibler la clientèle. Le marketing direct, la vente par correspondance, le marketing des bases de données illustrent la force que peuvent apporter les données internes à une entreprise.

Les *données externes* sont soit publiques (en libre accès ou à des coûts modiques) soit commerciales. Parmi les données *publiques* on trouve des informations publiées ou des informations accessibles on-line dans des bases de données spécialisées ou (et maintenant de plus en plus) sur Internet.

Les informations *commerciales* sont offertes par des services d'informations standardisées pour le marketing (des profils clientèle, mesure des ventes et des parts de marché, mesure d'exposition à des medias et de leur impact et plus rarement des études à la carte).

Recherche des informations primaires

Utiliser la connaissance du comportement du consommateur comme guide

Les modèles conceptuels du *comportement du consommateur* que ce soit le consommateur individuel ou le consommateur industriel sont un guide essentiel pour spécifier les informations nécessaires.

D'une manière simplifiée on suppose que le consommateur est modelé par des facteurs d'environnement (de son environnement), que par rapport à un produit ou service il est soumis à un ensemble de stimuli qui viennent de sources incontrôlables par l'entreprise (l'environnement) mais aussi à des stimuli générés par l'entreprise ou bien par le produit lui-même.

Facteurs

Parmi les **facteurs d'environnement** qui déterminent le comportement du consommateur on trouve les caractéristiques socio-démographiques (âge, sexe, occupation, état civil, niveau d'éducation, classe ou catégorie sociale, cycle de vie familial) mais aussi les caractéristiques psychosociales (personnalité, style de vie). Tous ces éléments sont importants pour la segmentation du marché.

Réponses aux stimuli

La **réaction aux stimuli** générés par le marché est progressive, les individus ne deviennent pas brusquement des acheteurs d'une marque (produit/service), ils évoluent en apprenant et traversent plusieurs phases. Il s'agit d'une étape cognitive (information, apprentissage, connaissance), d'une étape affective (ou se forment les attitudes et les préférences et où les motivations ont un rôle à jouer) et étape conative (d'action, achat). L'ordre de ces étapes diffère en fonction des situations d'achat (premier achat, achat répété, produit impliquant ou non-impliquant etc.). Pour l'entreprise il est important de savoir dans quelle phase se trouve le client par rapport au produit offert, afin de préparer et doser les stimuli à administrer et trouver un mix-marketing adapté pour faire avancer les sujets visés vers l'état d'acheteur ou d'acheteur fidèle. C'est pour cela que le recueil direct d'informations auprès de la clientèle devra permettre de déceler progressivement par

rapport au produit, le degré de connaissance, les attitudes, les intentions d'achat, les motivations, les habitudes et comportements d'achat.

Moyens de collecte d'informations primaires

Observation, enquête et expérimentation

Les données primaires sont obtenues par observation ou par communication directe. La communication suppose souvent la rédaction d'un questionnaire et c'est le seul moyen pour obtenir des informations sur les attitudes, connaissances, motivations et intentions. L'observation est plus adaptée pour procurer des données comportementales.

Le questionnaire

Le questionnaire est l'instrument le plus fréquemment utilisé pour le recueil direct des informations. Après une rédaction attentive qui essaye de couvrir tout les informations à obtenir, suivi par des phases de prétest et correction, le questionnaire est administré (par voie postale, téléphone ou télématique) à un échantillon représentatif de la population ciblée. Les questionnaires remplis sont collectés et dépouillés. Toutes les réponses sont codifiées et enregistrées dans des tableaux de données (appelés parfois bordereaux de dépouillement) où chaque ligne représente un questionnaire, comme dans la Figure 2.

(a) exemple de questionnaire :

Questionnaire No. (1)
 Courte présentation des objectifs de l'étude....
 1) Laquelle des marques suivantes préférez-vous? (2)

Marque A ☐ 1
 Marque B ☐ 2
 Marque C ☐ 3
 Autres ☐ 4

2) Qui vous a signalé cette marque (OUI=1 NON=0)

la Pub ☐ (3)

la presse spécialisée ☐ (4)

des amis ☐ (5)

autres sources ☐ (6)

3) Indiquez votre degré d'accord sur la possession des attributs suivants par les marques A,B et C: (1 = pas du tout d'accord, 2 = pas d'accord, 3 = d'accord 4 = tout à fait d'accord)

	Marque A		Marque B		Marque C	
	1	2	1	2	1	2
	3	4	3	4	3	4
Pas cher	OOOO(7)		OOOO(11)		OOOO(15)	
Bonne qualité	OOOO(8)		OOOO(12)		OOOO(16)	
Innovant	OOOO(9)		OOOO(13)		OOOO(17)	
Commode	OOOO(10)		OOOO(14)		OOOO(18)	

4) Estez vous ? (19) 5) Quel est votre age ?(20)

Femme ☐ 01
 Homme ☐ 02

(b) exemple de tableau de dépouillement d'enquête

(ICI lien vers ou intégrer question.html)

	De
	Anr

Figure 2. Le questionnaire: précodage, codage et

Le questionnaire présenté en Figure 2 est un exemple réduit et simplifié. Un vrai questionnaire (le cas CAMIP) est donné en annexe ainsi qu'une collection des types de questions qui tient compte d'un modèle hiérarchique du comportement du consommateur.

La codification des questions

Type de question	Nombres de colonnes (variables) par question
Questions nominales :	
Q. au choix multiple exclusif	une
Q. au choix multiple non exclusif	autant de colonnes que le nombre de modalités (codées 0 ou 1)
Q. ouvertes (autres...)	une et on augmente le nombre de modalités en fonction des réponses
Questions. ordinales	plusieurs modalités (tant de colonnes que propositions possibles)
Questions. quantitatives	une; si échelle d'attitude une colonne pour chaque item

Tableau 1.1 - Règles d'enregistrement des réponses aux questionnaires dans les tableaux de données d'enquête

Il est possible d'enregistrer les réponses au questionnaire directement sur l'ordinateur. Les questionnaires automatisés sur ordinateur ont l'avantage d'éliminer en grande partie les fautes de frappe, ils codent et enregistrent les données automatiquement dans des tableaux. Un exemple de questionnaire automatisé sur tableur est donné en annexe. Sur le Web de l'Internet le mécanisme des formulaires (Forms) facilite la saisie des réponses à l'aide d'objets d'interface adaptés (zones de texte, zone d'entrée, cases d'options, cases à cocher, listes à sélection unique ou multiple etc.). Les informations saisies sont captées sur le serveur de celui qui mène l'enquête et enregistrées dans une base de données ou autrement traitées par des programmes adaptés, qui profitent d'un autre standard qui s'est imposé sur Internet le CGI (Common Gateway Interface). En profitant des deux standards les Forms et les CGI il est relativement facile de mener des enquêtes interactives en temps réel et à distance. Des informations supplémentaires concernant la préparation des enquêtes sur Internet se trouvent dans l'annexe.

Nature des informations

• Types d'informations

Informations quantitatives résultent de l'observation des comportements des intervenants sur le marché, elles sont relativement objectives et contrôlables et répondent aux questions " combien ? " et " comment ? " (ex. notoriété des produits, possession des produits, quantités achetées, fréquences d'achat, lieux d'achat, modes d'information sur les produits etc.)

Informations qualitatives s'intéressent aux facteurs qui déterminent les comportements et sont de nature subjective. Leur recueil et interprétation sont en général plus difficiles (ex. les motivations, perceptions, les opinions et attitudes, les préférences etc.)

La mesure

En marketing le questionnaire sert aussi comme instrument de mesure.

La **mesure** est définie: comme étant formée de " règles pour attribuer des nombres à des objets, de

telle sorte qu'elles représentent des quantités d'attributs " (1. on mesure des attributs des objets, non pas les objets eux-mêmes; 2. la manière dans laquelle les nombres sont attribués n'est pas spécifiée, les nombres jouent un rôle de symbole, ce ne sont pas les propriétés intrinsèques des nombres, qui sont prises en compte mais les propriétés des attributs qu'ils représentent). Une autre définition considère la mesure "comme le moyen d'obtenir des symboles représentant les propriétés des personnes, d'objets, d'événements ou d'états, de telle sorte que ces symboles aient les mêmes relations entre eux que les choses représentées".

Les échelles de mesure

Les **échelles nominales**: ont comme seule propriété *l'identité*, la plus simple des propriétés d'une échelle de nombres. L'identité exprime l'appartenance des objets étudiés à une catégorie. Elles sont les plus pauvres en informations.

Les **échelles ordinales** ont une seconde propriété des échelles de nombres, *l'ordre*, qui est ajouté à la première. Elles permettent de classer les objets selon un ordre donné.

Elles apparaissent dans les mesures de préférences sous la forme d'échelles de classement, comparaison par paires, tiercé des préférences etc..

Les **échelles d'intervalle** ajoutent une troisième propriété des échelles de nombres, le fait que *l'intervalle* entre les nombres a un sens, ce qui signifie que les *différences* peuvent être comparées. (exemple: les échelles de température Celsius et Fahrenheit)

Les **échelles de ratio** se différencient par rapport aux échelles d'intervalle par le fait qu'elles possèdent un *zéro absolu*.

Le tableau de données: documentation, vérification, transformation des données.

Organisation d'un tableau de dépouillement d'enquête

Dans la Figure 3. est présentée une organisation possible sur tableur des données collectées par enquête.

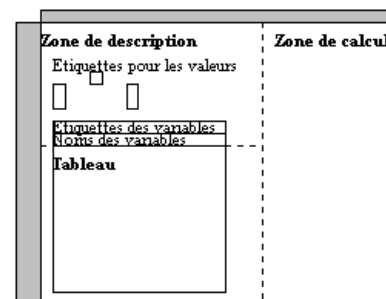


Figure 3. Organisation des données sur une feuille de calcul (sur tableur)

Variables et observations

Chaque colonne du tableau représente une variable (voir par exemple le tableau de la Figure 2.) . Pour faciliter la sélection des variables pour les éventuels traitements, elles doivent avoir des noms.

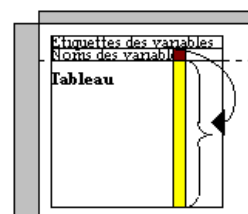


Figure 4. La place et le nom d'une variable dans le tableau de données (tableur)

Analyses préliminaires des données: tableaux de fréquence et statistiques descriptives

Les traitements usuels

Les premiers traitements appliqués aux données cherchent à résumer et à mettre en évidence les éventuelles erreurs et/ou données aberrantes qui ont échappé à la correction au moment de la saisie. Il s'agit des statistiques descriptives et de tableaux de fréquence (simples et croisés). Une grande partie des études marketing (90%) ne vont pas plus loin dans les analyses pour décrire le marché.

Tableaux de fréquence

Les tableaux de fréquence montrent le nombre des cas enregistré dans différents intervalles ou catégories. Les tableaux de fréquences simples (tris à plat) comptent les fréquences pour une seule variable. Ils sont souvent représentés sous forme d'**histogrammes**.

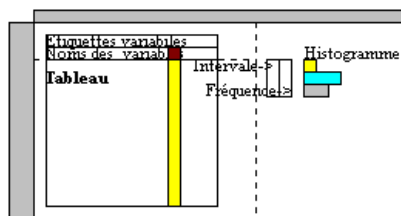


Figure 5. Le tableau de fréquence simple (ou tris à plat). Sur tableur il y a la fonction FREQUENCE(nom variable, matrice d'intervalles) qui le calcule.

voilà comment on calcule sur un tableur la fréquence et la fréquence cumulé des réponses à la première question de la mini-enquête présenté en Figure 3.

	D	E	F	G	H
1		Interv.	Freq.	Freq.Cum	<--Formule
2	Marque A	1	3	3	=F3
3	Marque B	2	4	7	=G3+F4
4	Marque C	3	7	14	..
5	Marque D	4	2	16	..

Tableau 1.2 - Calcul de la distribution et répartition de fréquence sur tableur

Anatomie d'un distribution de fréquence

A partir de la distribution de fréquence et la répartition des fréquences (ou la fréquence cumulée) on calcule un grand nombre de statistiques descriptives. La Figure 6. montre graphiquement comment des statistiques comme la moyenne, la médiane, les quartiles sont déduites à partir d'une distribution de fréquence. Elle montre aussi quel est l'intérêt d'utiliser la fréquence cumulée pour identifier certaines de ces mesures. Pour mieux reconnaître l'allure de la distribution de fréquence empirique on trace sur l'histogramme le **polygone de fréquence** en réunissant les points qui se trouvent au milieu sur le côté haut des barres.

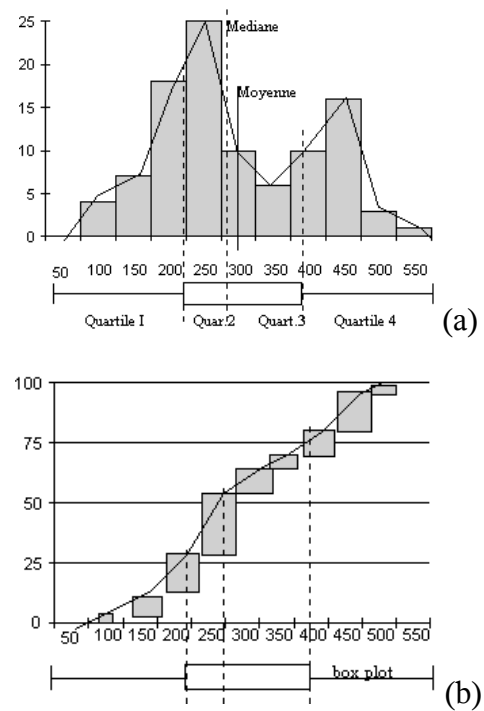


Figure 6. Anatomie d'une distribution de fréquence (a) et d'une répartition de fréquence (b) . Le lien entre un histogramme et un box plot

Les tableaux croisés

Les tableaux croisés (tris croisés) comptent la fréquence des cas ayant de caractéristiques (intervalles) communes sur deux ou plusieurs variables. Ils résument l'information contenue dans les variables juxtaposés et permettent d'analyser l'association et les liens entre les variables. Les tableaux croisés de deux variables seront traités au chapitre sur les analyses bivariés et les tableaux croisés multiples apparaîtront dans les chapitres qui portent sur l'analyse de données multidimensionnelle.