

ETUDES ET RECHERCHE MARKETING



S. MAYOL



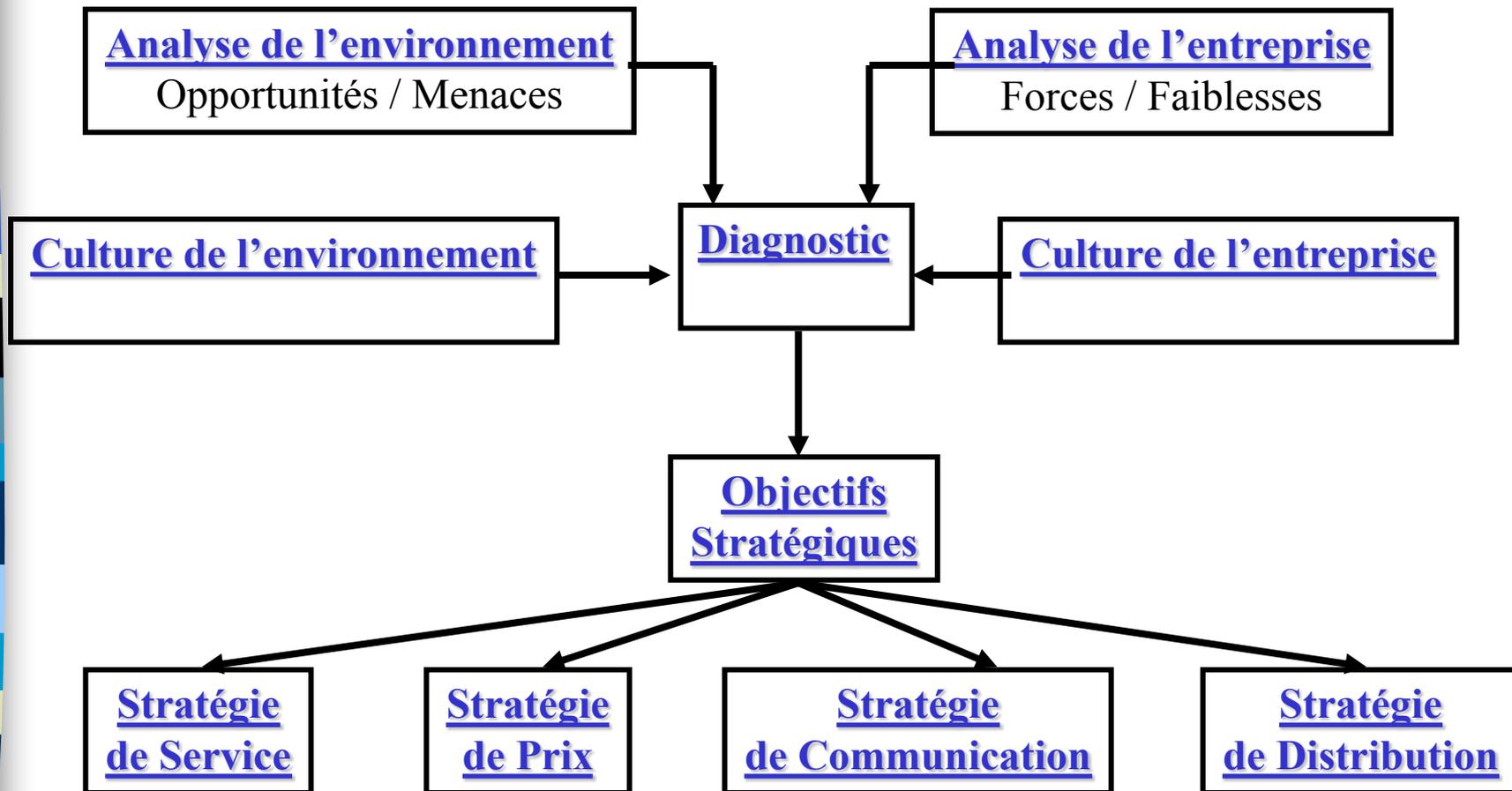
Introduction : démarche marketing

- **Comprendre** le marché

pour

- **Agir** sur le marché

Démarche marketing d'après la méthode LCAG





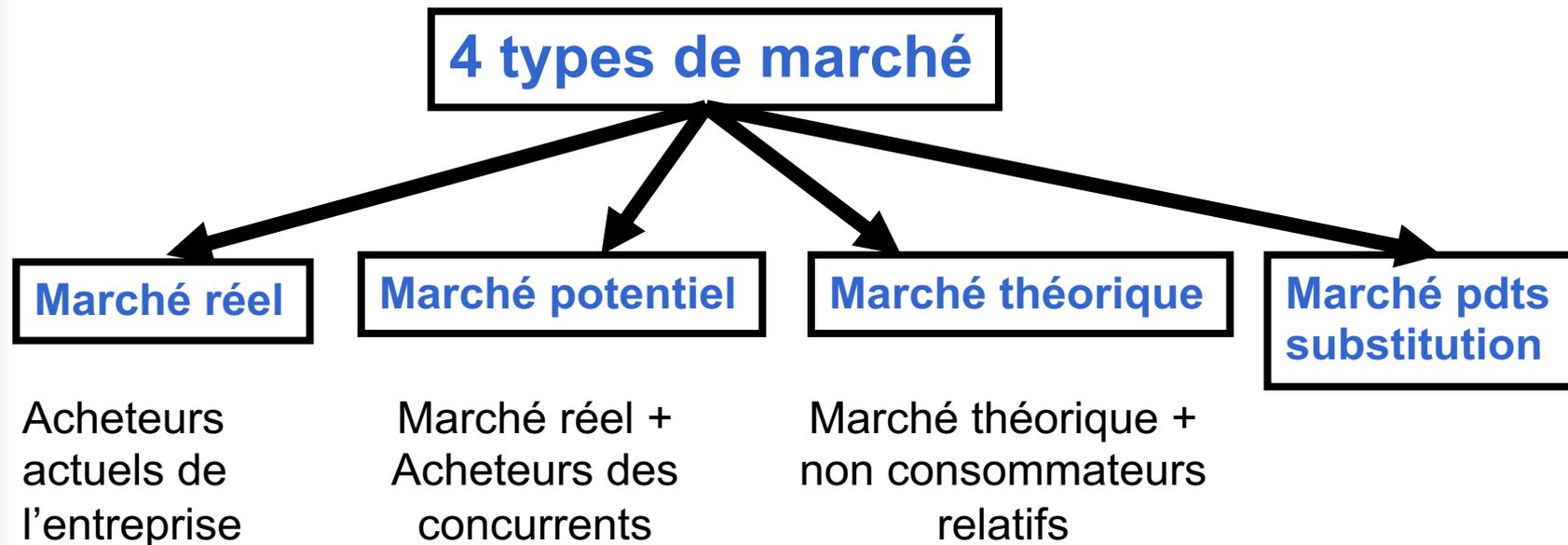
Introduction : la notion de marché - définition

- En économie : lieu de rencontre entre offre et demande

C'est-à-dire, endroit où acheteur et vendeurs se rencontrent pour pratiquer des échanges commerciaux

- En marketing : définition plus large : ensemble des agents économiques qui achètent e vendre un produit

Introduction : la notion de marché - définition



Introduction : la notion de marché - agents

CONSOMMATEURS	INTERVENANTS SUR L'ACTE D'ACHAT	OFFREUR S	AUTRES AGENTS
Conso actuels Prospects Non conso relatifs Non conso absolus	Influenceurs Prescripteurs Leaders opinion Entourage Médias Assoc. conso Décideurs Acteurs Payeurs	Entreprises offrant le même produit	Etat CCI Administ. Syndicats Assoc Pro Org. Inter



Introduction : Etude de marché - définition

Une étude de marché a pour
finalité de fournir de
l'information nécessaire à
l'amélioration de la décision

Introduction : Etude de marché - objectifs

- **Identifier opportunités commerciales**
(besoins non satisfaits, attributs à développer)
- **Expliquer le cpt du conso**
(motivations, freins, attitudes, préférences, processus d'achat)
- **Connaître le positionnement actuel de l'ent. et de ses produits et donc de définir le positionnement idéal**
- **Évaluer une action déjà entreprise**
(lancement de pdt, de marque, de campagne)
- **Faire des prévisions de vente pour estimer la demande future**
- **Mesurer la notoriété de l'entreprise et de ses produits**
- **Tester un produit ou une marque avant son lancement**



Introduction : les études marketing

- Ensemble des activités qui consistent à recueillir, analyser, exploiter les informations relatives à un problème marketing ou commercial afin de permettre aux responsables de prendre les meilleures décisions possibles



Introduction : les études marketing

- Problème au sens large: difficulté, besoin de comprendre, obtenir des informations
- Informations de 2 ordres :
 - Qualitative
 - Quantitative

Introduction : le marché des études

- Les études pèsent 16,75 milliards € dans le monde
- L'UE représente 37% du total
- La Grande Bretagne est le premier marché européen

	Acteurs en France	CA 2001 (K€)	Evol	Effectif
1	Taylor Nelson Sofrès	97000	+7,8%	550
2	Ipsos France	86338	+18,2%	581
3	ACNielsen	79876	+10,0%	550
4	TNS Sécodip	71971	+8,3%	837
5	Groupe GFK	47500	+5,6%	350
6	IMS Health	40398	+25,5%	185
7	IRI Secodip	36091	+10,1%	356
8	Médiamétrie	35800	+7,8%	316
9	CSA TMO	34000	+3,0%	100
10	Research International	28000	=	136

1ère partie : Les études marketing

S. MAYOL



I - le contenu des études marketing

1-1 Le domaine des études

- **Etude de l'environnement économique**
 - Domaines économiques (secteurs, branches...)
 - Domaines monétaires et financiers (inflation...)
 - Domaines sociaux (mentalités, conflits...)
 - Démographie
 - Technologie, technique
 - Domaine juridique (législation, réglementation...)

I - le contenu des études marketing

1-1 Le domaine des études

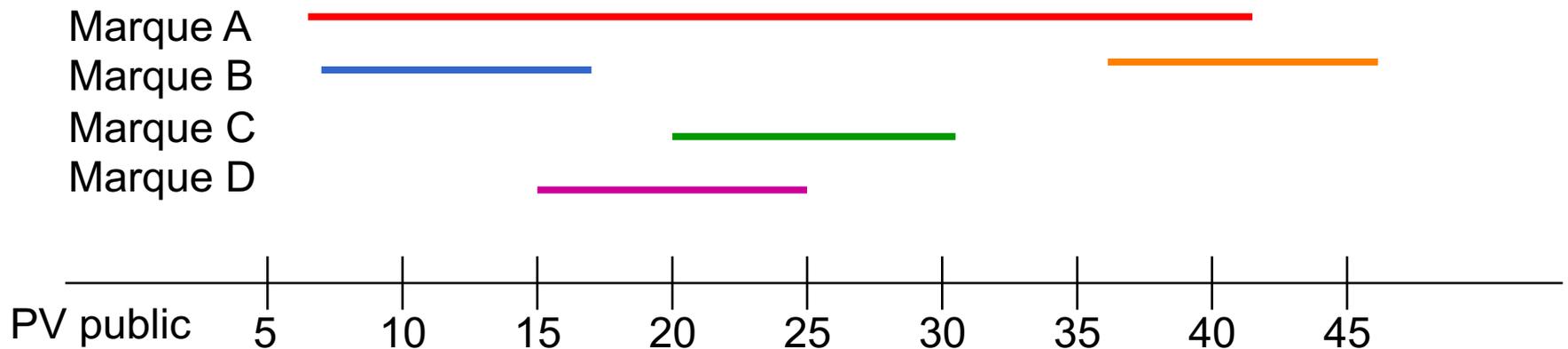
- Etude des consommateurs ou utilisateurs

Marché actuel de la concurrence	Marché actuel de l'entreprise	Non consommateurs relatifs	Non consommateurs absolus
Marché potentiel 			
Marché actuel de la profession 			
Marché théorique de la profession 			
Etudes et Recherche - S. Mayol			

I - le contenu des études marketing

1-1 Le domaine des études

- Etude de la distribution
- Etude de la promotion des ventes et de la publicité
- Etude de la concurrence
- Etude du produit





I - le contenu des études marketing

1-1 Le domaine des études

- Les études interviennent à chaque stade de la démarche marketing :
 - Marketing stratégique : objectifs, stratégie
 - Marketing opérationnel : suivi des actions, contrôle des performances



I - le contenu des études marketing

1-1 Le domaine des études

On distingue deux types d'études

- Les études qui visent à résoudre des problèmes stratégiques
- Les études liées aux actions commerciales ou au « marketing mix » (études à dominante opérationnelle)



I - le contenu des études marketing

1-2 L'étude de marché

- Constitue l'étude marketing la plus complète pouvant être menée par une entreprise
- Elle est caractérisée par un degré important de complexité liée à la variété des informations à recueillir : demande, offre, distribution, environnement



I - le contenu des études marketing

1-2 L'étude de marché

L'étude de la demande

Situation actuelle, évolutions passées et perspectives d'avenir

■ Informations **quantitatives** :

- Demande globale : ventes en volume, CA, saisonnalité, répartition géographique...
- Identification des segments, de leur situation et de leur évolution

■ Informations **qualitatives** :

- Intervenants dans le processus d'achat
- Facteurs explicatifs des comportements d'achat et de consommation
- Motivations, freins, attentes...

I - le contenu des études marketing

1-2 L'étude de marché

L'étude de l'offre

Informations sur l'entreprise et ses concurrents

- Informations **quantitatives** :
 - Structure : nombre d'entreprises, poids des entr. Les plus importantes
 - Performances commerciales : CA, PDM, notoriété, profitabilité,...
 - Positions concurrentielles
- Informations **qualitatives** :
 - Images des produits et des marques
 - Images institutionnelles
 - Positionnements
 - Stratégies marketing et caractéristiques en terme de marketing mix



I - le contenu des études marketing

1-2 L'étude de marché

L'étude de la distribution

- Caractériser les formes de distribution dominantes sur le marché étudié : formes de commerce, méthodes de vente, types de PDV, pratiques commerciales...
- Recueillir des informations analogues à celles de l'étude de l'offre



I - le contenu des études marketing

1-2 L'étude de marché

L'étude de l'environnement

- Éléments susceptibles d'influencer le marché
- Composantes technologique, démographique, juridique, socioculturelle, économique

I - le contenu des études marketing

1-3 Exemple : les tweens

- Pour les sociétés Publicis, Coca-Cola, Adidas, ... , les jeunes sont une cible privilégiée
- En Europe, aux Etats-Unis ou au Japon, les 12-25 ans forment un groupe de référence qui anticipe et amplifie les phénomènes de mode
- Très sensibles aux codes sociaux, ils se passionnent pour tel courant vestimentaire (ex. : Helly Hansen, Sean John, Lacoste...)

I - le contenu des études marketing

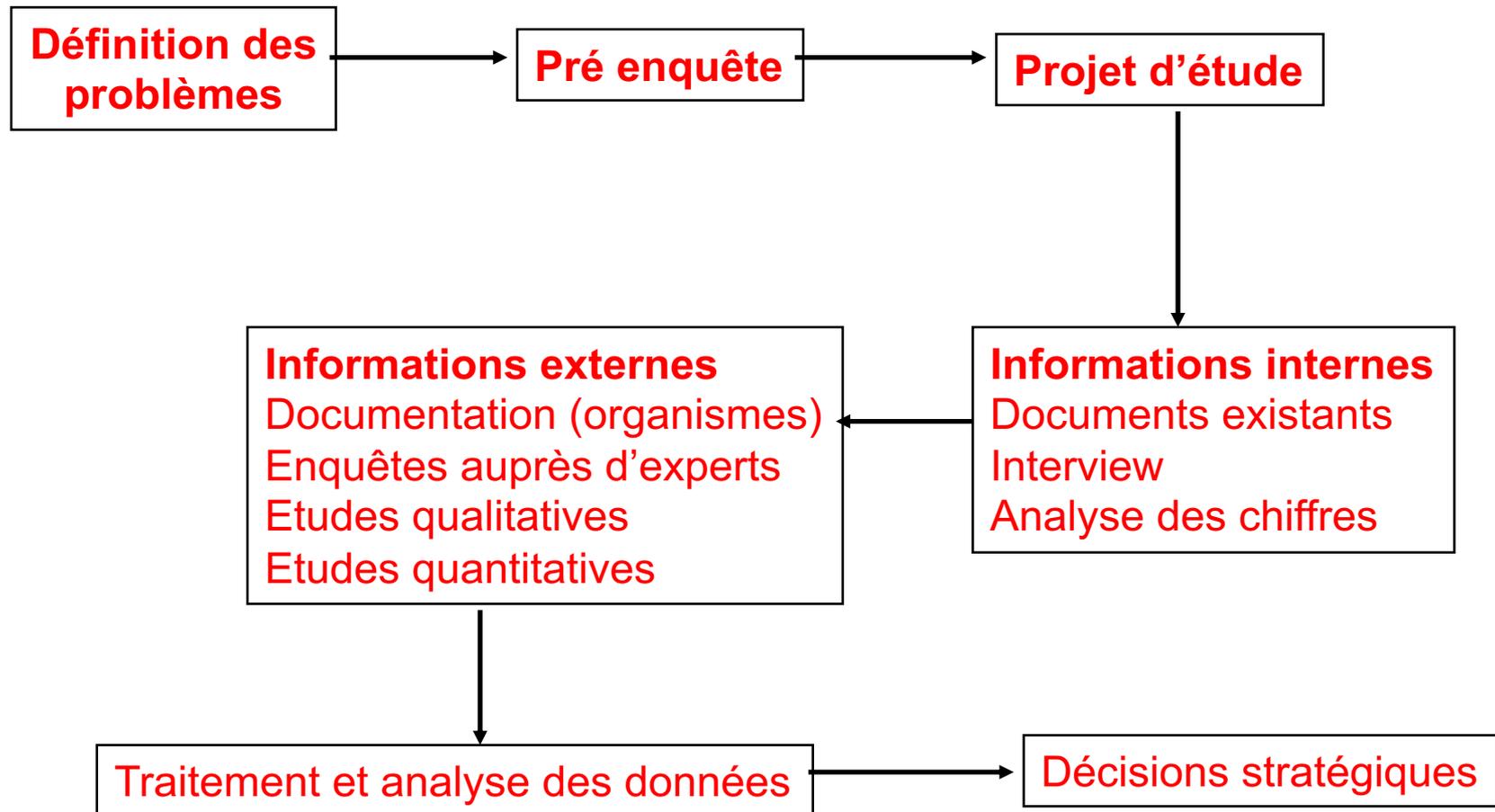
1-3 Exemple : les tweens

- En réalisant des études , les entreprises cherchent à mieux comprendre ces générations :
 - En Europe : 120 millions de personnes
 - 33 Md d'€ de pouvoir d'achat
- « Teen-agers and twenty something » = **tweens**

Les études sont une aide essentielle à la prise de décision marketing et commerciale

I - le contenu des études marketing

1-4 le déroulement d'une étude de marché





I - le contenu des études marketing

1-4 Le déroulement d'une étude de marché

- 1. Une étude part d'un problème commercial rencontré par une entreprise dans le cadre de son activité
 - Commanditaire
 - Cahier des charges
 - Appel d'offres
- 2. Une réunion préalable entre le commanditaire et les responsables de l'étude permet de définir :
 - Le problème à résoudre et la finalité de l'étude
 - Les objectifs de l'étude

I - le contenu des études marketing

1-4 Le déroulement d'une étude de marché

- 3. En fonction des objectifs et du budget, les responsables déterminent les sources d'information à privilégier et les techniques d'étude :
 - Informations primaires
 - Informations secondaires
- 4. La proposition d'étude indique :
 - Le plan d'étude
 - Des éléments complémentaires : budget, modalités de règlement, confidentialité des résultats...
- 5. Accord et signature de la proposition d'étude
- 6. Supports : questionnaire, guide d'entretien, plan d'animation, grille d'observation

II - les études documentaires

2-1 les sources internes

■ STATISTIQUES INTERNES :

- Chiffre d'affaires
- Marge par produit
- Nombre de clients
- Circuits, formes de distribution
- Catégories de clientèles
- Tonnage de production
- Régions géographiques
- Etc ...

II - les études documentaires

2-2 les sources externes

■ Organisations internationales

- OCDE (organisation de coopération et de développement économique)
 - Statistiques bimestrielles portant sur les pays membres
 - Études sur les problèmes économiques
 - Statistiques du commerce extérieur
- UE
 - Bulletin mensuel de statistiques
 - Rapports sur les principaux faits économiques
- ONU
 - Bulletin mensuel de statistiques internationales

II - les études documentaires

2-2 les sources externes

■ Organisations françaises

- INSEE (institut national de la statistique et des études économiques)
 - Économie, démographie, ménages, entreprises, consommation, niveau de vie, etc...
- CFCE (centre français du commerce extérieur)
 - Études par pays et par produits
- CREDOC (centre de recherches pour l'étude et l'observation des conditions de vie)
 - Consommation
- INED (institut national d'études démographiques)
 - Populations

II - les études documentaires

2-2 les sources externes

■ Organisations françaises

- Documentation française
 - Études documentaires et problèmes économiques
- Journal officiel
 - Lois, décrets, avis, rapports et nominations
- Ambassades
- Ministères
- Chambres syndicales



III - les études qualitatives

3-1 caractéristiques

■ OBJECTIFS:

- Rechercher les raisons d'un comportement,
- Identifier des variables explicatives,
- Trouver et comprendre les freins et les motivations,
- Expliquer une attitude
- Connaître l'ensemble des raisons qui poussent un consommateur de telle ou telle façon



III - les études qualitatives

3-1 caractéristiques

■ LIMITES:

- Si les études qualitatifs permettent de comprendre les facteurs explicatifs, elles ne permettent pas de savoir le poids relatif de ces différentes raisons qui devra être mesuré en second lieu par une étude quanti
- Difficulté d'être conscient de ses « motivations », de les verbaliser et de les hiérarchiser
- Risque de limiter les réponses à ce qui est socialement acceptable



III - les études qualitatives

3-2 Les techniques d'entretien

■ Entretien individuel en profondeur

- Explorer la face cachée d'un discours
- Peu de questions
- Libre parole aux enquêtés
- Répétitions, gestes, silences ou lapsus sont des informations supplémentaires
- On cherche à comprendre le sens réel des mots



III - les études qualitatives

3-2 Les techniques d'entretien

■ Entretien individuel en profondeur

- L'entretien non directif : s'appuie sur les méthodes d'investigation issues de la psychanalyse
- L'entretien semi directif : variante du précédent. Très courant en étude marketing. Le guide d'entretien récapitule les grands thèmes qui doivent être abordés. Si besoin l'enquêteur relance après un silence prolongé

III - les études qualitatives

3-2 Les techniques d'entretien

■ L'entretien semi directif

Types de relance	Objectifs	Exemples
Relance simple	Encourager le répondant	Reprendre les deux ou trois derniers mots de l'interviewé et laisser la phrase en suspens
Relance différée	Revenir sur un point abordé trop rapidement	« tout à l'heure, vous m'avez dit... » (reprendre 4 ou 5 mots)
Relance interprétative	Clôturer un thème	« si je vous comprends bien, pour vous, tout se passe comme si... »



III - les études qualitatives

3-2 Les techniques d'entretien

■ L'entretien semi directif

- Phase d'introduction : créer un climat de confiance, informations non retenues

(exemple : nous allons parler de produits de beauté, comment réagissez vous quand vos amis abordent le sujet)

- Phase de centrage du sujet : introduction du sujet par une question ouverte

(exemples : racontez moi vos dernières vacances ; supposez que vous gagniez une semaine en voyage organisé, comment réagiriez vous ?)

III - les études qualitatives

3-2 Les techniques d'entretien

■ L'entretien semi directif

- Phase d'approfondissement : revenir sur un des points évoqués précédemment, mais trop rapidement. Vérifier que tous les thèmes prévus dans le guide ont été abordés

(exemples: tout à l'heure, vous m'aviez dit que vous détestez les émissions de télé réalité... ; si je comprends bien, vous n'aimez pas les présentateurs de ces émissions)

- Phase de conclusion : ramener l'enquêté à la réalité. L'entretien a permis de collecter des informations intimes et personnelles.

(au fond, vous me disiez tout à l'heure que tous ces produits de beauté ne servaient pas à grand-chose... il doit bien avoir du vrai là dedans...)



III - les études qualitatives

3-2 Les techniques d'entretien

■ Les Réunions de groupe

- Réunion de 6 à 12 personnes
- Dans une salle spécifiquement aménagée
- Sous la direction d'un animateur
- Les débats sont enregistrés et filmés
- La réunion dure plusieurs heures
- Degré d'homogénéité des participants
- Participants sont rémunérés pour leur participation



III - les études qualitatives

3-2 Les techniques d'entretien

■ Les Réunions de groupe

Principe de la dynamique de groupe: les membres du groupe se stimulent les uns les autres grâce à l'échange d'idées et aux réactions qui s'ensuivent

4 grandes catégories d'entretiens de groupe:

III - les études qualitatives

3-2 Les techniques d'entretien

■ Les Réunions de groupe

- Groupe de réflexion (focus group): réunit des personnes choisies en fonction de leur expertise par rapport au problème à résoudre. Utilisé pour images de marque, recensement des critères de choix d'un produit, l'étude des freins et motivations. Ex : mères de famille pour couches culottes
- Groupe de discussion : moins structurée, met en présence des personnes non spécialistes du secteur. Utilisé pour des études de motivation vis-à-vis d'un produit.

III - les études qualitatives

3-2 Les techniques d'entretien

Les Réunions de groupe

- Groupe nominal : alternance de phases de discussion collective et de réflexions individuelles. Utilisée pour la recherche d'idées nouvelles, l'apprentissage d'un vocabulaire, la description de situations de consommation et l'identification de critères de choix
- Groupe Delphi : réunit 6 à 10 experts très compétents et physiquement éloignés les uns des autres. Ces experts ne se connaissent pas. Les experts ne savent pas qui a répondu à quoi. On fournit à chaque expert la réponse centrale du groupe (moyenne, médiane et intervalle de dispersion). On observe une convergence et un lissage des réponses du groupe

III - les études qualitatives

3-2 Les techniques d'entretien

Autres techniques

- Méthodes projectives : trouvent leurs racines dans la psychologie clinique. Repose sur les travaux de Freud : la projection est une activité psychique qui expulserait de sa conscience des sentiments répréhensibles pour les attribuer à autrui. Utilisés pour les freins et les motivations des consommateurs

Exemples de tests :

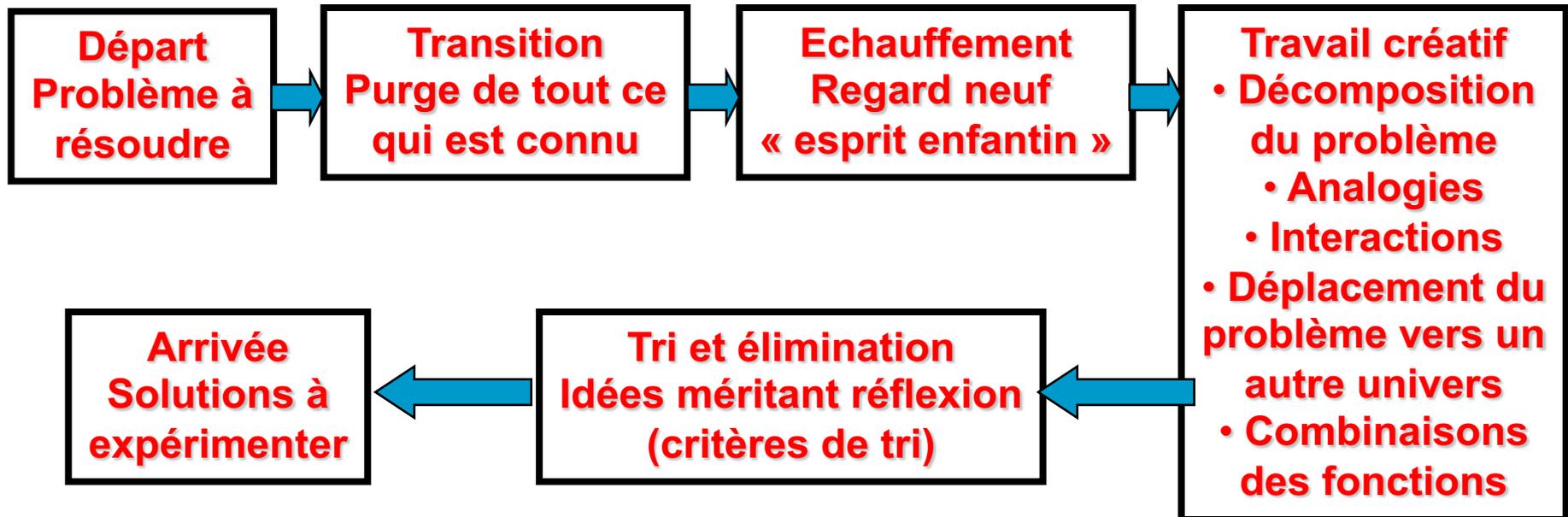
- Thematic Apperception Test
- Bande dessinée
- Portrait chinois
- Associations de mots

III - les études qualitatives

3-2 Les techniques d'entretien

■ Autres techniques

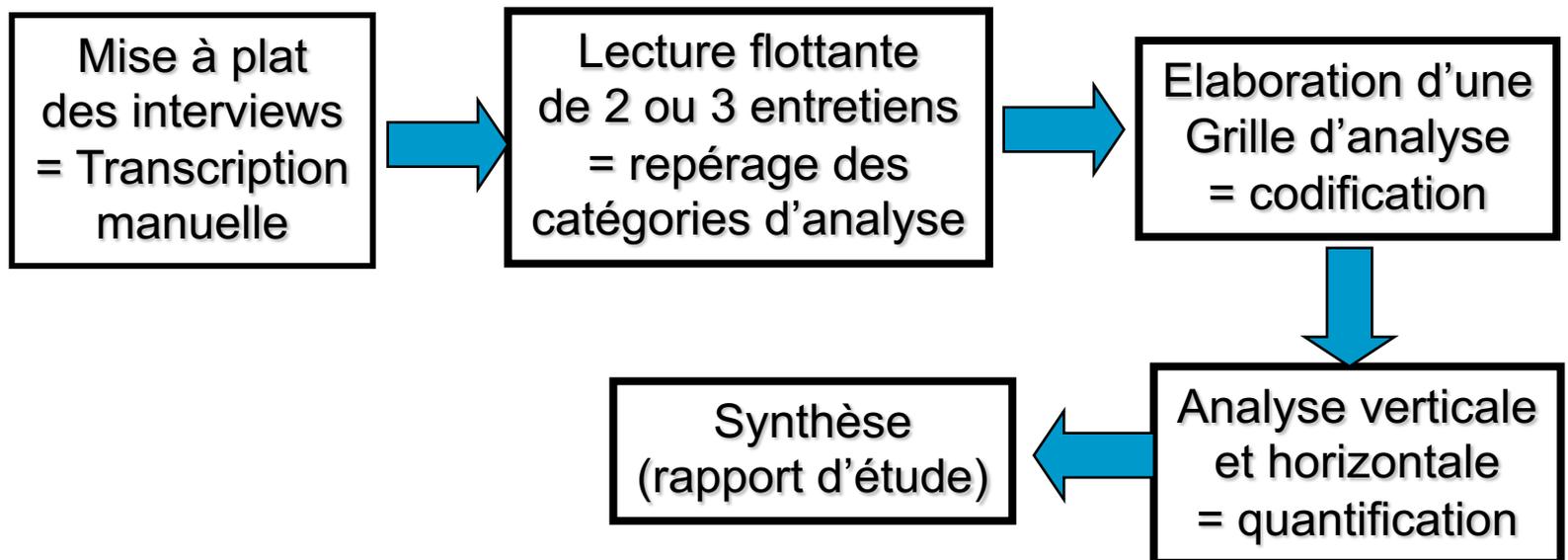
Les méthodes de créativité : utilisées pour rechercher des idées de nouveaux produits. La technique la plus connue est celle du brainstorming



III - les études qualitatives

3-3 L'analyse des données

■ L'ANALYSE DE CONTENU



III - les études qualitatives

3-3 L'analyse des données

■ L'ANALYSE DE CONTENU

- Transcription: c'est la copie sur un support papier des enregistrements audio ou vidéo des interviews. Chaque silence, répétition, lapsus est soigneusement noté
- Catégorisation: après lectures de quelques interviews, on repère les thèmes, les sous thèmes, les évocations. Ces éléments forment les catégories et les sous catégories de l'analyse de contenu

Catégorie	Odeur	Relation avec l'entourage
Sous catégorie	Fleuri Musqué Poivré	Séduction Attirance Sensualité



III - les études qualitatives

3-3 L'analyse des données

■ L'ANALYSE DE CONTENU

- Codification: les différentes catégories structurent la grille d'analyse. Le chargé d'études reprend chaque interview puis ventile les unités du discours (mots, phrases, thèmes) entre ces catégories
- Quantification: on calcule les fréquences d'apparition des différentes catégories

III - les études qualitatives

3-3 L'analyse des données

■ L'ANALYSE DE CONTENU

Exemple simplifié

Les quatre premières interviews d'une étude portant sur les perceptions d'un parfum donnent les résultats présentés dans le tableau ci-dessous.

✦ tableau 13 : la quantification d'une analyse de contenu ✦

Catégories X	Individu 1	Individu 2	Individu 3	Individu 4	Total
X1- Odeurs	oui*	oui	oui	oui	4/4 =
fleurie	2**	0	1	0	100 %
poivrée	1	0	1	1	
musquée	3	0	0	4	
cuir	1	0	1	0	
agréable	5	4	9	9	
désagréable	4	6	4	4	
Sous-total	16	10	16	18	60
X2 - Entourage	oui	oui*	oui	non	3/4 =
séduire	0	4	2	0	75 %
drague	0	3	2	0	
sortie amis	5	2	2	0	
attirance	0	5	4	0	
sensualité	1	0	10	0	
Sous-total	6	14	20	0	40
Totaux (X1 + X2)	22	24	36	18	100

* : oui = la catégorie a été citée au moins une fois par l'individu ; non = la catégorie n'a pas été abordée.

** : le chiffre indique le nombre de fois où la catégorie ou sous-catégorie apparaît chez l'individu *i*.

III - les études qualitatives

3-3 L'analyse des données

L'ANALYSE DE CONTENU

Analyse horizontale. L'odeur a été citée par 100 % des répondants, devant l'entourage (75 %). Si l'on accepte de ne pas donner le même poids à chaque individu, on observe que l'odeur fait l'objet de 60 % (= 60/100) des citations, contre 40 % (40/100) pour l'entourage. De même, lorsque l'odeur est mentionnée, la sous-catégorie « agréable » représente 45 % (= 27/60) des évocations, devant « désagréable » (18/60 = 30 %), etc.

Analyse verticale. Pour l'individu 1, c'est l'odeur qui prime (16/22 = 73 % des citations totales de l'interview) par rapport à l'entourage (6/22 = 27 %) ; pour l'individu 2, c'est l'inverse : 58 % (= 14/24) des unités de son discours concernent l'entourage, contre 42 % (= 10/24) pour l'odeur, etc.



III - les études qualitatives

3-3 L'analyse des données

L'ANALYSE FACTORIELLE DES CORRESPONDANCES

- Les fréquences d'apparition des catégories sont simples à calculer mais pauvres sur le plan marketing
- Les traitements multivariés améliorent ces résultats car ils facilitent la mise à jour de dimensions latentes qui structurent les données.
- Ils se fondent sur un calcul d'indices de distance géométrique entre les variables



III - les études qualitatives

3-3 L'analyse des données

L'ANALYSE FACTORIELLE DES CORRESPONDANCES

- L'AFC permet de révéler les associations existantes entre deux ensembles de variables.
- C'est un prolongement intéressant de l'analyse de contenu car elle se contente d'un tableau à double entrée (tableau de contingence) qui récapitule les croisements des différentes modalités de deux variables.

III - les études qualitatives

3-3 L'analyse des données

L'ANALYSE FACTORIELLE DES CORRESPONDANCES

Exemple : une analyse de contenu a donné une liste d'associations entre des caractéristiques d'un produit et des marques

Marques	Luxe	Ht Gamme	Gde marque	Ne sait pas
Lancel	6	13	16	0
Cartier	19	2	2	0
Lacoste	2	14	36	3
etc				

III - les études qualitatives

3-3 L'analyse des données

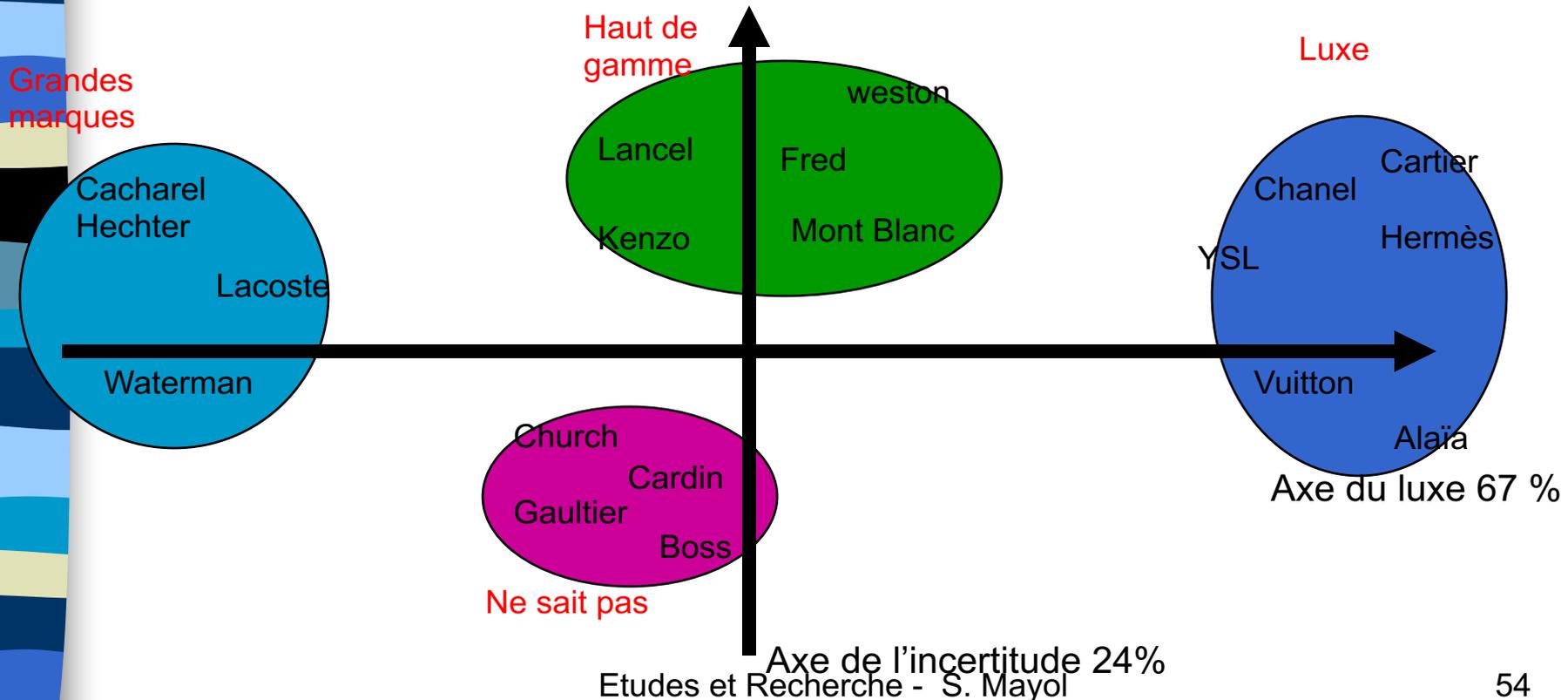
L'ANALYSE FACTORIELLE DES CORRESPONDANCES

- Lancel est vue comme une marque de luxe pour 6 citations, Cartier, pour 19
- L'analyse des différents couples de modalité des variables permet d'extraire par ordre d'importance décroissante les dimensions (axes du graphique) qui restituent le mieux l'information totale (« l'inertie ») contenue dans le tableau
- Le pourcentage d'inertie associé à chaque axe donne une indication sur la qualité d'ensemble de la représentation graphique

III - les études qualitatives

3-3 L'analyse des données

L'ANALYSE FACTORIELLE DES CORRESPONDANCES



III - les études qualitatives

3-3 L'analyse des données

L'ANALYSE FACTORIELLE DES CORRESPONDANCES

- L'axe horizontal résume 67% de l'information
- L'axe vertical ajoute 24% supplémentaires
- A eux seuls, ils expliquent 91% de l'inertie
- On nomme les axes à partir de l'analyse des contributions de chaque modalité, c'est à dire le poids (pourcentage) représenté par chaque modalité
- L'objectif de cette carte perceptuelle est d'identifier les attributs qui sous tendent l'image de sa marque et les concurrents qui sont les plus proches, donc dangereux
- Dans notre exemple, Boss, Cardin et Gaultier ont une image floue et brouillée et doivent se repositionner



III - les études qualitatives

3-4 L'analyse de contenu

Comment le phénomène étudié est défini? - les étapes

- Transcription des entretiens
- Définition de l'unité d'analyse
 - Objectif de l'étude
 - Permet l'interprétation du discours
 - Mots, phrase – expressions, thème
- Type d'analyse de contenu
 - Analyse lexicale : nature et richesse du vocabulaire
 - Analyse thématique : Découpage par thème



III - les études qualitatives

3-4 L'analyse de contenu

Comment le phénomène étudié est défini? - les étapes

- Construction d'une grille générale
 - Outil de dépouillement pour l'ensemble des interviews
 - Définir les thèmes a posteriori (Exclusifs, Pertinents et Homogènes)
- Codage des entretiens et remplissage la grille
- Analyse (fréquence d'utilisation, carte d'associations)

Catégories de thèmes	I1	I2	I3	Synthèse horizontale
Avant le vol ■ Réservation ■ Attente aéroport ■ Enregistrement ■ Embarquement				
Pendant le vol ■ Accueil ■ Service ■ Confort				
Après le vol ■ Exactitude ■ Livraison bagages ■ Transports ■ Services complémentaires				
Synthèse verticale				

→
Analyse horizontale

↓
Analyse verticale



III - les études qualitatives

3-5 Exemple : le rêve des français

- Cet exemple porte une étude réalisée par le Credoc en 1992 sur un échantillon représentatif de 995 français.
- L'analyse des réponses à la question "Si vous gagniez au loto, que feriez vous ?" donne une idée des rêves des français de cette époque.
- L'**analyse lexicale** permet de dégager les principaux thèmes évoqués

Mots et expressions les plus cités

On se fait une idée rapide du contenu des réponses en examinant les mots les plus cités. Le haut du lexique fait apparaître que les mots outils. Ils n'apportent rien sur le contenu mais sont très révélateurs de l'énonciation (importance du je par exemple).

Le lexique réduit ne laisse apparaître que les mots pleins. En lemmatisant on regroupe les différentes formes d'un même mot (enfant et enfants) et on réduit certaines ambiguïtés (voyage nom ou verbe ?).

Enfin la recherche des segments répétés achève la préparation du lexique.

Finalement la poursuite de l'analyse se fera sur un texte transformé par les étapes précédentes. On pourra enfin regrouper les mots de même signification directement ou en utilisant des dictionnaires de synonymes.

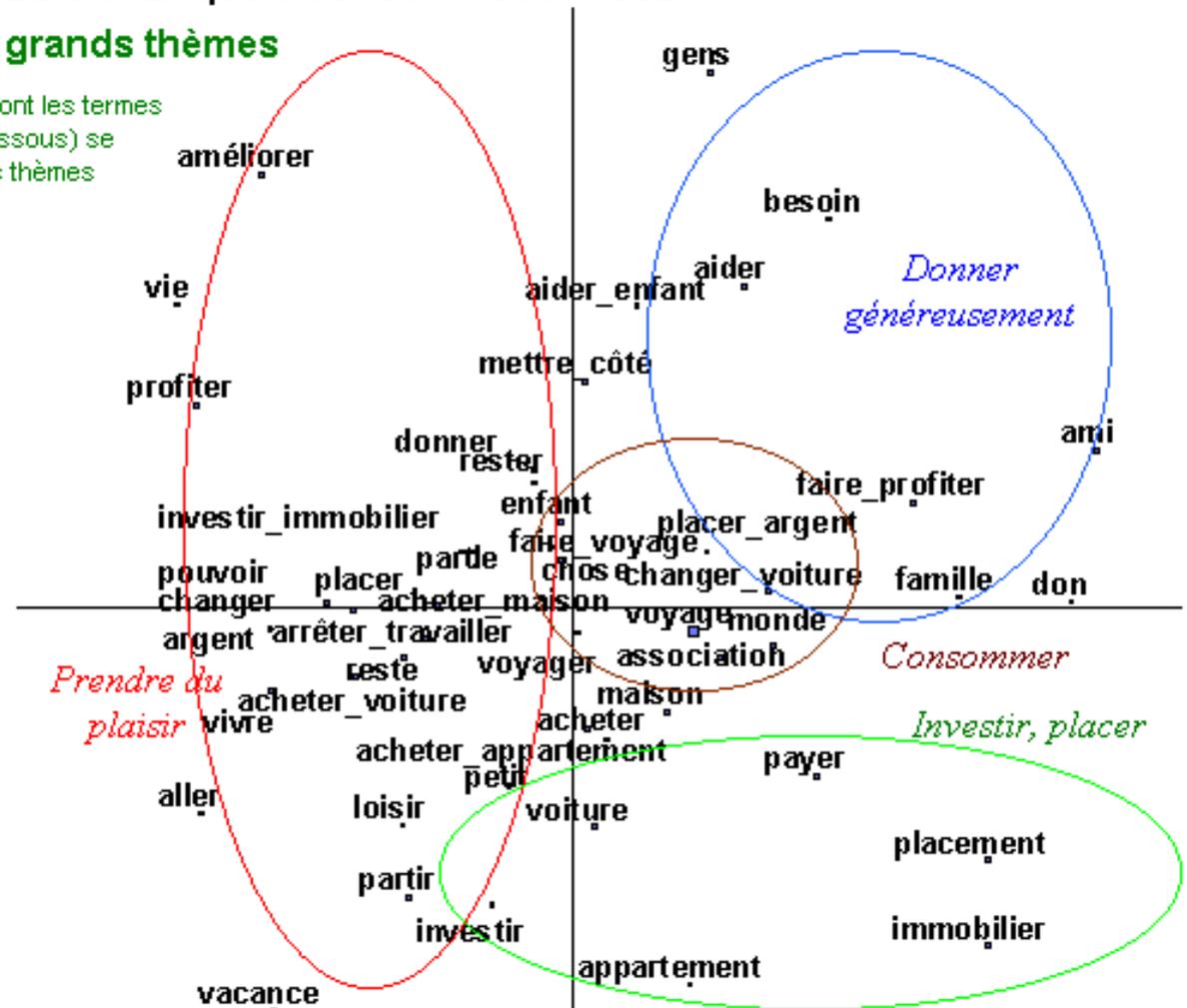
Tout ce travail se fait dans Lexica.

Lexique complet		Lexique réduit		Lexique lemmatisé		Segments répétés	
je	1182	maison	355	maison	363	acheter_maison	125
j	653	enfants	286	faire	357	placer_argent	52
de	613	voiture	198	acheter	350	faire_voyage	48
une	566	voyage	192	voyage	311	arrêter_travailler	46
en	464	ferais	182	enfant	288	donner_enfant	34
à	412	voyages	158	placer	237	investir_immobilier	33
le	385	achèterais	150	voiture	197	acheter_voiture	32
des	362	famille	122	donner	187	aider_enfant	31
un	362	argent	118	voyager	147	changer_voiture	29
maison	355	acheter	115	aider	122	partir_voyage	25
enfants	286	donnerais	113	famille	122	acheter_appartement	24
la	286	placerais	111	argent	117	mettre_de_côté	24
pour	280	faire	107	monde	93	partir_en_vacances	22
l	257	reste	97	travailler	91	tour_monde	21
mes	244	monde	95	partir	89	partager_enfant	19

Les associations lexicales : les grands thèmes

La carte ci contre (AFCM) visualise la manière dont les termes les plus fréquents du corpus (début de liste ci dessous) se trouvent associés. On peut en déduire les grands thèmes évoqués dans les réponses.

voyage	152	3%
voyager	125	2%
acheter_maison	125	2%
placer	90	1%
voiture	86	1%
maison	77	1%
placer_argent	52	<1%
acheter	49	<1%
faire_voyage	48	<1%
enfant	46	<1%
arrêter_travailler	46	<1%
placement	43	<1%
faire	40	<1%
donner	38	<1%
...	5018	83%



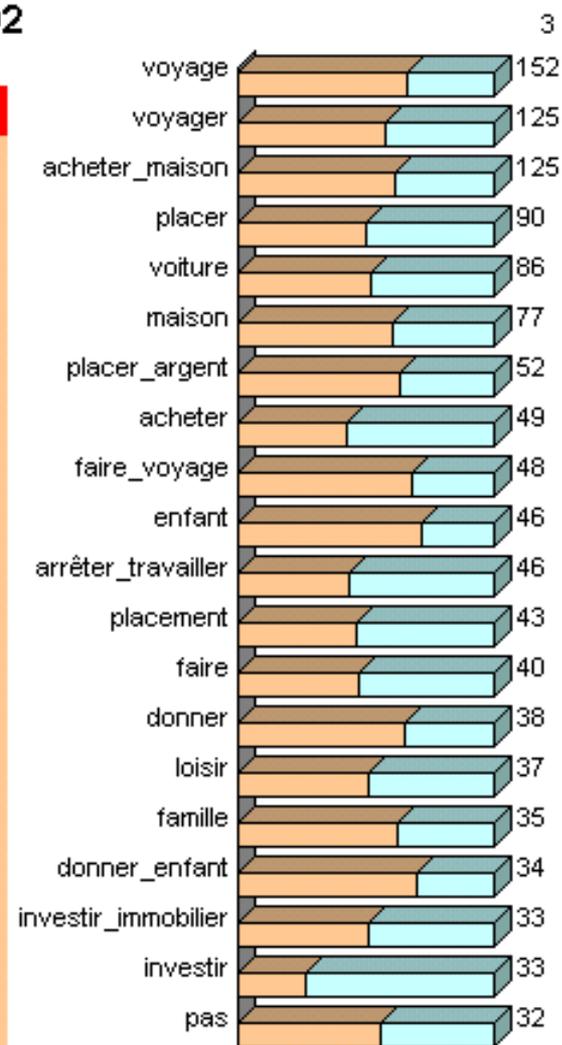
CREDOC: Enquête consommation 1992

Les effets du
contexte. Hommes
et femmes, des
rêves proches...

mais les femmes
pensent à leurs
enfants et sont plus
généreuses...

Les hommes		
voyager	53	2%
voyage	51	2%
acheter_maison	48	2%
placer	45	2%
voiture	41	1%
maison	30	1%
acheter	28	1%
arrêter_travailler	26	<1%
investir	24	<1%
placement	23	<1%
faire	21	<1%
placer_argent	19	<1%
loisir	18	<1%
investir_immobilier	16	<1%
être	16	<1%
aller	16	<1%
faire_voyage	15	<1%
...	2268	82%

Les femmes		
voyage	101	3%
acheter_maison	77	2%
voyager	72	2%
maison	47	1%
placer	45	1%
voiture	45	1%
placer_argent	33	1%
faire_voyage	33	1%
enfant	33	1%
donner	25	<1%
donner_enfant	24	<1%
aider_enfant	23	<1%
famille	22	<1%
acheter	21	<1%
mettre_de_côté	20	<1%
arrêter_travailler	20	<1%
acheter_voiture	20	<1%
...	2616	80%

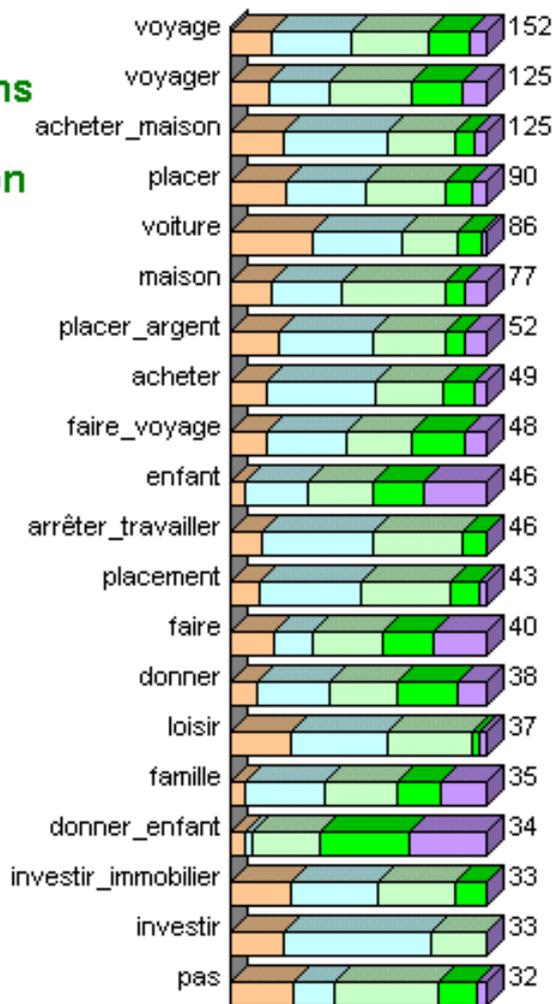


995 observations

Mots et expressions les plus citées selon l'âge

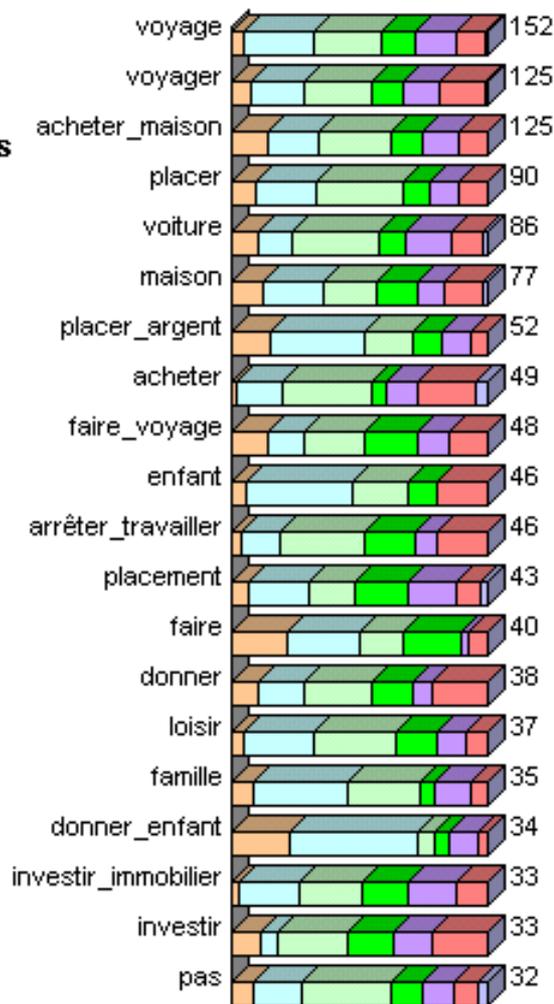
- 18-24
- 25-34
- 35-49
- 50-64
- 65+

CREDOC: Enquête consommation 1992



mots et expressions les plus citées selon le diplôme

- Aucun
- CEP BEPC
- CAP BEP
- Bac
- Bac+2 BTS
- Bac+4
- Refus
- Autres

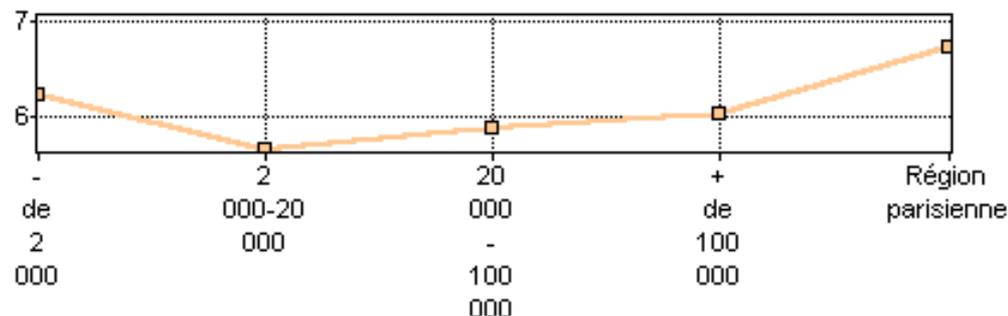


**L'énonciation :
influence de
l'identité du
répondant sur
la longueur de
sa réponse**

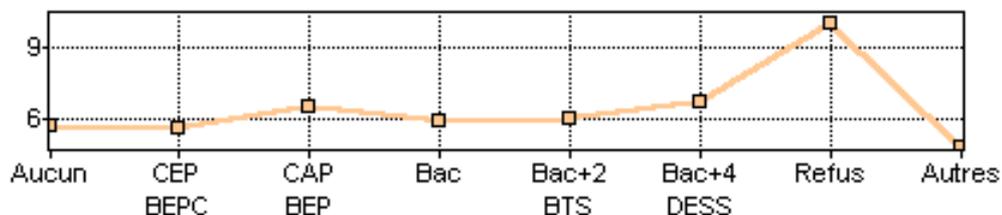
Longueur des
réponses en
nombre de
mots (sans
mots outils)

		MOTS ET SEGMENTS
HABITAT	Région parisienne	6,7
	- de 2 000	6,2
	+ de 100 000	6,0
	20 000 - 100 000	5,9
	2 000-20 000	5,6
SEXE	Homme	6,2
	Femme	6,0
AGE	18-24	6,3
	35-49	6,3
	25-34	6,0
	50-64	6,0
	65+	5,7
	DIPLOME	
	Bac+4 DESS	6,7
	CAP BEP	6,5
	Bac+2 BTS	6,1
	Bac	5,9
	Aucun	5,7
	CEP BEPC	5,6

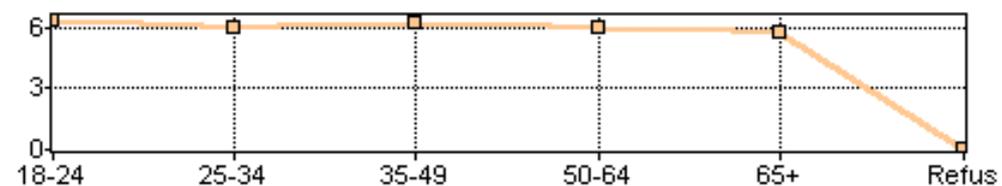
Longueur des réponses selon l'habitat



Longueur des réponses selon le diplôme



Longueur des réponses selon l'âge



IV - les études quantitatives

4-1 caractéristiques

- Un sondage est une enquête ponctuelle (réalisée une seule fois) qui consiste :
 - À extraire un échantillon au sein d'une population de base déterminée
 - À interroger les personnes faisant partie de cet échantillon avec des questionnaires pour extrapoler ensuite les réponses obtenues à l'ensemble de la population
- Un sondage mesure soit des proportions soit des valeurs
- Enquêtes « omnibus », baromètres ou « tracking »

IV - les études quantitatives

4-1 caractéristiques

- Le CREDOC a administré en face à face 1600 questionnaires portant sur les comportements alimentaires (stockage, préparation, consommation) pour 150 familles de produits. Cette enquête a permis de distinguer 7 groupes de consommateurs homogènes (segments)
- L'IFOP propose aux entreprises son étude omnibus *Papyfop* effectuée 2 fois par an auprès d'un échantillon



IV - les études quantitatives

4-1 caractéristiques

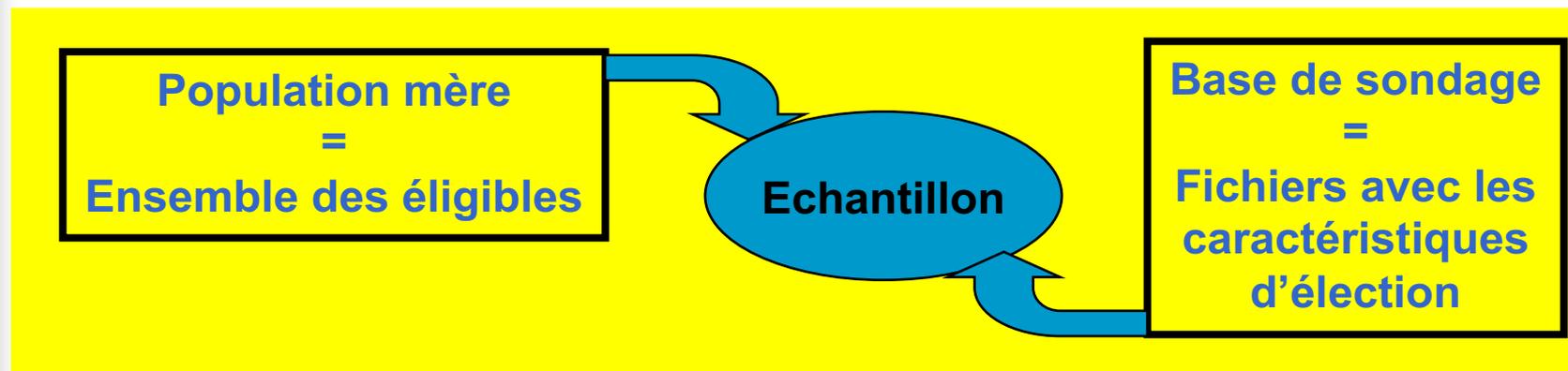
La réalisation d'une étude quanti commence par une réflexion d'ordre statistique (définition de la base de sondage, évaluation de la taille de l'échantillon, choix d'un mode de tirage), se poursuit par la conception et l'administration d'un questionnaire et s'achève par une phase de contrôle de la qualité du terrain

IV - les études quantitatives

4-2 détermination de la base de sondage

Délimiter la population mère

Elle découle de la population visée par l'étude. Elle regroupe l'ensemble des « éligibles », c'est-à-dire la totalité des personnes susceptibles d'être interviewées parce qu'elles possèdent la ou les caractéristiques correspondant aux objectifs de l'étude. Sauf cas particulier, la population mère ne peut pas être équivalente à la population française.



Population totale



IV - les études quantitatives

4-2 détermination de la base de sondage

■ Délimiter la population mère

Plusieurs obstacles surgissent:

- En l'absence d'orientation marketing claire, une infinité de population mère est possible (ex: tous les français de 7 à 77 ans), ce qui engendre des résultats trop généraux
- L'opérationnalisation des caractères statistiques d'éligibilité conduit souvent à exclure de façon parfois arbitraire certains individus de la population mère



IV - les études quantitatives

4-2 détermination de la base de sondage

Délimiter la population mère : exemple

Coca Cola étudie le marché potentiel français pour lancer un nouveau soft drink.

Aucune cible n'est définie

Population mère : ens. des personnes buvant ou susceptibles de boire un soft drink

En définissant les cibles potentielles, les contours de la population mère se précisent

IV - les études quantitatives

4-2 détermination de la base de sondage

Délimiter la population mère : exemple

Objectifs marketing	Cibles marketing potentielles=options de population mère	Opérationnalisation des caractères statistiques pour être éligible
Evaluer les ventes pour un nouveau produit lancé par Coca Cola en France	Cible 1: clients réguliers actuels de Coca Cola	Consommer du coca cola au moins une fois par semaine Résider en France
	Cible 2: nouveaux clients: jeunes, urbains et sportifs	Ne pas consommer de Coca Cola Age : 15 – 25 Habitat : ville de plus de 80 000 h
	Cible 3: nouveaux clients: femmes actives et aisées	Ne pas consommer de Coca Cola Sexe féminin, travail à plein temps Revenus annuel sup. à 30 K€



IV - les études quantitatives

4-2 détermination de la base de sondage

Concrétiser la base de sondage

- La population mère se matérialise par une base de sondage composée des fichiers à partir desquels on procède au repérage des éligibles
- L'échantillon est un groupe d'éligibles extrait de la base de sondage en utilisant une méthode de tirage
- Pb: la base de sondage ne correspond pas toujours à la population mère

Ex: fichiers France télécom (unité=foyer) et condition d'éligibilité=homme entre 20 et 60

IV - les études quantitatives

4-3 définition de la taille de l'échantillon

Eléments budgétaires

- Le budget contraint souvent la taille de l'échantillon
- La gestion du terrain (préparation et édition des questionnaires, achat de fichiers, tirage de l'échantillon, rémunération et déplacement des enquêteurs) représente plus de 50% du coût total d'une étude quanti
- Un budget de 15 000 € permet en moyenne d'interroger 500 à 1000 personnes
- Si la population est rare ou dispersée, le coût peut être doublé
- Si le rapport population mère de l'étude / population française est inférieur à 10%, on est en présence d'une population qualifiée de rare

IV - les études quantitatives

4-3 définition de la taille de l'échantillon

ASPECTS STATISTIQUES

- Seuls ces éléments devraient entrer en considération pour fixer la taille de l'échantillon.
- L'intérêt d'une étude quanti est de pouvoir généraliser les résultats observés dans l'échantillon à l'ensemble de la population mère
- Contraintes pour la composition de l'échantillon:
 - Tous les individus doivent être identifiables
 - Tous les individus doivent pouvoir être interrogés
- Estimation statistique de la précision obtenue
 - On détermine un intervalle de confiance pour chacune des variables considérées.
 - L'amplitude de cet intervalle donne la marge de fluctuation des résultats dans la population mère
 - Plus l'intervalle est étroit, plus le sondage est précis. La précision dépend directement de la taille de l'échantillon.
 - Si la taille de l'échantillon est grande, la précision est bonne et la généralisation des résultats se fait sans trop de risques.
 - La précision n'est pas influencée par la taille de la population mère

IV - les études quantitatives

4-3 définition de la taille de l'échantillon

■ Exemple

Sur un échantillon de 1000 personnes, on observe que 20% ont regardé l'émission de TF1, « star académie ». Comment généraliser ce taux d'audience sur la population mère

- Fixation d'un seuil de risque pour le calcul de l'intervalle de confiance.

Par tradition, on accepte un risque de $\alpha=0,05$. c'est-à-dire que dans 95% des cas ($1-\alpha$), le taux d'audience réel (celui qu'on obtiendrait si l'on interrogeait toute la population mère) tombe dans l'intervalle calculé

IV - les études quantitatives

4-3 définition de la taille de l'échantillon

■ Exemple

- Formule et calcul de résultat

$$p - z \sqrt{(pq/n)} \leq \pi \leq p + z \sqrt{(pq/n)}$$

p = audience observé dans l'échantillon

q = 1 - p

z = valeur dérivée de la loi normale centrée réduite, égale à 1,96 si $\alpha = 0,05$

π = audience réelle dans la population mère (inconnue)

n = taille de l'échantillon

$$0,20 - 1,96 \sqrt{((0,2 \times 0,8)/1000)} \leq \pi \leq 0,2 + 1,96 \sqrt{((0,2 \times 0,8)/1000)}$$

On obtient un intervalle de confiance égal à $17,5 \% \leq \pi \leq 22,5 \%$

La précision du sondage pour cette estimation de π est égale à $\pm 2,5$ points autour de la valeur observée dans l'échantillon (soit $20\% \pm 2,5\%$)

IV - les études quantitatives

4-3 définition de la taille de l'échantillon

- La formule donnée permet de calculer la taille de l'échantillon à interroger si l'on se fixe soi-même le degré de précision voulue, c'est-à-dire l'amplitude tolérée pour l'intervalle de confiance.

$$n = (z^2 pq) / e^2$$

Avec e = précision jugée suffisante pour généraliser les résultats de façon convenable = $(\pi - p)$

Si l'on admet qu'une précision de $\pm 3\%$ convient, quelle serait alors la taille de l'échantillon ? L'économie est importante car l'effectif à interroger se réduit à 683 personnes au lieu de 1000.

$$n = (1,96)^2(0,2)(0,8)/(0,03)^2 = 683$$

IV - les études quantitatives

4-4 extraction de l'échantillon

- Il existe différentes manières pour extraire un échantillon de la base de sondage

Méthodes aléatoires

L'échantillon est effectué à partir d'un tirage au hasard effectué sur un fichier nominatif de tous les individus appartenant à la base de sondage la validité de cette méthode suppose que tous les individus tirés au sort acceptent de répondre
aléatoire simple : il s'agit d'un tirage au sort simple des éligibles. Le tirage au sort est équiprobable

aléatoire stratifié : on définit des strates telles que les individus ne peuvent appartenir qu'à une seule strate. Au sein d'une même strate, chaque individu a la même probabilité d'être interrogé, mais au niveau global, la probabilité diffère si le taux de sondage diffère selon les strates

aléatoire en grappe : on réalise une partition de la base de sondage en combinant de nombreux critères. On constitue des grappes homogènes : chaque grappe devrait théoriquement être une « population mère miniaturisée ». On effectue un triage au sort d'une ou plusieurs grappes. Tous les individus des grappes sélectionnées sont interrogés.

IV - les études quantitatives

4-4 extraction de l'échantillon

■ Méthodes empiriques

Ces techniques reposent sur un choix raisonné des individus :

quotas : c'est le plus utilisé. On construit l'échantillon dans une population mère miniaturisée sur la base de critères jugés pertinents. Les personnes interrogées sont sélectionnées en fonction de la pertinence de leurs critères

Exemple : la pop. mère des consom. d'un pdt X est constituée de 50% d'hommes; parmi eux 30% sont des clients fidèles, 50% alternent entre deux marques et 20% sont infidèles. Si l'échantillon a une taille de 400 pers., il faudra interviewer 200 hommes dont 60 ($=400 \times 50\% \times 30\%$) acheteurs fidèles à une seule marque, 100 ($=400 \times 50\% \times 50\%$) clients fidèles à deux marques et 40 ($=400 \times 50\% \times 20\%$) acheteurs infidèles.

IV - les études quantitatives

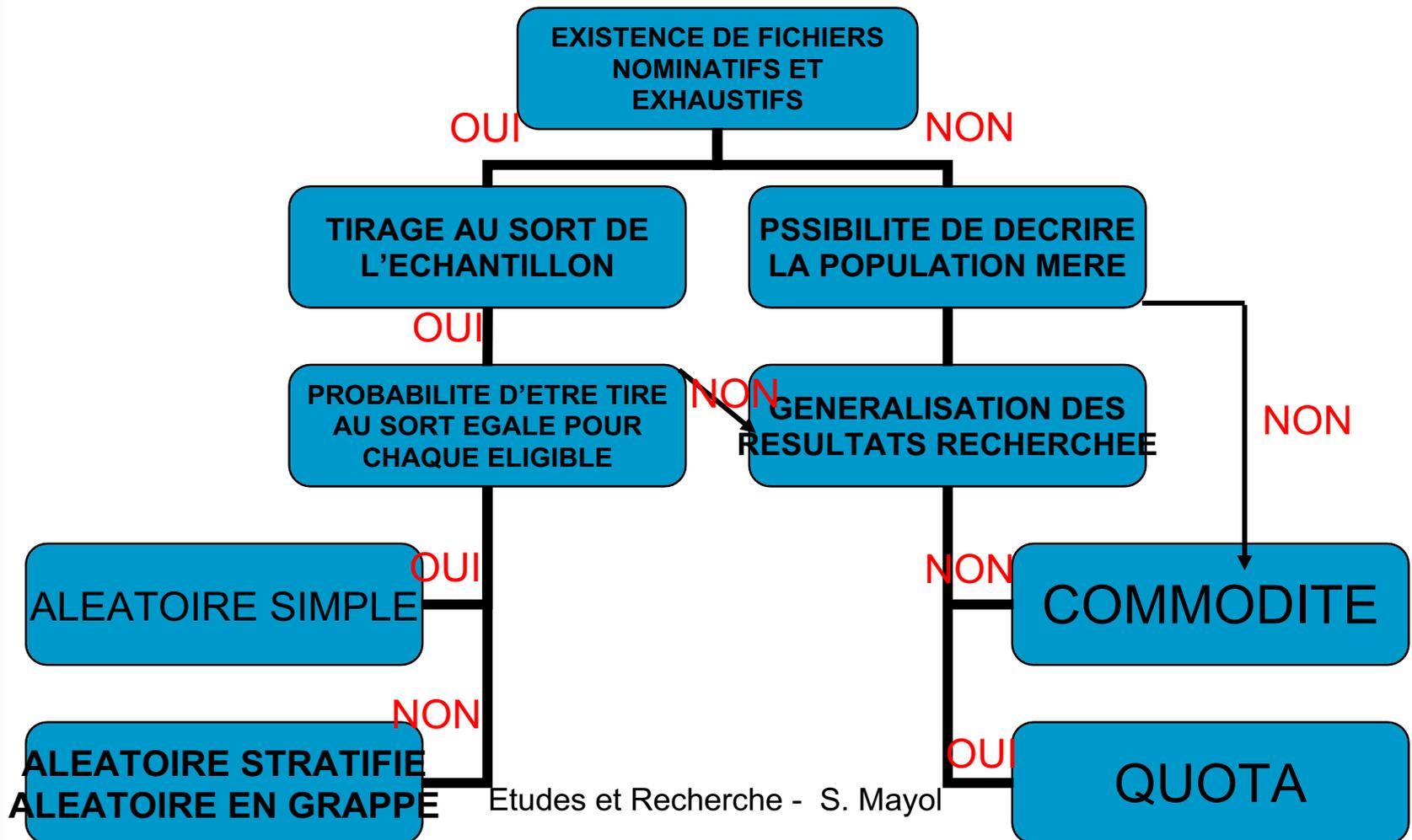
4-4 extraction de l'échantillon

commodité ou convenance : on laisse le hasard statistique agir supposant que la population mère se retrouvera dans l'échantillon. On teste cependant l'homogénéité des répondants et on prend un groupe test pour comparer les résultats.

Exemple : on teste une nouvelle formule de déodorants pour chiens auprès de deux cibles jugées, sur la base d'expériences antérieures, comme les plus réceptives. Un premier échantillon est composé de 50 maîtresses de caniches âgées de plus de 50 ans; un second échantillon regroupe 50 hommes ayant un labrador.

IV - les études quantitatives

4-4 extraction de l'échantillon



IV - les études quantitatives

4-5 construction du questionnaire

- Cette opération est délicate car toute erreur à ce niveau biaise les réponses et conduit à des interprétations erronées. Le questionnaire suit des principes simples à énoncer mais plus délicats à respecter:
Sur le fond, chaque question doit être comprise et acceptée par tous les répondants; les modalités d'enregistrement des réponses doivent être adaptées au recueil des informations détenues par le répondant.
Un questionnaire doit prendre en compte les trois grandes phases d'une interview:
- Phase de contact : présentation succincte de l'objet du questionnaire, remerciements de l'enquêté pour sa collaboration, enfin on pose des questions qualifiantes.

IV - les études quantitatives

4-5 construction du questionnaire

OBJECTIFS	EXEMPLES DE QUESTIONS
Présentation générale	
Présentation rapide de l'enquêteur et de la sté	« Bonjour, Caroline Martin, de l'institut »
Exposé (vague) du thème général de l'étude	« Nous réalisons aujourd'hui une étude sur les habitudes alimentaires des hommes »
Valorisation du répondant	« Nous serions très heureux d'entendre votre point de vue et remarques sur ce sujet »
Indication (approximative) de la durée du questionnaire	« Accepteriez vous de me consacrer 10 minutes pour répondre aux quelques questions suivantes »
Qualification des répondants	
Vérification de l'éligibilité (appartenance à la population mère)	Possédez-vous d'un congélateur chez vous ? Si non → STOP, remerciez
Catégorisation du répondant (uniquement si échantillon par quota)	Pour mieux vous situer, êtes vous... Prendre les critères retenus et vérifiez le remplissage

IV - les études quantitatives

4-5 construction du questionnaire

- Phase centrale : une ou deux questions de mise en route amène le répondant vers les thèmes centraux de l'étude. C'est une progression logique, en partant de questions générales pour aller vers des questions plus personnelles.
- Phase de clôture : les questions de situation permettent d'identifier le répondant sur la base des variables socio-démographiques usuelles (âge, sexe, profession, revenu etc...). Cette fiche signalétique sera très utile pour les traitements statistiques ultérieurs



IV - les études quantitatives

4-5 construction du questionnaire

Forme des questions

Une question peut être :

- Fermée** : les réponses sont prévues à l'avance et l'enquêteur coche les cases correspondantes. C'est la forme de question la plus adaptée à une enquête quanti.
- Ouverte** : le répondant est libre de formuler sa réponse avec ses propres mots. Cette approche convient peu à une analyse quanti car elle allonge la durée du questionnaire et complique l'analyse.

IV - les études quantitatives

4-5 construction du questionnaire

■ **Forme des questions**

LES DIFFERENTS TYPES D'ECHELLE DE MESURE D'UN QUESTIONNAIRE

DICHOTOMIQUE : choix d'une réponse possible parmi deux modalités

« considérez-vous votre téléviseur Sony comme un produit robuste »

oui

non

MULTICHOTOMIQUE : choix d'une ou plusieurs réponses parmi plusieurs modalités

« quels sont le ou les qualificatifs qui vous semblent le mieux convenir pour votre téléviseur Sony »

robuste

fragile

fiable

solide

RANGS : choix d'une classification des modalités présentées

« classez les marques suivantes de la plus robuste à la moins robuste »

Sony

Philips

Thomson

Toshiba

IV - les études quantitatives

4-5 construction du questionnaire

■ **Forme des questions**

LES DIFFERENTS TYPES D'ECHELLE DE MESURE D'UN QUESTIONNAIRE

SEMANTIQUE DIFFERENTIELLE : choix d'un échelon sur une échelle mettant en scène deux mots sémantiquement opposés

« où situez vous la marque Sony sur l'échelle suivante ? »

robuste fragile

SUPPORTS SEMANTIQUES : choix d'un échelon sur une échelle mettant en scène deux mots sémantiquement opposés complétés de mots intermédiaires

« comment évaluez vous la robustesse de votre téléviseur Sony ? »

très bonne bonne moyenne faible mauvaise très mauvaise

IV - les études quantitatives

4-5 construction du questionnaire

■ **Forme des questions**

LES DIFFERENTS TYPES D'ECHELLE DE MESURE D'UN QUESTIONNAIRE

LIKERT : choix du degré d'accord ou de désaccord avec une proposition

« indiquez votre degré d'accord ou de désaccord avec le jugement : mon téléviseur Sony est robuste »

Tout à fait d'accord plutôt d'accord plutôt pas d'accord pas du tout d'accord



ICONES : choix d'une réponse parmi plusieurs modalités présentées sous formes d'icônes ou de dessins

« comment jugez vous votre téléviseur Sony ? »

Robuste



fragile

IV - les études quantitatives

4-5 construction du questionnaire

■ Erreurs usuelles de formulation

Ambiguïté et imprécision de la question : c'est le cas lorsque :

- La réponse est trop imprécise (« regardez-vous souvent la télévision? »)
- plusieurs interprétations de la question sont possibles (« musique d'enfer »)
- lorsqu'on utilise des mots techniques ou sophistiqués (« quel est le positionnement de la Renault? »)
- lorsqu'on utilise des formes interro-négatives (« ne pensez vous pas que... »),
- lorsqu'on utilise des doubles négations (« êtes vous en désaccord avec le fait de ne pas voter? »)



IV - les études quantitatives

4-5 construction du questionnaire

■ Erreurs usuelles de formulation

■ Formulation orientée :

c'est le cas lorsqu'une question peut suggérer une réponse, par sa formulation

■ Nature du thème abordé :

certains sujets conduisent à des réponses biaisées parce qu'ils sont considérés comme des tabous : religion, argent, hygiène, santé, ,alcool, sexe, politique. L'interviewé donne une réponse socialement acceptable

IV - les études quantitatives

4-6 administration du questionnaire

■ Entretien en face à face

- On interroge généralement les gens à domicile, dans la rue ou à la sortie d'un magasin. La durée maximale du questionnaire doit être de 10-15 minutes
- Avantages : présence de l'enquêteur, interactivité
- Inconvénients : présence de l'enquêteur, durée de l'enquête, praticité, tabous sociaux

IV - les études quantitatives

4-6 administration du questionnaire

■ Courrier

- Envoyé par courrier (postal ou électronique) et complété directement par les répondants
- Avantages : choix du moment de réponse, durée de l'enquête, anonymat possible, taille du questionnaire, absence de l'enquêteur
- Inconvénients : absence de l'enquêteur, identité du répondant, questionnaires inutilisables, taux de réponse, date de retour.

IV - les études quantitatives

4-6 administration du questionnaire

■ Téléphone

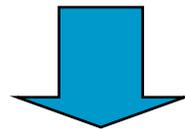
- Appel de l'enquêteur chez l'enquêté.
- Avantages : présence de l'enquêteur, interactivité
- Inconvénients : présence de l'enquêteur, durée de l'enquête, praticité, tabous sociaux, doute sur l'anonymat, intrusion chez les gens, choix du moment délicat

V - la distinction Qualitatif / Quantitatif

- Fonction du sujet
- Fonction des contraintes du terrain (coût mais aussi cibles)
- Complémentarité

Étude qualitative

Définir les dimensions du problème par des informations surtout verbales



Étude quantitative

quantifier les dimensions du problème par des informations surtout chiffrées

V - la distinction Qualitatif / Quantitatif

UNIVERS QUALITATIF



Basé sur le concept
d'écoute



« J'écoute »

(support = « oreille »)
technique individuelle,
groupe



**LA COMPREHENSION
PSHYCHO-SOCIO**

UNIVERS QUANTITATIF



Basé sur le concept *d'observation,
de déclaration*



« J'observe »

« Je pose des questions »
(support = grille d'observations,
questionnaire)
différentes techniques



**LA MESURE COMPREHENSIVE
du moment**

V - la distinction Qualitatif / Quantitatif

ETUDES QUALITATIVES



Objet : comprendre les individus un par un ou

les groupes d'individus structurés

pas de modèle, c'est l'émergence des structures de l'individu ou du groupe qui construit le modèle (le « pré supposé »)

ETUDES QUANTITATIVES



Objet : comprendre les individus en les groupant selon leur similarité observée

Pré supposé : existence d'un modèle, validation de ce modèle.

C'est une moyenne qui émerge ; les comportements originaux sont négligés

V - la distinction Qualitatif / Quantitatif

ETUDES QUALITATIVES



Individus : en fonction du problème, un (petit) nb d'individus ($10 < x < 50$) pour offrir *l'exhaustivité* des attitudes et comportements relatifs au phénomène



L'interview : basée sur un guide d'entretien/ d'animation

ETUDES QUANTITATIVES



Individus : un nombre choisi d'individus

représentativité statistique ou empirique (quota) sur un sujet donné



Interviews des individus grâce au *questionnaire*.

Le questionnaire est écrit en fonction de l'hypothèse :

« cette question se pose de façon pertinente ».

V - la distinction Qualitatif / Quantitatif

ETUDES QUALITATIVES



ANALYSE DES RESULTATS

Principes

- *Continuité* entre la prise d'informations (par le Chargé d'étude) et l'analyse (par le CE)
- *Liberté / guide* d'entretien, le CE peut adapter son guide au discours de l'interviewé.

ETUDES QUANTITATIVES



ANALYSE DES RESULTATS

Principes

- *Séparation* entre la prise d'infos (par l'enquêteur) et l'analyse (par CE)
- *Unanimité*, le questionnaire doit être compris de la même façon par tout le monde

V - la distinction Qualitatif / Quantitatif

Justification :

La méthode non directive

- Recherche d'un modèle que l'on ne connaît pas

L'adaptation du guide

- Hypothèse : les individus sont différents, le guide doit s'adapter à leur discours (leur vision des choses)

Justification de la directivité

- Validation d'un modèle déjà existant

L'adaptation du guide

- Hypothèse : les individus peuvent être agglomérés (tests statistiques)

V - la distinction Qualitatif / Quantitatif

ETUDES QUALITATIVES

2 conséquences



A. *Modéliser* le fonctionnement des individus

➤ Structurer des attitudes

B. *Interpréter*

➤ Poser les hypos sur les raisons qui sous-tendent les attitudes ou les comportements



Mettre en rapport *logique* les attitudes, les unes / aux autres

ETUDES QUANTITATIVES

3 conséquences



A. *Valider* des hypothèses émises en quali

B. *Modéliser* sur la base d'hypos connues et posées a priori

C. *Analyser* / à un référent (banque de données, chiffres déjà connus, ou comparer plusieurs objets entre eux).



Mettre en rapport *statistique* (par leur poids, leur importance) les variables, les unes / aux autres

V - la distinction Qualitatif / Quantitatif

ETUDES QUALITATIVES



- Privilégier les interactions entre les phénomènes
- Accepter la réalité comme complexe et difficilement déchiffrable



- On obtient des phénomènes marginaux
- Atteindre le maximum d'exhaustivité des attitudes et des comportements et savoir les expliquer par les motivations

ETUDES QUANTITATIVES



- Eliminer au maximum les interactions entre les phénomènes
- Accepter une certaine simplification de la réalité



- Elimination des phénomènes marginaux
- Atteindre l'exhaustivité des attitudes et des comportements par rapport à un réseau d'hypothèses



VI - analyse des résultats

6-1 dépouillement des résultats

- L'interprétation des résultats est effectuée avec rigueur et critique afin de prendre en compte les **biais** possibles :
 - Les achats réellement effectués sont tjs inférieurs aux intentions d'achat exprimées dans les sondages
 - Les résultats obtenus à des questions faisant appel à la mémoire sont souvent peu fiables
 - on répond plus facilement « oui » que « non »
 - Une lassitude de la personne interrogée peu entraîner des réponses peu fiables en fin de questionnaire



VI - analyse des résultats

6-1 dépouillement des résultats

- Terrain achevé
- Masse d'information brute qui doit être traitée de manière à répondre clairement aux questions posées lors de la définition des objectifs
- Plan :
 - Préparation de l'information
 - Construction des tableaux de distribution de fréquences
 - Calculs statistiques élémentaires



VI - analyse des résultats

6-1 dépouillement des résultats

- Regroupement dans un tableau des informations dispersées dans chacun des questionnaires
- Construction d'un tableau :
 - individus en ligne
 - Variables en colonnesÀ l'intersection d'une ligne et d'une colonne , on trouve la réponse d'un individu à une question

VI - analyse des résultats

6-1 dépouillement des résultats

- Matrice colonnes/lignes
- Possibilité de le faire avec un tableur (excel, par ex.)
- Exemple :

NUMINDIV	I1	I5	I16	I21	I52	I96	Etc.
0101	5	3	2	2	4	3	
0102							
...							
Etc.							



VI - analyse des résultats

6-1 dépouillement des résultats

- **Dépouillement terminé** : pile de questionnaires échangée contre un tableau contenant autant de lignes que de répondants et autant de colonnes que de questions
- Comment faire si enquête de 30 questions posée à 1000 personnes ?
 - Réorganiser l'information en construisant un **tableau de fréquences**



VI - analyse des résultats

6-2 l'analyse des données

■ La description des informations

- Les indices de base : ils permettent:
 - De situer le centre de gravité des observations
 - Moyenne
 - Médiane
 - D'informer sur la dispersion des informations par rapport au centre
 - Variance
 - Écart type



VI - analyse des résultats

6-2 l'analyse des données

■ Le tableau de fréquences

- Travail de regroupement ou de synthèse de données : tri ou tabulation
- Si une variable : tri à plat
- Si plusieurs variables : tris croisés

VI - analyse des résultats

6-2 l'analyse des données

■ Les tris à plat

- Il s'agit du calcul simple des fréquences relatives (f_i) de chaque variable.
- Chaque fréquence est égale au rapport du nombre d'individus (n_i) qui possèdent la modalité i de la variable X , sur le nombre d'individus étudiés (n)
- $F_i = n_i / n$
- Un tri à plat est toujours possible quelle que soit la nature de la variable.

VI - analyse des résultats

6-2 l'analyse des données

■ Les tris à plat : exemples

– Possédez vous un
téléphone portable ?

Q1	Oui	Non	Total (n)
ni	50	150	200
fi	(50/200) 25%	(150/200) 75%	(200/200) 100%

– Quel est le revenu
mensuel net moyen
de votre foyer?

Q2	- 2000 €	+ 2000 €	Total (n)
ni	110	90	200
fi	(110/200) 55%	(90/200) 45%	(200/200) 100%



VI - analyse des résultats

6-2 l'analyse des données

■ Les tris à plat

- Cette opération, encore appelée tabulation, consiste à réorganiser l'ensemble des valeurs prises par une variable
- Deux types de variables d'un point de vue statistique : qualitatives ou quantitatives (discrètes ou continues)



VI - analyse des résultats

6-2 l'analyse des données

■ Les tris à plat

■ 1. Variable qualitative

- Les différentes modalités ne sont pas directement mesurables par un nombre, mais représentent des « qualités » : nationalité, statut social
- Niveau de mesure ordinal ou nominal
- Exemple: hommes/femmes
- Chaque « bâton » correspond à une occurrence de la modalité ; calcul en proportion ou en pourcentage



VI - analyse des résultats

6-2 l'analyse des données

■ Les tris à plat

■ 1. Variable quantitative

- Directement mesurables par un nombre
- Taille, revenu, quantités consommées
- Discrète : seules valeurs possibles sont entières (nombre d'enfants, de salariés)
- Ou continues : valeurs possibles infinies au sein d'un intervalle donné (vitesse de la voiture, âge, poids...)



VI - analyse des résultats

6-2 l'analyse des données

- **Les tris à plat**

2.a. Variables quantitatives discrètes

- Pas de difficultés majeures
- Différence /qualitatives : nature des statistiques que l'on peut calculer



VI - analyse des résultats

6-2 l'analyse des données

■ Les tris à plat

2.b. Variables quantitatives continues

- En théorie, nombre infini de valeurs
- On crée des classes à l'intérieur desquelles les valeurs prises par la variable sont regroupées
- Attention à créer des classes pertinentes
- Exemple : série de CA sur échantillon de 200 entreprises

VI - analyse des résultats

6-2 l'analyse des données

■ Les tris croisés

- Permet d'analyser les relations entre 2, ou plus de 2, variables, considérées simultanément

A. Construction

- Construire un tableau à double entrée
« tableau de fréquences croisées »

Ou tableau de contingence

- Etapes de construction :
 - Comptage du nombre d'apparitions des couples de modalités
 - Vérification des totaux lignes et colonnes



VI - analyse des résultats

6-2 l'analyse des données

■ Les tris croisés

Ils combinent les réponses à deux questions

- Ils imposent la construction d'un tableau à double entrée (appelé tableau de contingence)
- Ce tableau ventile toutes les combinaisons des modalités de réponses prises deux à deux

VI - analyse des résultats

6-2 l'analyse des données

■ Les tris croisés

Questions	Q1 : possédez vous un téléphone portable		
	OUI	NON	TOTAUX
Q2 : revenus mensuels du foyer			
- 2000 €	25	85	110
+ 2000 €	25	65	90
TOTAUX	50	150	200

VI - analyse des résultats

6-2 l'analyse des données

■ Les tris croisés

Questions	Q1 : possédez vous un téléphone portable		
	OUI	NON	TOTAUX
Q2 : revenus mensuels du foyer			
- 2000 €	12,5%	42,5%	55%
+ 2000 €	12,5%	32,5%	45%
TOTAUX	25%	75%	100%

Tris croisés avec l'effectif total : 12,5% des répondants gagnent plus de 2000 € par mois et possèdent un mobile

VI - analyse des résultats

6-2 l'analyse des données

■ Les tris croisés

Questions	Q1 : possédez vous un téléphone portable		
Q2 : revenus mensuels du foyer	OUI	NON	TOTAUX
- 2000 €	27,8%	72,2%	100%
+ 2000 €	22,7%	77,3	100%
TOTAUX			

Tris croisés en ligne : 27,8% des répondants gagnant plus de 2000 € par mois possèdent un mobile

VI - analyse des résultats

6-2 l'analyse des données

■ Les tris croisés

Questions	Q1 : possédez vous un téléphone portable		
Q2 : revenus mensuels du foyer	OUI	NON	TOTAUX
- 2000 €	50%	56,7%	
+ 2000 €	50%	43,3%	
TOTAUX	100%	100%	

Tris croisés en colonne : 43,3% des répondants qui possèdent un mobile gagnent plus de 2000 € par mois

2ème partie : **L'analyse approfondie des** **données**

S. MAYOL



UTILISATIONS

La statistique est l'ensemble des méthodes scientifiques à partir desquelles sont recueillies, présentées, résumées et analysées les données.

Deux niveaux:

Stat. Descriptive sert à décrire des ensembles nombreux et dégager l'essentiel de l'information qui en résulte.

Stat. Théorique permet d'énoncer et d'élaborer des lois.

I - TABLEAUX STATISTIQUES

■ 1) CAS D'UNE SEULE VARIABLE

■ Le tableau brut se présente sous la forme suivante:

Individu	variable
1	x_1
2	x_2
.	.
.	.
.	.
n	x_n

■ Le nombre d'individus observé étant en général important, le tableau précédent ne permet pas d'analyser l'information obtenue. Il est donc nécessaire de créer un tableau plus synthétique où les observations identiques (possédant la même modalité) ont été regroupées.

modalité	effectif
C_1	n_1
C_2	n_2
.	.
.	.
.	.
C_k	n_k

- 
- Pour une variable **qualitative**, les modalités sont susceptibles d'être classées en groupes d'individus.
 - Pour une variable **quantitative**, les modalités ne sont pas classées
 - Exemple: le sexe est une variable qualitative alors que l'âge ne l'est pas. Néanmoins toutes les variables peuvent être transformées en variable qualitative, par regroupement. Ainsi, le variable âge peut permettre une classification des individus en classes d'âge.

2) CAS DE DEUX VARIABLES

Individu	variable1	variable2
1	x_1	y_1
2	x_2	y_2
.	.	.
.	.	.
.	.	.
n	x_n	y_n

Le tableau brut se présente sous la forme suivante

On désire créer un tableau appelé tableau de contingence donnant le nombre d'individus possédant simultanément la modalité i de variable1 et la modalité j de variable2 qui se présentera sous la forme suivante:

		Variable2		
		D_1	D_j	D_r
variable1	C_1	n_{11}	n_{1j}	n_{1r}

	C_i	n_{i1}	n_{ij}	n_{ir}

	C_k	n_{k1}	n_{kj}	n_{kr}

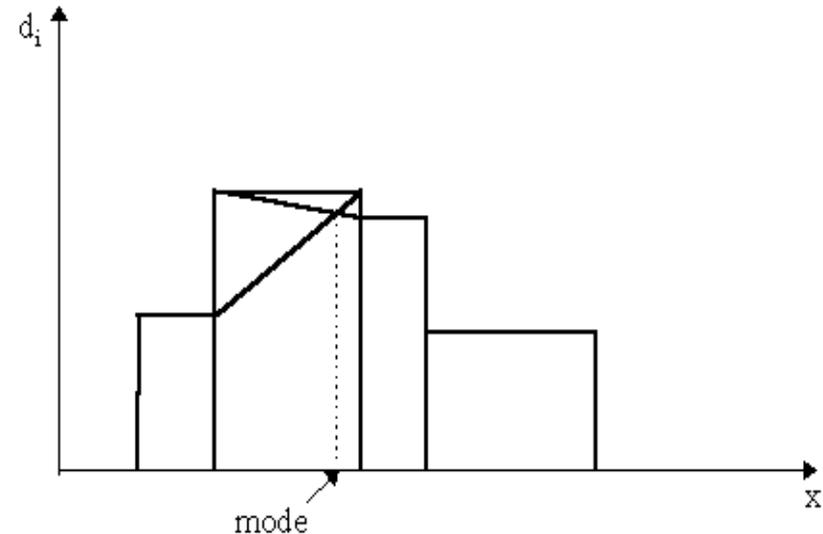
II - TENDANCE CENTRALE

1) MODE

C'est la valeur observée d'effectif maximum.

Variable discrète: classer les données par ordre croissant. Celle d'effectif maximum donne le mode.

Il est fortement conseillé d'utiliser le diagramme en bâtons pour déterminer le mode. En effet, deux valeurs consécutives x_i , x_{i+1} peuvent avoir le même effectif maximum; on parlera d'intervalle modal $[x_i, x_{i+1}]$. Il peut aussi y avoir un mélange de deux populations qui conduit à un diagramme en bâtons où apparaissent deux bosses; on considérera deux modes. Il est déconseillé, sauf raison explicite, d'envisager plus de deux modes.

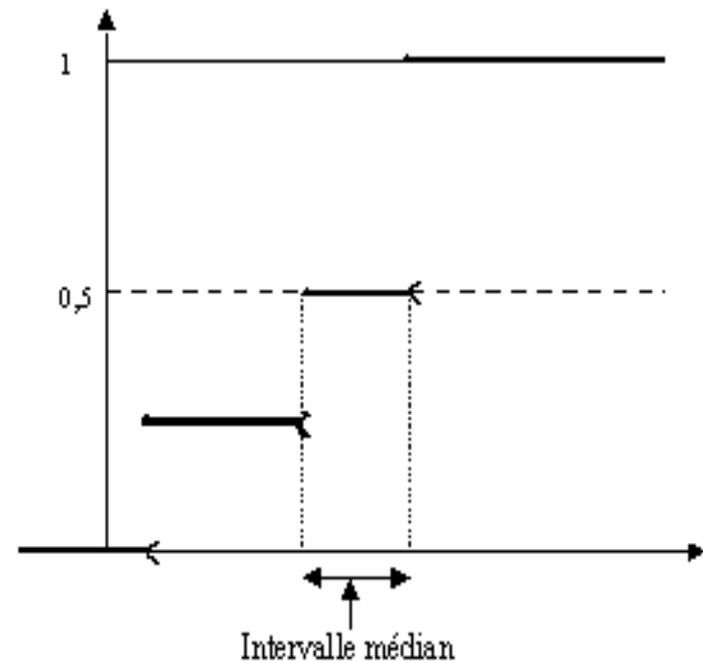
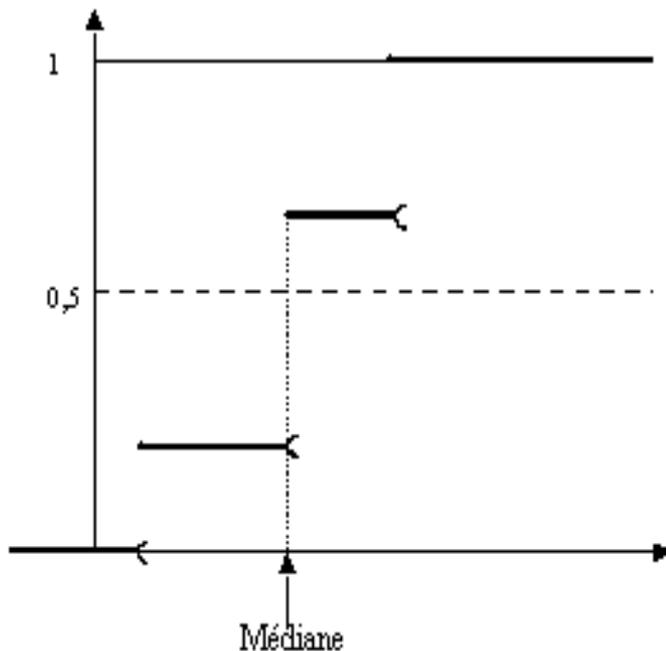


Variable classée: la classe modale correspond à la classe ayant l'effectif maximum. Il est fortement conseillé d'utiliser l'histogramme pour déterminer le mode. Comme pour le cas discret, on peut avoir deux classes modales. Toutes les valeurs de la classe pouvant à priori se réaliser, on ne se contentera pas de déterminer la classe modale. Une des valeurs de cette classe sera le mode. Certains auteurs préconisent par simplicité de prendre le centre de la classe modale. Il est préférable cependant de tenir compte des classes adjacentes de la manière suivante:

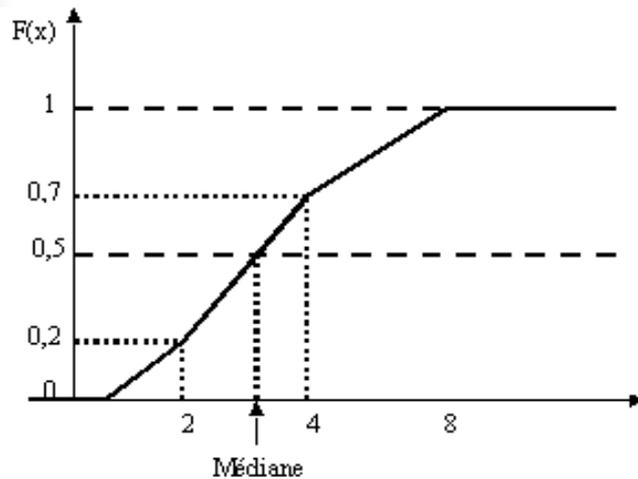
2) MEDIANE

Les valeurs étant rangées par ordre croissant, c'est la valeur de la variable qui sépare les observations en deux groupes d'effectifs égaux.

Variable discrète: la détermination peut s'obtenir à partir du tableau statistique en recherchant la valeur de la variable correspondant à une fonction cumulée égale à $n/2$ (effectif cumulé) ou $1/2$ (fréquence cumulée). Il est encore plus facile de lire sur les graphiques cumulatifs les abscisses des points d'ordonnée $n/2$ (effectif cumulé) ou $1/2$ (fréquence cumulée). Si tout un intervalle a pour image $n/2$ ($1/2$ pour la fréquence), on parlera d'intervalle médian (on peut prendre le milieu de l'intervalle comme médiane)



Variable classée: l'abscisse du point d'ordonnée $n/2$ ($1/2$ pour la fréquence) se situe en général à l'intérieur d'une classe. Pour obtenir une valeur plus précise de la médiane, on procède à une interpolation linéaire. La valeur de la médiane peut être lue sur le graphique ou calculée analytiquement.



d'où la valeur de la médiane.

De manière générale, si a et b sont les bornes de la classe contenant la médiane, $F(a)$ et $F(b)$ les valeurs de la fréquence cumulée croissante en a et b , alors

$$\frac{Mé - 2}{4 - 2} = \frac{0,5 - 0,2}{0,7 - 0,2}$$

$$Mé = a + (b - a) \times \frac{0,5 - F(a)}{F(b) - F(a)}$$

■ 3) MOYENNE ARITHMETIQUE

- Si x_i sont les observations d'une variable discrète ou les centres de classe d'une variable classée,

la moyenne arithmétique \bar{x} est égale à $\sum_{i=1}^k \frac{n_i x_i}{n} = \sum_{i=1}^k f_i x_i$

La moyenne arithmétique est un paramètre de tendance centrale plus utilisé que les autres de par ses propriétés algébriques:

- a) Pour plusieurs populations d'effectifs n_1, n_2, \dots, n_k , de moyennes respectives
moyenne globale = moyenne des moyennes

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k n_i \bar{x}_i}{n}$$

- b) **La moyenne arithmétique conserve les changements d'échelle et d'origine**

$$x: (x_i, n_i) \rightarrow y: (y_i = ax_i + b, n_i)$$

$$\bar{x} \rightarrow \bar{y} = a\bar{x} + b$$



■ 4) QUANTILES

Ce sont des caractéristiques de position.

Il y a 1 médiane $Mé$ qui sépare les observations en 2 groupes d'effectifs égaux

3 quartiles $Q1, Q2, Q3$ qui séparent les observations en 4 groupes d'effectifs égaux

9 déciles $D1, D2, \dots, D9$ qui séparent les observations en 10 groupes d'effectifs égaux

99 centiles $C1, C2, \dots, C99$ qui séparent les observations en 100 groupes d'effectifs égaux

La détermination de ces caractéristiques est identique à celle de la médiane.

Les quartiles sont obtenus lorsqu'on a cumulé 25, 50, 75% de la population

Les déciles sont obtenus lorsqu'on a cumulé 10, 20, ..., 90% de la population

Les centiles sont obtenus lorsqu'on a cumulé 1, 2, ..., 99% de la population

Remarque: la notion de déciles et de centiles n'a de sens que s'il y a beaucoup d'observations et donc essentiellement pour une variable classée.



III - DISPERSION

Comme leur nom l'indique, ces caractéristiques essaient de synthétiser par une seule valeur numérique la dispersion de toutes les valeurs observées.

1) ÉTENDUE

C'est la différence entre la plus grande et la plus petite observation

2) INTERVALLE INTER-QUARTILE

C'est la différence entre le troisième et le premier quartile

3) VARIANCE ET ECART-TYPE

Si x_i sont les observations d'une variable discrète ou les centres de classe d'une variable classée, la variance

$$V \text{ est égale à } \sum_{i=1}^k \frac{n_i (x_i - \bar{x})^2}{n} = \sum_{i=1}^k f_i (x_i - \bar{x})^2$$

$$\text{On a aussi } V = \sum_{i=1}^k \frac{n_i x_i^2}{n} - \bar{x}^2$$

c.à.d. moyenne des carrés - carré de la moyenne

- On utilise plus couramment l'écart-type qui est la racine carrée de la variance et qui a l'avantage d'être un nombre de même dimension que les données (contrairement à la variance qui en est le carré)
- La variance est un paramètre de dispersion plus utilisé que les autres de par ses propriétés algébriques:

- a) Pour plusieurs populations d'effectifs n_1, n_2, \dots, n_k , de moyennes respectives $\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k$, de variances V_1, V_2, \dots, V_k

Variance globale = variance des moyennes + moyenne des variances

$$V = \sum_{i=1}^k \frac{n_i}{n} (\bar{x}_i - \bar{\bar{x}})^2 + \sum_{i=1}^k \frac{n_i}{n} V_i$$

où $\bar{\bar{x}}$ représente la moyenne des moyennes

- b) changement d'échelle et d'origine

$$x: (x_i, n_i) \rightarrow y: (y_i = ax_i + b, n_i)$$

$$V_x \rightarrow V_y = a^2 V_x$$



- 4) COEFFICIENT DE VARIATION

$$CV = \frac{\sigma}{\bar{x}}$$

C'est un coefficient qui permet de relativiser l'écart-type en fonction de la taille des valeurs.
Il permet ainsi de comparer la dispersion de séries de mesures exprimées dans des unités différentes

IV - LA CORRELATION

- Lorsqu'on observe deux variables quantitatives sur les mêmes individus, on peut s'intéresser à une liaison éventuelle entre ces deux variables.
- Le coefficient de corrélation r permet de mesurer un lien entre deux variables continues X et Y , sans que celui-ci ne soit forcément un lien de causalité. Il est calculé avec la formule.

$$r_{XY} = \frac{\text{COV}(X;Y)}{\sqrt{(\text{var}(X) \text{ var}(Y))}}$$

Où la variance de X , $\text{var}(X)$, est égale à la moyenne de X moins la moyenne de X au carré, et la covariance de X et Y , $\text{cov}(X, Y)$ est égale à la moyenne de XY moins le produit de la moyenne de X et de la moyenne de Y .



L'interprétation de la valeur de r est la suivante

r proche de 0 = pas de lien entre X et Y

r proche de 1 = lien fort et dans le même sens

r proche de -1 = lien fort en sens contraire

Exemple

On mesure sur 7 enfants leur âge et le nombre d'heures passées par jour devant la télévision. Observe t-on un lien ?

X = AGE	5	6	7	13	15	18	10
Y = TEMPS	2	2	2	4	3	3	2

Les calculs sont les suivants

X	5	6	7	13	15	18	10
Y	2	2	2	4	3	3	2
X ²	25	36	49	169	225	324	100
Y ²	4	4	4	16	9	9	4
XY	10	12	14	52	45	54	20

Les moyennes sont

X	10,57
Y	2,57
X ²	132,57
Y ²	7,14
XY	29,57



D'où

$$\text{VAR}(X) = 132,57 - 10,57^2 = 20,85$$

$$\text{VAR}(Y) = 7,14 - 2,57^2 = 0,54$$

$$\text{COV}(X, Y) = 29,57 - 10,57 \times 2,57 = 2,4$$

$$r = 0,72$$

La corrélation est donc assez forte, et le temps passé devant la télévision semble augmenter avec l'âge. Cependant, il faut savoir si ce résultat est significatif, c'est à dire si ce coefficient de corrélation est significativement différent de 0. Le test z de Fisher permet de le savoir.

Dans ce test, on cherche à savoir si le coefficient de corrélation est différent ou non d'une certaine valeur $r_0 = 0$

Pour effectuer ce test, on calcule la quantité $(Z-Z_0)\sqrt{(n-3)}$ où n est le nombre d'individus et

$$Z_0 = 0,5 \ln \left(\frac{1+r_0}{1-r_0} \right)$$

et
$$Z_0 = 0,5 \ln \left(\frac{1+r}{1-r} \right)$$

Si cette quantité est supérieure à 2, les deux coefficients de corrélation r_0 et r sont significativement différents l'un de l'autre.

Ici $Z = 1,8$

Le coefficient de corrélation 0,72 n'est donc pas significativement différent de 0, ce qui signifie que le lien entre l'âge et le temps passé devant la télévision n'est pas significatif sur cet échantillon. La raison à cela est bien sûr la taille beaucoup trop faible de l'échantillon

V - LA REGRESSION

Faire une régression linéaire entre une variable à expliquer et des Variables explicatives X_1, X_2, X_3, \dots , c'est trouver la meilleure équation du type

$$Y = a_1X_1 + a_2X_2 + a_3X_3 + \dots + a_nX_n + b$$

Le t de Student permet de savoir si cette variable est significative.

Elle l'est lorsque la valeur absolue de t est supérieure à 1,96.

VI - LE KHI DEUX

Le test du khi 2, noté χ^2 , permet de savoir si un lien observé entre deux variables est significatif. Pour utiliser ce test, les variables doivent être qualitatives.

Testons par exemple le lien entre le genre des consommateurs et la possession d'un vêtement de marque Dyna, spécialisée dans le sport.

	Possèdent un Dyna	Ne possèdent pas de Dyna	Total
Hommes	10	40	50
Femmes	20	30	50
Total	30	70	100

Ce tableau indique, par exemple, que 20 femmes possèdent un Dyna

La proportion de femmes qui possèdent un Dyna est donc de 40%.

Chez les hommes, elle est de 20%.

Les femmes semblent donc plus susceptibles d'acheter cette marque.

Toutefois cette différence observée entre hommes et femmes pourrait uniquement résulter de fluctuations engendrées par un échantillon trop petit.

Il faut donc vérifier que cette différence n'est pas due à l'échantillonnage.

$$\chi^2 = n \left\{ \left[\frac{\sum_{\text{cases}} \text{case2}}{\sum \text{ligne} \times \sum \text{colonne}} \right] - 1 \right\}$$



Où

« case » = effectif de chacune des 4 cases du tableau

« ligne » et « colonne » = ligne et colonne du tableau où
se trouve cette case

n = nombre total d'individus

$$\chi^2 = 100 \left\{ \left[\frac{40^2}{50 \times 70} + \frac{30^2}{50 \times 70} + \frac{10^2}{50 \times 30} + \frac{20^2}{50 \times 30} \right] - 1 \right\}$$

$$\chi^2 = 4,76$$



Le lien entre deux variables est statistiquement significatif quand le χ^2 calculé est supérieur au χ^2 critique, qui dépend de la taille du tableau donné.

Quand les deux variables ont chacune deux modalités, comme c'est le cas dans cet exemple, le degré de liberté v est égal à (nombre de lignes - 1) X (nombre de colonnes - 1)

$$\text{Ici } (2-1) \times (2-1) = 1$$

Avec un $v=1$, on obtient un χ^2 critique = 3,84

$$\chi^2 \text{ observé} = 4,76 > \chi^2 \text{ critique} = 3,84$$

Le lien observé ici est donc significatif malgré la petite taille de l'échantillon.



Le test du χ^2 permet également de vérifier s'il y a accord entre des effectifs issus de la réalité et des effectifs supposés à partir d'une théorie.

Dans ce cas on calcul les effectifs théorique de chaque phénomène observé.

Le calcul du χ^2 s'effectue selon la formule

$$\chi^2 \text{ cal} = \sum \frac{(\text{effectif théorique} - \text{effectif observé})^2}{\text{Effectif théorique}}$$

Exemple : les données suivantes concernent la consommation d'un jus de fruit auprès d'un échantillon de 220 femmes.

	0-1 enfant	2 enfants	+2 ans	Total	%
Oui , régulièrement	5	30	8	43	19,5
Oui, occasionnellement	20	50	45	115	52,3
Non	12	40	10	62	28,2
Total	37	120	63	220	100

	0-1 enfant	2 enfants	+2 ans	Total
Oui , régulièrement	8*	23	12	43
Oui, occasionnellement	19	63*	33	115
Non	19	34	18*	62
Total	37	120	63	220

$$* 37 \times 19,5\% = 8$$

$$* 120 \times 52,3\% = 63$$

$$* 63 \times 28,2\% = 18$$


$$\chi^2 = \frac{(8-5)^2}{8} + \frac{(23-30)^2}{23} + \frac{(12-8)^2}{12} + \frac{(19-20)^2}{19} + \frac{(63-50)^2}{63} + \frac{(33-45)^2}{33} +$$
$$\frac{(10-12)^2}{10} + \frac{(34-40)^2}{34} + \frac{(18-10)^2}{18} = 14,79$$

Le nombre de degrés de libertés ici est égal à 4.

La lecture de la table du de Pearson indique que la probabilité d'indépendance est inférieure à 1%. Le nombre d'enfants explique

bien la propension à acheter des jus de fruits.



VII - L'ANALYSE EN COMPOSANTE PRINCIPALE (ACP)

Cette méthode réalise un exercice de sémantique en identifiant des groupes de variables associées à des réponses similaires.

Elle permet une radioscopie des données.

Elle n'opère que sur des variables quantitatives (échelles ou numériques)



L'analyse en composantes principales est une technique de statistique descriptive qui calcule les axes principaux du nuage des observations regroupées selon les modalités de la variable à évaluer.

Les cartes factorielles sont des représentations graphiques issues de cette analyse. Ce sont des graphiques en deux dimensions, chacune représentant un des axes.

Les critères d'évaluation et les modalités de la variable à évaluer sont positionnées sur ce graphique. L'interprétation de leur position relative donnera des indications sur les rapport qu'entretiennent les variables.



Le pourcentage de variance qui caractérise chacun des axes est indiqué au côté de chacun des axes. La fonction Choisir les axes permet de modifier les composantes de la carte.

L'analyse en composantes principales suppose le calcul préalable des coefficients de corrélations entre deux couples de critères. On peut également présenter ces coefficients sous la forme de la Matrice des corrélations ou du Diagramme des corrélations (graphique dans lequel les variables corrélées sont reliées par un trait dont l'épaisseur est fonction du coefficient de corrélation).

Analyse en composantes principales

Variables : VITESSE, CONFORT, SECURITE.

VITESSE	Nb. cit.	Fréq.
Non réponse	25	12,5%
Pas du tout	7	3,5%
Plutôt non	17	8,5%
Moyennement	61	30,5%
Plutôt oui	45	22,5%
Tout à fait	45	22,5%
TOTAL OBS.	200	100%

Moyenne = 3,59 Ecart-type = 1,09

SECURITE	Nb. cit.	Fréq.
Non réponse	25	12,5%
Pas du tout	4	2,0%
Plutôt non	14	7,0%
Moyennement	69	34,5%
Plutôt oui	78	39,0%
Tout à fait	10	5,0%
TOTAL OBS.	200	100%

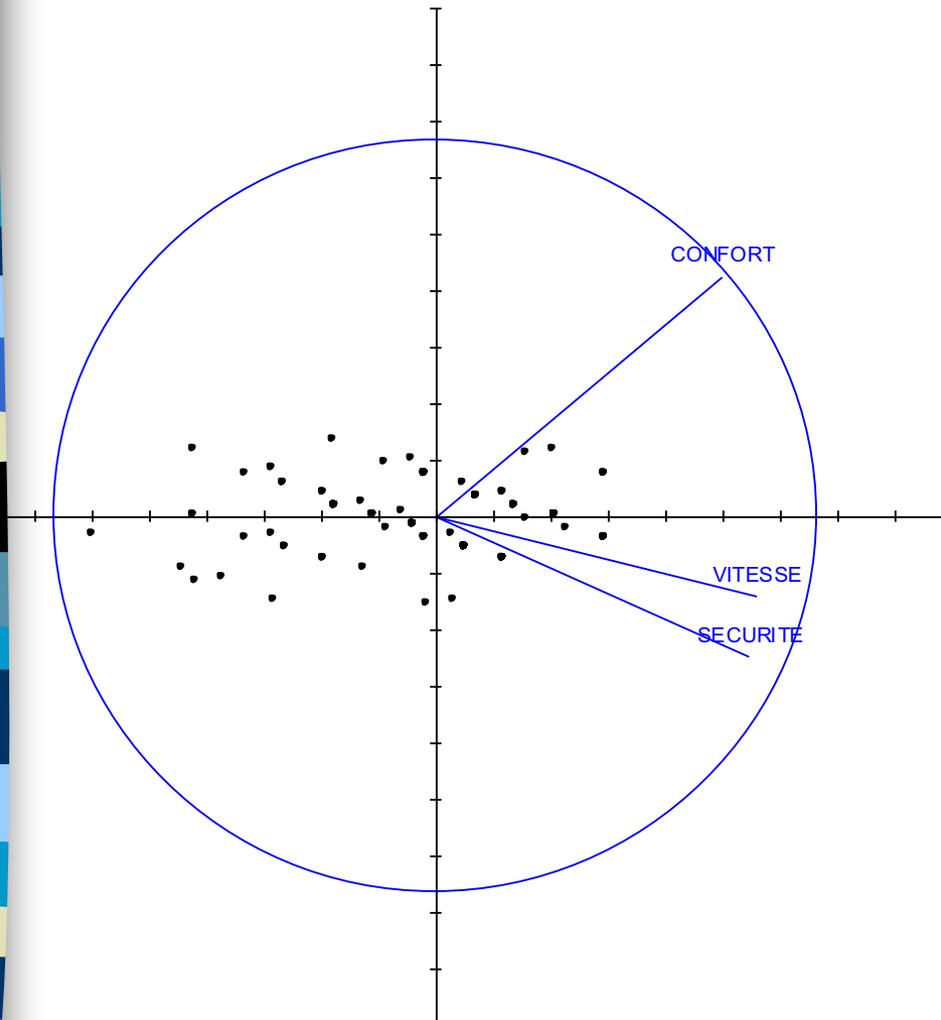
Moyenne = 3,43 Ecart-type = 0,81

CONFORT	Nb. cit.	Fréq.
Non réponse	25	12,5%
Pas du tout	5	2,5%
Plutôt non	11	5,5%
Moyennement	85	42,5%
Plutôt oui	69	34,5%
Tout à fait	5	2,5%
TOTAL OBS.	200	100%

Moyenne = 3,33 Ecart-type = 0,76

Analyse en composantes principales

Variables : VITESSE, CONFORT, SECURITE.



La carte montre les positions des 3 critères et les coordonnées des 175 observations.

25 observations ne sont pas prises en compte (non-réponse à au moins un des critères).

86.2% de la variance est expliquée par les deux axes représentés.

Chaque observation est représentée par un point.

Analyse en composantes principales

Variables :VITESSE, CONFORT, SECURITE.

	VITESSE	CONFORT	SECURITE
VITESSE	1,00		
CONFORT	0,47	1,00	
SECURITE	0,58	0,42	1,00

La carte montre les positions des 3 critères et les coordonnées des 175 observations.

25 observations ne sont pas prises en compte (non-réponse à au moins un des critères).

86.2% de la variance est expliquée par les deux axes représentés.

Chaque observation est représentée par un point.

Analyse en composantes principales

Variables : VITESSE, CONFORT, SECURITE.

	Axe 1 (+66.0%)	Axe 2 (+20.2%)
CONTRIBUTIONS POSITIVES	VITESSE +36,0%	CONFORT +69,0%
	SECURITE +34,0%	
CONTRIBUTIONS NEGATIVES		SECURITE -23,0%
		VITESSE -7,0%

La carte montre les positions des 3 critères et les coordonnées des 175 observations.

25 observations ne sont pas prises en compte (non-réponse à au moins un des critères).

86.2% de la variance est expliquée par les deux axes représentés.

Chaque observation est représentée par un point.



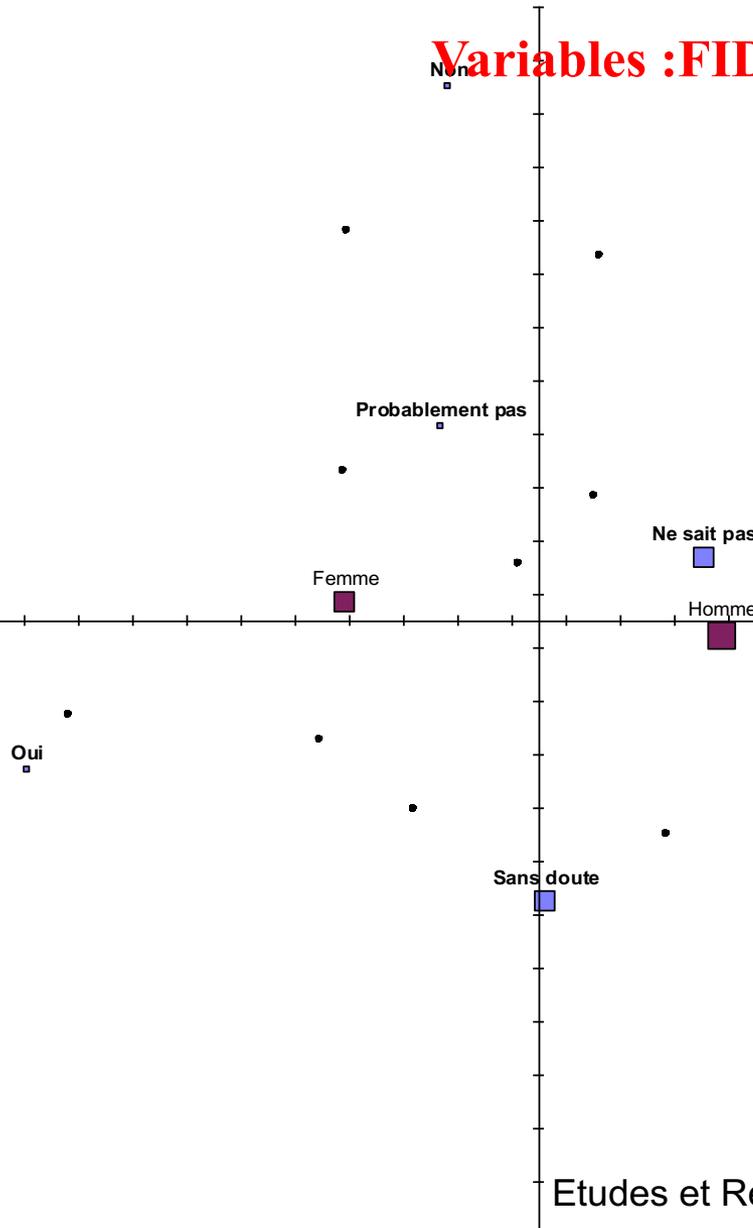
VIII - L'ANALYSE FACTORIELLE DES CORRESPONDANCES (AFC)

Le résultat de cette analyse est un graphique en deux dimensions, que l'on appelle la carte factorielle et qui donne les projections des modalités suivant les axes les plus représentatifs. Le pourcentage de la variance expliquée par chacun des axes est mentionnée en regard de chacun des axes. La fonction permet de choisir les deux facteurs (composantes) qui seront utilisées sur l'axe horizontal et l'axe vertical de la carte factorielle.

Le pourcentage de variance expliquée qui caractérise chacun des axes factoriels est indiqué sous les axes.

Analyse des correspondances multiples

Variables : FIDELITE, SEXE.



La carte montre les positions des 7 modalités et les coordonnées des 168 observations.

42.7% de la variance est expliquée par les deux axes représentés.

Les non-réponses ont été ignorées.

32 observations ne sont pas représentées (non-réponse à l'une au moins des questions).

Chaque observation est représentée par un point.

Analyse des correspondances multiples

Variables : **FIDELITE, SEXE.**

	F1	F2	F3	F4	F5	S1	S2
F1	15	-	-	-	-	5	10
F2	-	52	-	-	-	27	25
F3	-	-	19	-	-	9	10
F4	-	-	-	17	-	8	9
F5	-	-	-	-	65	37	28
S1	5	27	9	8	37	86	-
S2	10	25	10	9	28	-	82

La carte montre les positions des 7 modalités et les coordonnées des 168 observations.

42.7% de la variance est expliquée par les deux axes représentés..

Le tableau est le tableau des effectifs (tableau de Burt) pour les 7 modalités.

Les non-réponses ont été ignorées.

F1 : Oui F2 : Sans doute F3 : Probablement pas F4 : Non F5 : Ne sait pas S1 : Homme S2 : Femme

Analyse des correspondances multiples

Variables :FIDELITE, SEXE.

	Axe 1 (+22.7%)		Axe 2 (+20.0%)	
CONTRIBUTIONS POSITIVES	Homme	+24.3%	Non	+46.7%
	Ne sait pas	+14.7%	Probablement pas	+6.9%
	Sans doute	+0.0%	Ne sait pas	+2.4%
CONTRIBUTIONS NEGATIVES	Oui	-32.6%	Sans doute	-40.1%
	Femme	-25.6%	Oui	-3.3%
	Probablement pas	-1.6%	Homme	-0.3%

Le tableau donne, pour les 2 premiers axes factoriels, les contributions relatives (positives et négatives) des modalités.

Les non-réponses ont été ignorées.

IX - EXEMPLE

Aspirations des étudiants d'une école de management

- Questions d'identification
 - Sexe
 - Age
- "Si vous gagniez au loto, que feriez vous ?"

Un texte de 25 lignes

Un entretien non directif

- Ces données seront traitées à la prochaine séance.

IX - EXEMPLE

Construire des lexiques et recoder en thèmes

- Importer les données à partir d'un fichier texte
- Produire des listes de mots
 - Réduire
 - Supprimer des mots
 - Regrouper des mots (même racine, synonymie)
- Recodage à partir des lexiques
 - Création de variables de codage (les thèmes)

Pilote - vin.4DD

Aperçu à l'écran / Loto

Classement lexical / Loto

Supprimer	Marquer	Dictionnaire	Réponses	Editer
Grouper	Réduire	Contexte	Rép. n°1	Remplacer

Lexique	Occ.	Indi
je	136	74
m	16	15
en	53	41
vais	3	3
dans	22	21
la	23	20
Meuse	1	1
rendre	1	1
heureux	1	1
mes	27	24
enfants	17	15

LOTO

je m'en vais dans la Meuse / rendre
heureux mes enfants / une voiture, j'irais
en vacances.

Une des expr. marquées (OU booléen)

Corpus : 1821 Lexique : 419 Caractères séparateurs : 16 Maximum d'expressions différentes dans un questionnaire : 31

Aller à : < 1 >

Fermer Enregistrer les codifications

IX - EXEMPLE

Résultats

- Situer un mot dans son contexte à travers les concordances et les lexiques relatifs
- Extraire des VERBATIM en considérant le contexte du mot (profil)
- Produire des cartes d'associations lexicales
 - Tableau croisé pour décrire les classes par les thèmes (le Chi2)
 - Positionner les thèmes par l'analyse factorielle des correspondances

IX - EXEMPLE

Chi2 et tableau croisé

- Le test du chi2 est la certitude exprimée en pourcentage de la dépendance des deux variables.
- Plus 1-p est important plus les variables sont significativement dépendantes
- Si l'effectif est inférieur à l'effectif théorique, l'encadrement est en rouge sinon en bleu.

IX - EXEMPLE

L'analyse factorielle des correspondances

- Identifier un nombre de dimensions pour simplifier et interpréter un ensemble de données
- Variables nominales → Influence et importance des modalités
- Représentation des modalités → il met en correspondance (proximité et opposition) les liens éventuels entre les modalités des deux variables

Pilote - vin.4DD

Facteur 1/2 / Loto

PEM n°0/0

épargner
aider
enfants
+ oeuvres
donner
famille
amis

voiture
acheter
+ maison
vacances
+ voyage
argent
+ placement
+ vie

Contributions

Facteurs :		1	2		
% Information :		29,23%	13,01%		
Contributions :		Absol.	Relat.	Absol.	Relat.
voiture	1	0,48	0,02	14,17	0,36
+ placement	2	4,98	0,18	20,02	0,32
+ maison	3	14,86	0,67	6,04	0,12
+ vie	4	9,96	0,31	10,50	0,14
argent	5	0,02	0,00	10,83	0,30
épargner	6	4,34	0,12	12,38	0,15
+ voyage	1	0,33	0,03	0,78	0,03
vacances	2	1,07	0,06	0,17	0,00
acheter	3	7,71	0,37	16,61	0,36
enfants	1	0,66	0,02	0,67	0,01
donner	2	22,20	0,79	0,59	0,00
famille	3	4,24	0,17	0,63	0,01
+ oeuvres	4	18,99	0,75	0,00	0,00
amis	5	1,43	0,06	5,43	0,10
aider	6	8,66	0,38	1,11	0,02

↑ ↑ ↓ ↓ ↓ ↓
 Tri/départ Détail : Double-clic
 Détail Exporter Retour

A vertical toolbar on the left side of the application window. It contains several icons: a mouse cursor, a magnifying glass with a plus sign, a magnifying glass with a minus sign, a hand, a plus sign, a minus sign, a plus sign, a minus sign, a color palette, a circle with 'N°', the numbers '123', the letters 'ABC', and a grid icon labeled 'Rec'.

oeuvre

+ donner

amis

famille

+ voyages

placements

enfants

+ maison p

acheter

voiture

vacances

Navigation and editing icons: arrow, zoom in, zoom out, hand, pan, zoom in, zoom out, pan, fill, N°, 123, ABC, Rec.

oeuvres

amis

+ voyager

+ donner

placements

vacances

famille

enfants

+ voyages

voiture

+ maison
acheter

IX - EXEMPLE

L'analyse factorielle des correspondances

- Pourcentage d'information restituée par les axes
- Contribution des variables à la définition des axes
 - D'un point de vue géométrique
 - D'un point de vue numérique



3ème partie :

L'utilisation d'un logiciel de traitement des données : le Sphinx

S. MAYOL



I - PRESENTATION

- Le Sphinx est comme la loupe de Sherlock Homes : C 'est le **raisonnement** qui le pousse à n 'examiner que certains aspects d 'une enquête
- La statistique du Sphinx permet de **grossir** telle ou telle partie d 'une **enquête** mais ne dit pas **laquelle** il faut grossir.

3 ETAPES





3 ETAPES

- Élaboration du questionnaire
- Saisie des réponses
- Traitement et analyse



Élaboration du questionnaire

- **Modèles de questionnaire**
- **Bibliothèques de questions**
- **Définition de questions**
- **Questions tableau**
- **Renvois, filtres**
- **Mise en page automatique**

Questionnaire: *La liste des questions*

Le Sphinx Lexica - Automobiles [Elaboration du questionnaire]

Fichier Edition Stade Rédiger Organiser ?

LISTE DES QUESTIONS

Parties **A propos de votre voiture**

En-tête...

Chercher

Variable	Libellé	Modalités de réponse / Contrôles
A propos de votre voiture		
1. POSSESSION	Avez-vous une voiture à votre disposition ?	Oui ; Non ; <i>La réponse est obligatoire</i>
2. APPARTENANCE	A qui appartient-elle ?	A vous ; A votre société ; A quelqu'un
3. MARQUE	Quelle est la marque de votre voiture ?	Renault ; Peugeot ; Citroën ; Volksw
ab 4. MARQUE_AUTRE	Si "Autre", précisez :	
5. TYPE AUTO	S'agit-il d'une voiture achetée neuve ou d'occasion ?	Neuve ; Occasion ;
45 6. ANCIENNETE	Quel est son âge ?	
7. PUISSANCE	Quelle est sa puissance ?	- de 4 CV ; 4 à 5 CV ; 6 à 7 CV ; 8 à
8. CARBURANT	Quel est le carburant utilisé ?	Sans plomb ; Gasoil ; GPL ; Essenc
9. OPTIONS	Quel est son équipement ?	Jantes alu ; Vitres teintées ; Lecteur
Votre satisfaction		
10. VITESSE	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn
11. CONFORT	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn
12. SECURITE	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn
13. CONSOMMATION	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn
14. TENUE DE ROUTE	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn
15. POLLUTION	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn
16. PERF_ENTRETIEN	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn
17. ESPACE	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn
18. ESTHETIQUE	Etes-vous satisfait de la performance de votre voiture en ce	Pas du tout ; Plutôt non ; Moyenn

Questionnaire: *La définition des questions*

Question n° 3 [X]

Libellé :
Quelle est la marque de votre voiture ?

Variable :
MARQUE

Contrôles... Bibliothèque...

Barème...

Type

- Fermée (unique)
- Fermée (multiple)
- Fermée (échelle)
- Numérique
- Texte
- Code
- Date / heure
- Image

Modalités :

Renault ; Peugeot ; Citroën ; Volkswagen ; Fiat
; Ford ; Toyota ; BMW ; Mercedes ; Opel ; Volvo
; Autre ;

Modifier

12 modalités

Des réponses ont été saisies. Rectifier les modalités question par question.

Changer le type...

← →

OK Annuler

Questionnaire: *Mise en page et impression*

Impression des questionnaires - Automobiles (D:\SphinxME\Enquêtes\Automobiles)

A PROPOS DE VOTRE VOITURE

Une vraie-fausse enquête, démonstration des possibilités du Sphinx.

A propos de votre voiture

1. Avez-vous une voiture à votre disposition ?

1. Oui 2. Non

La réponse est obligatoire.

Aller à '23. CRITERES' si POSSESSION = "Non"

2. A qui appartient-elle ?

1. A vous 2. A votre société
 3. A quelqu'un de votre famille

3. Quelle est la marque de votre voiture ?

1. Renault 2. Peugeot 3. Citroën
 4. Volkswagen 5. Fiat 6. Ford
 7. Toyota 8. BMW 9. Mercedes
 10. Opel 11. Volvo 12. Autre

4. Si "Autre", précisez :

5. S'agit-il d'une voiture achetée neuve ou d'occasion ?

1. Neuve 2. Occasion

6. Quel est son âge ?

7. Quelle est sa puissance ?

1. - de 4 CV 2. 4 à 5 CV 3. 6 à 7 CV
 4. 8 à 9 CV 5. 10 CV et plus

8. Quel est le carburant utilisé ?

1. Sans plomb 2. Gasoil 3. GPL 4. Essence

9. Quel est son équipement ?

1. Jantes alu 2. Vitres teintées
 3. Lecteur CD 4. Climatisation
 5. Ordinateur de bord 6. Airbags
 7. Attelage 8. Intérieur cuir
 9. Autres

Vous pouvez cocher plusieurs cases.

Votre satisfaction

Etes-vous satisfait de la performance de votre voiture en ce qui concerne

- | | 1 | 2 | 3 | 4 | 5 |
|-----------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|
| 10. La vitesse | <input type="radio"/> |
| 11. Le confort | <input type="radio"/> |
| 12. La sécurité | <input type="radio"/> |

19. Donnez une note globale sur 20 de la qualité de votre voiture



II - LE QUESTIONNAIRE

6 TYPES DE QUESTIONS

- Fermées **uniques**
- Fermées **multiples**
- Fermées **échelles**
- Ouvertes **numériques**
- Ouvertes **textes**
- Ouvertes **codées**



QUESTION FERMEE UNIQUE

Modalité de réponse:

Une seule réponse possible parmi un **choix limité** de propositions

exemples: « oui » ou « non »

« rouge », « noir », « vert » ou bleu »



QUESTION FERMEE MULTIPLE

Modalité de réponse:

Plusieurs réponses possibles
parmi un **choix limité** de
propositions

exemple: « rouge », « noir »,
« vert », « bleu »



QUESTION FERMEE ECHELLE

Modalité de réponse:

Une seule réponse possible
parmi un choix propositions
graduées les unes par
rapport aux autres

exemple: « tout à fait », « plutôt »,
« plutôt pas », « pas du tout »



QUESTION OUVERTE NUMERIQUE

Modalité de réponse:

Réponse **libre** numérique sans propositions prédéfinies

**ATTENTION TRAITEMENT
DIFFICILE**



QUESTION OUVERTE TEXTE

Modalité de réponse:

Réponse **libre** texte **sans**
propositions prédéfinies

**ATTENTION TRAITEMENT
DIFFICILE**



QUESTION OUVERTE CODEE

Modalité de réponse:

Réponse **libre numérique** faisant
référence à une **nomenclature**

exemple : code postal

III - LA SAISIE DU QUESTIONNAIRE

- Un clic simple sur une ligne la **sélectionne**. Le menu « questionnaire » lui est applicable
- Un double clic sur une ligne fait apparaître une **boîte de dialogue** relatif à la question sélectionnée. On peut la **modifier**
- Un double clic sur la zone blanche au dessous de la liste fait apparaître un nouveau dialogue pour **créer une nouvelle question**
- Toute ligne sélectionnée peut être **glissée** vers le haut ou le bas pour **modifier sa place**



MODIFICATIONS APRES SAISIE

SEULES MODIFICATIONS POSSIBLES

- Changer une variable **fermée unique** en **fermée échelle** et réciproquement
- Faire des corrections de **texte**
- Ajouter une **modalité**

Les autres modifications seront
des **recodages**



Saisie des réponses

- **Saisie directe**
- **Saisie rapide**
- **Consultation, qualification**
- **Échantillonnage**
- **Redressement**
- **Gestion de panel**
- **Importation**



IV - SAISIE DES REPONSES

DEUX MODES DE SAISIE POSSIBLES:

- Une saisie dite « **directe** »

Se fait question par question. A chaque question, une fenêtre de saisie apparaît

- Une saisie dite « **rapide** »

Se fait sur l'ensemble des questions. Une seule fenêtre apparaît pour l'ensemble du questionnaire

Saisie rapide

Le Sphinx Lexica - Automobiles [Saisie des réponses]

A propos de votre voiture - Avez-vous une voiture à votre disposition ?

1. POSSESSION

2. APPARTENANCE

3. MARQUE

4. MARQUE_AUTRE

5. TYPE_AUTO

6. ANCIENNETE

7. PUISSANCE

8. CARBURANT

9. OPTIONS

10. VITESSE

11. CONFORT

12. SECURITE

13. CONSOMMATION

14. TENUE DE ROUTE

15. POLLUTION

16. PERF_ENTRETIEN

17. ESPACE

18. ESTHETIQUE

19. NOTE

20. KILOMETRAGE

21. ENTRETIEN

22. DEPENSE

23. CRITERES

24. FIDELITE

25. VOITURE IDEALE

26. SEXE

27. DATE_NAISSANCE

28. CSP

29. CODE POSTAL

POSESSION

La réponse est obligatoire.

Modalités de réponse :

1. Oui

2. Non

Ajouter une modalité...

Options...

Reprendre...

Compléter

Annuler

Terminer

Observation n° 201

Saisie directe

Etes-vous satisfait de la performance de votre voiture en ce qui concerne

	Pas du tout	Plutôt non	Moyennement	Plutôt oui	Tout à fait
10. La vitesse	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
11. Le confort	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
12. La sécurité	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
13. La consommation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>
14. La tenue de route	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. La pollution	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. L'entretien	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. L'espace	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. L'esthétique	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Votre satisfaction

Importation de données externes

Ouvrir un fichier de données



- Des données formatées Le fichier doit être un fichier ASCII (enregistré sous le format 'texte seulement' dans un tableur ou une base de données.).
- Un texte à analyser Le fichier doit être un fichier ASCII (texte seulement). Le texte peut contenir des annotations qui seront traduites en variables de contexte.
- Une base d'images Toutes les images (au format BMP) doivent être dans un même répertoire.
- Un tableau d'effectifs Les valeurs des cases du tableau pourront également être tapées directement dans une fenêtre.
- Une étude statistique Fichiers d'extension .SAV
- Une base de données Fichiers Access, DBase, Paradox, Excel, et toutes les bases de données pour lesquelles il existe un pilote 'ODBC'.
- Des données sur Internet Analyser un ensemble de pages (aspiration d'un site, résultat d'une recherche ou ensemble de liens)
- Une enquête sur Internet Télécharger une enquête publiée sur Internet.

- Fichier de données
- Contenu du Presse-Papiers



Traitement et analyses

- **Dépouillement automatique**
- **Édition de rapports**
- **Tri à plat**
- **Tri croisé**
- **Test statistiques**
- **Analyse multi-variée**
- **Re-codification**
- **Analyse lexicale**
- **Gestion de données**

Tableau à plat

Le Sphinx Plus² - Automobiles [Traitements et analyses]

Fichier Edition Stade Dépouiller Recoder Analyser Approfondir ?

TABLEAU A PLAT : MARQUE - Echantillon total (202 observations)

Tableau Graphique Valeurs Comparer Ordonner Regrouper Renommer Légende Carrousel Suivant
Tests Caract. Permuter Supprimer Revenir Dimensions Transposer Autre

MARQUE Quelle est la marque de votre voiture ?

MARQUE	Nb. cit.	Fréq.
▷ Renault	36	17,8%
▷ Peugeot	20	9,9%
▷ Citroën	21	10,4%
▷ Volkswagen	18	8,9%
▷ Fiat	12	5,9%
▷ Ford	14	6,9%
▷ Toyota	3	1,5%
▷ BMW	9	4,5%
▷ Mercedes	8	4,0%
▷ Opel	10	5,0%
▷ Volvo	3	1,5%
▷ Autre	21	10,4%
TOTAL OBS.	202	

La somme des pourcentages est inférieure à 100% du fait des suppressions.

Tableau croisé / AFC

Tableau Croisé : MARQUE ==> FIDELITE - Echantillon total (200 observations)

	▽	▽	▽	▽	▽	
FIDELITE MARQUE	Oui	Sans doute	Probablement pas	Non	Ne sait pas	TOTAL
▷ Renault	26,7% (4)	34,6% (18)	5,3% (1)	0,0% (0)	16,9% (11)	18,0% (34)
▷ Peugeot	13,3% (2)	9,6% (5)	10,5% (2)	0,0% (0)	13,8% (9)	10,0% (18)
▷ Citroën	26,7% (4)	5,8% (3)	21,1% (4)	5,9% (1)	9,2% (6)	10,5% (18)
▷ Volkswagen	13,3% (2)	15,4% (8)	5,3% (1)	5,9% (1)	9,2% (6)	9,0% (18)
▷ Fiat	0,0% (0)	1,9% (1)	15,8% (3)	23,5% (4)	6,2% (4)	6,0% (12)
▷ Ford	0,0% (0)	1,9% (1)	10,5% (2)	11,8% (2)	13,8% (9)	7,0% (14)
▷ Toyota	0,0% (0)	3,8% (2)	0,0% (0)	0,0% (0)	1,5% (1)	1,5% (3)
▷ BMW	13,3% (2)	3,8% (2)	0,0% (0)	0,0% (0)	7,7% (5)	4,5% (9)
▷ Mercedes	6,7% (1)	3,8% (2)	0,0% (0)	11,8% (2)	4,6% (3)	4,0% (8)
▷ Opel	0,0% (0)	5,8% (3)	10,5% (2)	0,0% (0)	7,7% (5)	5,0% (10)
▷ Volvo	0,0% (0)	0,0% (0)	5,3% (1)	5,9% (1)	1,5% (1)	1,5% (3)
▷ Autre	0,0% (0)	13,5% (7)	15,8% (3)	35,3% (6)	7,7% (5)	10,5% (21)
TOTAL	100% (15)	100% (52)	100% (19)	100% (17)	100% (65)	100% (168)

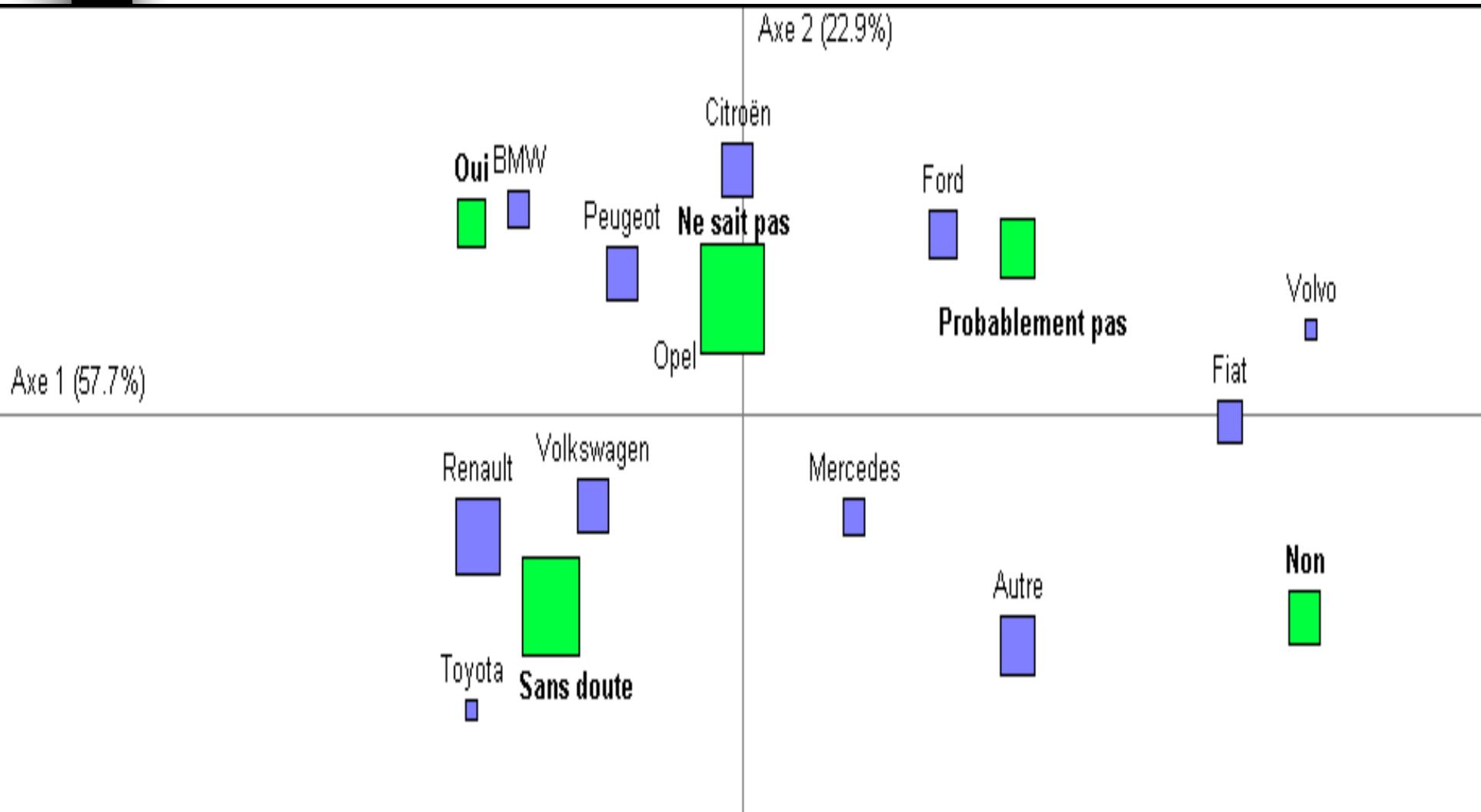
La dépendance est significative. $\chi^2 = 68,46$, $ddl = 44$, $1-p = 98,94\%$.

Les cases encadrées en bleu (rose) sont celles pour lesquelles l'effectif réel est nettement supérieur (inférieur) à l'effectif théorique.

Attention, 49 (81.7%) cases ont un effectif théorique inférieur à 5, les règles du χ^2 ne sont pas réellement applicables.

Les valeurs du tableau sont les pourcentages en colonne établis sur 200 observations.

Tableau croisé / AFC



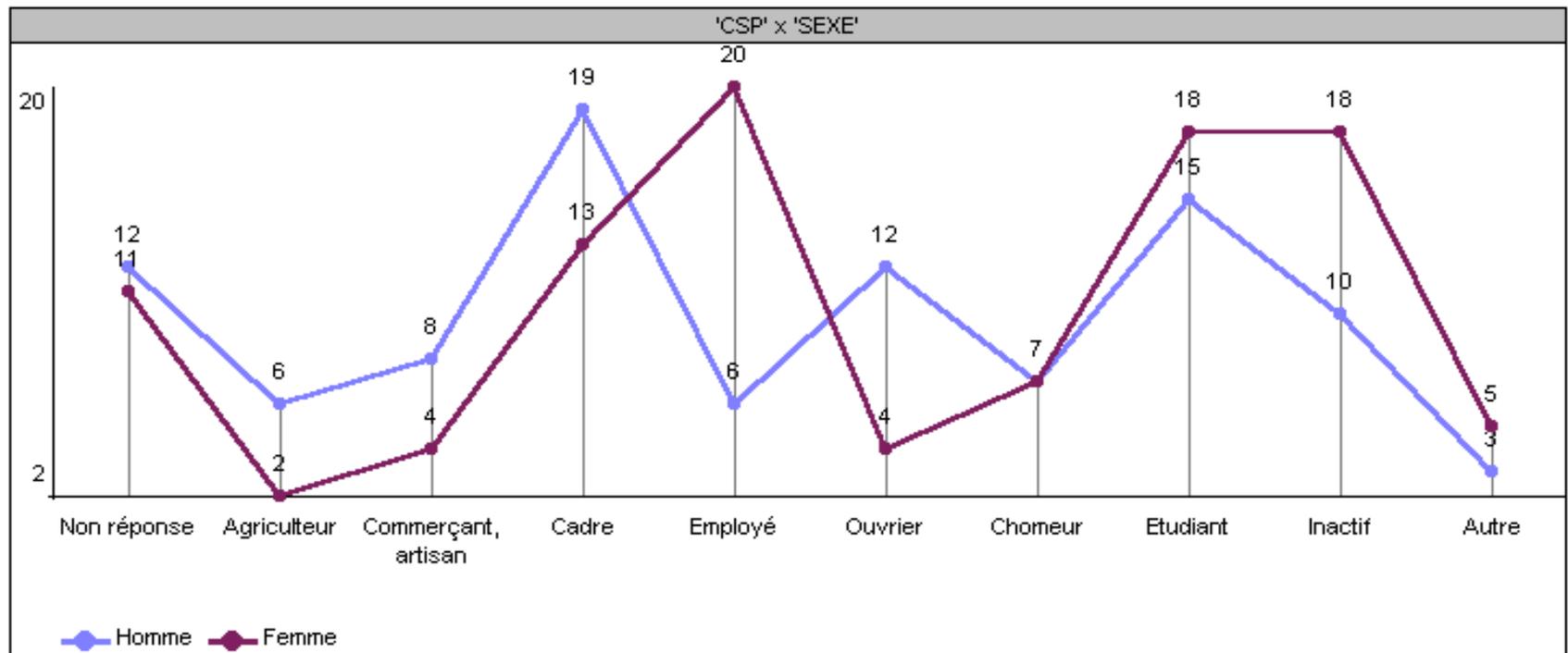
Graphiques

Le Sphinx Lexica - Automobiles [Traitements et analyses]

Fichier Edition Stade Dépouiller Recoder Analyser Approfondir ?

TABLEAU CROISE : CSP ==> SEXE - Echantillon total (200 observations)

Tableau Graphique Profils Ordonner Regrouper Renommer Légende Carrousel Suivant
AFC Options Caract. Permuter Supprimer Revenir Dimensions Transposer Autre



4ème partie :
Exemples : Le marché de
l'ultra frais et les salles de sport



S. MAYOL

Exemple n°1 : Le marché des salles de sport



S. MAYOL

Les salles de sport en France



Le marché
et
l'enseigne Moving

La pratique du sport

Questions posées :

Faites-vous de la musculation, du fitness ou de la gymnastique d'entretien, maintien en forme, que ce soit chez vous ou dans une salle de sport ?

Si non, pratiquez-vous d'autres sports ?

	2000	2002	2005
Pratiquants de fitness et autre	25%	28%	27%
Pratiquants d'autres sports	31%	33%	28%
Non sportifs	44%	39%	45%

⇒ Soit aujourd'hui entre 12 et 13 millions de personnes qui pratiquent la musculation, le fitness....

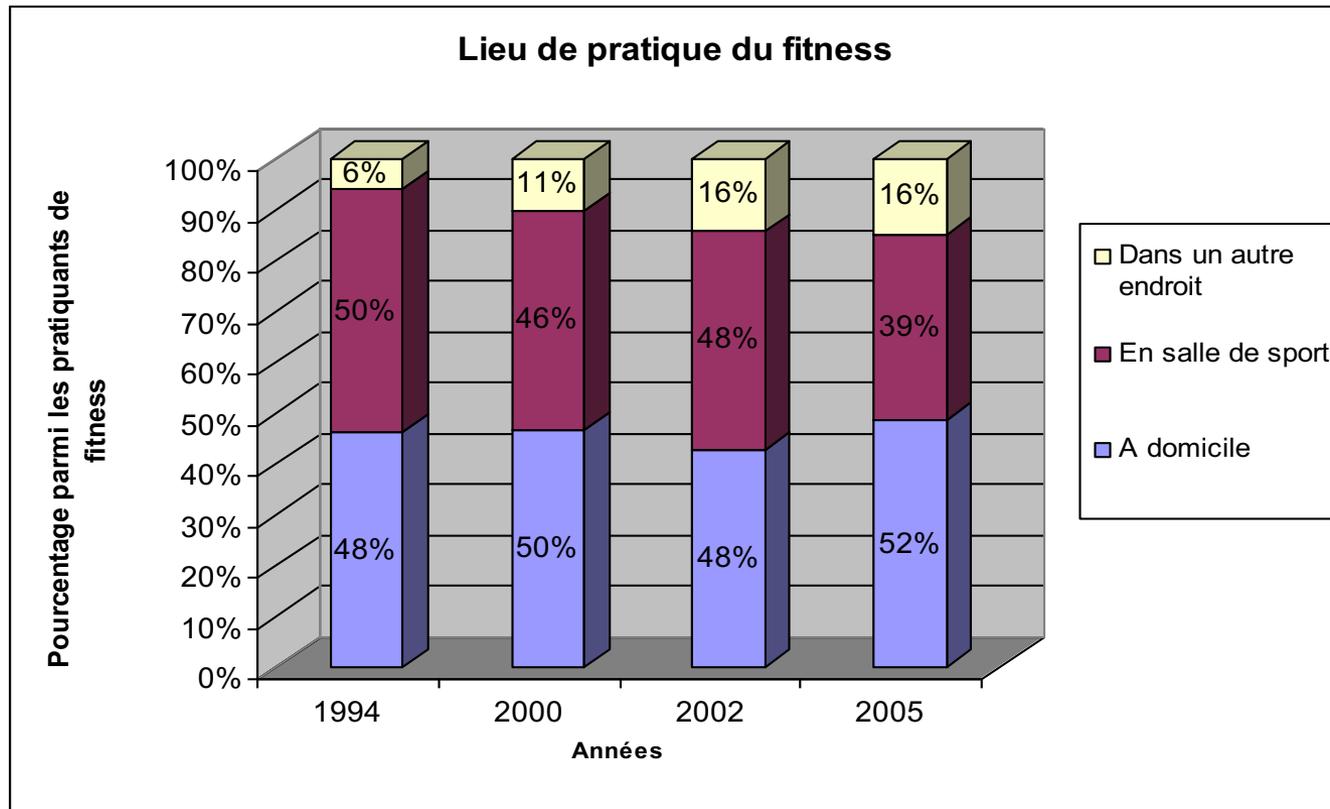
Détails des résultats par sexe et tranches d'âge

	Homme	Femme	15-24 ans	25-49 ans	50 ans et +
Pratiquants de fitness et autre	23%	31%	40%	25%	25%
Pratiquants d'autres sports	38%	20%	30%	31%	24%
Non sportifs	39%	49%	30%	44%	51%

Les femmes sont les premières amatrices de fitness alors que les hommes font plus de sport mais de manière plus diversifiée.

Le fitness attire aussi de plus en plus les jeunes.

Lieu de pratique du fitness



Il y a donc aujourd'hui environ 5 millions de Français inscrits dans une salle de sport. Néanmoins, on peut constater une **baisse de la part des salles de sport**. Celle-ci s'explique notamment par le développement de la vente de matériel permettant de pratiquer **à domicile** (appareils de musculation, guides divers sous format vidéo/informatique/écrit...) mais aussi par **l'essor de structures sportives au sein des entreprises** grâce à l'action du comité d'entreprise.



Les motivations à la pratique du fitness 1/2

La raison principale avancée par les pratiquants de fitness et autres activités proches, est de l'ordre de la **santé**. Il s'agit pour eux de se maintenir dans une bonne forme physique.

Cependant, ce critère est prédominant surtout pour les plus de 30 ans et croît avec l'âge.

Le critère d'**esthétisme** vient en second lieu. D'un côté, il y a ceux qui cherchent à **développer leur musculature**. Ce sont surtout les hommes et les jeunes. De l'autre côté, on retrouve ceux qui cherchent à **garder la ligne** voire à **perdre du poids**. Cette fois, ce sont plutôt les femmes et les jeunes. Ce constat est facilement expliqué par les psychologues par les différences de morphologies entre les sexes mais surtout les divergences de leurs préoccupations.



Les motivations à la pratique du fitness 2/2

L'aspect ludique à proprement parlé du sport n'arrive qu'en troisième position mais reste quand même un facteur de motivation important. Il s'agit de faire du sport par **goût du sport**, pour **se défouler**, éventuellement pour **se relaxer**...

Enfin, l'envie de **faire de nouvelles connaissances** par la pratique en groupe constitue une motivation supplémentaire, notamment pour l'inscription en salle de sport. Néanmoins, il est à noter que cette préoccupation concerne beaucoup moins la tranche d'âge 25-49 ans qui estime avoir assez d'opportunités ailleurs (notamment par leur emploi) de satisfaire leur sociabilité.

L'image des salles de sport

	ENSEMBLE		ADHERENTS A UNE SALLE DE SPORT	
	2002	2005	2002	2005
Très bonne image	15%	13%	21%	20%
Plutôt bonne image	59%	60%	67%	69%
Total de "bonne image"	74%	73%	88%	89%
Plutôt mauvaise image	13%	11%	7%	10%
Très mauvaise image	2%	2%	1%	0%
Total de "mauvaise image"	15%	13%	8%	10%

L'image des salles de sport est **plutôt bonne** y compris dans l'esprit de ceux qui n'y sont pas adhérents. Celle-ci est d'ailleurs restée plutôt stable ces dernières années.

Cependant, il reste des personnes réticentes notamment plus l'on progresse dans l'âge (alors que l'on ne décerne aucune différence majeure entre hommes et femmes). Ceci peut s'expliquer par la volonté de donner une **image jeune qui ne correspond pas nécessairement aux aspirations des seniors**



Les facteurs qui contribuent à la bonne image des salles de sport

1/2

Le premier critère à entrer en ligne de compte est la **diversité des activités** sportives proposées à la salle de sport. En effet, outre le fitness traditionnel sous toutes ses formes (aérobic, abdos-fessiers, exercices cardio-musculaires, step...), les clubs de sports tendent vers une diversification de leur offre : exercices de relaxation et détente (stretching, yoga, sauna...), espace beauté, espace spa...

Ensuite, c'est **le personnel** qui détermine l'image des salles de sport. Il s'agit d'un personnel jeune, compétent et enthousiaste. Face aux exigences croissantes des clients, les clubs ont fait appel à des professeurs n'ayant plus uniquement une formation sportive mais aussi d'enseignement afin de répondre de manière efficace à leur mission d'enseignement.

A égalité, on trouve **l'ambiance**. Il s'agit d'entretenir une ambiance détendue et jeune qui, néanmoins, comme nous l'avons déjà signalé, peut déplaire aux adhérents les plus âgés.



Les facteurs qui contribuent à la bonne image des salles de sport

2/2

Devancé de peu, on retrouve ensuite le **niveau d'équipement**. Ces dernières années, la politique des salles de sport s'est orientée vers une modernisation : nombreux renouvellements des appareils, achats de nouvelles machines, améliorations des espaces...

Outre ces quatre critères, suivent à quasi égalité le **bon aménagement des horaires** (il s'agit de coller à la demande notamment de ceux qui souhaitent en profiter après leurs journées de travail en proposant par exemple des créneaux nocturnes et durant le week-end), **la bonne hygiène des locaux** (importante dans un espace de sport collectif surtout étant la proportion de femmes – plus sensibles à ce genre d'arguments-) et enfin, **le suivi personnalisé** (via des conseils, des bilans réguliers, l'élaboration de programmes...).



Les facteurs qui contribuent à la mauvaise image des salles de sport 1/2

Incontestablement, ce sont les **prix** qui entravent le plus la fréquentation des salles de sport (c'est le cas pour plus de 60% ayant une mauvaise image des salles de sport).

En dépit des formules tarifaires diverses et variées proposées, il faut compter en moyenne 650 euros à l'année, 300 euros pour le trimestre et 25 euros pour la journée. Par une gamme de prix dégressifs, on remarque que les salles de sport tentent de fidéliser leurs clients mais ceux-ci restent trop prohibitifs. Ce genre d'arguments expliquent notamment que les plus jeunes (généralement les moins aisés) préfèrent pratiquer ces sports ailleurs.



Les facteurs qui contribuent à la mauvaise image des salles de sport 2/2

Loin derrière on retrouve la critique de **pratiques commerciales agressives pour la signature des contrats**. Effectivement, cette façon de procéder s'inscrit dans la politique de fidélisation des clients.

Mais il faut aussi évoquer le problème de la « surpopulation » des salles de sport. Le **sureffectif** pousse certains à se tourner vers la formule à domicile (parfois même en petit groupe).

Puis, c'est **l'atmosphère générale** qui est critiquée : l'envie à tout prix de faire « branché » et à la mode, l'aspect club de rencontre...

Par contre, l'encadrement en lui-même, les horaires, le personnel, les locaux... sont peu remis en question.



MOVING

Quelques chiffres (source : moving.fr)

créé en 1987

+ de 100000 adhérents dans 70 clubs

1er réseau européen de clubs de forme

CA: 250 M€

750 salariés

Principaux concurrents :

Club Med Gym

Forest Hill

Gymnasium



Quelle cible?

Marché doté d'un réel potentiel de croissance

4,6% de pénétration en France contre 6 à 10% en Europe et aux Etats-Unis

13% des Français ont l'intention d'adhérer prochainement à un club de forme

20% pratiquent des activités proches de celles du Moving



Moving de Versailles (quel marché?)

(source INSEE)

Population vieillissante (services adaptés)

>7,5% de + de 75 ans

<23 % de – de 20 ans

Forte proportion (23%) de ménages à une personne (marché favorable)



Expertise accrue du public

+ exigeant en termes de :

Services

Qualité

Prix

Souplesse de formules



Quel positionnement marketing?

Un positionnement marketing de leader

Référence sur les activités et innovations produits

Sur l'ambiance et les relations clients

Sur les offres commerciales

Avantage du réseau

Plan de fidélisation de la clientèle (formules longue durée, relations aux clients accrus, offres promotionnelles, veille marketing)



Importance de la communication

Promesse d'un « bien-être plaisir au quotidien »

Organe de communication interne : Le Moving Mag (disponible dans tous les clubs Moving)

Sur quels canaux agir?

Publicité à la télévision (seul club)

Relations avec la presse

Suivi personnalisé du client (bilan à l'entrée, programme individuel...)

Exemple n°2 : Le marché de l'ultra frais



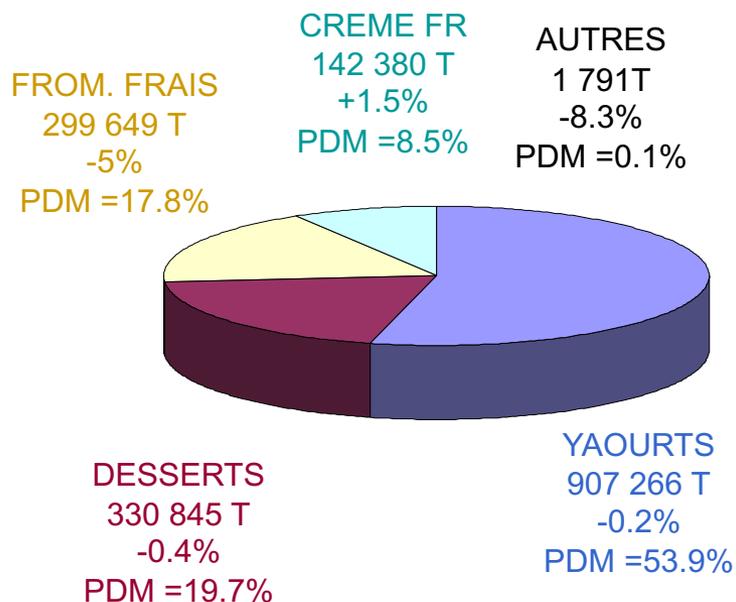
S. MAYOL

ULTRA FRAIS TOTAL PAR FAMILLE

POIDS DES FAMILLES EN VOLUME

Volume total UF : 1 681 931 T

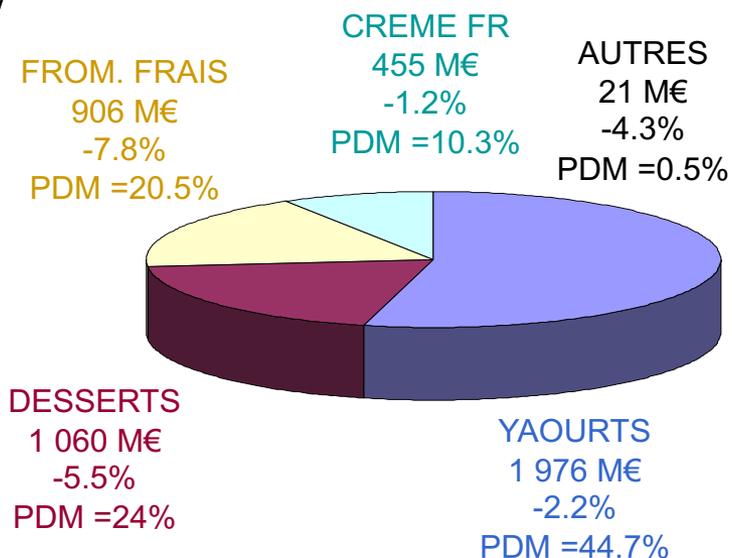
Evolution: - 1.0 %



POIDS DES FAMILLES EN VALEUR

Valeur totale UF : 4 418 M€

Evolution: - 4.1 %

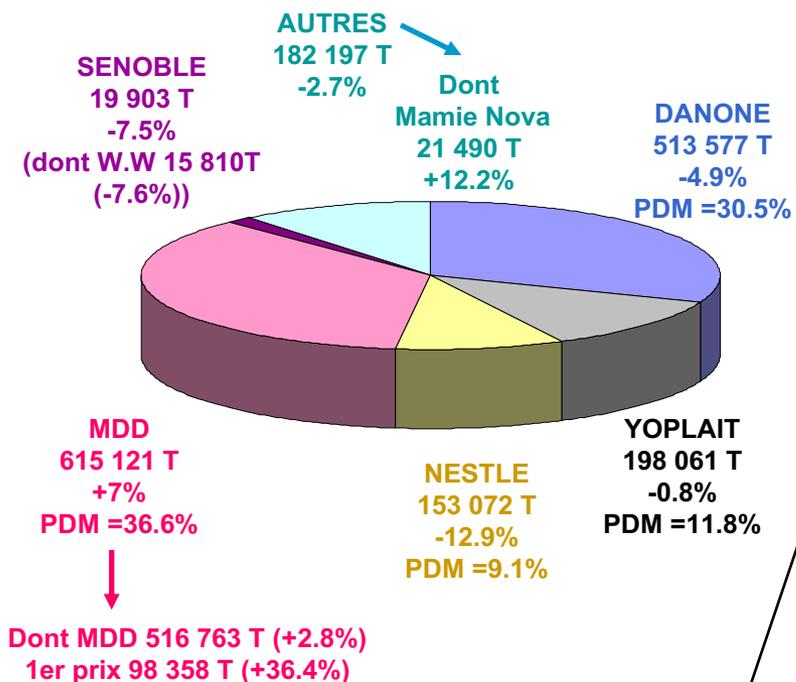


ULTRA FRAIS TOTAL PAR MARQUE

POIDS DES FAMILLES EN VOLUME

Volume total UF : 1 681 931 T

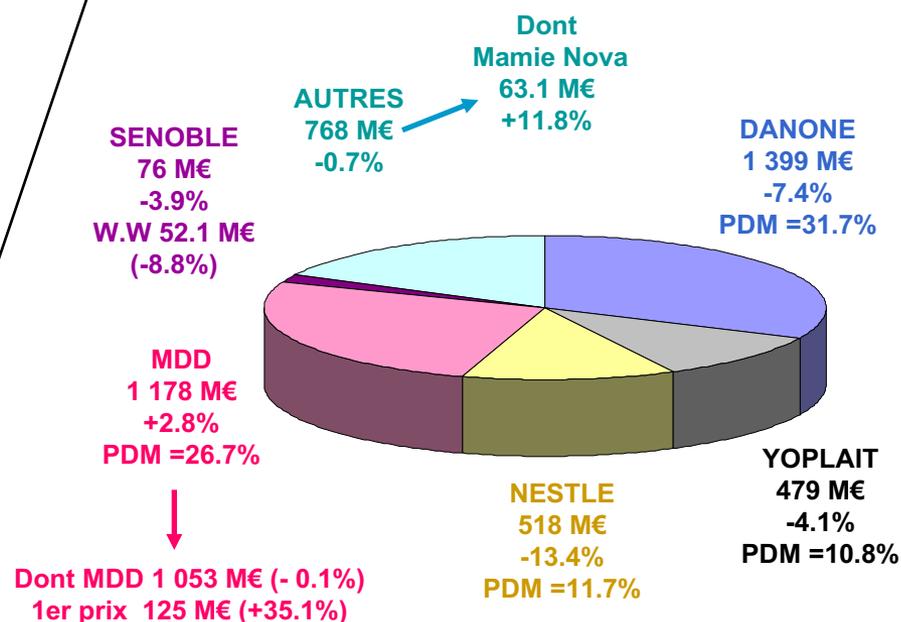
Evolution: - 1.0 %



POIDS DES FAMILLES EN VALEUR

Valeur totale UF : 4 418 M€

Evolution: - 4.1 %





YAOURTS

YAOURTS PAR MARQUE

POIDS DES FAMILLES EN VOLUME

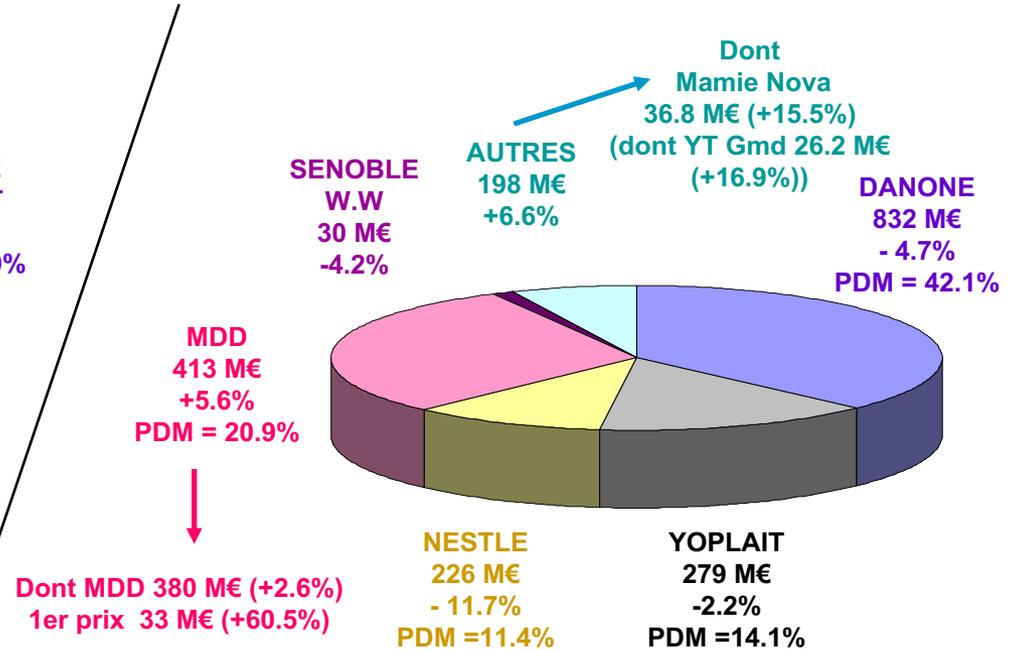
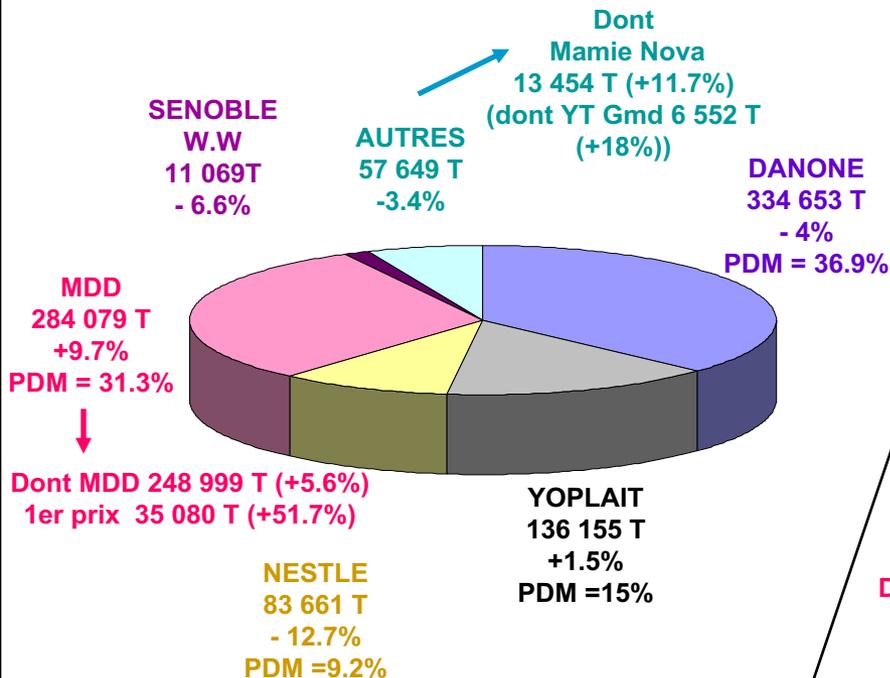
Volume total YAOURT : 907 266 T

Evolution : - 0.2%

POIDS DES FAMILLES EN VALEUR

Valeur totale YAOURT : 1 976.1 M€

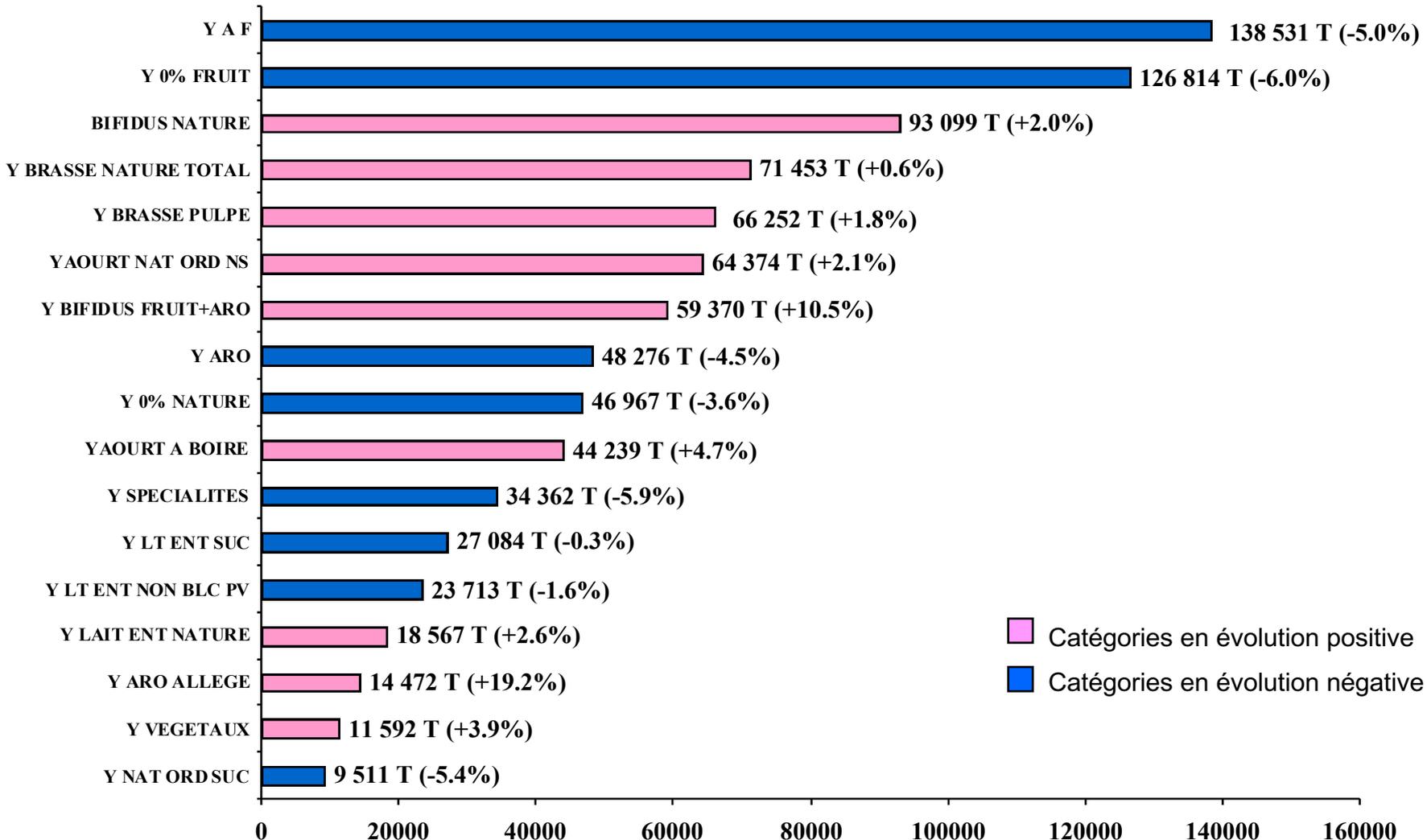
Evolution : - 2.2 %



YAOURTS EN VOLUME

POIDS DES YAOURTS : 907 266 T

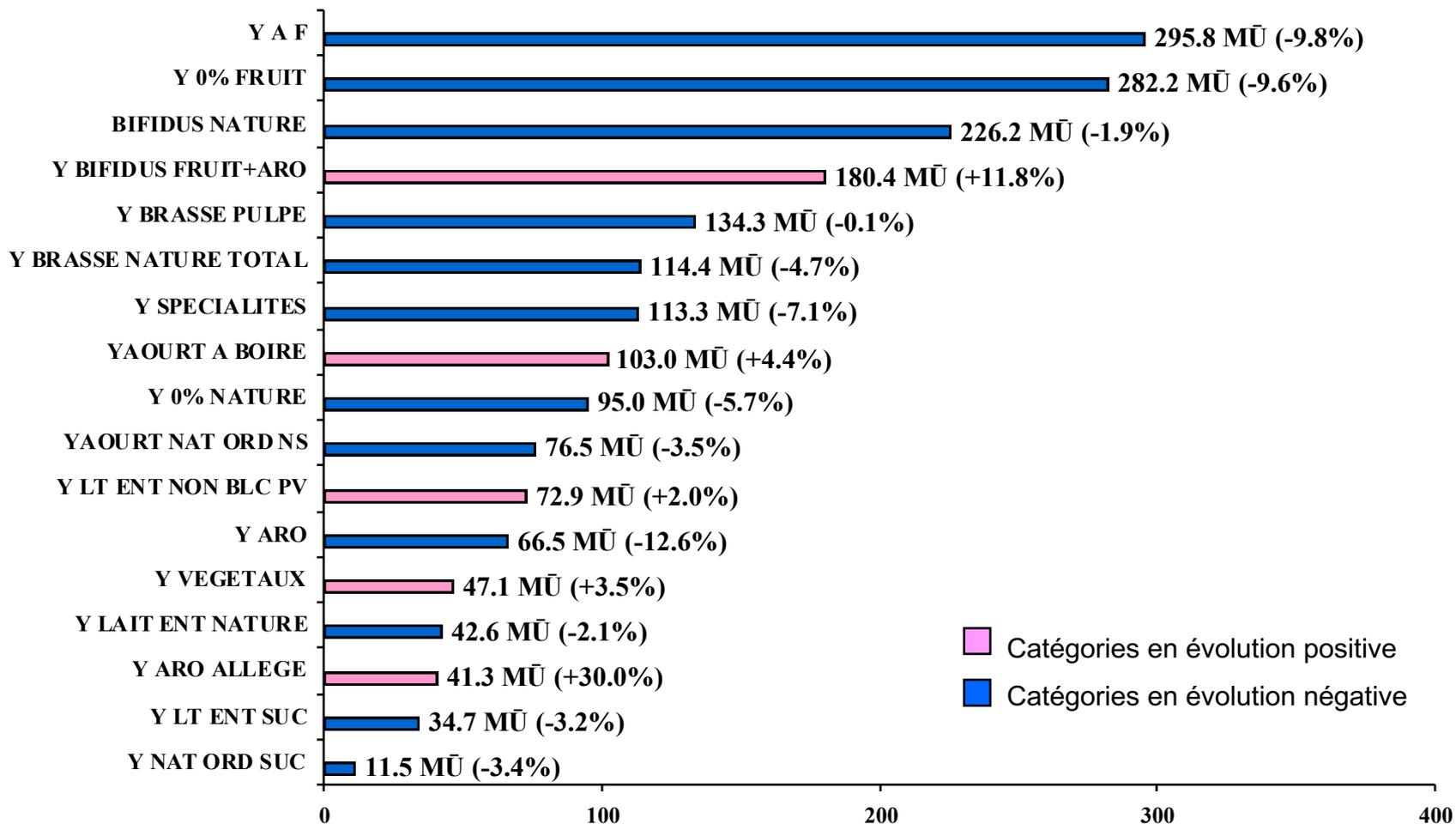
EVOLUTION: - 0.2 %



YAOURTS EN VALEUR

POIDS DES YAOURTS : 1 976.1 M€

EVOLUTION: - 2.2%



YAOURTS PROGRESSIONS

	Volume en Tonnes	Evolution volume	Evolution valeur	Tonnage MDD (en tonnes)	PDM MDD en volume
Y ARO ALLEGE	14 472	19.2%	30.0%		0.0%
Y BIFIDUS FRUIT+ARO	59 370	10.5%	11.8%	1 436	2.4%
YAOURT A BOIRE	44 239	4.7%	4.4%	3 672	8.3%
Y VEGETAUX	11 592	3.9%	3.5%		0.0%
Y LAIT ENT NATURE	18 567	2.6%	-2.1%	10 086	54.3%
YAOURT NAT ORD NS	64 374	2.1%	-3.5%	35 852	55.7%
BIFIDUS NATURE	93 099	2.0%	-1.9%	12 870	13.8%
Y BRASSE PULPE	66 252	1.8%	-0.1%	27 976	42.2%
Y BRASSE NATURE TOTAL	71 453	0.6%	-4.7%	40 200	56.3%

YAOURTS REGRESSIONS

	Volume en Tonnes	Evolution volume	Evolution valeur	Tonnage MDD (en tonnes)	PDM MDD en volume
Y LT ENT SUC	27 084	-0.3%	-3.2%		0.0%
Y LT ENT NON BLC PV	23 713	-1.6%	2.0%	8 456	35.7%
Y 0% NATURE	46 967	-3.6%	-5.7%	17 884	38.1%
Y ARO	48 276	-4.5%	-12.6%	29 500	61.1%
Y A F	138 531	-5.0%	-9.8%	45 065	32.5%
Y NAT ORD SUC	9 511	-5.4%	-3.4%	19 043	200.2%
Y SPECIALITES	34 362	-5.9%	-7.1%	1 692	4.9%
Y 0% FRUIT	126 814	-6.0%	-9.6%	22 810	18.0%

FOCUS SUR LES YAOURTS

Dans les **YAF** : 138 531 T (-5.0%) → Danone **Croc' Fruits** 10 409 T (-33.8%)
Danone **Recette Crémeuse** 22 001 T (-9.5%)
La Laitière **Fruits Pâtissiers** 10 408 T (+4.6%)

YAF 0% 126 814 T (-6.0%) → Taillefine **Crerosso** 7 493 T (+4.6%)
Sveltesse **Recette Dessert** 7 329 T (-31.1%)

Brassés Pulpés 66 252 T (+1.8%) → Danone **Velouté Fruix** 30 612 T (-1.1%)

Yaourt à Boire 44 239 T (+4.7%) → Nestlé **Yoco** 5 354 T (-17.8%)
Yoplait **Yop** 30 719 T (+9.8%)

FOCUS SUR LES YAOURTS

Bifidus Nature 93 099 T (+2.0%) → Danone **Actimel nature** 15 375 T (-19.6%)

Danone **Bifidus Nature** 61 054 T (+5.2%)

Bifidus Fruit + Aro 59 370 T (+10.5%)

→ Danone **Bio Fibres** 22 238 T (+24.6%)
dont Bio Figues 2 187 T et Bio Rhubarbe 909 T (Nov. 03)

Danone **Actimel Orange** 560 T (-31.7%)

Danone **Actimel Fraises** 5 578 T / **Multifruits** 2 045 T

Yaourt Santé 43 463 T (+12.5%) → Danone **Danacol** 4 068 T (Mars 2004)

(dont Actimel)

St Hubert Ilô 1 115 T (Nov. 2003)

Fruit d'Or Pro-Activ 2 887 T (Mars 2004)

Nestlé Geste Minceur 812 T (Mars 2004)

Yoplait Renforce 699 T (Avril 2004)

Danone Actimel 0% nature 7 915 T (+5.5%)

Danone Actimel Ananas 1 444 T (Sept. 2004)



DESSERTS

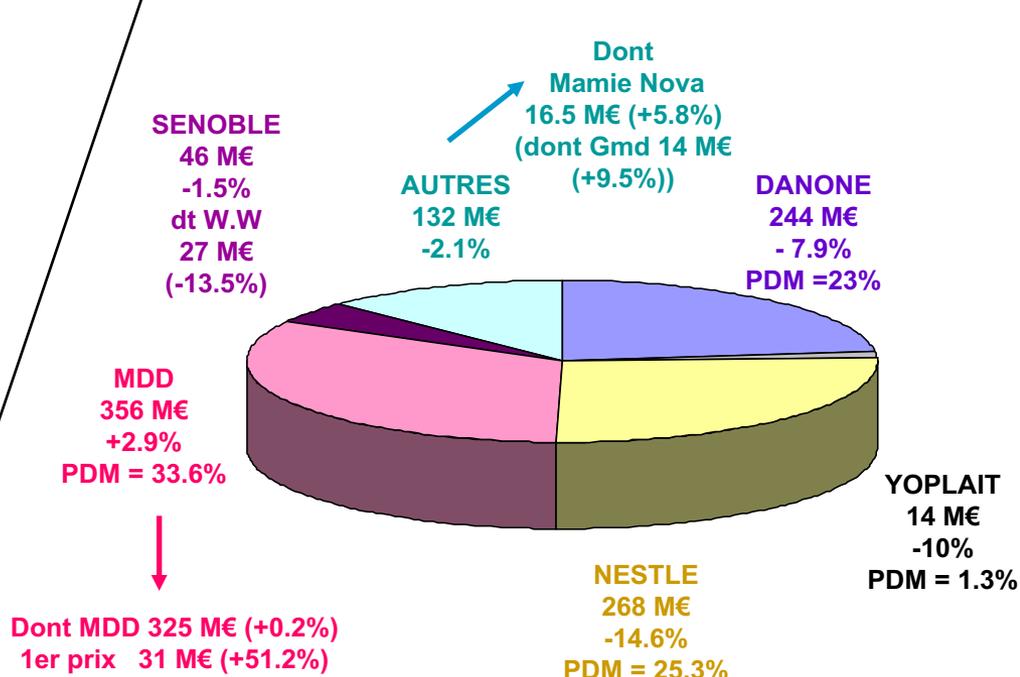
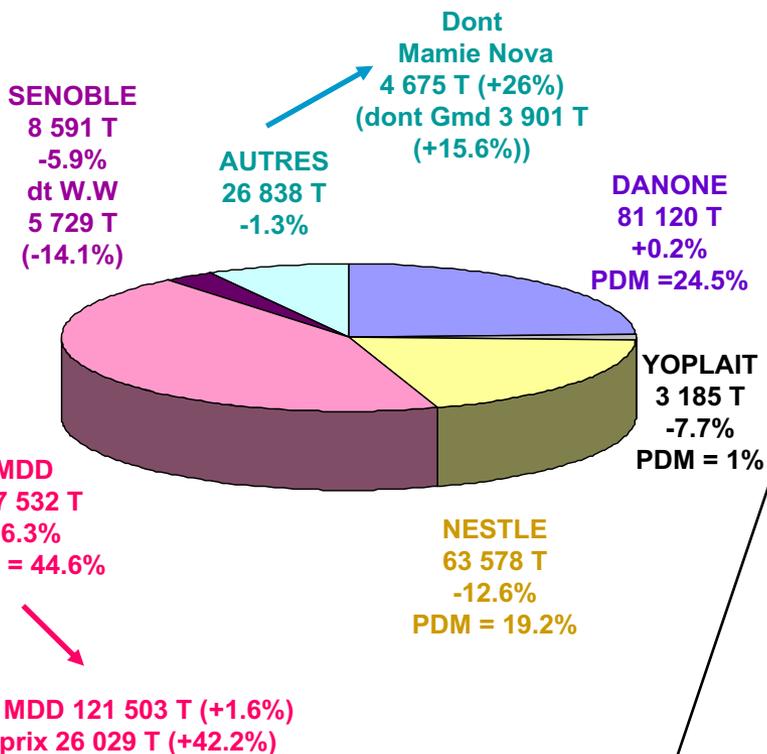
DESSERTS PAR MARQUE

POIDS DES FAMILLES EN VOLUME

Volume total DESSERT: 330 845 T
 Évolution: - 0.4%

POIDS DES FAMILLES EN VALEUR

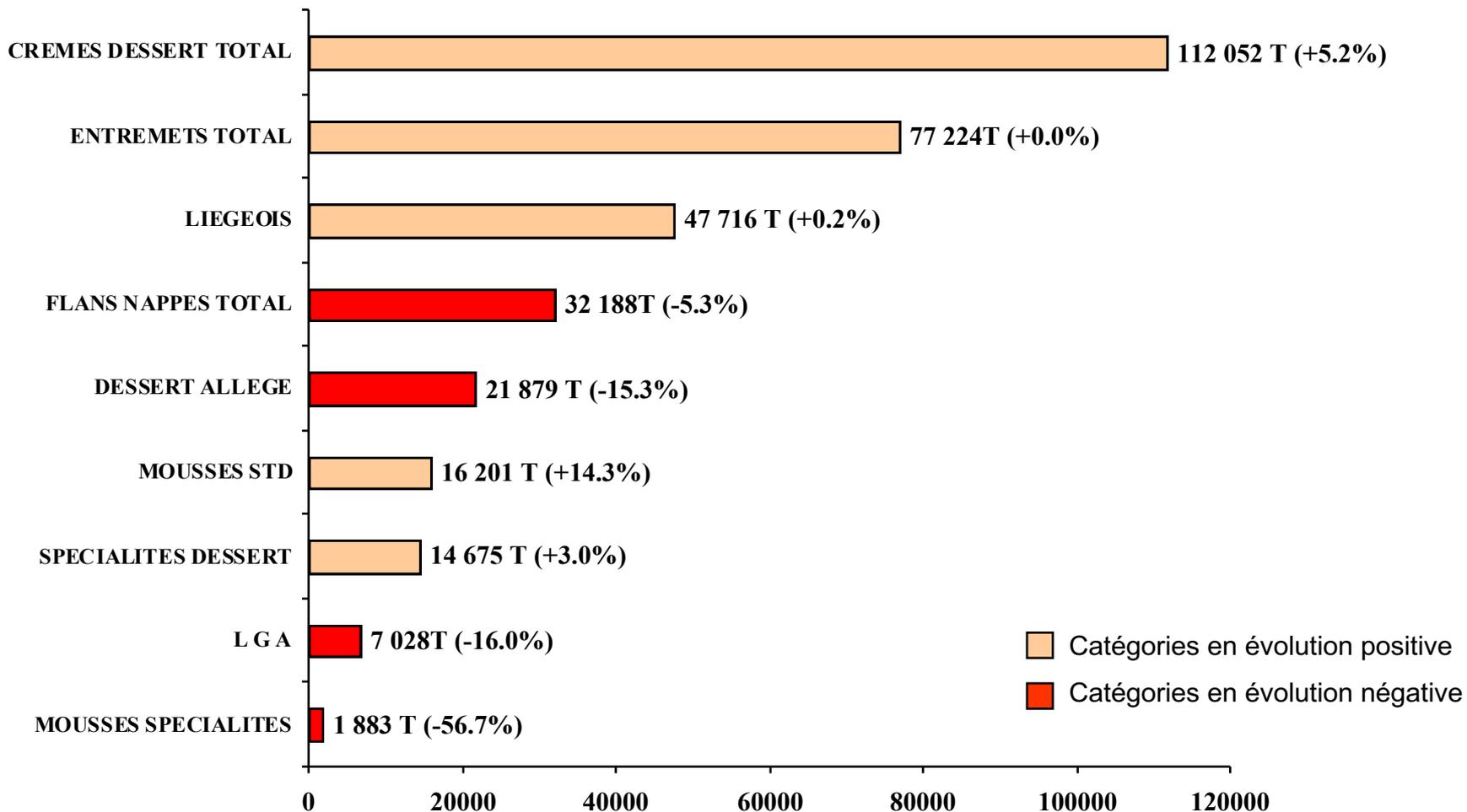
Valeur totale DESSERT : 1 059.5 M€
 Évolution: - 5.5 %



DESSERTS EN VOLUME

POIDS DES DESSERTS : 330 845 T

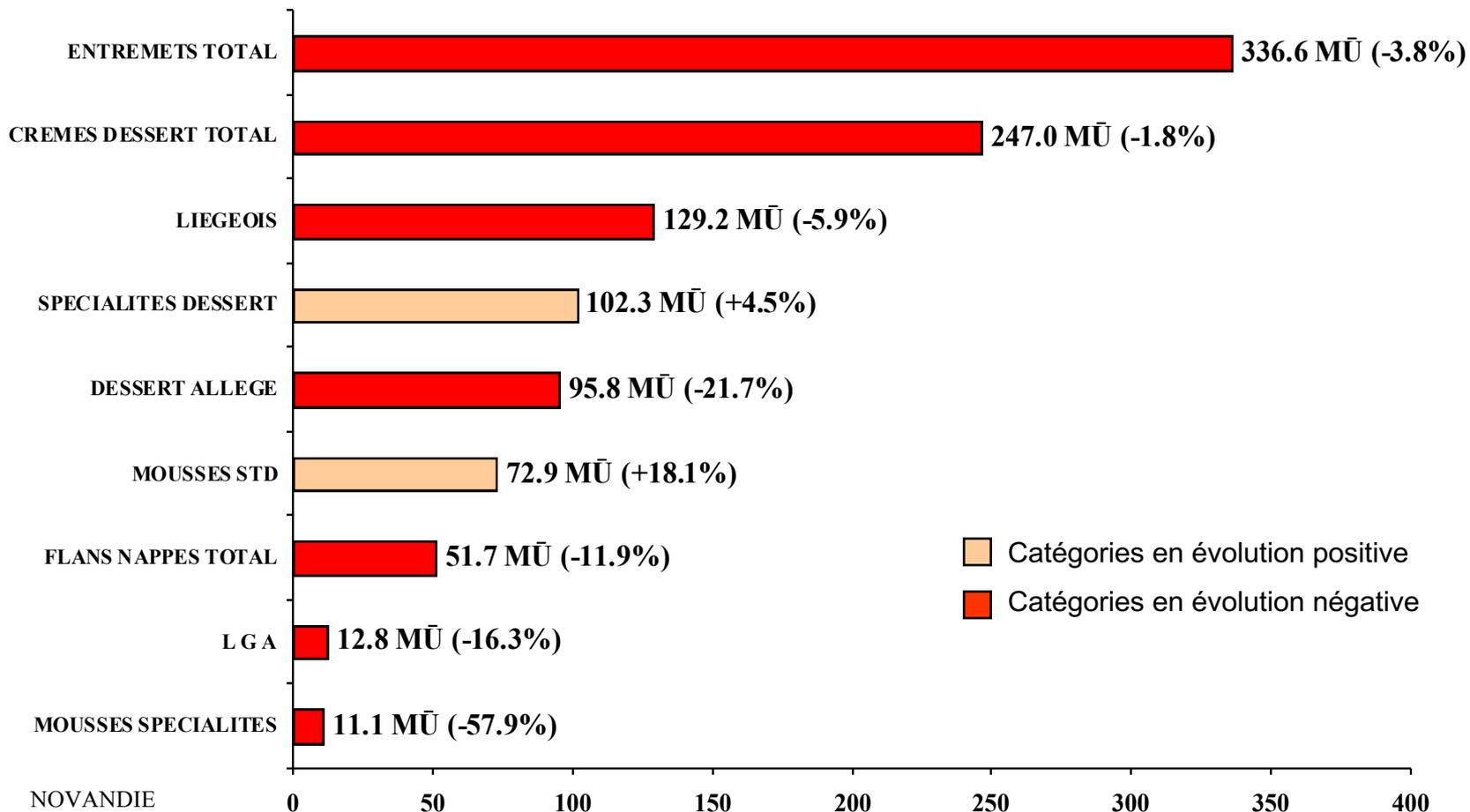
EVOLUTION: -0.4%



DESSERTS EN VALEUR

POIDS DES DESSERTS : 1 059.5 M€

EVOLUTION: - 5.5%



DESSERTS PROGRESSIONS

	Volume en Tonnes	Evolution volume	Evolution valeur	Tonnage MDD (en tonnes)	PDM MDD en volume
MOUSSES STD	16 201	14.3%	+ 18.1%	8 715	53.8%
CREMES DESSERT TOTAL	112 052	5.2%	- 1.8%	52 211	46.6%
SPECIALITES DESSERT	14 675	3.0%	+ 4.5%	5 299	36.1%
LIEGEOIS	47 716	0.2%	- 5.9%	26 704	56.0%
ENTREMETS TOTAL	77 224	0.0%	- 3.8%	26 476	34.3%

DESSERTS REGRESSIONS

	Volume en Tonnes	Evolution volume	Evolution valeur	Tonnage MDD (en tonnes)	PDM MDD en volume
FLANS NAPPES TOTAL	32 188	-5.3%	- 11.9%	22 389	69.6%
DESSERT ALLEGE	21 879	-15.3%	- 21.7%		
L G A	7 028	-16.0%	- 16.3%	4 380	62.3%
MOUSSES SPECIALITES	1 883	-56.7%	- 57.9%	8 715	462.9%

FOCUS SUR LES DESSERTS

- Mousse standard** 16 201 T (+14.3%) → **Nestlé Mousse standard** 6 596 T (+22.4%)
Pour info Nestlé **Mousse Spécialité** 1 883 T (-56.7%)
- Crème Dessert** 112 052 T (+5.2%) → **Danone Crème Dessert Danette** 56 945 T (+5.3%)
MDD Crème Dessert 52 211 T (+7.4%)
- Entremets** 77 224 T (+0.0%) → **Danone Entremets** 10 884 T (-9.4%)
La Laitère Petits pots de crème 13 611 T (-9.0%)
dont **Petit pot Chocolat** 4 896 T (-9.0%)
Petit pot Pistache 1 675 T (-21.2%)
- Liégeois** 47 716 T (0.2%) → **Nestlé Mousse Viennois Pistache** 335 T (lancement Fév/04)
- Desserts Allégés** 21 879 T (-15.3%) → **Danone** 8 062 T (+8.2%) et **Bridelice** 6 755 T (-13.4%)



FROMAGES FRAIS

FROMAGES FRAIS PAR FAMILLE

POIDS DES FAMILLES EN VOLUME

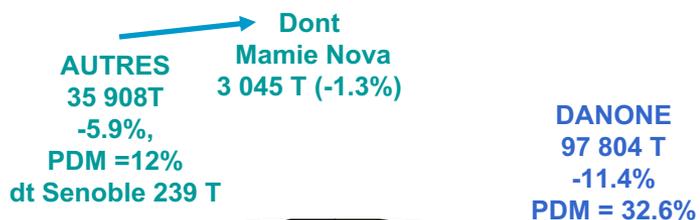
Volume total UF : 299 649 T

Evolution: - 5.0%

POIDS DES FAMILLES EN VALEUR

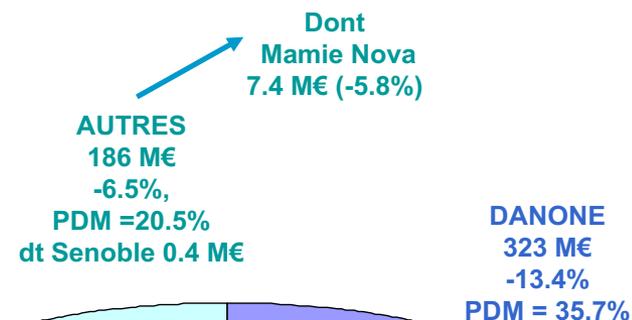
Valeur totale UF : 905.8 M€

Evolution: - 7.8%



MDD
114 882 T
+2.1%
PDM =38.3%

Dont MDD 94 777 T (- 1.9%)
1er prix 20 105 T (+41%)



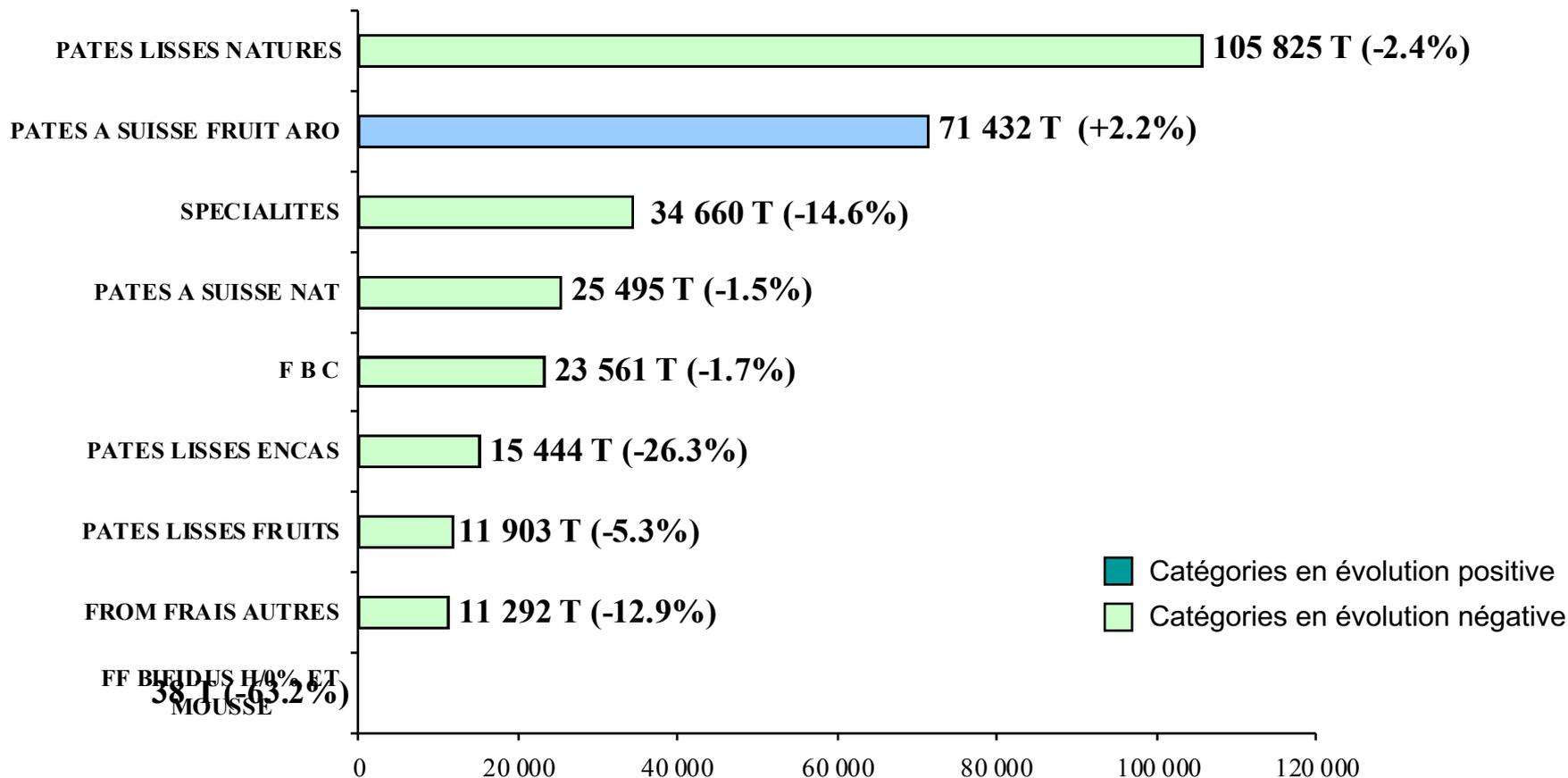
MDD
235 M€
-0.9%
PDM =25.9%

Dont MDD 200 M€ (- 3.4%)
1er prix 35 M€ (+ 76.8%)

FROMAGES FRAIS EN VOLUME

POIDS DES FROMAGES FRAIS : 299 649 T

EVOLUTION : - 5.0 %



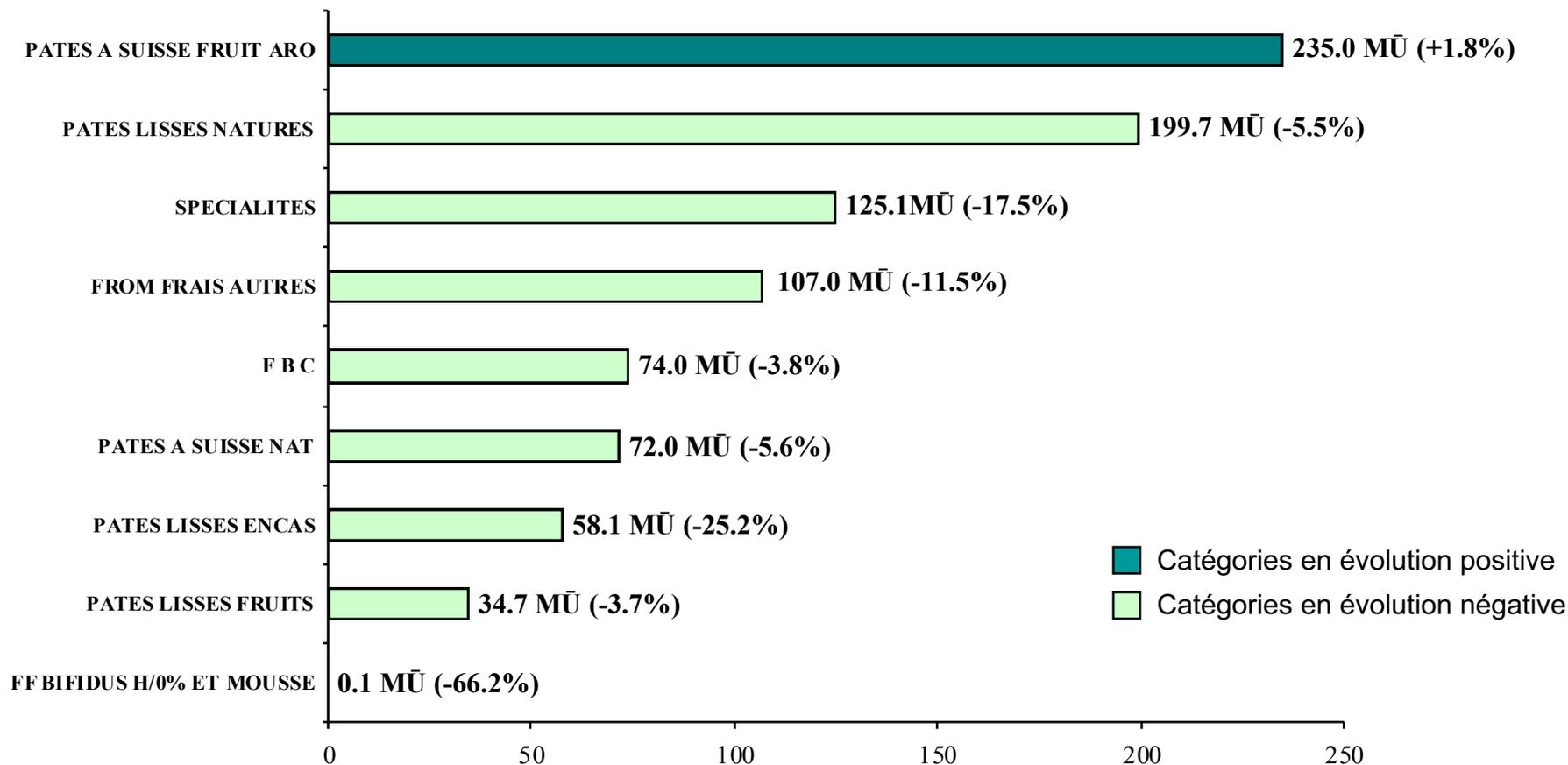
NB : Spécialités Hors Snacking = Gervita, Fjord, Perle de Lait

Pâtes Lisses Snacking = Petit Encas, Petite Faim, Encas MDD

FROMAGES FRAIS EN VALEUR

POIDS DES FROMAGES FRAIS : 905.8 M€

EVOLUTION: - 7.8 %



NB : Spécialités Hors Snacking = Gervita, Fjord, Perle de Lait

Pâtes Lisses Snacking = Petit Encas, Petite Faim, Encas MDD

FROMAGES FRAIS PROGRESSIONS

	Volume en Tonnes	Evolution volume	Evolution valeur	Tonnage MDD (en tonnes)	PDM MDD (en volume)
PATES A SUISSE FRUIT ARO	71 432	2.2%	+ 1.8%	18 787	26.3%

FROMAGES FRAIS REGRESSIONS

	Volume en Tonnes	Evolution volume	Evolution valeur	Tonnage MDD (en tonnes)	PDM MDD en volume
PATES A SUISSE NAT	25 495	-1.5%	- 5.5%	9 881	38.8%
F B C	23 561	-1.7%	- 3.8%	7 949	33.7%
PATES LISSES NATURES	105 825	-2.4%	- 5.5%		0.0%
PATES LISSES FRUITS	11 903	-5.3%	- 3.7%		0.0%
FROM FRAIS AUTRES	11 292	-12.9%	- 11.5%	1 985	17.6%
SPECIALITES	34 660	-14.6%	- 17.5%	1 479	4.3%
PATES LISSES ENCAS	15 444	-26.3%	- 25.2%	1 836	11.9%

FROMAGES FRAIS SNACKING ET H/SNACKING

POIDS DES « SNACKING » EN VOLUME

Volume total : 15 444 T

Évolution: - 26.3%

- **Danone Petit Encas : 13 607 T (-27.9%)**
 - dont *Stracciatella* 2 925 T (-36.0%)
 - Nature* 1 772 T (-24.9%)
- **MDD : 1 836 T (-12.4%)**

POIDS DES SPECIALITES « HORS SNACKING » EN VOLUME

Volume total : 34 660 T

Évolution: - 14.6 %

- **Danone**
 - *Gervita* : 7 924 T (-15.5%)
 - *Fjord* : 10 396 T (-14.1%)
 - *Taillefine Mousse 0%* : 1 923 T (-36.1%)
- **Yoplait *Perle de lait std* : 12 905 T (-9.1%)**
 - *Aloé Vera* : 195 T (Février 05)
- **MDD : 1 479 T (+1.4%)**



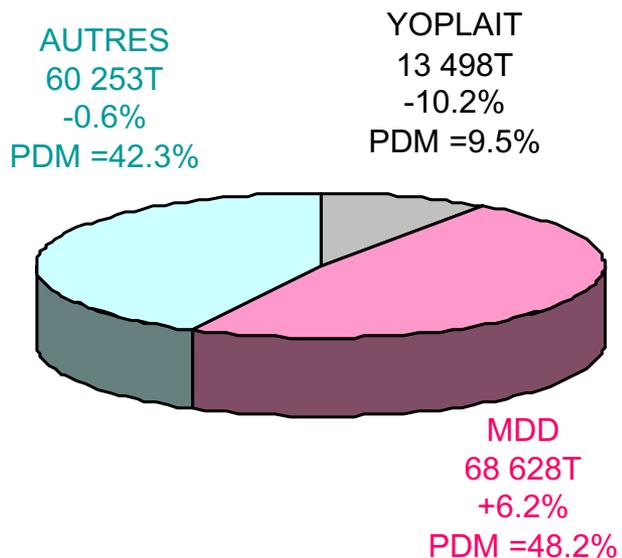
CREME FRAICHE

CREME FRAICHE

POIDS DES FAMILLES EN VOLUME

Volume total UF : 142 380 T

Evolution: + 1.5 %



POIDS DES FAMILLES EN VALEUR

Valeur totale UF : 454.9 M€

Evolution: - 1.2 %

