
**MODÉLISATION UPLIFT -
POUR AMÉLIORER LE CIBLAGE EN
TÉLÉMARKETING**

ACCÈSD - DESJARDINS

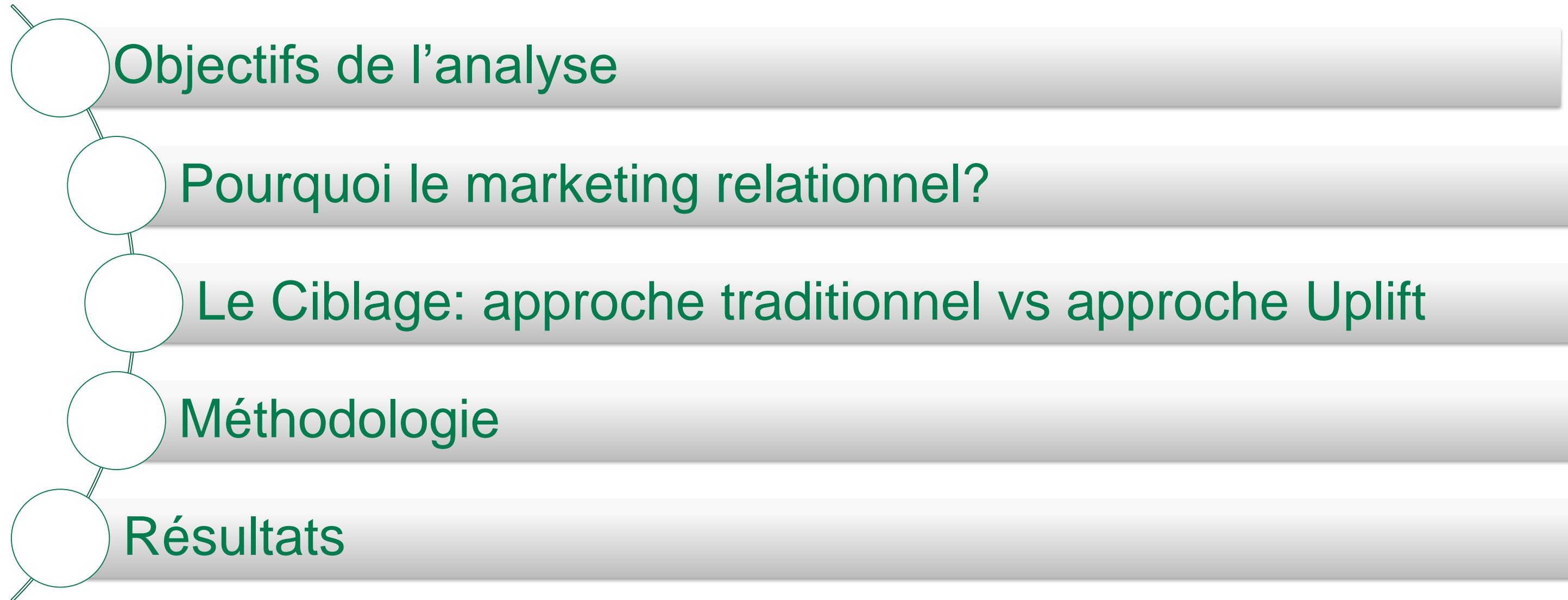
Milena Valverde - Finance et Intelligence d'Affaires.
VP AccèsD



Desjardins

Coopérer pour créer l'avenir

Agenda



Objectifs

- Déterminer le profil des membres qui ont plus de chance d'acquiescer le produit offert seulement s'ils sont contactés par l'équipe de télémarketing

Mise en contexte

Marketing relationnel

- Marketing relationnel - Marketing direct – marketing ciblé

communication directe avec les clients à partir d'une variété de canaux, à savoir le courriel, par la poste, télécommercialisation, textos, la publicité ciblée sur internet, etc.

- Importance des données des clients
potentiel des entités financières, accès aux informations sociodémographiques des clients et des informations financières (revenus, assurances, crédits, épargne)
- Datamining – Analytique – Modèles prédictifs pour augmenter **l'efficacité des campagnes de marketing**



LE BUT EST DE CIBLER
Les clients les plus susceptibles d'acquérir le produit offert seulement s'ils font l'objet d'une intervention marketing

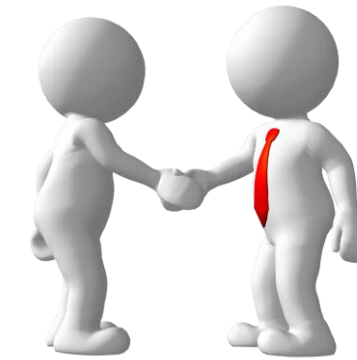
Processus télémarketing



Production
de listes de
ciblage



Appels aux
clients



Taux de
closing



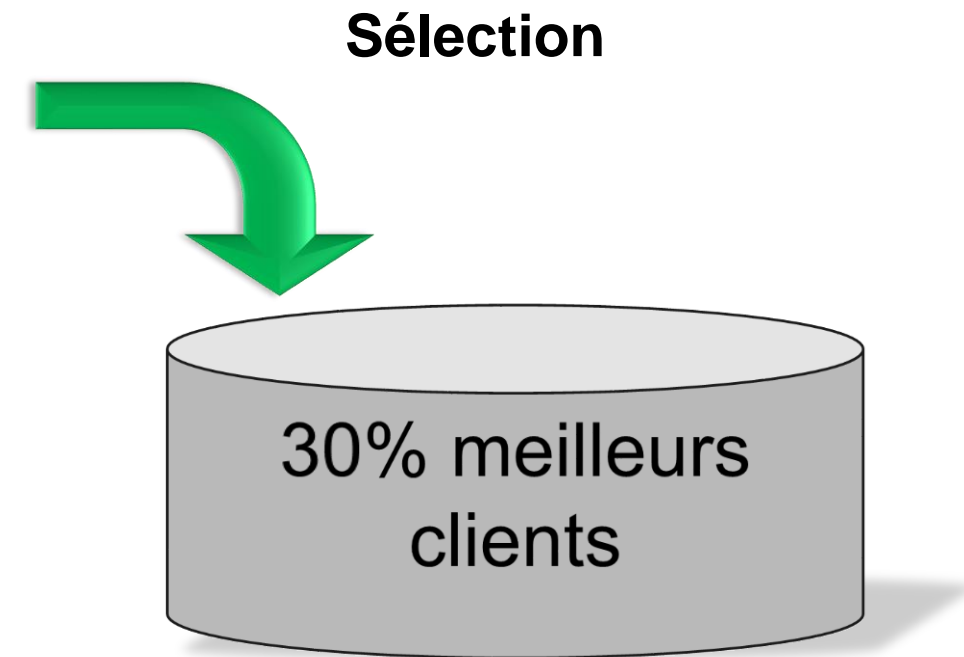
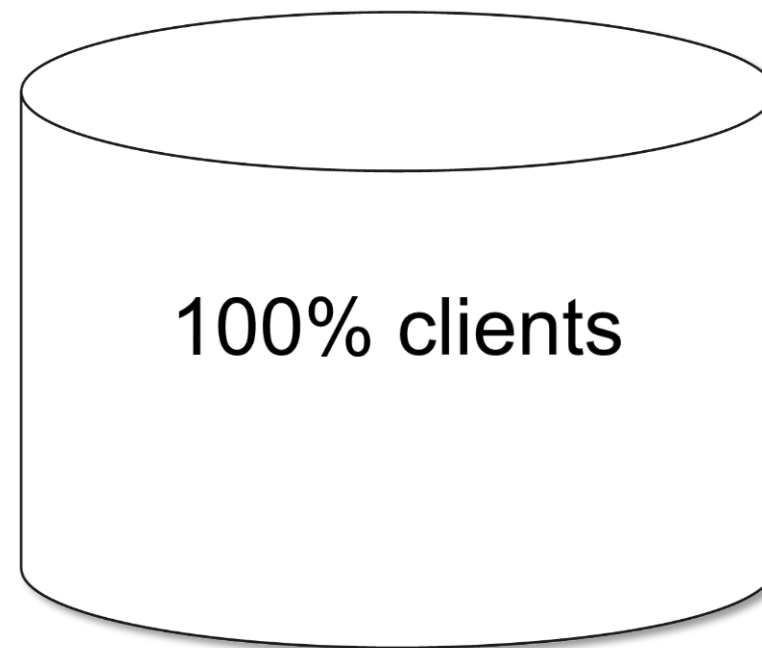
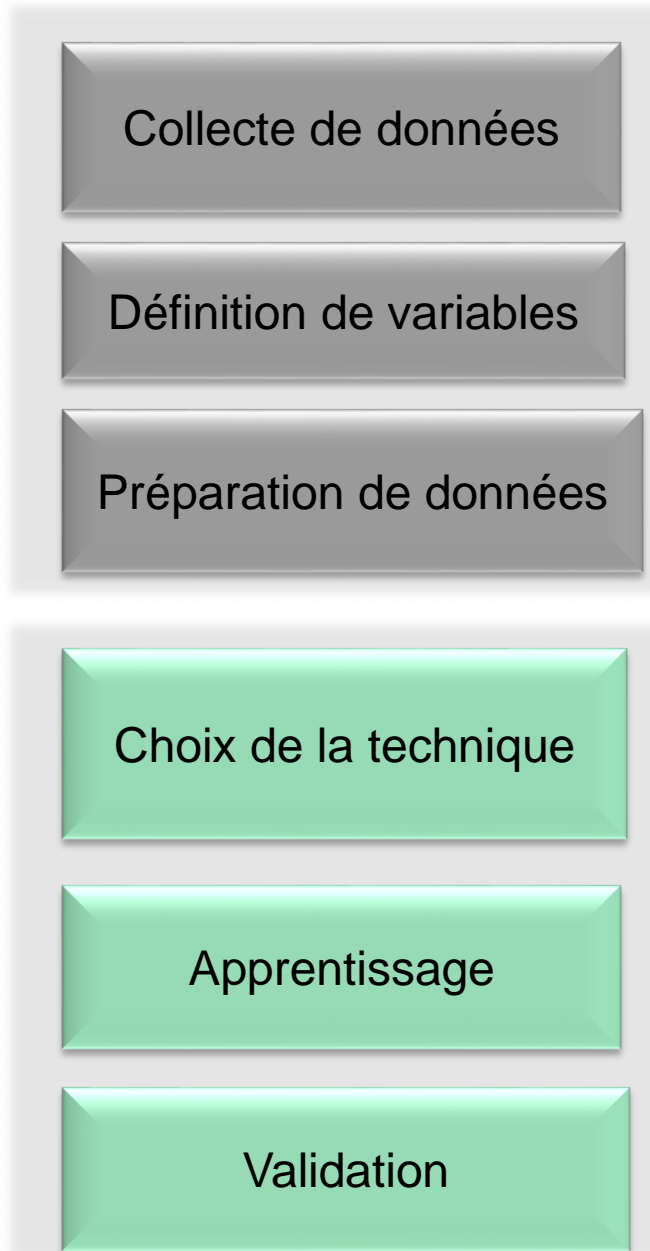
Augmenter le retour sur l'investissement en améliorant le ciblage.

Ciblage des clients

Les étapes du ciblage

LES ÉTAPES DU CIBLAGE

Modélisation



Techniques de ciblage

Il existe une subtile différence entre deux principes de ciblage d'une campagne commerciale:

Approche
traditionnel

- La sélection des personnes à forte probabilité d'achat, si elles sont intégrées dans une campagne de marketing.

Approche
Uplift

- La sélection des personnes à forte probabilité d'achat, **seulement** si elles sont intégrées dans une campagne de marketing.

Le second principe permet d'augmenter les résultats de ventes additionnelles générées par la campagne.

Approche traditionnelle:

On veut prédire:

Y = le membre a acquis le produit offert pendant la période d'intervention «Oui» ou «Non»

en fonction de plusieurs variables dépendantes: variables sociodémographiques, information financière, détention de produits.

Il s'agit de connaître les facteurs associés à un phénomène

Cette approche (scoring par la propension à consommer) a tendance à concentrer la sélection sur les personnes qui auraient acheté, quelles que soient les circonstances.

Ces modèles peuvent être améliorés **en utilisant un groupe témoin.**

Il serait mieux de mesurer l'impact d'intervention marketing?

- OUI, parce que le fait qu'un membre ait acquis le produit offert pendant ou à la suite d'un appel de télémarketing ne veut pas nécessairement dire que la campagne a eu un effet sur lui.
- On se demande, ce qui se serait arrivé si cet appel n'avait pas eu lieu ?

Est-ce que le membre aurait acquis le produit ?

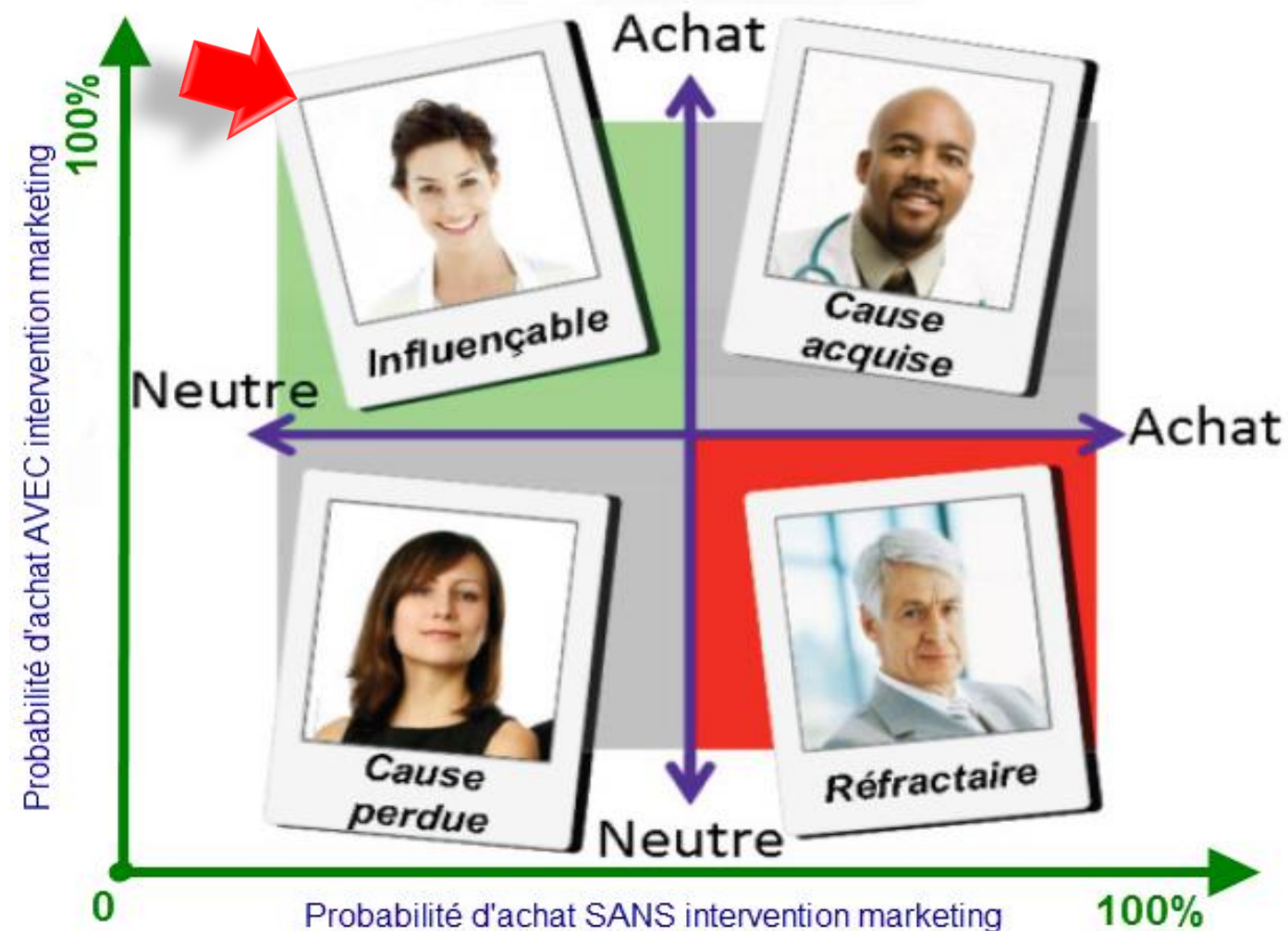


Approche Uplift

Segmentation des clients



Figure 2. Segments des clients selon leur réaction aux interventions marketing (9)



- **Influençable:**
Il est plus susceptible d'acheter s'il est traité
- **Cause acquise:**
Il est très probable qu'il achète s'il est traité ou pas.
- **Cause perdue:**
Il est peu probable qu'il achète s'il est traité ou pas
- **Réfractaire:**
Il est découragé d'acheter s'il est traité.

Importance d'utiliser un groupe témoin

Le groupe témoin ou contrôle est un ensemble de personnes qui répondent à tous les critères de sélection pour être ciblées, mais qui ont été délibérément omis de la campagne

Il fait partie des bonnes pratiques de ciblage en marketing:

- Il permet de mesurer l'efficacité des campagnes de Marketing
- Il permet de détecter leurs éventuels impacts négatifs

Groupe témoin

- Membres pas sollicités
- 255 107 membres

Groupe traitement

- Joints par télémarketing
- 24 404 membres

Approche Uplift pour trouver le segment des Influençables

«l'impact de l'intervention marketing»

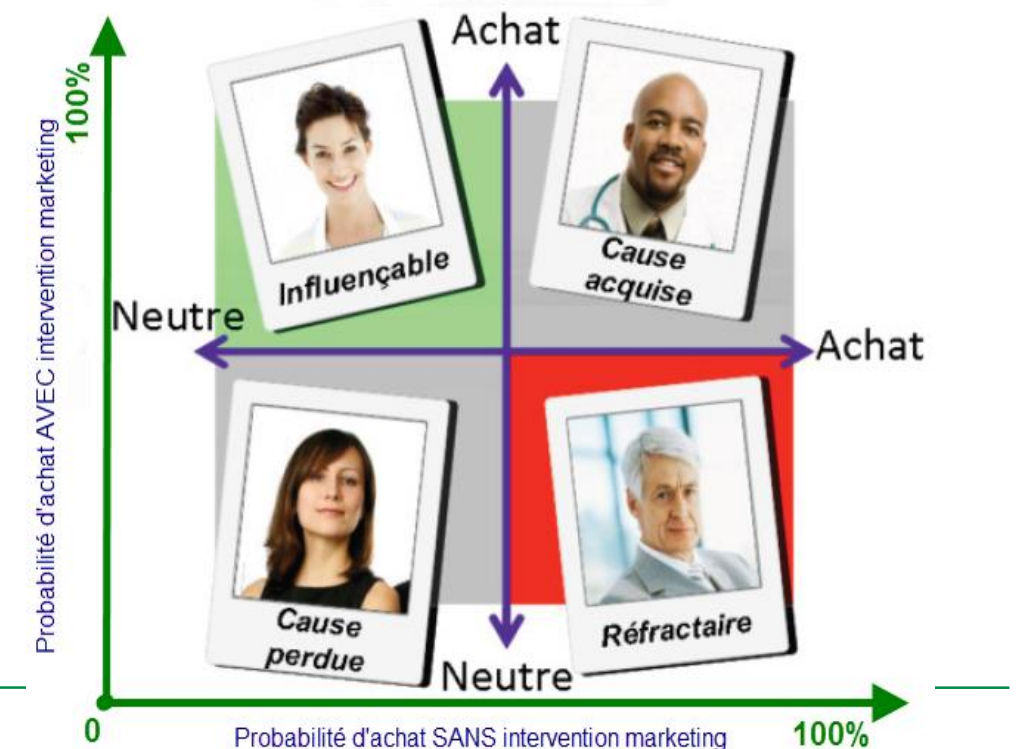
=

Chances d'acheter si le membre est joint

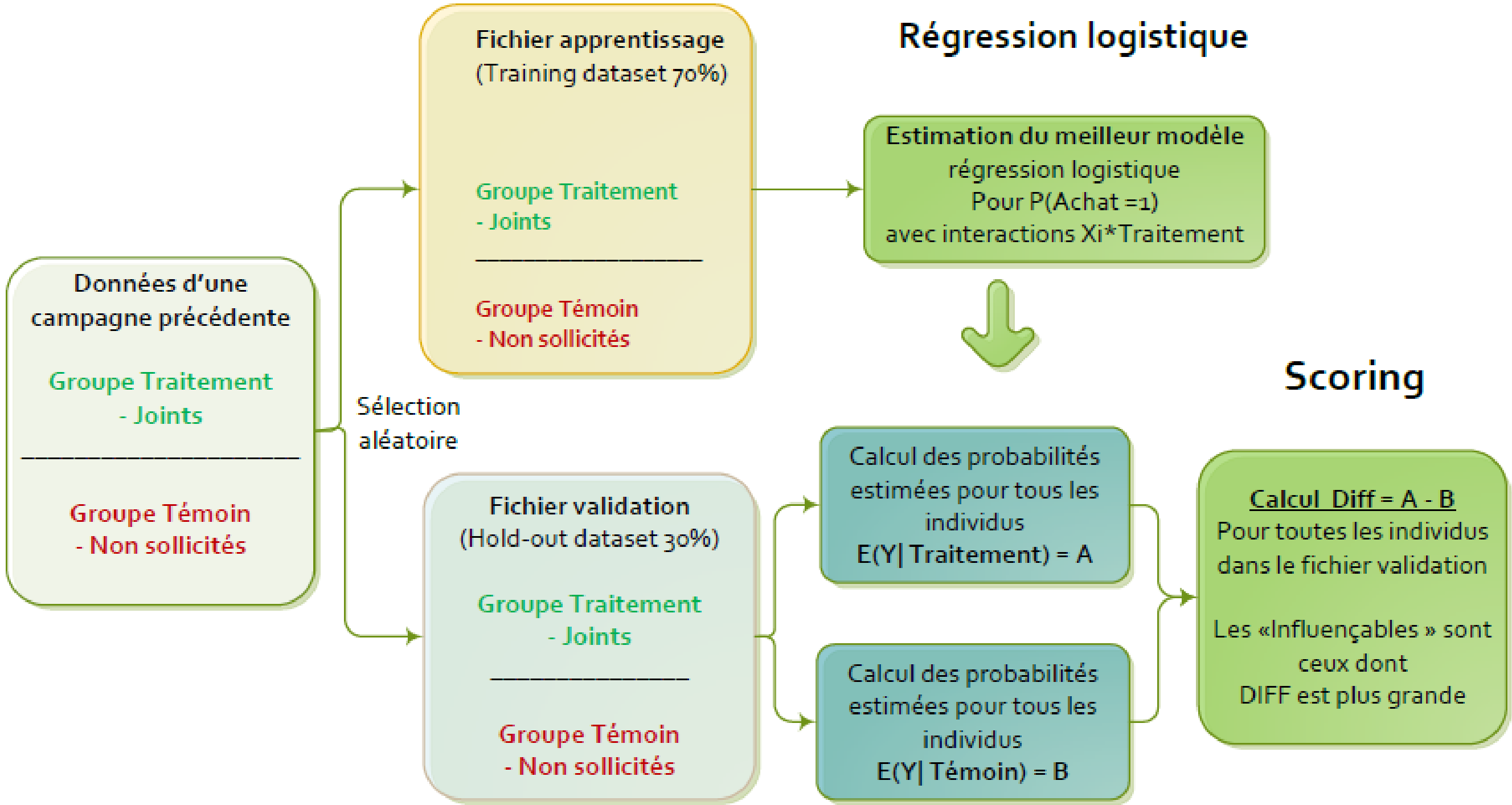
-

Chances d'acheter si le membre n'est pas sollicité

Les valeurs plus élevées de cette différence représentent les membres pour lesquels l'intervention marketing fait augmenter le plus de chance d'acquérir le produit offert. Donc, ils représentent le segment «**Influençables**».



Méthodologie Uplift (Lo, 2002)



Division de l'échantillon aléatoirement

```
/* DIVISER L'ENSEMBLE DE DONNÉES EN TRAINING DATASET AND HOLD-OUT DATASET*/
```

```
DATA TMP;
```

```
SET LIB.ANALYSE;
```

```
ALEAT =rand("Uniform");
```

```
RUN;
```

```
PROC SORT DATA = TMP; BY ALEAT; RUN;
```

Tri aleatoire

```
DATA TRAINING;
```

```
SET TMP;
```

```
IF _N_ < (279511*.70) THEN DATASET="Training";
```

```
else dataset="Validation";
```

```
run;
```

```
;
```

70% FichierTraining

Estimation du meilleur modèle

Variables disponibles

Variable réponse

- Le membre a acquis le produit offert pendant la période d'intervention ?
 - Oui (1)
 - Non (0)

Intervention Marketing

- Traitement ou témoin
 - Traitement (1)
 - Témoin (0)

Variables explicatives

- Âge
- Sexe
- Niveau de richesse
- Cycle de vie
- Détention de produits de financement
- Détention de produits d'épargne
- Détention de produits d'assurance
- Montants: crédits, épargne
- Joint par des autres canaux (par la poste, courriel,...)
- etc.

Sélection du meilleur modèle

- Régression logistique
- Inclure les variables explicatives et l'interaction de ces variables explicatives avec le traitement

Procédure SAS : sélection Stepwise

```
PROC LOGISTIC DATA= LIB.ANALYSE;  
CLASS  
SEX (REF=LAST)  OPR (REF=FIRST) .../param=REF  
;
```

```
MODEL Y (REF='0') =  
TRT  ANC  SEX  OPR  ...  
TRT*ANC  TRT*SEX  TRT*OPR  TRT*JEU .../  
selection=stepwise slentry=.5 slstay=.5
```

Inclure tous les interactions
variable *Traitement



```
;  
RUN;
```


Scoring

Calcul des probabilités
estimées pour tous les
individus
 $E(Y| \text{Témoin}) = B$

```
DATA TRUELIFT_1; SET LIB.HOLD; TRT=0; RUN;  
  
PROC LOGISTIC DATA=LIB.TRAIN ;  
CLASS.../ param=REF;  
MODEL ... ;  
score data= TRUELIFT_0 out=estimates0;  
RUN;
```

Calcul des probabilités
estimées pour tous les
individus
 $E(Y| \text{Traitement}) = A$

```
DATA TRUELIFT_1; SET LIB.HOLD; TRT=1; RUN;  
  
PROC LOGISTIC DATA=LIB.TRAIN ;  
CLASS.../ param=REF;  
MODEL ... ;  
score data= TRUELIFT_1 out=estimates1;  
RUN;
```

Calcul Diff = A - B

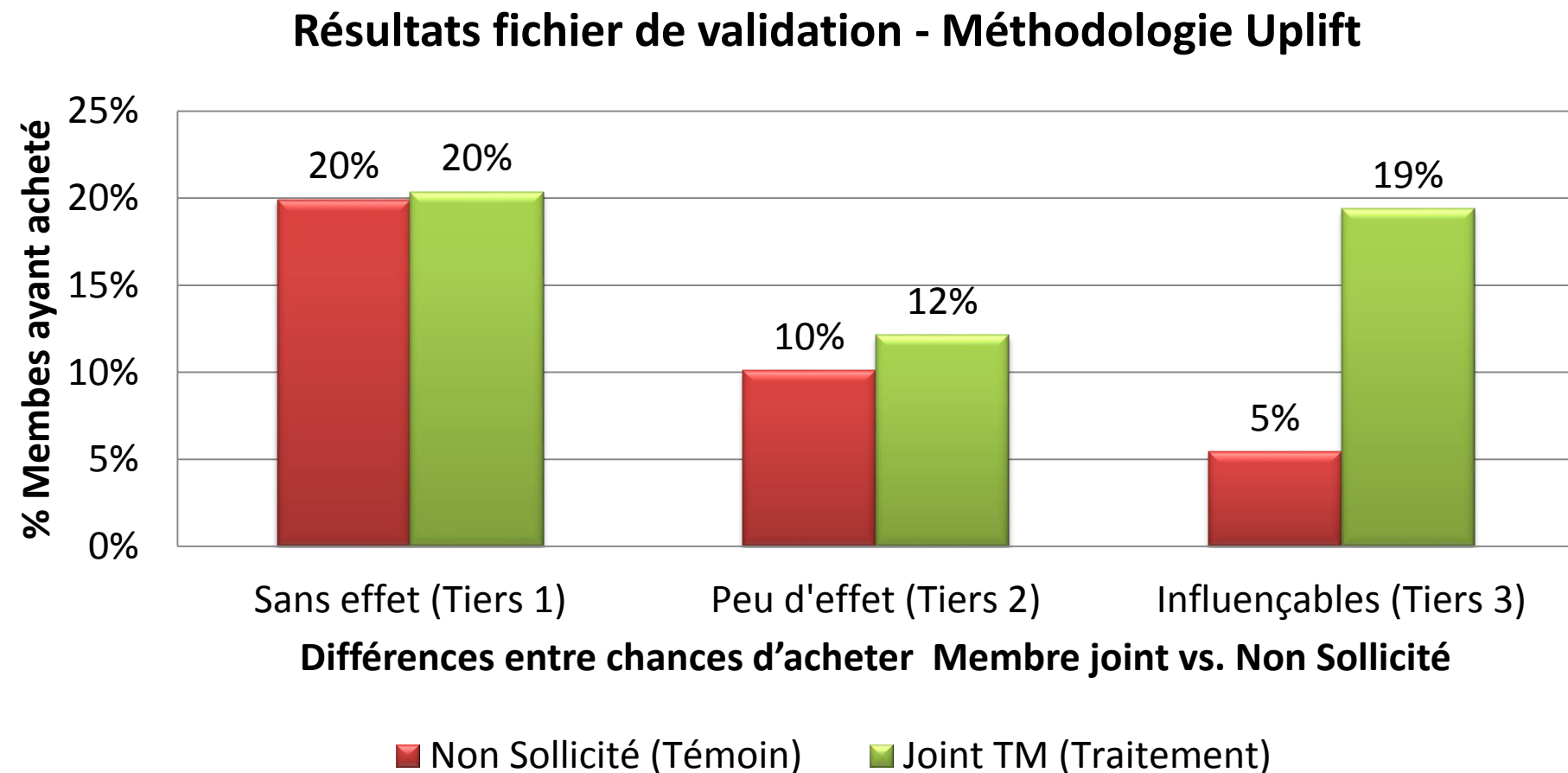
Pour toutes les individus
dans le fichier validation

Les «Influençables » sont
ceux dont
DIFF est plus grande

Résultats

Résultats scoring Méthodologie Uplift

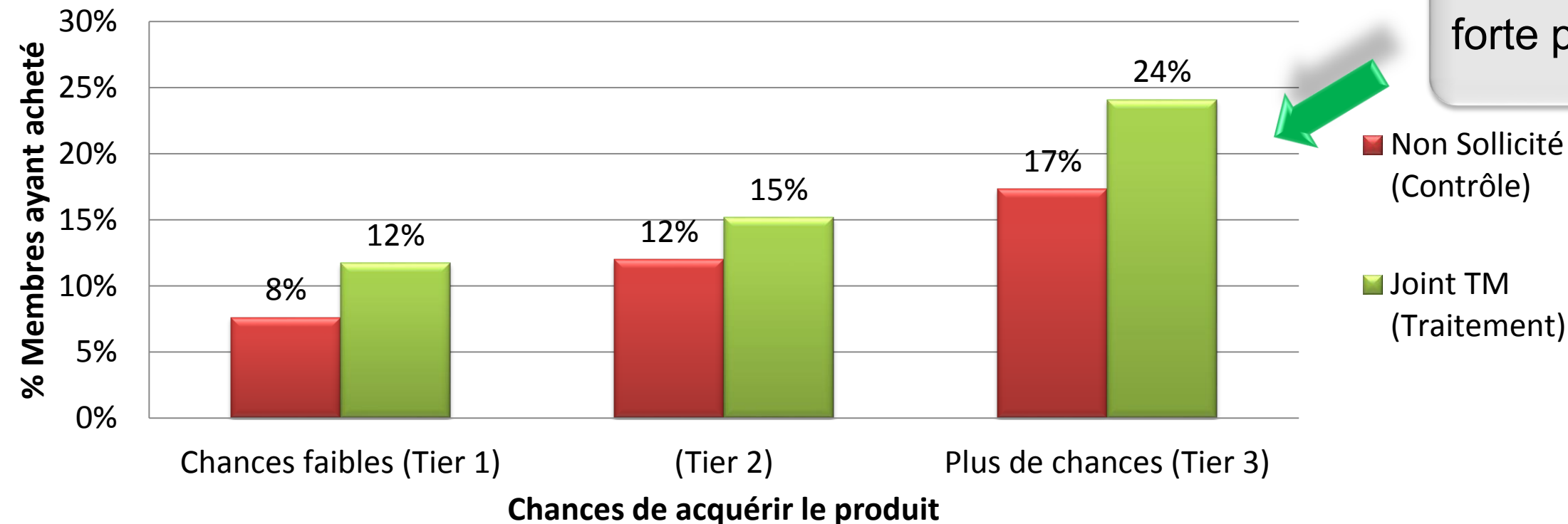
- fichier validation



Résultats scoring méthodologie traditionnelle

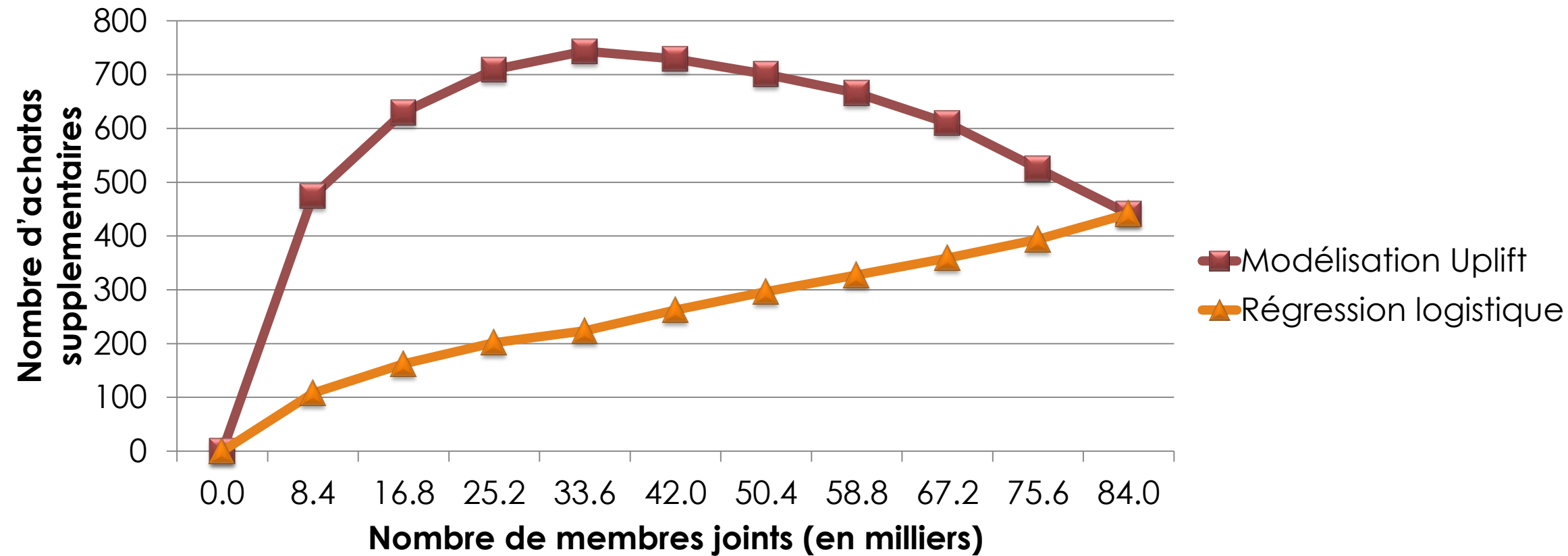
- fichier validation

Résultats fichier de validation - Méthodologie traditionnelle



- On remarque un taux de closing estimé de 24%, par contre, dans le groupe témoin les chances d'acquérir le produit est élevé (17%).
- On cible une portion de personnes qui auraient acheté même si l'intervention de marketing n'avait pas eu lieu

MESURE DE PERFORMANCE DU MODÈLE



Le coefficient Qini est une généralisation du coefficient Gini, adapté au cas où on utilise le groupe témoin. Essentiellement, on cherche à calculer les gains obtenus avec le modèle Uplift, donc le nombre de ventes supplémentaires en ciblant un groupe de clients.

Si on joint 8 400 membre et qu'on utilise:






- L'approche du modèle logistique, on aurait fait 108 ventes supplémentaires.
- L'approche modèle Uplift, on aurait fait 475 ventes supplémentaires.

Calcul de l'odds ratio interactions

Procédure SAS Exemple: Odds ratio des interactions

```
PROC LOGISTIC DATA= LIB.ANALYSE;  
...;  
lsmeans NRI*TRT/ ilink exp diff;  
slice NRI*TRT/ sliceby(NRI) diff exp;  
RUN;
```

Odds ratio des membres Joints versus Non Sollicités selon Niveau de richesse

Niveau de Richesse		Odds Ratio	P-value
1		5,7	✓ 0,0
2		3,1	✓ 0,0
3		2,5	✓ 0,0
4		1,8	✓ 0,0
5		1,4	👎 0,8



MESSAGES CLÉ

- Nous avons identifié un segment de clients qu'on doit sans doute cibler, les « **Influençables** », qui sont des clients plus susceptibles d'acquérir le produit offert seulement s'ils sont joints par le télémarketing.
- Les résultats nos emmènent aussi à **éviter de cibler** quelques segments des clients (*résultats confidentiels*)
- Nous avons remarqué l'importance d'inclure dans l'analyse le groupe témoin: Sans le groupe témoin, on ne serait pas capable de distinguer les «Influençables» des «Causes acquises». Cela veut dire qu'on ne pourrait pas savoir si le fait que le membre ait répondu positivement à l'intervention marketing est vraiment un effet de l'appel, ou si par contre c'est un effet de la consommation naturelle.

Références bibliographiques

- Y. Lo, 2002. The true lift model A Novel Data Mining Approach to Response Modeling. ACM SIGKDD Explorations Newsletter, 4(2):78–86.
- Generating Incremental Sales. Maximizing the incremental impact of cross-selling, up-selling and deep-selling through uplift modelling, Nicholas.Radcliffe.
- Julie Desjardins, L'analyse de régression logistique. Université de Montréal. Tutorial in Quantitative Methods for Psychology. 2005, Vol. 1(1), p. 35-41.

