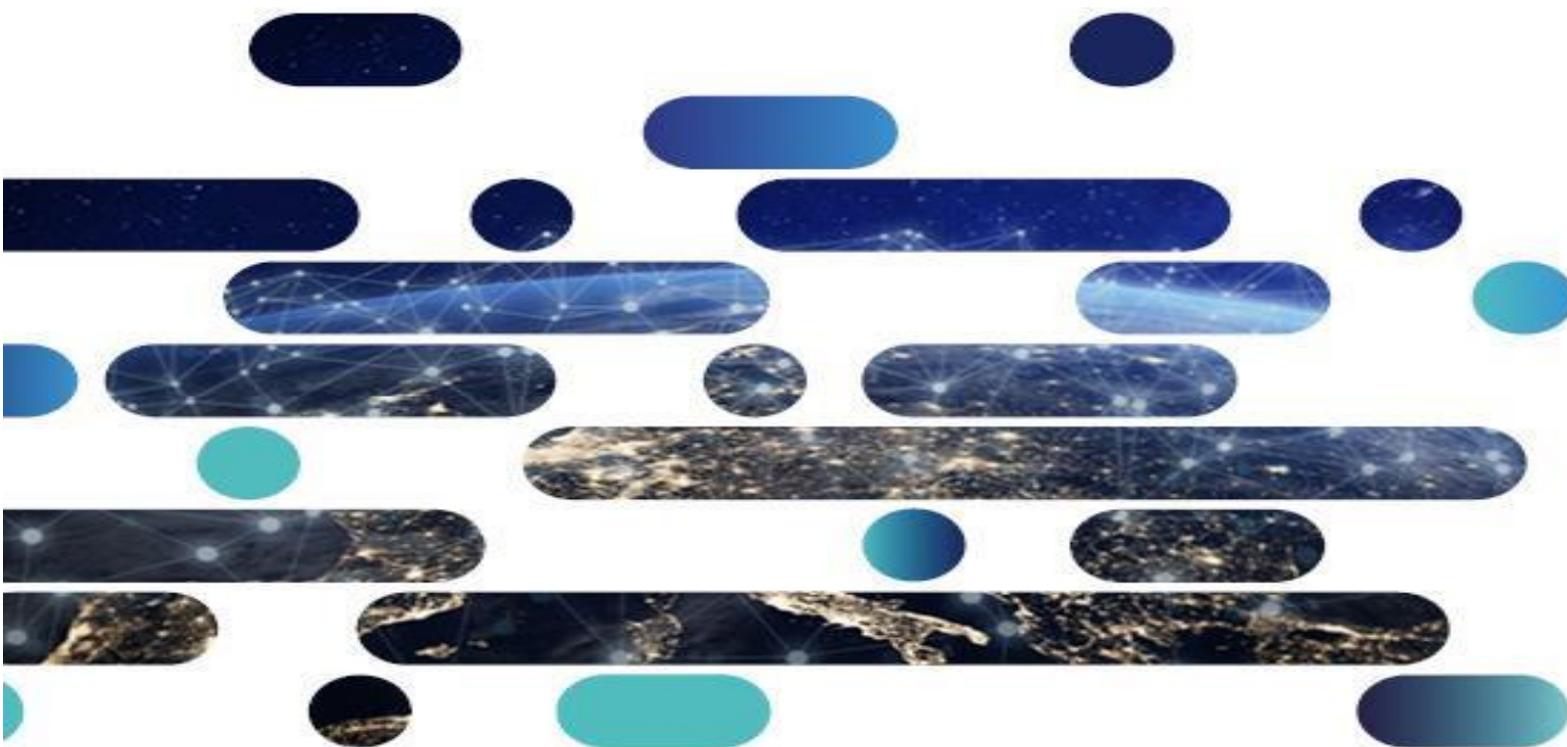


Les enjeux techniques et opérationnels d'une optimisation tarifaire.

Linda Krolikowski

Nabil Rachdi

- 1. Anti-sélection: mythe ou réalité?**
- 2. Biais de sélection et principe du jumelage par score de propension (PSM)**
- 3. Application du jumelage des sujets par score de propension**
- 4. Elasticité au prix individuelle et optimisation tarifaire dans un contexte concurrentiel**



Anti-sélection: mythe ou réalité?

Comment déterminer la marge?

Marge

Prix de vente

Fixé en selon la stratégie de l'entreprise

Coût de production

En assurance, on ne connaît le coût de production au moment de la vente du produit

- Une bonne estimation du coût de production est indispensable pour toutes les industries.
- En assurance, les modèles mathématiques aident les actuaires à maîtriser ce coût



De bons modèles ne suffisent pas, il faut tenir compte de l'environnement de marché.

Anti-sélection : mythe ou réalité?



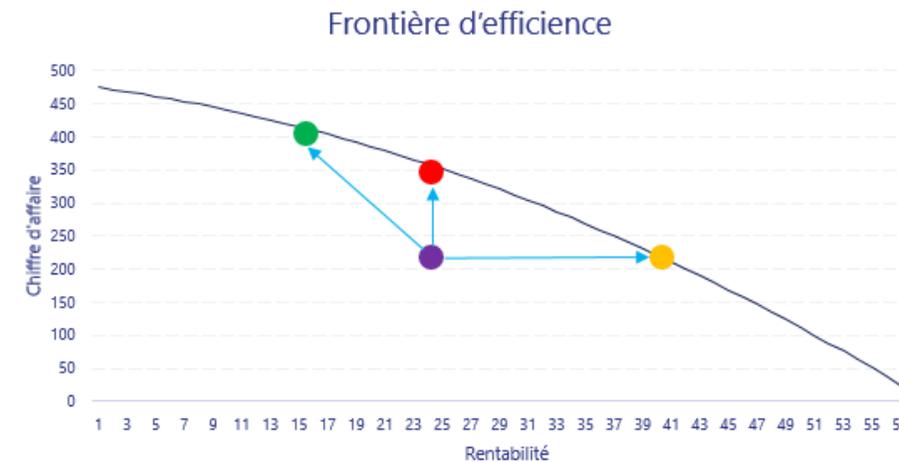
Simulation des indicateurs

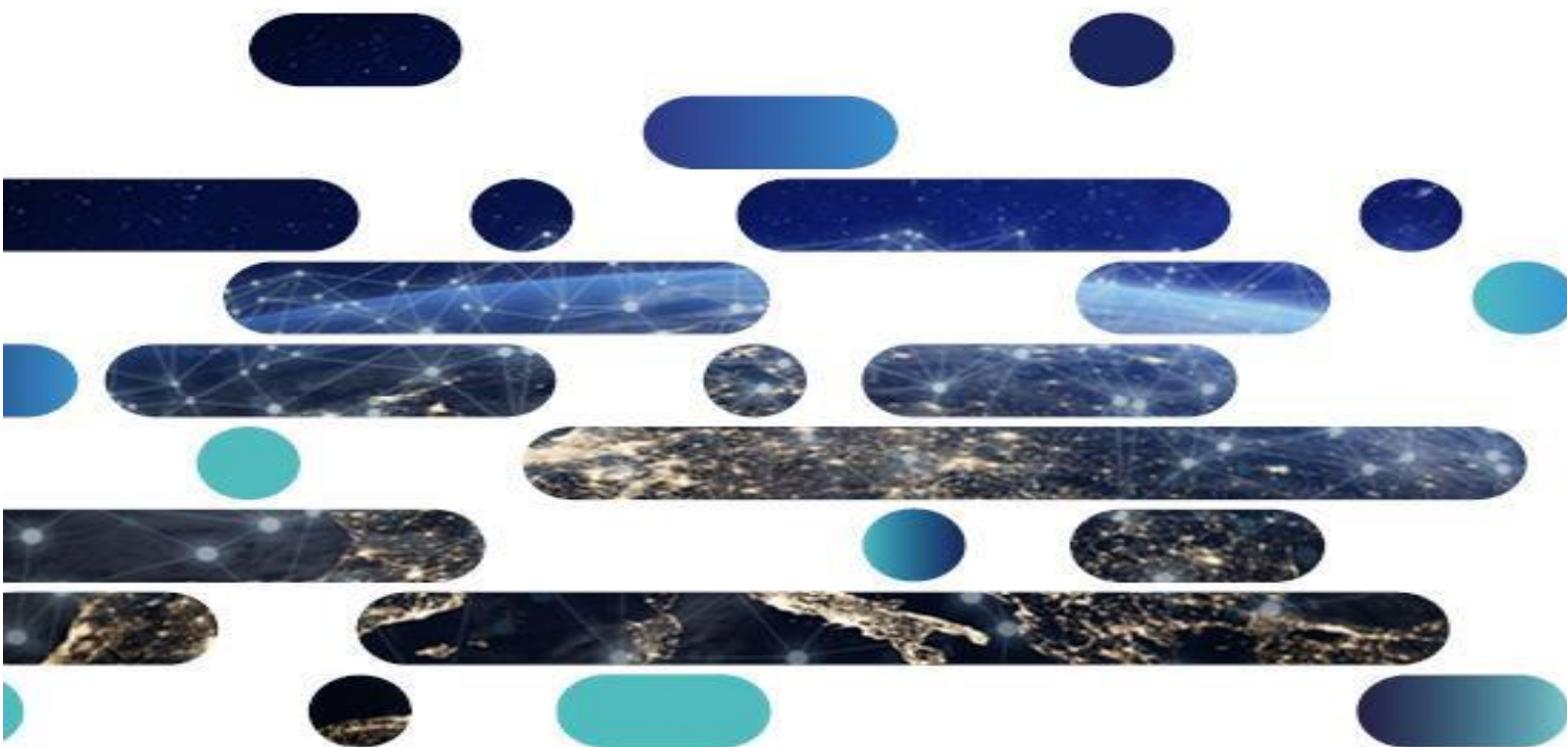
$$\text{Rentabilité}(P_i) = \sum_{i \leq n} (P_i - S_i) * \hat{f}(P_i, X_i)$$

Sous contrainte:
$$\sum_{i \leq n} \hat{f}(P_i, X_i) \geq \theta$$

- P_i : prix commercial HT que paie le client i où $i \leq n$ (n : nombre de prospects)
- S_i : la sinistralité du client i
- $\hat{f}(P_i, X_i)$: la probabilité de conversion du prospect i
- X_i : caractéristique du prospect i (âge, sexe, expérience, ...)

Ensemble des scénarios maximisant la rentabilité fixée





Principe du jumelage par score de propension

Principe du jumelage par score de propension

$$Y = \begin{cases} 1 & \text{si le devis est souscrit} \\ 0 & \text{sinon} \end{cases}$$



Probabilité de conversion

Quelle serait la probabilité de conversion du sujet si je modifie son tarif?

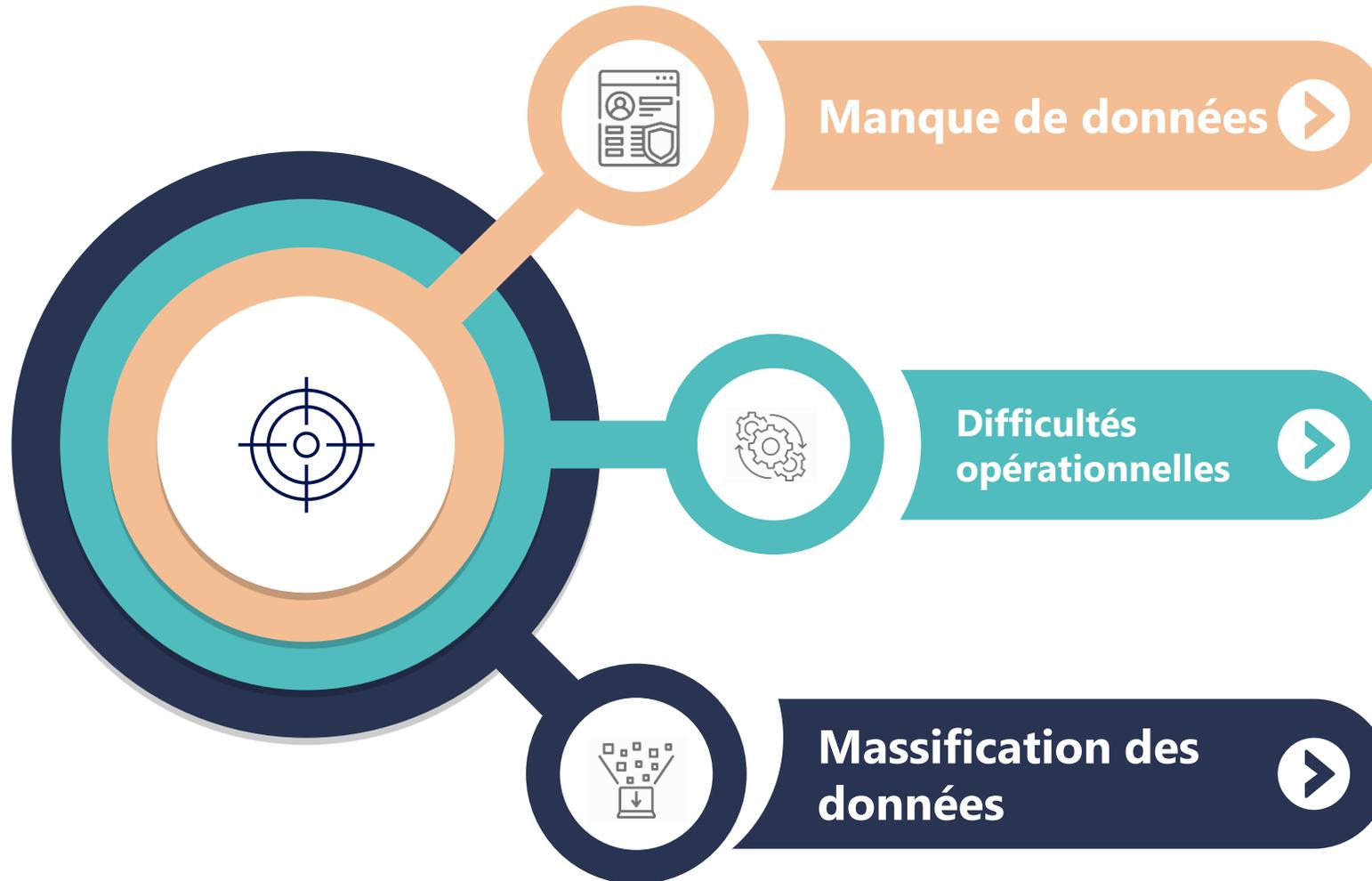
Propensity score matching



Mesure sans biais de l'élasticité-prix

$$Y_i = b_1 * X_{1i} + b_2 * X_{2i} + b_3 * P_i$$

Principe du jumelage par score de propension



Manque de données

Les profils de la base de données associés à la nouvelle politique tarifaire peuvent être éloignés en termes de caractéristiques du segment concerné, ce qui crée un biais de sélection.

Difficultés opérationnelles

Les A/B tests sont souvent compliqués à mettre en place et difficiles à exploiter.

Massification des données

Des techniques mathématiques issues de l'univers médical peuvent pallier aux contraintes opérationnelles des price tests. Il s'agit du jumelage de sujets par score de propension (PSM).

Jumelage des sujets : principe

Traitement : stratégie 1

Contrôle : stratégie 2

Taux de transformation moyen des individus de traitement (strat 2) < Taux de transformation moyen des individus de contrôle (strat 2)

Sujet	Stratégie	Âge	Csp	taux de transformation	Score de propension (probabilité d'appartenir à la stratégie 1)	Jumeau	taux de transformation_jumeau
1	1	22	Agriculteur	0.34	D = 1	4	0.30
2	1	24	Agriculteur	0.36		4	0.30
3	1	29	Ouvrier	0.43		5	0.38
4	2	26	Agriculteur	0.30	D = 0	-	-
5	2	33	Ouvrier	0.38		-	-
6	2	45	Ouvrier	0.40		-	-

➔ Il faut que le modèle qui estime le score de propension soit performant pour exclure les individus de contrôle dont le profil est trop éloigné de ceux du groupe de traitement.

➔ Il faut que des sujets du groupe de contrôle ressemblent aux profils du groupe de traitement, autrement le jumelage n'a pas de sens.

Jumelage des sujets : principe

Traitement : stratégie 1

Contrôle : stratégie 2

Taux de transformation moyen des individus de traitement (strat 2) = Taux de transformation moyen des individus de contrôle (strat 2)

Sujet	Stratégie	Âge	Csp	taux de transformation	Score de propension (probabilité d'appartenir à la stratégie 1)	Jumeau	taux de transformation_jumeau
1	1	22	Agriculteur	0.34	0.8	4	0.30
2	1	24	Agriculteur	0.36	0.75	5	0.38
3	1	29	Ouvrier	0.43	0.3	6	0.40
4	2	26	Agriculteur	0.30	0.95	-	-
5	2	33	Ouvrier	0.38	0.5	-	-
6	2	45	Ouvrier	0.40	0.2	-	-

➡ Il faut que le modèle qui estime le score de propension soit performant pour exclure les individus de contrôle dont le profil est trop éloigné de ceux du groupe de traitement.

➡ Il faut que des sujets du groupe de contrôle ressemblent aux profils du groupe de traitement, autrement le jumelage n'a pas de sens.

Jumelage des sujets : principe

Traitement : stratégie 1

Contrôle : stratégie 2

Taux de transformation moyen des individus de traitement (strat 2) << Taux de transformation moyen des individus de contrôle (strat 2)

Sujet	stratégie	âge	csp	taux de transformation	Score de propension	jumeau	taux de transformation_jumeau
1	1	22	Agriculteur	0.34	0.8	4	0.30
2	1	24	Agriculteur	0.36	0.75	4	0.30
3	1	29	Agriculteur	0.43	0.7	4	0.30
4	2	56	Ouvrier	0.30	0.3	-	-
5	2	43	Ouvrier	0.38	0.25	-	-
6	2	45	Ouvrier	0.40	0.2	-	-



Il faut que le modèle qui estime le score de propension soit performant pour exclure les individus de contrôle dont le profil est trop éloigné de ceux du groupe de traitement.



Il faut que des sujets du groupe de contrôle ressemblent aux profils du groupe de traitement, autrement le jumelage n'a pas de sens.

Jumelage des sujets : principe

Traitement : stratégie 1

Contrôle : stratégie 2

Sujet	stratégie	âge	csp	taux de transformation	Score de propension	jumeau	taux de transformation_jumeau
1	1	22	Agriculteur	0.34	0.8	4	0.30
2	1	24	Agriculteur	0.36	D = 1	4	0.30
3	1	29	Ouvrier	0.43	0.6	5	0.38
4	2	26	Agriculteur	0.30	0.73	-	-
5	2	33	Ouvrier	0.38	D = 0	-	-
6	2	45	Ouvrier	0.40	0.2	-	-

	Traitement	Contrôle avant matching	Contrôle après matching
Agriculteur	67%	33%	67%
Ouvrier	33%	67%	33%
	Traitement	Contrôle avant matching	Contrôle après matching
Age (moyenne)	25	34,67	28,33

Jumelage des sujets : principe

Individu i appartenant à la stratégie 1

→ Ligne initiale présente dans la base de données. Le taux de transformation du sujet i est estimé à 40% par le modèle.

Sujet	âge	CSP	INSEE	voiture	Stratégie	Tarif	Marge	Score de transformation
i	30	cadre	93100	peugeot	1	100	10	40%
i	30	cadre	93100	peugeot	2	120	18	30%
i	30	cadre	93100	peugeot	3	130	25	25%

Duplication du sujet i dans le modèle de transformation global :

- mêmes caractéristiques clients
- tarifs différents
- probabilité de conversion différente

Taux de transformation estimé par le modèle pour les jumeaux du sujet i appartenant aux stratégies 2 et 3 respectivement.

Principe du jumelage par score de propension



Estimation du taux de transformation

- Sur la base de devis initiale.
- De la manière la plus exacte possible par XGBoost.
- Incorporation de données concurrentielles.

Constitution des groupes de politique tarifaire.

- Segmentation de la marge relative par CAH.

Jumelage des sujets

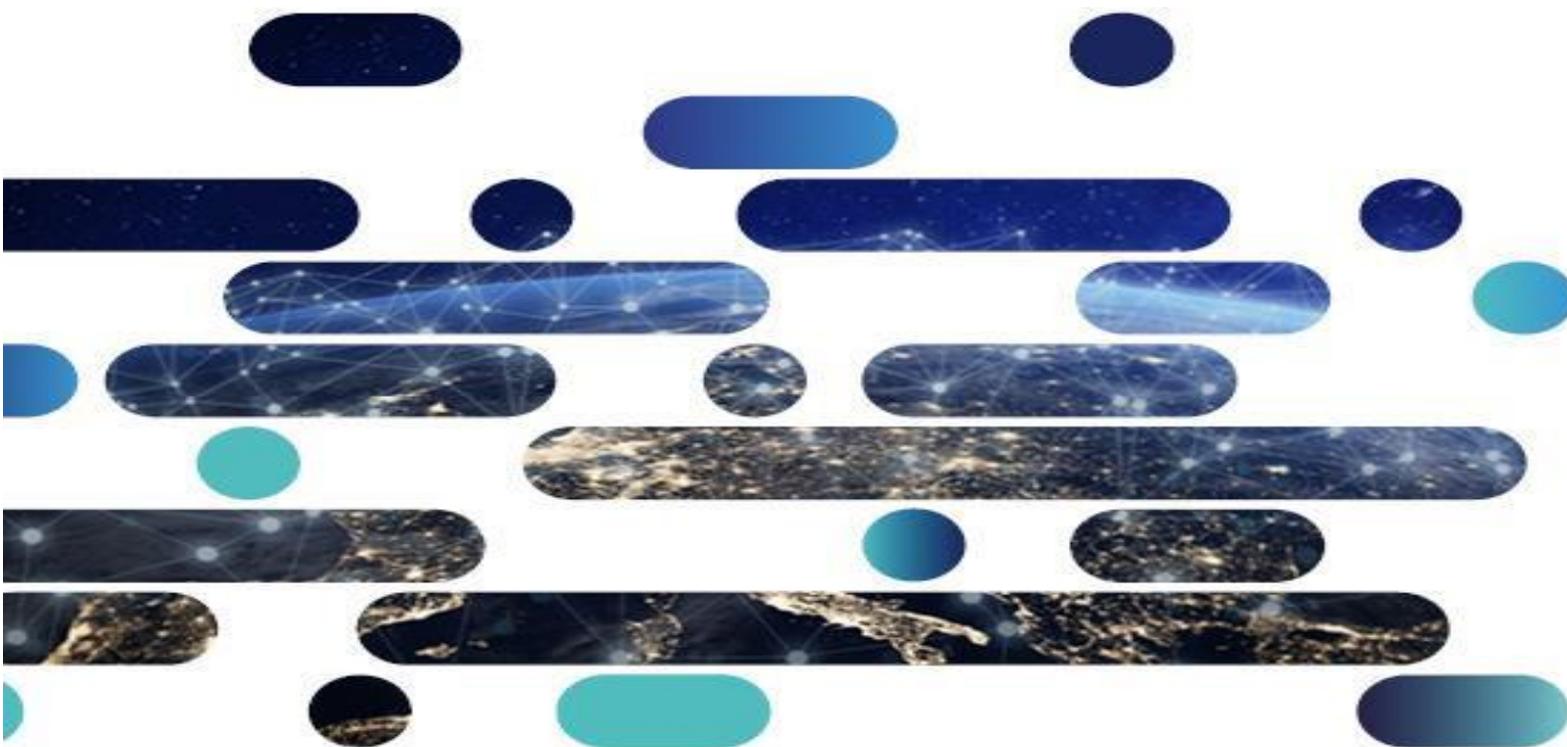
- Estimation du score de propension pour chaque couple de strategies.
- Appariement des données.
- Analyse des caractéristiques avant/après matching.

Modélisation de l'élasticité au prix

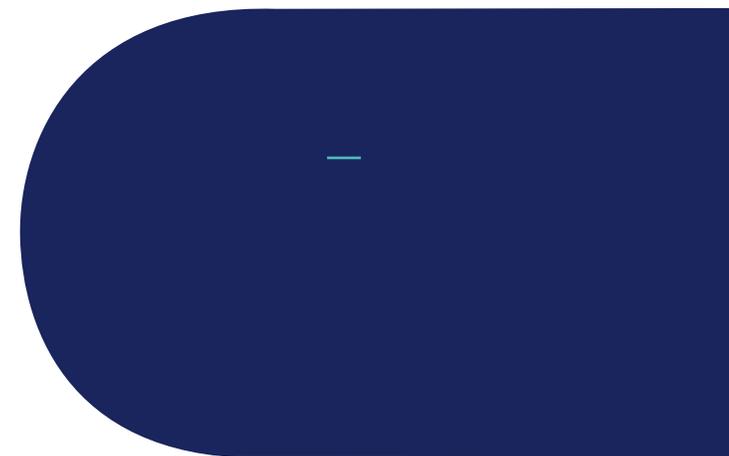
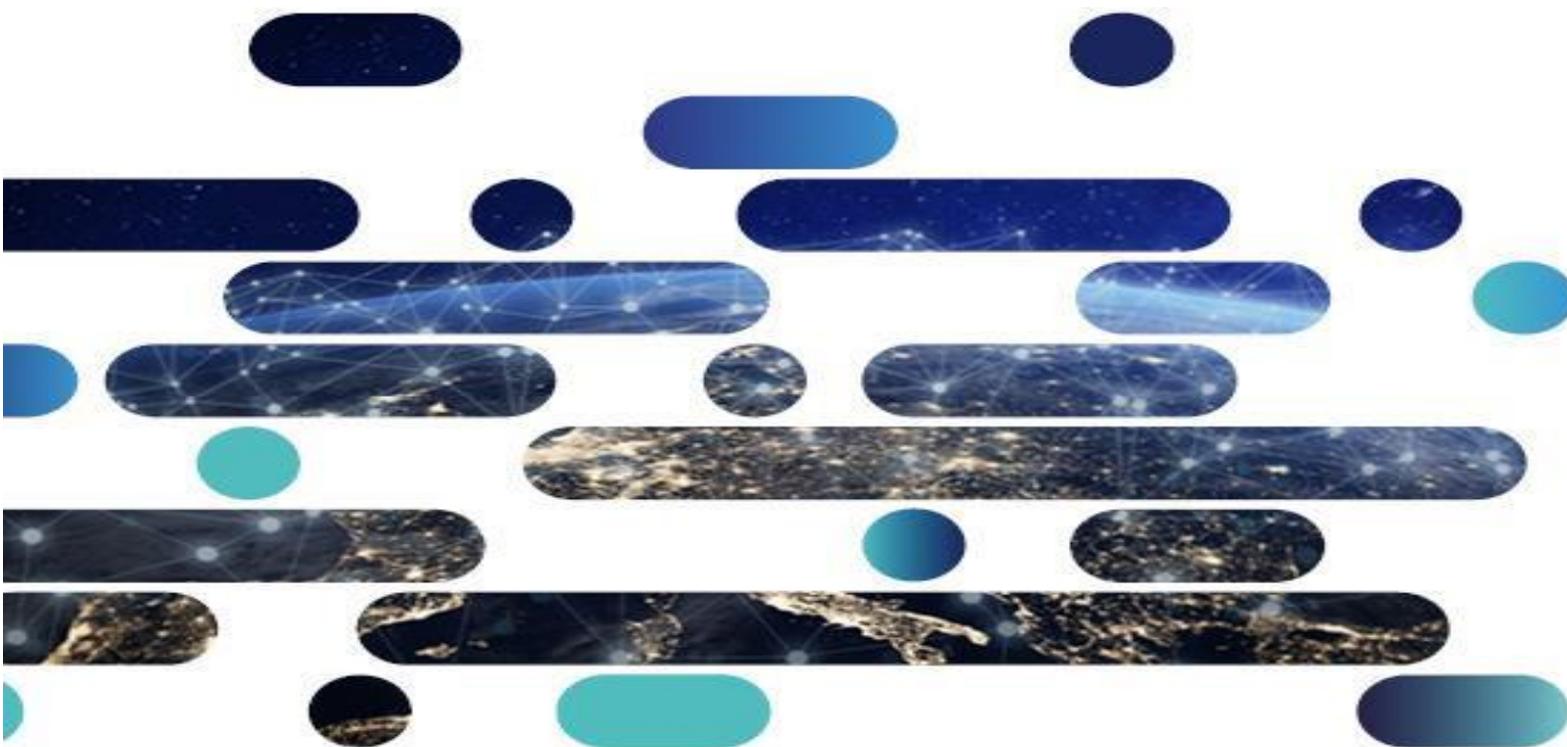
- Apprentissage sur la base devis alimentée des nouvelles données issues du jumelage.
- Modèle GLM pour satisfaire la propriété de dérivabilité.

Optimisation tarifaire





Application du jumelage des sujets par score de propension



Estimation du taux de transformation individuel sur la base devis initiale

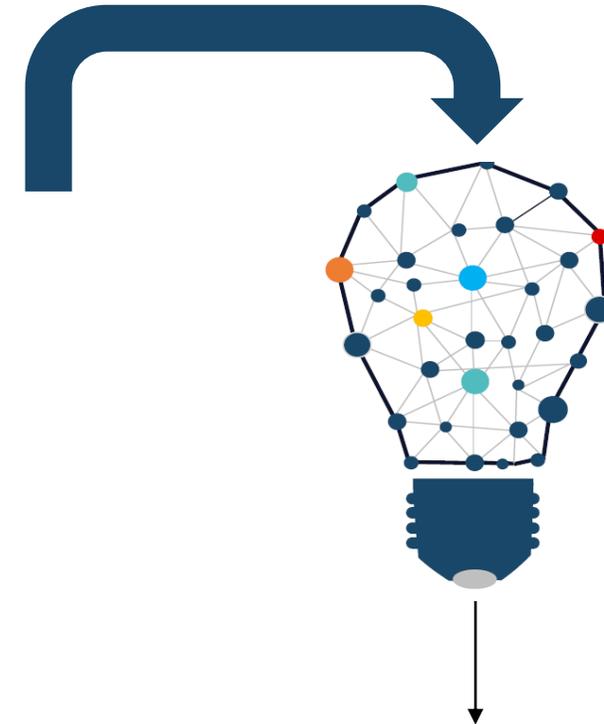
Modélisation du taux de transformation

Ce qui impacte la conversion individuelle

- Caractéristiques du conducteur, du véhicule
- Données sur le contrat
- Indices de compétitivité : tarif leader, tarif médian, marge (calculée à partir de la prime pure)

Benchmark marché sur certains profils +
mapping

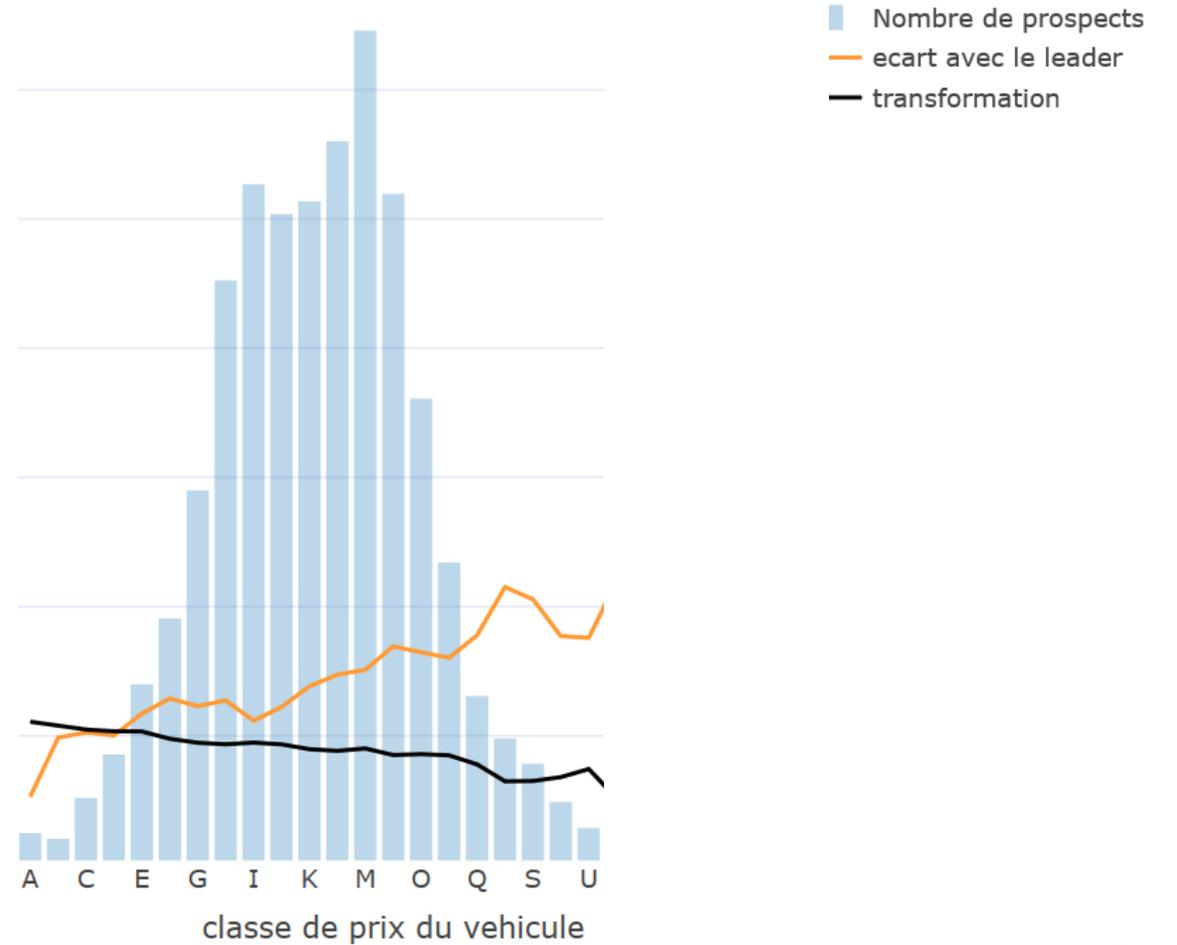
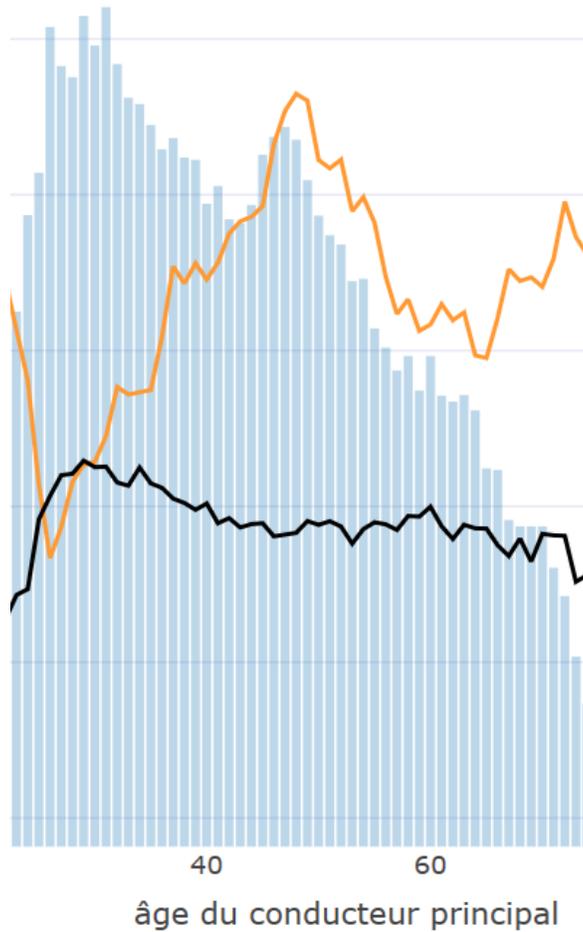
Modélisation XGBoost



Prédiction du tarif leader/médian pour chaque
prospect de la base devis de l'assureur

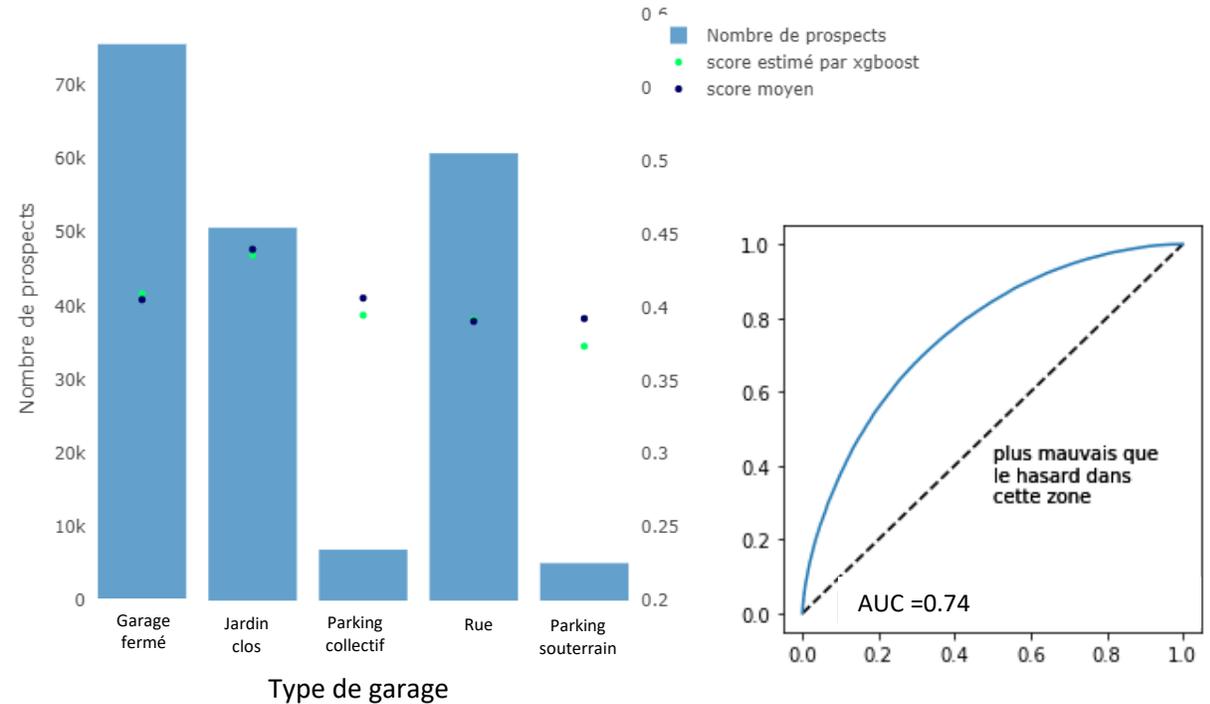
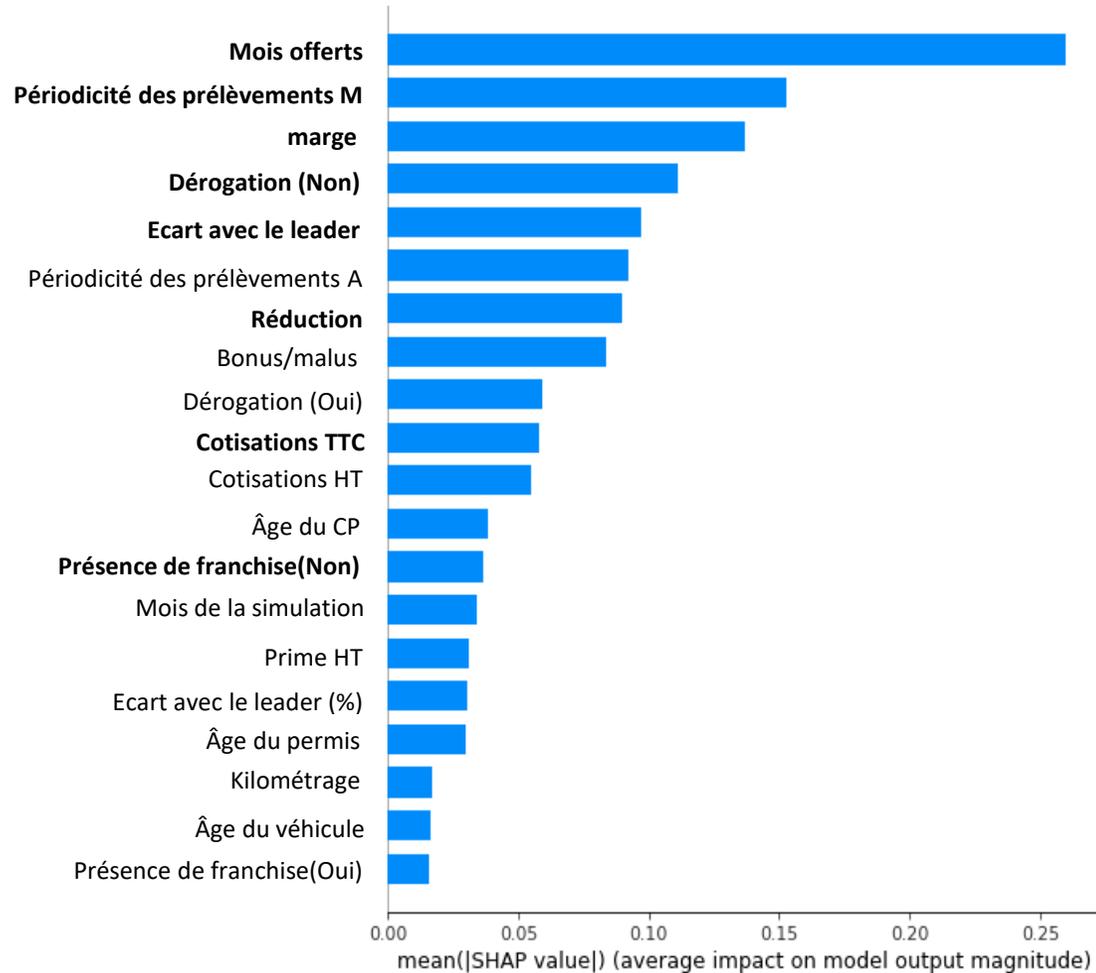
Modélisation du prix concurrent

Etude du positionnement

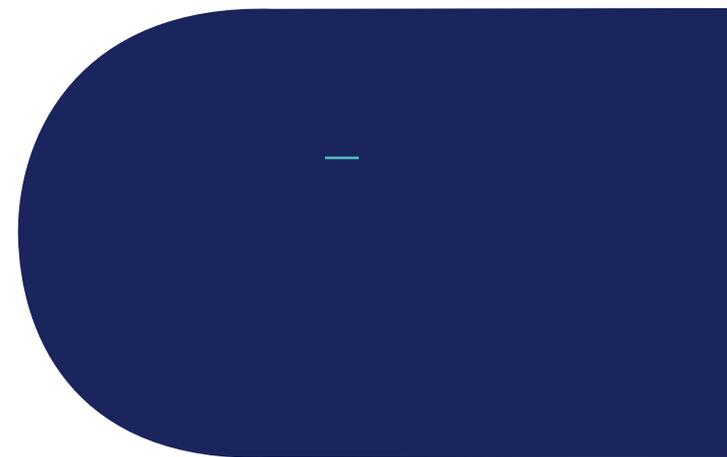
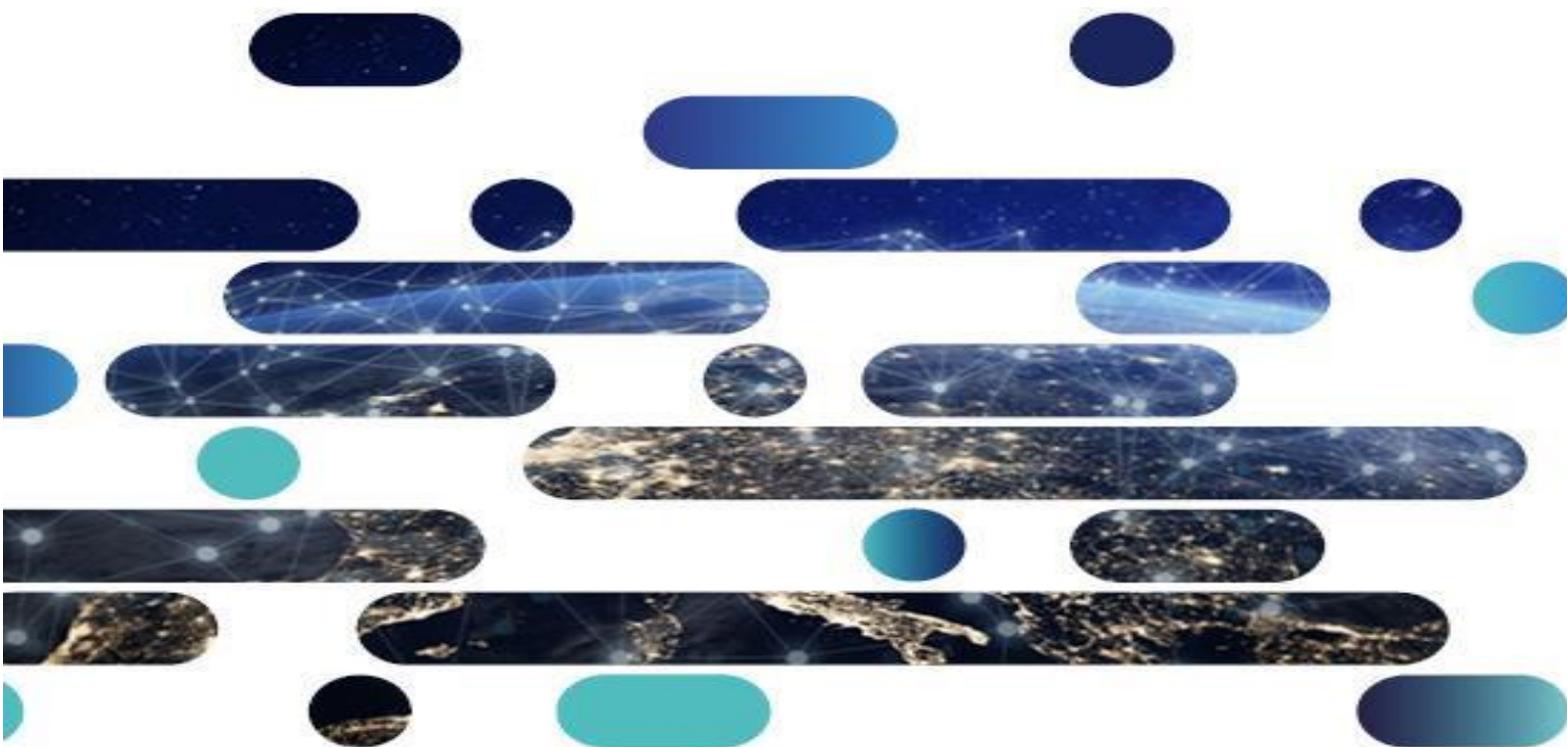


Modélisation du taux de transformation

Interprétabilité/performance - XGBoost



Critères	Performance sur la base d'apprentissage
Gini normalisé	48%
Score moyen estimé	37%

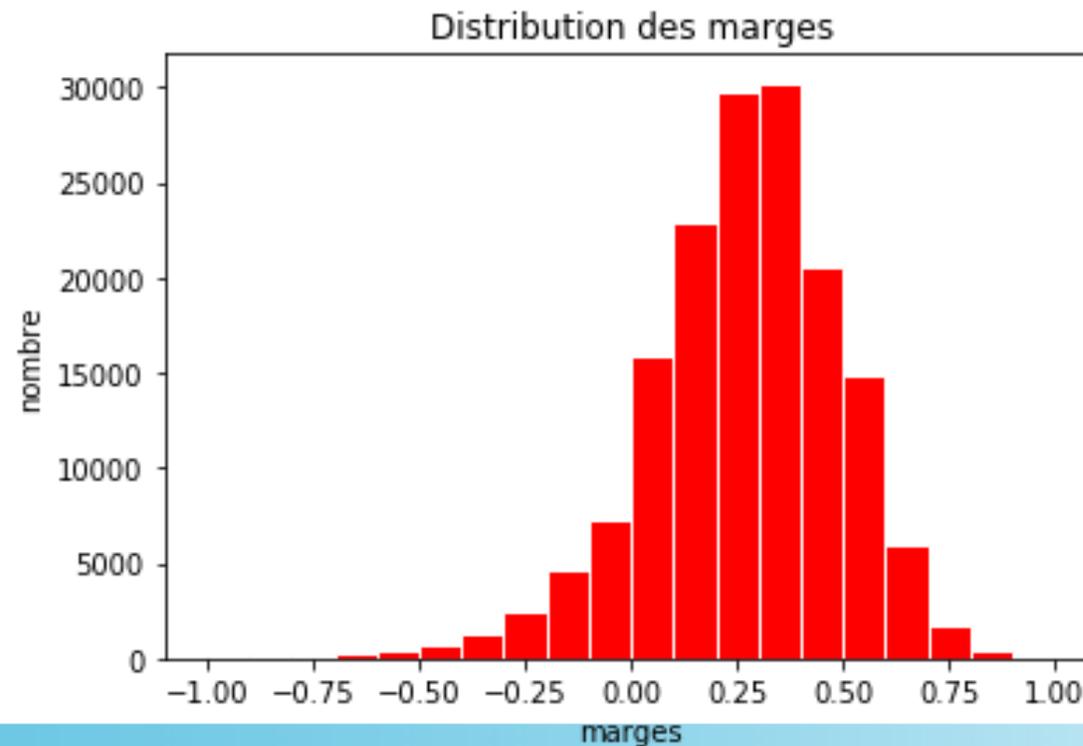


Constitution des groupes de politique tarifaire

Constitution des groupes de politique tarifaire

Création de la variable marge

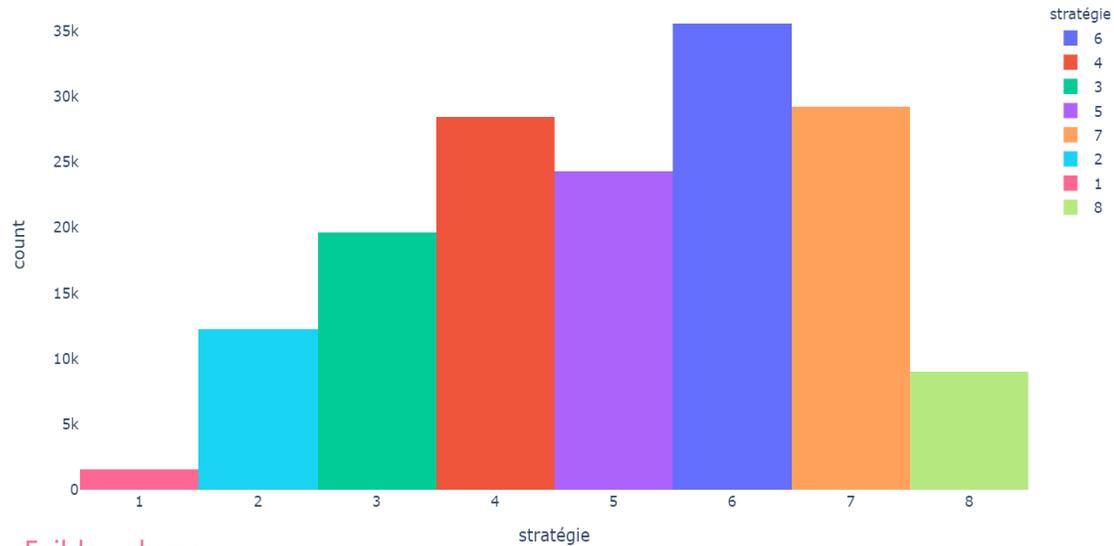
- I. Prime HT = prime pure + frais
- II. Marge = Mt cotisations HT – Prime HT
- III. Marge en % = Marge / Mt cotisations HT = -évolution(Mt cotisations HT , Prime HT)



Constitution des groupes de politique tarifaire

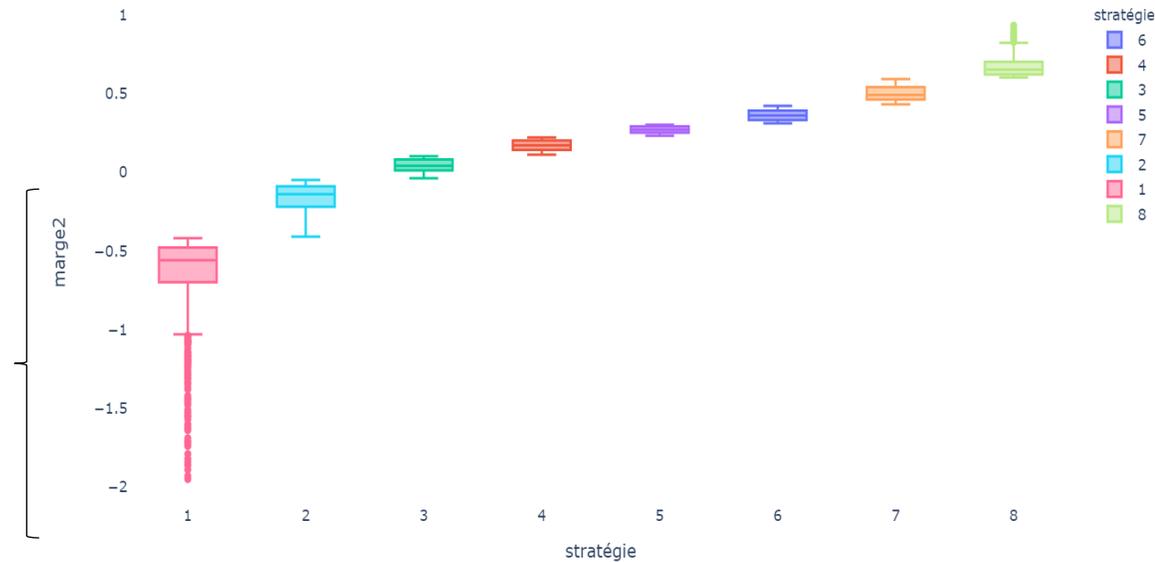
Classification ascendante hiérarchique

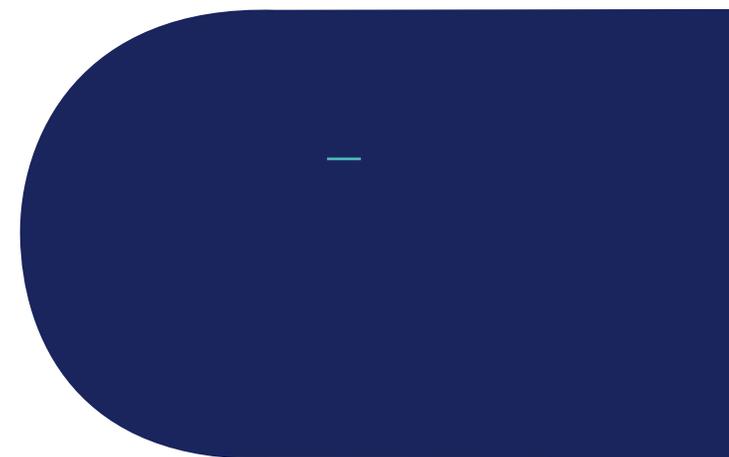
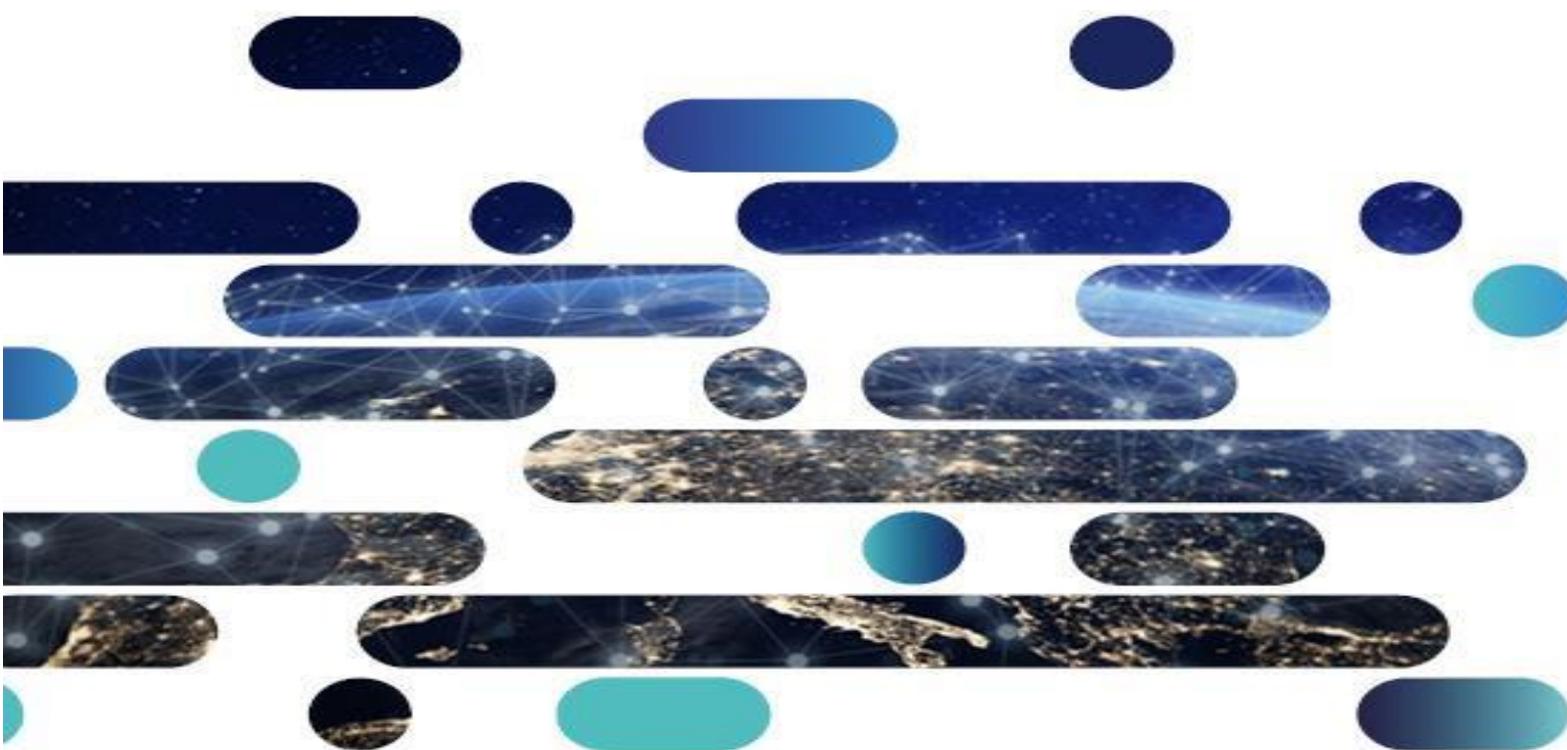
On utilise le critère de Ward qui minimise la variance intra-classes des sous-groupes. Au sein des dendrogrammes, les individus sont reliés par des traits verticaux dont la longueur est proportionnelle à leur distance.



Faible volume dans la classe 1

Marge négative

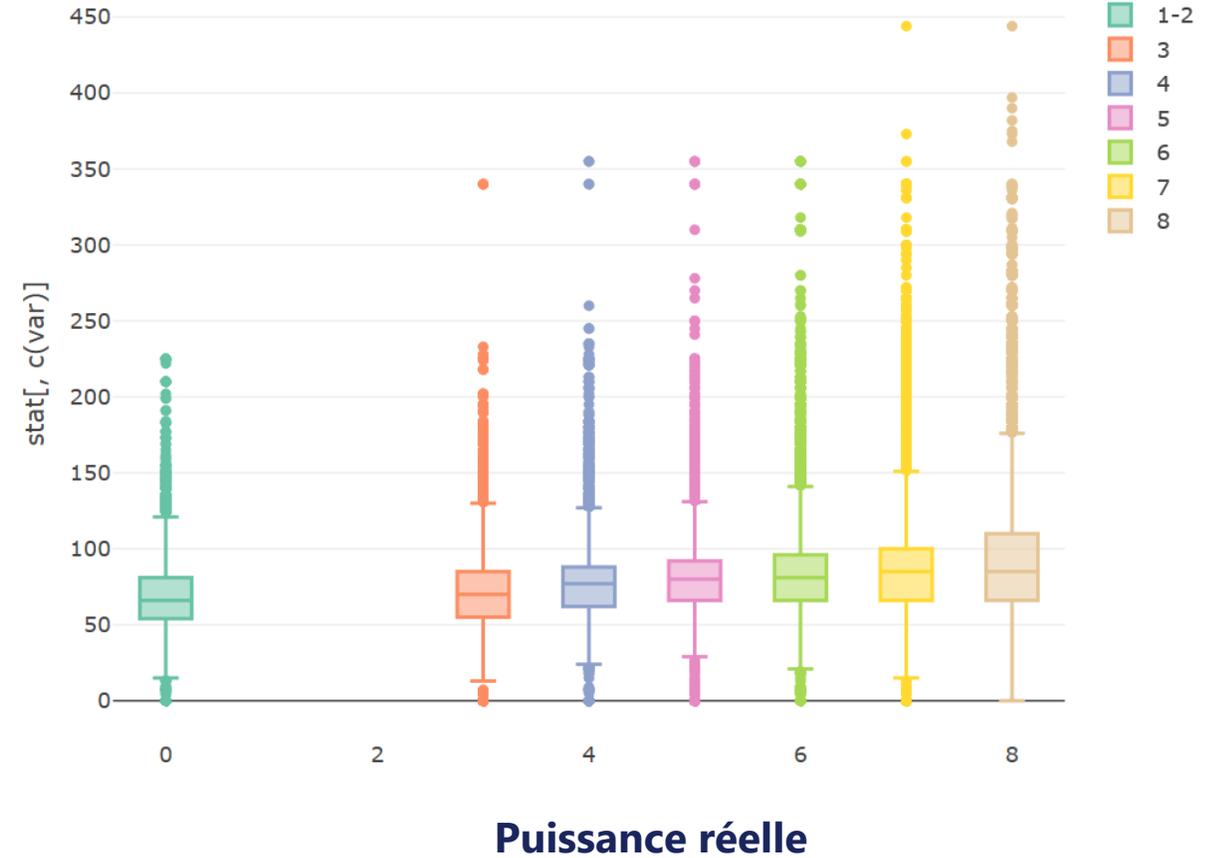
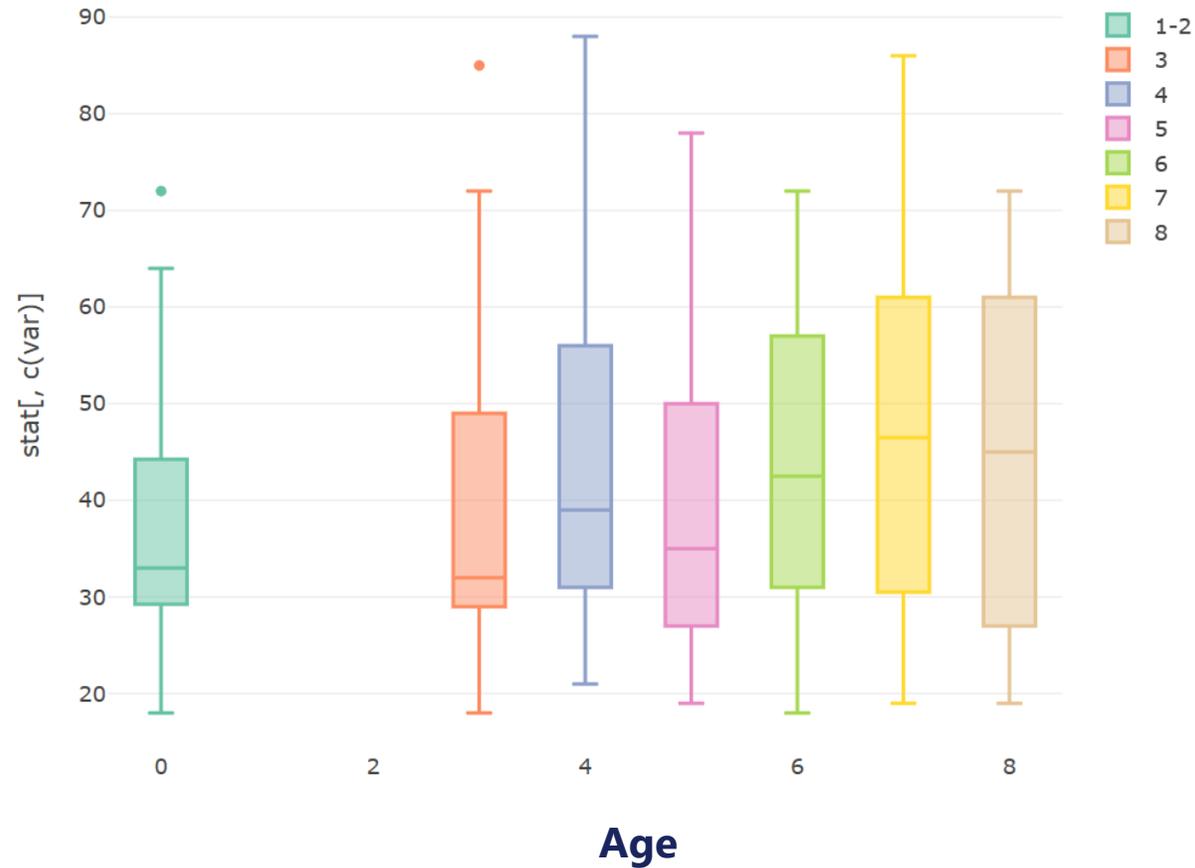




Jumelage des sujets par score de propension

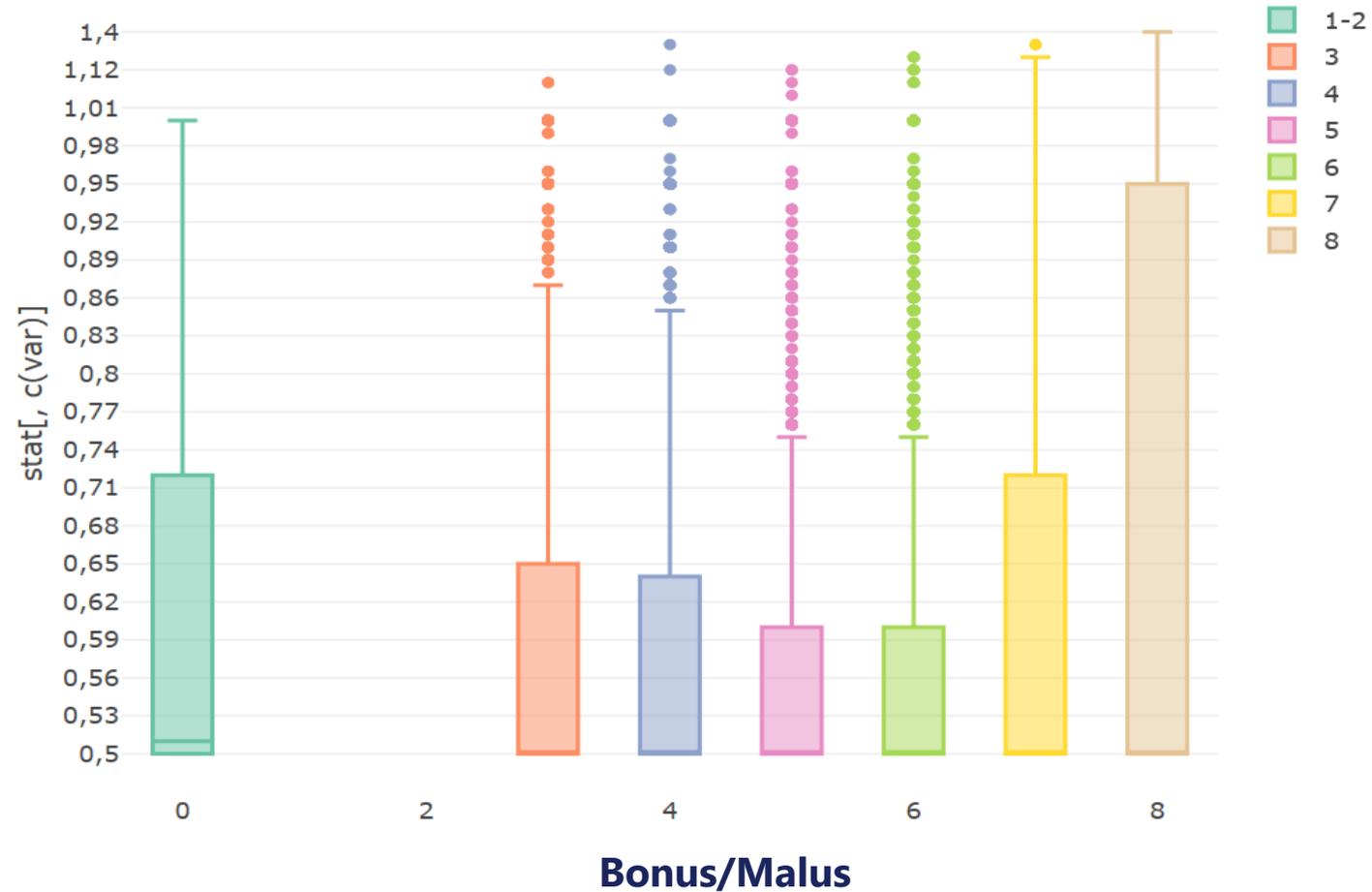
Jumelage par score de propension

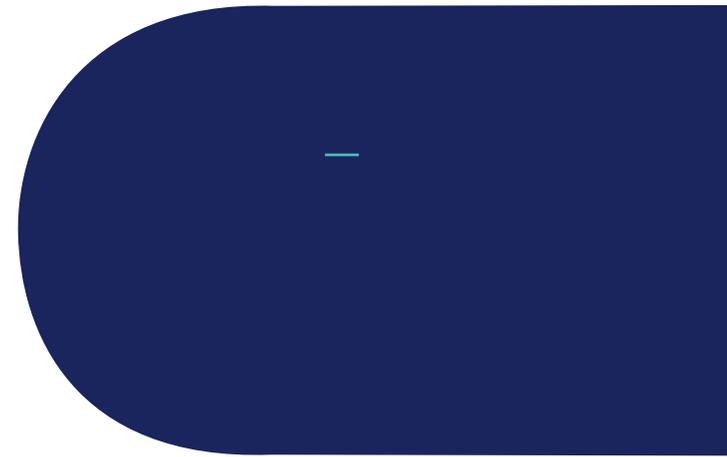
- Caractéristiques de chaque stratégie tarifaire : entre spécificité et support commun



Jumelage par score de propension

Caractéristiques de chaque stratégie tarifaire : entre spécificité et support commun





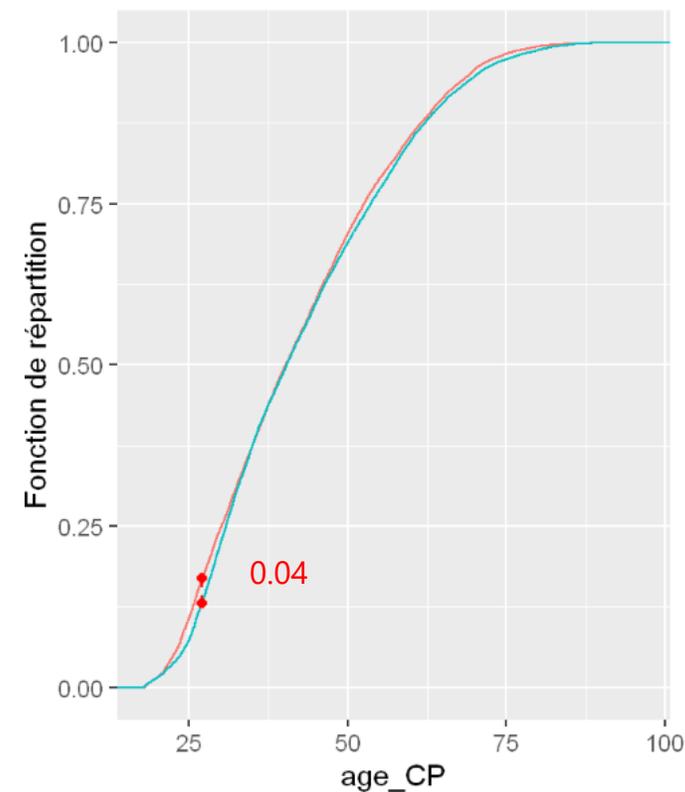
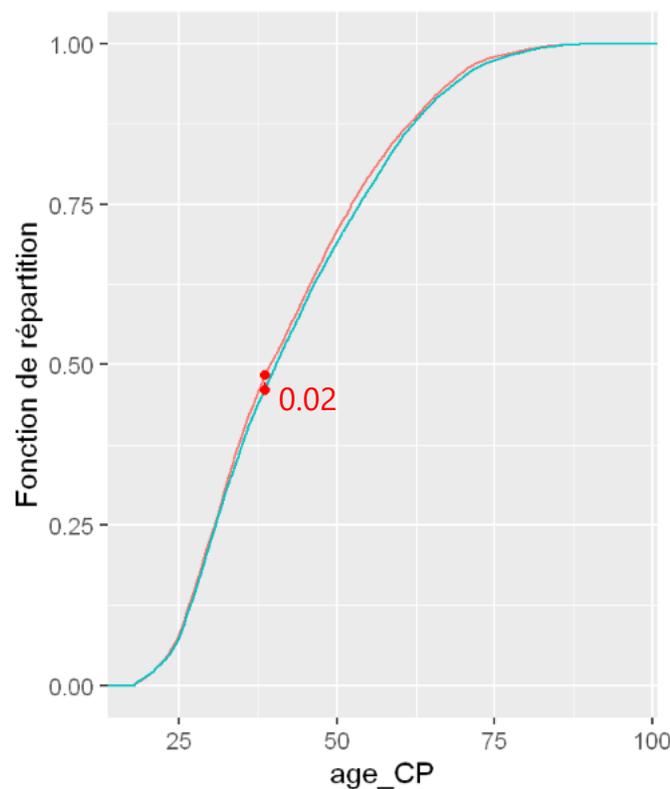
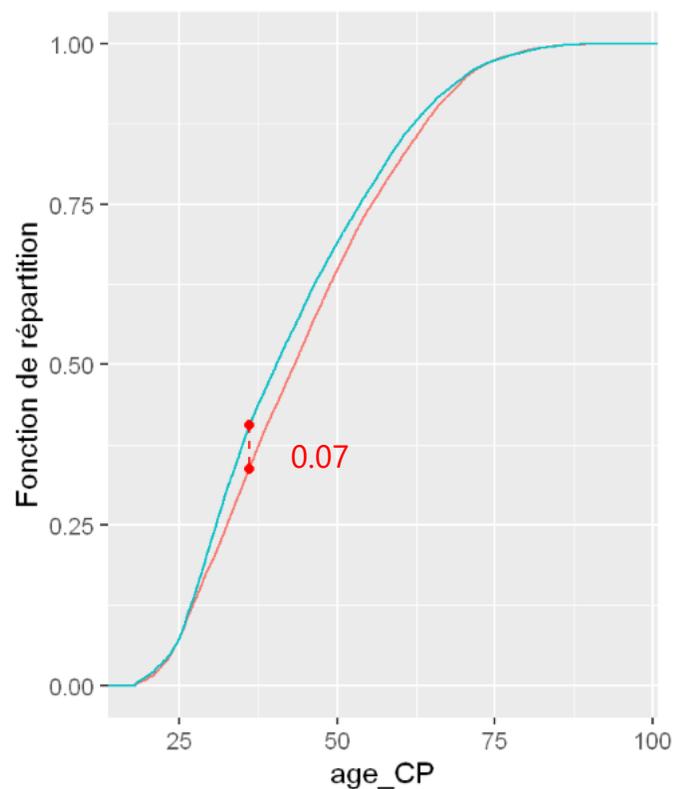
Groupe de traitement : stratégie 4

Stratégie 4 (marge faible)

Contrôle : stratégie 5 (marge moyenne)

Score de propension estimé par XGBoost

Score de propension estimé par GLM



— groupe de contrôle — groupe de traitement

— groupe de contrôle — groupe de traitement

— groupe de contrôle — groupe de traitement

Stratégie 4 (marge faible)

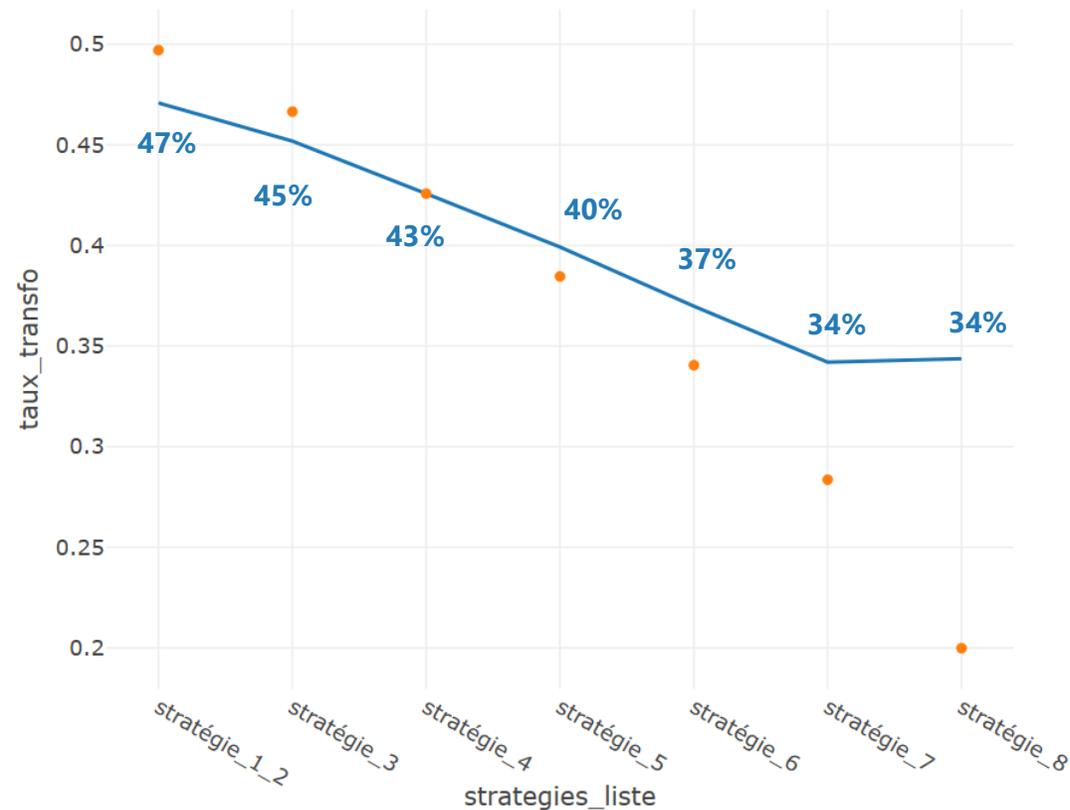
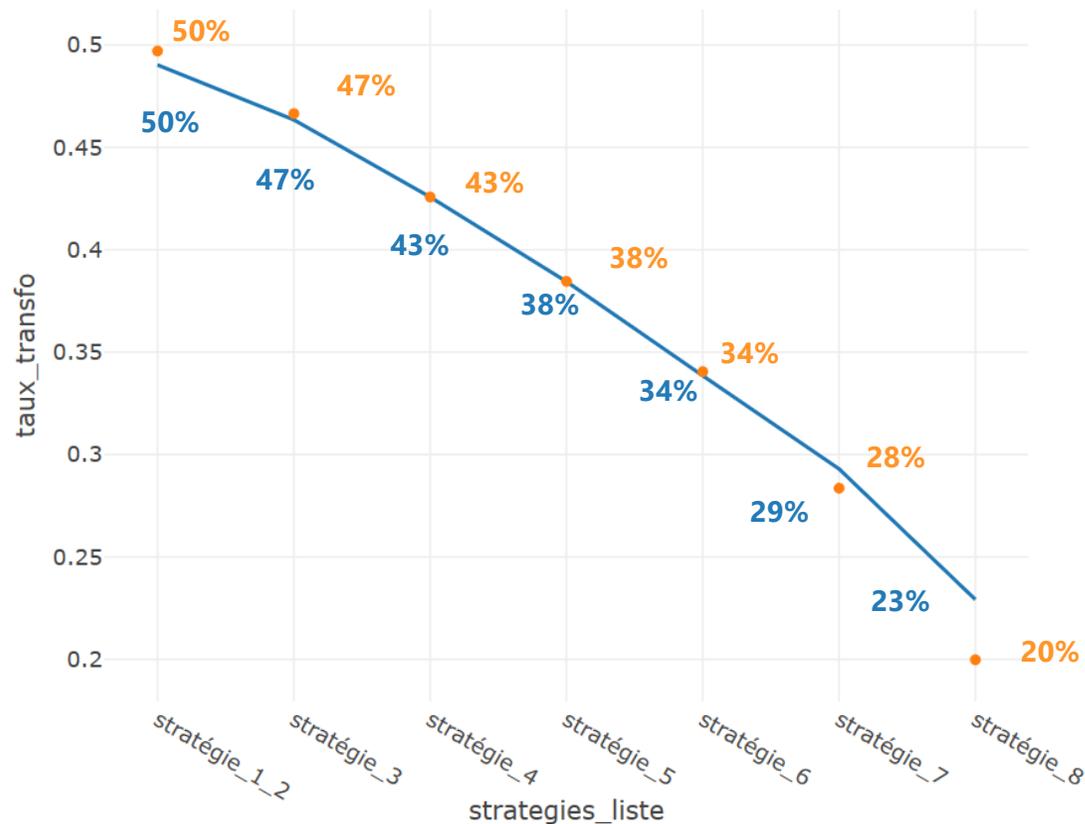
Taux de transformation moyen par stratégie
Taux de transformation moyen du groupe 4 pour les différentes stratégies

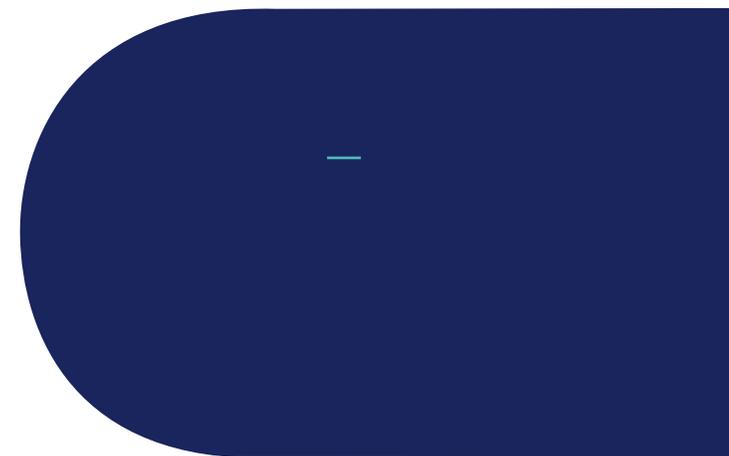
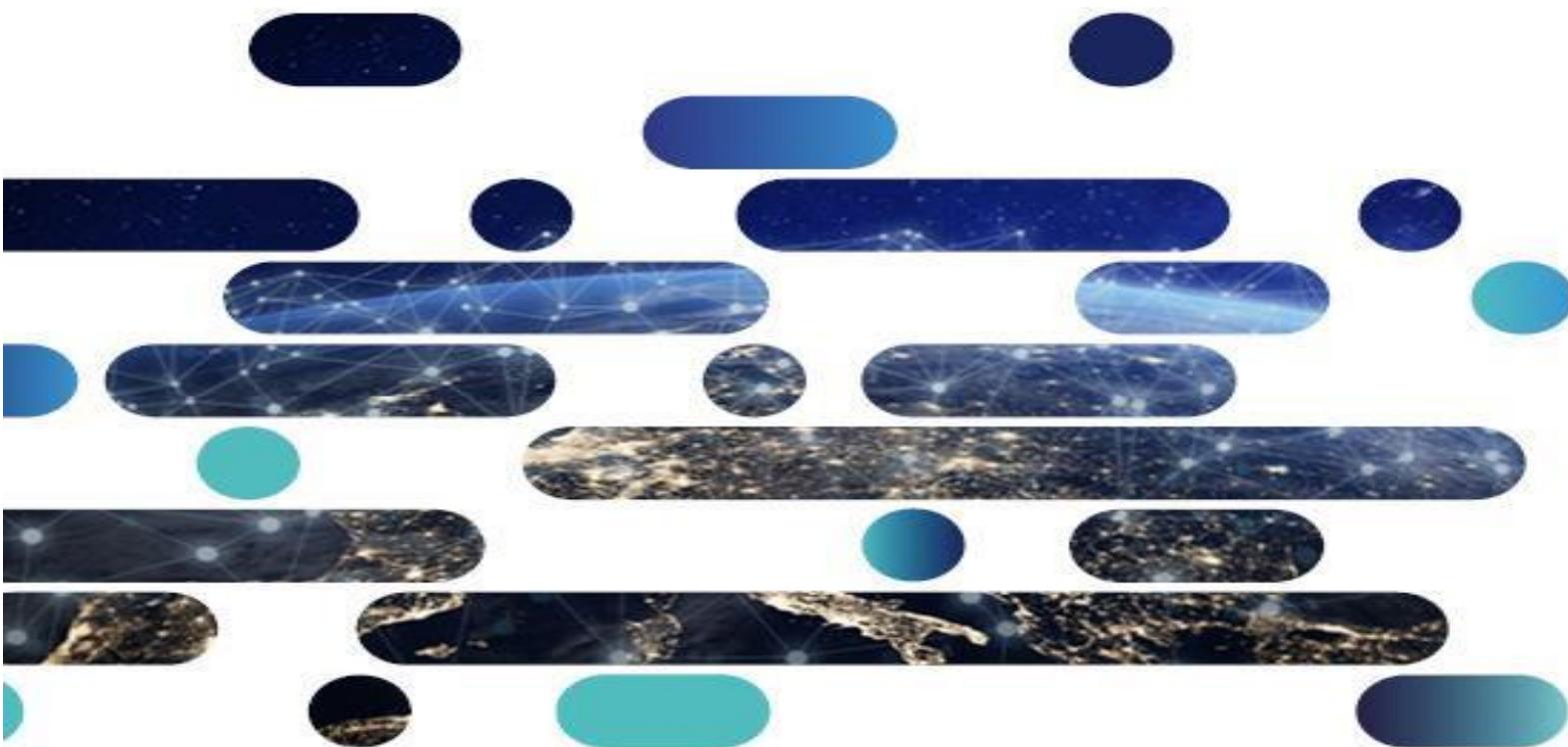
Score de propension estimé par **GLM**

Score de propension estimé par **XGBoost**

Courbe d'élasticité (ATT) pour la stratégie_4

Courbe d'élasticité (ATT) pour la stratégie_4

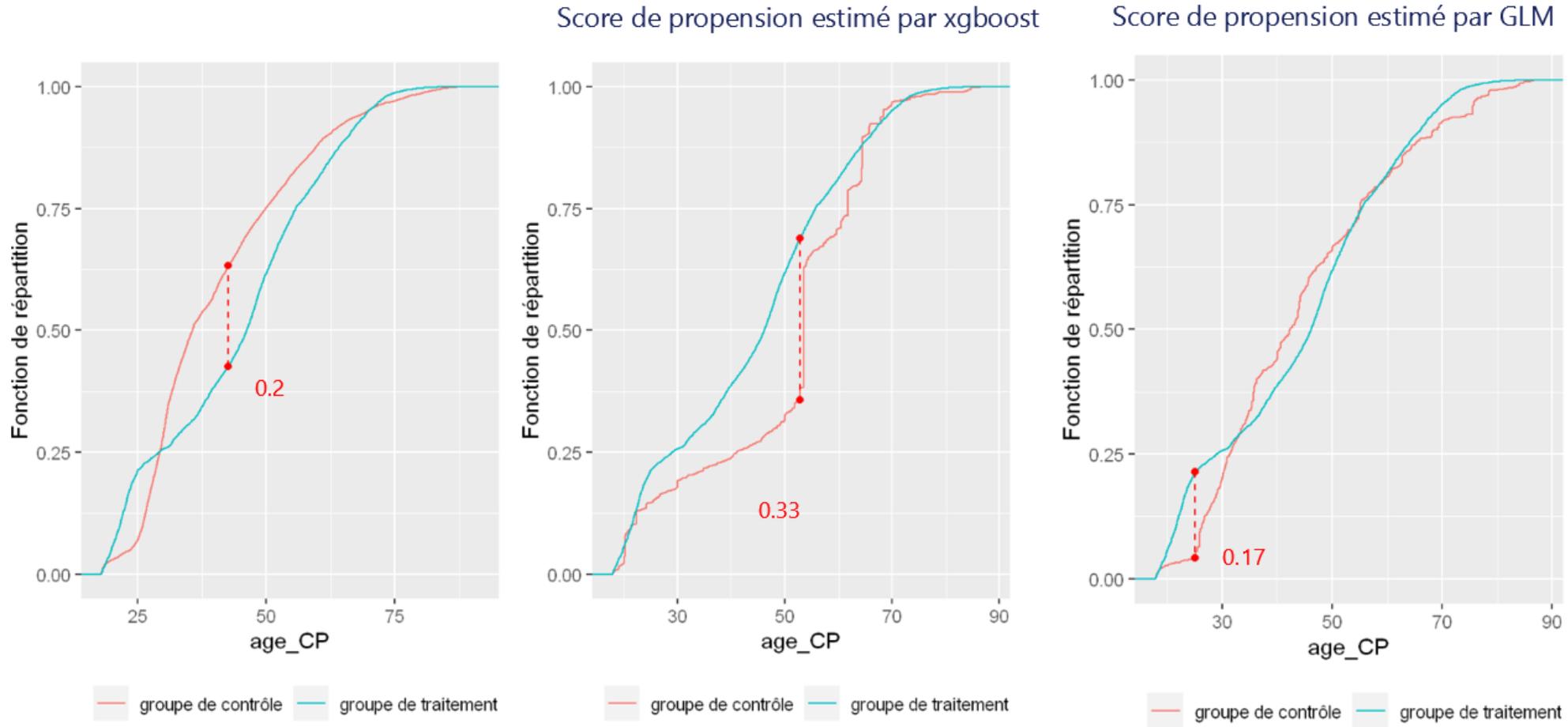




Limites à travers un exemple

Stratégie 8 (marge extrêmement élevée)

Contrôle : stratégie 1-2 (marge négative)





Modélisation de l'élasticité au prix individuelle et optimisation

Modélisation de l'élasticité au prix individuelle

Définition de l'équation d'optimisation

Maximiser:

$$\text{Rentabilité} = \sum_{i \leq n} (P_i - S_i) * \hat{f}(P_i, X_i)$$

Modularité + simulabilité

Sous contrainte:

$$\left\{ \begin{array}{l} \frac{1}{n} \sum_{i=1}^n \hat{f}(P_i, X_i) \geq \gamma \\ P_i \in [P_0 \pm 20\%] \forall i \leq n \end{array} \right. \text{ nouvelle contrainte}$$

- P_i : prix commercial que paie le client i où $i \leq n$ (n : nombre de prospects)
- S_i : la sinistralité du client i
- $\hat{f}(P_i, X_i)$: la probabilité de conversion du prospect i
- X_i : caractéristique du prospect i (âge, sexe, expérience, ...)



Maximiser la rentabilité totale

Sous contraintes:

- Taux de conversion moyen $\geq \bar{m}\%$
- Variation tarifaire $\leq 20\%$



Personnalisable

Modélisation de l'élasticité au prix individuelle

Echantillon de test – nouvelles données

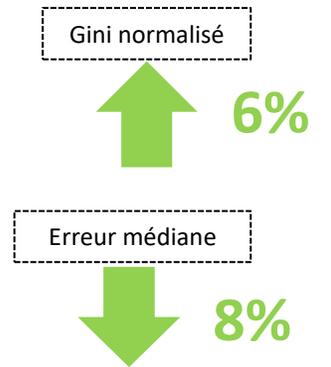
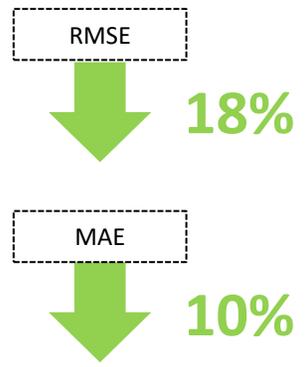
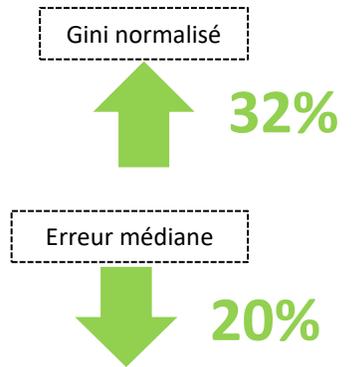
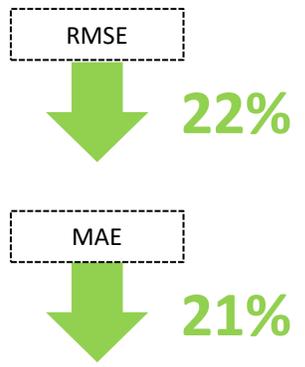
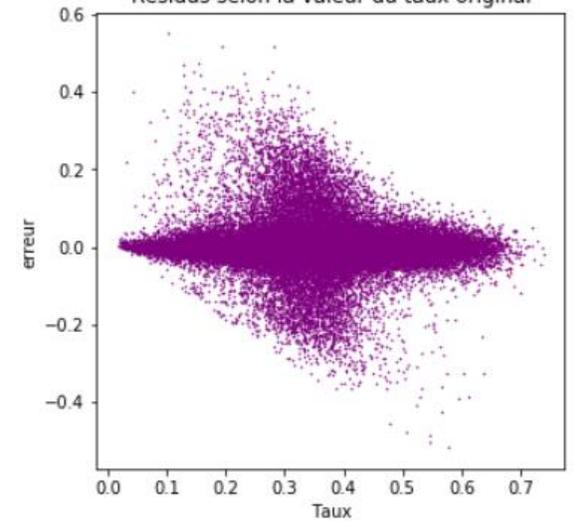
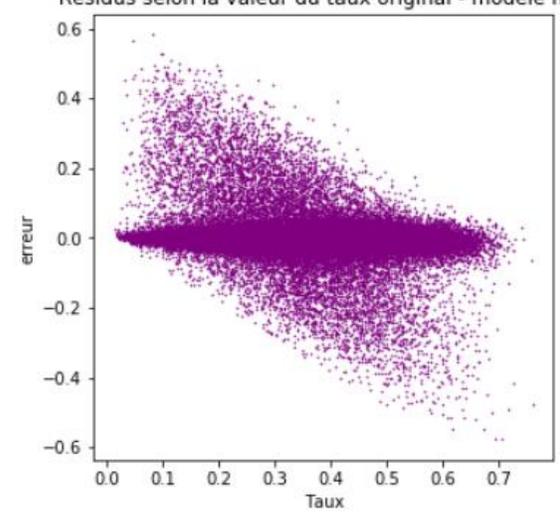
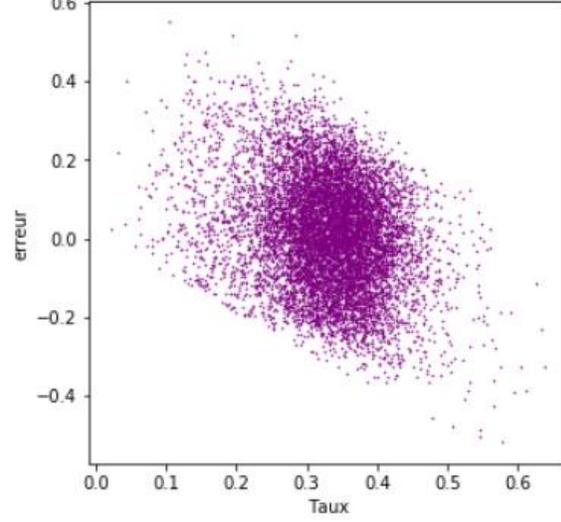
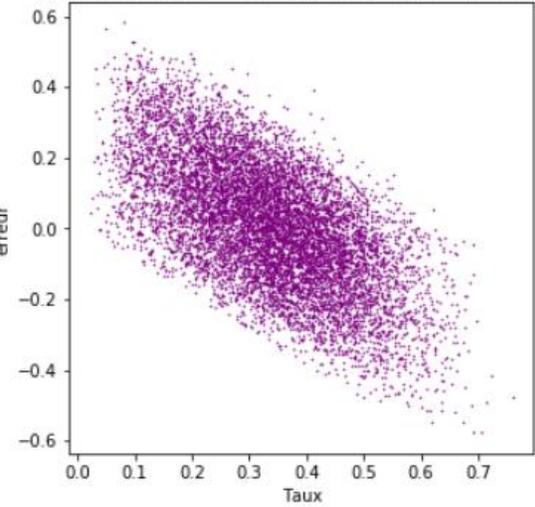
Echantillon de test – mixte

Résidus selon la valeur du taux original - modèle naïf

Résidus selon la valeur du taux original

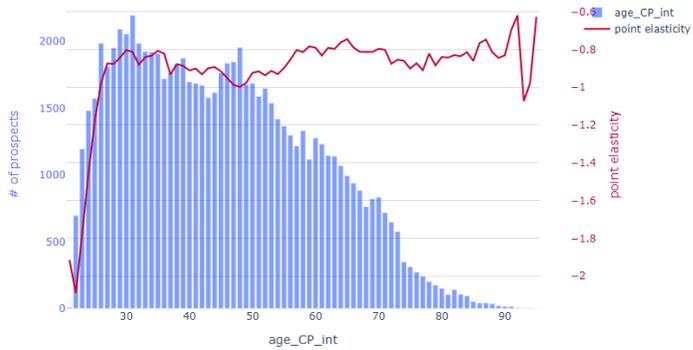
Résidus selon la valeur du taux original - modèle naïf

Résidus selon la valeur du taux original

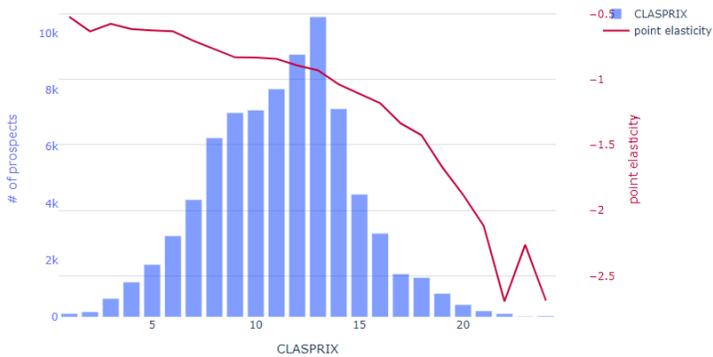


Modélisation de l'élasticité au prix individuelle

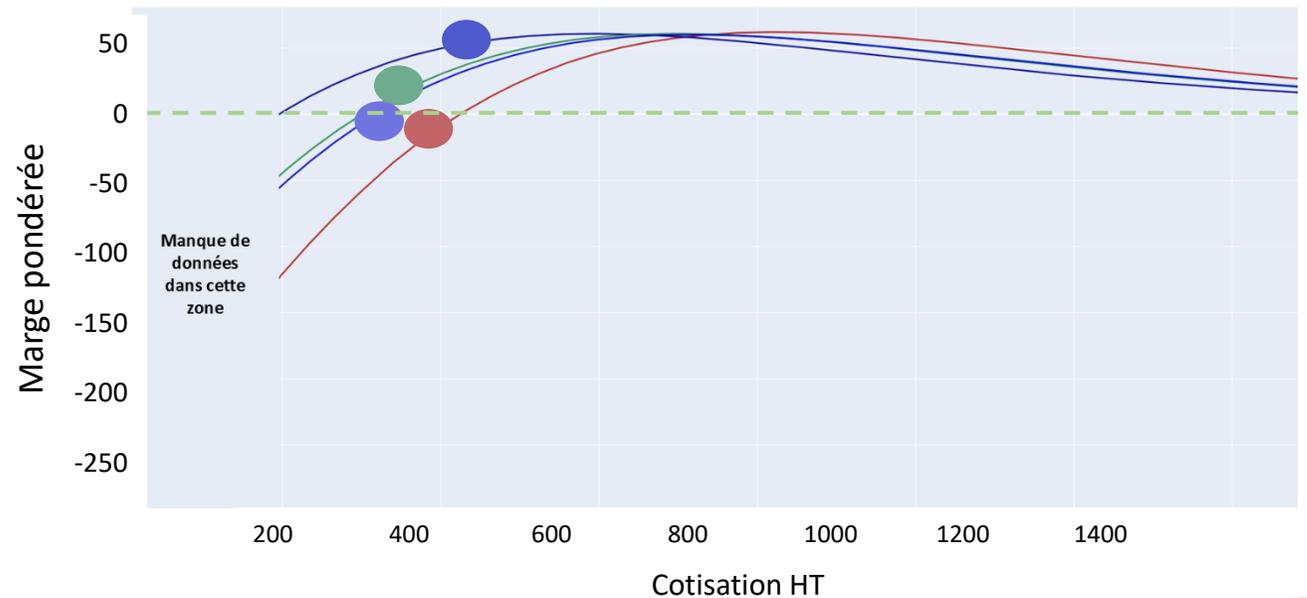
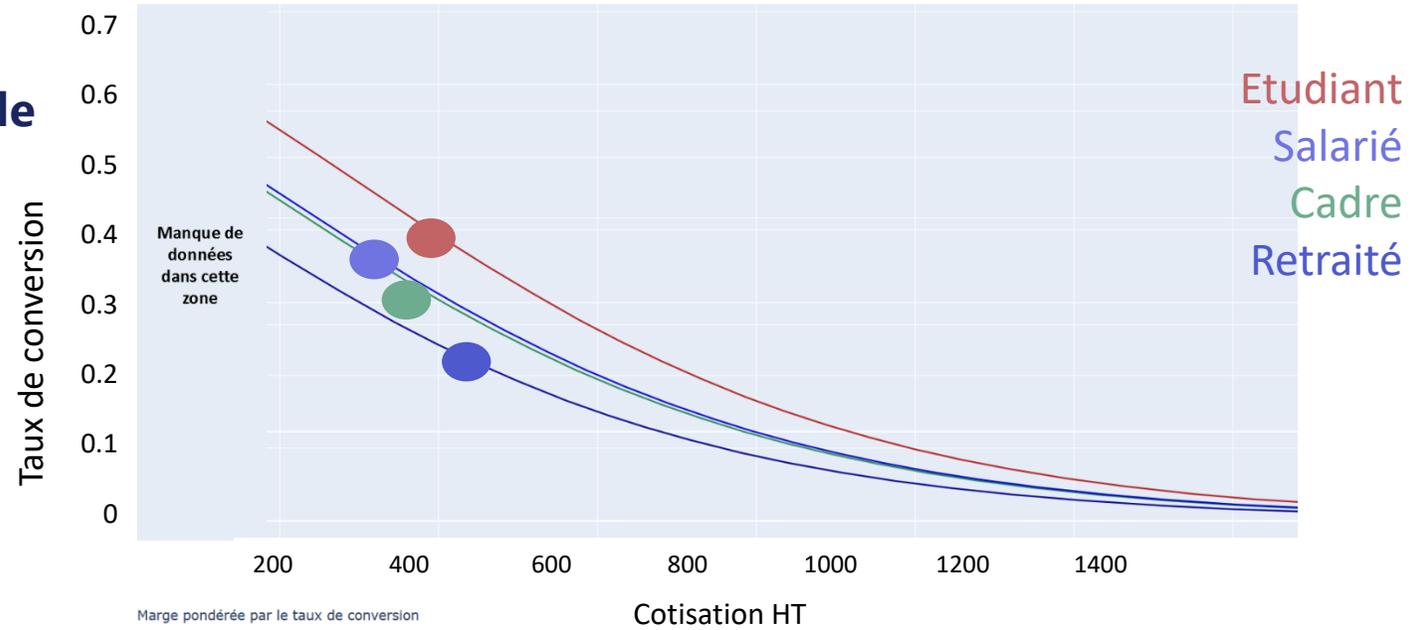
Elasticité au prix selon l'âge du conducteur



Elasticité au prix selon la classe de prix



Courbes de demande et de rentabilité -- CSP



Optimisation

L'équation lagrangienne:

$$L(P_1, \dots, P_n, \lambda) = \underbrace{- \sum_{i=1}^n (P_i - S_i) * \hat{f}(P_i, X_i)}_{\text{Rentabilité totale}} + \underbrace{\lambda * (n\gamma - \sum_{i=1}^n \hat{f}(P_i, X_i))}_{\text{pénalisation sur le taux de conversion}}$$

- $\lambda = 0$: pas de contrainte
- λ petit: plus d'importance pour la rentabilité
- λ grand: plus d'importance pour la contrainte

Algorithme sur-mesure :

Initialisation : $\lambda = 0, P_i = P_{init}$

Pour $\lambda \in [0, \lambda_{max}]$:

$$\rightarrow P_{i,n+1} = P_{i,n} - learning_rate * \frac{\delta L}{\delta P_{i,n}}$$

\rightarrow Si $P_i > P_{init} * 1.2$ alors $P_i = P_{init} * 1.2$

Si $P_i < P_{init} * 0.8$ alors $P_i = P_{init} * 0.8$

\rightarrow Si la contrainte de volume est vérifiée, on garde la solution, sinon on la retire.



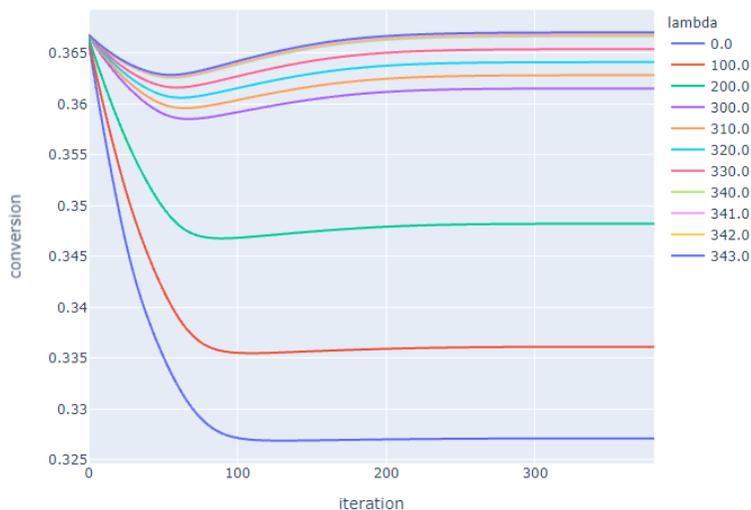
**Inversion coûteuse de la
matrice**
 \rightarrow **Résolution par descente
de gradient projetée**

Optimisation

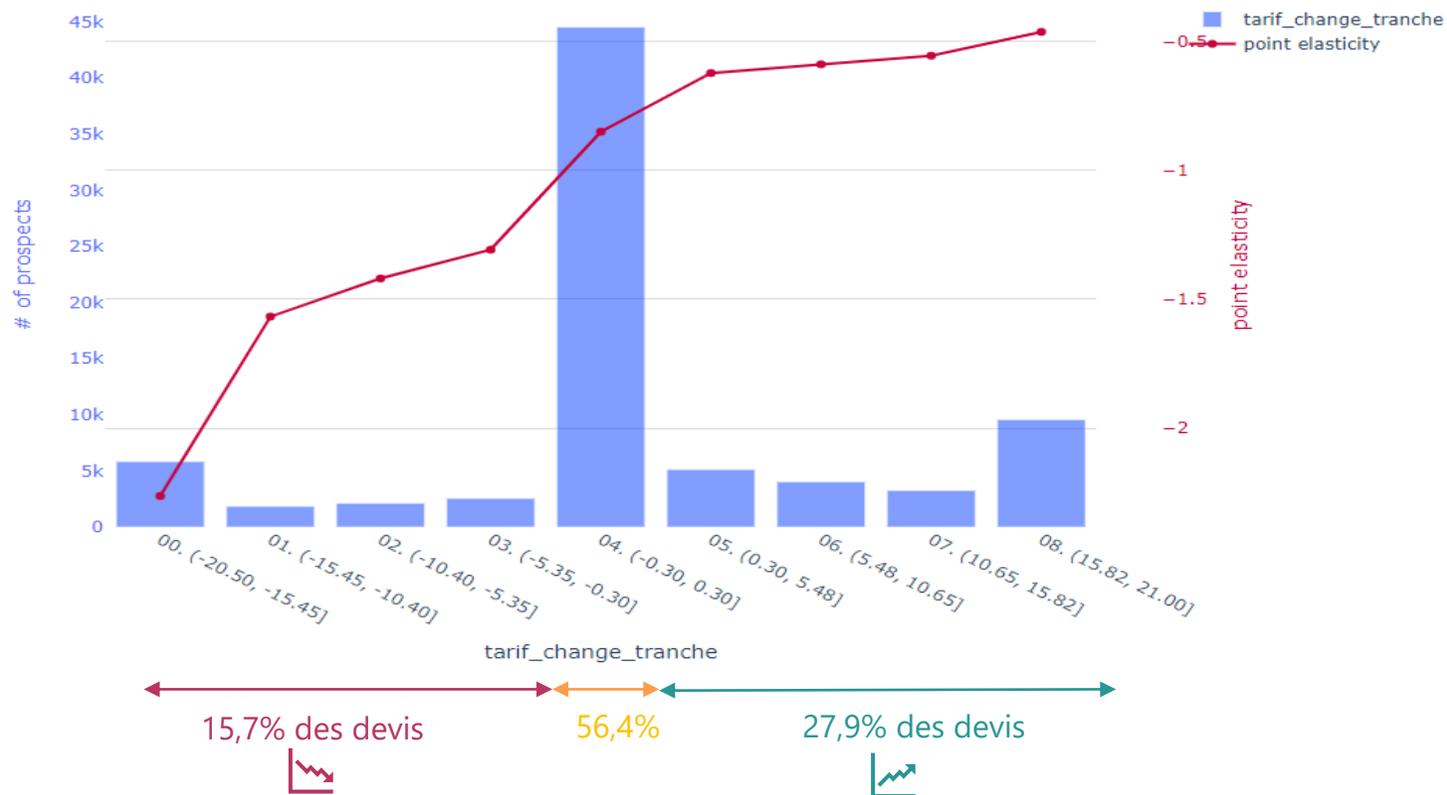
Après Optimisation

Périmètre: auto particulier, formule Tous Risques, devis 2019

- ✓ Marge totale: **+8,6%**
- ✓ Conversion: **+0%**
- ✓ Chiffre d'affaire: **+3%**
- ✓ Loss Ratio: **-1.89 pts**

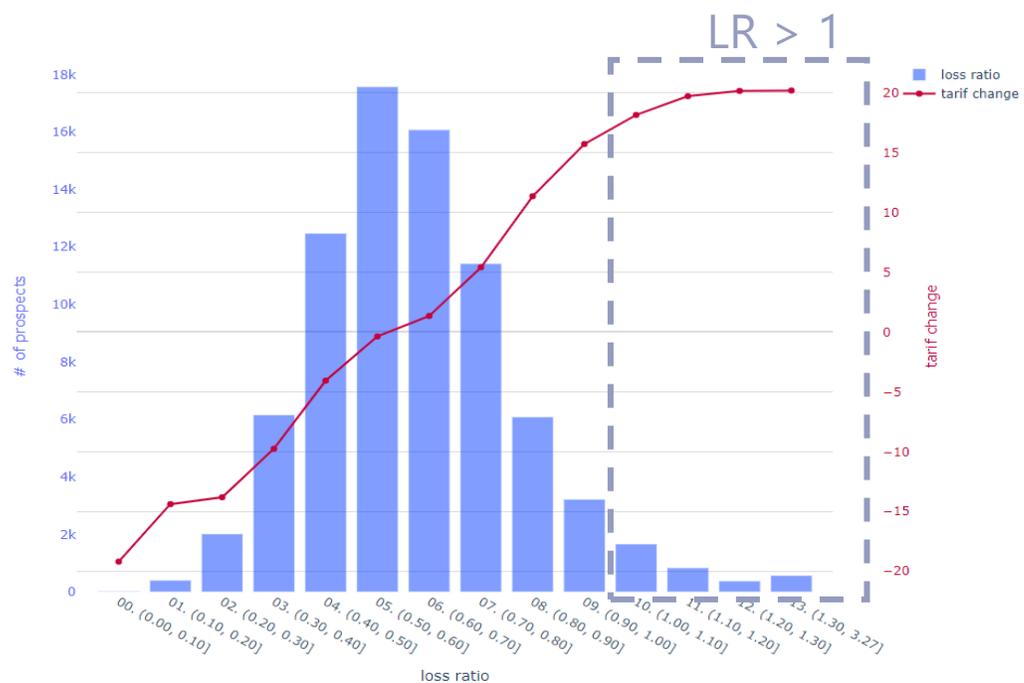


Variation du prix en pourcentage

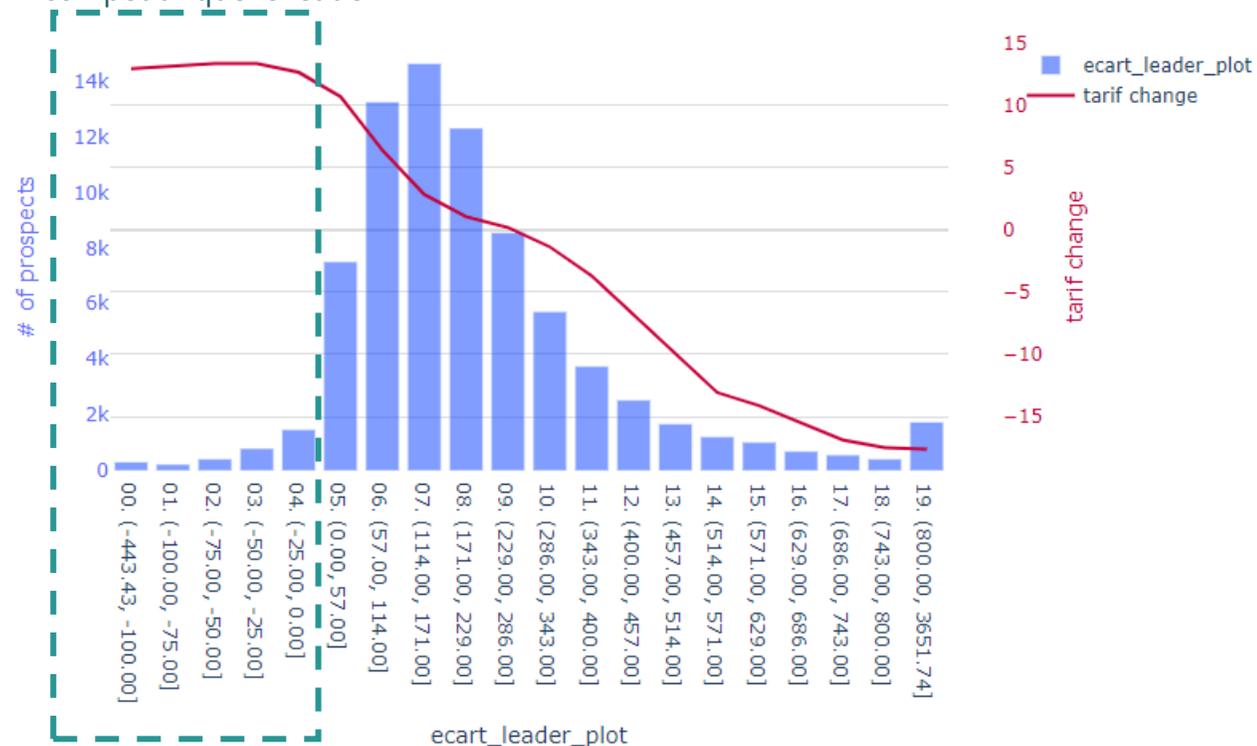


- Pas de changement tarifaire pour la moitié des devis (56,4%)
- Augmentation tarifaire pour les prospects peu sensible au prix (27,9%)
- Baisse de tarif pour les prospects très sensible au prix (15,7%)
- **Les variations du prix sont limitées à $\pm 20\%$**

Optimisation



+ compétitif que le leader

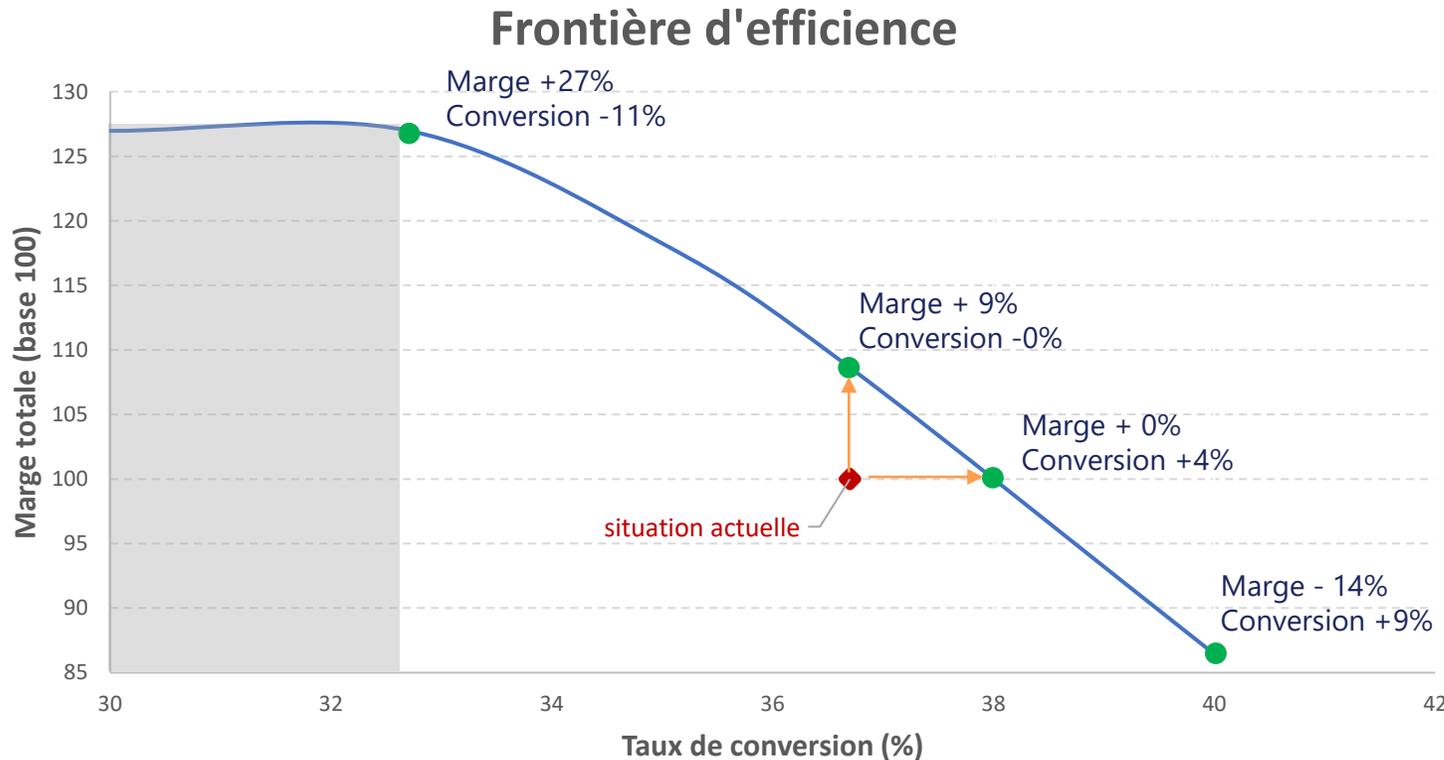


- ✓ Augmentation de tarif pour les prospects avec un LR élevé
- ✓ Baisse de tarif pour les prospects avec un LR faible

- ✓ Augmentation de tarif pour les prospects sur lesquels notre tarif est plus agressif
- ✓ Baisse de tarif pour les prospects sur lesquels nous ne sommes pas assez compétitifs

Optimisation

Frontière d'efficience



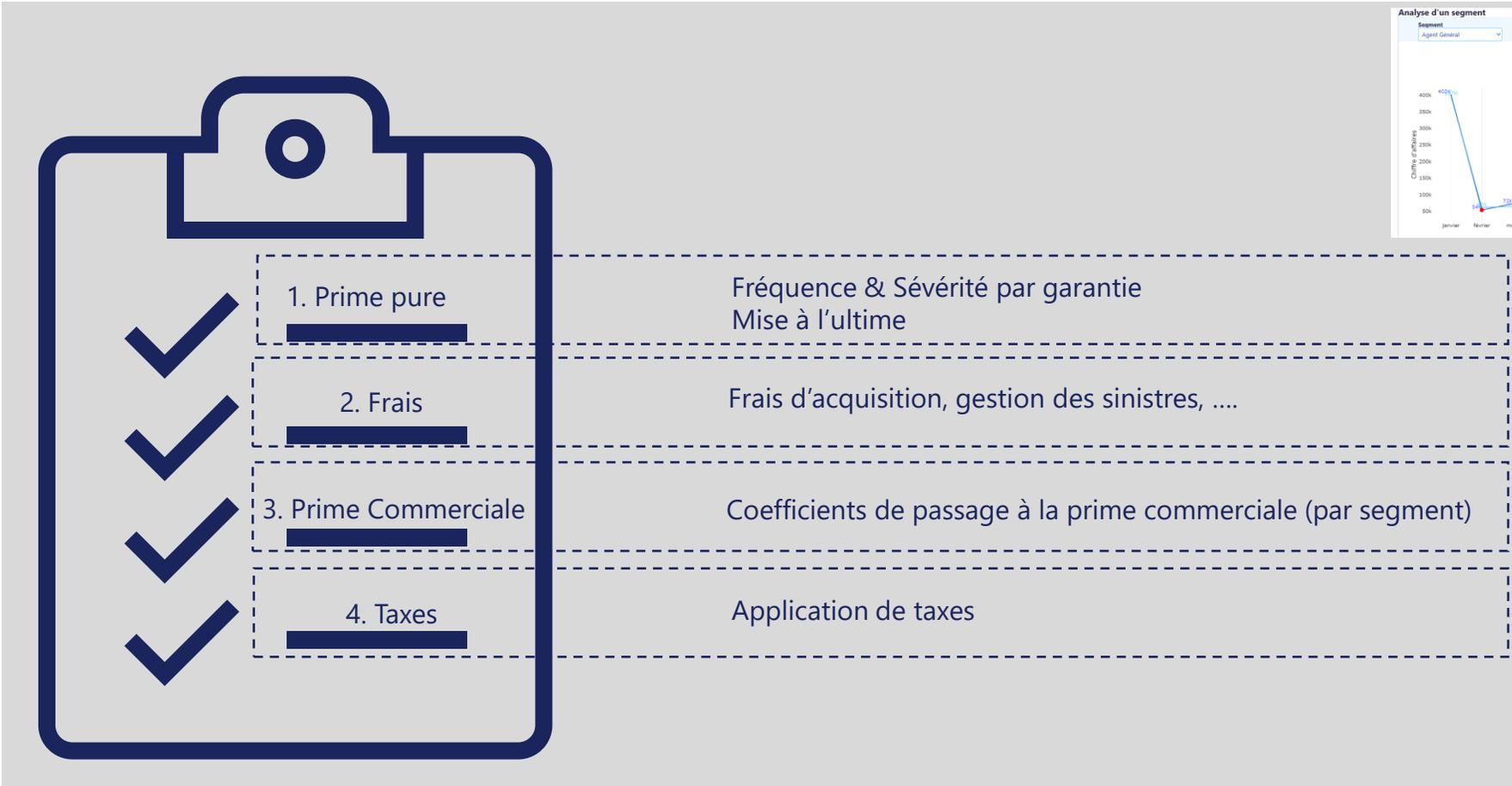
- ✓ Situation actuelle optimisable
- ✓ Les situations optimisées sont nombreuses
- ✓ Il est nécessaire d'arbitrer le compromis entre la marge et le taux de conversion selon la stratégie de l'entreprise.

Contraintes personnalisables :

- Bornes ou niveau d'évolution imposés sur un segment en particulier
- Contrainte sur le loss ratio, la valeur client
- Optimisation restreinte sur une partie du marché

Optimisation

Mise en production à partir d'une grille d'allocations de primes commerciales.



Suivi & pilotage...

Enregistrer l'image/zoomer sur l'image...



...En tenant compte de son environnement



FIN