

## > La méthode des quotas

La méthode des quotas vise à constituer un échantillon dont la structure sur certaines variables (sociodémographiques, comportementales...) est proche de celle de la cible étudiée. Elle évite ainsi de générer des écarts de pondération importants lors de l'étape statistique de redressement de l'échantillon sur les critères de quotas.

L'indicateur statistique d'efficacité du redressement  $Q$  permet d'estimer l'impact potentiel d'écarts de structure d'échantillon. Il est obtenu par le ratio :

$$Q = \frac{(\sum_{i=1}^n p(i))^2}{n(\sum_{i=1}^n p(i)^2)}$$

Où  $p(i)$  est le poids de l'individu  $i$  et  $n$  la taille de l'échantillon. L'indicateur varie de 0 et 1, il atteint son maximum de 1 lorsque tous les poids sont identiques.

Lorsque l'efficacité du redressement présente une qualité  $Q$ , la formule de la taille d'échantillon effective de Kish (1964, Survey Sampling, New-York Wiley) énonce que la qualité des estimations est équivalente à celle que l'on obtiendrait avec un échantillon de taille  $Q \times N$  :

- la variance de l'estimateur est divisée par  $Q$  et la longueur de l'intervalle de confiance est multiplié par  $1 / (\sqrt{Q})$
- Une qualité de redressement de 90% peut induire une dégradation de 5% de la fiabilité des résultats,
- Une qualité de redressement de 50% peut induire une dégradation de 40% de la fiabilité des résultats,

Une fois la population de base à interroger déterminée (en fonction de la problématique de l'étude), la méthode des quotas permet ainsi de s'assurer de la représentativité de l'échantillon en lui affectant une structure similaire à celle de la population de base qui limitera l'impact du redressement. Cela permet notamment de réduire les biais d'échantillonnage dus aux différences de taux d'acceptation ou de joignabilité selon les cibles (couples de 2 actifs vivant en banlieue pour une enquête téléphone, jeunes ou CSP- pour une enquête en ligne).

## Le choix des critères de quotas

Les critères de quotas doivent être :

- Pertinents et liés aux variables d'intérêt (par exemple : vote aux précédentes élections pour une étude politique, opérateur pour une étude sur des services de téléphonie...),
- Associés à des données de cadrage fiables et valides permettant de connaître les distributions recherchées sur la cible étudiée (la mention de la source ayant fourni les données de cadrage est indispensable),
- Aisément identifiables à partir du questionnaire pour permettre la validation en début d'interview,
- Non redondants entre eux : si deux quotas sont indépendants, l'éventuelle déformation de l'un n'implique pas la déformation de l'autre,
- En nombre adapté à l'enquête, entre 3 et 5 **critères** généralement, un trop grand nombre de quotas pouvant générer des problèmes de fin de quotas avec des profils quasi-impossibles à trouver : le contrôle d'autres variables liées au sujet de l'enquête peut toujours être opéré par post-stratification (redressement). Dans la pratique, certains instituts visent à ne pas excéder le seuil de 1 contrainte pour 40 individus statistiques (le nombre de contraintes est égal au nombre total de **modalités** de redressement moins le nombre de variables de redressement).
- Compatible avec la taille de l'échantillon : Les électeurs de Jacques Cheminade ont représenté 0,2% des voix lors de l'élection présidentielle de 2017 → Est-il pertinent de rechercher strictement 2 personnes ayant voté pour Jacques Cheminade lors de la présidentielle 2017 dans un échantillon de 1000 personnes pour une étude d'intention de vote ?

*Exemple de quotas pour un échantillon national représentatif de la population France Métropolitaine de 18 ans ou plus :*



Population totale 18 ans ou plus		%
		EE
<b>SEXE</b>		
<u>HOMME</u>		47,7
<u>FEMME</u>		52,3
<b>TOTAL</b>		<b>100,0</b>
<b>AGE</b>		
<u>18-24 ans</u>		10,7
<u>25-34 ans</u>		15,9
<u>35-49 ans</u>		26,4
<u>50-64 ans</u>		25,3
<b><u>(65ans et +)</u></b>		<b>21,7</b>
<u>65-74 ans</u>		10,9
<u>75 ans et +</u>		10,8
<b>TOTAL</b>		<b>100,0</b>
<b>UDA 9</b>		
<u>Region Parisienne</u>		18,5
<u>NORD</u>		6,3
<u>EST</u>		8,6
<u>BP EST</u>		7,7
<u>BP OUEST</u>		9,3
<u>OUEST</u>		13,6
<u>SUD OUEST</u>		11,3
<u>SUD EST</u>		12,1
<u>MEDITERRANEE</u>		12,7
<b>TOTAL</b>		<b>100,0</b>
<b>Profession de l'individu</b>		
<u>(Indépendants / Chefs d'entreprise)</u>		<u>4,7</u>
<u>Agriculteur</u>		1,1
<u>Artisan/commerçant/Chef d'entreprise de +10 sal</u>		3,6
<u>Cadres et profession intellectuelle supérieure</u>		10,0
<u>Professions intermédiaires</u>		13,9
<u>Employés</u>		16,7
<u>Ouvriers</u>		12,8
<u>Retraités</u>		27,7

Source : INSEE Enquête Emploi 2017

Dans notre exemple, nous devons donc, pour une étude auprès d'un échantillon de 1000 personnes, interroger 477 hommes, 159 personnes âgées de 25 à 34 ans, 277 retraités, etc.

Ces critères de sexe, âge, région et catégorie socioprofessionnelle peuvent être complétés par d'autres critères importants pour certains types d'enquêtes comme le vote à une élection précédente dans un sondage politique. Il peut être également pertinent d'envisager des quotas croisés (ex : sexe par tranche d'âge), selon la problématique de l'enquête.

Il est important d'utiliser une statistique de référence avec la définition appropriée, notamment pour la **personne de référence du ménage** (PRM) dans le Recensement Permanent de Population : [définition INSEE](#). La collecte du recensement utilise depuis 2018 une définition basée sur l'actif le plus âgé qui pourra être utilisée dès la publication des nouveaux chiffres démographiques.

Les études réalisées par téléphone, par internet, en face à face sur PC, tablettes ou mobiles, ont une gestion automatisée des quotas grâce au suivi logiciel dynamique des réponses données aux questions de profilage du répondant. Lorsque 477 hommes sont interrogés, le logiciel mettra automatiquement fin aux interviews réalisées avec des hommes de façon à assurer la représentativité de l'échantillon final (non sollicitation quand le critère est pré-identifié dans la base de contact ou stop entretien au moment du screening en début de questionnaire).

Cependant, le nombre de critères de quotas et leur niveau de détails imposent des contraintes au « Terrain » lors de la réalisation de l'enquête, contraintes qui peuvent soulever des problèmes d'allongement des temps de recherche, de faisabilité de la recherche et l'augmentation des coûts de production. Il est donc indispensable, lors de la définition du protocole d'enquête, de choisir les quotas adaptés à la problématique posée et aux besoins du client.