

Qu'est-ce-qu'un sondage ?

- **Un sondage est une enquête réalisée auprès d'un échantillon représentatif de la population étudiée. Les résultats obtenus auprès de l'échantillon sont ensuite extrapolés à la population totale.**
- 2 grands types de sondage existent :
 - **Le sondage politique est défini de manière légale depuis 2016. Un sondage politique est « une enquête statistique visant à donner une indication quantitative, à une date déterminée, des opinions, souhaits, attitudes et comportements d'une population par l'interrogation d'un échantillon ».** On peut distinguer 2 formes de sondage politique :
 - ✓ **Le sondage électoral pour connaître les intentions de vote des électeurs.** Il est réglementé par la loi. La réglementation principale porte sur la diffusion des résultats : depuis 2002, les sondages ne peuvent être publiés la veille et le jour de l'élection. Cette interdiction est contournée dans la mesure où elle ne s'applique que sur le territoire national et qu'elle n'empêche pas la publication de tels sondages sur des médias diffusant hors de France. Les sondages électoraux représentent 0,5% % du chiffre d'affaires des instituts de sondage.
 - ✓ **Le sondage d'opinion : donne l'avis la population sur des thèmes de société intéressant les sociologues, les économistes ou les hommes politiques (par exemple la réforme des retraites).** Il représente 9.5% % du chiffre d'affaires des instituts de sondage.
 - **Le sondage marketing : étude de marché qui, en grande majorité, n'est pas destinée à être rendue publique et qui s'intéresse aux comportements et opinions des consommateurs.** Il représente 75% du chiffre d'affaires des instituts de sondage.
- Le coût est très différent selon le type de sondage :
 - Un sondage politique publié dans un média coûte jusqu'à 12 000 euros ; un sondage marketing de 10 000 à 1.5 millions d'euros
 - Le coût très faible du sondage politique s'explique par le fait que c'est une **enquête omnibus : une enquête administrée pour le compte de plusieurs entreprises ou chaque entreprise se voit affecter une ou plusieurs des questions posées.**

Quels sont les principes du sondage ?

- Les sondages sont basés sur la théorie mathématique des probabilités. Celle-ci suppose que, pour connaître les événements qui peuvent survenir dans une population donnée, il suffit d'étudier une petite partie de celle-ci, à condition de respecter des règles rigoureuses de sélection de cette population pour assurer sa représentativité.
- Ainsi dans un sondage :
 - On étudie non pas l'intégralité, mais une partie de la population
 - On déduit alors les résultats de cette partie sur l'ensemble de la population avec une marge d'erreur
 - Cette marge d'erreur ne dépend pas de la taille de la population étudiée, mais de la taille de l'échantillon : avec 100 personnes interrogées, la marge d'erreur est de 10 %, avec 400 personnes, la marge d'erreur est de 5 %, avec 1000 personnes, la marge d'erreur est de 3 %, avec 10 000, la marge est de 1 %. Ainsi, plus la taille augmente, plus la marge d'erreur diminue. Cependant, cette baisse n'est pas proportionnelle : quand on passe d'un échantillon de 100 à 400 personnes, la taille est multipliée par 4, et la marge d'erreur est divisée par 2. Quand la taille passe de 1000 à 10 000, la taille est multipliée par 10, et la marge d'erreur seulement divisée par 3. Ainsi, la précision additionnelle est de plus en plus coûteuse ; les instituts de sondage se contentent donc généralement d'échantillons de 1 000 à 2 000 individus.

Comment construire un sondage ?

- **La construction de l'échantillon : en statistique, un échantillon est un ensemble d'individus représentatifs d'une population. La population de l'échantillon doit avoir les mêmes caractéristiques que la population totale.** Il existe deux méthodes pour obtenir un échantillon représentatif :

- La méthode aléatoire :
 - ✓ **Les modalités : on tire au hasard un certain nombre d'individus à partir d'un fichier complet de toute la population (le recensement de l'INSEE, la liste des abonnés à l'électricité ou au téléphone). Ensuite, ces personnes doivent être obligatoirement questionnées.**
 - ✓ Les intérêts de la méthode :
 - o La marge d'erreur est calculée de manière précise et scientifique
 - o Il n'y a pas de biais dans l'échantillon : il n'y a pas de sur ou de sous-représentation d'une catégorie de la population
 - ✓ Les limites de la méthode :
 - o Les listes exhaustives d'individus ou de ménages français sont très rares et ne sont pas accessibles pour la recherche publique. Seul l'INSEE peut utiliser, pour ces propres enquêtes, les listes des bulletins individuels du recensement. Or, ces listes ne sont pas mises à jour entre deux recensements.
 - o Il faut pouvoir trouver la personne sélectionnée : un nombre fréquent de passages à domicile est parfois nécessaire
 - o La personne sélectionnée peut refuser de répondre au sondage
- La méthode par quotas : est la plus employée par les instituts de sondage
 - ✓ **Les modalités : on construit un échantillon dans lequel les individus sont répartis selon les mêmes proportions que dans le groupe à étudier.** Il existe un grand nombre de caractéristiques possibles (hommes / femmes, âge, catégories socio-professionnelles, lieu de résidence, salariés du public / privé, etc). Le nombre de facteurs pris en compte dépend donc de la précision escomptée
 - ✓ Les intérêts de la méthode : c'est une méthode moins coûteuse et plus rapide qu'une méthode aléatoire. Les sondés sont interchangeables : il suffit de trouver une personne qui a les caractéristiques demandées (âge, sexe, profession,)
 - ✓ les limites :
 - o l'échantillon peut être biaisé car l'enquêteur n'interroge que les personnes présentes au moment de son passage.
 - o Elle ne permet pas de calculer scientifiquement la marge d'erreur du sondage. Les lois statistiques qui permettent de la déterminer ne valent théoriquement que pour les sondages aléatoires. En pratique, on considère cependant que la marge d'erreur des sondages par quotas est égale ou inférieure à celle des sondages aléatoires.

- **La construction du questionnaire :** la qualité du questionnaire est essentielle

- 2 grands types de questions :
 - ✓ La question fermée : **on propose une liste de réponses et le sondé doit en sélectionner une ou plusieurs**
 - o Les intérêts : la question fermée permet de tester une connaissance ou un comportement effectif. Le traitement des données est simple
 - o Les limites : elle peut donner l'impression que les choix sont limités et codifiés à l'avance.
 - ✓ La question ouverte : **le sondé répond librement et de manière développée**
 - o Les intérêts : la question ouverte permet d'évoquer l'ensemble des références spontanées du sondé. Ainsi, les réponses sont riches en informations.
 - o Les limites : les personnes interrogées ont davantage tendance à ne pas répondre aux questions ouvertes ou à répondre de manière superficielle. La question ouverte est plus difficile à analyser.
- Ces questions doivent être comprises par tous les sondés. Or une question n'a pas obligatoirement de signification claire. 2 difficultés :
 - ✓ 2 formulations apparemment proches peuvent être comprises de 2 manières différentes. Ainsi le terme « souvent » est vague : tous les jours ?, une fois par semaine ?, par an
 - ✓ Une même formulation peut avoir des sens différents selon les personnes interrogées ou le contexte : le terme insécurité est souvent compris par les catégories supérieures comme une menace de vol ; les personnes plus modestes le comprennent, en période de fort chômage, comme la peur de perdre un emploi.
- Il faut éviter les questions inductrices : **une question inductrice conduit la personne interrogée à répondre d'une certaine manière.** Ainsi, « êtes-vous d'accord » pousse le sondé à répondre positivement
- L'ordre des questions est aussi essentiel. **La position d'une question par rapport à d'autres peut induire une certaine réponse : c'est l'effet de halo.** Par exemple, une première question demande si « l'arrivée au pouvoir d'un parti a eu des effets économiques catastrophiques » et une deuxième « voterez-vous pour ce parti aux prochaines élections ». Dans ce cas,

la 1^{ère} question a un effet sur la réponse donnée à la deuxième : l'idée donnée dans la première question est que l'arrivée de ce parti a généré un ralentissement de la croissance économique. La personne ne va donc pas voter pour ce parti.

Comment interpréter les résultats ?

Les résultats sont alors utilisés pour tirer des conclusions. Pour les interpréter, il faut prendre certaines précautions :

- ☐ Il faut présenter les modalités de l'enquête : population étudiée, taille de l'échantillon, méthode d'échantillonnage utilisée (aléatoire ou quotas), la marge d'erreur
- ☐ Il faut ensuite préciser 2 taux de non-réponses :
 - le pourcentage et les caractéristiques des personnes qui ont refusé de répondre au sondage
 - le pourcentage de non-réponse à une ou plusieurs questions du sondage
- ☒ il faut indiquer si les données ont été redressées et comment. **Le redressement est une pratique de traitement de données pour corriger un biais de l'échantillon.** Il y en a deux formes :
 - pour corriger un biais structurel de l'échantillon : le redressement consiste à rendre l'échantillon conforme en proportions à la population de référence. Il affecte un poids (coefficient de pondération ou clé de correction) à chaque individu en fonction de ses caractéristiques.
 - pour corriger un biais observé a posteriori dans les réponses : dans le cas des sondages politiques, c'est le résultat final que l'on corrige, indépendamment de l'échantillon considéré. Par expérience, les sondeurs observent en effet un biais récurrent : le vote pour les partis d'extrême-droite ou d'extrême-gauche s'avoue difficilement devant l'enquêteur. En contrepartie, certaines formations, comme le PS, bénéficient toujours de plus d'intentions de vote que de votes réels