

OUTIL 12

Les données externes

“

Le statisticien peut aussi analyser des données externes à son organisation, et dont il n'est pas le producteur.

En quelques mots

En plus des données d'entreprise et des enquêtes par questionnaire, il est aussi possible d'accéder à des bases de données, gratuitement ou non. Des données d'enquête sont ainsi mises à disposition par **l'INSEE**, le **World Value Survey**, le **World Management Survey**, **Statista**... Ces enquêtes peuvent servir à étudier un sujet à un niveau national ou international, et ainsi à obtenir des résultats généraux qui peuvent ensuite être comparés à ceux de l'organisation, ou qui peuvent contribuer à éclairer certains aspects de l'organisation.

DIFFÉRENTES SOURCES DE DONNÉES EXTERNES

	Enquête internationale et intersectorielle sur les pratiques de management.
	Portail en ligne offrant des statistiques issues de données d'instituts, d'études de marché et d'opinion.
	Enquête internationale sur les valeurs des individus, actualisée régulièrement.
	Institut national de la statistique et des études économiques, met à disposition certaines de ses bases de données gratuitement.
	Réseau de diffusion et portail d'accès aux données françaises en sciences humaines et sociales.

? POURQUOI L'UTILISER ?

Objectif

Identifier différentes sources de données externes.

La liste présentée ci-contre n'est bien sûr pas exhaustive.

Contexte

Les données de sources externes peuvent être particulièrement utiles s'il n'est pas possible d'accéder à des données de l'entreprise et pas envisageable d'administrer un questionnaire, ou encore si l'analyse vise à obtenir des résultats génériques, dépassant le seul cadre de l'entreprise.

La diffusion de données d'enquête est de plus en plus courante et va sûrement se développer encore dans la mouvance de l'« open data ».

Ces données peuvent être de très bonne qualité. Elles peuvent permettre d'accéder à des informations intersectorielles ou internationales.

Les données d'enquêtes par questionnaire sont livrées avec le questionnaire et la méthodologie d'enquête. Il faut bien prêter attention à ces informations car elles donnent des éléments précieux pour évaluer la rigueur de l'enquête et le potentiel de généralisation des résultats. L'inconvénient des données de ce type est que l'analyste ne peut pas choisir les questions posées, la population interrogée, etc.

COMMENT L'UTILISER ?

Étapes

- 1. Chercher parmi différentes sources de données externes** celles qui se rapprochent le plus de votre objectif (périmètre, sujet d'enquête...). Ainsi, si votre enquête porte sur le travail et le management, l'enquête WMS sera sans doute la plus appropriée, notamment si vous souhaitez faire des comparaisons internationales. Si votre enquête porte sur un marché, la base Statista sera sans doute la plus proche de votre sujet.
- 2. Obtenir l'accès aux données** en respectant les règles et protocoles d'accès. Aujourd'hui, certains organismes de statistiques publiques ont mis en place des protocoles d'accès contraignants et onéreux (passage par des serveurs sécurisés avec identification par empreinte digitale par exemple), nécessaires pour garantir la sécurité des données.

Méthodologie et conseils

Google a récemment mis au point un moteur de recherche de bases de données, Google Dataset Search. Il permet d'identifier des bases de données à partir de mots-clés. Encore en version beta en 2019, ce moteur de recherche devrait devenir de plus en plus pertinent.

Les sources de données externes sont particulièrement utiles pour les études dépassant le cadre de l'entreprise : études sectorielles,

nationales, internationales. En revanche, leurs utilisateurs n'ont aucune maîtrise des questions posées et donc des variables présentes dans la base de données.

Avant de vous lancer...

- ✓ Il existe de nombreuses sources de données externes, gratuites ou non.
- ✓ Il faut veiller à ce que l'enquête identifiée corresponde aux objectifs visés par l'analyse : population enquêtée, questions posées, date de l'enquête...
- ✓ Cela peut permettre d'obtenir des résultats dépassant le cadre de l'entreprise, voire le cadre national.
- ✓ Cela peut également être utilisé dans le cadre d'une comparaison avec ce qui est observé au sein d'une organisation donnée.



COMMENT ÊTRE PLUS EFFICACE ?

Définir le sujet

La première étape pour trouver des données externes consiste à bien délimiter le sujet, et notamment les points sur lesquels vous pouvez accepter que les données s'éloignent un peu du sujet, et les points sur lesquels vous avez vraiment besoin de données précises.

Par exemple, si vous souhaitez obtenir une base de données sur le rapport au travail des individus, est-il plus important pour vous d'avoir une enquête contenant des variables précises (type d'emploi occupé, niveau hiérarchique, taille de l'équipe, de l'entreprise, type de contrat, catégorie professionnelle, etc.) ou bien d'avoir des informations sur l'importance générique qu'ils accordent au travail ?

Définir le périmètre souhaité

La seconde étape consiste à bien définir le périmètre souhaité. Le périmètre renvoie à la fois à la population étudiée (zone géographique,

âge, caractéristiques des répondants) et à la profondeur temporelle souhaitée.

Ainsi, vous pouvez souhaiter étudier le rapport au travail de l'ensemble des individus ayant entre 18 et 65 ans, ou bien au contraire inclure aussi des personnes plus jeunes, n'ayant éventuellement pas encore d'expérience professionnelle, ou encore des personnes plus âgées, ayant une expérience professionnelle mais ne travaillant plus.

En ce qui concerne la zone géographique, il faut savoir que certaines enquêtes sont nationales, alors que d'autres sont internationales. Enfin, il faut définir la profondeur temporelle et la date de production des données.

Accepter de faire évoluer le sujet en fonction des données disponibles

Contrairement au cas où vous pouvez construire votre propre questionnaire, vous n'avez pas de maîtrise sur les données externes que vous mobilisez : vous ne pouvez pas participer à la définition des variables, du périmètre, etc.

Il faut donc accepter de faire évoluer votre sujet en fonction de ces données.

Ainsi, si vous vous rendez compte que les données ne vous permettent pas de traiter le sujet « rapport des individus au travail en fonction de leur statut (salarié, inactif, chômeur, etc.) » car la base de données ne porte que sur des salariés, peut-être pouvez-vous le faire évoluer en « rapport des salariés au travail en fonction de leurs caractéristiques individuelles (genre, âge, etc.) ».

En tout état de cause, le fait de mobiliser des données externes suppose généralement une adaptation du sujet traité.

S'appropriier les données

Les données externes peuvent nécessiter un travail d'appropriation important, car vous n'avez pas participé à leur production.

Pour cela, vous pouvez vous aider des dictionnaires des variables et autres documents qui accompagnent la base de données, mais

également d'échanges et de contacts avec les producteurs de données qui pourront vous aiguiller sur certains points méthodologiques.



Cas d'entreprise

Exploiter des données externes

Beta est une petite entreprise souhaitant se développer dans le secteur des produits d'hygiène corporelle. Depuis plusieurs années, l'entreprise essaie d'augmenter ses ventes de produits, sans succès. Pour améliorer sa part de marché, l'entreprise souhaite :

- d'une part, modifier légèrement ses produits de façon à ce qu'ils correspondent mieux aux attentes des clients ;
- d'autre part, mener une campagne publicitaire permettant de convaincre les clients d'acheter les produits Beta.

L'entreprise se demande s'il est important de revoir en plus le packaging des produits. Elle hésite à entamer la démarche longue et coûteuse de commanditer une étude de marché à un prestataire externe.

Une recherche sur Google Dataset Search fait ressortir un ensemble d'indicateurs datant de 2015 disponibles sur le site Statista, et mentionnant le degré d'importance de chaque facteur de choix d'achat d'un produit d'hygiène corporelle.

Critère	Cité en premier par ...% de personnes
Mention sur la protection de la peau (hypoallergénique, etc.)	30%
Prix	23%
Marque	18%
Parfum	7%
Valeurs écoresponsables (bio, écologique, etc.)	7%
Caution médicale sur l'étiquette (testé par des médecins, etc.)	7%
Recommandation par un proche ou par les médias	4%
Packaging	1%

Cette enquête convainc donc l'entreprise de mener les actions suivantes :

- Revoir le parfum des produits pour leur donner une touche plus naturelle : un parfumeur est recruté pour définir les fragrances les plus adéquates, et des groupes de clients testeurs sont mis en place de façon à évaluer les parfums. Les parfums perçus comme les plus chimiques sont ainsi évacués, au profit de senteurs perçues comme plus naturelles.
- Employer plus de produits bio, quitte à augmenter le prix des produits. Les produits de la marque sont déjà en partie constitués d'ingrédients bio, mais en pourcentage insuffisant

pour obtenir une certification type Ecocert. L'objectif est donc d'atteindre un pourcentage d'ingrédients bio suffisant pour obtenir la certification.

- Ajouter sur l'étiquette les mentions « hypoallergénique » et « bio ».

Les produits de la marque ne contiennent effectivement pas de composants catégorisés comme allergènes.

- Axer la campagne publicitaire sur la dimension bio.

Le service marketing propose donc une campagne de publicité en ligne mettant en avant le respect de l'environnement et de la santé des clients.

Un an après la mise en œuvre de ces actions, l'entreprise Beta constate que sa part de marché est passée de 3 à 5 %.