

Analyse et exploitation des données

Table des matières

I. Techniques d'analyse de données	3
A. Connaissance client et démultiplication des données.....	3
B. Connaissance client et impact business	4
C. Base de données et retargeting.....	5
II. Analyse des données	6
A. Approches pour l'analyse de données	6
III. Exercice : Quiz	9
IV. Collecte de données	9
A. Différentes sources de données.....	9
V. Cartographier les données	13
A. Quoi, pourquoi, comment ?	13
VI. Exercice : Quiz	14
VII. Essentiel	15
VIII. Auto-évaluation	15
A. Exercice	15
B. Test.....	15
Solutions des exercices	16

I. Techniques d'analyse de données

Contexte

Dans un contexte digital, où presque toutes les entreprises du monde essaient d'intégrer le monde virtuel pour s'approcher davantage de leurs cibles et clients potentiels, l'expérience client est la clef des performances marketing et commerciales. Une marque se distingue auprès de ses clients en leur transmettant un message unique, différencié et cohérent lors de chaque étape de leur parcours d'achat.

C'est dans ce contexte de sur-mesure que la constitution et l'analyse des bases de données est stratégique pour les entreprises.

D'ailleurs, quand vous multipliez les points de contacts entre vous et vos clients, vous pourriez recueillir une très grande quantité d'informations qui constitueront votre base de données marketing.

Il est devenu donc important et surtout plus bénéfique d'utiliser l'analyse et les données prédictives pour prévoir le comportement des clients d'avance, lorsqu'ils interagissent avec votre marque.

D'ailleurs, l'objectif de l'analyse des bases de données client est de connaître le client de plus près pour :

- Une meilleure prise de décision de la façon d'attirer et de garder des clients,
- Une identification des clients à forte valeur ajoutée et l'élaboration d'une interaction efficace avec ces derniers.

A. Connaissance client et démultiplication des données

Connaître son client consiste à comprendre ses comportements d'achats, ses besoins, désirs et objectifs.

La digitalisation et la mise en place des stratégies *cross-canal* a entraîné une démultiplication des points de contacts riches en enseignement.

Exemple

- Les données collectées en temps réel par les différents cookies (Google Analytics ou MixPanel) du site internet de l'entreprise sur les parcours clients.
- Les statistiques journalières issues des interactions clients sur les différents réseaux sociaux (Facebook, Instagram, Pinterest, YouTube, etc.).
- Les pixels de suivi ou de traçage (comme celui de Facebook).
- Les données clients issues de leurs achats ou de leurs inscriptions aux *newsletters*.
- Les informations issues du service consommateur ou service après-vente.
- Les données des différentes campagnes marketing, qu'elles soient en magasin ou en ligne.
- Les données de panel distributeurs ou consommateurs disponibles périodiquement auprès des différents panels enseignes ou consommateurs de type AC Nielsen, IRI ou Kantar.

B. Connaissance client et impact business

C'est en analysant systématiquement ces données collectées que l'entreprise peut :

- Prendre des décisions sur la meilleure façon d'acquérir ou de fidéliser ses clients,
- Identifier quels sont les clients à forte valeur ajoutée et agir pro activement avec eux.

En exploitant ces données, l'entreprise peut mettre en place un plan marketing relationnel efficace. Elle peut ainsi influencer ses indicateurs clefs de succès (KPI), augmenter ses revenus en proposant des produits complémentaires répondant à ses besoins, mieux suivre sa relation client et détecter de potentiels problèmes comme le mécontentement de sa clientèle.

Exemple

Amazon est maître en la matière avec son système de suggestion très précis par rapport aux navigations clients, sa proposition systématique de lots basés sur les habitudes d'achat consommateurs et sa politique de présentation par avis consommateur.

Produits fréquemment achetés ensemble



Figure 1 - Illustration de ciblage par lot par Amazon

Au-delà du seul ciblage stratégique, analyser sa clientèle peut permettre un retour sur investissement supérieur de 115 % et des bénéfices supérieurs de 93 % (source : Segment¹).

1 <https://segment.com/>

C. Base de données et retargeting

Cette technique de marketing en ligne, appelée aussi **recyclage publicitaire**, illustre parfaitement le bénéfice du bon usage d'une base de données. Il s'agit de cibler les visiteurs d'une page ou d'un site avec une campagne de **publicité display** ciblée.

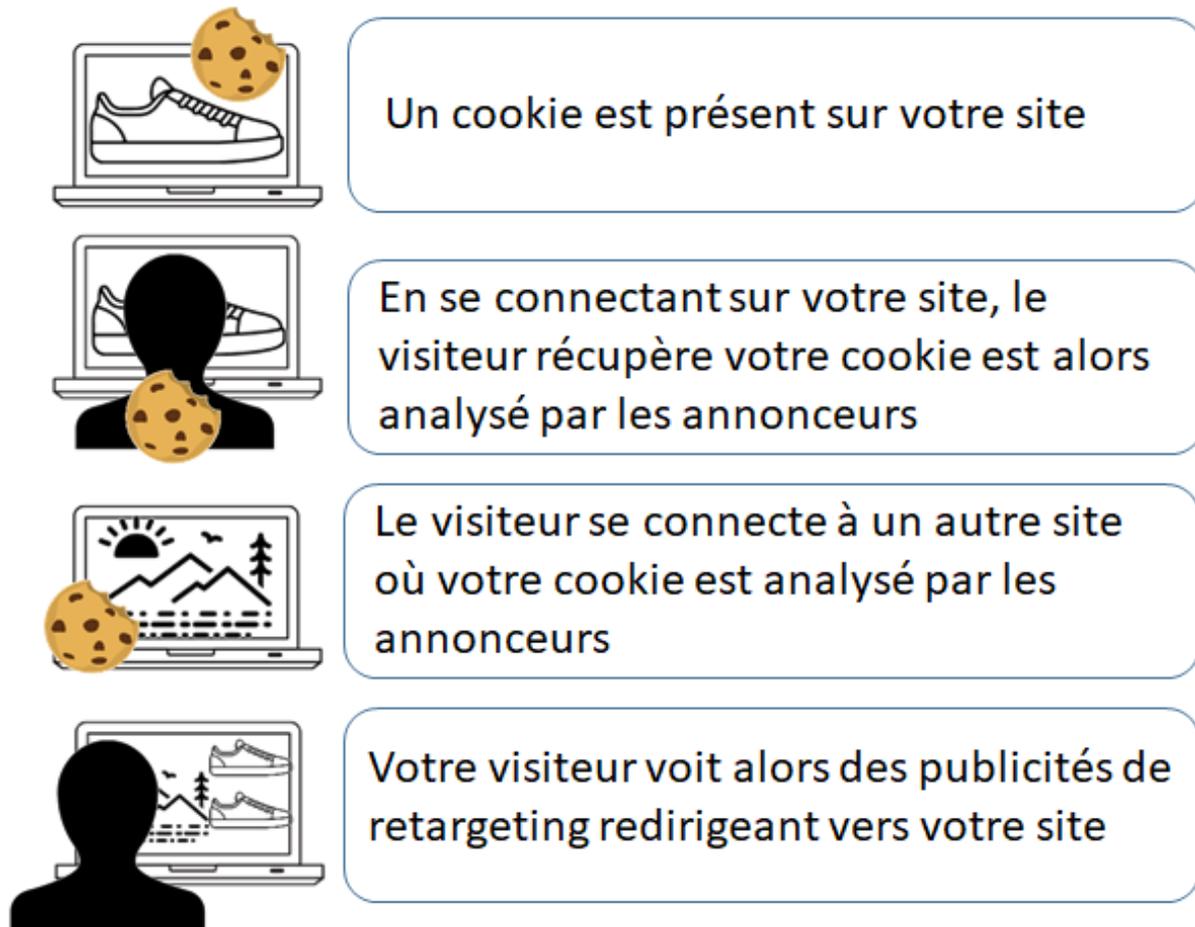


Figure 2 - Retargeting ou recyclage publicitaire

Cette méthode permet de convertir un visiteur à potentiel, réduire l'abandon de panier ou encourager le réachat. Et c'est un excellent moyen pour lutter contre la fatigue publicitaire de l'internaute tout en optimisant ses propres budgets publicitaires.

Que ce soit sous forme de bannières publicitaires, sur les réseaux sociaux ou par emailing, cette méthode ne peut être efficace sans la constitution et l'analyse précise d'une base de données qualifiée.

Si les entreprises ont de plus en plus conscience de l'importance d'analyser leurs données clients, bons nombres ne parviennent pas à constituer ou à maîtriser ses bases de données.

C'est pourquoi l'analyse et l'exploitation de data en entreprise constituent l'axe de cette formation. Elle a pour but de fournir les différents outils, à la fois méthodologiques et informatiques, pour une mise en œuvre efficace des données accessibles en entreprise spécifiquement et plus généralement dans les organisations. Elle étirent les outils suivants :

- Les approches pour l'analyse de données
- Les techniques de la collecte de données
- Comment cartographier les données

II. Analyse des données

A. Approches pour l'analyse de données

Définition

Les **méthodes quantitatives** s'allient au rassemblement de données générales structurées sur un grand nombre de personnes (généralement plus de 100). Elles sont très bénéfiques pour observer des phénomènes et mesurer les connexions entre différents facteurs.

Quant aux **méthodes qualitatives**, elles correspondent à l'analyse de matériaux non-structurés (texte, parole) d'un très petit nombre de personnes (généralement moins de 100). Cela sert à rassembler des vues assez complexes et obtenir un aperçu d'une situation précise.

La structuration du choix entre les 2 approches est très importante et sert à améliorer les résultats de la recherche qu'on a préalablement menée. De plus, dans certains cas, ces 2 méthodes peuvent être combinées.

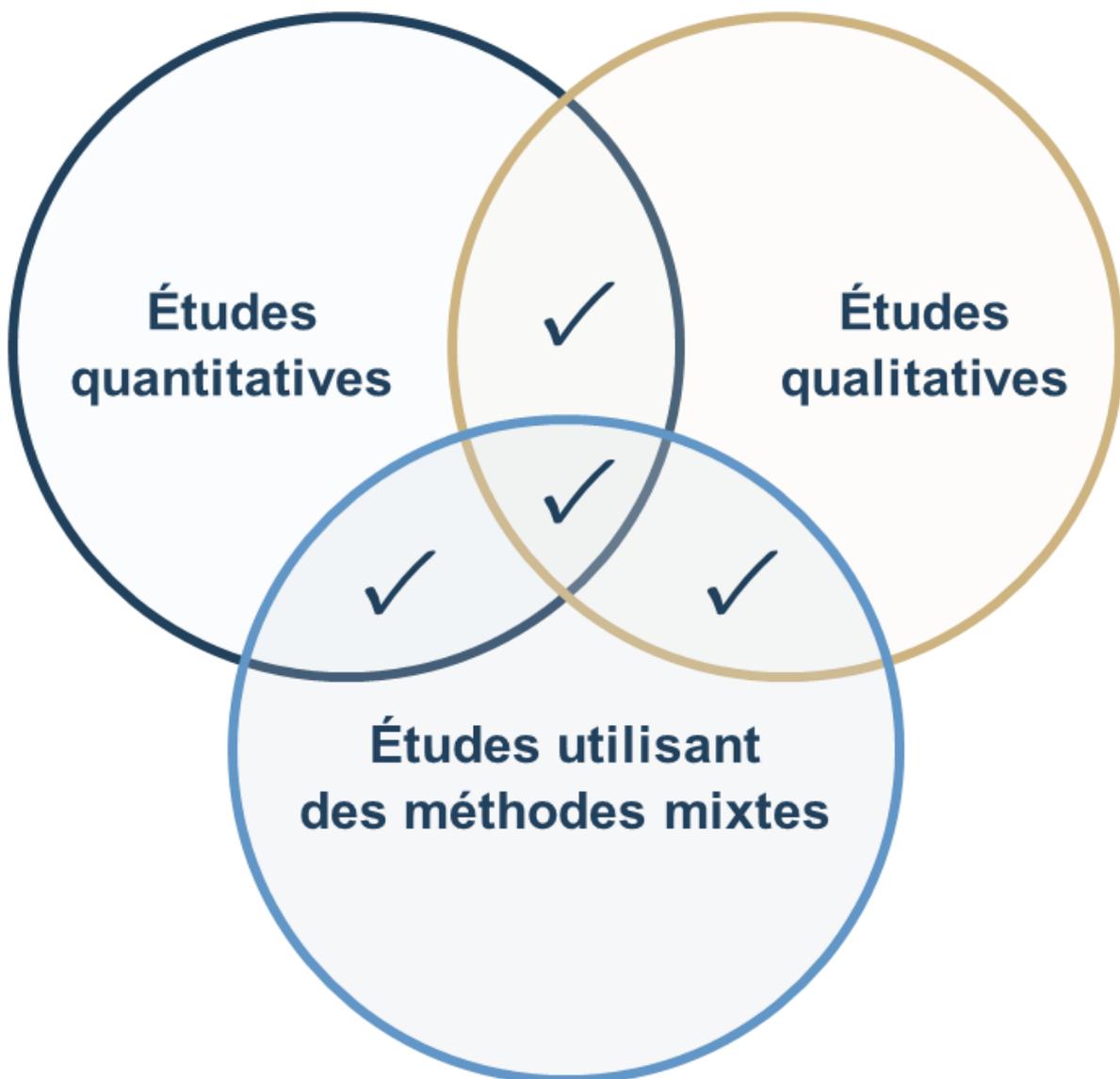


Figure 3 - L'approche quantitative et l'approche qualitative pour l'analyse de données

L'alliance entre une approche quantitative et une qualitative est possible. À titre d'exemple, vous pouvez mener des entretiens exploratoires (méthodes qualitatives) pour pouvoir définir les hypothèses qui permettent de créer des questionnaires, puis mener une gestion en ligne de ces derniers (méthodes quantitatives).

Méthode

Procédez comme suit :

1. Fixer les objectifs de l'étude qu'on va mener.
2. Déterminer le périmètre de l'étude : l'approche quantitative convient plus aux grands échantillons.
3. Lister les matériaux et données dont on dispose déjà : si un certain nombre de données structurées existent (disponibles sur le système d'information par exemple), les méthodes quantitatives peuvent être plus faciles à mettre en œuvre que les méthodes qualitatives.
4. Définir les ressources financières, humaines et informatiques disponibles.
5. Déterminer la méthode d'accès aux données nécessaires : par exemple, selon l'organisation, des fois, il est moins compliqué de mener des entretiens ou, au contraire, de partager des enquêtes en ligne.

Quels sont les objectifs ?

Plusieurs objectifs sont visés à travers la mise en œuvre d'une approche quantitative : décrire, comprendre, expliquer et prédire.

4 questions peuvent répondre à ces objectifs :

- **Décrire** : comment est la situation ?
- **Comprendre** : quelle est la situation ?
- **Expliquer** : pourquoi cette situation ?
- **Prédire** : quelle est la prochaine situation ?

Avant de commencer la mise en œuvre des outils statistiques, les objectifs qu'on aura à fixer doivent être clairement définis. En fait, la méthode utilisée sera différente en fonction du résultat choisi et le résultat ne sera pas le même.

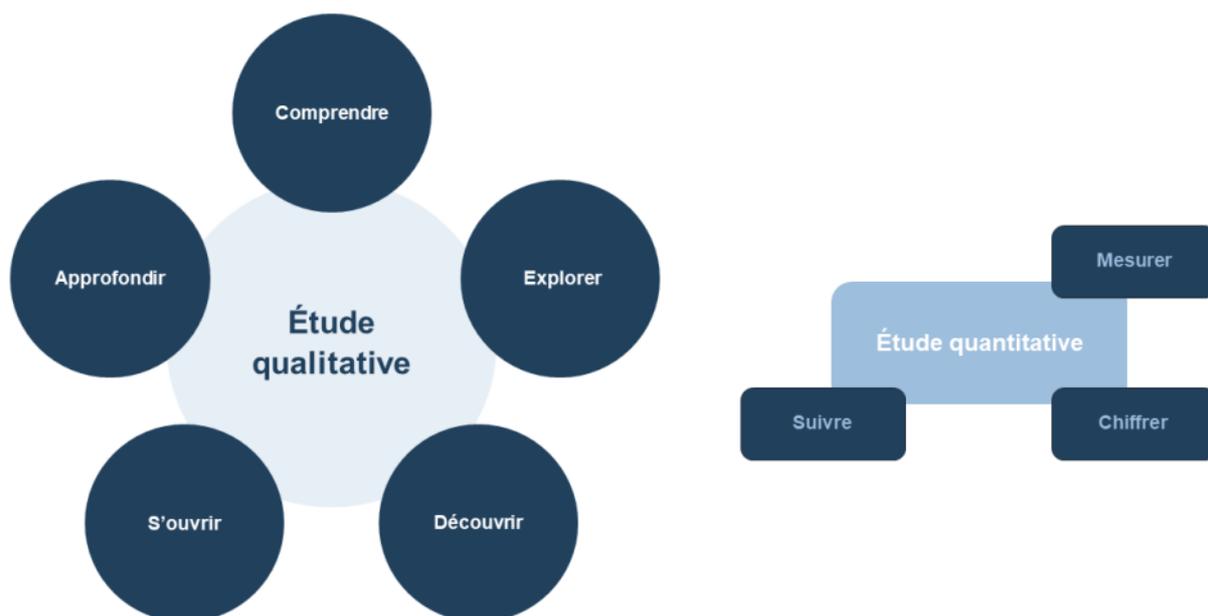


Figure 5 - Comment les objectifs conditionnent la démarche

Quels sont les moyens ?

Avant de commencer la procédure en mobilisant des statistiques ou une méthode quantitative, nous devons tout d'abord déterminer les moyens qu'on possède sur plusieurs niveaux :

- **Au niveau des compétences et connaissances** : de quelles connaissances et compétences je dispose sur les statistiques et les logiciels ? S'agit-il de connaissances qu'on devra acquérir rapidement ?
- **Au niveau des données** : quelles sont les bases de données dont je dispose ?
- **Au niveau financier** : à combien mon budget s'élève ?

Beaucoup de temps et d'argent doivent être investis en termes d'un projet d'étude statistique. Alors, il vaut mieux s'assurer avant même d'entamer l'étude que l'on dispose de tous les moyens nécessaires pour assurer la réussite du projet souhaité.

Voici les moyens dont on parle :

- Compétences et connaissances dans le **domaine statistique** : maîtrise des notions statistiques de base, maîtrise des logiciels de statistiques, notamment.
- **Bases de données** qui suffiront à traiter les problèmes rencontrés. Ces données ne doivent pas être agrégées à l'avance. L'agrégation à l'avance limite en effet les possibilités de références croisées et d'analyse.

Les différents logiciels de statistiques

Plusieurs logiciels de statistiques existent sur Internet aujourd'hui, on peut considérer Excel, R, SAS, SPSS et Stata les plus fréquentés par les internautes (liste non exhaustive).

Avant de choisir celui à utiliser, vous devez comprendre chaque fonctionnalité à part. Afin de faire un choix judicieux, vous pouvez lister les caractéristiques en suivant les critères ci-dessous :

- Rapport qualité / prix
- Facilité de la manipulation des interfaces
- Etc.

On compare les nombreux logiciels entre autres à partir de critères :

- Facilité d'usage (ce critère compte également pour les utilisateurs qui n'excellent pas en matière du codage) : Excel ou SPSS sont plus faciles à manipuler,
- Complexité des analyses possibles : la version basique d'Excel est particulièrement limitée en cela par rapport aux autres logiciels.

Conseil

Les meilleurs logiciels de statistiques et analyse de données :

- **MICROSOFT EXCEL**
- **STATA**
- **IBM SPSS STATISTICS**
- **Le logiciel R**
- **SAS**
- **STATISTICA**
- **STATISTIX**
- **MATLAB**

Exercice : Quiz

[solution n°1 p.17]

Choisissez la bonne réponse aux questions ci-dessous :

Question 1

L'analyse des données est utilisée à la fois en analyse prédictive et descriptive.

- Vrai
- Faux

Question 2

Pour quelle raison principale les entreprises ont recours à l'analyse des données ?

- Découvrir des tendances cachées dans les bases de données
- Réduire la quantité exponentielle des données dormant dans les entrepôts de données
- Utiliser les NTIC pour ne pas être obsolète

Question 3

Quel est l'un des intérêts de l'analyse des données ?

- L'avantage d'une longueur d'avance sur ses concurrents
- La conservation des données dans des entrepôts pour améliorer la connaissance du client et pour créer de nouvelles expériences client, publicité ou autre démarche marketing. Et donc avoir un avantage sur le concurrent

Question 4

Quelles sont les techniques utilisées par l'analyse des données ?

- Statistiques
- Publicités

Question 5

Le Datamining peut enfreindre le respect de la vie privée des personnes concernées.

- Vrai
- Faux

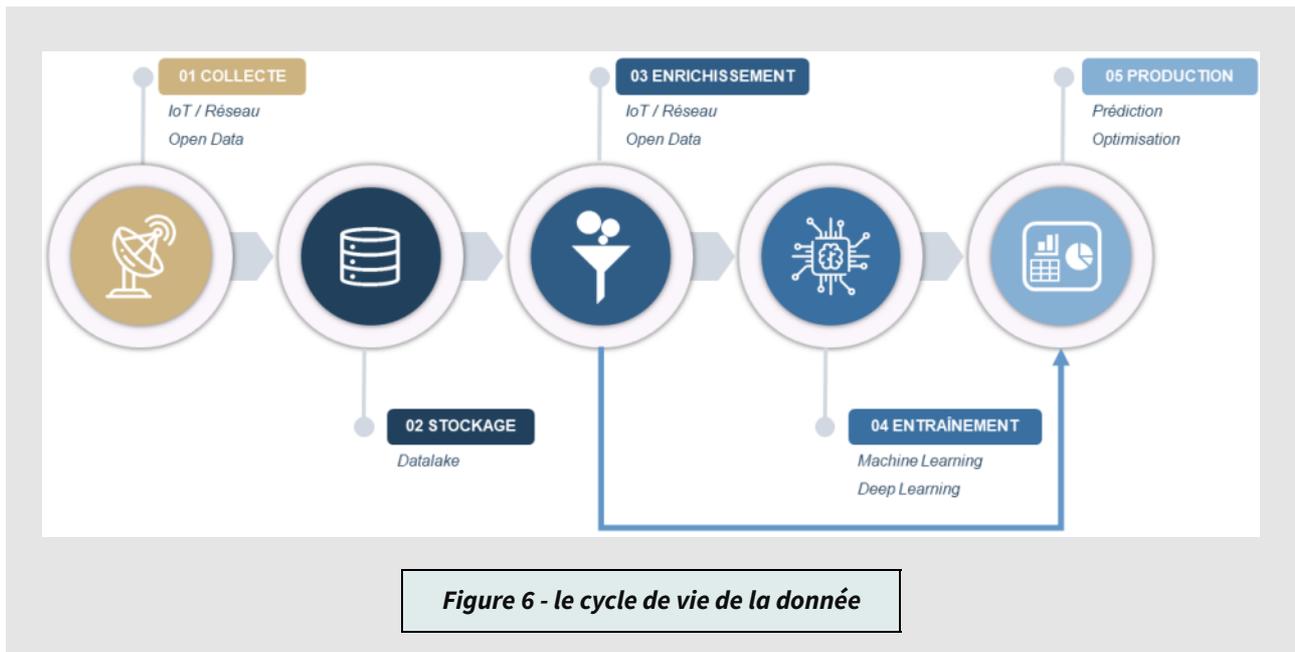
IV. Collecte de données

A. Différentes sources de données

Méthode

L'une des premières étapes de toute recherche statistique correspond à trouver des bases de données qui existent déjà. Par conséquent, dans une entreprise, il peut déjà y avoir une base de données d'employés, mais il existe également une base de données de clients sur les produits.

Nous ne devons pas non plus ignorer l'existence de bases de données statistiques publiques, qui peuvent fournir des informations intéressantes au niveau national ou international. Par ailleurs, si vous ne pouvez pas utiliser les données existantes pour la recherche, vous devez collecter les données vous-même.



Définition

Les données d'entreprise

Les entreprises ont généralement beaucoup de données sur leurs clients, produits, employés, etc. Ces données se stockent dans un **Système d'Information** (SI). Bien qu'il existe de nombreuses règles (anonymat, confidentialité, etc.) à respecter pour avoir accès à ces données, il est très intéressant de les étudier.

Comment l'utiliser ?

Le **Système d'Information des Employés** (SIE) contient des informations sur tous les employés de l'entreprise.

Premièrement, pour des raisons juridiques (telles que les rapports sociaux), ils peuvent être encouragés à créer et à maintenir ces bases de données.

Deuxièmement, à mesure que la quantité de données augmente, la numérisation de l'entreprise a exacerbé cette tendance.

Méthode

1. Définissez le thème et la population à étudier.
2. Essayez de déterminer s'il existe au sein de l'entreprise une base de données permettant de résoudre ce sujet : systèmes d'information, enquêtes menées, réseaux sociaux internes.
3. Respecter les règles de sécurité et les accords relatifs à l'utilisation de ces données.

Les données externes

Un accès gratuit à la base de données s'ajoute aux données d'entreprise et des enquêtes par **questionnaire**. Les données d'enquête peuvent être obtenues auprès de l'INSEE, World Value Survey, World Management Survey, Statista, etc.

Elles contribuent à l'étude des sujets au niveau national ou international, à obtenir des résultats de manière générale, qui peuvent être comparés aux résultats de l'entreprise mais aussi à clarifier certains aspects de l'organisation.



Source : <https://experts.mc2i.fr/>

Figure 7 - les ressources externes

Si les données ne sont pas accessibles depuis l'intérieur de l'organisation et qu'il n'est pas possible de gérer des questionnaires, ou si le but de l'analyse est de fournir des résultats généraux au-delà du périmètre de l'organisation, les données provenant de sources externes seront particulièrement utiles. La diffusion des données d'enquêtes devient de plus en plus courante et se développera certainement dans le mouvement des « *données ouvertes* ». Ces données peuvent être très bonnes en termes de qualité.

Ils peuvent donner accès à des informations interdépartementales ou internationales. Les données du questionnaire sont fournies avec le questionnaire et la méthode d'enquête. **Ces deux dernier point sont très importants car ils permettent d'évaluer la rigueur de de l'enquête et le potentiel de généralisation des résultats.**

Remarque

L'inconvénient de ce type de données est que l'analyste ne peut pas choisir la question à poser, la population à enquêter, etc.

Exemple

1. Essayez de trouver des sources de données externes différentes (périmètre, sujets d'enquête, etc.) qui sont les plus proches de vos objectifs.

Par exemple, si votre enquête concerne le travail et la gestion, une enquête WMS peut être la plus adéquate, surtout si vous comptez effectuer des comparaisons internationales. Dans le cas échéant, la base de données Statista peut être la base de données la plus convenable pour votre étude.

2. Gagnez l'accès aux données selon les règles et les protocoles mis en place. De nos jours, certaines organisations d'études statistiques publics ont instauré des protocoles d'accès restrictifs et coûteux (par exemple, via un serveur sécurisé avec fonction de reconnaissance d'empreintes digitales), ce qui est essentiel pour assurer la sécurité des données
3. Trouvez facilement les données de recherche qui vous intéressent dans plusieurs bases de données différentes. Si un individu dispose d'un identifiant personnel (adresse e-mail, numéro de téléphone, numéro d'employé, numéro de produit, etc.) partagé par les 2 bases de données, les différentes bases de données peuvent être mises en correspondance.

Cette opération doit être réalisée dans le cadre du respect des règles de confidentialité des données.

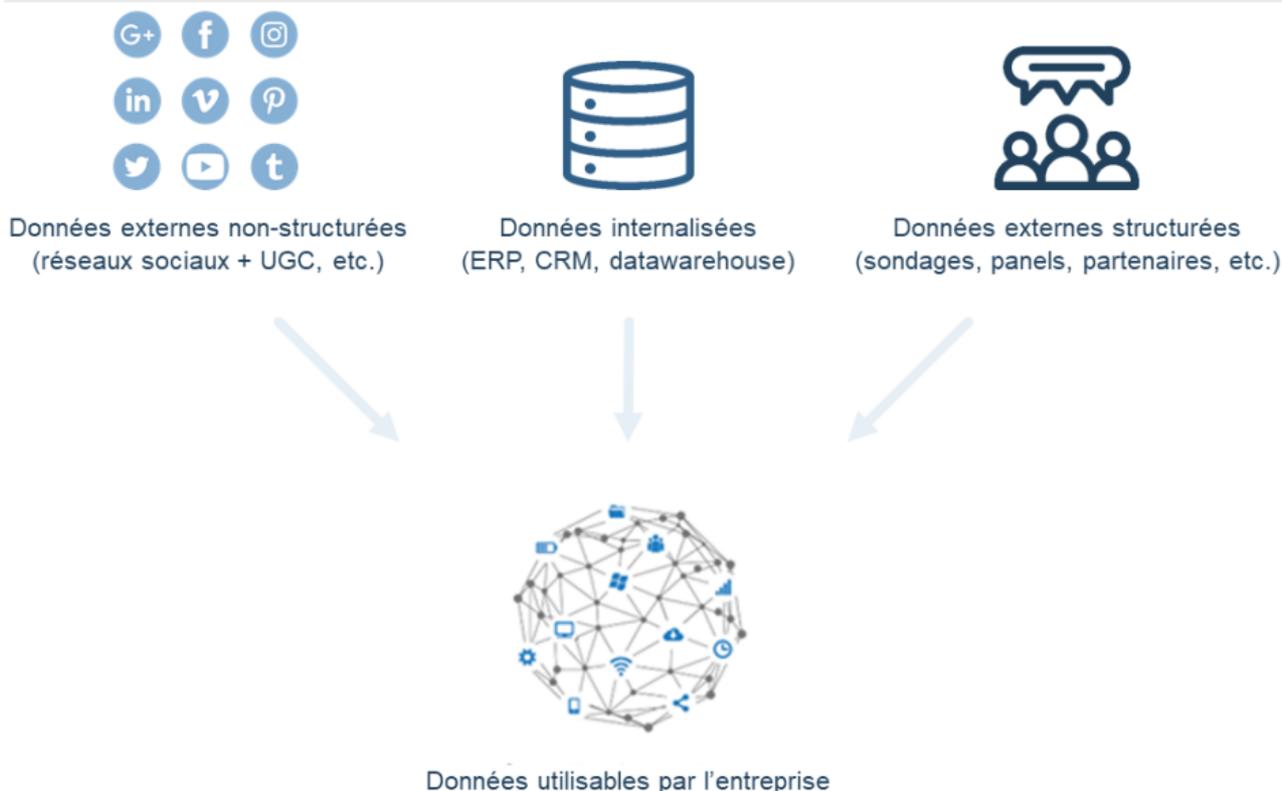


Figure 8 - la combinaison des ressources externes

Remarque

Il n'est pas recommandé de « copier et coller » les 2 bases de données et de les placer l'une à côté de l'autre. D'ailleurs, les personnes peuvent être classées dans un ordre aléatoire, ou simplement changer la disposition entre les 2 bases de données, ce qui peut entraîner des erreurs.

Par conséquent, il est impératif d'appairer les bases de données, ce qui y garantira la correspondance des identifiants. De toute façon, l'individu doit être identifié au moyen de variables publiques.

V. Cartographier les données

A. Quoi, pourquoi, comment ?

Comme les entreprises collectent des données en grande quantité, elles doivent être utilisées. C'est-à-dire qu'elles sont analysées, classées, bien ciblées et enfin qualifiées.

Cependant, comme elles ne se limitent pas à une base fixe et statique (comme un trophée dans une vitrine), elles doivent être rendues accessibles et à la portée de tous les employés, sans limites ni restrictions techniques. C'est le but de la cartographie des données : **établir un terrain commun pour tous les collaborateurs.**

Définition Qu'est-ce que la cartographie des données ?

En bref, le **mappage de données** est un moyen de visualiser les points de saisie et de traitement des données. Cette forme topographique est utilisée dans différents systèmes d'information de l'entreprise. Elle permet à tous les collaborateurs de la maîtriser, et le même langage « *data* » peut être autant utilisé entre les data scientists et les responsables informatiques qu'entre les services métiers qui s'y connaissent moins.

Méthode

La carte se crée en utilisant 3 principaux outils :

- **Outils sémantiques**, y compris la liste des données spécifiques à l'entreprise et des métadonnées des objets métier dans le vocabulaire métier, pour optimiser la compréhension et le contexte des données.
- **Modèle de données**, qui fournit des instructions précises sur la façon de modéliser les données et de les stocker dans divers systèmes de stockage (structurés, semi-structurés ou non-structurés). Un autre outil est également important : le dictionnaire de données techniques.
- **Outil de conception de traitement de flux de données**, qui nous renseigne sur comment transformer, manipuler et traiter les données à travers les différents SI (Système d'Information) de l'entreprise.

Remarque

En plus de ces 3 outils, nous devons également normaliser le format des conditions de disponibilité, d'accès et d'utilisation des données afin que les graphiques partagés ne rencontrent pas d'obstacles liés à la confidentialité et à la compréhension complète des données.

Pourquoi ce processus est-il capital ?

Avoir son œil sur l'ensemble de données existant dans l'entreprise : c'est la solution apportée par le processus de cartographie des données. L'objectif est de la rendre universelle (au moins au niveau de l'entreprise), même si elle est totalement accessible et compréhensible par tous les salariés, afin que chacun puisse identifier la source d'une donnée, mesurer son calcul, trouver d'éventuels licenciements, etc.

Exemple

Supposons ceci : vous cherchez l'adresse de facturation d'un client. Selon le glossaire de l'entreprise, la compréhension du terme « *adresse* » peut changer considérablement, et l'incapacité de suivre l'historique des mises à jour vous fera vous perdre dans le réseau complexe de bases de données - rien ne garantit que vous pourrez trouver ce que vous recherchez.

Ainsi, les données cartographiques peuvent établir un lien entre **l'idée technique** (le stockage et la conversion des données dans l'application SI) et **l'idée métier** (comment l'utiliser dans différents services).

Ce processus vise à briser les silos historiques et à décomposer les départements. Ou, en d'autres mots, à créer un espace public où tous les collaborateurs utiliseront le même langage, afin de fournir à chacun de meilleurs outils de gestion des opérations de données.

Comment tirer le meilleur parti de la cartographie des données ?

Une bonne data map nécessite un processus pédagogique : transformer une structure historiquement en silos (des cloisons techniques séparent chaque département, ce qui limite le flux commode d'information et d'utilisation au sein de l'entreprise) en un lieu vivant, évolutif et façonné par les salariés d'eux-mêmes, et tirer profit de leur expérience.

Une approche descendante du haut n'est pas la seule méthode, de ce fait, elle mettra toute cartographie obsolète à la corbeille juste après son lancement. La seule manière que cela peut s'effectuer est de manière fréquentative.

Ainsi, les projets et plans sont suivis de près, ce qui est au cœur d'une approche globale intégrée aux tâches quotidiennes.

Remarque

Chacun peut participer à la cartographie des données métiers granulaires à son propre niveau. Ce n'est que plus tard qu'on aura la possibilité d'allier les initiatives individuelles afin d'élargir le périmètre et générer des données agrégées basées sur des caractéristiques assez fréquentes. Une approche ascendante absolue.

Finalement, notre objectif est d'acquérir une carte contenant tous les services et applications du SI pour :

- Offrir une vue de tous les angles des données qui circulent dans la structure,
- Permettre à tous les employés d'en tirer profit.

Exercice : Quiz

[solution n°2 p.18]

Choisissez la ou les bonnes réponse(s) aux questions ci-dessous :

Question 1

Quels sont les outils qui permettent de créer la cartographie des données ?

- Outils sémantiques
- Modèle de données
- Outil de conception de traitement de flux de données

Question 2

Quels sont les objectifs qui s'associent à l'analyse de données ?

- Déterminer les facteurs qui expliquent quelques comportements
- Réduire la taille des données
- Augmenter la taille des données

Question 3

Quelles sont les différentes sources de données d'une entreprise ?

- Base de données des employés
- Base de données des clients
- Base de données des produits

Question 4

Est-ce que l'entreprise peut intégrer des données externes dans l'analyse de ses données ?

- Non
- Oui

Question 5

Quels sont les critères pour choisir un logiciel de statistique ?

- Prix
- Facilité de la manipulation des interfaces

VII. Essentiel

Plusieurs objectifs s'associent à l'analyse de données :

- Une meilleure perception de la situation ou du phénomène,
- Déterminer les facteurs qui expliquent quelques comportements,
- Définir des politiques et mesures plus efficaces en termes d'objectifs fixés.

Notre formation vise spécifiquement à vous faire comprendre des concepts de base et acquérir des compétences analytiques pour pouvoir réaliser vos différents objectifs. De plus, après la phase d'analyse, les résultats perçus et les recommandations formulées doivent être fournis de manière efficace et pédagogique.

VIII. Auto-évaluation

A. Exercice

Question 1

[solution n°3 p.19]

Quelles sont les différences entre les méthodes quantitatives et les méthodes qualitatives dans l'analyse de données ?

Question 2

[solution n°4 p.19]

Comment faire le retargeting ?

B. Test

Exercice 1 : Quiz

[solution n°5 p.19]

Choisissez la bonne réponse aux questions ci-dessous :

Question 1

Quelles peuvent être les applications de l'analyse des données spécifiques à l'industrie ?

- La segmentation de données
- L'hétérogénéisation de la variété des données

Question 2

L'analyse des données est synonyme d'opportunités tels que :

- La fouille de données sans l'aide de la main de l'homme
- L'interprétation des résultats de la fouille de données

Question 3

Quels sont les principaux défis à relever de l'analyse des données ?

- Réduire l'intelligence de l'exploration de données
- Augmenter la vitesse de la compilation de données

Question 4

Quelles sont les limites techniques de l'analyse des données ?

- Elle ne permet pas de décrire les données
- Elle ne peut traiter que des données chiffrées

Question 5

Quelles sont les caractéristiques des données provenant de la fouille du Web ?

- Leur flot est limité
- Leurs paramètres ne se modifient jamais avec le temps

Solutions des exercices

Exercice p. 9 Solution n°1**Question 1**

L'analyse des données est utilisée à la fois en analyse prédictive et descriptive.

- Vrai
- Faux
-  L'analyse des données (aussi appelée Analyse Exploratoire des Données ou AED) est une famille de méthodes statistiques dont les principales caractéristiques sont d'être multidimensionnelles et descriptives.

Question 2

Pour quelle raison principale les entreprises ont recours à l'analyse des données ?

- Découvrir des tendances cachées dans les bases de données
- Réduire la quantité exponentielle des données dormant dans les entrepôts de données
- Utiliser les NTIC pour ne pas être obsolète
-  L'analyse des quantités énormes de données qui viennent du marché, et cela en temps réel, permet aux entreprises de prendre les bonnes décisions stratégiques et de mener des actions intelligentes.

Question 3

Quel est l'un des intérêts de l'analyse des données ?

- L'avantage d'une longueur d'avance sur ses concurrents
- La conservation des données dans des entrepôts pour améliorer la connaissance du client et pour créer de nouvelles expériences client, publicité ou autre démarche marketing. Et donc avoir un avantage sur le concurrent
-  Le terme Entrepôt De Données ou EDD (ou base de données décisionnelle ; en anglais, *data warehouse*) désigne une base de données utilisée pour collecter, ordonner, journaliser et stocker des informations provenant de base de données opérationnelles et fournir ainsi un socle à l'aide à la décision en entreprise.

Question 4

Quelles sont les techniques utilisées par l'analyse des données ?

- Statistiques
- Publicités
-  La publicité car elle permet de récupérer des données (nb de clic, nb de personne ayant acheté, etc.) pour les chiffrés et en faire des statistiques. La méthode utilisées est donc les statistiques appliqué à la communication marketing de l'entreprise (la publicité).

Question 5

Le Datamining peut enfreindre le respect de la vie privée des personnes concernées.

- Vrai
- Faux

Q Le traitement des données avec le datamining transforme les données sous format anonyme avant le traitement pour respecter la vie privée des personnes concernées.

Exercice p. 14 Solution n°2

Question 1

Quels sont les outils qui permettent de créer la cartographie des données ?

- Outils sémantiques
- Modèle de données
- Outil de conception de traitement de flux de données

Q La carte se crée en utilisant 3 principaux outils : outils sémantiques, y compris la liste des données spécifiques à l'entreprise et des métadonnées des objets métier dans le vocabulaire métier. Modèle de données, qui fournit des instructions précises sur la façon de modéliser les données et de les stocker dans divers systèmes de stockage. Outil de conception de traitement de flux de données, qui nous renseigne sur comment transformer, manipuler et traiter les données à travers les différents SI de l'entreprise.

Question 2

Quels sont les objectifs qui s'associent à l'analyse de données ?

- Déterminer les facteurs qui expliquent quelques comportements
- Réduire la taille des données
- Augmenter la taille des données

Q Les objectifs : une meilleure perception de la situation ou le phénomène. Déterminer les facteurs qui expliquent quelques comportements. Définir des politiques et mesures plus efficaces en termes des objectifs fixés.

Question 3

Quelles sont les différentes sources de données d'une entreprise ?

- Base de données des employés
- Base de données des clients
- Base de données des produits

Q Les entreprises ont généralement beaucoup de données sur leurs clients, produits, employés, etc. Ces données se stockent dans un Système d'Information (SI).

Question 4

Est-ce que l'entreprise peut intégrer des données externes dans l'analyse de ses données ?

- Non
- Oui

Q Un accès gratuit à la base de données s'ajoute aux données d'entreprise et des enquêtes par questionnaire. Les données d'enquête peuvent être obtenues auprès de l'INSEE, World Value Survey, World Management Survey, Statista, etc.

Question 5

Quels sont les critères pour choisir un logiciel de statistique ?

- Prix
- Facilité de la manipulation des interfaces
-  Afin de faire un choix d'un logiciel de statistique, vous pouvez lister les caractéristiques suivant les critères ci-dessous : rapport qualité / prix et la facilité de la manipulation des interfaces.

p. 15 Solution n°3

Les études quantitatives et qualitatives sont des méthodes complémentaires que vous pouvez associer dans vos sondages pour obtenir des résultats à la fois détaillés et à large portée.

Pour faire simple, les données quantitatives fournissent les chiffres qui valident les points généraux de votre étude alors que les données qualitatives apportent les détails et le contexte pour en comprendre toutes les implications.

p. 15 Solution n°4

Comment fonctionne le retargeting ? Le retargeting fonctionne grâce à l'utilisation de cookies qui permettent de tracer la navigation des internautes ayant visité votre site web, et de les identifier sur d'autres sites. Lorsqu'un internaute visite votre site, un cookie est téléchargé sur son navigateur.

Exercice p. 15 Solution n°5**Question 1**

Quelles peuvent être les applications de l'analyse des données spécifiques à l'industrie ?

- La segmentation de données
- L'hétérogénéisation de la variété des données
-  La segmentation d'un marché consiste à le découper analytiquement en sous-marchés homogènes. Cette analyse se pratique en particulier mais pas uniquement dans le domaine du marketing.

Question 2

L'analyse des données est synonyme d'opportunités tels que :

- La fouille de données sans l'aide de la main de l'homme
- L'interprétation des résultats de la fouille de données
-  L'analyse de données, connue aussi sous l'expression de fouille de données, forage de données, prospection de données, data mining, ou encore extraction.

Question 3

Quels sont les principaux défis à relever de l'analyse des données ?

- Réduire l'intelligence de l'exploration de données
- Augmenter la vitesse de la compilation de données

Q De toute évidence, l'un des principaux défis de l'analyse de données à relever est de parvenir à stocker, à analyser toutes les informations et d'augmenter la vitesse de la compilation de données.

Question 4

Quelles sont les limites techniques de l'analyse des données ?

- Elle ne permet pas de décrire les données
- Elle ne peut traiter que des données chiffrées
- Q Si les données sont cryptées alors l'analyse des données ne peut pas comprendre le sens de ces données.

Question 5

Quelles sont les caractéristiques des données provenant de la fouille du Web ?

- Leur flot est limité
- Leurs paramètres ne se modifient jamais avec le temps
- Q Une copie de données traités est toujours enregistrée dans le disque lors de la fouille de données (sauf dans le cas de Big Data).