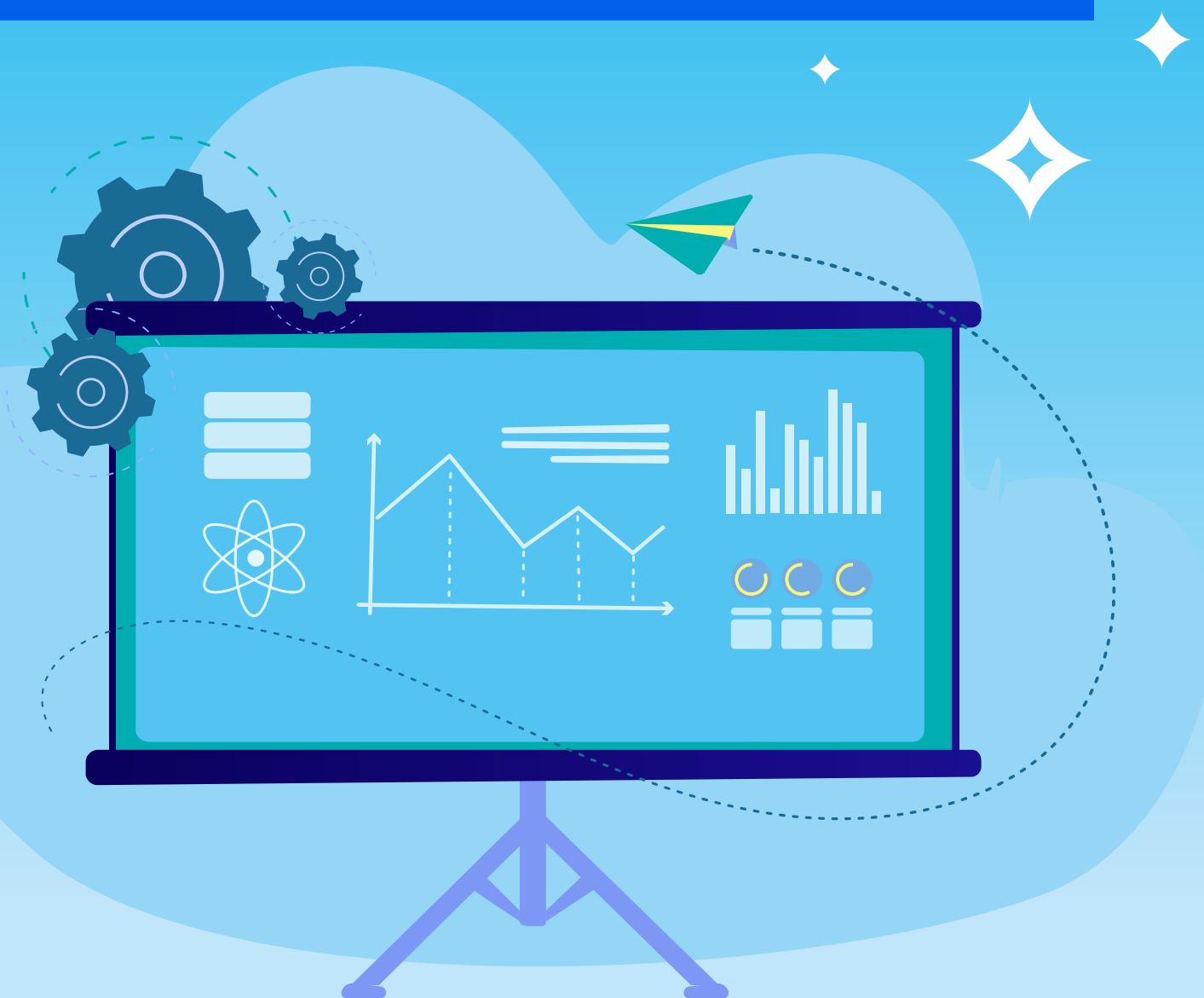


[LIVRE BLANC]

DE LA THÉORIE À LA PRATIQUE : COMMENT LA DATA SCIENCE AMÉLIORE VOTRE SUPPLY CHAIN



Si les entreprises sont aujourd'hui de plus en plus nombreuses à reconnaître la data science comme une solution puissante au service de l'organisation et de la planification,

le sujet reste mal maîtrisé et les termes qui y font référence sont souvent employés de façon approximative. Dans ce contexte, il est important de reprendre les bases afin de bien comprendre ce qu'est la data science, pour être en mesure d'en saisir ensuite toutes les opportunités, notamment au service de la Supply Chain. Quel que soit votre domaine d'activité, ce livre blanc vous éclaire ainsi sur les enjeux et les utilisations de la data science pour améliorer votre Supply Chain.



01

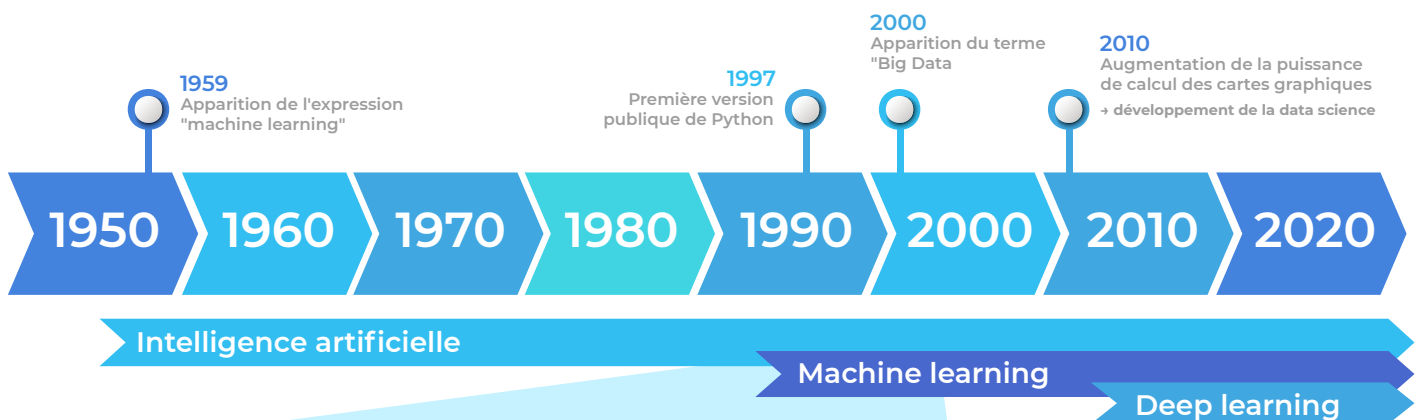
QU'EST-CE QUE LA DATA SCIENCE ?



Historique

Contrairement à ce que l'on pourrait penser, la data science n'est pas un sujet récent. **Bien qu'une conjecture de facteurs l'ait rendue très à la mode dernièrement, la data science existe en fait depuis les années 50.** Dès 1947, le mathématicien anglais Alan Turing a en effet donné une conférence sur l'intelligence artificielle. Ce n'est ensuite

qu'à partir de 2007 que l'engouement pour la data science a connu une forte croissance, principalement en raison de la conjonction de trois forces : le stockage, le traitement et la communication. L'évolution des dispositifs techniques ont en effet permis à ces secteurs de se développer, ce qui a engendré un intérêt exponentiel pour la data science.



Exemple :

Je souhaite estimer le prix d'une maison. En fonction de la période, différentes méthodes de data science sont utilisées.

Intelligence artificielle

Je crée un programme qui calcule le prix de la maison à partir de la superficie et du prix par m². Par exemple, $\text{prix} = \text{superficie} \times 3000$

Machine learning

Je crée un programme qui compare la maison que je dois estimer à une base de données qui référence les caractéristiques de plein d'autres maisons : superficie, localisation, nombre de pièces, ...

Deep learning

Je n'ai plus besoin de fournir de données structurées à mon programme, un réseau neuronal analyse toutes les données à disposition pour calculer le prix de la maison.

Définition

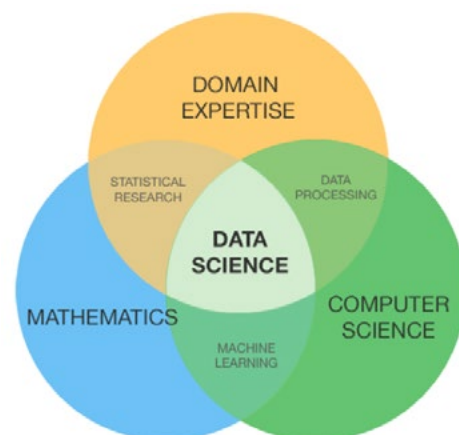
Le champ de la data science est vaste et recoupe des termes souvent mal employés, car méconnus. Alors, qu'entend-on par « data science » ? Qu'est-ce que l'intelligence artificielle ? Le machine learning ? Et le deep learning ?

Réponses...

Data science

La data science est un vaste domaine qui englobe toutes les applications et méthodes de la science des données.

Gestion des données, explication des modèles, prévisions des comportements, ... la data science regroupe tous les sous-domaines de l'étude des données, en s'appuyant sur de nombreuses méthodes telles que la statistique, l'analyse de clusters, les techniques d'apprentissage, etc.



Palmer, Shelly. Data Science for the C-Suite. New York: Digital Living Press, 2015. Print.

Intelligence artificielle

L'intelligence artificielle peut être définie comme l'imitation de l'intelligence humaine à l'aide de procédés informatiques. Basée sur la capacité des ordinateurs à comprendre les données, l'IA apporte une réponse à un problème au même titre que le ferait un expert, en prenant des décisions à partir des différents modèles de données. Plus les données sont nombreuses, plus l'IA est précise car elle a la capacité d'apprendre en créant des modèles d'entraînement sur les données.

Machine learning ou apprentissage automatique

Le Machine learning est un domaine de l'intelligence artificielle qui repose sur les statistiques. Il est utilisé dans des situations où l'on dispose d'une grande quantité de données. La machine commence par apprendre, autrement dit par extraire de la connaissance des données, puis elle applique ces connaissances sur de nouveaux éléments. Le Machine learning permet alors de découvrir des patterns afin d'effectuer des prédictions. L'ordinateur se sert en effet des données qu'il a analysées pour **comprendre la nature de données inconnues**. Différents algorithmes de Machine learning peuvent alors être utilisés pour résoudre des problèmes de prédiction, de classification, de régression, etc.

Deep learning

Le Deep learning est une application du Machine learning utilisé avec des réseaux de neurones. Même si la puissance de calcul du Machine learning est robuste, elle n'est pas suffisante pour certaines applications. C'est alors qu'intervient l'apprentissage profond. Les réseaux neuronaux utilisés par le Deep learning permettent en effet de soumettre chaque élément des données à une série de questions de type vrai/faux. En prenant en compte toutes les caractéristiques des données, le Deep learning permet de les classer avec une grande précision.

Toutes ces méthodes de la data science peuvent être utilisées à des fins différentes. Parmi elles : l'optimisation des chaînes d'approvisionnement. Aussi, la question de la planification se pose au sein de toutes les chaînes d'approvisionnement. In fine, l'objectif est que le produit qui sera vendu se trouve au bon endroit, au bon moment. Afin de s'en assurer, il est primordial pour les entreprises d'évaluer au mieux le futur,

et de réduire au maximum les facteurs d'incertitude. **Grâce à l'augmentation de la puissance de calcul, à l'afflux de données et à la disponibilité d'outils, il est aujourd'hui possible de faire la différence en matière d'anticipation.** Les algorithmes sont en effet plus puissants et plus facilement implémentables qu'il y a une dizaine d'années. Dès lors, pourquoi ne pas utiliser la data science au service de votre Supply Chain ?

Data Scientist, le métier de demain ?

Dans une interview pour McKinsey Quarterly, Hal Varian, économiste en chef chez Google, déclare en 2009 que les statisticiens auront le « job le plus sexy » des 10 prochaines années. Il considère en effet que l'extraction d'informations pertinentes à partir de données brutes constituera l'un des métiers les plus importants des prochaines années.

Si l'on considère aujourd'hui le Data Analyst comme un statisticien agile, expert en bases de données et IT, le métier se matérialise souvent par une équipe dont les membres disposent chacun de compétences en matière de data science, plutôt que par une personne maîtrisant l'ensemble de ces compétences. Et demain ? Face à l'évolution de la data science, il y a de grandes chances pour que le métier se spécialise. On peut en effet imaginer la généralisation de Chief Data Scientist Officers, qui chapoteront les projets de data science et orchestreront le travail de Data Designers, Data Builders, Data Meaning Managers, ... Le développement de la data science laisse en effet penser que les métiers se qualifieront pour occuper une place plus importante dans les entreprises.

02

LA DATA SCIENCE POUR RÉPONDRE AUX ENJEUX ACTUELS DE LA SUPPLY CHAIN



La Supply Chain est un secteur en cours de transformation, notamment grâce à l'utilisation de la data science et l'exploitation de ses données. La data science se révèle en effet particulièrement adaptée pour anticiper les ventes et optimiser les processus, en offrant à l'ensemble des équipes une certaine maîtrise face à la variabilité de leur activité. Ainsi, la data science répond aux différents enjeux de la Supply Chain :

- **Exploiter les données** : les entreprises disposent souvent d'une grande quantité de données souvent inexploitées. Historiques de vente, codes clients, saisonnalité, ... L'extraction des informations présentes dans les données constitue un défi important pour les entreprises qui cherchent à mieux comprendre les mécanismes de vente.
- **Être capable d'anticiper sans subir les aléas de la Supply Chain** : ruptures de stock, délais de production, délais d'approvisionnement, ... La data science permet d'anticiper les risques afin de s'en prémunir efficacement.
- **Faire des groupes de produits pour prévoir de manière plus large** : les gammes de produits toujours plus larges et à courte durée de vie nécessitent un travail de catégorisation afin d'adapter sa stratégie Supply Chain aux caractéristiques produits. Grâce à des techniques de clustering, la data science est d'une aide précieuse en la matière.
- **Améliorer la performance opérationnelle** : productivité, qualité, stocks, fin de vie, ... En optimisant les opérations, la data science permet à l'entreprise de gagner en performance.
- **Réduire les stocks** : le surstock est synonyme d'immobilisation de cash. Grâce à l'analyse des données, la data science est un moyen efficace de trouver le juste équilibre.
- **Réduire les coûts** : dégradation de la performance, désorganisation, litiges, ... Une mauvaise gestion de la Supply Chain implique nécessairement des coûts supplémentaires. La data science aide alors l'entreprise à prendre des décisions stratégiques, en lui fournissant une vision d'ensemble des coûts engendrés.
- **Améliorer le taux de service et la satisfaction client** : ces indicateurs sont inhérents au pilotage de la Supply Chain de l'entreprise, que la data science améliore.
- **Rassembler les équipes autour de l'information** : différents métiers entrent en jeu au moment de mettre en place des stratégies de vente. Les équipes Supply Chain, marketing, commerce ou encore finance doivent donc collaborer. A ce titre, la data science rassemble les différents métiers autour du partage de l'information.
- **Prendre les bonnes décisions** : in fine, la data science est un support utile à la prise de décisions. Elle permet à l'entreprise d'avoir une vision à 360° de ses activités passées, présentes et futures, ce qui l'accompagne dans la prise de décisions stratégiques.
- **Communiquer sur les chiffres pour justifier les prises de décisions** : grâce à l'exploitation des données, l'entreprise est à même d'argumenter chacune de ses décisions. La data science permet ainsi de convaincre en s'appuyant sur des chiffres, plutôt que de chercher à persuader en se basant sur des conjectures.

La continuité et l'efficacité de la chaîne d'approvisionnement nécessite une réponse rapide et précise à chacun de ces enjeux. A ce titre, l'étendue du potentiel de la data science peut être un allié de taille pour la Supply Chain. Il est bien sûr possible de fonctionner sans l'aide de la data science. Mais au vu de sa valeur ajoutée, pourquoi s'en passer ?

03

CAS D'APPLICATION DE LA DATA SCIENCE AU SERVICE DE LA SUPPLY CHAIN : L'ANALYSE COMPORTEMENTALE



De nombreuses applications de la data science peuvent intervenir au sein de la Supply Chain. L'analyse comportementale en est une. Appliquée à l'industrie, **l'analyse comportementale est une méthode qui consiste à analyser la façon dont se comportent les différents produits vendus** par l'entreprise : nombre de ventes, saisonnalité, cycle de vie, ... Cette analyse permet ensuite de prédire le comportement qu'adoptera un produit, en fonction de la similarité de ses caractéristiques avec celles des produits déjà connus. L'analyse des comportements se décline à travers différents procédés, et couvre différents objectifs :

Comment ?

- **Clustering** client/produit/classe
- Assortiments, analyse de courbes
- Analyse de la rupture de tendances et des changements de cycle de vie
- Analyse des tops et flops
- Identification de la saisonnalité
- Analyse des facteurs de vente

Pour qui ?

- Si l'analyse comportementale est adaptée aux entreprises de toutes tailles, les procédés mis en place par Colibri s'appliquent tout particulièrement aux PME/ETI dans l'industrie et dont les produits sont distribués en France.

Pour quoi ?

- **Mieux prévoir**
- Faire parler les données en vue d'améliorer ses process
- Anticiper un futur top ou flop

Le clustering, une méthode d'identification de tendances en fonction des caractéristiques produits

Clustering : définition et enjeux

Le clustering est une discipline particulière du Machine Learning et de l'analyse comportementale. Il a pour objectif de **séparer les données en groupes homogènes ayant des caractéristiques communes**. Chaque produit est analysé selon différents axes : profil, attributs, ventes, saisonnalité, ... Les produits qui partagent les mêmes caractéristiques et ont un comportement sensiblement identique sont ensuite regroupés dans des catégories, aussi appelées clusters. Cette connaissance des similarités permet alors de **prédire le comportement d'un produit selon ses points communs avec d'autres produits**.

De fait, un produit passe par plusieurs phases lors de son cycle de vie, comme en témoignent les historiques de ventes. Distinguer la phase de lancement de la phase active ou de la phase de fin de vie donne alors la possibilité de réaliser une analyse fine du comportement du produit. Le clustering permet ainsi de tirer partie des similarités entre produits, afin d'obtenir des planifications précises, notamment concernant l'introduction de nouveaux produits sur le marché. **Connaître les caractéristiques qui déterminent l'appartenance du produit à un cluster facilite en effet la planification des ventes et prépositionne ainsi les prévisions.**

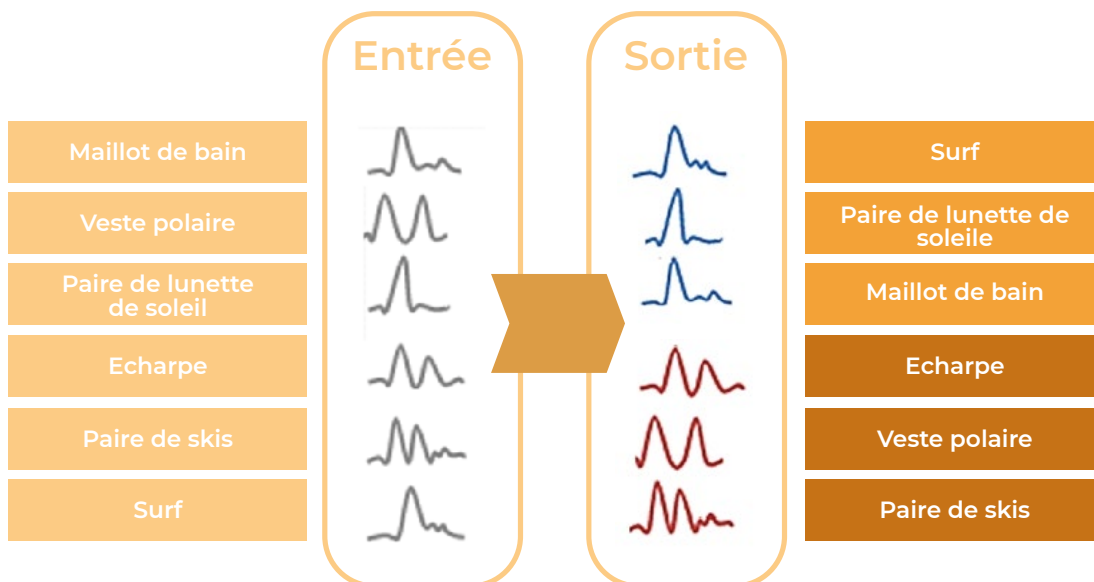
Avantages du clustering

- **Supply Chain** : facilitation de la prévision et accélération du processus grâce aux hypothèses réalisées sur les groupes de données.
- **New Product Introduction** : le nouveau produit est assigné à un cluster en fonction de ses caractéristiques, ce qui permet de faire une supposition sur le profil de lancement du produit et donc de mieux gérer les stocks de lancement.
- **Marketing** : suivant la caractéristique du produit, la partition de données permet de faire des suppositions sur les réactions du produit aux événements promotionnels.

Application : le clustering, une approche pertinente pour anticiper les phénomènes de saisonnalité

De nombreuses entreprises sont soumises à des **phénomènes de saisonnalité qui impactent la vente** des produits qu'elles proposent. Le clustering est alors une méthode efficace pour saisir toutes les spécificités de ces saisonnalités, en vue d'anticiper au mieux leurs conséquences sur les ventes et donc, d'adapter la production en fonction.

Prenons un exemple très simple : une entreprise fabrique et vend des équipements de sport. Parmi eux, 6 produits : un maillot de bain, une veste polaire, une paire de lunettes de soleil, une écharpe, une paire de skis et un surf. L'entreprise dispose de toutes les courbes de vente de chacun des produits. **L'algorithme analyse alors la forme des courbes, puis les produits qui ont un profil identique sont regroupés dans différents clusters :**



Deux clusters apparaissent ici : un cluster est formé avec les équipements sportifs d'été, tandis qu'un second regroupe les produits d'hiver. La saison semble déterminer ici les différents clusters. **L'entreprise peut alors se pencher sur les caractéristiques de chaque produit afin de dégager des tendances pertinentes pour son activité.**

Le clustering peut bien sûr être utilisé à des niveaux de détails beaucoup plus élevés, ce qui permet de réaliser des clusters très précis. Aussi, **la Supply Chain peut s'appuyer sur ces informations pour gagner en justesse et en efficacité dans sa planification.**

04

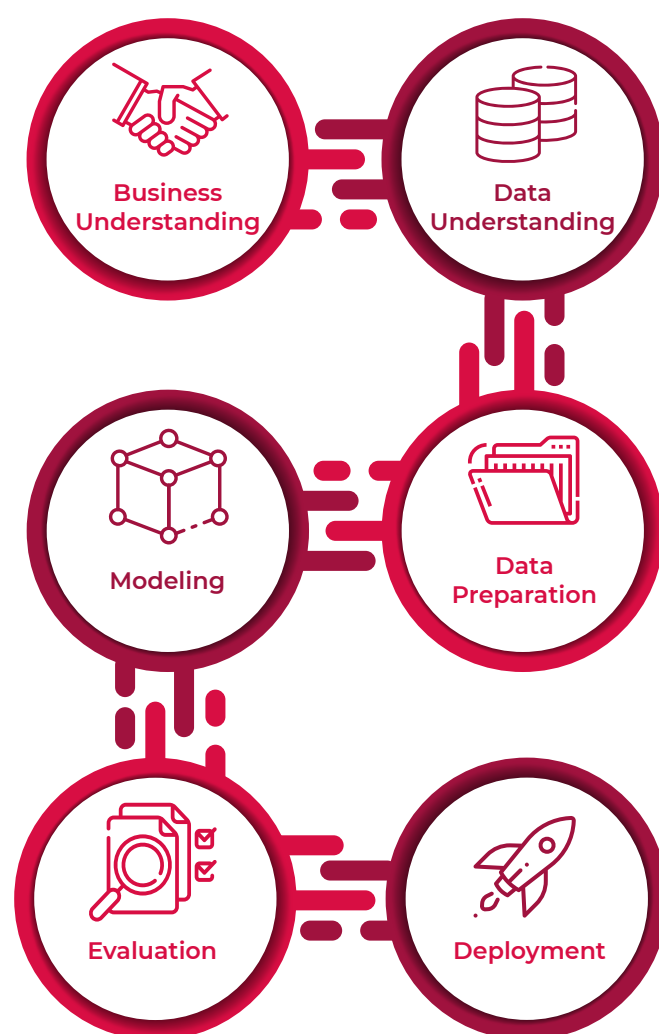
**COLIBRI, DE LA FACULTÉ
D'INTÉGRER LA DATA
SCIENCE À SES PACKS
DE SERVICES**



Depuis son lancement en 2014, Colibri a à cœur de faire évoluer son outil constamment, afin de résoudre les problématiques rencontrées par les entreprises.

Afin d'offrir à nos utilisateurs des outils d'optimisation de la Supply Chain, nous misons avant tout sur la simplicité et la lisibilité.

Notre objectif est de faire en sorte que toutes les informations dont disposent les entreprises ne forment pas une boîte noire opaque mais au contraire, **que les données deviennent un levier de performance.** Aussi, nos experts accompagnent les entreprises dans l'exploitation de leurs données, de la récupération à l'analyse, en passant par l'ordonnancement. Grâce à sa connaissance des méthodes de data science, **Colibri combine les technologies en vue d'améliorer les process de la Supply Chain.**



L'offre data science de Colibri comporte 4 grandes fonctions :

• Analyse des comportements :

- Rupture de tendances, changement cycle de vie
- Clustering client/produit/classe
- Assortiments, analyse de courbes
- Analyses des tops et flops
- Identification de la saisonnalité
- Identification de facteurs importants de ventes

• Analyse de l'impact de certaines données :

- Analyse des données tierces susceptibles d'impacter le business
- Analyse de l'impact des fêtes calendaires

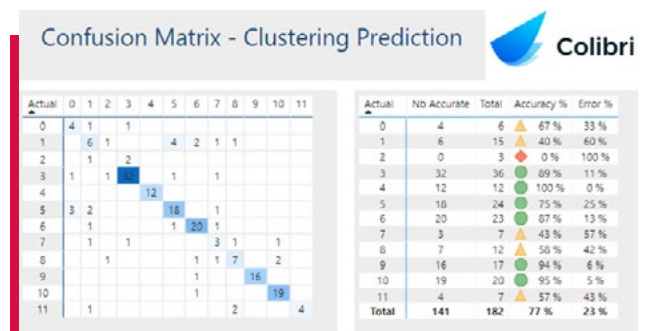
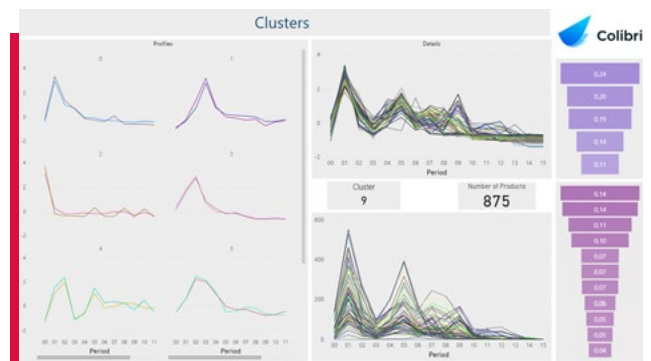
• Suivre et réagir aux lancements / fin de vie

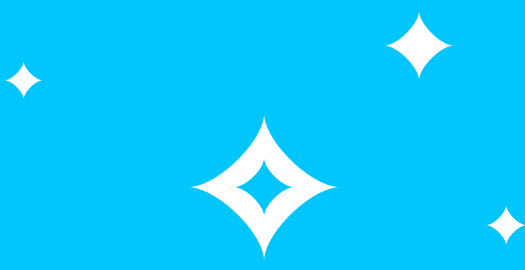
- Identification des critères déterminants lors d'un lancement produit

- Analyse de ce que dit une phase de lancement sur le cycle de vie du produit
- Définition de la famille/typologie à laquelle associer un produit

• Analyse des prix

- Analyse de l'élasticité des prix
- Analyse de l'élasticité des ventes ou promotions
- Analyse des poids financiers





Solution de Demand et Supply Planning nouvelle génération,

Colibri vous permet de gagner en efficacité, en précision et en compétitivité. De l'industrie à l'agroalimentaire, en passant par la distribution et le retail, Colibri s'adapte parfaitement à chaque métier. Du prototypage du projet jusqu'à son utilisation finale, vous bénéficiez de l'accompagnement de nos experts, afin de vous garantir une application optimale. Rejoignez l'écosystème Colibri, porté par des clients moteurs et une croissance forte !

[Demandez une démo](#)

[Contactez-nous](#)



Colibri