

LES MÉTIERS DE LA CONSO

# DATA SCIENTIST

et autres fonctions de la data



LES MÉTIERS DE LA CONSO  
**DATA SCIENTIST** & CO

**Julien Guitard, 40 ans**  
Chief Data and Science Officer  
chez Lucky cart



Polytechnicien issu de l'ENSAE, Julien Guitard a ensuite conduit une thèse sur l'évaluation scientifique des politiques publiques tout en travaillant à l'Insee. Puis, il rejoint la direction du Trésor à Bercy en tant qu'économiste et conseiller ministériel. L'année 2012 marque un virage dans sa carrière puisqu'il passe dans le privé, intégrant le Boston Consulting Group en tant que consultant. Après deux années au BCG, il met à profit ses compétences scientifiques en devenant conseiller spécialisé data chez Bluestone (EY). En 2017, il rejoint Lucky cart en tant qu'associé et directeur scientifique.

## AVANT TOUTE CHOSE, LA DATA C'EST QUOI ?

J'ai l'habitude de dire que la data remonte aux Sumériens pour bien signifier qu'elle a toujours existé... ou presque. En matière de commerce, il y a toujours eu la volonté de mieux connaître ses clients ou ses ventes. Avec le commerce moderne, cela a pris la forme des cartes de fidélité, des logiciels de CRM et des panels basés sur l'extraction des données de caisses. Pendant longtemps, on ne parlait pas de "data" mais de données ou de statistiques. Le terme s'est imposé avec l'accélération du phénomène depuis une dizaine d'années. Les évolutions technologiques et plus généralement les évolutions scientifiques (algorithmes, intelligence artificielle, machine learning, etc.) ont permis tout à la fois de manipuler un plus grand volume de données et d'ouvrir grand le champ des applications possibles. Il y a 20 ans, lorsque j'étais à l'Insee, nous manipulions des fichiers de 20 000 lignes. Aujourd'hui nous sommes capables de traiter un, voire plusieurs milliards d'enregistrements, ce que nous faisons chez Lucky cart sur la base des seuls tickets de caisse des enseignes.

## POURQUOI NE TRAITER "QUE" LES DONNÉES DE CAISSES ?

Pour la personnalisation des activations promotionnelles, la donnée de transaction est la meilleure data qu'il soit. Nul besoin de faire des croisements avec des sources socio-démographiques par exemple. Auparavant, on parlait en cohortes, panels, segments, aujourd'hui nous travaillons au niveau de l'individu. Un individu anonymisé et ramené à son historique de consommation, la somme de ses tickets de caisse.

## CONCRÈTEMENT, CE QUE CETTE DATA PERMET DE FAIRE ?

Elle permet de mieux comprendre et d'anticiper chaque consommateur pour lui proposer une opération promotionnelle lui correspondant au mieux à un instant T. Typiquement, à partir de 5 tickets d'historique, l'algorithme peut produire des scénarii avec probabilités pour une opération promotionnelle.



Derrière une activation promotionnelle comme on peut voir ci-dessus, il y a l'API ou interface de programmation d'application de Lucky cart qui vient se plugger sur la plateforme de l'enseigne partenaire afin de faire tourner "l'usine algorithmique" avec la base de données de l'enseigne alimentée en temps réel par le flux des caisses.

Pour donner un exemple concret, dans le cadre d'une opération de remboursement à 100% du panier d'achat, la data permet de personnaliser les bannières vues sur le Drive en adaptant par exemple le seuil de produit à acheter pour participer à l'opération et la probabilité de gain en quasi-temps réel. De façon très simplifiée, prenons un client A qui est acheteur de la marque activée. De fait, il va avoir un seuil au-dessus de ses habitudes avec une probabilité attractive pour lui. Si par contre, il a fait de récents achats sur cette marque, le seuil va être adapté. Quant au client B qui n'est pas acheteur de la catégorie, il aura un seuil bas et une probabilité attractive pour l'inciter à essayer. Le profil créé à partir de la data inclut l'intégralité des catégories que le consommateur achète, sa façon d'acheter (quantité par achat, fréquence et métriques dérivées du machine learning et de l'IA), ainsi que la manière dont il a réagi aux précédentes activations Lucky cart. Cette personnalisation s'adapte tout au long de la campagne. Cela permet de gérer beaucoup plus efficacement l'effet d'aubaine et de stockage et de proposer des offres plus pertinentes pour chaque consommateur.

## LES DIFFÉRENTS MÉTIERS DE LA DATA ?

Je dirige l'équipe en charge de concevoir, développer, opérer et entretenir "l'usine algorithmique" de notre solution. Lorsque j'ai rejoint Lucky cart, j'ai construit cette équipe data autour de trois fonctions :

- **les data scientists** : ils conçoivent les algorithmes (exemple : comment faire pour prédire ce que le shopper va acheter ?).
- **les data ingénieurs** : leur rôle est de faire tourner "l'usine algorithmique" avec les datas volumineuses des enseignes en s'appuyant sur les technologies cloud et internes.
- **les data analysts** : ils font la passerelle entre les sujets purement data et les sujets business en lien avec les équipes commerciales. Ils font parler la donnée pour la rendre intelligible au travers d'indicateurs et de rapports facilitant la prise de décision.

A présent, l'équipe se compose de 5 data scientists (dont moi), 3 data analysts et 4 data ingénieurs.

## DU DIRECTEUR DATA & SCIENCE AU PATRON...



Julien Guitard

... il n'y a qu'un pas, Julien étant associé et membre du comité de direction de Lucky cart depuis son arrivée. Preuve s'il en est du rôle capital de la data au sein de l'entreprise.



Romain Charles  
CEO

## AU QUOTIDIEN, VOTRE TRAVAIL ?

C'est un mix de tâches de fond et de management. Je continue à coder moi-même, car il me paraît important de montrer l'exemple et d'être en prise directe avec notre "usine algorithmique". Sur le plan managérial, j'organise le travail de l'équipe avec des méthodes plus ou moins agiles. J'ai également un rôle de coach qui me tient à cœur. Tant sur les questions liées directement au métier que sur les soft-skills (NDLR : *communication, gestion du temps, créativité...*).

## LE PLUS DIFFICILE À GÉRER ?

D'être toujours en tension entre les développements court terme et long terme au service de notre "usine". Le plus dur est de trouver le bon équilibre. Si on se concentre trop sur le court terme, on passe à côté d'évolutions pertinentes pour notre plateforme et si on est trop focus sur le long terme, nous ne sommes pas assez opérationnels et réactifs, ce que ne manquerait pas de nous rappeler l'équipe commerciale. Aussi, je dois constamment arbitrer entre ces deux échelles de temps.

## SI VOUS AVIEZ UNE BAGUETTE MAGIQUE...

Déjà, je ne l'utiliserai pas sur les personnes, mais sur les machines. On pourrait ainsi rêver de la suppression des limites technologiques et de l'avènement de "l'algorithmique café" comme j'ai l'habitude de dire à mes équipes : l'algorithmique n'aurait plus besoin de nous et on pourrait passer notre temps à boire des cafés. Si notre "usine algorithmique" apprend sans cesse au gré de nos développements et des campagnes d'activation menées tout au long de l'année, sa totale autonomie n'est pas pour tout de suite. Il faut des garde-fous humains. Et, puis, nous avons encore beaucoup d'optimisations et de recherches à faire... pour au moins 3 ou 4 ans... avant de passer au café !!!

## **VOUS MANIPULEZ DES BASES GIGANTESQUES AVEC CHAQUE JOUR DES MILLIONS DE NOUVELLES INFORMATIONS. UN JOUR SUBMERGÉ PAR LA DATA**

J'ai appris à être humble avec les chiffres. Si crainte il doit y avoir, c'est plus d'être submergé par les questions que nous pouvons être amené à nous poser. À trop vouloir chercher des corrélations (par exemple, les consommateurs d'une compote donnée achètent tel ou tel type d'apéritif), on risque de ne plus être opérationnel et de passer à côté de véritables gisements de valeur. Il faut cadrer tout cela.