Thomas Choussat

Pierre-Louis Clavel

Myriam OUNISSI

Amina Kahlane



L2 MIASHS

TD 3  
Groupe 21

Années : 2021-2022

Mémoire Préprofessionnalisation

**DATA SCIENTISTE**

Table des matières

INTRODUCTION……………………………………………………………………………. 3

[I. PRESENTATION : Le métier choisi](#_Toc92036779)

[II. A PRIORI : présenter ce métier à premier vu](#_Toc92036780)

[III. QUESTION](#_Toc92036781)

CORPS DU MÉMOIRE……………………………………………………………………….5

[I. RECHERCHE](#_Toc92036782)

[II. ANALYSE](#_Toc92036783)

CONCLUSION……………………………………………………………………………….11

ANNEXES…………..………………………………………………………………………..12

[I. BIBLIOGRAPHIE ; REFERENCES DOCUMENTAIRES](#_Toc92036784)

[II. INTERVIEWS](#_Toc92036785)

Pré- Professionnalisation

**PROJET : Data Science**

INTRODUCTION

# PRESENTATION : Le métier choisi

Ce mémoire est sur le métier de data scientiste. Nous avons choisi de prendre ce métier car c’est un domaine que nous ne connaissons pas ou très peu pour la plupart. Nous avions envies d’en savoir davantage sur le métier pour nous faire un avis sur celui-ci et pour savoir s’il est susceptible de nous intéresser dans le futur

Le Data Scientiste un Spécialiste des statistiques mais surtout de bien d’autres choses, nous aborderons donc toutes les questions qu’on peut se poser pour ce domaine.

# A PRIORI : présenter ce métier à premier vu

**Amina :** D’après mes connaissances une entreprise collecte certaines données de ses clients pour les traiter. C’est à ce moment donc qu’un data scientiste va intervenir pour donner un sens à ces données. Je suppose qu’il utilise souvent de la programmation mais également des statistiques.

**PL :** Pour moi, Data Scientiste c’est quelqu’un qui s’occupe de gérer les données d’une entreprise et de les traiter. J’imagine que c’est un métier qui demande beaucoup de compétences en informatique et que le métier consiste à rester sur son ordinateur toute la journée.

**Thomas :** La raison pour laquelle j’ai choisi un groupe faisant ce sujet est justement la considération de l’aspect flou de la réalité de ce métier. Pour moi un data scientiste est un professionnel qui traite et manipule des données. Je suppose que c’est un métier qui nécessite des connaissances en informatique et je suppose donc aussi en mathématique au vu de la taille des bases de données. J’imagine que c’est un métier qui se pratique au sein d’une entreprise et surtout via un ordinateur.

**Myriam :**

Avant de prendre ce sujet, je ne connaissais que de nom le métier et je ne voyais pas la différence entre les métiers dans la data (comme la différence entre data scientiste et data Analyste). Les préjugés que j’avais étaient que c’était un métier qui demandait de longues études, qui était bien payé et qui consistait à analyser et donner son expertise pour faire avancer un projet au niveau des demandes grâce à des données.

# QUESTION

Pour comprendre au mieux ce métier, nous nous sommes posés toutes les questions qu’on pouvait avoir en tête, en essayant dans un premier temps d’y répondre (réponse : écrite en bleu) sans la moindre recherche ou documentation.

**Le métier, fonction, rôle, spécificités…**

* Quel seraient les secteurs d’activités ? Dans quelles entreprises demande-t-on un data scientiste ?

Surement le secteur tertiaire principalement marchand ou dans le domaine de la recherche.

* Quelle est la fonction de ce métier ? Quel rôle occupes-t ’il dans l’entreprise ?

Un data scientiste serait une personne qui donnerait son expertise grâce aux données de l’entreprise afin de les aider sur un projet. (Cela est assez flou pour nous).

* Quelle serait une journée type ? Que fait le data-scientiste ? A quoi sert-il ?

Une grosse partie pourrais être de récupérer, de classer les données (peut-être) et de les traiter au vu du nom de l’intitulé du métier.

* Avec qui travailles-t-il ? Quel est son positionnement dans l’entreprise ? Qui dépend de lui et de qui dépendes-t-il ?

Il dépend forcément du client et de son directeur, mais nous pensons que c’est plus un domaine de collaboration, (plus que de positionnement). Nous pensons qu’il s'agit d’une fonction support au vu de la nécessité d’avoir un processus qui crée des données.

* Le salaire ?

Nous pensons que le revenu doit être quelque peu supérieur par rapport à la moyenne nationale. Pas forcément plus important que par rapport aux autres métiers nécessitant une formation importante.

**Le parcours/Trajectoire du pro, le côté personnel, les compétences acquises…**

* Comment devenir Data Scientiste ? Quel parcours Professionnel et quelles études sont à faire ?

Il faut probablement sortir d’une école d’ingénieur ou peut-être une formation équivalente.

* Trajectoire: Est-ce que ça ouvre d’autre porte sur d’autre métier (directeur de …)

Nous pensons que oui, il y a toujours la possibilité d’évoluer dans une entreprise, mais dans le cas de notre métier, il ne devrait pas y avoir forcément de suite logique après être devenu data scientiste.

* Quelle sont les compétences requises pour être un bon Data Scientiste ? Les Qualité demander ?

Communication, patience (pour comprendre la demande du client) et persévérance. Avoir aussi de bonnes connaissances en statistique et en mathématique.

* Les difficultés dans ce métier ? Points positifs et négatifs ?

Le stress pourrait être un point négatif ou aussi la redondance de faire toujours les mêmes calcules, les mêmes demandes etc.

**Le métier en lui-même**

* Temps Libre : Votre métier vous laisse-t-il beaucoup de temps libre (en plus des vacances) ? Pouvez-vous consacrer du temps à votre famille ?

Un data scientiste doit surement être lourd en termes de charge de travail.

* Le métier de Data Scientiste existe depuis combien de temps ? Est-ce que le métier/domaine continuera d’évoluer dans les prochaines années ?

Nous pensons que c’est un métier assez récent, pour certains d’entre nous c’est même un métier qui n’aurait pas plus de 30 ans. Avec l’évolution constante de la technologie, le métier pourrait demander d’autres compétences, cependant, un data scientiste sera toujours utile pour vérifier la fiabilité des chiffres.

* Ce métier est-il recherché en ce moment ? Quels sont les profils recherchés lors du recrutement d’un Data Scientiste ?

Au vu, selon nous, de ses longues études, du fait que les sociétés de divers domaines ont besoins d’un data scientiste ainsi que l’expansion récente de ce domaine, nous pensons que ce métier doit être assez demandé.

* Faut-il avoir un bon niveau en Anglais ? Est-il demandé de parler plusieurs langues ?

Pour communiquer avec des clients étrangers, nous pensons qu’il faudrait surement un bon niveau d’anglais. Et c’est toujours mieux de parler plusieurs langues si on est dans une grosse société qui demande de bouger internationalement.

CORPS DU MEMOIRE

# RECHERCHE

Dans un second temps, nous traiterons les précédentes questions seulement à l’aide de nos recherches : documentation, interview ou bibliographies.

Ainsi, de nouvelles questions pourraient s’ajouter au fil de nos recherches (les nouvelles questions seront en ***italique***) pour mieux connaître ce domaine.

**Le métier, fonction, rôle, spécificités…**

* Quel seraient les secteurs d’activités ? Dans quelles entreprises demande-t-on un data scientiste ?

Les différents secteurs d’activité sont : la finance, l’informatique ainsi que l’assurance. Ainsi, toutes les entreprises reliées à ces secteurs ont besoin de data scientiste.

Hugo Térrien, l’un de nos interviewers dit qu’il « travaille en partie pour le luxe et le tertiaire, dernièrement l’immobilier et un cabinet de conseil. ». On peut donc dire que c’est un métier qui travaille dans des secteurs accès divers.

* Quelle est la fonction de ce métier ? Quel rôle occupe-t-il dans l’entreprise ?

Spécialiste des statistiques, de l'informatiques et du marketing, le Data Scientiste recueille, traite, analyse et fait parler les données massives, autrement appelées “big data”, dans le but d'améliorer les performances d'une entreprise.

Son rôle consiste à analyser de manière efficace les données pour répondre aux problématiques de l’entreprise.

* *Nouvelle question :* Quelle est la différence entre data scientiste et data analyste ?

Le data analyste retraite et met en forme les données afin que d’autres personnes puissent les utiliser ainsi que les connaissances que transmettent ces données. C’est un autre métier. En revanche un Data scientiste va utiliser les données. La différentiation entre les métiers est assez nouvelle et est différenciée plutôt dans les grandes structures.

* Quelle serait une journée type ? Que fait-il ?

Comme mentionné par le premier professionnel interviewé, le Data scientist doit d’abord identifier le problème du client en se familiarisant avec les données. Ensuite, il traite ces données et les manie pour tenter de résoudre le problème.

“On prend les données fournies par les clients, on les nettoie, on les traite, dès fois on doit transformer les données, parce que par moment elles ne sont pas utilisables ni pour nous, ni pour les modèles.”

Pour faire “parler” ces données, le Data Scientiste utilise des logiciels et des algorithmes. Les plus connus sont Hadoop, Spark, Apache Storm, Flink et Hive.

* Avec qui travailles-t-il ? Quel est son positionnement dans l’entreprise ? Qui dépend de lui et de qui dépends-t-il ?

Pour être un bon data scientist il faut savoir travailler en équipe. En fonction des problématiques, il peut collaborer avec des experts (chef de projet digitale par exemple)

* Le salaire ?

D’après l’étude menée par le site internet Glossdoor, le salaire annuel moyen d’un Data Scientiste dépend de deux facteurs : l’expérience et la région.

A Paris, un data scientiste débutant touche en moyenne entre 45 et 55k euro par ans. Après les deux ans d’expériences, il peut atteindre 70k. Enfin un sénior avec plus de 5 ans d’expériences peut atteindre les 100K par ans.

Dans les grandes villes en général un débutant touche entre 45 et 50k. En étant un data scientiste confirmé son salaire varie entre 51 et 68k. Enfin après les 5 ans d’expériences il touchera entre 69 et 80K.

Enfin dans les régions on peut constater que les salaires sont généralement légèrement inférieurs que à ceux des grandes villes.

**Le parcours/Trajectoire du pro, le côté personnel, les compétences acquises…**

Comment devenir Data Scientiste ? Quel parcours Professionnel et quelles études sont à faire ?

Le métier de Data scientiste est généralement la suite d’études longues ou d’une reconversion professionnelle après des études dans un domaine proche. Les profils les plus courant sont des profils ayant fait un bac+5 dans les domaines de l’informatique et mathématique. Une part non-négligeable des Data scientiste ont poursuivi leurs études par un doctorat. Différentes écoles d’ingénieur proposent des parcours afin de devenir data scientiste par exemple Ensae, qui propose une formation post bac +5 ou Télécom SudParis – ENSAE ParisTech. Ces voies sont en revanche peu représentées parmi les Data-scientiste.

L’université n’est pas en reste avec des Masters comme le Master 2 Data science de Université Paris-Saclay ou le master parcours Traitement de l'information et data-science en entreprise (TIDE) de Paris à Paris 1.

D’après ce que nous avons vu avec nos interviews, la demande des reformations vers le métier de data scientiste sont possibles avec pôle emploi mais d’autres existent. Seulement, un excellent niveau mathématique est déjà nécessaire.

* *Nouvelle question* : Peut-t-on devenir un data scientiste après avoir fait une licence MIASHS ? peut t’on faire un master qui accèderait à ce métier après une licence MIASHS ?

La formation que nous effectuons actuellement, est une des licences qui sont ouvertes pour les masters sur la data science. La formation de Paris 1 Panthéon Sorbonne , TIDE dépend même de l'UFR 27. Le métier nous est donc un élément dans l’ensembles des possibles pour nous, cependant la forte demande génère une forte compétition pour les formations.

* Trajectoire: Est-ce que cela ouvre d’autres portes sur d’autres métiers ?

Une des problématiques du métier inhérente à la nouveauté du domaine dans lequel il évolue, c’est la spécialisation du métier. Le métier de DATA Scientiste est en train de disparaître au profit de ses spécialisations qui sont en train d’émerger. Un data-scientiste professionnel va pouvoir se spécialiser dans plusieurs domaines.

La suite logique avec d'expérience est de devenir « senior » Data scientiste, cependant des évolutions de carrière possible peuvent naturellement être, pour certains, de devenir chef de projet ou devenir lead Data scientiste et manager une équipe. Mais ces évolutions de carrière ne sont pas toujours possibles. Car elle demande gérer une équipe.

* Quelles sont les compétences requises pour être un bon Data Scientiste ? Les qualités demandées ?

Le point le plus récurent et le plus important est la nécessité de posséder un très bon niveau en mathématique, il est souvent rappelé lors nos interviews et de nos recherches. Les statistiques en probabilité, au même titre qu’une maîtrise de l’algèbre linéaire certes mineur, sont-elles aussi nécessaires.

Une autre partie des compétences demander sont en programmation. Il faut donc savoir manier les langages couramment utilisés dans ce domaine comme : majoritairement le langage python, le langage analytique R ou SAS, ainsi qu’une connaissance du SQL pour interagir avec les bases de données.

Evidemment, de très forte connaissance dans la science des données, à proprement parlé, sont primordiaux.

Certaines qualités peuvent être appréciées, notamment la rigueur nécessaire par l’aspect scientifique du métier. D’autre part, une certaine maîtrise de l’anglais est nécessaire afin de pouvoir aisément communiquer et comprendre. Afin de créer des modèles efficaces, il est nécessaire d’avoir une bonne compréhension afin de pouvoir comprendre ce qui est modélisé. D’autant plus que la communication, comme beaucoup d’autres métiers, est primordiale pour pouvoir travailler en équipe ou avec un client.

* Les difficultés dans ce métier ? Points négatifs ?

Il y a un ensemble de difficultés liées au métier :

D’abord, il s’agit d’un métier complexe dont l’accès peut être difficile, tant par la limite de l’offre de formation de qualité, que par sa complexité pratique.

De plus la constante évolution des technologies nécessite un apprentissage constant. Bien que cela ait des cotés gratifiants et enrichissants, cela est aussi très chronophage et nécessite une passion pour le domina à toute épreuve, mais surtout ne prend jamais fin.

*« Mon apprentissage n’a d’autres fruits que me faire sentir combien il me reste à apprendre »,* Montaigne

Un point négatif qui revient souvent est dans les petite structures et moyenne la nécessité de faire une partie du travail d’un Data miner afin de pouvoir correctement travailler. La nécessiter de traité les données, comme la préciser notre interviewé

*La relative nouveauté du métier pose des problèmes Il y a un flou sur ce qui est désigné comme data scientiste par le grand publique qui a tendance à additionner les casquettes sous une même signification.*

*Cela peut aller du Data Miner, en passant par le data scientiste, jusqu’à ingénieur en machine Learning qui sont des métiers demandant des expertises différentes. Qui peux mener a des problèmes qui peuvent demander bien plus que l’expertise du professionnel ainsi que surestimer les possibilités et la puissance des solutions, ainsi que leur fondement et origine.*

*Mais aussi dans des cas opposée les clients peuvent surestimer leur volume de connaissance sur le sujet Comme le dit notre interviewé “il faut de la diplomatie pour communiquer avec des clients qui ne comprennent pas forcement ce qu’on fait.”*

Dans les aspects positifs on peut noter l’aspect créateur quasiment Prothométhéen ,

Dans l’acte de création qui apporte de l’importance et ce qu’on fait et une certaine félicité quand le travail est terminé. Par exemple

Selon l’un des interviewers le point positif est la « satisfaction » d’avoir trouvé une méthode ou une solution au problème.

**Le métier en lui-même**

* Temps Libre : le métier est-il chronophage ? [statut de cadre]

Le métier de data scientist est un job à temps plein qui bénéficie de congés payés comme tous les autres métiers de cadre. La première personne interviewée nous a indiqué qu’elle possédait entre 35 et 40h de travail, ce qui lui laisse du temps libre pour faire autre chose.

* Le métier de Data Scientiste existe depuis combien de temps ? Est-ce que le métier/domaine continuera d’évoluer dans les prochaines années ?

Le métier de data scientiste, en tant que tel, est en train de disparaitre, du a une spécialisation qui s'opère. Le data scientiste a tendance à se spécialiser dans différents domaines.

* Ce métier est-il recherché en ce moment ? Quels sont les profils recherchés lors du recrutement d’un Data Scientiste ?

De par sa spécialisation, le métier de Data Scientiste est très recherché. Beaucoup de grandes entreprises recherchent activement ce type de métier, comme Google, Linkedin ou encore des starts ups. Les profils recherchés sont des personnes, ayant déjà travaillé dans le domaine du big data, à l’aise en algorithmie et en science des données.

* Faut-il avoir un bon niveau en Anglais ? Est-il demandé de parler plusieurs langues

Le Data scientiste est, pour la plupart du temps, en contact avec des clients. Le métier de data scient étant présent à l’international, l’acquisition de l’Anglais est donc requise pour faciliter la communication avec un client étranger, comme d’autres métiers. Une autre langue peut être utile mais n’est pas nécessaire. La première personne interviewée nous a confié qu’il pouvait pratiquer son métier sans une bonne maitrise de l’Anglais, mais que cela le ralentissait considérablement dans son métier.

# ANALYSE

*Généralisation des recherches : avec tous ces informations comment décrirais-t-on ce métier. Qu’est-ce que nous a apporté les interviews de différents de nos recherches documentaires.*

conclusion

*Comparaison des résultats obtenus avec les premières impression, explicitées dans l’introduction. Avis sur le métier.*

Pour conclure, nous pouvons dire que nos premières impressions étaient assez similaires à nos recherches. C’est un métier bien payé qui demande de longues études, des connaissances en statistiques, en mathématiques ou encore en programmation. Mais ce dont on cherchait le plus à comprendre et ce qui nous a pousser à choisir ce domaine c’était de comprendre quel sont réellement les fonctions et en quoi cela consiste d’être un data scientiste.

Ainsi, de nos recherches on a compris que le Data Scientiste recueille, traite, analyse et fait parler les données massives et on a aussi pu comprendre comment il le faisait.

**Avis sur ce métier :**

**Amina :** La data science est donc l’analyse et l’interprétation de données numérique complexe. Personnellement, grâce à cette étude j’ai découvert un métier qui englobait toutes les matières qui m’intéressait le plus et m’a donc donné l’envie d’approfondir encore plus la recherche pour plus tard

**PL :** Je pense qu’il faut avoir un goût pour le traitement des données et apprécier la programmation pour se lancer dans ce métier/domaine. Ce n’est pas mon cas.

**Thomas :** Bien qu’au globale le métier est un certain nombre d’attrait, l’aspect théorique et mathématique génèrent chez moi un certain engouement, l’aspect créateur semble en plus un gros plus.

Cependant une partie de l’informatique notamment le formatage des données suscite moins d’engouement.

Je pense que la piste est à approfondir.

**Myriam :** Personnellement, je n’aurai jamais pensé à prendre ce métier car je n’ai pas de grande affinité avec la statistique et les données en générale, de plus je ne connaissais pas plus que le nom du métier. Mais, en l’étudiant j’ai appris qu’une grande partie était informatique (un domaine qui m’intéresse beaucoup plus) et aussi le fait que ce ne soit pas si pénible qu’on ne l’aurait pensé.

ANNEXE

*bibliographie; références documentaires; verbatim des interviews*

# BIBLIOGRAPHIE ; REFERENCES DOCUMENTAIRES

**Sites pour les recherches :**

https://www.mba-esg.com/actus/data-analyst

<https://www.lebigdata.fr/salaire-data-scientist>

<https://www.lesechos.fr/idees-debats/leadership-management/data-scientist-cest-quoi-ce-metier-1244919>

<https://www.lebigdata.fr/13-competences-necessaires-devenir-data-scientist>

https://datascientest.com/metier-data-scientist

<https://formations.pantheonsorbonne.fr/fr/catalogue-des-formations/master-M/master-econometrie-statistiques-KBURDRPJ//master-parcours-traitement-de-l-information-et-data-science-en-entreprise-tide-formation-initiale-et-apprentissage-KBUREJV4.html>

**Ressource papier :**

Les métiers de l’intelligence artificielle / Fabrice Mateo Paris : l’Etudiant éditions, DL 2020 BU

Les métiers de la data Regache, Camille Edité par les Éditions de L'Etudiant – 2017 [ Bibliothèque municipale]

Office national d'information sur les enseignements et les professions (France)

Edité par Onisep – 2017 [bibliothèque municipale]

# INTERVIEWS

On a eu l’occasion d’interviewer deux personnes qui ont été, ou sont dans le domaine du data scientiste. On a pu voir quelques extraits dans la partie recherche afin de répondre à nos questions, voici donc l’entièreté de celle-ci.

1. ***Première Interview***

*Interview de Hugo Térrien un data scientiste depuis 2 ans questionner par Thomas C. le dimanche 21 novembre 2021. Suite à la demande de l’interviewé les questions prévues ont été envoyé en avance.*

**T- On peut démarrer l’interview s’il vous plaît ?**

**H**- Oui, je t’en prie. On peut commencer.

**T - Quel est le secteur de l’entreprise ?**

**H -** Je travaille dans une Start-up, Onogone, c’est une entreprise qui est le dans nord-est de Paris 20 personne à peu près, expansion rapide, On travaille pour l’industrie, automatisation, contrôle qualité et on travaille en partie pour le luxe et le tertiaire, dernièrement l’immobilier et un cabinet de conseil.

**T-Quelle est votre position dans l'entreprise, notamment dans le processus de production ?**

**H**- Je suis data scientiste et développeur. Sous la supervision d'un lead data scientiste, je suis donc chargé de concevoir et d'implémenter les modèles de machine

learning, puis de les intégrer dans l'architecture de nos services côté serveur.

**T-Quelle est votre journée type au travail ?**

**H-**Ma journée type c’est dur, je fais plutôt des répétitions d’étapes qui défile selon l’avancement du projet, c’est une boucle qu’on répète.

Tout d’abord on doit identifier la problématique métier, le problème en français, c’est comprendre les notions qui entoure le domaine.

**T - Une problématique métier ?**

**H –**Petite précision, nous on est sur la technique, on discute un peu avec les clients à propos de tout ce qui environne ce sujet, Ce n’est pas très claire, dis comme ça. J’ai une idée : par exemple le contrôle de qualité de l’entreprise demande une solution pour un détection automatique des défauts sortis à la sortie de la machine. Une fois la problématique identifiée. On tente de faire des mesures mathématiques cohérentes, qu’est-ce qu’on va mesurer en nombre, est qu’on réussit à tous les avoir on exploit les donnée, le client nous fournit les données en général et de vraiment créer des indicateurs de performance efficace.

**T -**Est-ce que vous pouvez ralentir ?

**H-**oui, désoler

On prend les données fournis par les clients, on les nettoie, on les traite, des fois on doit transformer les données, parce que par moment elles ne sont pas utilisables ni pour nous, ni pour les modèles.

Pour arriver à faire nos modèles, on utilise des méthodes déjà connue mais ces méthodes ne doivent pas juste être utiliser, on doit les transformer ou les combiner dès fois. Il y a besoin de plusieurs étapes quand on construit un modèle qu’on ne considère pas trop mal, comme est-ce que sur des données de testes il peut donner des résultats pas trop mal.

Si on juge le modèle comme bon avec les nouvelles données, on regarde où il se trompe et on cherche à savoir pourquoi il se trompe. Il y a pleine de raisons différentes à cela : ses qu’on a analysé le pourquoi on revient à l’analyse .

C’est très **intinératif** , cyclique, on explore, on se trompe jusqu’à tomber sur des choses, et à la fin on l’intègre.

**T – Vous faites aussi l’intégration [intégration du modèle dans un logiciel du client] ?**

**H**- L’intégration n’est pas systématique, ce n’est pas par essence au data scientiste de le faire.

Mais en générale, quand je dis intégrer ce n’est pas complet, juste pour que le modèle soit utilisé par le logiciel du client, on n'est pas full stack développeur chaqu’un a son métier et ses compétences.

**T - Vous travaillez par équipe ?**

**H**- Euh … On ne travaille pas toujours en équipe. Je travaille au sein d’une équipe mais c’est variable. Quelquefois on travaille sur des sujets différents et d’autre fois on travaille ensemble, ça dépend.

**T-Quel sont les qualités et compétence nécessaire pour effectuer ce métier ?**

**H**- Sans surprise un niveau suffisant en mathématique, pour appliquer ce que tu connais.

Il faut être capable de lire des Projets de recherche, faire de la veille sur les nouveaux papiers [de recherche]même si tout le monde ne le fait pas tout le temps.

Il faut aussi un bon niveau en programmation, si tu ne sais pas coder tu ne sais rien faire, au moins l’algorithmique, le reste c’est moins centrale.

Euh, de la curiosité il faut s’intéresser suffisamment au problème du client.

Il faut de la rigueur, la boucle itérative dont je te parler tout à l’heure il faut la suivre proprement, Il faut savoir dans l’ensemble même si ça ne marche pas il faut comprendre et analyser puis recommencer la boucle. Tant que tu comprends tu ne perds pas ton temps,

**T - Est ce que la communication c’est important dans ce métier ?**

**H -**Alors … De la communication il en faut forcément pour pouvoir partager avec les collèges et avec le client, mais le niveau de technicité doit être adapté selon la personne. Avec les collèges généralement ils savent ce que tu fais donc on doit plus détailler contrairement au client où là c’est plus dur. Il faut être diplomate pour expliquer que c’est de l’exploration de la boucle intinératif, n’est pas linéaire, même si tu as fait un truc qui ne marche pas tu avances quand même.

Le plus dur et de tenter d’expliquer au client qui pense savoir qu’on a tort, sans trop rentré sur le niveau technique qui les perdrait.

**T- Quelles études supérieures avez-vous effectuées ?**

**H**- J’ai fait un Master en mathématique appliqué à Université Catholique [de l'Ouest] à Angers, J'y ai étudié essentiellement les statistiques, la programmation et l'optimisation combinatoire. Après mes études, je ne savais pas quoi faire. Il n’y avait pas de cursus popularisé à l’époque. Donc j'ai travaillé quelques années dans la restauration. C'est à cette période que j'ai appris à coder en python. Courant 2018, parce que le domaine était en pleine essor, j’ai entendu parlez de la data science donc je me suis renseigné.

Et vu que ça regroupait les domaines que j’aimais, en faisant des recherches je suis tombé sur Kagles qui m'a permis de me confronter à des cas d'écoles...

**H**- Tu sais ce que sais Kagles ?

**T**- Une sorte de banque de défis autour de la Data science non ?

**H**-Plus ou moins, j’ai essayé les défis les plus simple au début, j’ai apprécié, du coup j’ai essayé de voir si je pouvais travailler dans ce domaine.

J'ai postulé à une formation en data-science en partenariat avec Pôle Emploi qui m'a permis d'être embauché là où je suis actuellement.

**T-Comment et pourquoi êtes-vous devenu Data Scientist ?**

**H-** La conception d’algorithme pour moi c’est un jeu, les mathématique et algorithmique la data science c’est assez proche c’est le même domaine. Je suis comme un gamin devant un Puzzle ou une énigme c’est marrant mais ce n’est pas rationnel. En ce qui me concerne je n’ai pas pensé à devenir scientiste.

**T-Ce métier vous laisse t ’il du temps libre en dehors de vos vacances ? êtes-vous satisfait de cela ?**

**H**- Dans l'ensemble, mes semaines de travail tournent dans les environs de 35h à 40h par semaine avec des journées qui commencent vers 9h-9h30, et qui finissent vers 18h. Donc ça me laisse le temps pour d'autres activités dans la journée. J'en suis satisfait, beaucoup plus que du rythme que j'avais quand je travaillais dans la restauration.

**T- Et vous n'avez pas des urgences parfois ?**

**H**- C’est déjà arrivé d’avoir des appels tard de clients avec qui on travaille régulièrement, mais on sait poser les limites. Après 18h tu peux répondre mais si tu n'es pas là cela n’est pas grave.

**T- Pour vous quels sont les point positifs et négatifs de votre métier ?**

**H**- Alors les points positifs comme l’histoire du puzzle, quand on résoudre une tâche et fait une synthèse c’est la meilleure partie. Ce qui est cool aussi c’est que toute la boite travaille sur le même sujet. De plus il y a la satisfaction lorsqu'on a la preuve sous les yeux que note dernière intuition était la bonne.

Pour les points négatifs plus spécifique le pré-traitement des données c’est fastidieux repercer les données mal connecter, les cases vide c’est très répétitif.

Dès fois, le fait d’avoir tester pas mal de chose et que ça ne marche pas c’est un peu découragé

Expliquer au client c’est vraiment dure aussi, on a souvent affaire à des clients qui crois savoir c’est dure d’expliquer qu’ils ont tort, en plus, le client est en position de force par rapport à nous.

**T- Comment pensez-vous que le métier/domaine continuera d’évoluer dans les prochaines années ?**

**H-** Je ne sais pas si j’ai une vision d’ensemble assez large, c’est clair qu'il n'y a pas assez de data scientiste et donc beaucoup de recrutement en ce qui concerne tout la partie maintenance dont j’ai oublié de parler de ça tout à l’heure. L’ajout de nouvelles données, nettoyage, réentrainement des modèles. Même si cela commence déjà être automatisées, donc il est probable que cette partie soit de moins en moins lourde pour nous. Je suppose que le monde de la data science à mon échelle (mais aussi de l'informatique en général) sera plus dépendant des services fournis par des géants comme Amazon, Google ou Microsoft, qui peuvent décharger les plus petites entreprises de tâches pourtant capitales par exemple OCR. [optical character recognition]

**T - Quels sont les profils recherchés lors du recrutement d’un Data Scientiste & Quelle formation faut-il réaliser ?**

**H**- Le profil rechercher disons qu'avant il fallait savoir coder et avoir des compétences en mathématiques. Quand j'étais en études, il n'y avait pas de cursus dédié spécifiquement à la data science ou du moins pas à ma connaissance. Aujourd'hui, certaines universités en proposent ou autre comme celle de pôle emploi m’as permis d’arriver là. Je crois qu’il y a des écoles d’ingé[nieurs] …. Oui, j’ai une collègue qui a fait ça. Je pense que maintenant le cursus ne peut plus faire l'impasse sur les statistiques et la programmation.

Les profils recherchés aujourd’hui, je le vois avec mes collèges qui sont soit des bacs+5, soit des doctorants. Il semble que des derniers aient un certain avantage sur les premiers dans le recrutement du fait du Crédit Impôts Recherche... [allègement fiscal pour le recrutement de neo-doctorant notamment ceux affecter au R&D]

**T-Quoi comme type de doctorat, en entreprise ?**

**H**- Je ne sais pas si mon entreprise est représentative, car l’accent est mis sur la recherche. Mon patron préfère les doctorants qui on pas travailler en entreprise, mais en générale je pense que ça fasse une différence.

**T -Faut-il avoir un bon niveau en Anglais ou une troisième langue ?**

Alors, je n'ai pas un bon niveau d'anglais pour le moment. Ça ne s'est pas révélé handicapant lorsque j'ai commencé à travailler à Onogone dans la mesure où, même si nous travaillions pour des Américains, je n'étais généralement pas à communiquer directe avec eux. Aujourd'hui, mon travail est en relation directe avec des Italiens et mon faible niveau en langue étrangère rend la communication plus lente et moins précise. Donc je dirai : oui il faut un bon niveau d'anglais, ou du moins un anglais professionnel correct. J'ajoute que dans mon cas, j'ai rarement besoin de présenter directement mes résultats aux clients mais que si c'est le cas, un Anglais propre est pratiquement indispensable. J'imagine qu'une troisième langue peut être un plus mais je pense que l'anglais est suffisant en général

**T-** Merci pour cette interview.

**H-** De rien …

1. ***Deuxième Interview***

*Dimanche 11 décembre 2021, Interview de Asma Ons C. (une ancienne Data Scientiste) par Myriam O.*

*Le métier, fonction, rôle, spécificités…*

**M- Quel métier exercez-vous aujourd’hui ? Qu’est-ce que c’est ? (En quoi ça consiste ?)**

**A**- A ce jour, je suis Chef de projet/ Consultante environnement. Je travaille sur des sujets divers et variés concernant l’accessibilité à l’eau, l’épuration des eaux, sur les ressources en eaux leur qualité et éventuellement l’asséchement des nappes phréatiques, sur les zones humides. Tout dépend du projets client.

Actuellement, je suis à mon compte depuis quelques années.

1. **Quelle est l’activité principale de cette entreprise ? Quel est son secteur ?**
2. Par ailleurs, il y a quelques années je travaillais en tant que Business manager senior et data scientiste dans des entreprises de développement informatique (Société de service du numérique SSII), j’ai travaillé chez Umanis, Ividata. Des sociétés travaillaient essentiellement dans la BI, BIG DATA, Data Science…
3. **Que faites-vous principalement de vos journées ? A quoi ressemble votre quotidien ?**
4. Mon quotidien consistait au développement du portefeuille, de l’administratif, le recrutement (car une grande part de mon ancien travail est de développer une équipe de consultant BIG DATA, Data Scientist, Architecte BIG DATA, Intégrateur, pour le compte de client tels que la société générale, EDF…, dans le but de travailler sur leur projet informatique. Aussi, lors de mon travail, je coachais mes équipes et par moment les former.
5. **Quelle était votre journée type quand vous étiez data scientiste ? Quel est pour vous le plus difficile et quelle est pour vous la plus grande différence entre une journée type d’aujourd’hui et d’avant en tant que data scientist ? Qu’est-ce que pour vous être data scientiste ? De qui vous dépend ? qui dépend de vous ? (Avec votre métier actuel et celui de data scientist) (Redondance dans les questions)**
6. Globalement, un Data Scientiste, il travaille dans une équipe constituée de Scrum master, Product owner, consultant/ développeur BIG DATA, Architecte BIG DATA, le travail de Data scientist vient en aval. En effet, après avoir intégré les données, nettoyé, gouverné les données etc… le Data scientist vient à la fin dans le but de faire parler les données en utilisant des logiciels et des algorithmes (tel que R. Le plus compliqué pour un data scientist c’est de faire parler les données et des fois, le lien entre les données ne se fait de façon évidente.
7. **Pourquoi avez-vous changé de métier ?**
8. J’ai changé de métier car j’ai perdu le sens dans ce que je faisais. Il n’y avait plus (?)

*Le parcours/Trajectoire du pro, le côté personnel, les compétences acquises…*

1. **Parlez-nous de votre parcours professionnel. Comment êtes-vous devenu Data Scientist ? Avez-vous pratiqué un ou plusieurs métiers avant de devenir Data Scientist. Si oui, le(s)quel(s)? Est-ce que votre parcours avait pour but de faire ce métier ?**
2. Je ne suis plus Data Scientiste. Mais encore aujourd’hui je travailler beaucoup et échanger avec des data scientist. À ce jour, je suis Chef de projet, consultante Environnement en Suisse. Je suis diplômée d’un master en Sciences de l’environnement et d’un MBA en conseil et entrepreneuriat. J’ai travaillé aussi en tant que Business manager junior puis Senior pendant 6ans.

Sinon, en général, un Data scientist est un métier qui est de plus en plus recherché aujourd’hui particulièrement, car il y a encore quelques années les sociétés avaient des difficultés avec leurs bases de données. Il s’agissait pour elles tout d’abord d’assurer la migration de leurs bases de données vers de nouvelles bases de données pouvant accueillir beaucoup de données (BIG DATA). Et par la suite, le data scientist entrait en jeu. Beaucoup de Data Scientist sont aussi des anciens développeurs BIG DATA....

1. **Quelles sont les compétences, que vous avez acquises pendant vos études supérieures, qui vous servent dans votre métier de data scientiste (si tu veux tu peux mettre les différences avec les compétences de ton métier actuel) ?**
2. La curiosité, la niaque

*Le métier en lui même*

1. **D’après votre expérience personnelle, quelles sont les compétences requises pour être un bon Data Scientist ?**
2. Être à l’écoute, esprit d’équipe, de la créativité et de la curiosité vis-à-vis de nouveaux logiciels étant donné que rien n’est figée et que c’est évolutif.
3. **A votre avis, le métier de Data Scientist existe depuis combien de temps ? Pensez-vous que le métier/domaine continuera d’évoluer dans les prochaines années ?**
4. C’est un métier qui existe depuis fin des années 90 sous un autre nom, lors du boom informatique. Par ailleurs, de nos jours, tous les noms de métier ont évolué et changé. C’est un métier qui continuera d’exister (je ne pense pas que nous allons arrêter d’utiliser internet, PC, Smartphone ou autre ; tous ces outils permettent de collecter les données dans le but de les traiter) ou et qui évoluera surement en fonction des technologies.
5. **Ce métier est-il recherché en ce moment ? Quels sont les profils recherchés lors du recrutement d’un Data Scientist ? Quelle formation ou parcours faut-il réaliser ?**
6. Il s’agit d’un métier demandé par ailleurs de par mon expérience dans le secteur du recrutement…, très peu de jeune diplômé en Data science sont recruter car les clients cherchent souvent les collaborateurs avec une double casquette c’est-à-dire un developpeur/ consultant BIG DATA et la casquette Data scientist. Car il faut se rendre à l’évidence, que sur le marché du travail, avoir des doubles compétences est toujours meilleurs mais aussi, de nos jours, il y a d’anciens développeur et consultant BIG DATA qui bifurque vers cette voie.
7. **Faut-il avoir un bon niveau en Anglais ? Est-il préférable de connaître une troisième langue ?**
8. Il est toujours primordial d’exceller en anglais non seulement en terme pour les tutos mais pour s’expatrier ailleurs et s’ouvrir plus d’opportunité. Aussi, de nos jours il est possible de suivre des formations et des masters en ligne avec EDX (Havard, MIT et autre) et essentiellement les cours sont en anglais.
9. **Est-ce que la théorie (ce qu’on nous apprend en cours pour atteindre ce métier) est différente de la pratique (ce que te demande ton métier) pour le métier de data scientist ?**
10. Il faut forcément de la théorie pour pouvoir faire ce métier mais comme tout métier c’est surtout la pratique qui t’en apprend le plus.
11. **S’il y en a, quelles sont pour toi, l’évolution ou les évolutions possibles (ou souhaitables) avec ce domaine ? Est-ce que ça ouvre l'autre porte sur d’autre métier (directeur de …)?**
12. Vu que je ne le suis plus je dirais qu’il est toujours possible d’évoluer et même de changer de métier. Mais le plus important est surtout de se diversifier comme je l’ai dit les clients cherchent souvent les collaborateurs avec une double casquette c’est-à-dire un développeur/ consultant BIG DATA et la casquette Data