

Presentation d'un métier de consultant

Camille Bonamy

Next Decision

2019/10/11

Plan

- Introduction
- Qu'est ce qu'un consultant ?
- Les emplois autour de la Data Science
- Quelques références fonctionnelles
- La question du coût du travail
- Bonus

Introduction

Qui ?

Camille Bonamy

- Issu de la promo 2017-2018 Master IS
- Stage de fin d'étude pour Next Decision chez un client : AS24, filiale du groupe Total sur le thème de la détection de fraudes
- En poste de Data Scientist chez Next Decision

Quoi ?

Les missions que j'ai effectué ont pour thème :

- la qualité des données
- la visualisation des données (reporting et app shiny) et des indicateurs
- la création d'indicateurs (e.g par réduction des dimensions ou test de significativité)
- la modélisation et la prévision (retard de livraison, qualité de produit, étude des commandes)

Qu'est ce qu'un consultant ?

Définition

Un consultant est un prestataire de services en conseil. Les consultants sont souvent regroupés au sein de sociétés de conseil, ou bien interviennent de façon indépendante. Il dispose d'un savoir-faire avéré dans un domaine bien précis sans pour autant être spécialiste.

[<https://fr.wikipedia.org/wiki/Consultant>]

L'Entreprise

Next Decision est un "Cabinet d'experts en Informatique Décisionnelle & Big Data, Master Data Management (MDM), et Elaboration Budgétaire".

- Les consultants sont donc amenés à se déplacer chez les clients (pour recueillir le besoin, travailler avec les métiers, faire des présentations, ...) sur une période donnée.
- Le travail à distance en agence et le télé-travail sont également utilisés.
- Un consultant expérimente ainsi le fonctionnement de différentes entreprises et travail sur des problématiques variées.

Les emplois autour de la Data Science

Bestiaire non exhaustif

Un projet peut être mené par

- un Data Scientist
- un Data Analyst
- un Data Engineer
- un Data Architect (en support)

Ceux-ci interagissent avec :

le support

- IT
 - BI (Business Intelligence)
 - "Infra" (Infrastructure)
 - Réseau
 - ...

pour le compte d'utilisateurs finaux

- Les "métiers" : marketing, commerciaux, logistiques, production, ...

Quelques références fonctionnelles R-istes et +

De manière général après 5 mn de blocage

- <https://stackoverflow.com>

Pour garder une trace, diffuser et rendre sa recherche reproductible

- apprendre : <https://happygitwithr.com/>
- utiliser : Github, Gitlab, ...

Pour manipuler les données, se connecter à des bases, sites...

- <https://www.tidyverse.org/>

ML et +

- <https://r4ds.had.co.nz/introduction.html> (pour apprendre)
- <https://www.rfordatasci.com/> (pour apprendre)
- <http://topepo.github.io/caret/index.html>
- <https://mlr.mlr-org.com/>

Des viz pour convaincre et persuader

- <http://www.ggplot2-exts.org/gallery/>
- <https://github.com/rfordatascience/tidytuesday>

La question du coût du travail

Moyenne Nantaise supposée pour un développeur consultant

- 28k€ brut annuel
- Des écarts importants selon le secteur, le poste et la taille de l'entreprise

Paris

- Coût de la vie estimé (INSEE) de 20% à 25% plus cher -->> si linéaire
33.6k€ - 35k€ brut

Take home message

- Les missions que vous effectuerez légitimeront votre status (expertise Data Science, Viz, Base de donnée et Big Data ?)
- Trouver une entreprise avec une équipe en place avec de l'expérience peut être un plus
- Une petite entreprise à d'avantage de chances de vous donner plus de responsabilités

Merci pour votre attention

Présentation créée avec xaringan : <https://github.com/yihui/xaringan>

Bonus

De l'open source français

- <http://larmarange.github.io/analyse-R/> (apprendre en français)
- <https://github.com/dreamRs/esquisse> ; <https://github.com/ThinkR-open/golem>

Comment créer un package R en ~ 2 mn

- <https://thinkr.fr/creer-package-r-quelques-minutes/>

Comment créer un module Python

- <https://stackabuse.com/creating-and-importing-modules-in-python/>

Docker pour exporter vos travaux et 'mettre en production'

- <https://colinfay.me/docker-r-reproducibility/>

RStudio et +

- <https://www.youtube.com/watch?v=bmaigtpKyiM> (tuto tidyverse)
- <https://rstudio.cloud/> et <https://rstudio.com/products/shinyapps/> (connect pour deployer son app shiny)

Un morceau de communauté R en France

- <https://github.com/frrrenchies/frrrenchies> (annuaire)
- <https://kutt.it/rgrrr> (slack)

Bonus #2

Questions d'entretien de stage :

- Dans quel package se trouve la fonction **lm** ?
- Quel est votre package préféré pour entraîner des modèles ?
- Avez vous effectué des projets ? Sont-ils sur git ?
- Avez vous un profil Kaggle ?
- Avez vous déjà participé à un challenge data ?
- Avez vous travaillé avec des données issues de problématiques réelles ?
(Comprendre données non propre)
- Quels sont les étapes pour mener un projet de data science ?
- Pouvez vous expliquer en terme simple un algorithme de machine learning (type random forest)

Suivi de test (papier...) SQL, Python et R.