

A/ sujet du test :

1/ à votre disposition : des données

Voici les open-data proposées par la ville de Paris :

<https://opendata.paris.fr/pages/home/>

2/ vous choisissez :

a/ votre thème

quel qu'il soit (commerce, culture, environnement...)

b/ votre sujet :

en dehors des analyses déjà faites

EXEMPLE: "arbres remarquables"

c/ vous choisissez votre méthode de récupération des données

idéalement par api

si vous êtes bloqué à cette étape : ok pour download

d/ vous choisissez votre SGBD et vous y installez vos données

les logiciels de bureautique (excel, google sheet, open office....) sont acceptés
vous pouvez aussi utiliser une "vraie" sgbd si vous vous sentez : mysql, elastic, bigquery...

Ou un outil de statistiques (R, matlab...)

si vous ne les maîtrisez pas => faites le choix de la sécurité, à savoir un outil de bureautique

e/ vous choisissez un angle d'étude

c'est le cas de figure classique pour un data scientist :

on doit expliquer des phénomènes, mais on ne sait pas ce que l'on cherche exactement

et le client final n'est pas toujours apte à formuler son besoin correctement

il s'agit donc d'une mise en situation réaliste

EXEMPLE pour les "arbres remarquables" :

=> chercher des corrélations (entre taille/longévité et type d'arbre, leur répartition par quartier en fonction de la densité et la pollution...)

vous êtes libre : le seul impératif est de prouver des choses à votre interlocuteur (moi en l'occurrence) et d'être dans la véracité : évidemment pas d'erreur de calcul, mais surtout : pas en essayant d'éviter les erreurs d'interprétation (vous pouvez très bien formuler des choses non 'certaines' à condition d'y apporter cette mention)

vous pouvez utiliser un logiciel de statistique (R, matlab...)

ou aussi envisager de développer de “mini-algorithmes” pour vous aider à charger, traiter et exporter vos données

le langage de programmation est totalement libre

ATTENTION :

- votre étude ne peut pas se résumer à une simple analyse descriptive des données qui se contenterait de présenter différents charts (ex : répartition des arbres par taille),
- Il faut chercher à prouver quelque chose à travers les données (ex: corrélation entre la taille des arbres et la pollution de l'air ?) => les charts sont ainsi des moyens intermédiaires étayant votre démarche intellectuelle sans être une fin en soi

f/ vous projetez vos analyses sous forme de dataviz

les dataviz de bureautique (excel, google sheet, open office....) sont acceptées
le niveau au-dessus également (par exemple data-studio <-> google sheet, ou elastic <-> kibana...)

vous pouvez vous appuyer sur un logiciel de statistique (R, matlab...)

ainsi que les outils de dataviz avancés (type metabase/Seektable/PowerBi...) si vous vous sentez

vous êtes libres :

* le seul impératif est de dénouer par la dataviz une explication complexe (ex : illustrer une corrélation contre-intuitive)

* ne seront pas acceptées des dataviz triviales - comme représenté basiquement une évolution mois/mois d'une simple variable

* le choix des charts a son importance (map, nuage, bubble, diagramme...), mais attention à ne pas en faire trop (avec des diagrammes inutilement complexes ou à pousser l'esthétisme trop loin) : il faut privilégier **l'efficacité, l'efficience, la simplicité et le pragmatisme**

B/ modalités

1/ délai :

Date d'envoi du présent mail + 5 jours

2/ rendu

à cette date m'envoyer votre projet / votre support quel que soit le support :

- * accès à un serveur, à des interfaces, à un site web....

- * partage google drive

- * code source des “mini algorithmes” le cas échéant

- * copies d'écran

- * documents (powerpoint....) avec votre raisonnement, votre contexte, votre conclusion

=> il me faut le résultat complet et fini

=> la forme m'importe peu, il n'y a aucun critère d'évaluation du test concernant le support

=> seules les données devront être soignées ainsi que la pédagogie et le contenu des explications qui devront être très claires (comme si j'étais un novice)

3/ en fonction du rendu, je programmerai un entretien pour que vous me présentiez votre étude

4/ prendre du plaisir avant tout !

Si vous ressentez ce test comme une contrainte, il n'y aura pas de mal à passer votre tour pour ce poste, c'est qu'il n'est pas fait pour vous :)

[C/ me contacter...](#)

... évidemment !

je suis à votre entière disposition

plutôt par email

je suis VOTRE support pour ce test, n'hésitez surtout pas à me solliciter