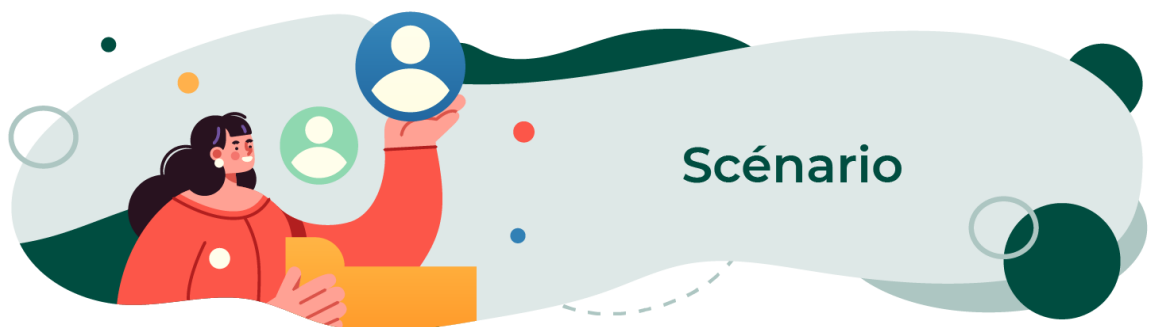


# Réalisez une étude de santé publique avec R ou Python

## Mission

90 heures

Version antérieure à la mise à jour du 29 mars 2023



Félicitations ! Vous avez récemment rejoint une équipe de chercheurs de la [Food and Agriculture Organization of the United Nations \(FAO\)](#), en tant que data analyst.



# Food and Agriculture Organization of the United Nations

Food and Agriculture Organization of the United Nations

Son rôle ? C'est l'un des organes qui composent l'ONU et dont l'objectif est « d'aider à construire un monde libéré de la faim ».

Le responsable de l'équipe, Marc, un chercheur en économie de la santé, est particulièrement enthousiaste à l'idée de votre arrivée au sein de son équipe. En effet, ce recrutement tombe à pic car Julien, l'ancien data analyst, a récemment été muté dans une autre équipe, et la vôtre vient de se voir confier la mission de réaliser une étude de grande ampleur sur le thème de la sous-nutrition dans le monde.

Dès votre arrivée, vous trouvez un mail de Marc dans votre boîte mail :

**Objet** : Bienvenue

**De** : Marc

**À** : Vous

Hello et bienvenue au sein de la FAO. J'espère que tu te plairas parmi nous. En tous cas, sache que je suis très content de t'accueillir au sein de mon équipe.

Tu es probablement au courant, mais nous avons une grosse tâche qui nous attend, et nous comptons particulièrement sur toi pour nous donner un panorama de l'état de la malnutrition dans le monde, à partir des

données que tu trouveras dans le dossier en pièce jointe. Je t'invite d'ailleurs fortement à lire le [Lexique des données](#) qui s'y trouve, écrit par Julien. Dans ce document, tu trouveras le descriptif des différents fichiers ainsi que des notes qui te sont adressées. Nous souhaitons que tu poursuives ses recherches en répondant aux questions qu'il t'a laissé.

Les données sur lesquelles j'aimerais particulièrement avoir des informations sont, pour l'année 2017 :

- la proportion de personnes en état de sous-nutrition ;
- le nombre théorique de personnes qui pourraient être nourries. Tu devrais pouvoir calculer ça à partir de la disponibilité alimentaire mondiale ;
- idem pour la disponibilité alimentaire des produits végétaux ;
- l'utilisation de la disponibilité intérieure, en particulier la part qui est attribuée à l'alimentation animale, celle qui est perdue et celle qui est concrètement utilisée pour l'alimentation humaine. Je crois que Julien avait trouvé un moyen de facilement calculer ces proportions.

Ce sont les informations principales dont on aimerait que tu puisses nous rendre compte, mais toute idée additionnelle est bonne à prendre !

Il y a aussi Mélanie, une autre chercheuse de l'équipe qui avait quelques demandes. Elle devrait prendre contact avec toi rapidement à ce sujet.

N'hésite pas à venir me voir si tu as la moindre question !

Marc.

PJ : [données FAO.zip](#)

À peine le temps de préparer votre café matinal, que vous recevez un message de Mélanie par messagerie instantanée :

**Mélanie :** Hello, j'espère que tout va bien et que t'intègres bien. En tout cas, je te souhaite à mon tour la bienvenue dans notre équipe.

**Vous :** Merci ! Je prends mes marques petit à petit. Marc m'a dit que tu avais une demande particulière ?

**Mélanie :** Tout à fait ! Je souhaiterais avoir une étude un peu plus fine pour chacun des pays. Par exemple, j'aimerais que tu puisses nous donner les pays pour lesquels la proportion de personnes sous-alimentées est la plus forte en 2017, ceux qui ont le plus bénéficié d'aide depuis 2013, ceux ayant le plus/le moins de disponibilité/habitant, etc., et toutes les infos que tu trouverais utiles pour mettre en relief les pays qui semblent être le plus en difficulté, au niveau alimentaire.

**Vous :** OK pas de soucis, je vais regarder ça.

**Mélanie :** Super, bon courage !

**Vous :** Merci, à bientôt !

Vous gardez en tête toutes ces requêtes, et attaquez immédiatement votre « exploration ».

Dans ce projet, vous êtes libre d'utiliser le langage de programmation de votre choix entre **R et Python** pour mener à bien votre analyse. Vous pouvez utiliser Jupyter (pour R et pour Python) ou R Markdown (pour R).



- Le notebook R Markdown ou Jupyter vous ayant permis de calculer/trouver l'ensemble des résultats demandés.

Pour faciliter votre passage devant le jury, déposez sur la plateforme, dans un dossier zip nommé "**Titre\_du\_projet\_nom\_prénom**", votre livrable nommé comme suit : **Nom\_Prénom\_n° du livrable\_nom du livrable\_date de démarrage du projet**. Cela donnera :

- *Nom\_Prénom\_1\_notebook\_mmaaaa*

Par exemple, votre livrable peut être nommé comme suit :

*Dupont\_Jean\_1\_notebook\_012023.*



Durant la présentation orale, l'évaluateur interprétera le rôle de Marc, votre responsable direct. La soutenance est structurée de la manière suivante :

- **Présentation du notebook (15 minutes)**
- **Discussion (10 minutes)**
  - L'évaluateur jouera le rôle de Marc. Il pourra revenir sur certains points pour vous questionner sur vos choix, et vérifier vos connaissances sur certaines fonctions du langage que vous avez choisi.
- **Debrief (5 minutes)**
  - À la fin de la soutenance, l'évaluateur arrêtera de jouer le rôle de Marc pour vous permettre de débriefer ensemble.

Votre présentation devrait durer 15 minutes (+/- 5 minutes). Puisque le respect des durées des présentations est important en milieu professionnel, les présentations en dessous de 10 minutes ou au-dessus de 20 minutes peuvent être refusées.

## Compétences évaluées

- Utiliser des librairies spécialisées pour les traitements data
- Créer un environnement de développement
- Manipuler des DataFrames