Vous êtes intégré à une nouvelle équipe de chercheurs de la [Food and Agriculture Organization of the United Nations (FAO)](https://fr.wikipedia.org/wiki/Organisation_des_Nations_unies_pour_l%27alimentation_et_l%27agriculture), l'un des organes qui compose l'ONU et dont l'objectif est d' « aider à construire un monde libéré de la faim ».

Votre équipe est chargée de réaliser une étude de grande ampleur sur le thème de la sous-nutrition dans le monde.

Le problème de la faim est complexe et peut avoir de multiples causes, différentes selon les pays. L’étape préliminaire de cette étude sera donc d’établir un “état de l’art” des recherches déjà publiées, mais également de mener une étude statistique destinée à orienter les recherches vers des pays particuliers, et de mettre en lumière différentes causes de la faim.

### **Les données**

Les données sont disponibles dans le dossier data et sont constituées de 5 fichiers:

* fr\_animaux.csv : multiples indicateurs de production des produits animaux en 2013
* fr\_population.csv: population mondiale par pays en 2013
* fr\_vegetaux.csv: multiples indicateurs de production des produits végétaux en 2013
* fr\_céréales.csv: quantité de céréales produites au niveau mondial en 2013
* fr\_sousalimentation.csv: nombre de personnes sous alimentées dans le monde de 2013 à 2017.

#### **Identifiez les grandes tendances**

Une fois les fichiers CSV téléchargés, il vous faudra les charger dans Python, afin d’identifier les grandes tendances et vous familiariser avec le jeu de données.

Créez un dataframe contenant les informations de population de chaque pays. Calculez le nombre total d’humains sur la planète. Critiquez votre résultat. En cas d’anomalie, analysez et effectuer les corrections nécessaires.

Question 1 : donnez le résultat de votre calcul pour l'année 2013.

Parmi les documents sur les Bilans alimentaires que vous avez téléchargés, il y a des informations redondantes. En effet, pour un pays donné, certaines de ces informations peuvent se calculer à partir d'autres :

* Production
* Importations
* Exportations
* Variation de stock
* Disponibilité intérieure
* Semences
* Pertes
* Nourriture, aussi appelée Disponibilité alimentaire
* Aliments pour animaux
* Traitement
* Autres utilisations

Question 2 : Identifiez ces redondances, en donnant votre réponse sous forme de formule mathématique (pas besoin de coder ici :soleil: ). C'est une équation à 3 termes de type \(a\_1 + a2 + [...] = b\_1 + b\_2 + [...] = c\_1 + c\_2 + [...]\) ) faisant intervenir chacune des 11 quantités données ci dessus. Illustrez cette équation avec l'exemple du blé en France. Pour avoir un indice, cliquez sur "Définitions et Standards" sur [**la page de téléchargement**](http://www.fao.org/faostat/fr/#data/FBS) des données.

Question 3 : Calculez (pour chaque pays et chaque produit) la disponibilité alimentaire en kcal puis en kg de protéines.

Vous ferez cela à partir de ces informations :

* Population de chaque pays ;
* Disponibilité alimentaire donnée pour chaque produit et pour chaque pays en kcal/personne/jour.
* Disponibilité alimentaire en protéines donnée pour chaque produit et pour chaque pays en g/personne/jour.

Question 4 : A partir de ces dernières informations, et à partir du poids de la disponibilité alimentaire (pour chaque pays et chaque produit), calculez pour chaque produit le ratio "énergie/poids", que vous donnerez en kcal/kg. Vous pouvez vérifier la cohérence de votre calcul en comparant ce ratio aux données disponibles sur internet, par exemple en cherchant la [**valeur calorique d'un oeuf**](https://fr.wikipedia.org/wiki/%C5%92uf_(aliment)).

Indication : La disponibilité alimentaire en kcal/personne/jour est calculée par la FAO en multipliant la quantité Nourriture par le ratio énergie/poids (en kcal/kg), puis en le divisant par la population du pays puis par 365. Ici, on vous demande juste de retrouver le ratio énergie/poids que la FAO a utilisé dans son calcul.

En suivant la même méthodologie, calculez également le pourcentage de protéines de chaque produit (pour chaque pays). Ce pourcentage est obtenu en calculant le ratio "poids de protéines/poids total" (attention aux unités utilisées). Vous pouvez vérifier la cohérence de votre calcul en comparant ce ratio aux données disponibles sur internet, par exemple en cherchant la teneur en protéines de l'[avoine](https://fr.wikipedia.org/wiki/Avoine_cultiv%C3%A9e).

Question 5 : Citez 5 aliments parmi les 20 aliments les plus caloriques, en utilisant le ratio énergie/poids.

Étonnamment, il arrive que ce ratio soit différent en fonction du pays. Il faudra donc réaliser pour chaque aliment une moyenne sur les différents pays. Vous créerez donc une nouvelle table grâce à une agrégation. **Attention à bien retirer les valeurs égales à 0 afin de ne pas fausser le calcul de la moyenne**.

Citez 5 aliments parmi les 20 aliments les plus riches en protéines.

Question 6 : Calculez, pour les produits végétaux **uniquement**, la disponibilité intérieure mondiale exprimée en kcal.

Question 7 : Combien d'humains pourraient être nourris si toute la disponibilité intérieure mondiale de produits végétaux était utilisée pour de la nourriture ? Donnez les résultats en termes de calories, puis de protéines, et exprimez ensuite ces 2 résultats en pourcentage de la population mondiale.

Question 8 : Combien d'humains pourraient être nourris si toute la disponibilité alimentaire en produits végétaux, la nourriture végétale destinée aux animaux et les pertes de produits végétaux étaient utilisés pour de la nourriture ? Donnez les résultats en termes de calories, puis de protéines, et exprimez ensuite ces 2 résultats en pourcentage de la population mondiale.

Question 9 : Combien d'humains pourraient être nourris avec la disponibilité alimentaire mondiale ? Donnez les résultats en termes de calories, puis de protéines, et exprimez ensuite ces 2 résultats en pourcentage de la population mondiale.

Question 10 : A partir des données téléchargées qui concernent la sous-nutrition, répondez à cette question : Quelle proportion de la population mondiale est considérée comme étant en sous-nutrition ?

Établissez la liste des produits (ainsi que leur code) considérés comme des céréales selon la FAO.

Repérez dans vos données les informations concernant les céréales (par exemple en créant une colonne de type booléen nommée "is\_cereal").

Question 11 : En ne prenant en compte que les céréales destinées à l'alimentation (humaine et animale), quelle proportion (en termes de poids) est destinée à l'alimentation animale ?

Sélectionnez parmi les données des bilans alimentaires les informations relatives aux pays dans lesquels la FAO recense des personnes en sous-nutrition.

Repérez les 15 produits les plus exportés par ce groupe de pays.

Parmi les données des bilans alimentaires au niveau mondial, sélectionnez les 200 plus grandes importations de ces produits (1 importation = une quantité d'un produit donné importée par un pays donné)

Groupez ces importations par produit, afin d'avoir une table contenant 1 ligne pour chacun des 15 produits. Ensuite, calculez pour chaque produit les 2 quantités suivantes :

* le ratio entre la quantité destinés aux "Autres utilisations" (Other uses) et la disponibilité intérieure.
* le ratio entre la quantité destinée à la nourriture animale et la quantité destinée à la nourriture (animale + humaine)

Question 12 : Donnez les 3 produits qui ont la plus grande valeur pour chacun des 2 ratios (vous aurez donc 6 produits à citer).

Question 13 : Combien de tonnes de céréales pourraient être libérées si les USA diminuaient leur production de produits animaux de 10% ?

#### **Enrichissez votre analyse**

Lors de votre présentation, vous présenterez les pistes étudiées jusqu'à maintenant grâce aux données de la FAO. En réalité, vous n'êtes pas le/la premier·e à étudier ce phénomène (bien heureusement!). Veillez donc à bien confirmer vos intuitions par des recherches sur internet : Pour une bonne analyse, le data analyst doit comprendre le domaine qu'il étudie. On appelle cela les "connaissances métier".

Notamment, cherchez les réponses à ces questions :

* Combien de personnes décèdent des causes de la faim ?
* Quelles sont les prévisions de population mondiale en 2050 ?

**Sources :**

Le paradoxe de la faim dans le monde :

<https://www.youtube.com/watch?v=iswD-wgR6Yg>

Peut-on régler le problème de la faim dans le monde ?

<https://www.youtube.com/watch?v=wRNHe2keeLw>

Le dérèglement climatique bouleverse le monde

<https://www.youtube.com/watch?v=0SXf8E9mxkc>

Petite initiation au problème des accaparements de terres

<https://www.youtube.com/watch?v=a4vtfl-zv6Q>

La bourse ou la faim ?

<https://www.youtube.com/watch?v=9sz0YgriE7Q>

Quand la boucherie, le monde pleure

[Quand la boucherie, le monde pleure - #DATAGUEULE 55](https://www.youtube.com/watch?v=KriTQ0aTrtw)