

# Mise en situation

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

Le problème de la faim est en effet complexe et peut avoir de multiples causes, différentes selon les pays. Votre équipe est chargée de réaliser une étude de grande ampleur sur le thème de la sous-nutrition dans le monde.

## Télécharger les données

L'interface du site de la FAO permet de sélectionner les types de données que vous souhaitez télécharger. Comme le volume de celles-ci est important, l'interface permet un premier filtrage selon les **pays** (ou *groupements de pays et régions du monde*), les **années**, les **éléments**, les **produits** (et groupements de produits : céréales, produits animaux, etc.).

**N.B :** La FAO sépare ses données en 2 catégories de produits alimentaires distincts : les produits issus des **animaux** et ceux issus de **végétaux**.

Il faudra prendre un peu de temps pour se familiariser avec cette interface. Avant de télécharger les fichiers CSV, il est possible de tester le filtrage grâce à l'option de prévisualisation (bouton “Montrer les données”).

Je vous conseille de se limiter aux cinq tables suivantes (toutes les années depuis 2013) :

1. Une table des bilans alimentaires pour les produits animaux
2. Une table des bilans alimentaires pour les produits végétaux
3. Une table donnant la population de chaque pays
4. Une table donnant les informations sur le nombre de personnes en sous-nutrition par pays
5. Une liste des produits considérés comme “céréales”

Pour les tables 1 et 2, il faut bien faire attention à ne sélectionner QUE “Produits Végétaux > (Liste)” et “Produits animaux > (Liste)” (voir les captures d'écran ci-dessous). Sinon, il risque d'y avoir de la redondance dans les données (exemple : une ligne contenant “viande bovine” et une ligne contenant “viande”. La ligne “viande” étant une agrégation de “viande bovine” et d'autres types de viandes).

Attention, les données de la FAO changent parfois suite à des actualisations. Il est possible alors que vos résultats d'analyse diffèrent suivant les dates de téléchargement. Mettez une stratégie pour avoir toujours des résultats à jour.

The screenshot shows the FAO Faostat web interface. The URL is [www.fao.org/faostat/en/#data/FBS](http://www.fao.org/faostat/en/#data/FBS). The interface is divided into several sections:

- COUNTRIES**: A list of countries with checkboxes. A red box highlights the "Select All" button.
- ELEMENTS**: A list of elements with checkboxes. A red box highlights the "Select All" button.
- ITEMS**: A list of items with radio buttons. A red box highlights the "Vegetal Products > (List)" option.
- YEARS**: A list of years with radio buttons. A red box highlights the "2013 (2013 and last available year)" option.
- Output Type**: A dropdown menu set to "Table".
- Thousand Separator in 'Show Data'**: Radio buttons for "None", "Comma", and "Period".
- Output Formatting Options**: Checkboxes for "Flags", "Codes", "Units", and "Null Values".

On the right side, there is a sidebar with the following sections:

- Food Balance Sheets**: A description of the data and a link to "Show More".
- Bulk Downloads**: A table of download options.
- Last Update**: December 12, 2017.
- Related Documents**: A link to "Update history".
- Definitions and stand...**: A link to "Definitions and stand...".
- Metadata**: A link to "Metadata".

Download Option	Size
All Data	15.84 MB
All Data Normalized	78.96 MB
All Area Groups	3.38 MB
Africa	2.7 MB
Americas	2.51 MB
Asia	2.88 MB
Europe	2.92 MB
Oceania	499 KB

## COUNTRIES REGION SPECIAL GROUPS

Filter results

- ☒ Afghanistan
- ☒ Albania
- ☒ Algeria
- ☒ Angola
- ☒ Antigua and Barbuda
- ☒ Argentina

Select All

Clear All

Afghanistan x Albania x Algeria x Angola x  
Antigua and Barbuda x Argentina x Armenia x  
Australia x Austria x Azerbaijan x Bahamas x  
Bangladesh x Barbados x Belarus x Belgium x

## ITEMS ITEMS AGGREGATED

Filter results e.g. population

- ☒ Population
- ☐ Wheat and products
- ☐ Rice (Milled Equivalent)
- ☐ Barley and products
- ☐ Maize and products
- ☐ Rye and products

Select All

Clear All

Population x

## ELEMENTS

Filter results e.g. total population - both sexes

- ☒ Total Population - Both sexes
- ☐ Production Quantity
- ☐ Import Quantity
- ☐ Stock Variation
- ☐ Export Quantity
- ☐ Domestic supply quantity

Select All

Clear All

Total Population - Both sexes x

## YEARS

Filter results e.g. 2013

- ☒ 2013 (2013 and last available year)
- ☐ 2012
- ☐ 2011
- ☐ 2010
- ☐ 2009
- ☐ 2008

Select All

Clear All

2013 x

Output Type

☒ Table ☐ Pivot

Thousand Separator in 'Show Data'

☒ None ☐ Comma ☐ Period

Output Formatting Options

☒ Flags ☐ Codes ☐ Units ☐ Null Values

## Food Balance Sheets

Food Balance Sheet presents a comprehensive picture of the pattern of a country's food supply during a specified reference period. The food balance... [Show More](#)

Food and Agriculture Organization of the United Nations (FAO)

## Bulk Downloads

All Data	15.84 MB
All Data Normalized	78.96 MB
All Area Groups	3.38 MB
Africa	2.7 MB
Americas	2.51 MB
Asia	2.88 MB
Europe	2.92 MB
Oceania	499 KB

## Last Update

December 12, 2017

## Related Documents

[Update history](#)[Definitions and stand...](#)[Metadata](#)

COUNTRIES REGIONS SPECIAL GROUPS

Filter results e.g. afghanistan

- ☒ Afghanistan
- ☒ Albania
- ☒ Algeria
- ☒ American Samoa
- ☒ Andorra
- ☒ Angola

Select All

Clear All

Afghanistan x Albania x Algeria x

American Samoa x Andorra x Angola x

Antigua and Barbuda x

ITEMS

Filter results e.g. availability > (list)

- ☐ -- Rail lines density (total route in km per 100 squar...
- ☐ -- Gross domestic product per capita, PPP, dissemi...
- ☐ -- Prevalence of undernourishment (percent)
- ☒ -- Number of people undernourished (million)
- ☐ -- Prevalence of severe food insecurity in the total ...
- ☐ -- Number of severely food insecure people (million)

Select All

Clear All

-- Number of people undernourished (million) x

ELEMENTS

Filter results e.g. value

- ☒ Value

Select All

Clear All

Value x

YEARS

Filter results e.g. 2017 / 2016-2018

- ☒ 2017 / 2016-2018
- ☒ 2016 / 2015-2017
- ☒ 2015 / 2014-2016
- ☒ 2014 / 2013-2015
- ☒ 2013 / 2012-2014
- ☐ 2012 / 2011-2013

Select All

Clear All

2013 / 2012-2014 x 2014 / 2013-2015 x

2015 / 2014-2016 x 2016 / 2015-2017 x

2017 / 2016-2018 x

Output Type

☒ Table

☐ Pivot

Thousand Separator in 'Show Data'

☒ None

☐ Comma

☐ Period

Output Formatting Options

☒ Flags

☐ Codes

☐ Units

☐ Null Values

## Suite of Food Security Indicators

For detailed description of the indicators below see attached document: Average Dietary Supply Adequacy; Average Value of Food Production; Share of... [Show More](#)

Food and Agriculture Organization of the United Nations (FAO)

### Bulk Downloads

All Data	325 KB
All Data Normalized	661 KB
All Area Groups	54 KB
Africa	77 KB
Latin America and the Caribbean	52 KB
Northern America and Europe	60 KB
Asia	76 KB
Oceania	20 KB

### Last Update

October 16, 2018

### Related Documents

[Update history](#)

[Definitions and stand...](#)

[Metadata](#)

COUNTRIES REGIONS SPECIAL GROUPS

Filter results e.g. afghanistan

☒ World + (Total)

☐ World > (List)

☐ Africa + (Total)

☐ Africa > (List)

☐ -- Eastern Africa + (Total)

☐ -- Eastern Africa > (List)

Select All

Clear All

World + (Total) X

ITEMS

ITEMS AGGREGATED

Filter results e.g. population

☐ Animal Products > (List)

☐ Cereals - Excluding Beer + (Total)

☒ Cereals - Excluding Beer > (List)

☐ Starchy roots + (Total)

☐ Starchy Roots > (List)

☐ Sugar Crops + (Total)

Select All

Clear All

Cereals - Excluding Beer > (List) X

ELEMENTS

Filter results e.g. total population - both sexes

☐ Total Population - Both sexes

☒ Production Quantity

☐ Import Quantity

☐ Stock Variation

☐ Export Quantity

☐ Domestic supply quantity

Select All

Clear All

Production Quantity X

YEARS

Filter results e.g. 2013

☒ 2013

(2013 and last available year)

☐ 2012

☐ 2011

☐ 2010

☐ 2009

☐ 2008

Select All

Clear All

2013 X

Output Type

☒ Table

☐ Pivot

Thousand Separator in 'Show Data'

☒ None

☐ Comma

☐ Period

Output Formatting Options

☒ Flags

☒ Codes

☒ Units

☐ Null Values

## Food Balance Sheets

Food Balance Sheet presents a comprehensive picture of the pattern of a country's food supply during a specified reference period. The food balance... [Show More](#)

Food and Agriculture Organization of the United Nations (FAO)

### Bulk Downloads

All Data	15.84 MB
All Data Normalized	78.96 MB
All Area Groups	3.38 MB
Africa	2.7 MB
Americas	2.51 MB
Asia	2.88 MB
Europe	2.92 MB
Oceania	499 KB

### Last Update

December 12, 2017

### Related Documents

[Update history](#)

[Definitions and stand...](#)

[Metadata](#)

# Entrer vos données dans une BD relationnelle

Lors de la première réunion, vous avez été désigné pour mettre en place la base de données que votre équipe pourra requêter (en SQL) à souhait pour réaliser cette étude.

**Indice : Pour exporter un dataframe de R ou Python vers une base de données, l'une des solutions est de passer via un fichier CSV.**

Sous Python, utilisez `to_csv` : `df.to_csv("export.csv", index=False)`

Sous R, utilisez `write_csv` : `readr::write_csv(df, file="export.csv")`

Avant de passer à la création de la BD, je vous conseille vivement de travailler sur un dataframe principal (qui fusionne *animal* et *vegetal*) de manière à obtenir une colonne pour chaque élément (Production, Importation, Disponibilité intérieure, Semences, etc.) de cette manière :

```
1 # Import des librairies
2 import numpy as np
3 import pandas as pd
4 import seaborn as sns
5
6 # Import des données des bilans alimentaires
7 veg = pd.read_csv("data/vegetal.csv")
8 ani = pd.read_csv("data/animal.csv")
9
10 # Ajout de la variable origin
11 ani["origin"] = "animal"
12 veg["origin"] = "vegetal"
13
14 # On regroupe veg et ani en un unique dataframe, via une union
15 temp = ani.append(veg)
16
17 # Suppression de ani et veg
18 del ani, veg
19
20 # On renomme les colonnes de temp
21 temp.columns = ["xx", "xx2", "country_code", "country", "xx3", 'element'
22                , 'item_code', 'item', 'xx4', "year", "unit", "value", 'xx5', 'xx6'
23                , 'origin']
24
25 # Transformation de temp en table pivot
26 data = temp.pivot_table(
27     index=["country_code", "country", "item_code", "item", "year", "origin"],
28     columns = ["element"], values=["value"], aggfunc=sum)
29 # On renomme les colonnes (attention l'ordre peut changer selon vos données !)
30 data.columns = ['domestic_supply_quantity', 'export_quantity', 'feed',
31                'food', 'food_supply_kcalcapitaday',
32                'food_supply_quantity_kgcapitayr', 'import_quantity', 'other_uses',
33                'processing', 'production', 'protein_supply_quantity_gcapitaday',
34                'seed', 'stock_variation', 'waste']
35
36 data = data.reset_index()
37 data.head()
```

```
data.head()
```

	country_code	country	item_code	item	year	origin	Domestic supply quantity	Export Quantity	Feed	...	Food supply (kcal/capita/day)	Food supply quantity (kg/capita/yr)	Import Quantity	Other uses	Processing	Production	Protein supply quantity (g/capita/day)	Seed	Stock Variation	Waste
0	1	Armenia	2511	Wheat and products	2013	vegetal	554.0	1.0	93.0	...	1024.0	130.60	361.0	0.0	10.0	312.0	30.52	30.0	-118.0	32.0
1	1	Armenia	2513	Barley and products	2013	vegetal	198.0	0.0	137.0	...	0.0	0.00	9.0	26.0	7.0	189.0	0.00	14.0	0.0	15.0
2	1	Armenia	2514	Maize and products	2013	vegetal	102.0	NaN	96.0	...	0.0	0.03	82.0	NaN	NaN	21.0	0.01	0.0	NaN	7.0
3	1	Armenia	2515	Rye and products	2013	vegetal	1.0	NaN	1.0	...	1.0	0.12	0.0	NaN	NaN	1.0	0.02	0.0	0.0	0.0
4	1	Armenia	2516	Oats	2013	vegetal	6.0	NaN	4.0	...	2.0	0.37	1.0	NaN	NaN	5.0	0.09	0.0	NaN	0.0

Vous pouvez pour ceci vous aider de la méthode [pivot\\_table](#), aussi via [cet exemple](#).

## Votre base devra contenir ces différentes tables :

### 1 Table population

- Une table appelée **population**, contenant la *population de chaque pays pour chaque année comprise entre 2013 et l'année courante*. Elle devra contenir 4 colonnes : **country**, **country\_code**, **year**, **population**.

#### 1.1 Q : Taille de la population mondiale ?

En se basant sur le dataframe construit avec Python ou R (à partir des CSV brutes), calculer la taille de la population mondiale (pour chaque année) ? **Indice** : pour 2013 ; 6 997 326 000 humains. Refaire le même calcul avec SQL ? **Indice** : même chiffre que la 1<sup>è</sup> Q. Constat que vos résultats selon la FAO sont alors bien au-dessus de la taille réelle de la population mondiale (par ex La Banque Mondiale). Trouver la source d'anomalie et effectuer les corrections nécessaires.

**Indice** : Il faut remarquer que la Chine est présente 2 fois dans les tables. C'est un problème que vous devriez corriger. Il faut exclure les lignes dont le code\_pays est 351 car il correspond à l'agrégation des pays 41, 96, 241 et 128.

### 2 Table dispo\_alim

- Une table appelée **dispo\_alim** contenant pour chaque pays, pour chaque produit, et pour chaque année, les informations suivantes :
  - La nature du produit (deux valeurs possibles : “animal” ou “végétal”)
  - Disponibilité alimentaire en tonnes
  - Disponibilité alimentaire en Kcal/personne/jour
  - Disponibilité alimentaire de protéines en g/personne/jour
  - Disponibilité alimentaire de matières grasses en g/personne/jour

Elle devra contenir ces colonnes : pays, code\_pays, année, produit, code\_produit, origin, dispo\_alim\_tonnes, dispo\_alim\_kcal\_p\_j,dispo\_prot, dispo\_mat\_gr .

## 2.1 Céréales

### 2.1.1 Q : Liste des céréales selon la FAO ?

Établir la liste des produits (ainsi que leur code) considéré comme des céréales selon la FAO et créer une colonne de type booléen nommée "is\_cereal" ?

#### Indice :

	Produit	Code Produit
<b>0</b>	Blé	2511
<b>11</b>	Riz (Eq Blanchi)	2805
<b>21</b>	Orge	2513
<b>33</b>	Maïs	2514
<b>45</b>	Millet	2517
<b>54</b>	Céréales, Autres	2520
<b>116</b>	Seigle	2515
<b>129</b>	Avoine	2516
<b>153</b>	Sorgho	2518

	country_code	country	item_code	item	year	origin	is_cereal
<b>0</b>	1	Armenia	2511	Wheat and products	2013	vegetal	True
<b>1</b>	2	Afghanistan	2511	Wheat and products	2013	vegetal	True
<b>2</b>	3	Albania	2511	Wheat and products	2013	vegetal	True
<b>3</b>	4	Algeria	2511	Wheat and products	2013	vegetal	True
<b>4</b>	7	Angola	2511	Wheat and products	2013	vegetal	True



## 2.1.2 Q : Proportion des céréales pour l'alimentation animale ?

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 ? **Indice : 46 %**

## 2.2 Comment mesure-t-on la disponibilité alimentaire ?

### 2.2.1 Q : Dispo alim par pays et par produit ?

Calculer (pour chaque pays et chaque produit) la disponibilité alimentaire en kcal puis en kg de protéines ?

**Indice :** Vous ferez cela à partir de ces informations :

- Population de chaque pays ;
- Disponibilité alimentaire ("Food Supply" en anglais) donnée pour chaque produit et pour chaque pays en kcal/personne/jour.
- Disponibilité alimentaire en protéines ("Protein supply quantity" en anglais) donnée pour chaque produit et pour chaque pays en g/personne/jour.

	country_code	country	item_code	item	year	origin	Domestic supply quantity	Export Quantity	Feed	Food	...	Other uses	Processing	Production	Protein supply quantity (g/capita/day)
0	1	Armenia	2511	Wheat and products	2013	vegetal	554.0	1.0	93.0	389.0	...	0.0	10.0	312.0	30.52
1	1	Armenia	2513	Barley and products	2013	vegetal	198.0	0.0	137.0	0.0	...	26.0	7.0	189.0	0.00
2	1	Armenia	2514	Maize and products	2013	vegetal	102.0	0.0	96.0	0.0	...	0.0	0.0	21.0	0.01
3	1	Armenia	2515	Rye and products	2013	vegetal	1.0	0.0	1.0	0.0	...	0.0	0.0	1.0	0.02
4	1	Armenia	2516	Oats	2013	vegetal	6.0	0.0	4.0	1.0	...	0.0	0.0	5.0	0.09

Protein supply quantity (g/capita/day)	Seed	Stock Variation	Waste	population	food_supply_kcal	food_supply_kgprotein
30.52	30.0	-118.0	32.0	2977000	1.112684e+12	33163184.60
0.00	14.0	0.0	15.0	2977000	0.000000e+00	0.00
0.01	0.0	0.0	7.0	2977000	0.000000e+00	10866.05
0.02	0.0	0.0	0.0	2977000	1.086605e+09	21732.10
0.09	0.0	0.0	0.0	2977000	2.173210e+09	97794.45

## 2.2.2 Q : Ratio "énergie/poids" ?

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 ?

**Indice 1 :** 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.

**Indice 2 :** La disponibilité alimentaire en kcal/personne/jour est calculée par la FAO en multipliant la quantité Nouriture (Food) 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.

## 2.2.3 Q : % de protéines de chaque produit ?

En suivant la même méthodologie, calculez le pourcentage de protéines de chaque produit (pour chaque pays).

**Indice :** 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.

## 2.3 Top 20 des aliments

### 2.3.1 Q : les plus caloriques

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

**Indication :** 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.

	ratio_kcal/kg	protein_%
item		
Fish, Body Oil	9003.715504	0.000000
Palmkernel Oil	8797.863661	0.069823
Palm Oil	8719.497538	0.137376
Maize Germ Oil	8683.086371	0.019466
Ricebran Oil	8658.062895	0.053281
Oilcrops Oil, Other	8645.450087	0.545248
Cottonseed Oil	8606.458261	0.061716
Soyabean Oil	8583.012799	0.060750
Rape and Mustard Oil	8492.322250	0.103028
Sunflowerseed Oil	8400.620285	0.042969
Olive Oil	8385.934335	0.000000
Groundnut Oil	8360.672341	0.012826
Coconut Oil	8302.647132	0.132093
Sesameseed Oil	7684.501021	0.000000
Fats, Animals, Raw	7603.648814	1.223827
Fats, Fish, Raw	7543.554004	0.700000

### 2.3.2 Q : les plus protéinés

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

**Indice :**

	ratio_kcal/kg	protein_%
item		
Soyabeans	2359.183293	28.515649
Groundnuts (Shelled Eq)	5622.381431	25.157275
Peas	3444.040776	22.321921
Pulses, Other and products	3347.672355	21.863897
Beans	3396.950228	21.653185
Sesame seed	5491.906604	17.611839
Offals, Edible	1130.422533	17.558732
Meat, Other	1122.542422	17.463164
Oilcrops, Other	3093.396839	15.728229
Infant food	3196.495430	14.658151
Bovine Meat	1685.444556	14.374246
Mutton & Goat Meat	1947.892555	13.584788
Cephalopods	488.881696	13.353059
Poultry Meat	1299.811793	12.626174
Pimento	3079.026152	11.983274

## 2.4 Disponibilité mondiale

### 2.4.1 Q : En végétaux uniquement

Calculer, pour les produits végétaux uniquement, la disponibilité intérieure mondiale exprimée en kcal et en Kg protéines pour chaque année et tracer la viz correspondante ?

**Indice :**

	dom_sup_kcal	dom_sup_kgprot
year		
2013	1.307838e+16	3.252844e+11

### 2.4.2 Q : Tous végétariens ?

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.

**Indice :**

Population potentiellement nourrie par la disponibilité intérieure en produits issus de végétaux (en termes calorifiques) : 14.33 Miliards,  
soit 205.0% de la population mondiale.  
Population potentiellement nourrie par la disponibilité intérieure en produits issus de végétaux (en termes de protéines) : 15.97 Miliards,  
soit 228.0% de la population mondiale.

### 2.4.3 Q : Rien ne se perd, tout se transforme

Combien d'humains pourraient être nourris si toute la disponibilité alimentaire en produits végétaux (Food), la nourriture végétale destinée aux animaux (Feed) et les pertes de produits végétaux (Waste) étaient utilisées 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.

### 2.4.4 Q : Tous bien nourris ?

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.

### Indice pour les deux dernières questions :

Population potentiellement nourrie par la disponibilité alimentaire, la nourriture animale et les pertes de produits végétaux (en termes calorifiques) : 9.61 Miliards,  
soit 137.0% de la population mondiale.  
Population potentiellement nourrie par la disponibilité alimentaire, la nourriture animale et les pertes de produits végétaux (en termes de protéines) : 9.66 Miliards,  
soit 138.0% de la population mondiale.  
- - - - -  
Population potentiellement nourrie par la disponibilité alimentaire mondiale (en termes calorifiques) : 8.07 Miliards,  
soit 115.0% de la population mondiale.  
Population potentiellement nourrie par la disponibilité alimentaire mondiale (en termes de protéines) : 10.18 Miliards,  
soit 145.0% de la population mondiale.

## 3 Table sous\_nutrition

Une table appelée **sous\_nutrition**, contenant le nombre de personnes en sous-alimentation pour chaque pays et pour chaque année. Elle devra contenir 4 colonnes : **pays**, **code\_pays**, **année**, **nb\_personnes**.

### 3.1 Q : Proportion de la pop en sous-nutrition ?

Quelle proportion de la population mondiale est considérée comme étant en sous-nutrition ?

### Indice :

Personnes en sous-nutrition en 2013: 690800000.0  
soit 9.872342663468874% de la population mondiale.

### 3.2 Q : Liste des pays en sous-nutrition

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

**Indice : liste non exhaustive (et spécifique à l'année 2013)**

	country_code	country
0	1	Armenia
1	2	Afghanistan
4	7	Angola
11	16	Bangladesh
13	19	Bolivia (Plurinational State of)

### 3.3 Q : Liste des produits les plus exportés

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

```
array(['Tomatoes and products', 'Coffee and products',  
      'Milk - Excluding Butter', 'Pineapples and products', 'Soyabeans',  
      'Fruits, Other', 'Maize and products', 'Pelagic Fish',  
      'Wheat and products', 'Vegetables, Other', 'Bananas',  
      'Sugar (Raw Equivalent)', 'Palm Oil', 'Rice (Milled Equivalent)',  
      'Cassava and products'], dtype=object)
```

### 3.4 Q : Les plus grandes importations ?

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

### Indice : liste non exhaustive

	country_code	country	item_code	item	year	origin	Domestic supply quantity	Export Quantity
<b>7372</b>	169	Paraguay	2615	Bananas	2013	vegetal	54.0	8.0
<b>7369</b>	165	Pakistan	2615	Bananas	2013	vegetal	33.0	86.0
<b>6158</b>	70	French Polynesia	2601	Tomatoes and products	2013	vegetal	1.0	0.0
<b>7364</b>	155	Vanuatu	2615	Bananas	2013	vegetal	16.0	0.0
<b>6166</b>	83	Kiribati	2601	Tomatoes and products	2013	vegetal	0.0	0.0
<b>6168</b>	86	Grenada	2601	Tomatoes and products	2013	vegetal	0.0	0.0
<b>7359</b>	147	Namibia	2615	Bananas	2013	vegetal	0.0	0.0
<b>6129</b>	25	Solomon Islands	2601	Tomatoes and products	2013	vegetal	0.0	0.0
<b>6240</b>	189	Saint Lucia	2601	Tomatoes and products	2013	vegetal	0.0	0.0
<b>7318</b>	93	Haiti	2615	Bananas	2013	vegetal	270.0	0.0
<b>7281</b>	35	Cabo Verde	2615	Bananas	2013	vegetal	10.0	0.0

### 3.5 Q : Regrouper les importations par produit

Grouper ces importations par produit, afin d'avoir une table contenant 1 ligne pour chacun des 15 produits. Ensuite, calculer 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)

### Indice :

		country_code	year	Domestic supply quantity	Export Quantity	Feed	Food	Food supply (kcal/capita/day)	Food supply quantity (kg/capita/yr)	Import Quantity	Other uses
item_code	item										
2577	Palm Oil	21056	336171	50409.0	50673.0	0.0	15214.0	8939.0	371.14	48858.0	35191.0
2514	Maize and products	21697	344223	955798.0	128786.0	546116.0	125184.0	30799.0	3677.13	122692.0	189387.0
2532	Cassava and products	20576	330132	256664.0	36052.0	87056.0	100637.0	9016.0	3274.98	38371.0	35987.0

		country_code	year	Domestic supply quantity	Export Quantity	Feed	Food	feed_vs_food
item_code	item							
2514	Maize and products	21697	344223	955798.0	128786.0	546116.0	125184.0	0.813520
2555	Soyabeans	20844	328119	267445.0	106912.0	17476.0	10649.0	0.621369
2763	Pelagic Fish	21695	342210	41025.0	23135.0	18817.0	21687.0	0.464571

### 3.6 Q : Top 3 produits

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

		country_code	year	Domestic supply quantity	Export Quantity	Feed	Food	Food supply (kcal/capita/day)	Food supply quantity (kg/capita/yr)	Import Quantity	Other uses
item_code	item										
2514	Maize and products	21697	344223	955798.0	128786.0	546116.0	125184.0	30799.0	3677.13	122692.0	189387.0
2555	Soyabeans	20844	328119	267445.0	106912.0	17476.0	10649.0	948.0	105.56	102082.0	638.0
2763	Pelagic Fish	21695	342210	41025.0	23135.0	18817.0	21687.0	3016.0	1210.95	28600.0	376.0

protein_%_x	ratio_kcal/kg	protein_%	dom_sup_kcal	dom_sup_kgprot	food_feed_kcal	food_feed_kgprot	is_cereal	other_uses_%	feed_vs_food
1066.404319	515152.764142	1240.511146	2.879427e+15	6.933790e+10	2.022351e+15	4.869913e+10	171.0	0.198145	0.813520
2281.251895	384546.876830	4648.050736	6.309518e+14	7.626368e+10	6.635203e+13	8.020026e+09	0.0	0.002386	0.621369
1959.204352	146083.969405	2018.574180	3.525350e+13	4.871294e+09	3.480579e+13	4.809431e+09	0.0	0.009165	0.464571

### 3.7 Q :

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

### 3.8 Q :



En Thaïlande, quelle proportion de manioc est exportée ? Quelle est la proportion de personnes en sous-nutrition ?

## 4 Table equilibre\_prod

- Une table appelée **equilibre\_prod** contenant pour chaque pays, pour chaque produit, et pour chaque année, les quantités suivantes :
  - Disponibilité intérieure
  - Aliments pour animaux
  - Semences
  - Pertes
  - Transformés
  - Nourriture
  - Autres utilisations
- 1. Elle devra contenir ces colonnes : pays, code\_pays, année, produit, code\_produit, dispo\_int, alim\_an, semences, pertes, transfo, nourriture, autres\_utilisations.

### 4.1 Q :

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 (*Production*)
- Importations (*ImportQuantity*)
- Exportations (*ExportQuantity*)
- Variation de stock (*StockVariation*)
- Disponibilité intérieure (*DomesticSupply*)
- Semences (*Seed*)
- Pertes (*Waste*)
- Nourriture (*Food*), aussi appelée Disponibilité alimentaire (*FoodSupply*)
- Aliments pour animaux (*Feed*)
- Traitement (*Processing*)
- Autres utilisations (*Otheruses*)

Identifiez ces redondances en donnant votre réponse sous forme de formule mathématique 😊. Illustrez cette relation avec l'exemple du blé en France (indice = 20298).

**Indice :** Cliquez sur "Définitions et Standards" sur la page de téléchargement des données. Il s'agit de la "balance" des produits alimentaires pour chaque pays. C'est une équation à 3 termes faisant intervenir chacune des 11 quantités données ci-dessus :

$$a_1 + a_2 + [\dots] = b_1 + b_2 + [\dots] = c_1 + c_2 + [\dots]$$

$\text{Production} + \text{Importations} - \text{Exportations} + \text{Variation de stock}$
---

$$=$$

$$\text{Disponibilité intérieure}$$

$$=$$

$$\text{Semences} + \text{Pertes} + \text{Nourriture} + \text{Aliments pour animaux} + \text{Traitement} + \text{Autres utilisations}$$

Remarque importante : Il arrive que l'égalité entre les termes 2 et 3 de l'équation ne soit pas toujours vraie (surtout quand domestic supply est négatif), mais les impacts sont négligeables sur le reste des calculs.

## 5 Q : Query SQL

Écrivez les requêtes SQL permettant de connaître... :

1. Les 10 pays ayant **le plus haut ratio disponibilité alimentaire/habitant** en termes de protéines (en kg) par habitant, puis en termes de kcal par habitant.
2. Pour chaque année disponible, les 10 pays ayant **le plus faible ratio disponibilité alimentaire/habitant** en termes de protéines (en kg) par habitant. Le nombre de lignes de la table renvoyée sera donc égal à 10 fois le nombre d'années disponibles.
3. La quantité totale (en kg) de produits perdus par pays et par année. La table renvoyée contiendra donc une ligne par couple (pays, année).
4. Les 10 pays pour lesquels la proportion de personnes sous-alimentées est la plus forte.
5. Les 10 produits pour lesquels le ratio Autres utilisations/Disponibilité intérieure est le plus élevé.

### 5.1 Enrichir votre analyse

Chercher sur le site de la FAO les réponses à ces questions :

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

## 6 Modalités de la soutenance

- 7 min Mise en contexte auprès de votre équipe. Donnez les chiffres clés (nombre de personnes qui décèdent à cause de la faim, chiffres de la sous-nutrition mondiale, évolution au cours des années précédentes, évolution possible au cours des années suivantes). Donnez ensuite différentes causes de la faim.
- 7 min Approfondissez l'une de ces causes (ou plusieurs de ces causes) en présentant les résultats de vos calculs. Vous conclurez cette partie en répondant à ces 2 questions par exemple :
- La faim dans le monde résulte t'elle d'un manque de production, ou de problèmes technologiques ?
  - Quelles sont les prévisions de population en 2050 ? Aura t'on besoin d'augmenter drastiquement la production alimentaire ?
  - Il n'est pas nécessaire de répondre scolairement à chacune des questions du projet durant la soutenance, mais vous devez présenter les chiffres les plus importants que vous avez trouvé en les incluant dans votre analyse.
- 3 min Détaillez des données téléchargées : source, combien de fichiers CSV, à quoi correspondent-ils, à quoi correspond chaque ligne d'un fichier, etc. (il est possible de montrer dans votre présentation des premières lignes des fichiers en les détaillant). Détaillez de la même manière les principales tables (dataframes ou tables sql) utilisées lors de l'analyse.
- 3 min Détaillez quelques requêtes avec au moins :
- Une agrégation
  - Une jointure (en justifiant le type de jointure : interne, externe gauche, etc.)
  - Une restriction
- Pour chacune des opérations d'algèbre relationnelle détaillée, indiquez les clés primaires des tables avant et après opération. Il est conseillé de donner des captures d'écran des tables pour expliquer les opérations utilisées
- 5 min Montrez le résultat de chacune des requêtes (sous forme de capture d'écran). Détaillez le code SQL de 2 d'entre elles.
- 5 min Questions-réponses de l'examinateur