

Traitement en R

1- Table 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^{ère} Q.** Constaté 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 Céréales

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" ?

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

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

3- Table Céréales

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 %**

4- Table dispo_alim

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

5- Table dispo_alim

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.

6- Table dispo_alim

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.

7- Table dispo_alim

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 créez 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

8- Table dispo_alim

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

	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	4400.540400	17.462464

9- Table dispo_alim et végétaux

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

10- Table dispo_alim et végétaux

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.

11- Table dispo_alim et végétaux

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 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.

12- Table dispo_alim

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.

13- Table 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.

14- Table 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

15- Table sous_nutrition

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

16- Table sous_nutrition

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)

17- Table sous_nutrition

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 destinée 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)

		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

18- Table sous_nutrition

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

19-

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

20-

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

21- Table équilibre_prod

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 :

$$a1+a2+[\dots]=b1+b2+[\dots]=c1+c2+[\dots] \quad a1+a2+[\dots]=b1+b2+[\dots]=c1+c2+[\dots]$$

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.

1. 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.

1. La quantité totale (en kg) de produits perdus par pays et par année. La table renvoyée contiendra donc une ligne par couple (pays, année).

1. Les 10 pays pour lesquels la proportion de personnes sous-alimentées est la plus forte.

1. Les 10 produits pour lesquels le ratio Autres utilisations/Disponibilité intérieure est le plus élevé.