



**Food and Agriculture
Organization of the
United Nations**

P4 – Réalisez une étude de santé publique avec Python

30 août 2021

Rappel du contexte et présentation des étapes

CONTEXTE

Etude de santé publique sur le thème de la précarité alimentaire dans le monde.

Données à disposition : 4 fichiers (population, sous-nutrition, dispo alimentaire et aide alimentaire)
Lexique de données de Julien. Lien [FAO](#) sur les bilans alimentaires et les listes de produits.

Méthode choisie : langage Python + libraires Pandas, Numpy, Matplotlib et Seaborn pour l'exploration des données.
Jupyter Notebook

Cadre de contrainte : apporter réponses aux questions soulevées par Marc et Mélanie, en s'aidant des indications de Julien

ETAPES

Exploration préliminaires & préparation des données pour l'exploitation.

Traitement des données et visualisations pour apporter les réponses attendues.

Ce sont les résultats qui seront présentés ici.
Le code est disponible en intégralité sur le notebook si besoin.

Présentation des Résultats



QUESTIONS POSEES

La proportion de personnes en état de sous-nutrition dans le monde

Liste des pays à plus forte proportion de population sous-alimentée

CADRE

Pour l'année 2017

Remarque : données manquantes sur la population en sous-alimentation de 99 pays

Précarité Alimentaire : proportion population dans le monde

METHODE

Regrouper les données population et population en sous nutrition par pays et par année

Joindre les deux regroupements filtrés sur l'année 2017 pour ensuite avoir un tableau avec les sommes et faire le calcul attendu.

RESULTAT

7,1 % de la population mondiale est sous-alimentée en 2017

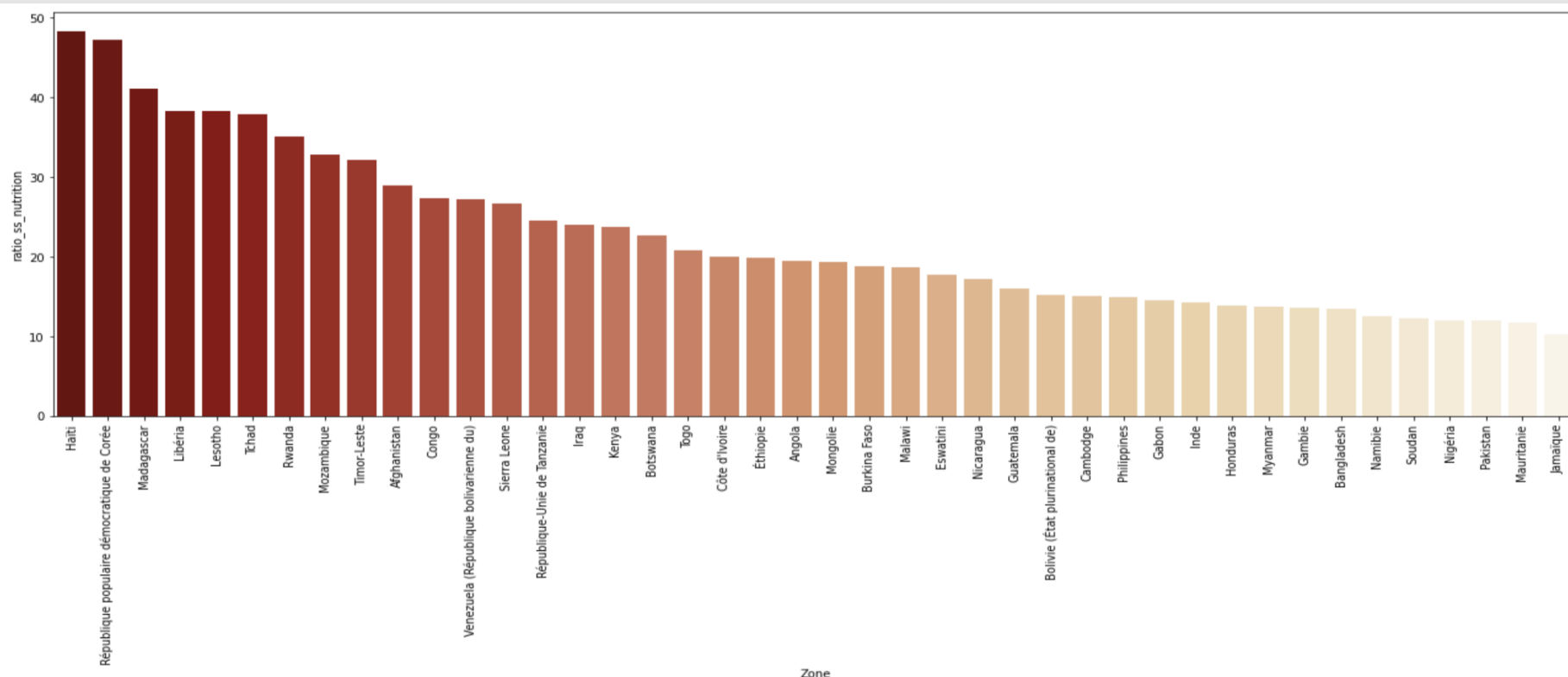
Base : population sous-alimentée de 535700000 pour population totale de 7548134111

Précarité Alimentaire : Classement des pays par proportion de population sous-alimentée

METHODE

Reprise du regroupement global en gardant le détail par pays, et calculer la proportion par pays

Résultat : Haïti, Rep. Pop. Démocratique de Corée, Madagascar, Libéria, Tchad, Rwanda, Mozambique, Timor-Leste, Afghanistan, Congo, Venezuela, Sierra Leone, Tanzanie, Iraq, Kenya, etc.



Nombre théorique de personnes pouvant être nourries en 2017

QUESTIONS POSEES

Nombre théorique de personnes nourries sur la base de la dispo alimentaire mondiale

Même question, sur la base de dispo alimentaire mondiale d'origine végétale

CADRE

Pour la population mondiale de 2017

Base de calcul: dispo alimentaire en kcal/jour/personne

Remarques :

données de dispo alimentaires du bilan alimentaire de 2013.
le besoin énergétique quotidien est estimé à 2500 kcal/jour/personne
(tiré de FAO : « La sous-alimentation dans le monde »).

Nombre théorique de personnes pouvant être nourries : sur la dispo alimentaire mondiale globale

APPROCHE

Obtenir ce nombre théorique via la population mondiale de 2017 et la disponibilité alimentaire globale en kcal/ jour

La dispo alimentaire est disponible en kcal/jour/ personne. Il faut donc regrouper les données par pays. Puis rapprocher avec les données de population de 2017 de chaque pays pour obtenir la dispo totale par pays en kcal/jour

Enfin, faire la somme des dispo par pays et faire le calcul final.

Le résultat sera donné en nombre de personnes ainsi qu'en pourcentage de la population mondiale de 2017.

RESULTAT

Pour un besoin de 2500 kcal/personne/ jour, 8367593851 de personnes pourraient être nourries.

Soit, pour une population mondiale de 7548134111 de personnes, un ratio de 110,86 %

Nombre théorique de personnes pouvant être nourries : sur la dispo alimentaire mondiale d'origine végétale.

APPROCHE

Même approche et traitement que pour la question sur la dispo alimentaire globale.

La manipulation préalable consiste à filtrer les données de dispo alimentaire sur l'origine végétale.

Regrouper et rapprocher les données filtrées, faire la somme des dispo par pays et faire le calcul final.

Le résultat sera donné en nombre de personnes ainsi qu'en pourcentage de la population mondiale de 2017.

RESULTAT

Pour un besoin de 2500 kcal/personne/ jour, 6904305685 de personnes pourraient être nourries.

Soit, pour une population mondiale de 7548134111 de personnes, un ratio de 91.47%

Utilisation de la disponibilité intérieure

QUESTIONS POSEES

Utilisation de la disponibilité intérieure mondiale, avec focus sur la part de l'alimentation animale, la part de l'alimentation humaine et les pertes.

Sur indication de Julien, un focus sera fait sur cette question cette fois sur la disponibilité alimentaire mondiale en céréales.

Un classement des pays en fonction de la disponibilité alimentaire par habitant sera établie.

Sur indication de Julien, un focus sera fait sur l'utilisation du manioc.

CADRE

Le lexique de données de Julien, ainsi que la notice des bilans alimentaires FAO permettent de reconstituer :
Dispo alimentaire = nourriture + aliments pour animaux + pertes + semences + traitement + autres utilisations.

La notice des bilans alimentaires donne la liste de céréales suivante : Blé, maïs, mil, orge, riz (paddy), seigle, avoine, sarrasin, sorgho, quinoa, autres céréales, mélanges de céréales.

Nous verrons que cette liste n'est pas conforme à celle présente dans les données à disposition, feront les ajustements nécessaires et un point sur les conséquences éventuelles sur les chiffres obtenus.

Utilisation de la disponibilité intérieure globale

APPROCHE

Une vérification est faite sur la formule qui permet de retrouver la dispo intérieure à partir des utilisations. Les écarts constatés sont peu significatifs en regard des volumes considérés, ce qui valide la formule.

Les données de dispo alimentaire sont regroupées par pays, et filtrées sur les utilisations

Les utilisations sont sommées et le calcul du pourcentage de chaque utilisation par rapport au total de dispo intérieure est calculé

ratio_utilisation

Nourriture	49.46%
alim_animaux	13.23%
autres_util	8.77%
Pertes	4.60%
Semences	1.57%
Traitement	22.36%

RESULTATS

La nourriture représente quasiment la moitié de l'utilisation de la dispo intérieure, versus 13 % pour l'alimentation pour animaux.

Les pertes en représentent presque 5 %

A noter les 22 % destinés à la transformation (nourriture humaine, pour animaux ou à autre fin qu'alimentation)

Utilisation de la disponibilité intérieure : focus sur les céréales

APPROCHE

Même approche générale que pour la dispo globale.

Les données de dispo alimentaire sont regroupées par pays, filtrées sur les utilisations et également sur les produits qui correspondent à la liste de céréales

Les utilisations sont sommées et le calcul du pourcentage de chaque utilisation par rapport au total de dispo intérieure en céréales est calculé

ratio_util_cereales	
Nourriture	42.74%
alim_animaux	36.28%
Pertes	4.45%
autres_util	9.75%
Semences	2.85%
Traitement	3.93%

RESULTATS

L'alimentation humaine représente presque 43 % de l'utilisation des céréales, < ratio de la dispo intérieure globale.

L'utilisation pour l'alimentation animale est en revanche plus importante, avec 36 % céréales vs 13 % dispo globale.

Les pertes sont quasi équivalentes.

Utilisation des céréales : point d'attention sur la constitution de la liste

CONTEXTE

Lors de mes recherches sur les céréales, en particulier sur FAOSTAT, j'ai constaté que dans les bilans alimentaires disponibles à la visualisation et à l'extraction, les céréales n'ont pas forcément les mêmes libellés que dans nos données à deux exceptions près : l'avoine et les « céréales, autres ».

Les autres libellés semblent avoir été modifiés dans nos données par rapport au fichier extrait de FAOSTAT, en supprimant la mention « et produits ».

La vigilance est de mise, car si le filtre sur les produits est établi dans nos données à disposition, sur la base des libellés des données brutes extraites de FAOSTAT, seules les données de « Avoine » et « Céréales, autres » remonteront.

Je propose ici de faire les calculs avec cette liste tronquée, pour illustrer l'impact très significatif sur les chiffres obtenus.

ratio_util_cereales_tronque	
Nourriture	16.68%
alim_animaux	69.46%
Pertes	2.20%
autres_util	0.88%
Semences	10.62%
Traitement	0.15%

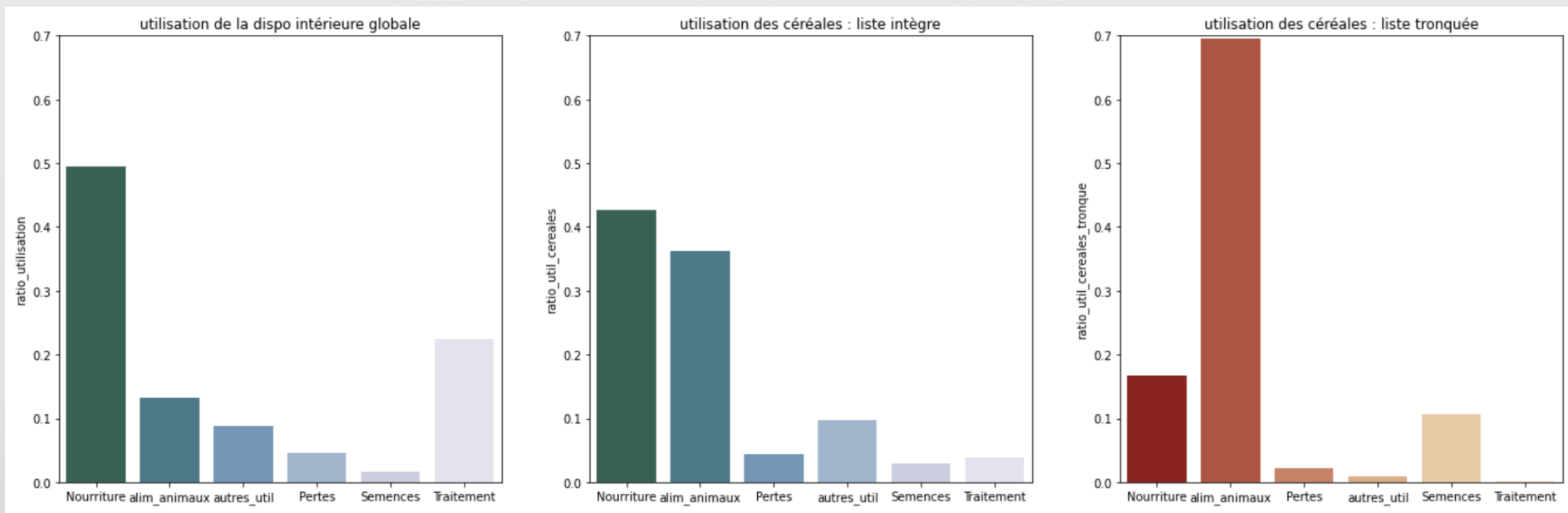
RESULTATS

L'alimentation humaine ne représente plus que 17 % de l'utilisation des céréales.

L'utilisation pour l'alimentation animale monte à 69 % de l'utilisation des céréales, ce qui change radicalement le poids vs l'alimentation humaine.

Cela peut mener à des interprétations et conclusions erronées !

Utilisation de la dispo intérieure : visualisation des résultats



CONCLUSIONS

Il apparaît que les céréales sont utilisées en grande proportion (36%) pour l'alimentation animale

L'énoncé de l'indication de Julien portait bien sur la liste globale des céréales.

Toutefois la recherche permet d'observer que l'avoine et les « céréales autres » sont utilisées à 69 % pour l'alimentation animale, vs 16.7 % pour la nourriture humaine.

disponibilité intérieure : classement des pays selon la disponibilité alimentaire par habitant.

APPROCHE 1 : croisement avec les données de précarité alimentaire

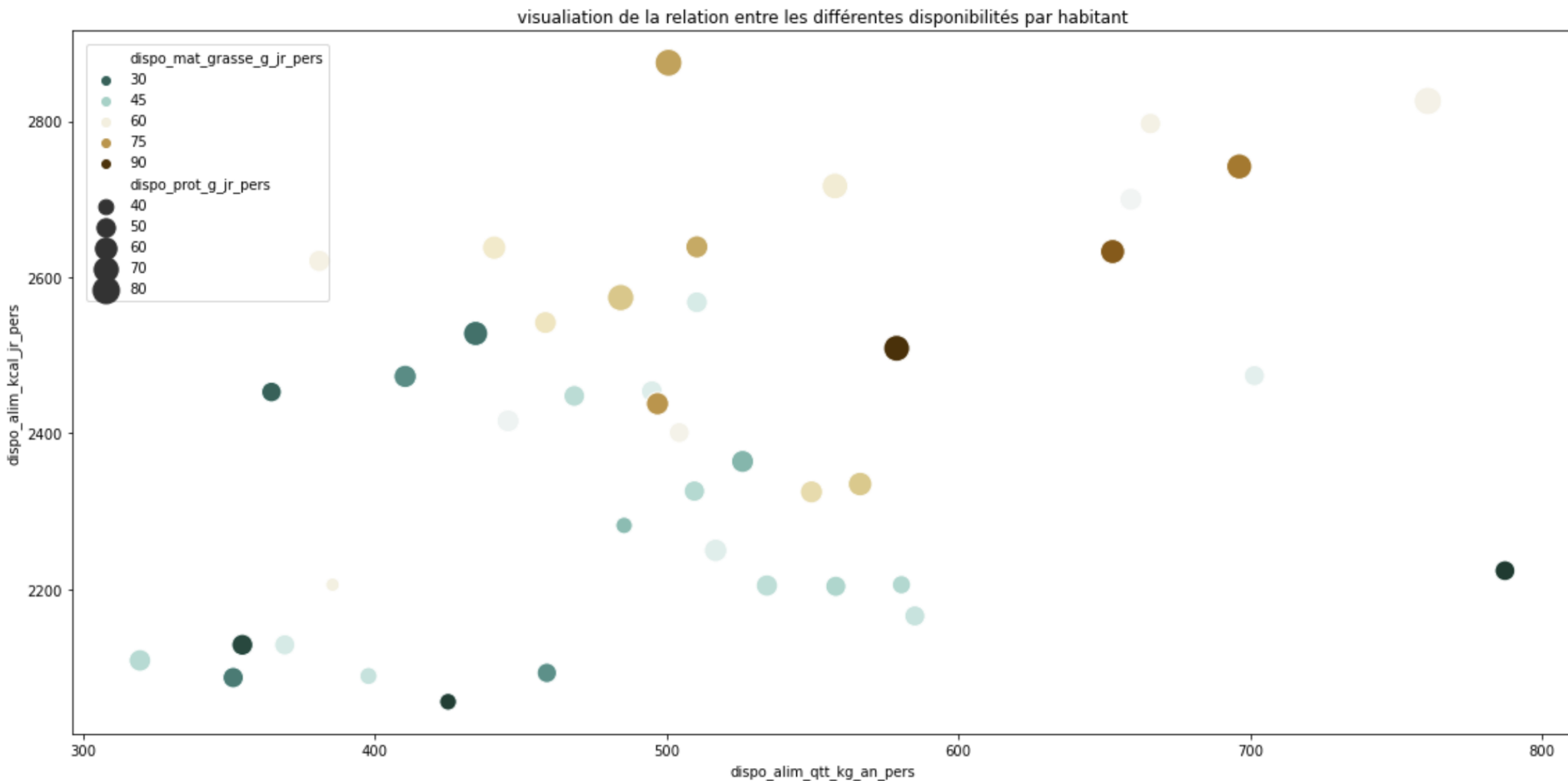
Grouper les données de dispo alimentaire par pays, en ne conservant que les 4 types de dispo: en kcal / pers /an, en kg/ pers /an, en g de protéines / pers / jour, en g de lipides/pers / jour.

Joindre ces données groupées avec les données de précarité alimentaire par pays, puis visualiser les données de dispo pour déterminer si une relation nette existe entre elles.

Choix d'un type de dispo pour classer les pays en fonction de sa quantité.

Zone	ratio_ss_nutrition	dispo_alim_kcal_jr_pers	dispo_alim_qtt_kg_an_pers	dispo_mat_grasse_g_jr_pers	dispo_prot_g_jr_pers
Haïti	48.259182	2089.0	397.78	48.92	47.70
République populaire démocratique de Corée	47.188685	2093.0	458.93	36.41	54.99
Madagascar	41.062924	2056.0	425.07	24.49	46.69
Libéria	38.279742	2206.0	385.48	59.85	37.66
Lesotho	38.249438	2528.0	434.51	32.26	73.48
Tchad	37.957606	2109.0	319.45	47.10	62.75
Rwanda	35.055619	2224.0	787.35	23.77	56.79
Mozambique	32.810898	2282.0	485.40	41.90	45.68
Timor-Leste	32.173531	2129.0	369.10	51.08	57.51
Afghanistan	28.928718	2087.0	351.41	33.50	58.26
Congo	27.393535	2206.0	580.47	46.62	51.41
Venezuela (République bolivarienne du)	27.208586	2633.0	652.86	82.44	72.96
Sierra Leone	26.707893	2401.0	504.31	58.81	56.11
République-Unie de Tanzanie	24.515033	2204.0	557.97	46.12	57.76
Iraq	23.966268	2542.0	458.43	64.17	63.60

disponibilité intérieure : classement des pays en précarité alimentaire selon la disponibilité alimentaire par habitant.



disponibilité intérieure : classement des pays en précarité alimentaire selon la disponibilité alimentaire par habitant.

RESULTATS APPROCHE 1

Il apparait une relative corrélation positive entre quantité en poids par personne et dispo en kcal/pers.
Moins nette, entre la dispo en kg et la dispo en lipides.
C'est plus difficile de statuer sur une éventuelle corrélation entre dispo en kg et dispo en g de protéines.

Ces observations, ramenées à la définition de la dispo en kg/pers/an de la notice FAO des bilans alimentaires, permettent de sélectionner la dispo en kg/pers/ an comme indicateur pertinent pour obtenir le classement des pays attendu.

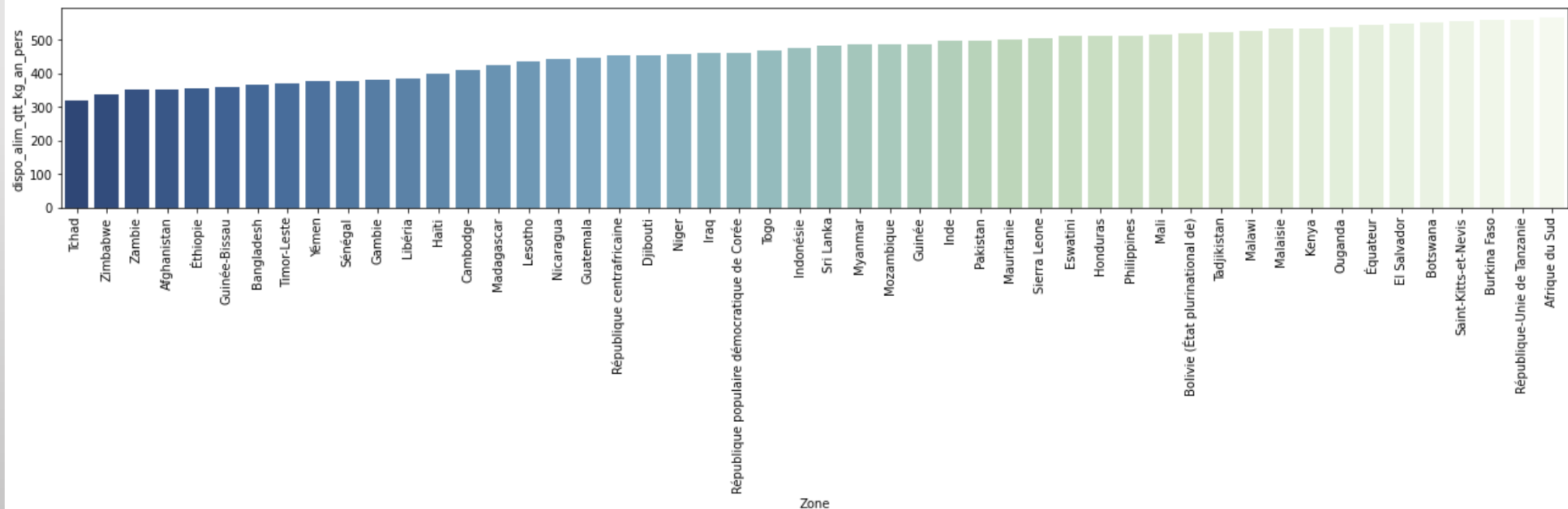
C'est l'objet de la deuxième approche proposée.

disponibilité intérieure : classement des pays en précarité alimentaire selon la disponibilité alimentaire par habitant.

APPROCHE 2

On écarte le rapprochement avec les données de précarité alimentaire. En effet, les pays pour lesquels nous n'avons cette donnée sont de facto exclus du classement.

Le focus est fait sur la dispo alimentaire en kg / pers / an, et une visualisation de la donnée est proposée pour afficher le classement.



disponibilité intérieure : classement des pays en précarité alimentaire selon la disponibilité alimentaire par habitant.

RESULTATS APPROCHE 2

Classement : Tchad, Zimbabwe, Zambie, Afghanistan, Ethiopie, Guinée-Bissau, Bangladesh, Timor-Leste, Yémen, Sénégal, Gambie, Liberia, Haïti, Cambodge, Madagascar, Lesotho, Nicaragua, Guatemala, Rep. Centrafricaine, etc.

On peut mettre en regard ce classement avec celui obtenu sur la base de la proportion de population en sous-nutrition :

Rappel : Haïti, Rep. Pop. Démocratique de Corée, Madagascar, Libéria, Tchad, Rwanda, Mozambique, Timor-Leste, Afghanistan, Congo, Venezuela, Sierra Leone, Tanzanie, Iraq, Kenya, etc.

CONSTATS

L'objet n'est pas ici de chercher une corrélation entre les deux données, mais à la fois de voir les pays communs aux deux listes, et d'inclure des pays absents du classement selon la prop. de population en sous-nutrition.

Ces deux facteurs s'aggravent mutuellement.

Aide Alimentaire : classement des pays en fonction de la quantité reçue depuis 2013

APPROCHE

Les données de l'aide alimentaire sont regroupées pays, et sommées pour avoir le cumul.
A noter que pour certains pays, la donnée n'est pas forcément disponible pour chaque année de la période observée.

Ces données sont rapprochées des données de populations en sous-nutrition et de disponibilités alimentaires par habitant

Aide Alimentaire : classement des pays en fonction de la quantité reçue depuis 2013

Zone	population	ratio_ss_nutrition	aide_alim_T	dispo_alim_kcal_jr_pers	dispo_alim_qtt_kg_an_pers	dispo_mat_grasse_g_jr_pers	dispo_prot_g_jr_pers
Éthiopie	106399924.0	19.830841	1381294.0	2129.0	354.55	25.89	60.39
Yémen	27834819.0	NaN	1206484.0	2217.0	375.41	47.81	58.90
Soudan	40813397.0	12.250879	669784.0	2335.0	566.31	69.27	70.69
Kenya	50221142.0	23.695200	552836.0	2205.0	534.36	47.96	61.87
Bangladesh	159685424.0	13.463972	348188.0	2453.0	364.54	29.92	55.86
Niger	21602382.0	NaN	276344.0	2549.0	456.15	52.53	80.93
Tchad	15016753.0	37.957606	267966.0	2109.0	319.45	47.10	62.75
Pakistan	207906209.0	11.928456	231072.0	2438.0	496.85	74.88	65.45
République populaire démocratique de Corée	25429825.0	47.188685	187412.0	2093.0	458.93	36.41	54.99
Zambie	16853599	nan	3026	1924.0	350.12	42.05	55.22
Tadjikistan	8880268	nan	2028	2199.0	520.41	60.38	63.38
Sao Tomé-et- Principe	207089	nan	1826	2398.0	663.93	77.67	53.10
Équateur	16785361	8.936358	1362	2346.0	544.91	94.64	61.30
Chine, continentale	1421021791	nan	1298	3112.0	901.90	95.09	98.05
Égypte	96442591	4.769677	1122	3518.0	803.79	63.09	103.22
Vanuatu	285510	nan	802	2833.0	740.96	104.98	69.69
Timor-Leste	1243258	32.173531	116	2129.0	369.10	51.08	57.51
Géorgie	4008716	7.483693	70	2902.0	649.44	64.53	80.50
Bolivia (État plurinational de)	11192855	15.188261	6	2250.0	516.79	53.07	65.96

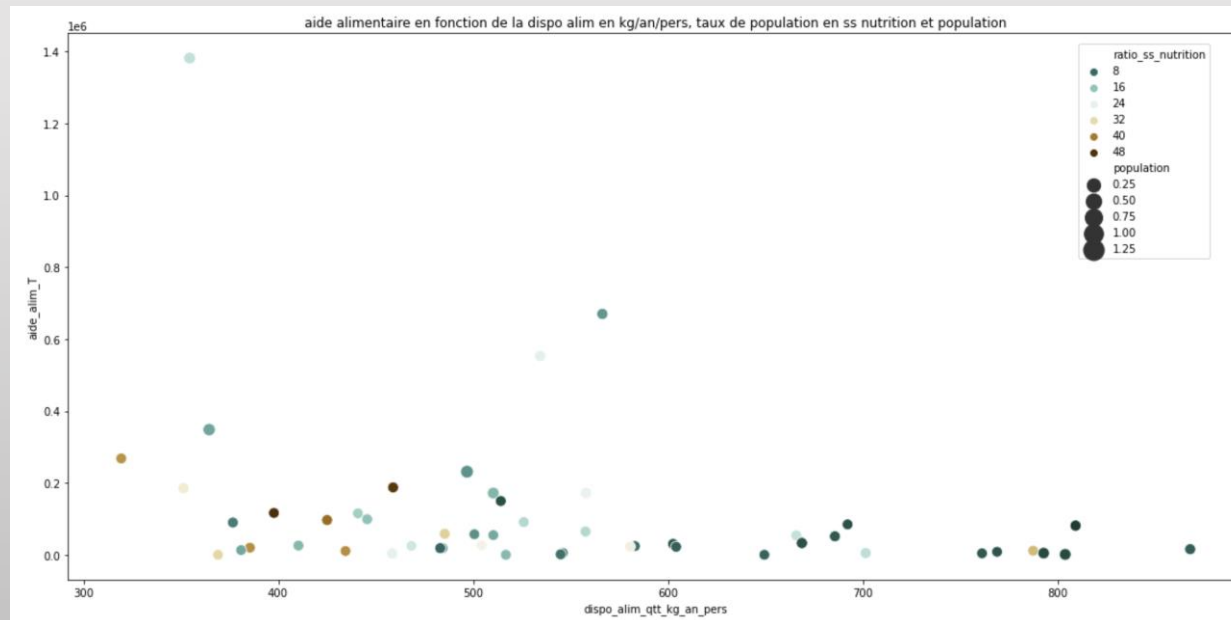
Aide Alimentaire : classement des pays en fonction de la quantité reçue depuis 2013

OBSERVATIONS

Les extraits du dataframe obtenu, qui présente les pays ayant reçu le plus / le moins d'aide alimentaire depuis 2013 sont donnés slide 21(précédent)

On voit que la quantité d'aide reçue n'est pas forcément liée à la proportion de population en sous-nutrition, ni aux disponibilités alimentaires par habitant même si l'aide semble plus fréquente quand la dispo alim en kg/an/pers est < 600 kg/an/pers

Ici une visualisation des données du dataframe, qui confirme qu'une partie des critères qui déterminent l'aide alimentaire sont extérieurs aux données à notre disposition, en gardant à l'esprit que la donnée d'aide reçue est parfois manquante.



Utilisation de la disponibilité intérieure : zoom sur la production et les exportations de manioc

APPROCHE : partir d'une analyse plus large pour retrouver les résultats sur la Thaïlande

Dans son lexique de données Julien indique un point d'intérêt sur l'utilisation du manioc en Thaïlande, en particulier concernant le ratio des exportations sur la production de ce produit.

Dans un premier temps, il faut regrouper par pays le dataframe de disponibilité alimentaire filtré sur le produit manioc. La proposition est de créer un subset avec en particulier les colonnes : production, dispo intérieure, exportations, nourriture, aliments pour animaux, importations.

Les ratio d'intérêt sont calculés : exportation et importation / production, et nourriture et aliments pour animaux / disponibilité intérieure. On rapproche ces données avec les données des pays en sous-nutrition. Le tableau final obtenu est présenté slide suivant (25)

Utilisation de la disponibilité intérieure : zoom sur la production et les exportations de manioc

Zone	ratio_ss_nutrition	ratio_exp_prod	ratio_imp_prod	ratio_alim_anim_dispo	ratio_nourr_dispo
Viet Nam	6.870989	91.96%	0.00%	36.20%	38.49%
Costa Rica	4.040442	84.83%	2.76%	0.00%	68.75%
Thaïlande	8.958268	83.41%	4.14%	28.74%	13.90%
Mexique	6.731992	20.83%	20.83%	0.00%	66.67%
Équateur	8.936358	17.39%	13.04%	0.00%	90.91%
Honduras	13.787233	8.33%	4.17%	0.00%	95.65%
Paraguay	8.737362	5.86%	0.00%	52.70%	32.17%
Nicaragua	17.228293	5.78%	0.72%	48.85%	25.19%
Panama	7.305013	4.55%	9.09%	13.04%	73.91%
Sri Lanka	7.572878	3.96%	7.92%	24.13%	65.71%
Indonésie	8.917406	3.43%	4.60%	1.98%	48.45%
Cambodge	14.991184	3.17%	0.00%	0.00%	4.94%
Pérou	6.996499	0.59%	0.34%	15.06%	73.69%
Inde	14.200590	0.29%	1.09%	0.00%	95.04%
République-Unie de Tanzanie	24.515033	0.19%	0.00%	13.49%	75.79%
Philippines	14.927796	0.17%	11.69%	5.40%	86.97%
Rwanda	35.055619	0.14%	3.19%	0.00%	92.69%
Colombie	5.520362	0.09%	1.08%	2.66%	82.32%
Ghana	6.867786	0.06%	0.01%	25.15%	35.02%
Bénin	7.158710	0.05%	0.00%	32.76%	31.21%
Cameroun	6.513047	0.02%	0.00%	20.67%	47.32%
Nigéria	11.945100	0.02%	0.03%	49.60%	38.93%

Utilisation de la disponibilité intérieure : zoom sur la production et les exportations de manioc

RESULTATS

La Thaïlande, le Viet Nam et le Costa Rica exportent plus de 83 % de leur production de Manioc, avec des ratio de population en précarité alimentaire de 8,9 %, 6,8 % et 4 % respectivement.

A noter également le Mexique et l'Equateur qui en regard de leur taux de population en précarité alimentaire exportent resp. 20,8 % et 17,4 % de leur production de Manioc

il est également à noter qu'en plus des exportations, quelques pays attribuent une part significative de la production de manioc à l'alimentation pour les animaux, voir importent également du Manioc

A ce titre, la Thaïlande cumule les paradoxes : avec un taux de population en précarité alimentaire de presque 9%, exporte 83,4% de sa production de Manioc, en importe, et utilise une part plus significative de sa disponibilité intérieure de ce produit pour l'alimentation pour animaux (28,7%) que pour la nourriture humaine (13,9%).

En conclusion : pays les plus en difficulté au niveau alimentaire.

QUESTION POSEE

L'objet est de mettre en relief les pays les plus en difficulté au niveau alimentaire.

Les critères sont à trouver dans les résultats obtenus tout au long de l'exploration des données.

APPROCHE

Etablir des critères sur les indicateurs mis en place : aide alimentaire cumulée depuis 2013, population la plus touchée par la sous-nutrition en 2017, pays pour lesquels la disponibilité alimentaire est faible.

Croiser entre eux les trois tableaux de données correspondant, en appliquant les critères choisis pour filtrer les pays qui semblent le plus en difficulté.

En conclusion : pays les plus en difficulté au niveau alimentaire.

LES TROIS CRITERES

La méthode choisie (à partir des tables des quartiles) permet de faire varier de façon simple les valeurs jugées pertinentes pour ces critères

Aide alimentaire : il a été vu que cette donnée n'est pas forcément reliée à la précarité alimentaire du pays : on va considérer les 75 % de pays dont l'aide est > au 1^{er} quartile de la donnée (9727,5 Tonnes)

Ratio de population en sous-nutrition : ici, il s'agit de se concentrer sur les 50 % de pays au-dessus de la médiane de la donnée (9,9 %)

Disponibilité alimentaire en kg/an/pers : il a été vu que l'aide reçue est globalement plus importante pour les 25 % des pays dans le quartile le plus bas de la donnée de disponibilité (< 545 kg/an/pers)

En conclusion : pays les plus en difficulté au niveau alimentaire.

APPROCHE 1

On met en place un critère de condition strict sur les trois critères (« & ») dans la sélection finale des pays.

RESULTATS APPROCHE 1

Ethiopie, Kenya, Bangladesh, Tchad, Pakistan, Rep. Dem. De Corée, Afghanistan, Philippines, Haïti, Nicaragua, Guatemala, Madagascar, Malawi, etc.

Cette approche donne une première liste de pays. Elle est problématique car comme nous l'avons vu lors de l'exploration, un critère strict de sélection sur la population en sous-nutrition exclut mécaniquement les pays en difficulté pour lesquels nous n'avons pas cette donnée à disposition.

L'objet d'une deuxième approche est d'apporter une solution à cette exclusion.

En conclusion : pays les plus en difficulté au niveau alimentaire.

APPROCHE 2

Le critère strict est conservé pour la sélection du groupe aide alimentaire & disponibilité en kg/an/personne

En revanche, la sélection entre ce groupe de critères, et le ratio de population en sous-nutrition est assoupli (« | ») pour inclure dans la liste les pays pour lesquels nous n'avons pas la donnée

RESULTATS APPROCHE 2

Ethiopie, Yemen, Soudan, Kenya, Bangladesh, Niger, Tchad, Pakistan, Rep. Dem. De Corée, Afghanistan, Tanzanie, Philippines, Mali, Haïti, Nicaragua, Guatemala, Madagascar, Malawi, Senegal, République Centrafricaine, Burkina Fasso, etc.

L'association des critères est un peu moins rigoureuse, mais donne une liste étendue avec des pays comme le Yemen et le Soudan par exemple, et correspond mieux à la réalité.

Nous retiendrons donc cette deuxième approche. Il semble plus difficile de limiter cette liste aux premiers pays pour conclure, tant les situations de chaque pays est spécifique et parfois aggravée par des facteurs extérieurs aux données à notre disposition : guerres, catastrophes naturelles (tremblements de terre, typhons, destruction de cultures par les invasions d'insects), changements climatiques, etc.

MERCI

Laurent Carn

carnlaurent@gmail.com



**Food and Agriculture
Organization of the
United Nations**