

Projet 3 : Réalisez une étude de santé publique

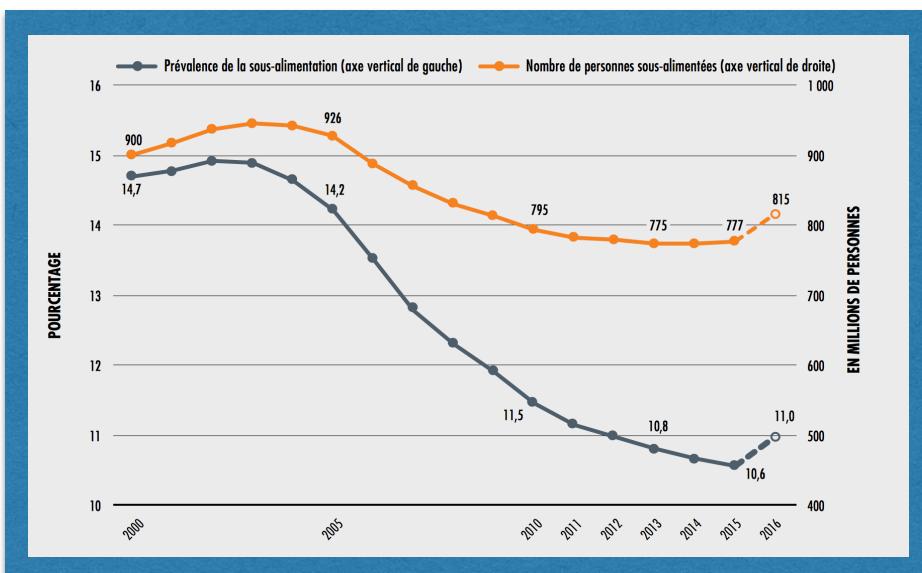
Année 2018-2019



СОВЕРШЕННО
СМОКАНИЕ

I-Mise en contexte

Nombre de personnes
qui décèdent à cause de la faim
en 2015 :
9,1 millions de personnes



Sous nutrition mondiale en 2017:
815 millions de personnes

Evolution possible : Sous-nutrition

2015

777 millions

2017

815 millions

Donnez le résultat de votre calcul pour l'année 2013, ainsi que pour la dernière année disponible au jour où vous effectuez ce projet. ?

```
population_2013=pd.read_csv("population_2013.csv")
Total_pop_2013 = (population_2013['Valeur'].sum())*1000#On réalise la somme de la colonne Valeur
population_2017=pd.read_csv("population_2017.csv")#Lecture du fichier
Total_pop_2017 = (population_2017['Valeur'].sum())*1000#On additionne la colonne Valeur
print ("Population totale en 2013 =",Total_pop_2013)
print( "Population totale en 2017=",Total_pop_2017)
```

```
Population totale en 2013 = 7213211612.0
```

```
Population totale en 2017= 7550036043.999999
```

Il y aurait pu avoir une anomalie si je n'avais pas exclu la Chine dès le départ dans mon jeu de données

Les causes de la faim



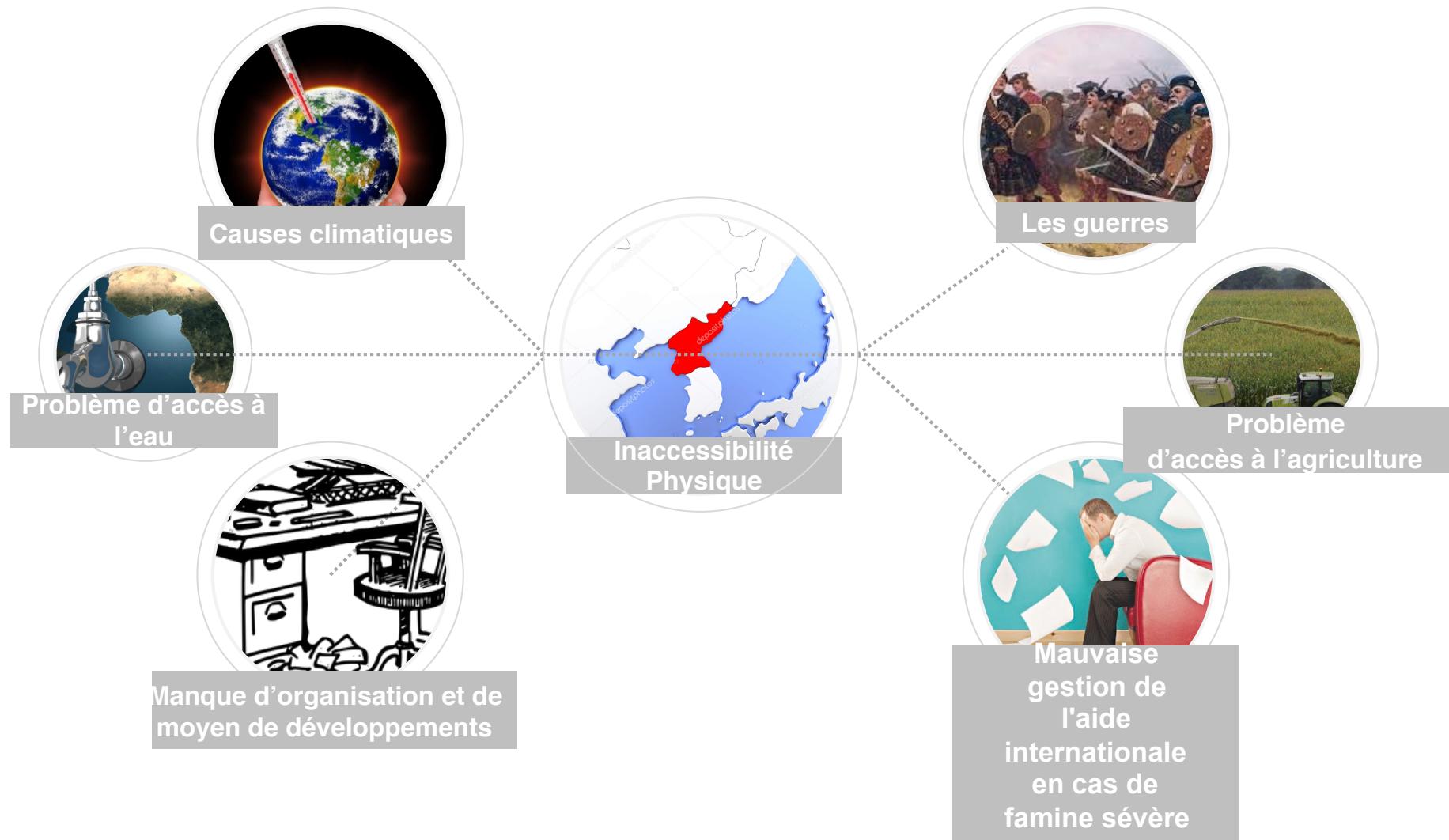
Inaccessibilité économique



Inaccessibilité physique



Inaccessibilité Physique



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 ?

0	Afghanistan	8.1
1	Afrique du Sud	2.6
2	Albanie	0.2
3	Algérie	1.8
4	Allemagne	<2

```
: sous_nutrition['Valeur'] = pd.to_numeric(sous_nutrition['Valeur'], errors='coerce')
#fill
```



```
: sous_nutrition=sous_nutrition.fillna(0)
sous_nutrition=sous_nutrition.replace([np.inf, -np.inf], 0)#np.inf = plus infini -np.inf=moins infini
sous_nutrition.head()
#Regarder si la chine
```

:

	Zone	Valeur
0	Afghanistan	8.1
1	Afrique du Sud	2.6
2	Albanie	0.2
3	Algérie	1.8
4	Allemagne	0.0

```
: Sous_nutrition=((sous_nutrition['Valeur'].sum())*(1000000))
print((Sous_nutrition/Total_pop_2013)*100)
```

12.158245829652502

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.

```
print ("ratio total ",(BA_veget["DI en Kcal"].sum()/(2350*365))/(Total_pop_2013))#calorie par jour 2350
print ("ratio total ",(BA_veget["DI en Kcal"].sum()*(100)/(2350*365))/(Total_pop_2013))#Donne un ratio en %,
```

```
ratio total  1.9985658346615374e-09
ratio total  1.9985658346615374e-07
```

```
BA_veget[ 'DI en kg prot']=BA_veget[ 'Disponibilité intérieure ']*BA_veget[ 'Ratio Proteine '#0,83 g/kg/jour
print("Pourcentage de la population avec la DI en kg prot",BA_veget["DI en kg prot"].sum()/((0.83*75/1000)*365)/(Total_
#75 est le poid moyen d'une personne
```

```
Pourcentage de la population avec la DI en kg prot 1.7770904301137787e-09
```

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.

```
BA_veget["q8"] = BA_veget["Aliments pour animaux"] + BA_veget["Nourriture"] + BA_veget["Pertes"]
```

```
BA_veget.head()
```

Ité de rassemblement /jour	Disponibilité de protéines en quantité (g/personne/jour)	Disponibilité intérieure ...	Pertes	Production	Semences	Traitement	Variation de stock	Ratio Energie	Ratio Protéine	DI en Kcal	DI en kg prot	q8
0.01	0.02	41000000.0	...	2000000.0	3000000.0	0.0	0.0	0.0	2.969761e-07	5.939521e-12	12.176018	0.000244
0.01	0.03	2000000.0	...	0.0	0.0	0.0	0.0	0.0	5.791033e-06	1.737310e-10	11.582066	0.000347
0.00	0.00	0.0	...	0.0	0.0	0.0	0.0	0.0	0.000000e+00	0.000000e+00	0.000000	0.000000
0.02	0.05	82000000.0	...	0.0	0.0	0.0	0.0	0.0	5.649788e-07	7.062235e-12	46.328264	0.000579
0.00	0.00	3000000.0	...	0.0	0.0	0.0	0.0	0.0	0.000000e+00	0.000000e+00	0.000000	3000000.0

```
BA_veget['DA_q8 en kg prot']=BA_veget["q8"]*BA_veget['Ratio Proteine']
print("Pourcentage de la population avec la DA en kg prot",BA_veget["DA_q8 en kg prot"].sum() / ((0.83*75/1000)*365)/(Total_pop_2010))
```

Pourcentage de la population avec la DA en kg prot 1.248546464938344e-09

```
BA_veget['DA_q8 en kg cal']=BA_veget["q8"]*BA_veget['Ratio Energie']
print("Pourcentage de la population avec la DA en kg cal",BA_veget["DA_q8 en kg cal"].sum() / ((2350)*365)/(Total_pop_2010))
```

Pourcentage de la population avec la DA en kg cal 1.4904554398782e-09

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 ?

```
cereale=pd.read_csv("cereale.csv")
Lst_cereale=cereale[ 'Produit' ].unique()#liste des cereales uniquement
```

```
df7=Bilan_alim.loc[Bilan_alim[ 'Produit' ].isin(Lst_cereale), : ]
```

```
df7.head()
```

Pays	Produit	Année	Aliments pour animaux	Autres Utilisations	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité alimentaire en quantité (kg/personne/an)	Disponibilité de matière grasse en quantité (g/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Disponibilité intérieure ...
7	Afghanistan	Blé	2013	0.000000e+00	0.0	1369.0	160.23	4.69	36.91 5.992000e+15 ... 7.750
12	Afghanistan	Céréales, Autres	2013	0.000000e+00	0.0	0.0	0.00	0.00	0.00 0.000000e+00 ... 0.000
32	Afghanistan	Maïs	2013	2.000000e+14	0.0	21.0	2.50	0.30	0.56 3.130000e+14 ... 3.100
34	Afghanistan	Millet	2013	0.000000e+00	0.0	3.0	0.40	0.02	0.08 1.300000e+13 ... 1.000
40	Afghanistan	Orge	2013	3.600000e+14	0.0	26.0	2.92	0.24	0.79 5.240000e+14 ... 5.200

5 rows × 23 columns

```
df7[ 'Aliments pour animaux' ].sum()/(df7[ 'Aliments pour animaux' ].sum())+(df7[ 'Nourriture' ].sum()))
```

0.45914025686646043

En Thaïlande, quel proportion de manioc est exportée ? Quelle est la proportion de personne en sous-nutrition

In [83]:

```
export_manioc=print(BA_veget.loc[(BA_veget['Pays']=='Thaïlande') & (BA_veget['Produit'] == "Manioc"),['Exportations - Quantité']] )
```

```
Exportations - Quantité      0.539903
dtype: float64
```

In [84]:

```
sous_nutr_thailande=sous_nutrition.loc[sous_nutrition['Zone']=='Thaïlande',['Valeur']].sum()*1000000
population_Thailande=population_2013.loc[population_2013['Zone']=='Thaïlande',['Valeur']].sum()*1000
(sous_nutr_thailande/population_Thailande)*100 #recuperer population thailande
```

Out[84]: Valeur 8.951755

```
dtype: float64
```

La faim dans le monde résulte t'elle d'un manque de production ?

NON

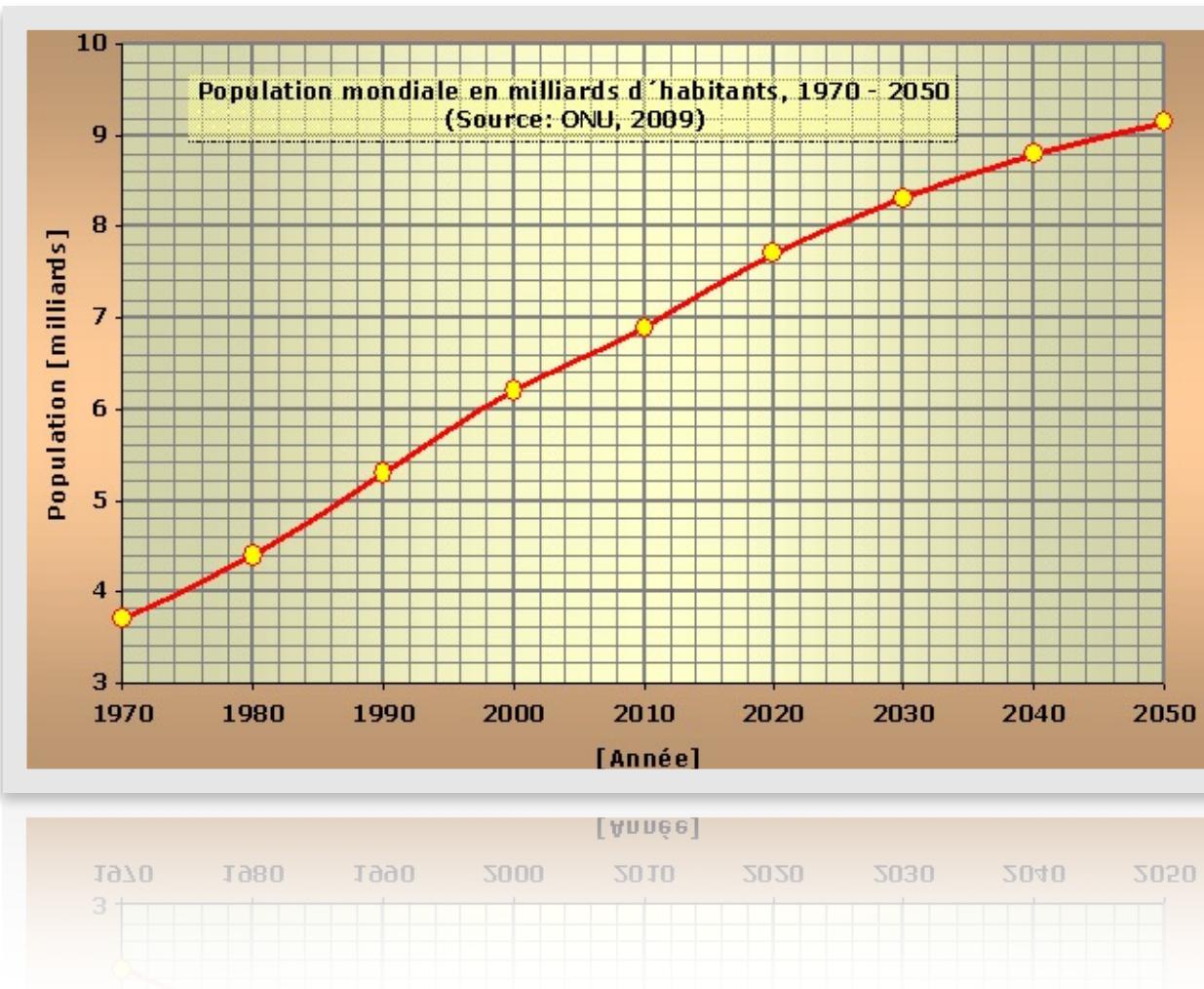


La faim dans le monde résulte d'elle de problèmes technologiques

Oui



La population en 2050



2050 : 9,8 milliards d'habitants par l'ONU en 2015

Augmentation drastique de la production alimentaire

III-Le détail des sources

j'ai utilisé 6 jeux de données

Nom du pays,
population totale, année

Population 2013

Population 2017

Produit	Année	Aliments pour animaux	Autres Utilisations	Disponibilité alimentaire (Kcal/personne/jour)	autre
					(kg/personne/jour)
Abats Comestible	2013	0.0	0.0	5.0	
Agrumes, Autres	2013	0.0	0.0	1.0	
Aliments pour enfants	2013	0.0	0.0	1.0	
Ananas	2013	0.0	0.0	0.0	
Bananes	2013	0.0	0.0	4.0	

Autres utilisations	Disponibilité alimentaire (Kcal/personne/jour)	Disponibilité alimentaire en quantité (kg/personne/an)	Demande (g/personne/jour)
0.0	1369.0	160.23	
0.0	0.0	0.00	
0.0	21.0	2.50	
0.0	3.0	0.40	
0.0	26.0	2.92	

Sous nutrition

Zone	Code Élément	Élément	Code Produit	Produit	Code année	Année	Unité
Afghanistan	511	Population totale	3010	Population-Estimations	2013	2013	1000 personnes
Afrique du Sud	511	Population totale	3010	Population-Estimations	2013	2013	1000 personnes
Albanie	511	Population totale	3010	Population-Estimations	2013	2013	1000 personnes
Algérie	511	Population totale	3010	Population-Estimations	2013	2013	1000 personnes
Allemagne	511	Population totale	3010	Population-Estimations	2013	2013	1000 personnes

Bilan alimentaire

Produit	Année	Aliments pour animaux	Autres Utilisations (Kg)
Agrumes, Autres	2013	0.0	0.0
Aliments pour enfants	2013	0.0	0.0
Ananas	2013	0.0	0.0
Bananes	2013	0.0	0.0
Bière	2013	0.0	0.0

Cereale

	Zone	Valeur
0	Afghanistan	8.1
1	Afrique du Sud	2.6
2	Albanie	0.2
3	Algérie	1.8
4	Allemagne	0.0

Pays, élément,
produit, année, valeur

Les tables SQL

	pays	code_pa	annee	produit	code_pro	dispo_int	alim_ani	semenes	pertes	transfo
111	Afrique du Sud	202	2013	Manioc	2532	570000000.0	0.0	0.0	0.0	0.0
112	Afrique du Sud	202	2013	Maïs	2514	10480000000.0	4715000000.0	33000000.0	405000000.0	43000000.0
113	Afrique du Sud	202	2013	Miel	2745	3000000.0	0.0	0.0	0.0	0.0
114	Afrique du Sud	202	2013	Millet	2517	13000000.0	8000000.0	0.0	0.0	0.0
115	Afrique du Sud	202	2013	Miscellanées	2899	0.0	0.0	0.0	0.0	0.0
116	Afrique du Sud	202	2013	Mollusques, Autres	2767	7000000.0	0.0	0.0	0.0	0.0
117	Afrique du Sud	202	2013	Noix	2551	17000000.0	0.0	0.0	0.0	0.0
118	Afrique du Sud	202	2013	Oeufs	2744	497000000.0	0.0	67000000.0	51000000.0	0.0
119	Afrique du Sud	202	2013	Oignons	2602	501000000.0	0.0	0.0	59000000.0	0.0
120	Afrique du Sud	202	2013	Olives	2563	1000000.0	0.0	0.0	0.0	0.0
121	Afrique du Sud	202	2013	Oranges, Mandarines	2611	422000000.0	0.0	0.0	36000000.0	0.0

equilibre_prod

dispo_alim

	pays	code_pays	annee	population
1	Afghanistan	2	2012	30696958.0
2	Afghanistan	2	2013	31731688.0
3	Afghanistan	2	2014	32758020.0
4	Afghanistan	2	2015	33736494.0
5	Afghanistan	2	2016	34656032.0
6	Afghanistan	2	2017	35530081.0
7	Afrique du Sud	202	2012	52998212.99999999
8	Afrique du Sud	202	2013	53767395.99999999
9	Afrique du Sud	202	2014	54539571.0
10	Afrique du Sud	202	2015	55291225.0
11	Afrique du Sud	202	2016	56015473.0
12	Afrique du Sud	202	2017	56717155.99999999
13	Albanie	3	2012	2920039.0
14	Albanie	3	2013	2918978.0

	pays	code_pa	annee	produit	code_pro	dispo_int	alim_ani	semenes	pertes	transfo
111	Afrique du Sud	202	2013	Manioc	2532	570000000.0	0.0	0.0	0.0	0.0
112	Afrique du Sud	202	2013	Maïs	2514	10480000000.0	4715000000.0	33000000.0	405000000.0	43000000.0
113	Afrique du Sud	202	2013	Miel	2745	3000000.0	0.0	0.0	0.0	0.0
114	Afrique du Sud	202	2013	Millet	2517	13000000.0	8000000.0	0.0	0.0	0.0
115	Afrique du Sud	202	2013	Miscellanées	2899	0.0	0.0	0.0	0.0	0.0
116	Afrique du Sud	202	2013	Mollusques, Autres	2767	7000000.0	0.0	0.0	0.0	0.0
117	Afrique du Sud	202	2013	Noix	2551	17000000.0	0.0	0.0	0.0	0.0
118	Afrique du Sud	202	2013	Oeufs	2744	497000000.0	0.0	67000000.0	51000000.0	0.0
119	Afrique du Sud	202	2013	Oignons	2602	501000000.0	0.0	0.0	59000000.0	0.0
120	Afrique du Sud	202	2013	Olives	2563	1000000.0	0.0	0.0	0.0	0.0
121	Afrique du Sud	202	2013	Oranges, Mandarines	2611	422000000.0	0.0	0.0	36000000.0	0.0

	pays	code_pa	produit	code_pro	année	dispo_alim_tonnes	dispo_ali	dispo_mi	dispo_pr	origin
1	Afghanistan	2	Abats Comestible	2736	2013	54578.50336	5.0	0.2	0.77	animal
2	Afghanistan	2	Agrumes, Autres	2614	2013	40933.87752	1.0	0.01	0.02	vegetal
3	Afghanistan	2	Aliments pour enfants	2680	2013	1903.90128	1.0	0.01	0.03	vegetal
4	Afghanistan	2	Ananas	2618	2013	0.0	0.0	0.0	0.0	vegetal
5	Afghanistan	2	Bananes	2615	2013	85675.55760000001	4.0	0.02	0.05	vegetal
6	Afghanistan	2	Beurre, Ghee	2740	2013	37126.07496	23.0	2.61	0.03	animal
7	Afghanistan	2	Bière	2656	2013	2855.85192	0.0	0.0	0.0	vegetal
8	Afghanistan	2	Blé	2511	2013	5084368.36824	1369.0	4.69	36.91	vegetal
9	Afghanistan	2	Boissons Alcooliques	2658	2013	0.0	0.0	0.0	0.0	vegetal
10	Afghanistan	2	Café	2630	2013	0.0	0.0	0.0	0.0	vegetal
11	Afghanistan	2	Coco (Incl Coprah)	2560	2013	0.0	0.0	0.0	0.0	vegetal

population

sous_nutrition

	pays	code_pays	annee	nb_personnes
1	Afghanistan	2	2012	7200000.0
2	Afghanistan	2	2013	8100000.0
3	Afghanistan	2	2014	9000000.0
4	Afghanistan	2	2015	9900000.0
5	Afghanistan	2	2016	10500000.0
6	Afrique du Sud	202	2012	2400000.0
7	Afrique du Sud	202	2013	2600000.0
8	Afrique du Sud	202	2014	2900000.0
9	Afrique du Sud	202	2015	3200000.0
10	Afrique du Sud	202	2016	3400000.0
11	Albanie	3	2012	200000.0
12	Albanie	3	2013	200000.0

IV-Le détail des opérations d'algèbre relationnelle

Exemple d'agrégation

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

```
ratio=Bilan_alim['Ratio Energie'].groupby(Bilan_alim['Produit']).mean()  
ratio.sort_values(ascending=False)#on va ranger le ratio par ordre décroissant  
#: renvoie un dataframe avec les lignes triées de telle sorte que la colonne 'C' soit dans l'ordre croissant
```

Produit	
Huile de Soja	6.833624e-06
Huile de Tournesol	6.552116e-06
Graisses Animales Crue	6.376613e-06
Huil Plantes Oleif Autr	6.176561e-06

IV-Le détail des opérations d'algèbre relationnelle

Exemple de jointure : Interne

```
pop2013=population_2013[['Zone','Année','Valeur']]
pop2013.columns=['Pays','Année','Population']
Bilan_alim=Bilan_alim.merge(pop2013, left_on=['Pays','Année'], right_on=['Pays','Année'])
```

	Disponibilité de matière grasse en quantité (g/personne/jour)	Disponibilité de protéines en quantité (g/personne/jour)	Disponibilité intérieure	Exportations - Quantité	Importations - Quantité	Nourriture	Pertes	Production	Semences	Traitements	Variation de stock	Population
2	0.20	0.77	53000000.0	0.0	0.0	53000000.0	0.0	53000000.0	0.0	0.0	0.0	31731.688
9	0.01	0.02	41000000.0	2000000.0	40000000.0	39000000.0	2000000.0	3000000.0	0.0	0.0	0.0	31731.688
6	0.01	0.03	2000000.0	0.0	2000000.0	2000000.0	0.0	0.0	0.0	0.0	0.0	31731.688
0	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	31731.688
0	0.02	0.05	82000000.0	0.0	82000000.0	82000000.0	0.0	0.0	0.0	0.0	0.0	31731.688

IV-Le détail des opérations d'algèbre relationnelle

Exemple de restriction

```
sous_nutr_thailande=sous_nutrition.loc[sous_nutrition['Zone']=="Thaïlande",['Valeur']].sum()*1000000
population_Thailande=population_2013.loc[population_2013['Zone']=="Thaïlande",['Valeur']].sum()*1000
(sous_nutr_thailande/population_Thailande)*100 #recupérer population thailande
```

IV-Le détail des opérations d'algèbre relationnelle

Les 10 pays ayant le plus haut ratio **disponibilité alimentaire/habitant** en termes de protéines (en kg) par habitant

The screenshot shows a MySQL query interface with the following details:

```
1 SELECT pays, SUM(dispo_prot) / 1000 AS dispo_prot_kg
2 FROM dispo_alim
3 GROUP BY pays
4 ORDER BY (dispo_prot_kg)
5 DESC LIMIT 10
```

Below the query, there is a toolbar with various icons for database management. To the right of the toolbar, the text "Nombre de lignes cha" is partially visible.

	pays	dispo_prot_kg
1	Islande	0.13306
2	Chine - RAS de Hong-Kong	0.12907
3	Israël	0.128
4	Lituanie	0.12436
5	Maldives	0.12232
6	Finlande	0.11756

IV-Le détail des opérations d'algèbre relationnelle

Les 10 pays ayant le plus haut ratio **disponibilité alimentaire/habitant** en termes termes de kcal par habitant.

```
1 SELECT pays, SUM(dispo_alim_kcal_p_j) AS dispo_kcal
2 FROM dispo_alim
3 GROUP BY pays
4 ORDER BY(dispo_kcal) DESC
5 LIMIT 10|
```

Table Formulaire

Nombre de lignes chargées : 10

	pays	dispo_kcal
1	Autriche	3770
2	Belgique	3737
3	Turquie	3708
4	États-Unis d'Amérique	3682
5	Israël	3610
6	Irlande	3602

IV-Le détail des opérations d'algèbre relationnelle

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 SELECT année,pays,SUM(dispo_prot)/1000 AS dispo_prot_kg
2 FROM dispo_alim
3 GROUP BY année,pays
4 ORDER BY(dispo_prot_kg)
5 ASC LIMIT 10
```

Nombre de lignes chargées : 10

	année	pays	dispo_prot_kg
1	2013	Libéria	0.03766
2	2013	Guinée-Bissau	0.04405
3	2013	Mozambique	0.04568
4	2013	République centrafricaine	0.04604
5	2013	Madagascar	0.04669
6	2013	Haïti	0.0477

IV-Le détail des opérations d'algèbre relationnelle

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

The screenshot shows a database interface with a query editor at the top and a result table below.

Query:

```
1 SELECT pays,annee, sum(pertes)
2 FROM equilibre_prod
3 Group by pays,annee
4
5
```

Result Table:

	pays	annee	sum(pertes)
1	Afghanistan	2013	1135000000.
2	Afrique du Sud	2013	2193000000.
3	Albanie	2013	276000000.
4	Algérie	2013	3753000000.
5	Allemagne	2013	3781000000.
6	Angola	2013	4799000000.

Below the table, a message says "Nombre de lignes chargées : 174".

IV-Le détail des opérations d'algèbre relationnelle

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

```
1 SELECT sous_nutrition.pays,sous_nutrition.nb_personnes/population.population AS ratio
2 FROM sous_nutrition,population
3 WHERE sous_nutrition.pays=population.pays AND sous_nutrition.annee=population.annee AND sous_nutrition.annee='2016'
4 GROUP BY sous_nutrition.pays,sous_nutrition.nb_personnes
5 ORDER BY(ratio)
6 DESC LIMIT 10
```

Nombre de lignes chargées : 10

	pays	ratio
1	République centrafricaine	0.60940826240075
2	Zimbabwe	0.46438587568502
3	Haïti	0.46094275330694
4	Zambie	0.44601446895046
5	République populaire démocratique de Corée	0.43360655802326
6	Madagascar	0.42981293376209

IV-Le détail des opérations d'algèbre relationnelle

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

The screenshot shows a MySQL query interface with the following details:

Query:

```
SELECT produit, AVG(autres_utilisations/dispo_int) AS Ratio_AU_DI
FROM equilibre_prod
GROUP BY produit
ORDER BY Ratio_AU_DI DESC
limit 10
```

Results:

	produit	Ratio_AU_DI
1	Alcool, non Comestible	0.96551724137931
2	Plantes Aquatiques	0.92066115702479
3	Huile de Palme	0.65252019281057
4	Huil Plantes Oleif Autr	0.54369391758518
5	Huile de Palmistes	0.53418347963942
6	Huile de Colza&Moutarde	0.46264905520177

l'huile de palme est principalement utilisée:

- dans l'oléochimie (savon, cosmétiques...) (19 %)
- en tant qu'agrocarburant