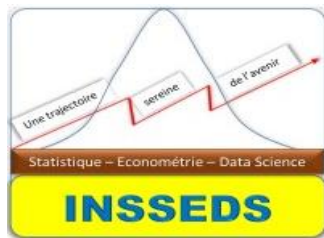


MINISTRE DE L'ENSEIGNEMENT SUPERIEUR  
ET DE RECHERCHE SCIENTIFIQUE



Institut Supérieur de Statistique  
D'Econométrie et de Data Science

REPUBLIQUE DE COTE D'IVOIRE



Union-Discipline-Travail

MASTER 1

**STATISTIQUE – ECONOMETRIE – DATA SCIENCE**

MINI PROJET

**STATISTIQUES DESCRIPTIVE (ANALYSE UNI-BIVARIEES)**

**ANALYSE DES PRIORITÉS D'AIDE HUMANITAIRE :  
CLASSIFICATION DES PAYS SELON DES  
INDICATEURS SOCIO-ÉCONOMIQUES ET  
SANITAIRES**

ANNEE ACADEMIQUE

2024 – 2025

Nom: YOBO

Prenom(s): BAYE GUY ANGE HENOC

Enseignant – Encadreur

AKPOSSO DIDIER MARTIAL

## Avant-Propos

Dans le cadre du master en Statistique, Économétrie et Data Science à l'INSSEDS (Institut Supérieur de Statistique, d'Économétrie et de Data Science), la réalisation d'un mini-projet à la fin de chaque module constitue une exigence pour valider les crédits. C'est dans ce contexte académique que se situe la présente étude.

Il est important de préciser que cette analyse est entièrement descriptive et se limite à répondre à un ensemble de questions préalablement identifiées. Ainsi, en tant qu'auteur, nous n'avons pas la prétention d'avoir exploré tous les aspects nécessaires pour conférer à cette étude le caractère d'une véritable recherche thématique.

De plus, toutes les conclusions présentées dans ce rapport n'engagent que l'auteur et ne reflètent ni l'opinion d'autres personnes, ni celle de l'INSSEDS.

## SOMMAIRE

Avant-Propos .....	1
INTRODUCTION .....	5
I- PREPARATION DES DONNEES.....	6
1- Dictionnaire des données.....	6
2- Presentation du jeux de données .....	7
3- Analyse de la base de données .....	8
3.1- Visualisation et traitement des valeurs manquantes.....	8
3.2- Visualisation et traitement des valeurs aberrantes et/ou extrêmes ...	9
II- STATISTIQUES DESCRIPTIVE UNIVARIEE ET BIVARIEES .....	10
1- Analyses Univariées .....	10
1.1- Parametres statistiques.....	10
1.2- Interprétation des paramètres du tableau statistique simple .....	11
1.3- Les Histogrammes des variable d'intérêt .....	14
2- ANALYSE BIVARIEES .....	16
III- CATEGORISATION CONTEXTUEL DES PAYS EN TENANT COMPTE DES ANALYSES UNI_ BIVARIEES .....	18
1- Regroupement continental.....	18
2- Liste des pays nécessitants de l'aide (avec facteurs) .....	19
a- Revenu net par personne (revenu).....	19
b- PIB par habitant (pib_par_hab).....	20
c- Espérance de vie (esperance_vie).....	20
d- Taux de fécondité total (total_fertilite).....	21
e- Taux de mortalité infantile (enfant_mort).....	21
3- Les Cibles de Croissance : Pays Prioritaires pour le PDG .....	22
Conclusion.....	25
ANNEXE .....	26

## LISTE DES TABLEAUX

Tableau 1: Dictionnaires de données.....	6
Tableau 2: Extrait de la base de données.....	7
Tableau 3: statistiques simple des variables quantitatives.....	10
Tableau 4: correlation lineaire .....	16
Tableau 5: extrait revenu inferieur.....	19
Tableau 6: extrait revenu pib_par_hab.....	20
Tableau 7: extrait esperence_vie inferieur .....	20
Tableau 8: extrait total_fertilite inferieur .....	21
Tableau 9: taux de mortalité .....	21

## LISTE DES GRAPHIQUES

Figure 1: Visualisation et traitement de données manquante.....	8
Figure 2: Visualisation des valeurs aberrantes .....	9
Figure 3: Winsorisation des valeurs.....	9
Figure 4: variable d'intérêt.....	15
Figure 5: corrélation lineaire .....	16

## INTRODUCTION

Dans un monde où les inégalités socio-économiques et les crises humanitaires persistent, les organisations non gouvernementales (ONG) jouent un rôle crucial dans l'amélioration des conditions de vie des populations vulnérables. HELP International, une ONG humanitaire internationale, s'engage à lutter contre la pauvreté et à fournir une assistance essentielle aux pays en difficulté, en particulier lors des catastrophes naturelles et des situations de crise. La collecte de 10 millions de dollars, soit près de 6 milliards de F.CFA, représente une opportunité unique pour l'ONG d'amplifier son impact. Cependant, pour maximiser l'efficacité de cette aide, il est impératif de cibler les pays où les besoins sont les plus pressants. Une étude rigoureuse est donc nécessaire pour identifier les pays qui bénéficieront le plus de cette allocation.

Le PDG de HELP International doit faire face à un dilemme crucial : comment décider de l'allocation des ressources dans un contexte où de nombreux pays présentent des besoins variés ? Les défis socio-économiques et sanitaires qui touchent ces pays ne sont pas uniformes et nécessitent une approche analytique pour prioriser l'aide. Cette étude se penche sur la problématique de la sélection des pays cibles, en identifiant les critères pertinents pour évaluer les besoins et en fournissant une analyse approfondie des données disponibles.

Les résultats de cette étude devraient permettre d'identifier une liste de pays prioritaires pour l'intervention de HELP International, en mettant en lumière les indicateurs clés qui déterminent le besoin d'assistance. Nous nous attendons également à fournir des recommandations concrètes sur l'allocation des ressources financières, afin d'optimiser l'impact de l'aide humanitaire.

Pour répondre à ces objectifs, nous adopterons une méthodologie en plusieurs étapes. D'abord, nous effectuerons un prétraitement des données en nettoyant et en standardisant les informations socio-économiques et sanitaires recueillies. Ensuite, nous utiliserons des analyses statistiques descriptives pour comprendre la distribution des indicateurs parmi les pays concernés.

Enfin nous identifierons les facteurs les plus influents dans le développement, nous établirons un modèle qui permettra de classer les pays selon leur degré de vulnérabilité, facilitant ainsi la prise de décision pour l'allocation des ressources.

I- **PREPARATION DES DONNEES**

La préparation des données consiste en l’ensemble des tâches effectuées pour collecter, nettoyer, transformer, combiner et organiser des données brutes en données de qualité, prêtes pour l’analyse.

**1- Dictionnaire des données**

Un dictionnaire de données est un document qui présente une description complète de chaque variable utilisée dans une analyse statistique ou économétrique. Il détaille les propriétés, les caractéristiques et le contexte de chaque variable, ainsi que leur signification.

VARIABLE	NATURE	DESCRIPTION	MODALITES
Pays	Qualitative	Nom du pays	caractere
enfant_mort	Quantitative	Décès d'enfants de moins de 5 ans pour 1000 naissances vivantes	Numérique
exportations	Quantitative	Exportations de biens et services par habitant. Donné en pourcentage du PIB par habitant	Numérique
dep_sante	Quantitative	Dépenses totales de santé par habitant. Données en pourcentage du PIB par habitant	Numérique
importations	Quantitative	Importations de biens et services par habitant. Donné en pourcentage du PIB par habitant	Numérique
revenu	Quantitative	Revenu net par personne	Numérique
taux_inflation	Quantitative	La mesure du taux d'inflation annuel du PIB total	Numérique
esperance_vie	Quantitative	: Le nombre moyen d'années qu'un nouveau-né vivrait si les tendances de mortalité actuelles devaient rester les mêmes	Numérique
total_fertilite	Quantitative	Le nombre d'enfants qui naîtraient à chaque femme si les taux de fécondité par âge actuels devaient rester les mêmes.	Numérique
pib_par_hab	Quantitative	Le PIB par habitant. Calculé comme le PIB total divisé par la population totale	Numérique

Tableau 1: Dictionaries de données

2- Presentation du jeux de données

	pays	enfant_mort	exportations	dep_sante	importations	revenu	inflation	esperance_vie	total_fertilite_fer	pib_par_hab
1	Afghanistan	90.2	10.0	7.58	44.9	1610	9.44	56.2	5.82	553
2	Albania	16.6	28.0	6.55	48.6	9930	4.49	76.3	1.65	4090
3	Algeria	27.3	38.4	4.17	31.4	12900	16.10	76.5	2.89	4460
4	Angola	119.0	62.3	2.85	42.9	5900	22.40	60.1	6.16	3530
...	...	...	...	...	...	...	...	...	...	...
163	Vanuatu	29.2	46.6	5.25	52.7	2950	2.62	63.0	3.50	2970
164	Venezuela	17.1	28.5	4.91	17.6	16500	45.90	75.4	2.47	13500
165	Vietnam	23.3	72.0	6.84	80.2	4490	12.10	73.1	1.95	1310
166	Yemen	56.3	30.0	5.18	34.4	4480	23.60	67.5	4.67	1310
167	Zambia	83.1	37.0	5.89	30.9	3280	14.00	52.0	5.40	1460

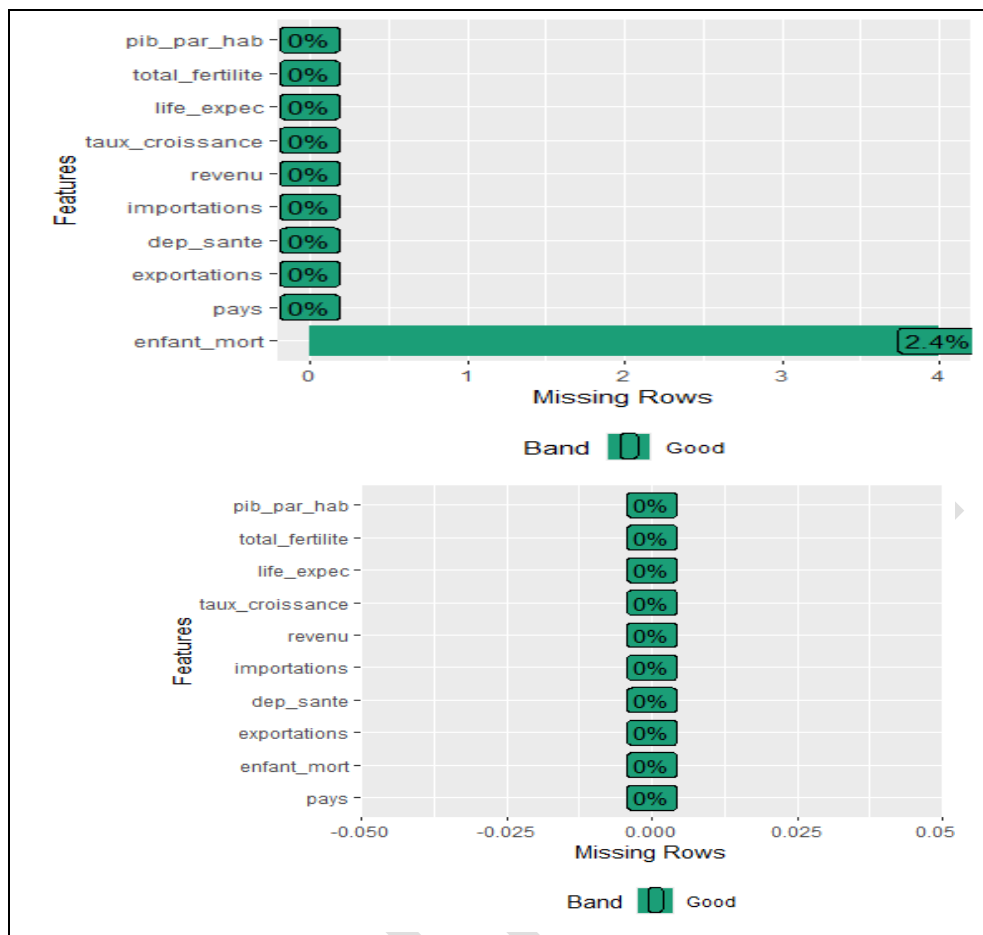
Tableau 2: Extrait de la base de données      Source : HELP international

L’ensemble de données présente 167 observations et 10 variables. Il semble qu’il y ait des valeurs manquantes dans les données, À l’exception de la colonne de pays avec type d’objet, le reste des colonnes a des valeurs numériques.

### **3-Analyse de la base de données**

#### **3.1- Visualisation et traitement des valeurs manquantes**

Dans cette partie, nous allons essayer de détecter visuellement si toute fois notre jeu de donnée présente des valeurs manquantes, et les traiter.



Le jeu de donnée contient des valeurs manquantes qui ont été traité sur l'image ci-contre et les Valeur manquantes se situe au niveau de la variable "enfants morts"

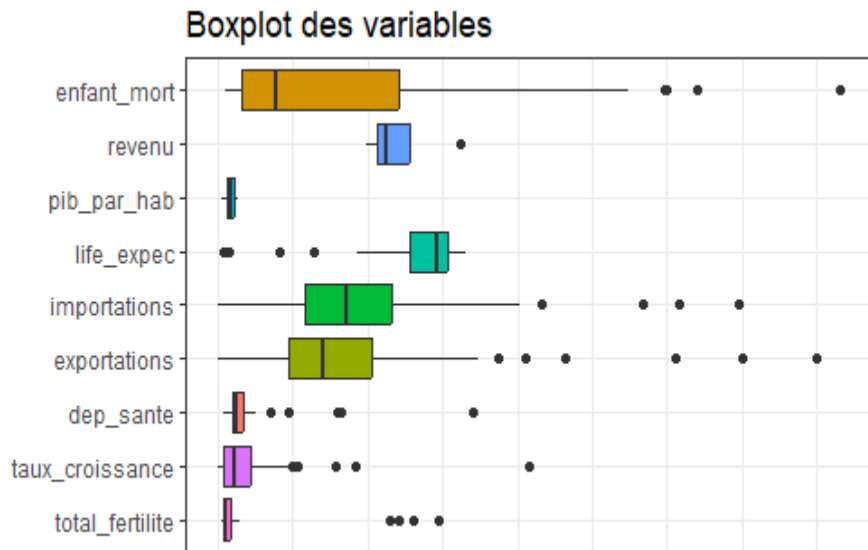
Figure 1: Visualisation et traitement de données manquante

Ces valeurs peuvent résulter d'erreurs de mesure, de saisie, de calcul ou simplement de valeurs extrêmes réelles dans les données. Les valeurs atypiques peuvent avoir un impact considérable sur les analyses statistiques en influençant les résultats des statistiques descriptives comme la moyenne et l'écart-type, ainsi que sur les tests d'hypothèse. Il est donc essentiel de repérer et de gérer ces valeurs extrêmes avant d'entreprendre une analyse statistique



### 3.2- Visualisation et traitement des valeurs aberrantes et/ou extrêmes

#### ➤ Détection des valeurs aberrantes et/ou extrêmes



Une valeur aberrante est une valeur qui s'écarte fortement des valeurs des autres observations, anormalement faible ou élevée. Dans le cas général elle peut modifier l'interprétation de la Moyenne, etc...

#### ➤ Winsorisation

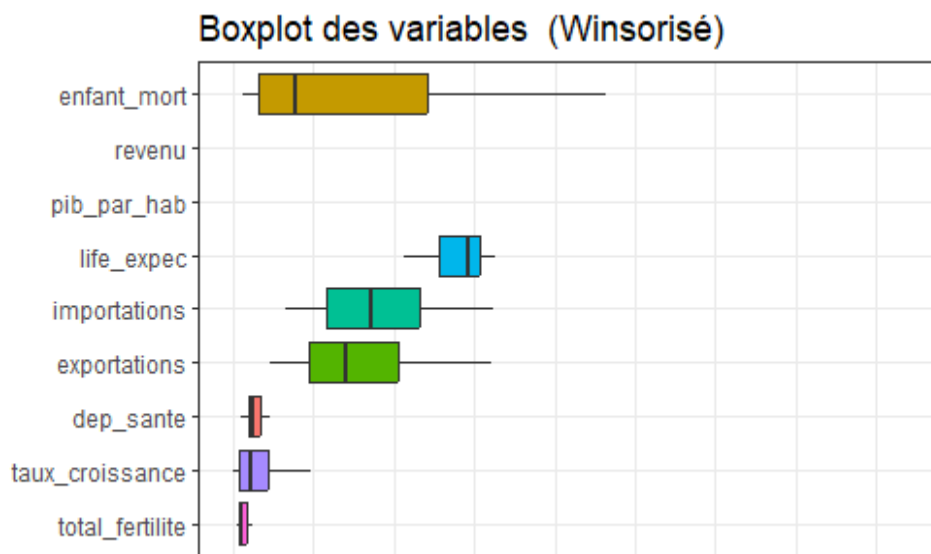


Figure 3: Winsorisation des valeurs

La winorization est une méthode statistique qui permet de traiter les valeurs extrêmes dans un jeu de données. Elle consiste à remplacer les valeurs extrêmes par des valeurs plus proches qui ne sont pas considérées comme extrêmes.

Nb : Nous utiliserons les graphiques avec les valeurs aberrantes / extrêmes pour la suite des analyses, car nous estimons que celles-ci peuvent représenter, en quelque sorte, des pays en situation d'extrême difficulté.

II- **STATISTIQUES DESCRIPTIVE UNIVARIEE ET BIVARIEES**

Cette partie nous permettra sans doute d’avoir certaines informations sur chaque variable et aussi de déterminer les tendances générales de chacune des variables.

1- **Analyses Univariées**

1.1- **Parametres statistiques**

Dans cette partie, nous allons aborder une analyse totalement univariée des variables. Nous allons cependant décrire quelques principales caractéristiques de nos variables quantitatives. Ce qui nous permettra d’avoir des détails plus instructifs avant d’entamer la modélisation statistique. Tableau des résumés numériques de nos variables. Ainsi que des graphiques.

STATISTIQUES SIMPLE									
VARIABLE	E-mort	expor	sante	impor	revenu	T_crois	Life_es	t_ferti l	pib
PARA METRES DE TENDANCE									
Minimum	2.60	0.11	1.81	0.07	49.60	-4.21	2.04	1.15	1.47
Médiane	19.20	35.00	6.33	42.90	9940.00	5.45	73.10	2.41	4640.0
Moyenne	37.80	41.35	7.76	45.62	17023.16	130.49	69.30	4.39	12901.54
Maximum	208.0	200.0	85.1	174.0	125000.0	11400.0	82.8	74.0	105000
PARAMETRES DE DISPERSION									
Etendu	205.4	199.89	83.29	173.93	124950.4	11404.21	80.76	72.85	104998.53
Ecart-type	40.2516	27.93612	7.643689	24.7656	19359.93	998.8153	12.674	9.581274	18367.13
Coeff _variation	106.4961	67.55477	98.4905	54.28396	113.727	765.4556	18.28901	218.2344	142.3639
PARAMETRES DE FORME									
Skewness	1.46	2.33	7.17	1.79	2.17	9.58	-2.57	6.05	2.17
Kurtosis	1.75	8.93	63.93	6.07	6.57	98.62	9.47	36.18	5.19

Tableau 3: statistiques simple des variables quantitatives

## **1.2- Interprétation des paramètres du tableau statistique simple**

### **➤ PARAMETRE DE TENDANCE**

☐ Mortalité infantile (enfant mort): La médiane (19.20) indique que la moitié des pays à un taux de mortalité infantile inférieur à 19.20 décès pour 1000 naissances, ce qui est relativement positif. Cependant, la moyenne (37.80) étant plus élevée suggère que certains pays ont des taux beaucoup plus élevés, tirant ainsi la moyenne vers le haut.

☐ Exportations (exportations): La médiane (35.00) montre que la moitié des pays exporte moins de 35% de leur PIB par habitant, suggérant une dépendance modérée vis-à-vis des exportations. La moyenne (41.35) supérieure à la médiane indique la présence de pays ayant des exportations particulièrement élevées.

☐ Dépenses de santé (dep\_sante): La médiane (6.33) signifie qu'au moins la moitié des pays dépense moins de 6.33% de leur PIB par habitant pour la santé, ce qui peut refléter un investissement modeste. La moyenne (7.76) plus élevée indique que certains pays investissent beaucoup plus dans le secteur de la santé.

☐ Revenu (revenu): La médiane (9940.00) révèle que la moitié des pays a un revenu net par personne inférieur à 9940 dollars, ce qui peut indiquer un niveau de pauvreté. En revanche, la moyenne (17023.16) étant bien plus élevée suggère la présence de pays avec des revenus nettement supérieurs, ce qui pourrait masquer les difficultés économiques de certains.

☐ Taux d'inflation (taux\_inflation): La médiane (5.45) indique qu'au moins la moitié des pays à un taux d'inflation inférieur à 5.45%, suggérant une certaine stabilité économique. Cependant, la moyenne (130.49) très élevée par rapport à la médiane souligne qu'il existe des pays avec des taux d'inflation extrêmement élevés, affectant la stabilité économique.

☐ Espérance de vie (esperance\_vie): La médiane (73.10) montre que la moitié des pays ont une espérance de vie supérieure à 73.10 ans, ce qui indique une bonne santé générale dans plusieurs pays. La moyenne (69.30) étant plus basse peut être influencée par des pays où l'espérance de vie est très faible.

☐ Taux de fertilité (total\_fertilite): La médiane (2.41) indique que la moitié des pays à un taux de fécondité inférieur à 2.41 enfants par femme, suggérant un niveau de reproduction en dessous du seuil de remplacement. En revanche, la moyenne (4.39) plus élevée indique que quelques pays ont des taux de fécondité beaucoup plus élevés.

□ PIB par habitant (pib\_par\_hab): La médiane (4640.0) révèle que la moitié des pays à un PIB par habitant inférieur à 4640 dollars, ce qui peut refléter un niveau de vie modeste. La moyenne (12901.54) étant beaucoup plus élevée indique la présence de pays avec un PIB par habitant très élevé, ce qui peut donner une fausse impression de prospérité générale.

### ➤ paramètres de dispersion

☑ Mortalité infantile (enfant\_mort) : Étendue de 205.4, indiquant une forte variabilité. CV de 106.50% montre que les taux diffèrent largement.

□ Exportations (exportations): Étendue de 199.89, signifiant des disparités. CV de 67.55% indique une variabilité modérée.

□ Dépenses de santé (dep\_sante): Étendue de 83.29, montrant des différences significatives. CV de 98.49% indique une variabilité élevée.

□ Revenu (revenu): Étendue de 173.93, révélant des disparités économiques. CV de 54.28% montre une variabilité faible.

□ Taux d'inflation (taux\_inflation): Étendue très élevée (124950.4) et CV de 113.73% signalent une grande variabilité, avec des pays en crise.

□ Espérance de vie (esperance\_vie): Étendue de 11404.21 et CV très élevé (765.46%) montrent des différences marquées, soulignant des inégalités importantes.

Taux de fertilité (total\_fertilite) : Étendue de 80.76 et CV faible (18.29%) indiquent une cohérence relative entre les pays.

□ PIB par habitant (pib\_par\_hab) : Étendue de 104998.53 et CV très élevé (218.23%) montrent de grandes disparités économiques.

## ➤ paramètres de forme

### Skewness (Asymétrie)

La skewness mesure la symétrie de la distribution des données :

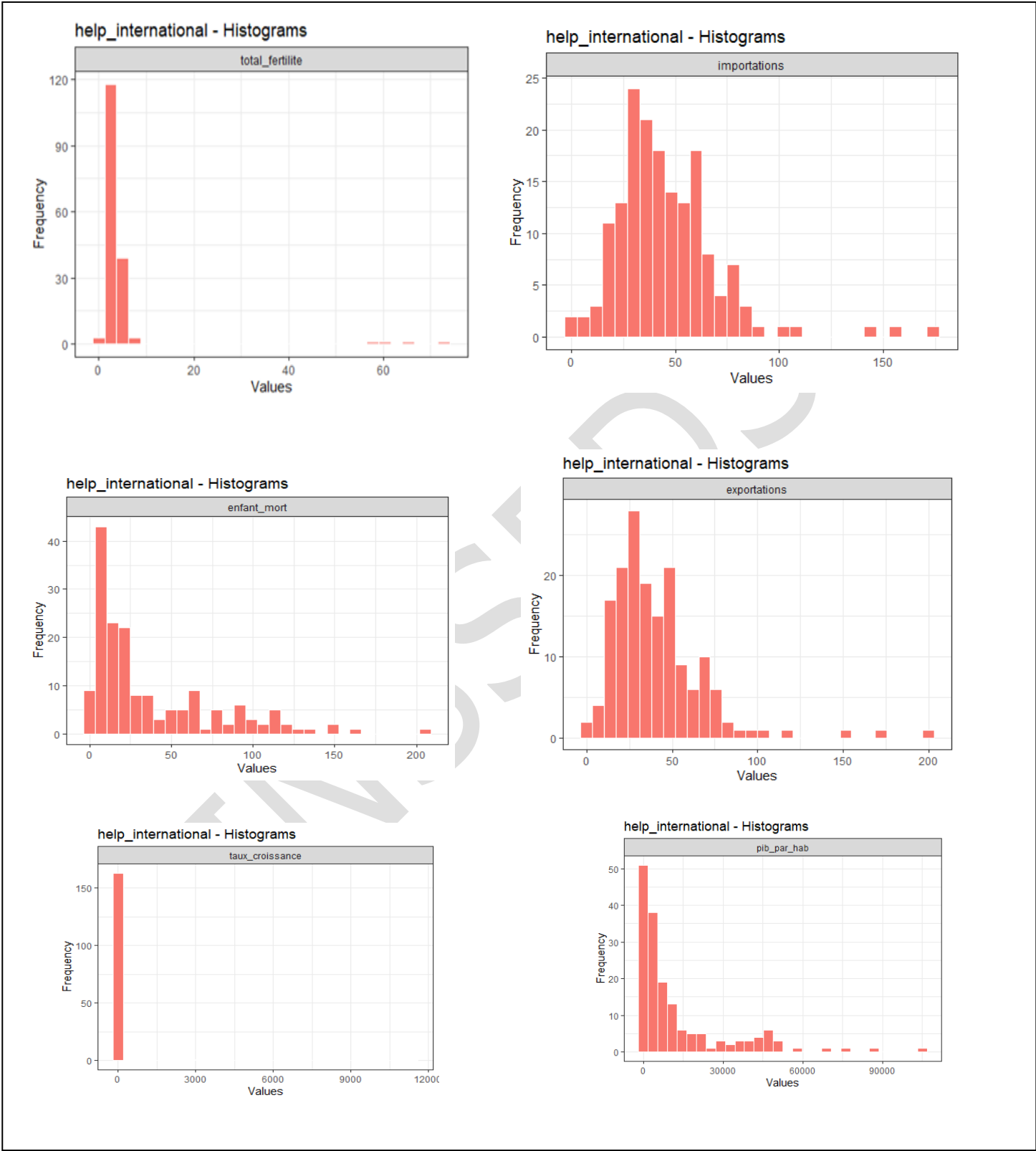
- Skewness positive (supérieure à 0) indique que la distribution est asymétrique vers la droite (queue de distribution plus longue vers les valeurs positives). Par exemple:
  - 1.46, 2.33, 1.79, et 2.17 montrent des distributions légèrement à modérément asymétriques vers la droite, ce qui suggère que des valeurs élevées tirent la moyenne vers le haut.
  - 7.17 et 9.58 indiquent une asymétrie très marquée à droite, ce qui peut être dû à la présence de valeurs extrêmes ou outliers très élevés.
- Skewness négative (inférieure à 0) indique une asymétrie vers la gauche (queue de distribution plus longue vers les valeurs négatives). Par exemple:
  - -2.57 montre une asymétrie marquée à gauche, ce qui signifie que des valeurs faibles influencent la moyenne à la baisse.

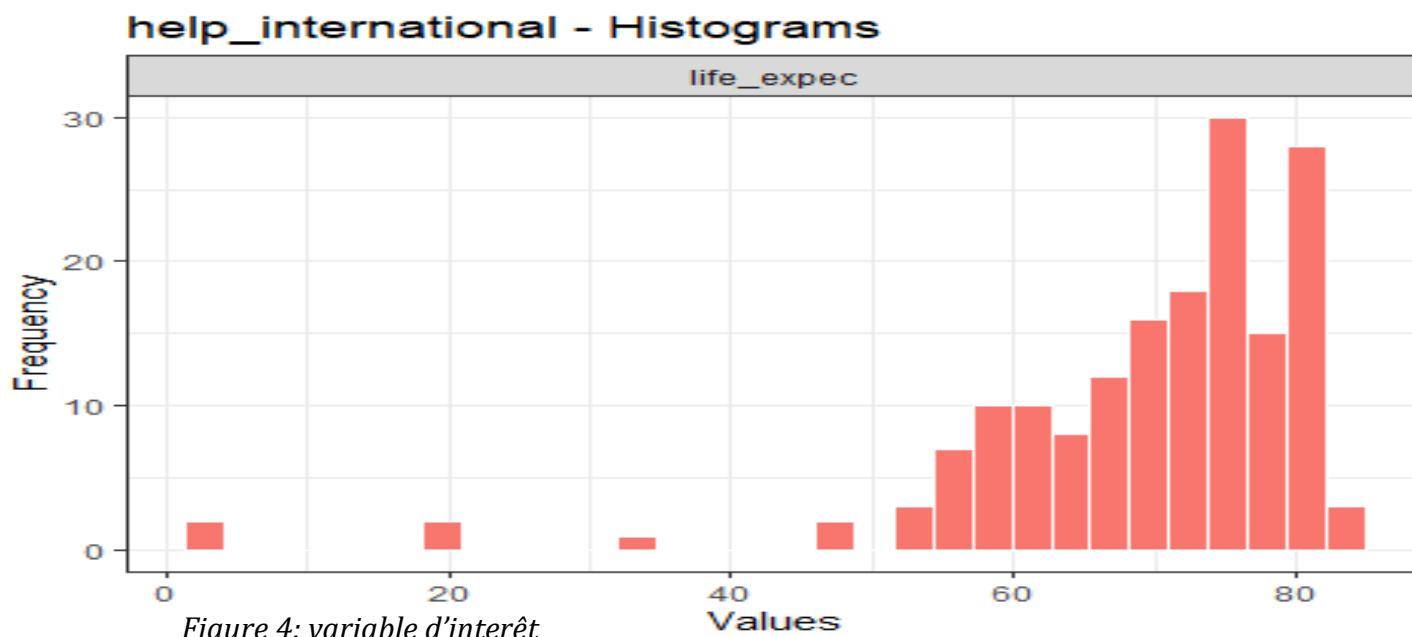
### Kurtosis (Aplatissement)

La kurtosis mesure la forme des "queues" d'une distribution, c'est-à-dire la probabilité de valeurs extrêmes par rapport à une distribution normale :

- Kurtosis positive (supérieure à 3) indique une distribution avec des queues plus épaisses, ce qui signifie qu'il y a une probabilité plus élevée de valeurs extrêmes :
  - 1.75 est proche de la normale (kurtosis de 3), ce qui suggère que la distribution est assez équilibrée.
  - 6.07, 6.57, et 5.19 montrent des distributions modérément "pointues" avec des queues plus épaisses que la normale, ce qui suggère une présence modérée de valeurs extrêmes.
  - 8.93, 36.18, 63.93, et surtout 98.62 indiquent des distributions fortement pointues avec des queues très lourdes, ce qui signifie une forte probabilité de valeurs extrêmes (outliers).

**1.3- Les Histogrammes des variable d’intérêt**





#### Histogramme de la fertilité totale :

Ce graphique montre la distribution de la fertilité totale dans la population. On observe une distribution asymétrique avec une majorité de valeurs inférieures à 40. Cela indique que la plupart des pays ont des taux de fertilité relativement faibles, avec quelques outliers ayant des taux plus élevés.

#### Histogramme des importations :

Cet histogramme représente la fréquence des différentes valeurs d'importations. On remarque une distribution bimodale avec deux pics principaux autour de 10 et 20. Cela suggère une répartition hétérogène des niveaux d'importation entre les pays.

#### Histogramme de la mortalité infantile :

Ce graphique illustre la distribution de la mortalité infantile. La majorité des valeurs se concentrent dans la plage 0-50, indiquant que la plupart des pays ont des taux de mortalité infantile relativement faibles. Cependant, on observe également quelques valeurs plus élevées, traduisant des disparités entre les pays.

#### Histogramme des exportations :

Cet histogramme montre la fréquence des différentes valeurs d'exportations. On constate une distribution asymétrique avec un pic principal autour de 20. Cela suggère que la majorité des pays ont des niveaux d'exportation similaires, avec quelques outliers ayant des niveaux plus élevés.

#### Histogramme du taux de croissance :

Ce graphique représente la distribution du taux de croissance économique. On observe une majorité de valeurs entre 0 et 12000, indiquant que la plupart des pays ont des taux de croissance relativement faibles. Cependant, la présence de valeurs plus élevées montre une certaine hétérogénéité au sein de la population.

#### Histogramme du PIB par habitant :

Cet histogramme illustre la distribution du PIB par habitant. La majorité des valeurs se situent dans la plage 0-30000, suggérant que la plupart des pays ont des niveaux de revenu par habitant relativement modestes. Néanmoins, on constate également l'existence de valeurs plus élevées, révélant des écarts de niveau de vie entre les pays.

Dans l'ensemble, ces graphiques mettent en évidence des tendances et des disparités au sein des indicateurs économiques et sociodémographiques analysés. Une analyse plus approfondie de ces données pourrait permettre d'identifier les facteurs sous-jacents et de formuler des recommandations pertinentes.

2- ANALYSE BIVARIEES

✓ Matrice de corrélation

Une matrice de corrélation est un tableau qui montre la force et la direction des relations linéaires entre plusieurs variables. Cette analyse (matrice de corrélation) établira une la liaison deux (2) à deux (2) des variables. Ce qui nous permettra d’avoir une idée sur les variables qui sont liés à la variable cible et de détecter les variables explicatives pouvant être liées entre elles.

	enfant_mort	exportations	dep_sante	importations	revenu	taux_croissance	life_expec	total_fertilite	pib_par_hab
enfant_mort	1.0000000	-0.33194991	-0.20851288	-0.13110043	-0.5229308	0.2736903	-0.88629614	0.8449601	-0.4831951
exportations	-0.3319499	1.00000000	-0.09231375	0.75187429	0.5275646	-0.1231677	0.33230942	-0.3438243	0.4275053
dep_sante	-0.2085129	-0.09231375	1.00000000	0.07821018	0.1427045	-0.2504706	0.22197147	-0.2062584	0.3641016
importation	-0.1311004	0.75187429	0.07821018	1.00000000	0.1329567	-0.2499945	0.06288076	-0.1691871	0.1245822
revenu	-0.5229308	0.52756464	0.14270449	0.13295674	1.0000000	-0.1415135	0.60860822	-0.5002222	0.8948517
taux_croissanc e	0.2736903	-0.12316772	-0.25047056	-0.24999453	-0.1415135	1.0000000	-0.22535666	0.2977480	-0.2181943
life_expec	-0.8862961	0.33230942	0.22197147	0.06288076	0.6086082	-0.2253567	1.00000000	-0.7555299	0.5986947
total_fertilite	0.8449601	-0.34382428	-0.20625835	-0.16918712	-0.5002222	0.2977480	-0.75552994	1.0000000	-0.4555476
pib_par_hab	-0.4831951	0.42750526	0.36410156	0.12458225	0.8948517	-0.2181943	0.59869469	-0.4555476	1.0000000

Tableau 4: correlation lineaire

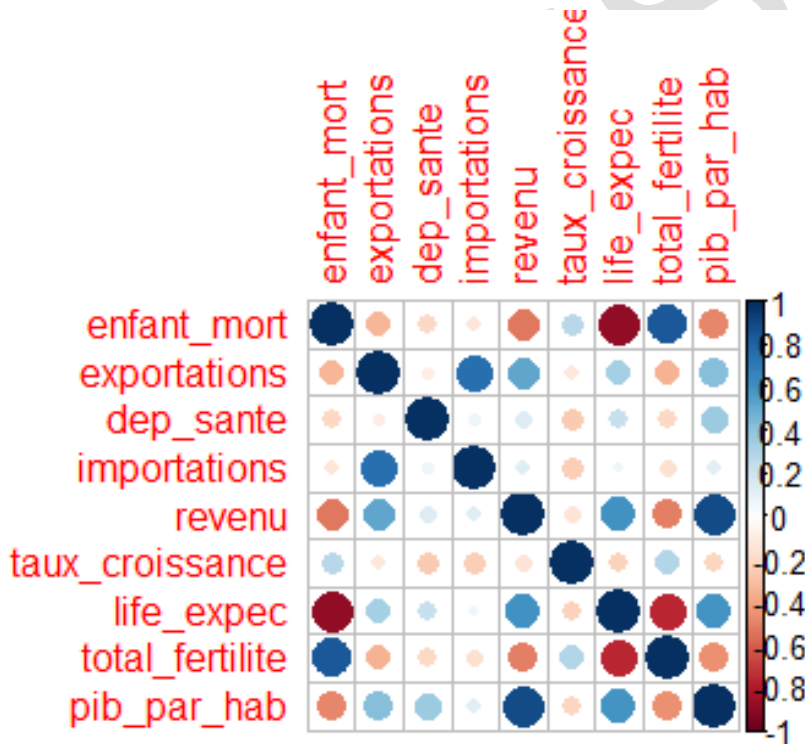
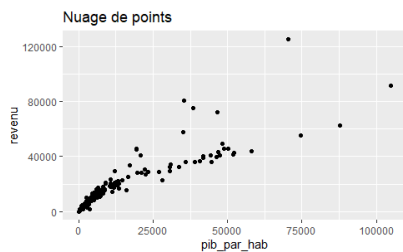


Figure 5: corrélation lineaire

Le tableau ci-contre nous présente une corrélation deux à deux des variables quantitatives. Les données colorées en vert représentent la ligne diagonale de la matrice. Cette corrélation est parfaite car toute variable est corrélée à elle-même. D’une manière générale, nous remarquons une très faible corrélation entre les variables.

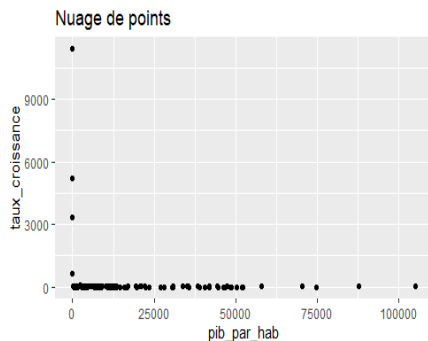
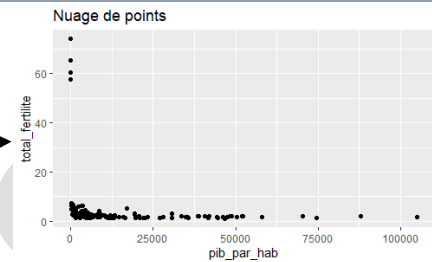


## ✓ Relation entre les variables



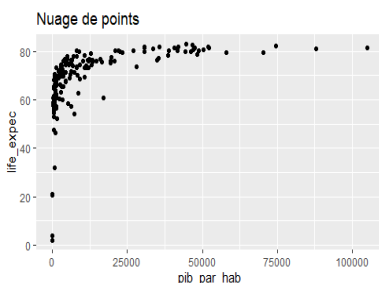
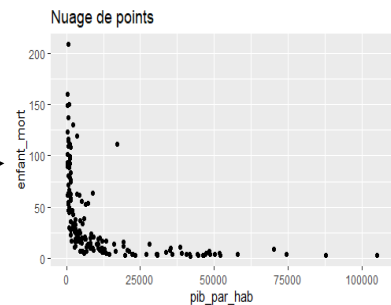
Ce graphique met en évidence une tendance générale selon laquelle les pays avec un PIB par habitant plus élevé tendent à avoir un taux de fertilité plus bas. Cette tendance peut être utile pour identifier des priorités d'aide pour les pays ayant un faible PIB et une forte croissance démographique, ce qui peut exercer une pression sur les ressources.

ce graphique montre un schéma typique observé dans les études socio-économiques : les pays à faibles revenus tendent à avoir des taux de fertilité élevés, tandis que les pays plus développés économiquement affichent des taux de fertilité plus bas.



Ce graphique montre qu'il existe une tendance où les pays à PIB faible peuvent avoir des taux de croissance très élevés, souvent en raison de rattrapages économiques ou de transitions rapides. En revanche, les pays à PIB plus élevé tendent à avoir des taux de croissance plus modérés, signe de maturité économique.

ce graphique montre une relation inverse marquée entre le PIB par habitant et le taux de mortalité infantile. Les pays disposant d'un PIB par habitant faible affichent généralement des taux de mortalité infantile élevés, tandis que ceux avec un PIB par habitant élevé connaissent une mortalité infantile très faible. Cette tendance souligne l'impact positif du développement économique sur la santé infantile, indiquant que l'amélioration du niveau de vie est associée à une réduction significative de la mortalité infantile.



Ce graphique illustre que l'augmentation du PIB par habitant contribue à une amélioration de l'espérance de vie, mais qu'au-delà d'un certain seuil, cet effet se réduit. Cela suggère que si la richesse économique favorise la santé et la longévité, d'autres facteurs que le PIB pourraient influencer l'espérance de vie dans les pays les plus développés.

L'analyse bivariable effectuée permet d'identifier des priorités d'intervention, d'évaluer l'impact potentiel des actions de l'ONG, et de justifier des décisions stratégiques. Elle fournit une base solide pour formuler des recommandations éclairées et maximiser l'efficacité de l'aide humanitaire.

III- **CATEGORISATION CONTEXTUEL DES PAYS  
EN TENANT COMPTE DES ANALYSES  
UNI BIVARIEES**

1- **Regroupement continental**

continent	Enfant-mort	exports	Dep-sante	imports	revenu	inflation	life_expec	total_ferti	Pib-par-hab
Africa	82.155319	1424.041128	137.449087	1327.895745	5627.510638	10.730553	61.504255	4.671277	2644.617021
Asia	29.773333	7621.005078	558.497751	5769.554846	20924.444444	10.522711	72.651111	2.594667	11731.444444
Europe	6.184615	18336.020256	2746.897897	16562.588974	29930.256410	2.803615	77.497436	1.599744	28534.102564
North America	30.020000	4009.195733	1448.965947	5027.166933	15960.666667	3.277867	72.440000	2.309333	13522.133333
Oceania	25.711111	2992.762222	1039.541111	3634.203333	11247.777778	3.252222	68.933333	3.340000	11651.111111
South America	22.258333	2265.883333	555.213500	2013.119167	12429.166667	11.160833	74.450000	2.409167	7690.000000

- ☐ D'après les graphiques et le tableau ci-dessus :
- L'Afrique enregistre le taux de mortalité infantile le plus élevé, tandis que l'Europe présente le taux le plus faible.
  - L'Europe affiche les niveaux les plus élevés en matière d'exportations, d'importations et de dépenses de santé, alors que l'Afrique enregistre les plus faibles.
  - Les revenus les plus élevés se trouvent en Europe et en Asie, tandis que les plus bas sont en Afrique.
  - L'espérance de vie la plus élevée est observée en Europe, alors que l'Afrique a la plus faible.
  - L'inflation est la plus forte en Afrique, en Asie et en Amérique du Sud, tandis qu'elle est la plus basse en Amérique du Nord, en Océanie et en Europe.
  - L'Afrique a le taux de fécondité le plus élevé, alors que l'Europe enregistre le plus bas.
  - Enfin, l'Europe à le PIB par habitant le plus élevé, alors que l'Afrique a le plus faible.

D'après les résultats, il apparaît que plus les niveaux d'importations, d'exportations, de dépenses de santé, d'espérance de vie et de PIB sont élevés, plus un pays tend à être développé. En revanche, des taux élevés de mortalité infantile, d'inflation et un revenu total faible semblent indiquer un moindre niveau de développement.

Ainsi, avant de procéder à la modélisation, on peut envisager une division des pays de l'ensemble de données en trois groupes : les pays moins développés, principalement situés en Afrique et en Amérique du Sud ; les pays à développement intermédiaire, regroupant certains pays d'Asie et du Pacifique ; et les pays développés, majoritairement localisés en Europe et en Amérique du Nord. De ce fait, il est probable que les pays africains et sud-américains nécessitent une aide financière plus importante.

## 2- Liste des pays nécessitants de l'aide (avec facteurs)

Plusieurs pays ont été sélectionnés sur la base de certains critères, selon différentes catégories, notamment au niveau financier et économique (PIB, revenu, inflation). Tous les pays dont le revenu et le PIB par habitant sont inférieurs à la moyenne mondiale (selon le *FMI & Banque Mondiale*) sont classés comme pays pauvres. Sur le plan démographique et sanitaire, tous les pays dont le taux d'espérance de vie est inférieur à la moyenne (OMS) sont classés parmi ceux qui nécessitent une aide. Ceux dont le taux de mortalité infantile est supérieur à la moyenne ont besoin d'aide.

### ❖ Facteurs Économiques et Financiers

#### a- Revenu net par personne (revenu)

Pays	Enfant Mort	Exportations	Dép Santé	Importations	Revenu
Afghanistan	90.2	10.000	7.58	44.9000	1610.0
Albanie	16.6	28.000	6.55	48.6000	9930.0
Algérie	27.3	38.400	4.17	31.4000	12900.0
Angola	119.0	62.300	2.85	42.9000	5900.0
Arménie	18.1	20.800	4.40	45.3000	6700.0
Azerbaïdjan	39.2	54.300	5.88	20.7000	16000.0
Bangladesh	49.4	16.000	3.52	21.8000	2440.0
Barbade	14.2	39.500	7.97	48.7000	15300.0
Biélorussie	5.5	51.400	5.61	64.5000	16200.0
Belize	18.8	58.200	5.20	57.5000	7880.0
Bénin	111.0	23.800	4.10	37.2000	1820.0
Bhoutan	42.7	42.500	5.20	70.7000	6420.0
Bolivie	46.6	41.200	4.84	34.3000	5410.0
Bosnie-Herzégovine	6.9	29.700	11.10	51.3000	9720.0
Botswana	52.5	43.600	8.30	51.3000	13300.0
Brésil	19.8	10.700	9.01	11.8000	14500.0
Bulgarie	10.8	50.200	6.87	53.0000	15300.0
Burkina Faso	116.0	19.200	6.74	29.6000	1430.0
Burundi	93.6	8.920	11.60	39.2000	764.0
Cambodge	44.4	54.100	5.68	59.5000	2520.0
Cameroun	108.0	22.200	5.13	27.0000	2660.0
Cap-Vert	26.5	32.700	4.09	61.8000	5830.0
République Centrafricaine	149.0	11.800	3.98	26.5000	888.0
Tchad	150.0	36.800	4.53	43.5000	1930.0
Chine	15.7	26.300	5.07	22.6000	9530.0

Tableau 5: extrait revenu inferieur

La plupart des pays figurants sur la liste ont des revenus inférieurs à la Moyenne donc conçois des difficultés et un manque de meilleur accès aux services de santé et d'éducation, ce qui défavorise le développement.

b- PIB par habitant (pib\_par\_hab)

Pays	Enfant mort	Exportations	Dép.Santé	Importations	Revenu	Taux Croissance	Vie Espérance	Total Fertilité	PIB par Hab
Afghanistan	90.2	10.000	7.58	44.9000	1610.0	9.440	56.20	5.82	553.00
Albania	16.6	28.000	6.55	48.6000	9930.0	4.490	76.30	1.65	4090.00
Algérie	27.3	38.400	4.17	31.4000	12900.0	16.100	76.50	2.89	4460.00
Angola	119.0	62.300	2.85	42.9000	5900.0	22.400	60.10	6.16	3530.00
Antigua et Barbuda	10.3	45.500	6.03	58.9000	19100.0	1.440	76.80	2.13	12200.00
Argentine	14.5	18.900	8.10	16.0000	18700.0	20.900	75.80	2.37	10300.00
Arménie	18.1	20.800	4.40	45.3000	6700.0	7.770	73.30	1.69	3220.00
Azerbaïdjan	39.2	54.300	5.88	20.7000	16000.0	13.800	69.10	1.92	5840.00
Bangladesh	49.4	16.000	3.52	21.8000	2440.0	7.140	70.40	2.33	758.00
Biélorussie	5.5	51.400	5.61	64.5000	16200.0	15.100	70.40	1.49	6030.00
Belize	18.8	58.200	5.20	57.5000	7880.0	1.140	71.40	2.71	4340.00
Bénin	111.0	23.800	4.10	37.2000	1820.0	0.885	61.80	5.36	758.00
Bhoutan	42.7	42.500	5.20	70.7000	6420.0	5.990	72.10	2.38	2180.00
Bolivie	46.6	41.200	4.84	34.3000	5410.0	8.780	71.60	3.20	1980.00
Bosnie-Herzégovine	6.9	29.700	11.10	51.3000	9720.0	1.400	76.80	1.31	4610.00
Botswana	52.5	43.600	8.30	51.3000	13300.0	8.920	57.10	2.88	6350.00
Brésil	19.8	10.700	9.01	11.8000	14500.0	8.410	74.20	1.80	11200.00
Bulgarie	10.8	50.200	6.87	53.0000	15300.0	1.110	73.90	1.57	6840.00
Burkina Faso	116.0	19.200	6.74	29.6000	1430.0	6.810	57.90	5.87	575.00
Burundi	93.6	8.920	11.60	39.2000	764.0	12.300	57.70	6.26	231.00
Cambodge	44.4	54.100	5.68	59.5000	2520.0	3.120	66.10	2.88	786.00
Cameroun	108.0	22.200	5.13	27.0000	2660.0	1.910	57.30	5.11	1310.00
Cap-Vert	26.5	32.700	4.09	61.8000	5830.0	0.505	72.50	2.67	3310.00
République Centrafricaine	149.0	11.800	3.98	26.5000	888.0	2.010	47.50	5.21	446.00
Tchad	150.0	36.800	4.53	43.5000	1930.0	6.390	56.50	6.59	897.00
Chili	8.7	37.700	7.96	31.3000	19400.0	8.960	79.10	1.88	12900.00
Chine	15.7	26.300	5.07	22.6000	9530.0	6.940	74.60	1.59	4560.00
Colombie	18.6	15.900	7.59	17.8000	10900.0	3.860	76.40	2.01	6250.00
Comores	88.2	16.500	4.51	51.7000	1410.0	3.870	65.90	4.75	769.00
Congo (1)	NA	116.000	41.10	7.9100	49.6	609.000	20.80	57.50	6.54
Cor					5190.000	20.70	60.40		4.95

Le PIB par habitant est un indicateur clé du niveau de vie et du bien-être économique d'un pays. Une croissance du PIB par habitant indique généralement une amélioration des conditions économiques et des opportunités pour les citoyens. Cependant le tableau ci-contre nous présente une part des pays ayant des pays dans des situation défavorable

Tableau 6: extrait revenu pib\_par\_hab

❖ Facteurs Démographiques et Sanitaires

c- Espérance de vie (esperance\_vie)

Pays	Enfant Mort	Exportations	Dép. Santé	Importations	Revenu	Taux Croissance	Vie Espérance	Total Fertilité	PIB par Hab
Afghanistan	90.2	10.000	7.58	44.9000	1610.0	9.44	56.20	5.82	553.00
Angola	119.0	62.300	2.85	42.9000	5900.0	22.4	60.10	6.16	3530.00
Azerbaïdjan	39.2	54.300	5.88	20.7000	16000.0	13.8	69.10	1.92	5840.00
Benin	111.0	23.800	4.10	37.2000	1820.0	0.88	61.80	5.36	758.00
Botswana	52.5	43.600	8.30	51.3000	13300.0	8.92	57.10	2.88	6350.00
Burkina Faso	116.0	19.200	6.74	29.6000	1430.0	6.81	57.90	5.87	575.00
Burundi	93.6	8.920	11.60	39.2000	764.0	12.3	57.70	6.26	231.00
Cambodge	44.4	54.100	5.68	59.5000	2520.0	3.12	66.10	2.88	786.00
Cameroun	108.0	22.200	5.13	27.0000	2660.0	1.91	57.30	5.11	1310.00
République Centrafricaine	149.0	11.800	3.98	26.5000	888.0	2.01	47.50	5.21	446.00
Tchad	150.0	36.800	4.53	43.5000	1930.0	6.39	56.50	6.59	897.00
Comores	88.2	16.500	4.51	51.7000	1410.0	3.87	65.90	4.75	769.00
Congo1	NA	116.000	41.10	7.9100	49.6	609.0	20.80	57.50	6.54
Congo2	NA	63.900	85.10	2.4600	54.7	5190.0	20.70	60.40	4.95
Côte d'Ivoire	111.0	50.600	5.30	43.3000	2690.0	5.39	56.30	5.27	1220.00
Guinée Équatoriale	111.0	85.800	4.48	58.9000	33700.0	24.9	60.90	5.21	17100.00
Érythrée	55.2	4.790	2.66	23.3000	1420.0	11.6	61.70	4.61	482.00
Fidji	24.1	57.800	4.86	63.9000	7350.0	4.23	65.30	2.67	3650.00
Gabon	63.7	57.700	3.50	18.9000	15400.0	16.6	62.90	4.08	8750.00
Gambia	81.2	22.800	5.60	17.7000	1660.0	4.30	65.50	5.71	562.00

Tableau 7: extrait esperance\_vie inferieur

Une espérance de vie élevée est généralement le signe d'un système de santé efficace, d'un environnement sain et d'un bon niveau de vie. Cependant le tableau nous présente des pays dont l'espérance de vie est alarmant

d- Taux de fécondité total (total\_fertilite)

pays	Enfant mort	Exportation	Dépenses sant	Importation	Revenu	Taux croissan	Espérance vi	Total fécondit	PIB par habitant
Afghanistan	90.2	10.00	7.58	44.90	1610.0	9.44	56.20	5.82	553.00
Angola	119.0	62.30	2.85	42.90	5900.0	22.4	60.10	6.16	3530.00
Benin	111.0	23.80	4.10	37.20	1820.0	0.88	61.80	5.36	758.00
Burkina Faso	116.0	19.20	6.74	29.60	1430.0	6.81	57.90	5.87	575.00
Burundi	93.6	8.92	11.60	39.20	764.0	12.3	57.70	6.26	231.00
Cameroun	108.0	22.20	5.13	27.00	2660.0	1.91	57.30	5.11	1310.00
République Centrafricaine	149.0	11.80	3.98	26.50	888.0	2.01	47.50	5.21	446.00
Tchad	150.0	36.80	4.53	43.50	1930.0	6.39	56.50	6.59	897.00
Comores	88.2	16.50	4.51	51.70	1410.0	3.87	65.90	4.75	769.00
Congo1	NA	116.00	41.10	7.91	49.6		609.20.80	57.50	6.54
Congo2	NA	63.90	85.10	2.46	54.7		5190.20.70	60.40	4.95
Côte d'Ivoire	111.0	50.60	5.30	43.30	2690.0	5.39	56.30	5.27	1220.00
Guinée Équatoriale	111.0	85.80	4.48	58.90	33700.0	24.9	60.90	5.21	17100.00
Érythrée	55.2	4.79	2.66	23.30	1420.0	11.6	61.70	4.61	482.00
Gambie	80.3	23.80	5.69	42.70	1660.0	4.30	65.50	5.71	562.00
Guinée	109.0	30.30	4.93	43.20	1190.0	16.1	58.00	5.34	648.00
Guinée-Bissau	114.0	14.90	8.50	35.20	1390.0	2.97	55.60	5.05	547.00
Irak	36.9	39.40	8.41	34.10	12700.0	16.6	67.20	4.56	4500.00
Libéria	89.3	19.10	11.80	92.60	700.0	5.47	60.80	5.02	327.00
Macédoine	NA	10.40	39.80	7.09	58.1		11400.2.04	74.00	1.47

Tableau 8: extrait total\_fertilite inferieur

e- Taux de mortalité infantile (enfant\_mort)

Pays	Enfant Mort	Exportations	Dépenses Santé	Importation	Revenu
Afghanistan	90.2	10.000	7.58	44.9000	1610
Angola	119.0	62.300	2.85	42.9000	5900
Azerbaïdjan	39.2	54.300	5.88	20.7000	16000
Bangladesh	49.4	16.000	3.52	21.8000	2440
Bénin	111.0	23.800	4.10	37.2000	1820
Bhoutan	42.7	42.500	5.20	70.7000	6420
Bolivie	46.6	41.200	4.84	34.3000	5410
Botswana	52.5	43.600	8.30	51.3000	13300
Burkina Faso	116.0	19.200	6.74	29.6000	1430
Burundi	93.6	8.920	11.60	39.2000	764
Cambodge	44.4	54.100	5.68	59.5000	2520
Cameroun	108.0	22.200	5.13	27.0000	2660
République Centrafricaine	149.0	11.800	3.98	26.5000	888
Tchad	150.0	36.800	4.53	43.5000	1930
Comores	88.2	16.500	4.51	51.7000	1410
Côte d'Ivoire	111.0	50.600	5.30	43.3000	2690
Guinée Équatoriale	111.0	85.800	4.48	58.9000	33700
Érythrée	55.2	4.790	2.66	23.3000	1420
Gabon	63.7	57.700	3.50	18.9000	15400
Gambie	80.3	23.800	5.69	42.7000	1660
Ghana	74.7	29.500	5.22	45.9000	3060
Guinée	109.0	30.300	4.93	43.2000	1190
Guinée-Bissau	114.0	14.900	8.50	35.2000	1390
Haïti	208.0	15.300	6.91	64.7000	1500
Inde	58.8	22.600	4.05	27.1000	4410

Tableau 9: taux de mortalité

Des taux de fécondité élevés engendrent une pression sur les ressources du pays et affectent la capacité d'investir dans l'éducation et la santé. Un contrôle des naissances efficace peut contribuer à un meilleur développement économique.

Un taux élevé de mortalité infantile indique souvent des conditions sanitaires défavorables, un manque d'accès aux soins de santé, et des problèmes de nutrition. Cela peut avoir des implications sur la croissance de la population et le développement socio-économique.

### **3-Les Cibles de Croissance : Pays Prioritaires pour le PDG**

Lors de la contextualisation, et en tenant compte des analyses effectuées, nous avons pu sélectionner et catégoriser les pays ayant besoin d'aide, comme la plupart de certains pays sur certains continents. Cependant, nous nous sommes davantage axés sur les pays présentant un cas extrême de pauvreté, des pays qui, même parmi les pays pauvres, figurent parmi ceux touchés par la pauvreté extrême. Nous avons pris en compte plusieurs facteurs (économiques, financiers, démographique et sanitaire), ce qui nous a donné les résultats suivants :

#### **Aides Économiques et Financières**

- **Burundi** : PIB par habitant faible, faible revenu net par personne, et un déséquilibre entre importations et exportations.
- **Mozambique** : Faible PIB par habitant et des dépenses de santé faibles.
- **Tchad** : Taux d'inflation potentiellement élevé et un PIB par habitant très bas.
- **Sierra Leone** : Faible revenu net par personne et déséquilibre commercial.
- **Malawi** : PIB par habitant très bas et problèmes économiques chroniques.
- **Niger** : Taux de fertilité très élevé et un système de santé à améliorer.

#### **Aide Démographiques et Sanitaires**

- **Afghanistan** : Taux de mortalité infantile très élevé, faible espérance de vie.
- **Nigeria** : Forte mortalité infantile et des dépenses de santé insuffisantes.
- **Centrafrique** : Très faible espérance de vie et forte mortalité infantile.
- **Angola** : Taux de mortalité infantile élevé malgré une certaine croissance économique.

#### **Aide Financière, Démographiques et Sanitaires**

- **Équateur** : PIB par habitant modeste, mais également des défis sanitaires.
- **Nigeria** : Nécessité d'un soutien dans les deux domaines, avec des indicateurs préoccupants dans les deux catégories.
- **Tchad** : En raison de sa situation économique et de ses défis en matière de santé.
- **Afghanistan**

✓ **Répartition des ressources financières.**

**Après la sélection des pays nécessitants de l'aide, une répartition des ressources s'est faite En tenant compte des différents facteurs, et qui est aussi une suggestion :**

Pays à cibler : Les pays identifiés, les plus vulnérable sont ceux où la mortalité infantile est élevée et où les ressources pour la santé sont limitées. Ces pays doivent être prioritaires dans l'allocation des fonds.

Proposition de Répartition des Fonds :

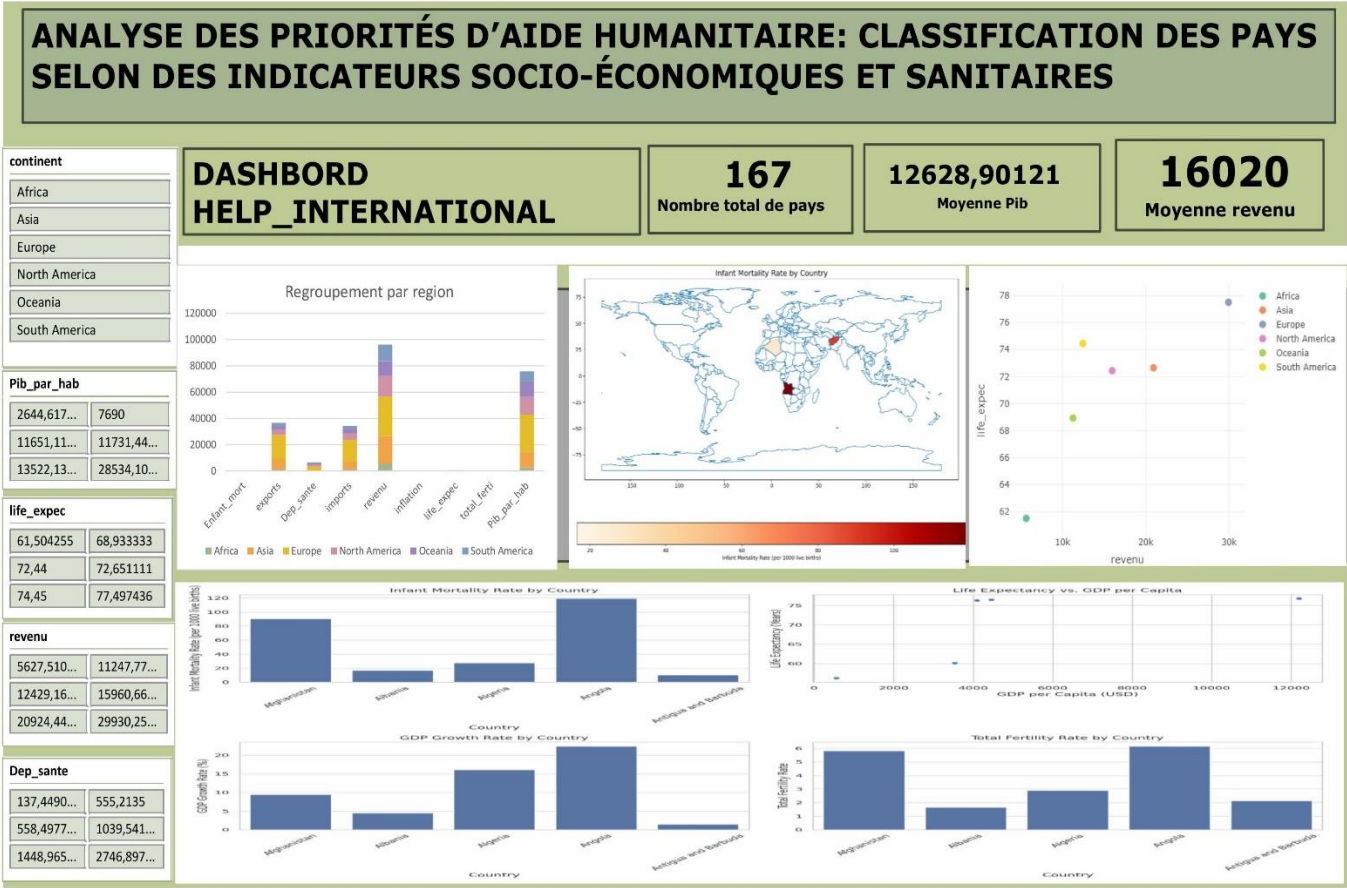
- ❖ 40% des fonds pour des initiatives de santé publique et la réduction de la mortalité infantile dans ces pays.
- ❖ 30% pour des projets de développement économique (augmentation des revenus) dans les pays à faible revenu.
- ❖ 20% pour l'amélioration de l'espérance de vie par des programmes de nutrition et d'accès à l'eau potable.
- ❖ 10% pour des interventions spécifiques selon les besoins particuliers de chaque pays (par exemple, infrastructures médicales).

Nb : Ce partage des fonds répond aux principaux besoins de développement observés : réduire la mortalité infantile, stimuler l'économie, améliorer la santé publique, et adapter les interventions aux besoins locaux. Il s'agit d'une approche équilibrée et stratégique pour maximiser l'impact des investissements dans ces pays.



❖ **TABLEAU DE BORD**

Ce tableau de bord interactif présente une analyse approfondie des indicateurs économiques et de santé de différents pays, offrant une visualisation claire des tendances et facilitant la comparaison des performances à l'échelle continentale et selon les niveaux de développement.





## Conclusion

L'analyse statistique pour HELP International visait à identifier les pays ayant les besoins humanitaires les plus urgents, en s'appuyant sur des indicateurs socio-économiques et sanitaires. Dans un contexte de crise et d'inégalités croissantes, le principal défi pour l'organisation était donc de cibler les pays où l'aide serait la plus efficace. Les résultats montrent que les taux de mortalité infantile, avec une médiane de 19,2 décès pour 1 000 naissances, sont particulièrement élevés dans certains pays africains tels que la Sierra Leone, le Tchad et le Niger. De plus, les écarts d'espérance de vie sont notables, des pays comme l'Afghanistan, la Centrafrique et le Nigeria ayant une espérance de vie bien inférieure à la moyenne mondiale, tout comme leurs revenus par habitant, ce qui révèle des niveaux de vie précaires. Par ailleurs, il apparaît que des pays en Afrique et en Amérique du Sud affichent un PIB par habitant largement inférieur à la moyenne mondiale, tandis que le Niger et le Tchad présentent des taux de fécondité élevés, mettant sous pression leurs ressources et infrastructures.

Ainsi, il est recommandé de cibler l'aide économique vers des pays aux indicateurs faibles en termes de PIB par habitant et de revenu, avec des déséquilibres commerciaux comme le Burundi, le Mozambique et le Tchad. Pour un impact durable, renforcer les infrastructures de santé dans des pays où la mortalité infantile est préoccupante et l'espérance de vie basse, tels que l'Afghanistan, la Centrafrique et le Nigeria, serait stratégique. De plus, dans les pays aux taux de fécondité élevés, des programmes de sensibilisation et de contrôle des naissances pourraient alléger la pression démographique sur les ressources. Cependant, certaines limites de l'analyse doivent être prises en compte : des données manquantes pour certaines variables et certains pays, de même que l'influence des valeurs extrêmes conservées pour représenter des situations de grande difficulté, réduisent en partie la précision des résultats. Ces valeurs extrêmes, bien qu'elles reflètent des situations de grande difficulté, peuvent fausser certaines tendances statistiques.

En outre, l'absence d'analyses longitudinales limite cette étude, puisqu'elle se base uniquement sur des données ponctuelles, sans suivre l'évolution des besoins dans le temps. Pour améliorer la précision de l'analyse, il serait pertinent de réaliser des études longitudinales, qui permettraient de suivre les indicateurs au fil des années et d'affiner les priorités d'aide en fonction de leur évolution. De plus, l'intégration de nouveaux indicateurs, tels que l'accès à l'eau potable, la qualité de l'éducation, et la stabilité politique, enrichirait les critères de sélection et permettrait de mieux cibler les besoins des populations. Enfin, l'adoption de modèles prédictifs permettrait d'anticiper les futurs besoins et d'ajuster l'aide humanitaire de manière proactive, en tenant compte des projections économiques et sanitaires.

# ANNEXE

## Reference

### Webiographie

- + Organisation des Nations Unies (ONU). (2023). *Rapport mondial sur le développement humain*. Récupéré de <https://hdr.undp.org/fr/content/human-development-reports>
- + Banque mondiale. (2023). *Indicateurs du développement dans le monde*. Récupéré de <https://databank.worldbank.org/>
- + Organisation mondiale de la santé (OMS). (2023). *Rapport sur la santé dans le monde*. Récupéré de <https://www.who.int/>
- + Institut National de la Statistique et des Études Économiques (INSEE). (2022). *Analyse démographique et économique des pays*. Paris, France.
- + **Fonds monétaire international (FMI)**. (2023). *Perspectives économiques mondiales*. Récupéré de <https://www.imf.org>
- Ce rapport du FMI permet de mieux comprendre les tendances économiques globales et les défis spécifiques auxquels les pays en développement sont confrontés.
- + **Organisation de Coopération et de Développement Économiques (OCDE)**. (2023). *Base de données sur les indicateurs de bien-être*. Récupéré de <https://www.oecd.org>
  - L'OCDE compile des données sur les dépenses de santé, le bien-être et l'espérance de vie, qui sont des indicateurs cruciaux pour les analyses comparatives entre pays développés et en développement.
  - **Le coinstat**, cours sur la statistique descriptive univariée et bivarie.
  - **OMS - Taux de mortalité** : [Rapport de l'OMS sur la mortalité](#)
  - **Statistiques de l'OMS** : [Données de l'OMS](#) , [Rapport de l'OMS sur l'espérance de vie](#)

### Bibliographie

**Mr Akposso**, 2024 cours sur la statistique descriptive uni et bivariée.  
**François Husson**, Cours sur la statistique descriptive univariée et bivarie.

## SOURCE DU CODE << R >>

### ✓ PREPARATION DES DONNEES

#### Importation et presentation du jeux de données

```
help <- read.csv("C:/Users/HP/Desktop/help_international.csv", sep=";", stringsAsFactors=TRUE)
presentation du jeux de données
head(help), tail(help)
la structure du jeu de données
str(help)
visualisation des valeurs manquantes
library(DataExplorer) plot_missing(help)
Traitement des valeurs manquantes
median(help$enfant_mort, na.rm = TRUE) help$enfant_mort[is.na(help$enfant_mort)] <- median(help$enfant_mort,na.rm=TRUE) plot_missing(help)
Traitement des valeurs extrêmes et/ou aberrantes library(tidyr) library(tidyverse) help %>% pivot_longer(cols = c(enfant_mort, life_expec, total_fertilite,
exportations, importations, revenu, pib_par_hab, taux_croissance, dep_sante), names_to = "Attributes", values_to = "values") %>%
ggplot(aes(x=reorder(Attributes, values, FUN=median), y=values, fill=Attributes)) + geom_boxplot(show.legend=FALSE) + labs(title="Boxplot des variables ") +
theme_bw() + theme(axis.title.y=element_blank(), axis.title.x=element_blank()) + ylim(0, 210) + coord_flip()
```

#### Winzorisatoin library

```
(dplyr) library(tidyr) library(ggplot2) library(psych) help_winsorized <- help %>% mutate(across(c(enfant_mort, life_expec, total_fertilite,
exportations, importations, revenu, pib_par_hab, taux_croissance, dep_sante), ~ winsor(., trim = 0.05))) help_winsorized %>%
pivot_longer(cols = c(enfant_mort, life_expec, total_fertilite, exportations, importations, revenu, pib_par_hab, taux_croissance,
dep_sante), names_to = "Attributes", values_to = "values") %>% ggplot(aes(x = reorder(Attributes, values, FUN = median), y = values,
fill = Attributes)) + geom_boxplot(show.legend = FALSE) + labs(title = "Boxplot des variables (Winsorisé)") + theme_bw() +
theme(axis.title.y = element_blank(), axis.title.x = element_blank()) + ylim(0, 210) + coord_flip()
```

### ✓ STATISTIQUE DESCRIPTIVES

- Analyse univariee

#### Resumer Numerique des variables

```
henoc.qt.resume <- function(vecteur){
res1<-min(vecteur)
res2<-max(vecteur)
library(RVAideMemoire)
res3<-mod(vecteur)
res4<-median(vecteur,na.rm=TRUE)
res5<-mean(vecteur,na.rm=TRUE)
res6<-quantile(vecteur,na.rm=TRUE)
res7<-cv(vecteur)
res8<-var(vecteur,na.rm=TRUE)
res9<-sd(vecteur,na.rm=TRUE)
library(moments)
res10<-skewness(vecteur)
interpskew<- ifelse(res10<0,'distribution etalee a gauche','distribution etalee a droite')
res11<-kurtosis(vecteur)
interpkurt<- ifelse(res11<3,'distribution platikurtique','distribution leptokurtique')
return(list(minimum=res1,maximum=res2,mode=res3, mediane=res4,moyenne=res5,quantile=res6, coefficient_variation=res7,variance=res8,ecart_type=res9,
coefficient_assymetrie=res10,interpretation_skewness=interpskew, coefficient_applatissement=res11,interpretation_kurtosis=interpkurt))
}
```

#### Representation graphique

```
henoc.qt.graph<-function(vecteur){
require(ggplot2)
require(gridExtra)
# Diagramme en b?tons
res1 <- ggplot(data.frame(vecteur), aes(x = vecteur)) +geom_bar(stat = "count") + labs(title = "Diagramme en b?tons", x = "Valeurs", y = "Effectifs")
# Polygone des fr?quences cumul?es
res2 <- ggplot(data.frame(vecteur), aes(x = vecteur)) + stat_ecdf() +labs(title = "Polygone des fr?quences cumul?es", x = "Valeurs", y = "Fr?quences cumul?es")
# Histogramme
res3 <- ggplot(data.frame(vecteur), aes(x = vecteur)) +geom_histogram(aes(y=..density..), bins = 30, color = "black", fill = "green") +
geom_density(alpha = .2, fill = "red") +
labs(title = "Histogramme", x = "Valeurs", y = "Densit?")
# Bo?te ? moustaches
res4 <- ggplot(data.frame(vecteur), aes(x = "", y = vecteur)) +geom_boxplot(fill = "green") +labs(title = "Bo?te ? moustaches", x = "", y = "Valeurs")
grid.arrange(res1, res2, res3, res4, ncol = 2)}
```

- Analyse bivariees

## Liaison

```
henoc.2qt.liason<-function(vecteur1,vecteur2){
  res1<- cor(vecteur1,vecteur2,method="pearson")
  res2<- cor(vecteur1,vecteur2,method="spearman")
  res3<- cor(vecteur1,vecteur2,method="kendall")

  res4<- cor(vecteur1,vecteur2,method="pearson")^2
  inter1<- ifelse(res4<0.10,'liaison tr?s faible',ifelse(res4<0.40,'liaison faible',ifelse(res4<0.60,'liaison moyenne',ifelse(res4<0.80,'liaison forte','liaison tres
forte'))))
  res5<- lm(vecteur1~vecteur2)$coefficients
  res6<- cor.test(vecteur2, vecteur1)
  inter2<- ifelse(res6$p.value<0.05,'liaison significative','liaison non significative')
  rem<- "Si la liaison n'est pas significative, Ne pas tenir compte de son intensite"
  return(list(Correlation_Pearson=res1, Correlation_Spearman=res2, Correlation_Kendall=res3, Coefficient_Determination=res4,
Interpretation_Liaison=inter1, Coefficients_Droite_Regression=res5, Resultat_Test_Liaison=res6, p.value=res6$p.value, Significacite_Liaison=inter2,
Remarque=rem))
}
```

## ANALYSE BI

```
# Analyse de corrélation entre les variables
cor_matrix <- cor(help[, -1], use = "complete.obs")print(cor_matrix)
# Visualisation de la matrice de corrélation
library(corrplot)
corrplot(cor_matrix, method = "circle")dependance entre les variable .

```{r}
library(ggplot2)
ggplot(help, aes(x = pib_par_hab, y = revenu)) +
  geom_point() + xlab("pib_par_hab") + ylab("revenu") + ggtitle("Nuage de points")
ggplot(help, aes(x = pib_par_hab, y = enfant_mort)) + geom_point() + xlab("pib_par_hab") + ylab("enfant_mort") + ggtitle("Nuage de points") ggplot(help,
aes(x = pib_par_hab, y = total_fertilite)) + geom_point() + xlab("pib_par_hab") + ylab("total_fertilite") + ggtitle("Nuage de points")
ggplot(help, aes(x = pib_par_hab, y = taux_croissance)) +geom_point() + xlab("pib_par_hab") + ylab("taux_croissance") +
  ggtitle("Nuage de points") ggplot(help, aes(x = pib_par_hab, y = life_expec)) +geom_point() + xlab("pib_par_hab") +
```

## ✓ CATEGORISATION

### CATEGORISATION ET CONTEXTUALISATION

```
*** problemes economique et financier
PIB
moyenne_pib_par_hab <- mean(help$pib_par_hab , na.rm = TRUE)
pays_en_difficulté <- help[help$pib_par_hab < moyenne_pib_par_hab, ]
print("Pays en difficulté :")
print(pays_en_difficulté)
## revenu
moyenne_revenu <- mean(help$revenu, na.rm = TRUE)
pays_en_difficulté <- help[help$revenu < moyenne_revenu, ]
print("Pays en difficulté :")
print(pays_en_difficulté)
## inflation
moyenne_taux_croissance <- mean(help$taux_croissance , na.rm = TRUE)
pays_en_difficulté <- help[help$taux_croissance >= moyenne_taux_croissance, ]
print("Pays en difficulté :")
print(pays_en_difficulté)
**** probleme sanitaire et demographiques
#mortalité
moyenne_enfant_mort <- mean(help$enfant_mort , na.rm = TRUE)
pays_en_difficulté <- help[help$enfant_mort >= moyenne_enfant_mort, ]
print("Pays en difficulté :")
print(pays_en_difficulté)
#esperence vie
moyenne_life_expec <- mean(help$life_expec, na.rm = TRUE)
pays_en_difficulté <- help[help$life_expec < moyenne_life_expec, ]
pays_developpés <- help[help$life_expec >= moyenne_life_expec, ]
print("Pays en difficulté :")
print(pays_en_difficulté)
##total_fertilite
moyenne_total_fertilite <- mean(help$total_fertilite , na.rm = TRUE)
pays_en_difficulté <- help[help$total_fertilite >= moyenne_total_fertilite, ]
print("Pays en difficulté :")
print(pays_en_difficulté)
```