



21 MARS 2025

SAE PROJET

ANALYSE STATISTIQUE DU BONHEUR DANS LE MONDE

TITULAIRES : M. LABIOD, M. GOMEZ-GARCIA

GROUPES 11-12

BUT SCIENCE DES DONNÉES 1



AMILIA BALIT
ENZO CHEN
MICHAEL DAI
HUGO MA
REINE MANTOUADI KINA

01



IUT de Paris - Rives de Seine
Université Paris Cité



SOMMAIRE

- Introduction
- Problématique
- Configuration de l'environnement de travail
- Chargement et fusion des fichiers
- Ajout de la variable "continent"
- Nettoyage des données
- Exportation finale des données





INTRODUCTION

Base de données du World Happiness Report de 2005 à 2024,
avec des données couvrant 165 pays



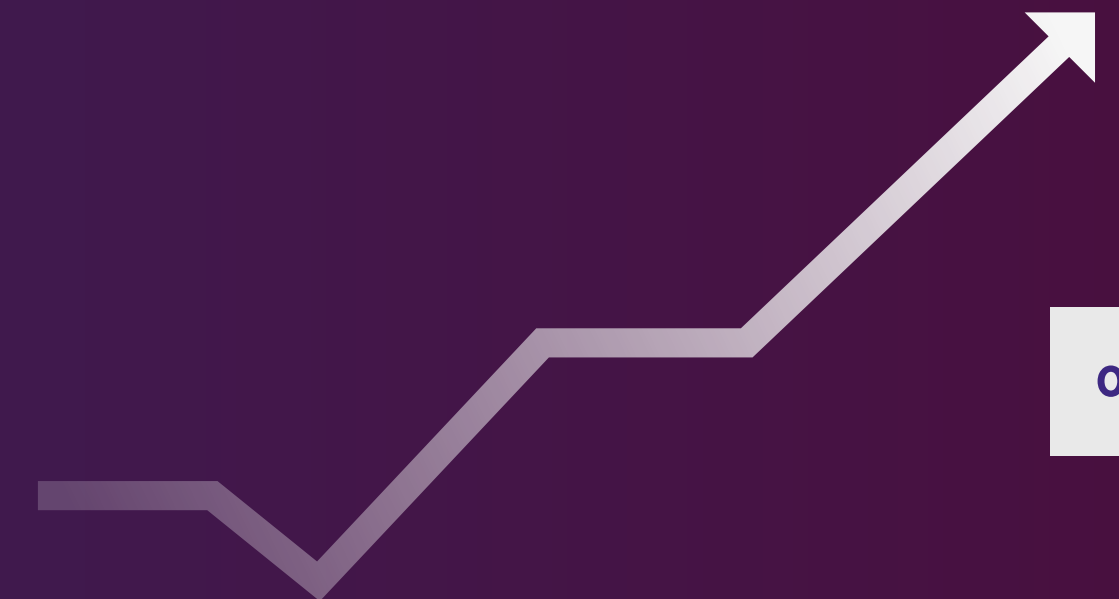
BONHEUR ?

VARIABLES ÉTUDIÉES ?



PROBLÉMATIQUE




L'AUGMENTATION DU PIB PAR HABITANT ENTRAÎNE-T-ELLE UNE AMÉLIORATION DU BONHEUR, QUEL QUE SOIT LE CONTEXTE GÉOGRAPHIQUE ET TEMPOREL ?



CONFIGURATION DE L'ENVIRONNEMENT DE TRAVAIL

Activation des bibliothèques nécessaires avec la commande **library()**.

Les packages sont installés avec **install.packages()** et comprennent :

- **readxl** : Pour lire les fichiers Excel 
- **dplyr** : Pour la manipulation et le nettoyage des données 
- **writexl** : Pour exporter les résultats au format Excel 

Définition du répertoire de travail avec la commande **setwd()**,



CHARGEMENT ET FUSION DES FICHIERS

Importation des données

3 fichiers excel -> read_excel()

Ajout d'une variable

Variable "Year" -> mutate()

Rennommage

Ladder score -> Life Ladder

FUSION DES DONNÉES EN 1 SEULE BASE ->
BIND_ROWS()



AJOUT DE LA VARIABLE "CONTINENT"



Nom français	Capitale	Continent
Afghanistan	Kaboul	Asie
Afrique du Sud	Pretoria	Afrique
Åland(les Îles)	–	Europe
Albanie	Tirana	Europe
Algérie	Alger	Afrique
Allemagne	Berlin	Europe
Andorre	Andorre-la-Vieille	Europe
Angola	Luanda	Afrique
Anguilla	–	Amérique N.
Antarctique	–	Antarctica
Antigua-et-Barbuda	Saint John's	Amérique N.
Arabie saoudite	Riyad	Asie
Argentine	Buenos Aires	Amérique S.

EXTRAIT DU CODE R

```
country_continent_mapping <- country_continent_mapping %>%  
  mutate(Nom_anglais = case_when(  
    Nom_anglais == "United States of America" ~ "united states",  
    Nom_anglais == "United Kingdom of Great Britain and Northern Ireland" ~ "united kingdom",  
  
    Nom_anglais == "Taiwan (Province of China)" ~ "taiwan province of china",  
    .... )
```

left_join() -> fusion des deux bases



NETTOYAGE DES DONNÉES

Valeurs manquantes (NA) :

- Colonnes avec $<15\%$ de NA : imputées par moyennes conditionnelles (pays).
- Colonnes avec $>15\%$ de NA : supprimées pour assurer la fiabilité.

Valeurs aberrantes :

- Valeurs extrêmes détectées via IQR, remplacées par la médiane pour éviter les biais.



LIEN PROGRAMME R



EXPORTATION FINALE DES DONNÉES

-> DATA SET PROPRE ET HOMOGÈNE

exporté au format Excel grâce à la fonction **write_xlsx()**

A	B	C	D	E	F	G	H	I	J	K	L
Country.name	year	Life.Ladder	Log.GDP.per.capita	Social.support	Healthy.life.expectancy.at.birth	Freedom.to.make.life.choices	Generosity	Perceptions.of.corruption	Positive.affect	Negative.affect	Continent
australia	2005	7,340688229	10,66205788	0,96789217	69,80000305	0,9349733	0,001449357	0,390415937	0,769770324	0,23801209	Océanie
belgium	2005	7,262290478	10,74380779	0,934874713	68,40000153	0,923843026	0,001449357	0,597554445	0,676885664	0,26037994	Europe
brazil	2005	6,636771202	9,4351511	0,882922947	63,09999847	0,882186115	0,001449357	0,744994044	0,769921243	0,301779836	Amérique du Sud
canada	2005	7,418048382	10,70729923	0,961552441	70,5	0,957306266	0,248032972	0,502681196	0,78329885	0,233278155	Amérique du Nord
czechia	2005	6,439256668	10,32168579	0,918759465	67,09999847	0,865234971	0,001449357	0,900732756	0,638764262	0,257948965	Europe
denmark	2005	8,01893425	10,84901237	0,972371519	68,30000305	0,97113502	0,001449357	0,236521706	0,776689112	0,15367195	Europe
egypt	2005	5,167754173	8,989677429	0,847842455	61,40000153	0,81736201	0,001449357	0,743971888	0,688714862	0,345555186	Afrique
france	2005	7,093392849	10,63676929	0,940338254	70,69999695	0,89481926	0,001449357	0,687850833	0,681277752	0,225094199	Europe
germany	2005	6,619549751	10,69079208	0,963490367	69,90000153	0,846623778	0,001449357	0,781006813	0,684764624	0,197262481	Europe
greece	2005	6,006309986	10,45363522	0,836539447	69,59999847	0,734171808	0,001449357	0,860563099	0,597686529	0,263643026	Europe
hungary	2005	5,193933487	10,10293579	0,929628253	65	0,696874499	0,001449357	0,902810693	0,578233123	0,290326953	Europe
iran	2005	5,308190346	9,497694016	0,765978038	64,30000305	0,651167691	0,001449357	0,636490345	0,514641285	0,456109405	Asie
italy	2005	6,853783607	10,697855	0,928000689	70,59999847	0,802194953	0,001449357	0,943912327	0,606102049	0,294698149	Europe
japan	2005	6,515817165	10,55191517	0,927711964	72,40000153	0,867779255	0,001449357	0,698929727	0,685981691	0,15315105	Asie
jordan	2005	6,294660091	9,28282547	0,920012951	65,80000305	0,750211722	0,001449357	0,669726729	0,629833221	0,239559561	Asie
lebanon	2005	5,49124527	9,572590828	0,796278358	65,09999847	0,703205824	0,001449357	0,945177019	0,558134913	0,292149812	Asie
mexico	2005	6,580657959	9,791635513	0,902807653	64,40000153	0,813745499	0,001449357	0,764249027	0,763287425	0,218942866	Amérique du Nord
netherlands	2005	7,463979244	10,80906963	0,947357953	70,69999695	0,901007771	0,001449357	0,57134223	0,700739443	0,232794717	Europe
pakistan	2005	5,224657536	8,25220871	0,590945721	53,20000076	0,629995883	0,001449357	0,844436169	0,652143024	0,237265706	Asie
poland	2005	5,587209225	9,843979836	0,921527624	66,19999695	0,782473147	0,001449357	0,982930899	0,610625982	0,282439232	Europe
romania	2005	5,048648357	9,733084679	0,837685466	64,5	0,800120592	0,001449357	0,956884563	0,576348126	0,345686913	Europe
saudi arabia	2005	7,079644203	10,6748333	0,867819488	61,20000076	0,750211722	0,001449357	0,505149066	0,680691183	0,242553025	Asie
spain	2005	7,152785778	10,54381371	0,961042881	70,40000153	0,916164696	0,001449357	0,777272284	0,694322646	0,240642577	Europe
sweden	2005	7,376315594	10,72415447	0,951469898	71	0,964395404	0,001449357	0,743971888	0,742479682	0,150766104	Europe
turkiye	2005	4,718733788	9,800280571	0,819936395	66,09999847	0,623114944	0,001449357	0,876998603	0,478679717	0,271501123	Asie
united kingdom	2005	6,983556747	10,66725636	0,978839755	69,09999847	0,922354519	0,001449357	0,398456872	0,779467702	0,261732221	Europe
venezuela	2005	7,169620991	9,316228867	0,955278456	65,5	0,838198006	0,001449357	0,719800055	0,802928805	0,23301369	Amérique du Sud
argentina	2006	6,312925339	9,936885834	0,938462794	65,81999969	0,733003676	-0,159916088	0,851799488	0,748283088	0,328230023	Amérique du Sud
armenia	2006	4,289310932	9,020847321	0,681876779	63,84000015	0,520197809	-0,232203081	0,849513113	0,45341751	0,469418824	Asie
austria	2006	7,122211456	10,83645248	0,936350405	69,5	0,941382349	0,298436075	0,490111172	0,746153116	0,173811689	Europe
azerbaijan	2006	4,727870941	9,154187202	0,85441488	60,58000183	0,771528184	-0,236623272	0,774117172	0,468755603	0,275695145	Asie
bangladesh	2006	4,318909168	7,939726353	0,672002196	59,11999893	0,611664236	0,055184368	0,785916209	0,458654761	0,320792794	Asie
belarus	2006	5,657649994	9,489098549	0,917898655	60,06000137	0,707080483	-0,249204546	0,708274722	0,535119176	0,269400358	Europe
benin	2006	3,329801559	7,843532562	0,444781214	51,95999908	0,580069304	-0,011397413	0,789862454	0,520846486	0,30910027	Afrique
bolivia	2006	5,373986244	8,670664787	0,834279954	60,90000153	0,770134747	-0,045580111	0,794484198	0,708094001	0,431944966	Amérique du Sud
botswana	2006	4,739367008	9,494558334	0,883035779	48,84000015	0,823774636	-0,198090717	0,723239183	0,642786026	0,225759342	Afrique
burkina faso	2006	3,801490784	7,327224731	0,796405017	49,43999863	0,588338017	0,029684732	0,797700584	0,677514315	0,265572011	Afrique
cambodia	2006	3,568744659	7,746443272	0,793081462	57,63999939	0,750211722	0,253480405	0,829181135	0,652143024	0,341023296	Asie
cameroon	2006	3,851072073	8,089927673	0,689600766	47,84000015	0,653422654	-0,017281633	0,907067657	0,588019311	0,270873696	Afrique
chad	2006	3,434800625	7,369199753	0,724308193	47,08000183	0,3061031899	0,025769988	0,961073756	0,571167529	0,262727261	Afrique
chile	2006	6,062851906	9,869904518	0,835543633	67,77999878	0,744291902	0,163423926	0,633629858	0,75241518	0,347656876	Amérique du Sud

