

ACADEMY

*Analyse exploratoire de données de systèmes
éducatifs pour l'expansion à l'international*

Marion Dedieu

09/2023

01. CONTEXTE
02. JEU DE DONNÉES
03. ANALYSE
04. CONCLUSION

SOMMAIRE

01.

CONTEXTE

Academy
Problématique
Objectifs



EDTech France



ACADEMY

FORMATION EN LIGNE
NIVEAU LYCÉE + UNIVERSITÉ

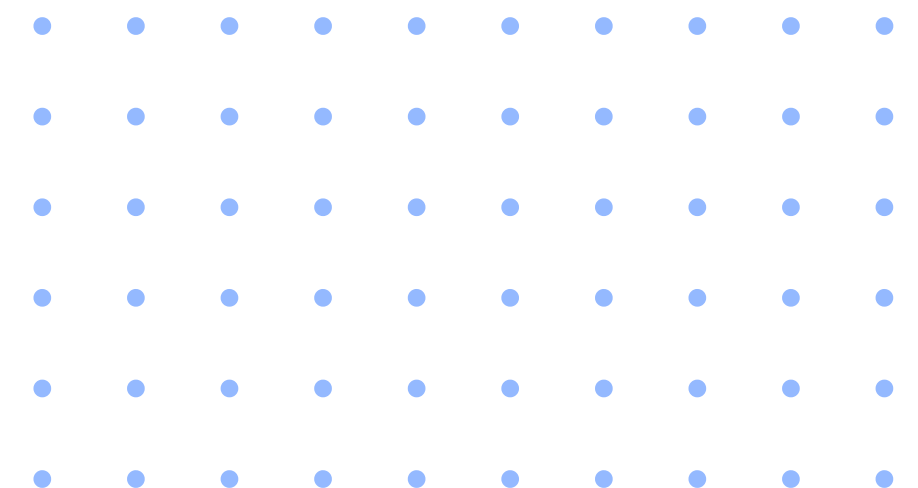


PROBLEMATIQUE

Expansion à l'international

OBJECTIFS

Pré-analyse du jeu de données :
Pertinence / Qualité
Exploration des pays à fort potentiel



02.

JEU DE DONNÉES

Source
Description
Qualité
Indicateurs

SOURCE

Données sur l'éducation de
la Banque mondiale

5 fichiers de données au
format CSV

Essentiel des données :
EdStatsData.csv



WORLD BANK GROUP

DESCRIPTION & QUALITÉ

70 colonnes
886 930 lignes

COLONNES

Pays
Indicateurs
Relevés annuels

TYPES

4 colonnes Texte
66 colonnes Numérique

VALEURS

242 pays/régions
3665 indicateurs

COMPLÉTUDE

13.90%

PAS DE DOUBLONS

INDICATEURS HISTORIQUES

GDP PER CAPITA, PPP
(CURRENT INTERNATIONAL \$)

Capacité financière moyenne des populations.
Influence sur la demande et la tarification des services.

INTERNET USERS
(PER 100 PEOPLE)

Accès à Internet essentiel pour l'apprentissage en ligne.

PERSONAL COMPUTERS
(PER 100 PEOPLE)

Ordinateur personnel indispensable pour la formation en ligne.

% OF STUDENTS IN SECONDARY
EDUCATION, GENERAL PROGRAMMES

Demande potentielle pour le niveau lycée (voie générale).

% OF STUDENTS IN SECONDARY
EDUCATION, VOCATIONAL PROGRAMMES

Demande potentielle pour le niveau lycée (voie professionnelle).

% OF ALL STUDENTS IN TERTIARY
EDUCATION, IN ISCED 5

Demande potentielle pour l'enseignement supérieur.

INDICATEURS PROSPECTIFS

% OF THE TOTAL POPULATION BY
HIGHEST LEVEL OF EDUCATIONAL
ATTAINMENT, POST SECONDARY

Population potentielle intéressée par des cours en ligne de niveau universitaire

MEAN YEARS OF SCHOOLING, AGE 15–19

L'augmentation de cette valeur suggère une amélioration de l'accès à l'éducation pour les jeunes.

MEAN YEARS OF SCHOOLING, AGE 20–24

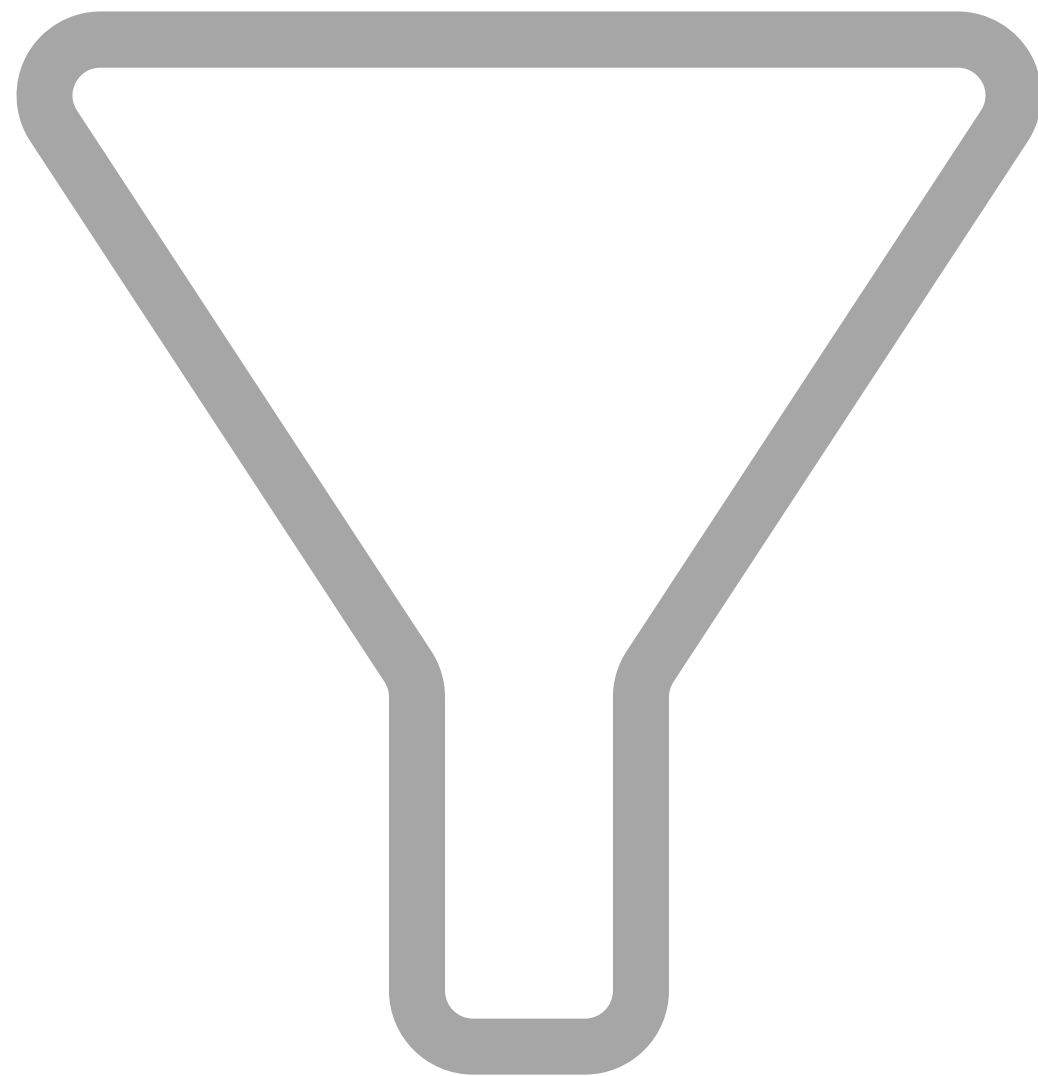
L'augmentation de ce chiffre peut indiquer un accès accru à l'éducation supérieure

03.

ANALYSE

*Méthode
Résultats*

MÉTHODE DE CLEANING



Filtre sur les indicateurs choisis

Indicateurs historiques et prospectifs

Filtre sur taux de remplissage > 2%

Retrait de 41 pays

Création d'un dataframe historique

1970 à 2017

Dernière valeur non nulle

Conservation de la dernière valeur historique

EDA

Données historiques

Données prospectives

Analyse

Moyenne
Médiane
Ecart-type
Min, Max
Q1, Q3
Boîte à moustaches

Scoring

Qcut
Score de 0 à 1
Corrélations
Score global pondéré

Ranking

Score min > 60%
Top 15

Analyse

(Sur 2015)
Moyenne
Médiane
Ecart-type
Min, Max
Q1, Q3
Boîte à moustaches

Evolution

2015–2030

Ranking

Sur évolution
moyenne

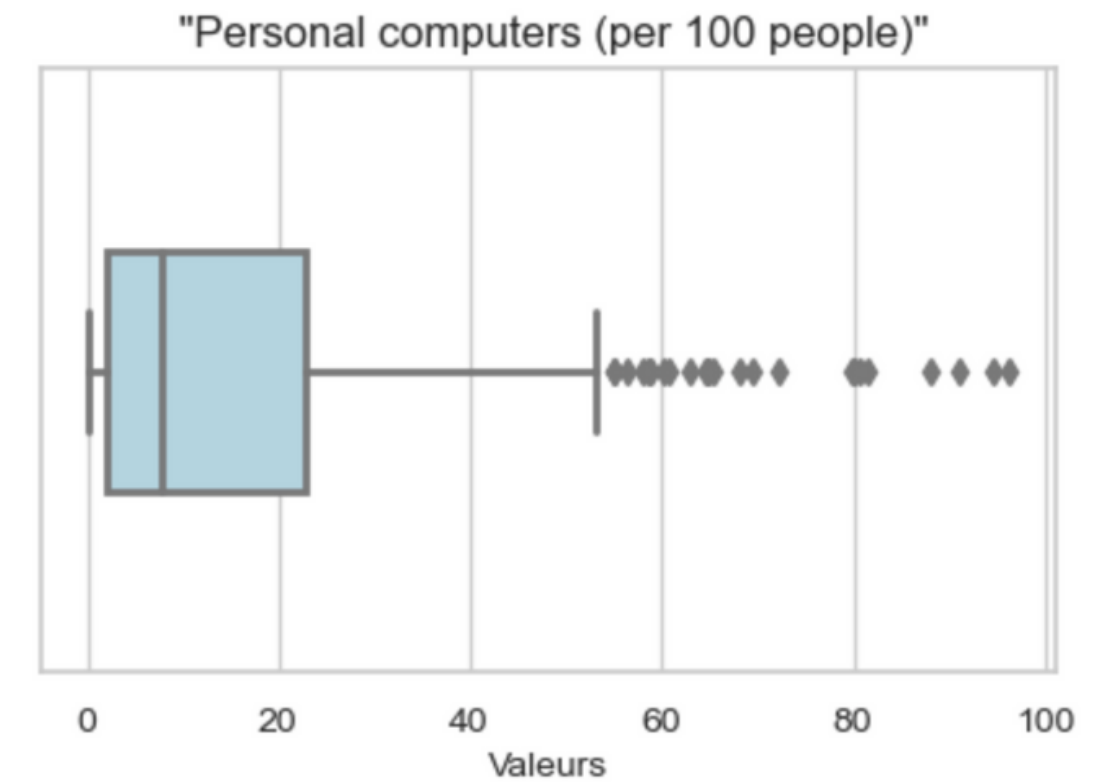
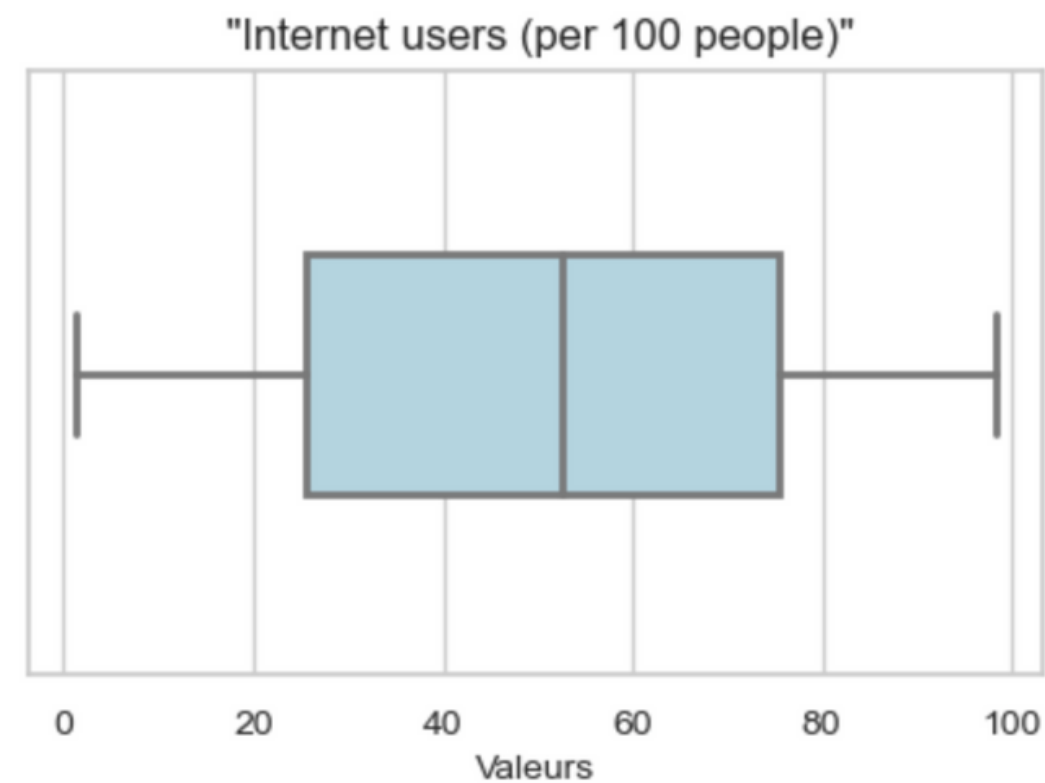
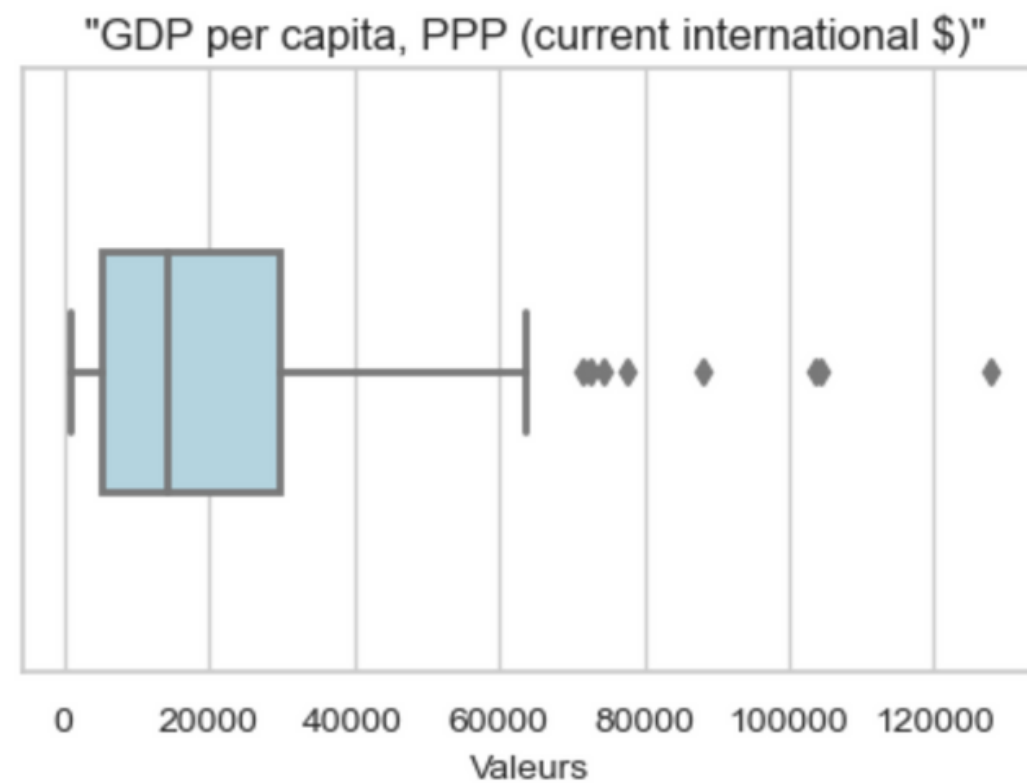
DONNÉES HISTORIQUES

INDICATEURS HISTORIQUES

PIB (PPA)

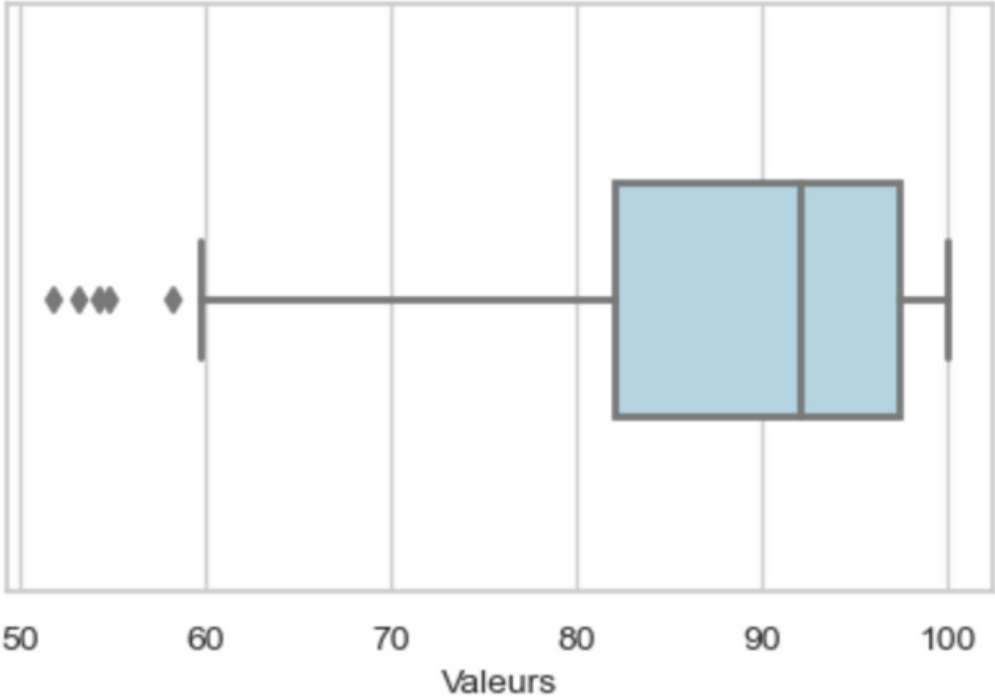
% Internet

% Ordinateurs



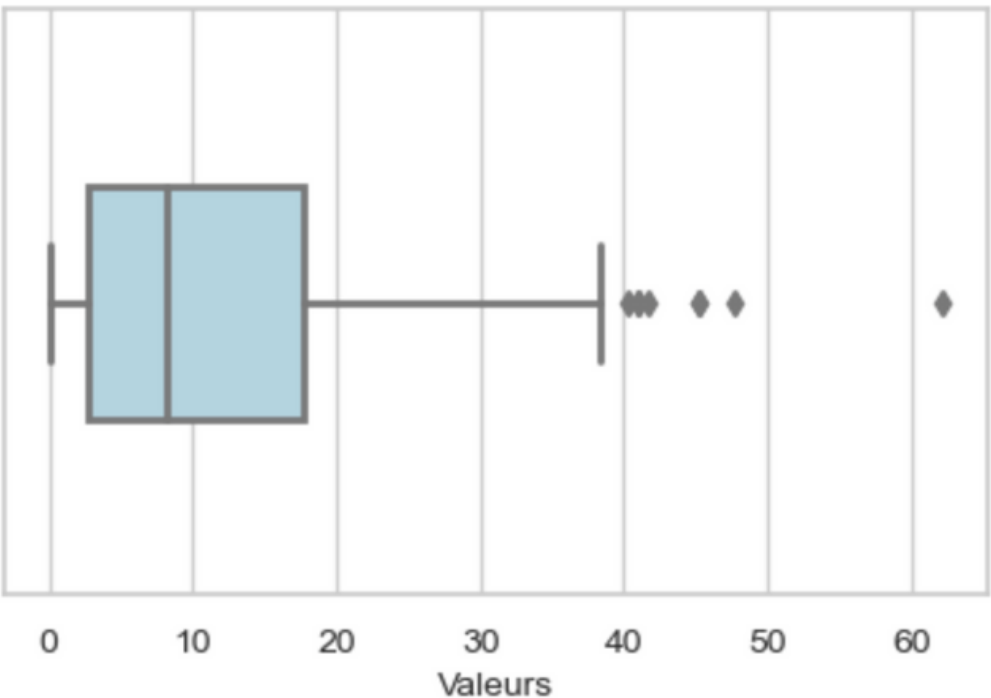
Lycée voie générale

"Percentage of students in secondary education enrolled in general programmes, both sexes (%)"



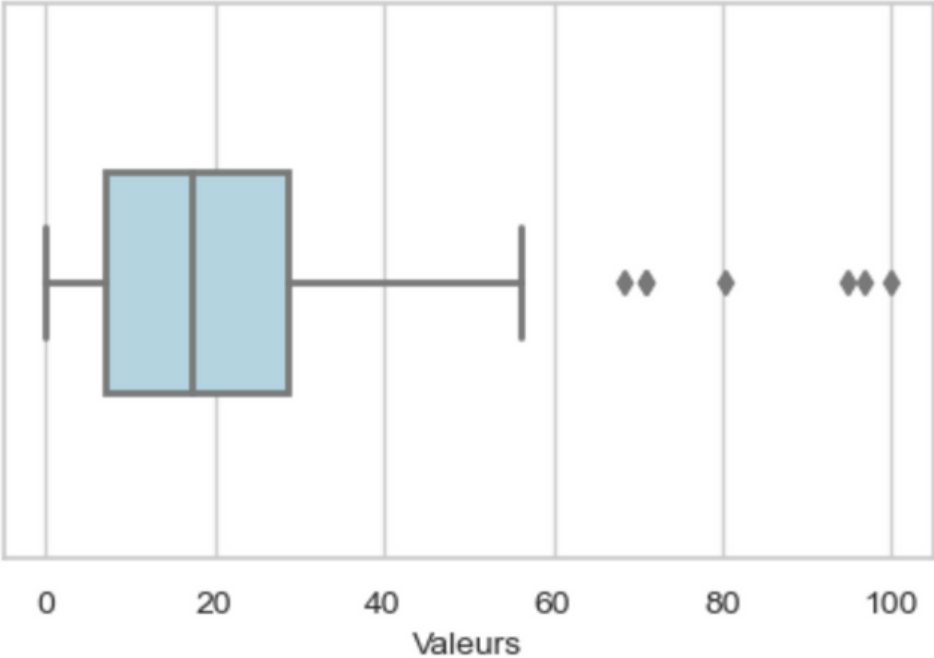
Lycée voie professionnelle

"Percentage of students in secondary education enrolled in vocational programmes, both sexes (%)"

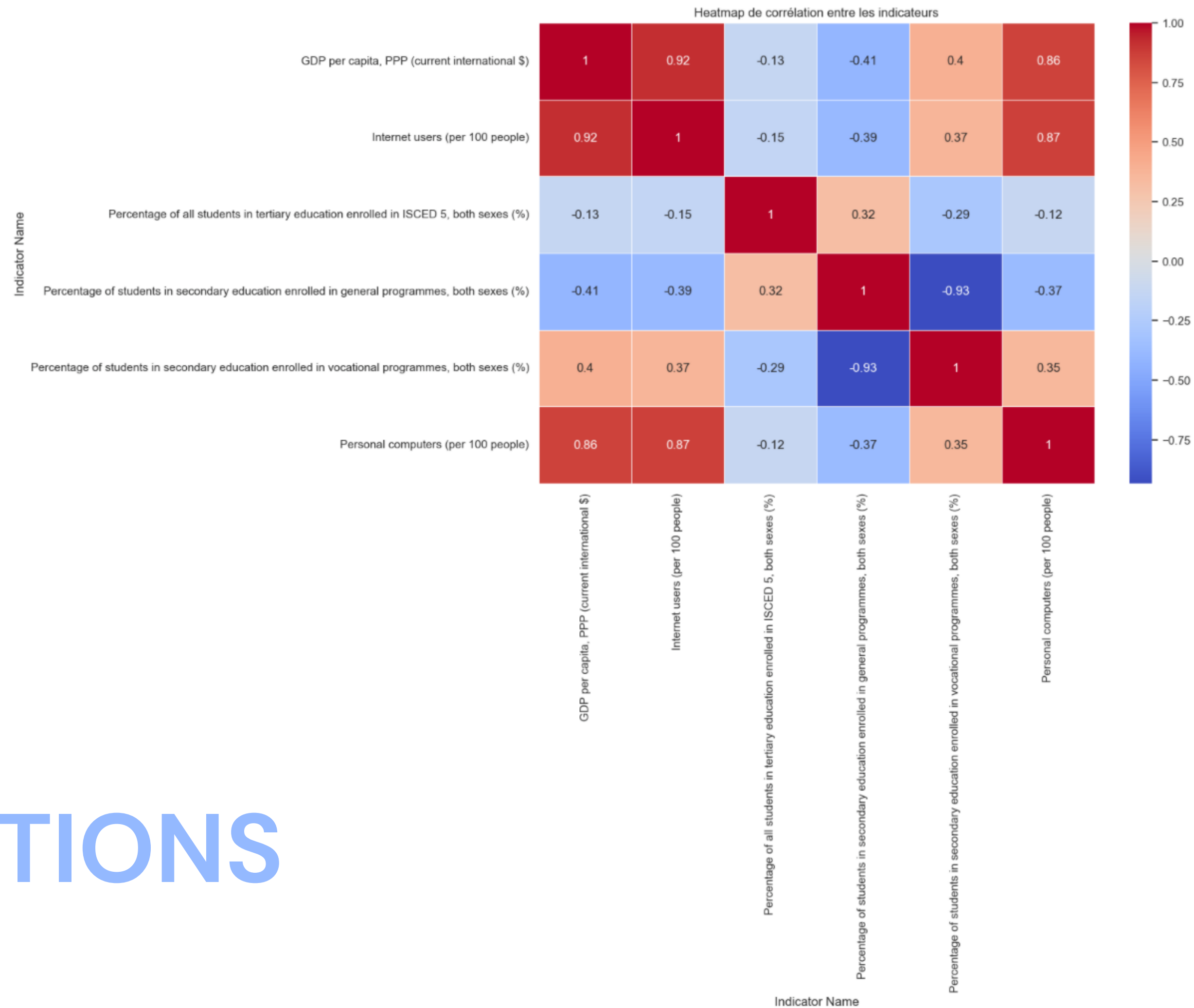


Enseignement supérieur

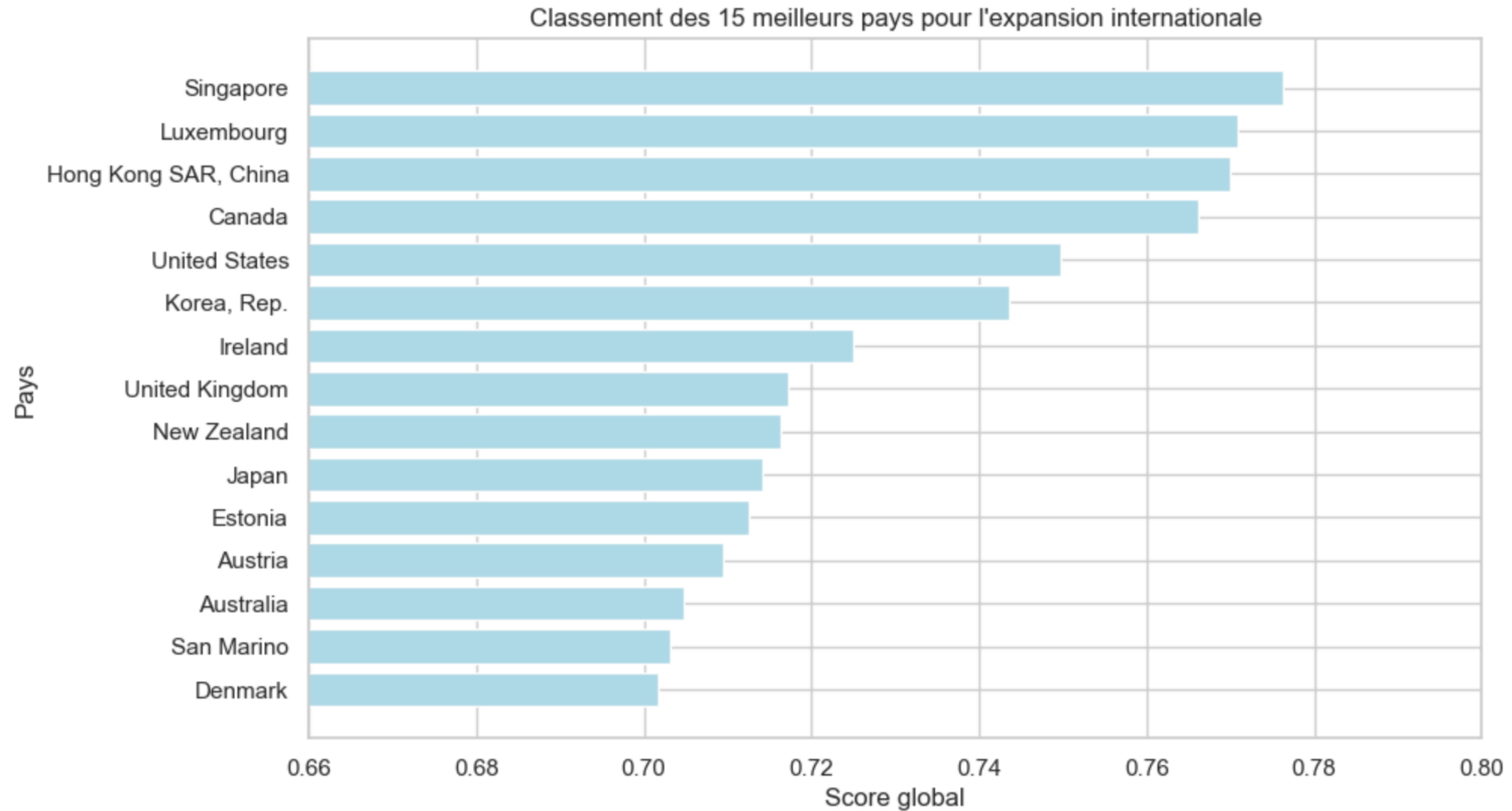
"Percentage of all students in tertiary education enrolled in ISCED 5, both sexes (%)"



CORRÉLATIONS



CLASSEMENT POUR L'EXPANSION



RÉGIONS POUR L'EXPANSION

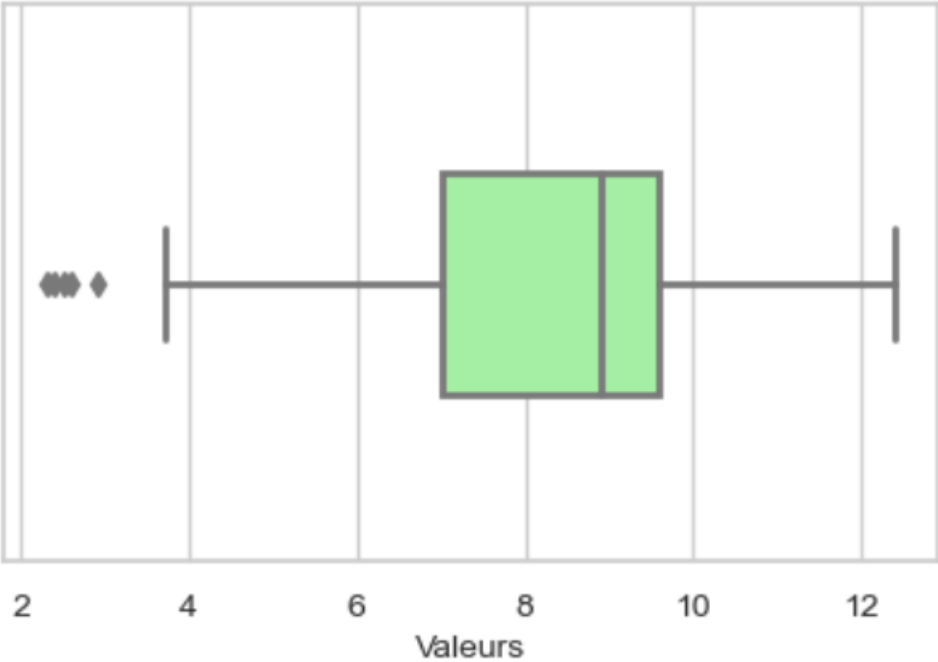
Top 15 des meilleurs pays pour l'expansion internationale



DONNÉES PROSPECTIVES

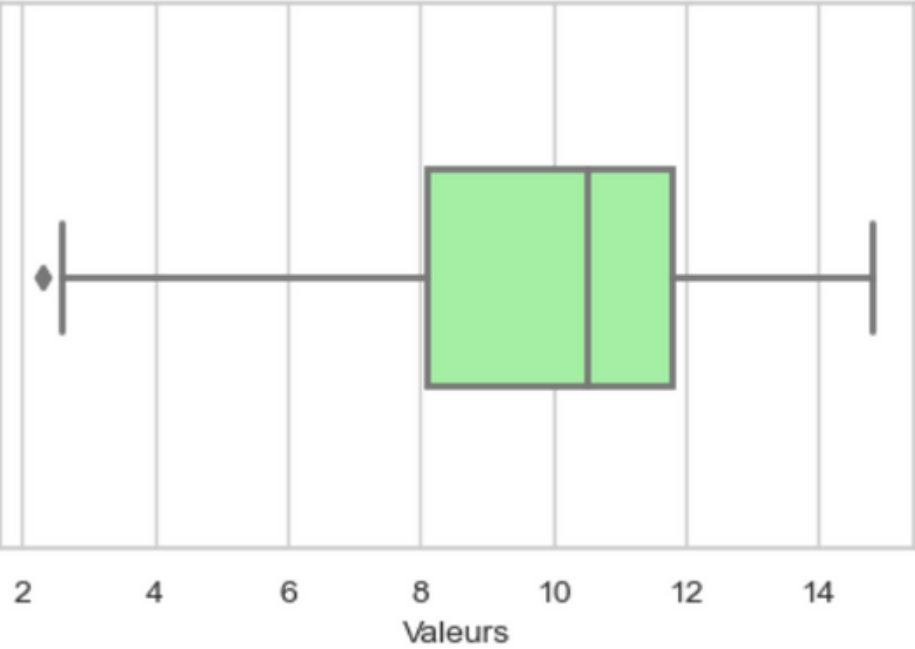
Durée scolarisation (15–19 ans)

"Wittgenstein Projection: Mean years of schooling. Age 15-19. Total"



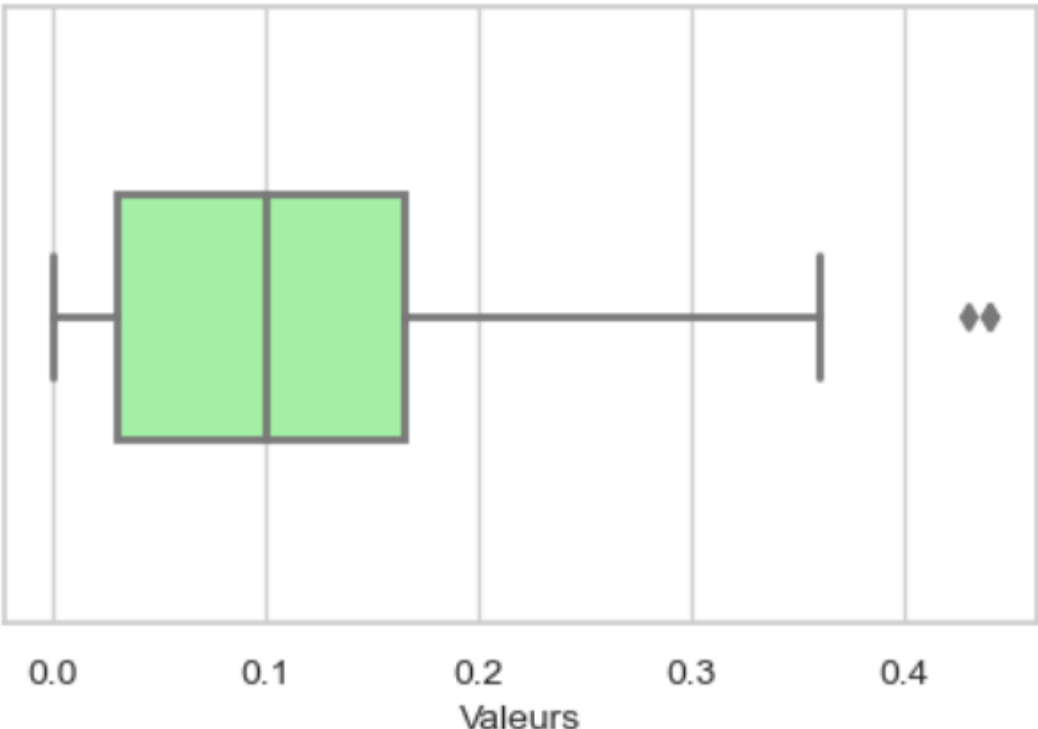
Durée scolarisation (20–24 ans)

"Wittgenstein Projection: Mean years of schooling. Age 20-24. Total"

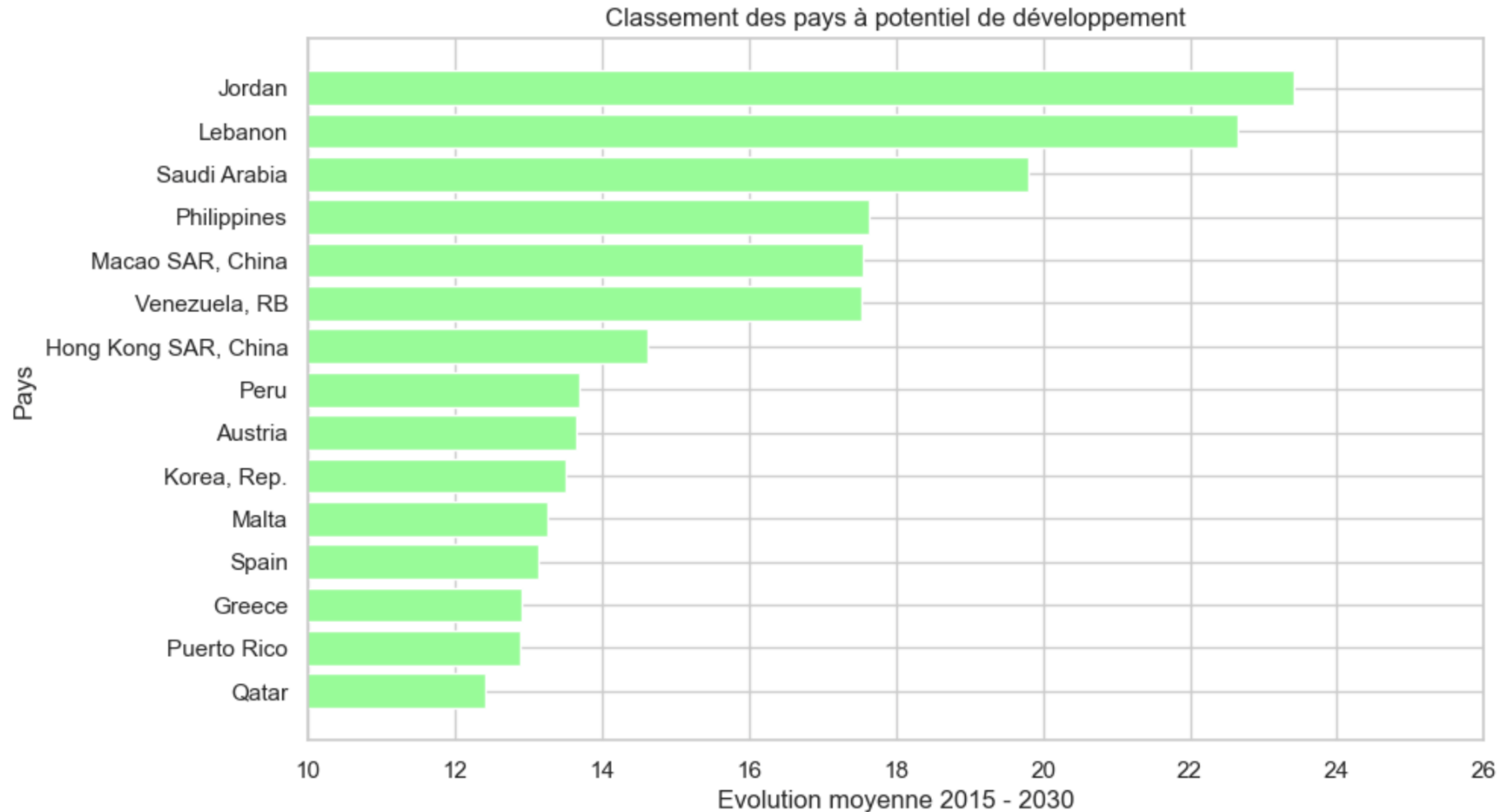


% population qui a atteint le niveau universitaire

"Wittgenstein Projection: Percentage of the total population by highest level of educational attainment. Post Secondary. Total"

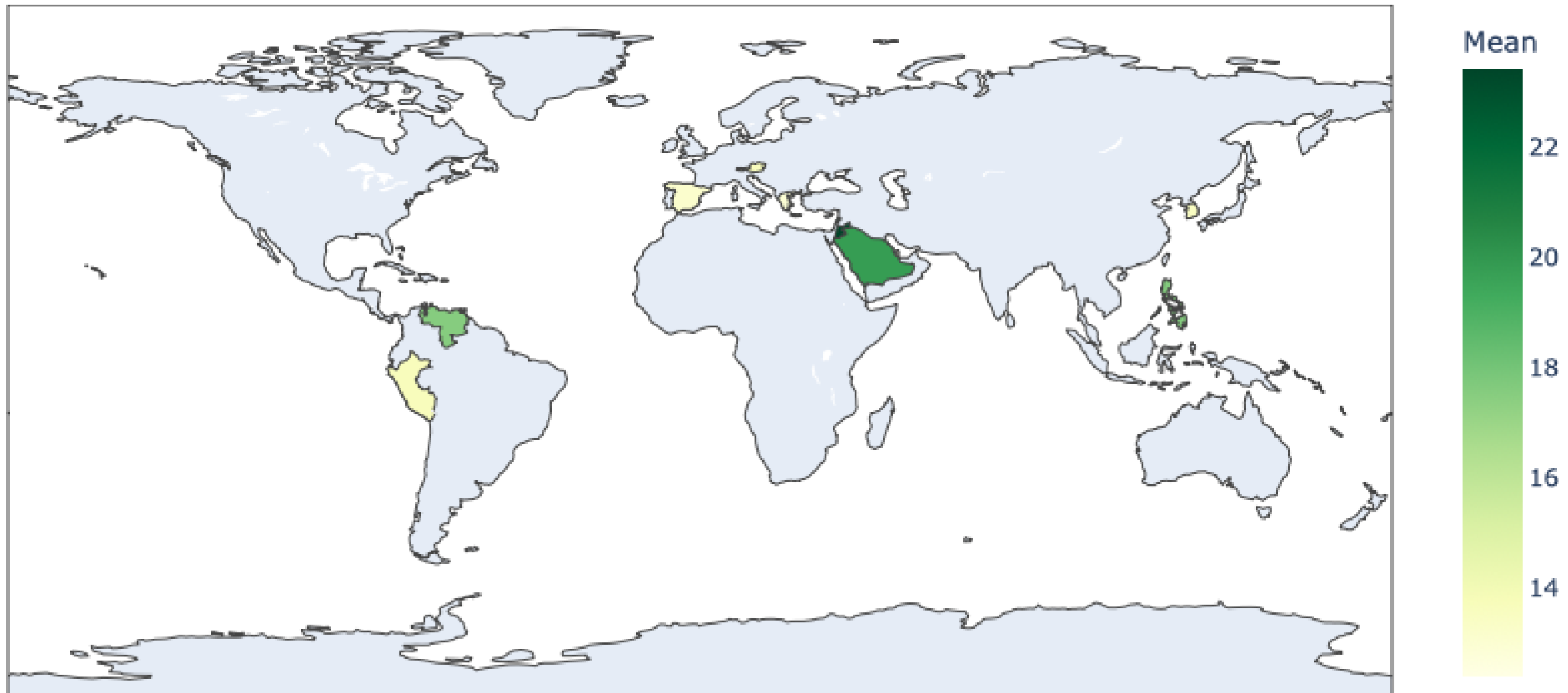


CLASSEMENT PAYS À POTENTIEL



RÉGIONS À POTENTIEL

Top 15 des pays à potentiel de développement

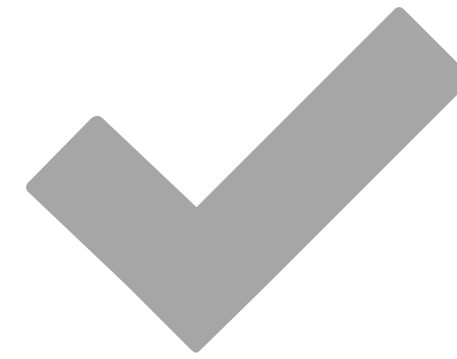


04.

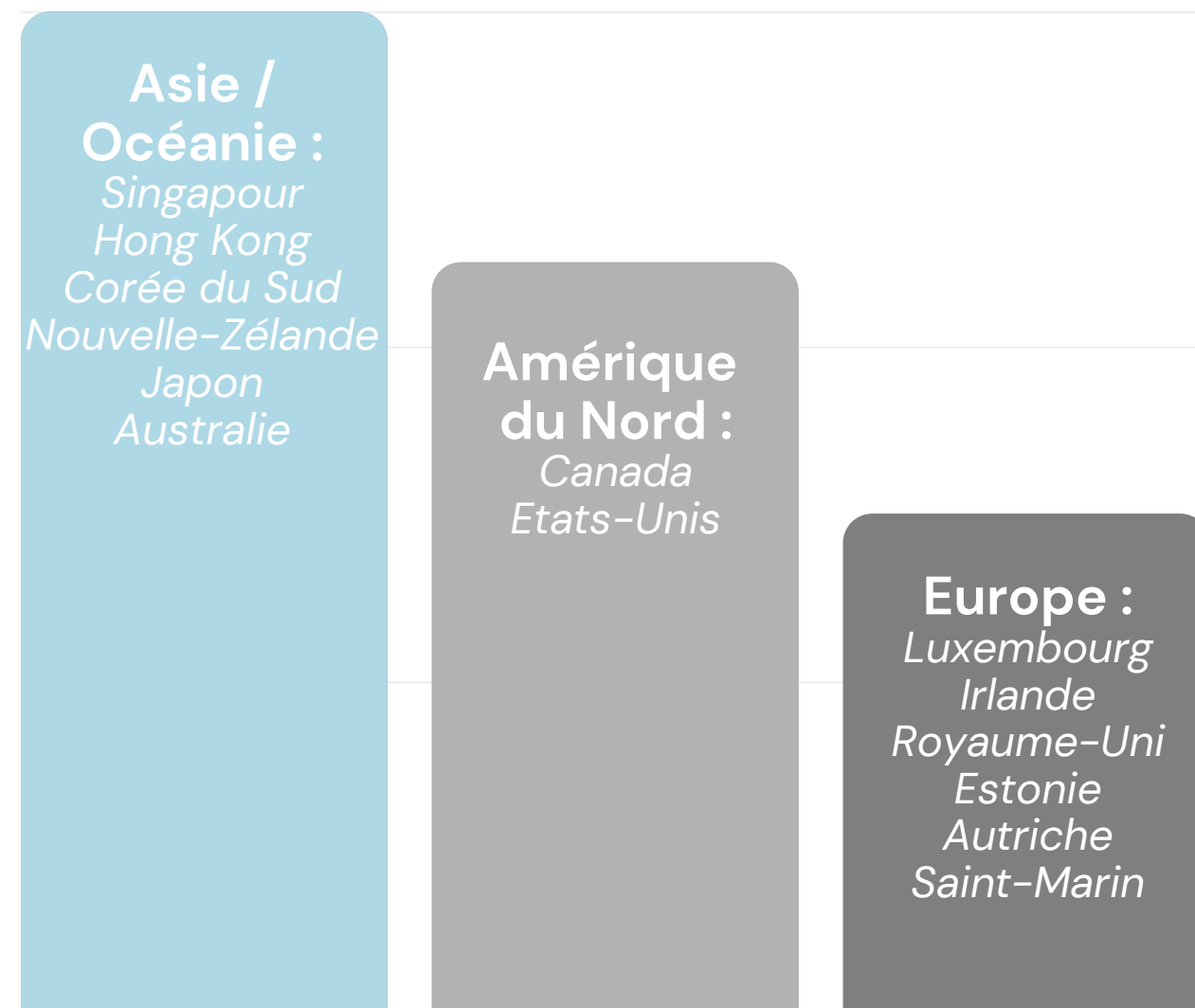
CONCLUSION

*Pertinence du jeu de données
Nuances sur l'analyse*

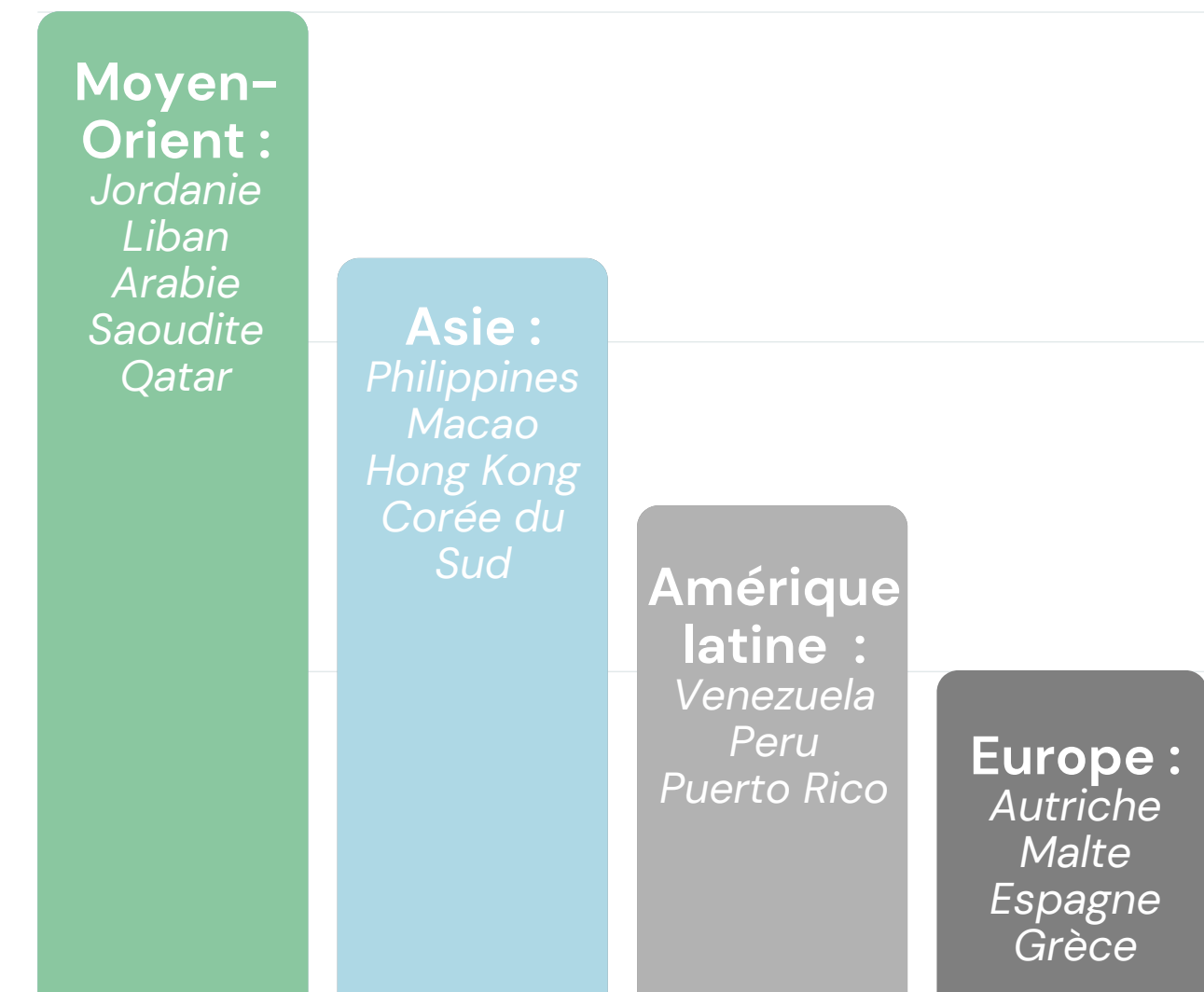
DATA SET PERTINENT

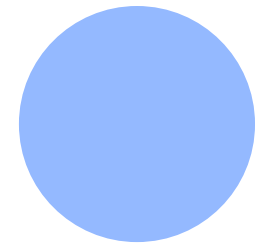


TOP 3 EXPANSION

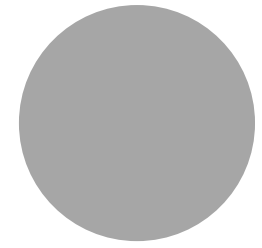


TOP 4 POTENTIEL

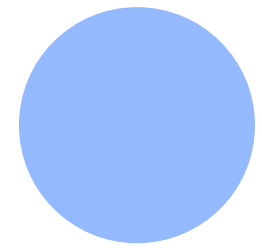




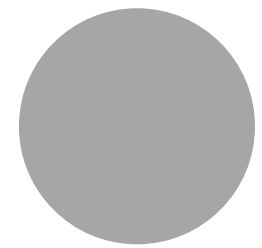
LANGUES DE FORMATION
/ FUSEAUX HORAIRES



MÉTHODE DES DERNIÈRES
VALEURS NON NULLES



CHOIX DES INDICATEURS



MÉTHODE DE SCORING

NUANCES
SUR L'ANALYSE

MERCI !

Des questions ?