

Projet 2

# Analysez des données de systèmes éducatifs

Mauyves NKONDO

Data Scientist



# Problématique :

Analyser les données de la Banque mondiale pour déterminer quels pays possèdent les principaux potentiels, l'extension actuelle et future de l'Académie et déterminer dans quels pays opérer en priorité.

## Questions résolues

- Quels sont les pays avec un fort potentiel de clients pour nos services ?
- Pour chacun de ces pays, quelle sera l'évolution de ce potentiel de clients ?
- Dans quels pays l'entreprise doit-elle opérer en priorité ?

# Introduction :

## **Partie 1 : Importation des données**

### **a : Importation des librairies**

## **Partie 2 : Nettoyage des données**

### **a : Traitement de la base de donnée EdStats Country-Series.csv**

### **b : Traitement de la base de donnée EdStats Country.csv**

### **c : Traitement de la base de donnée EdStats data.csv**

### **d : Traitement de la base de donnée EdStats FootNote.csv**

### **e : Traitement de la base de donnée EdStats Series.csv**

## **Partie 3 : Sélections des Indicateurs**

### **a : Choix et Présentation des Indicateurs retenus**

## **Partie 4 : Score**

## **Partie 5 : Restituer les données**

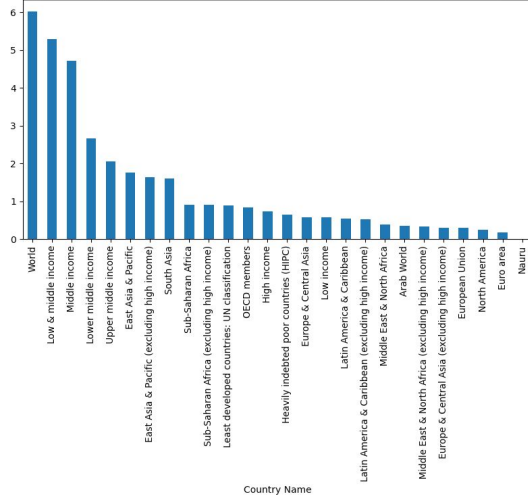
### **a : donnée régionale**

### **b : Listes des payes retenus et leurs scores**

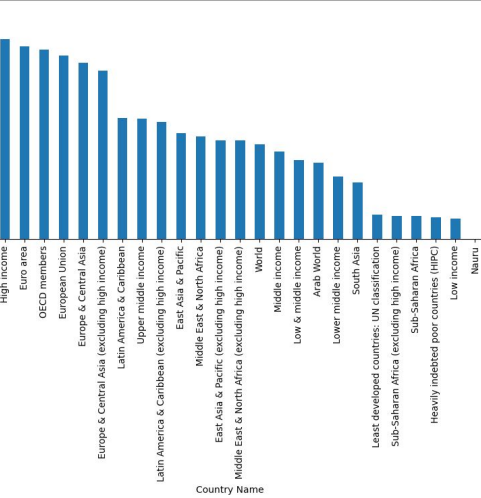
## **Partie 6 : Pays en forte croissance**

## **Partie 7 : Conclusion**

Population de l'age officiel pour l'enseignement superieur, les deux sexes (nombre)



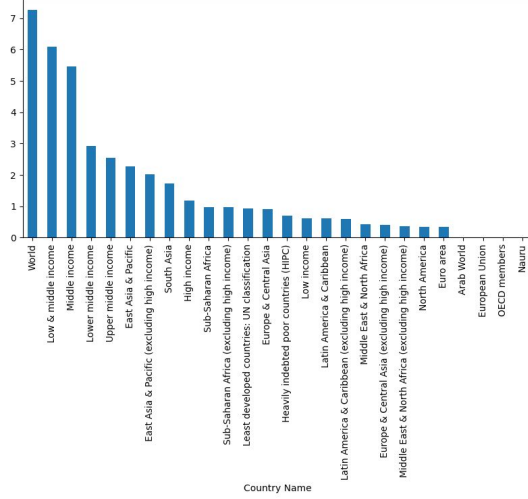
Taux brut de scolarisation tertiaire, hommes et femmes(%)



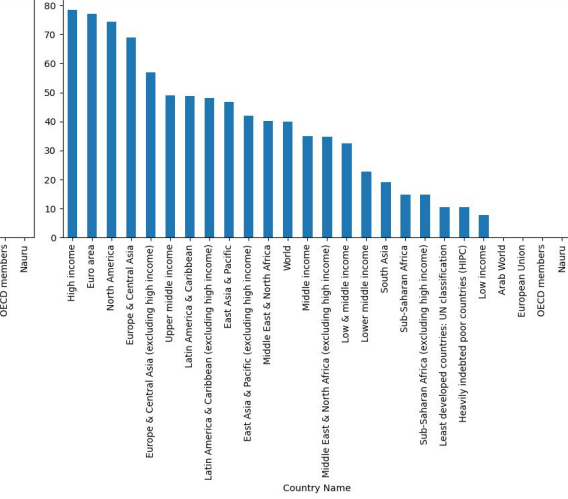
Sur cette présentation des quatre (4) indicateurs :

- **SP.TER.TOTL.IN** ( Population de l'age officiel pour l'enseignement superieur )
- **SE.TER.ENRR** ( Taux brut de scolarisation tertiaire )
- **UIS.E.3** ( Population totale dans l'enseignement secondaire )
- **UIS.E.5.B** ( Population totale dans l'enseignement supérieur )

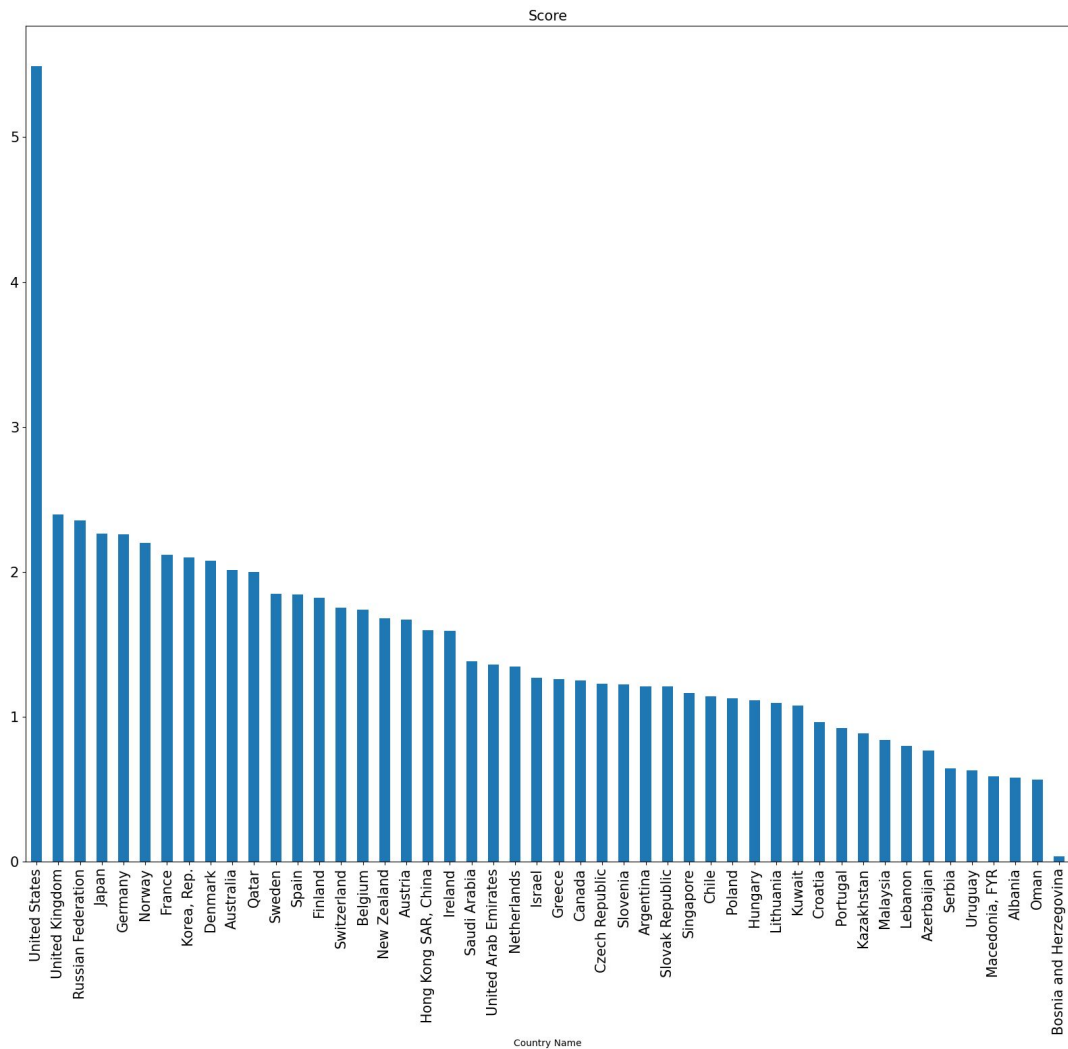
Perspectives de la population mondiale aux nations unies



Population âgée de 15 à 74 ans au cours des 3 derniers mois. - Internet dual-up



Nous allons analyser leur évolution sur les graphes de matplotlib.



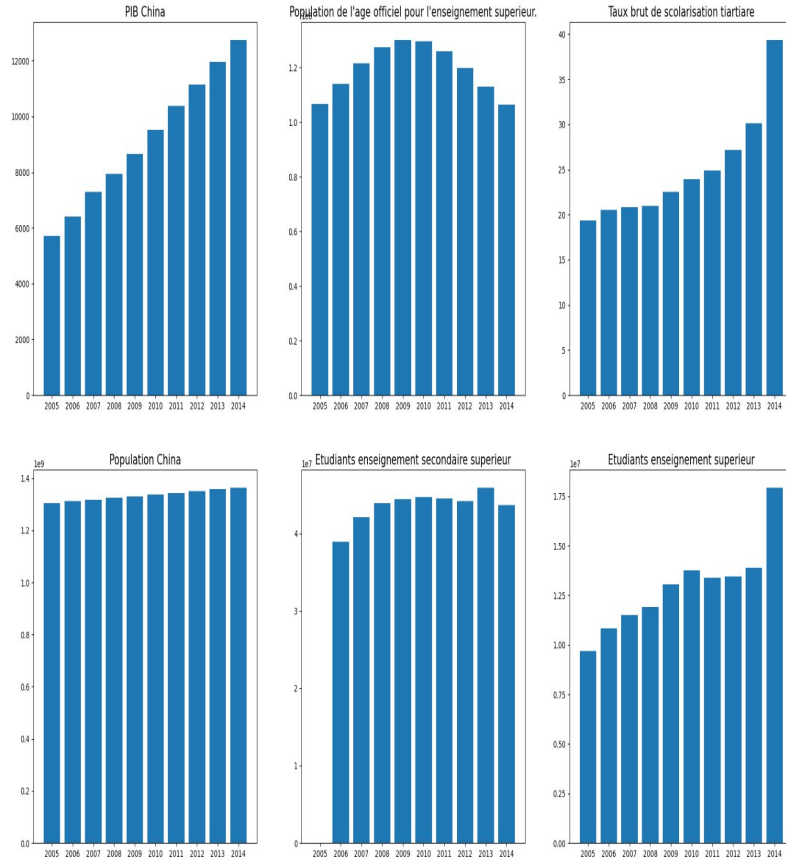
Sur la base de l'analyse réalisée plus haut sur l'échelle graphique des pays choisis, nous présentons les Etats-Unis est en tête pour la seconde place trouve le Royaume-Uni et en troisième place trouve la Fédération de Russie.

L'entreprise doit d'abord investir aux Etats-Unis, puis voir les pays du Royaume Uni et ensuite les pays représentés sur le graphique.

Ce plan aura des retombées positives pour la société.

La Chine est en forte croissance avec sa population et son évolution sur les graphiques matplotlib.

Depuis les années d'apparition de l'Internet et des réseaux sociaux, la Chine connaît une croissance rapide et sera rentable sur l'investissement de nos jours.



# Conclusion :

Le jeu de données aide dans la prise de décision à formuler des recommandations à l'entreprise EdTech selon notre étude que nous venons de réaliser.

- Les États-Unis se classent au premier rang pour la deuxième place, le score prédit pour le Royaume-Uni est plus élevé que celui de la Chine. Le potentiel du taux de scolarisation dans l'enseignement secondaire est favorable au Royaume-Uni, en stagnation dans l'enseignement supérieur alors que la tendance est en croissance pour la Chine. Le Royaume-Uni sera privilégié.
- Les années antérieures à 1990 sont moins pleines, et nous ne sommes pas intéressés parce qu'Internet n'a fait son apparition qu'en 1991.
- 
- Des années après 2015 avaient déjà été retirées, c'est-à-dire que nos indicateurs étaient vides. Cependant, nous excluons également 2015, qui est beaucoup moins bien informé qu'en 2014.

En outre, compte tenu de l'évolution rapide des nouvelles technologies, il peut être utile de garder les années antérieures à 2005 (apparition de YouTube, Facebook, la montée des réseaux sociaux et des plateformes de diffusion en continu) parce qu'ils ne sont pas représentatifs des nouvelles technologies, et le risque.

Par conséquent, nos indicateurs "technologiques" (accès Internet et accès aux ordinateurs) diminuent.

NB : Seulement des années après 2005 (inclus) et avant 2015 (exclus) sont retenues.