

# PROJET 2 Analysez des données de systèmes éducatifs

Emmanuel LEPLANOIS



# Problématique



- ▶ **Mark**, votre manager, vous a convié à une réunion pour vous présenter le projet d'**expansion à l'international** de l'entreprise. Il vous confie **une première mission d'analyse exploratoire**, pour déterminer si les données sur l'éducation de la banque mondiale permettent d'informer le projet d'expansion.
- ▶ Voici les différentes questions que Mark aimerait explorer, que vous avez notées durant la réunion :
  - 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 ?

# Jeu de données

- Issu de la Banque Mondiale :

- <https://datacatalog.worldbank.org/dataset/education-statistics>

Disponible en téléchargement direct à [https://s3-eu-west-1.amazonaws.com/static.oc-static.com/prod/courses/files/Parcours\\_data\\_scientist/Projet+-+Donn%C3%A9es+%C3%A9ducatives/Projet+Python\\_Dataset\\_Edstats\\_csv.zip](https://s3-eu-west-1.amazonaws.com/static.oc-static.com/prod/courses/files/Parcours_data_scientist/Projet+-+Donn%C3%A9es+%C3%A9ducatives/Projet+Python_Dataset_Edstats_csv.zip)

Le fichier zip contient 5 fichiers csv :

- EdStatsData.csv

- EdStatsFootNote.csv

- EdStatsSeries.csv

- EdStatCountry.csv

- EdStatsCountry-Series.csv

Nous allons principalement utiliser le fichier EdStatsSeries qui contient les données par indicateur et par pays.

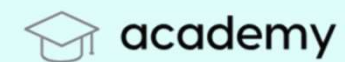
# Qualité des données



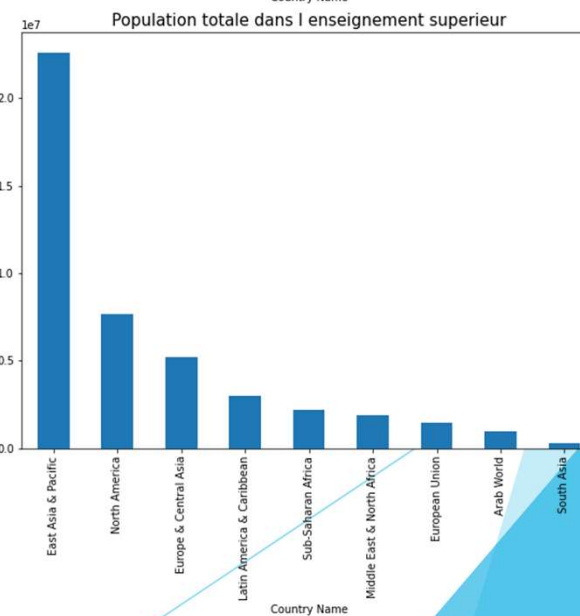
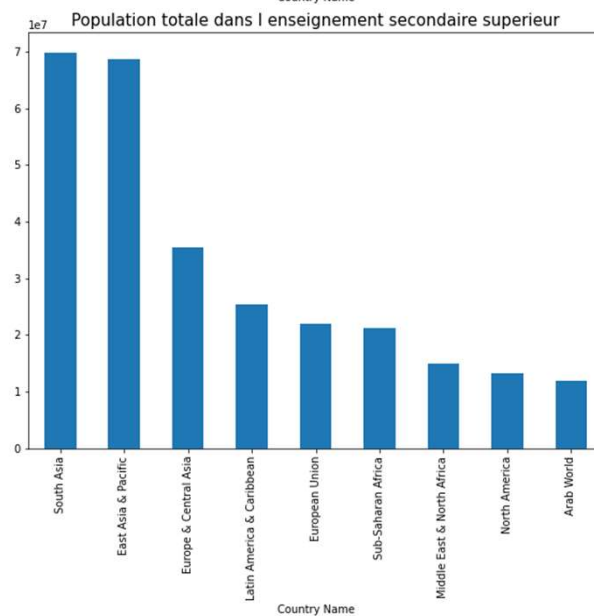
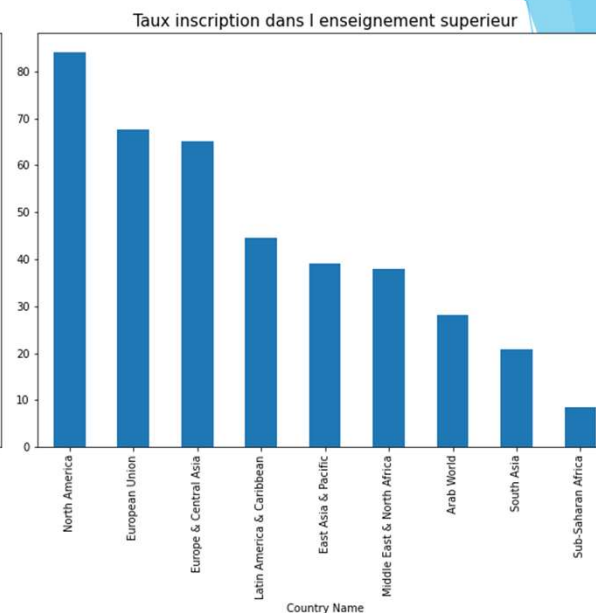
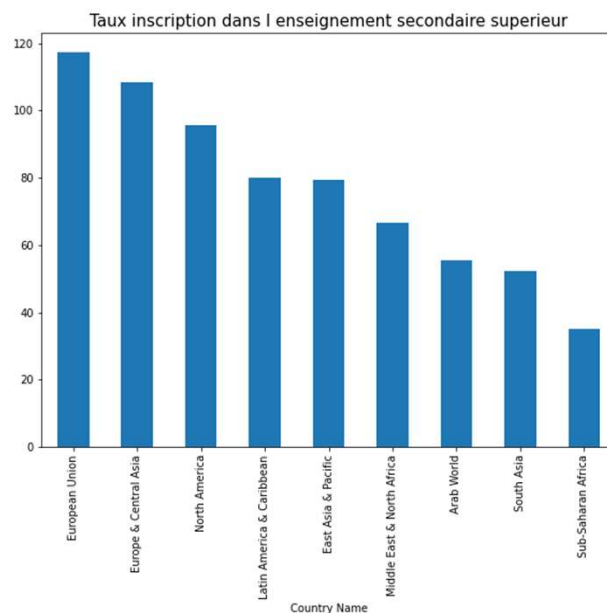
- ▶ Beaucoup de lignes ne sont pas renseignées (33% aucune valeur)
- ▶ On choisit de garder les données que sur dix ans donc de 2005 à 2015
- ▶ Les données pour 2017 et 2020 sont peu valorisées
- ▶ Sélection de pays ayant un PIB/hab > 10 000, taux internet > 60% et population > 2 000 000

# Choix des indicateurs

UIS.E.5.B	Etudiant-es dans premier cycle enseignement supérieur
UIS.E.3	Etudiant-es dans enseignement secondaire supérieur
NY.GDP.PCAP.PP.KD	PIB par habitant (en \$ constant 2011)
SE.TER.ENRR	Taux inscription enseignement supérieur (%)
SE.SEC.ENRR.UP	Taux inscription enseignement secondaire supérieur (%)
IT.NET.USER.P2	utilisateur Internet pour 100 personnes
SP.POP.TOTL	Population totale

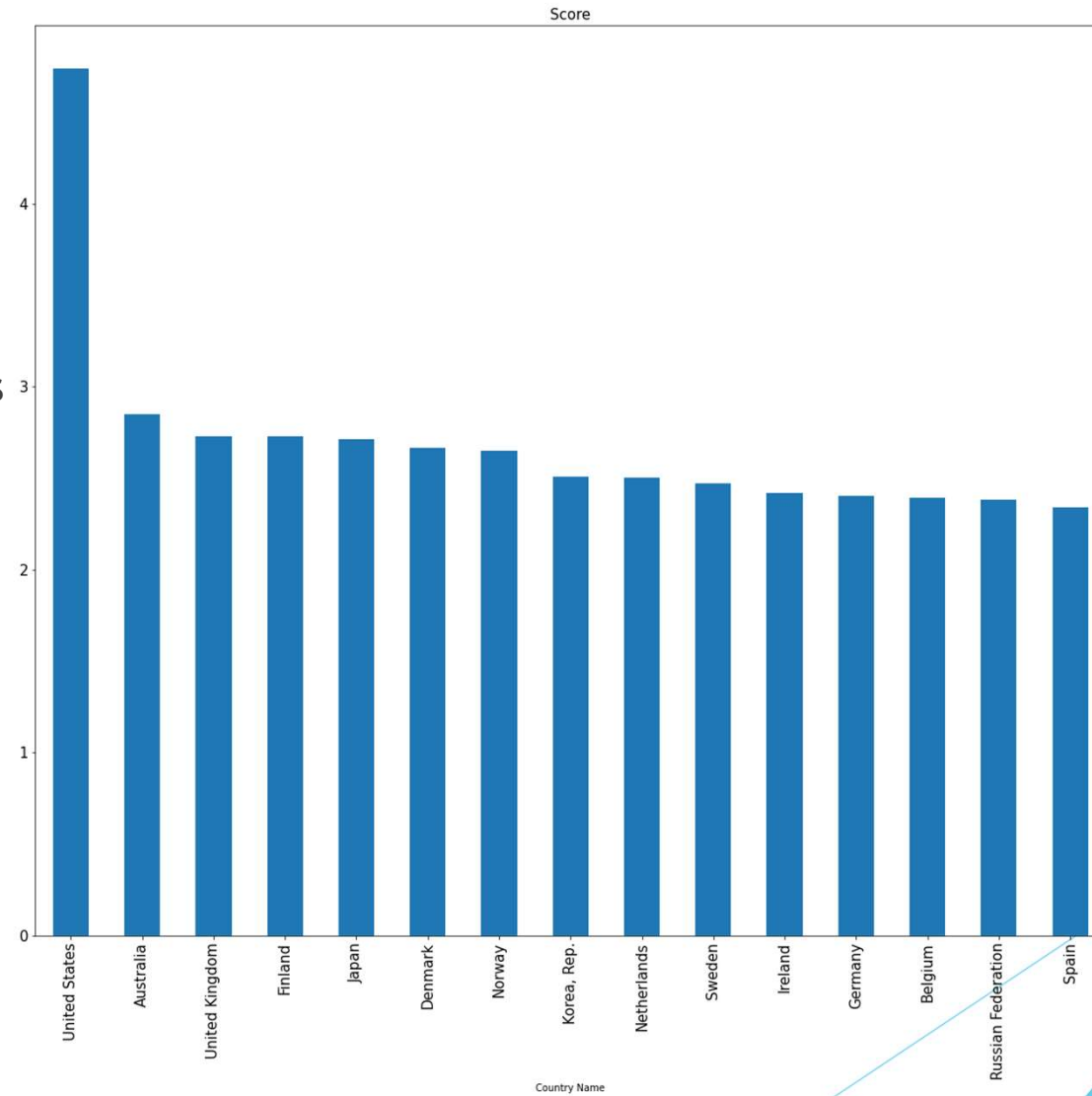


# Indicateurs régionaux



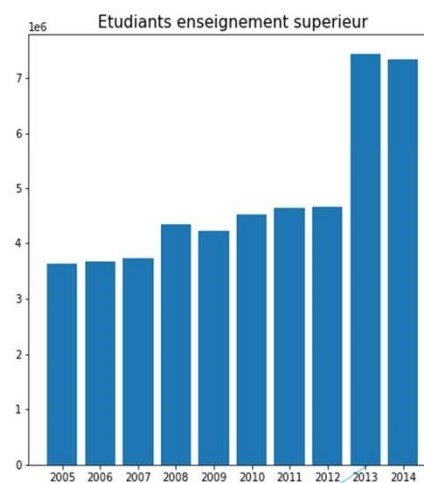
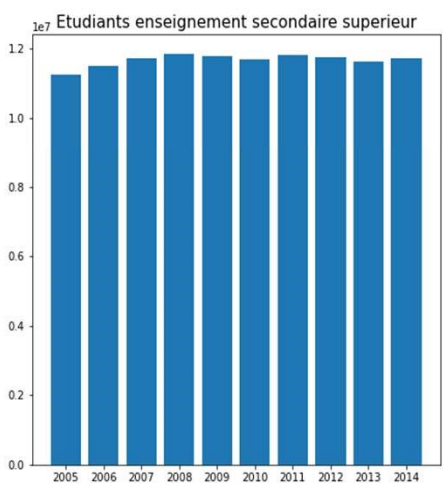
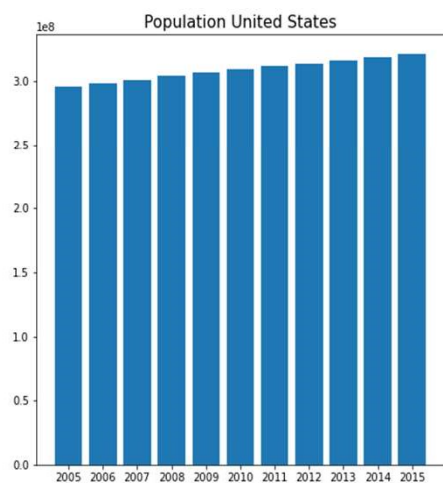
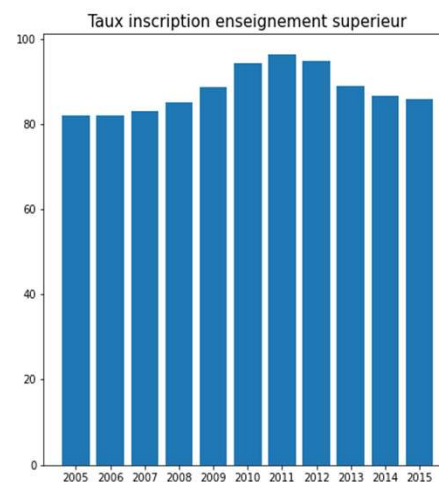
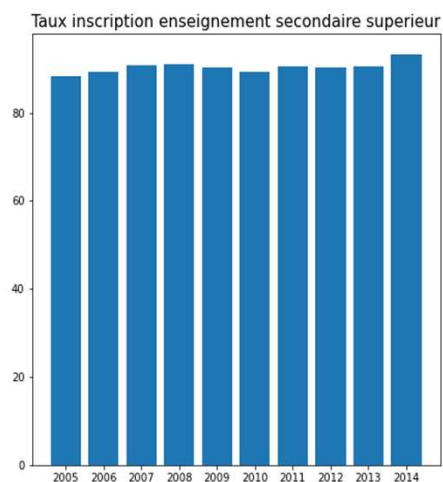
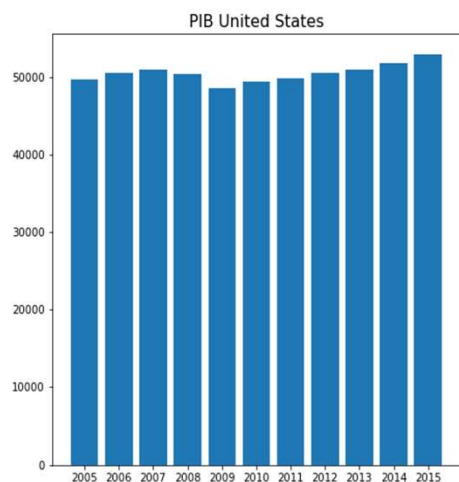
# Pays

- Création d'un score a partir des indicateurs normalisés.



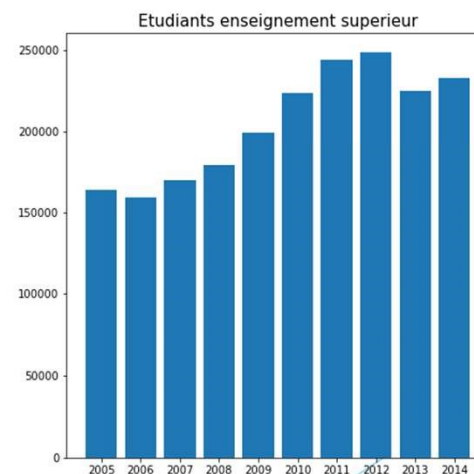
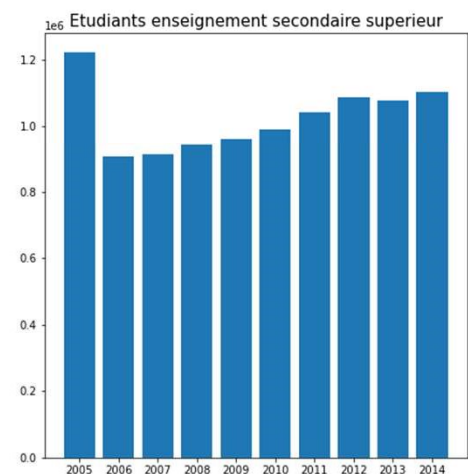
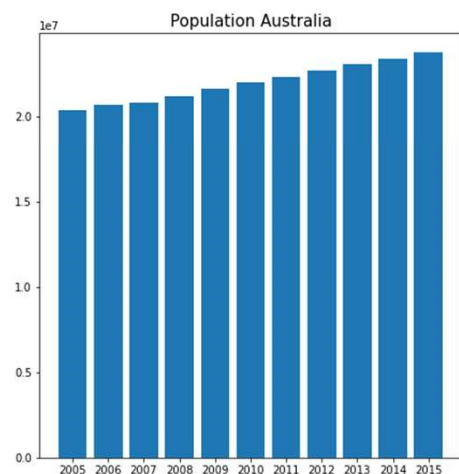
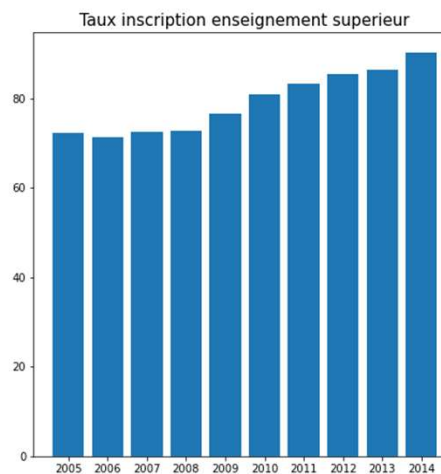
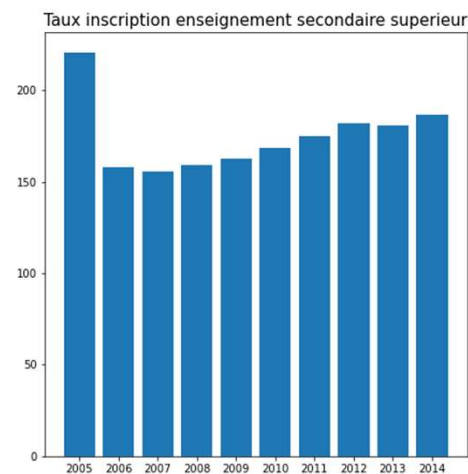
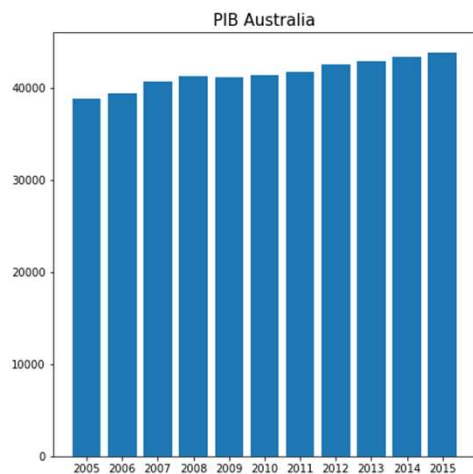


# Evolution indicateur USA

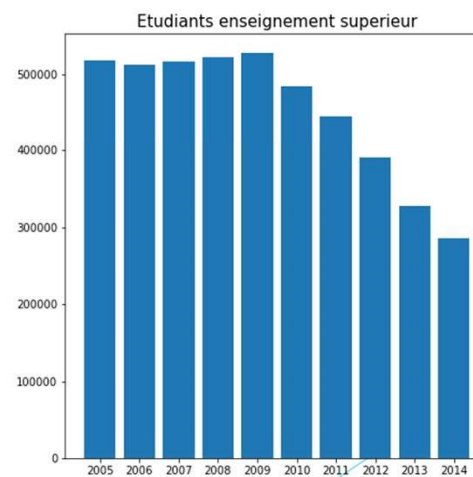
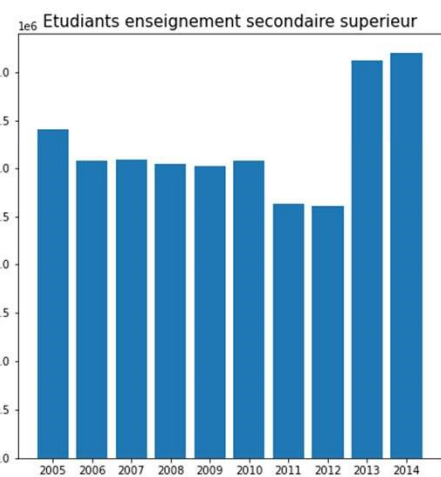
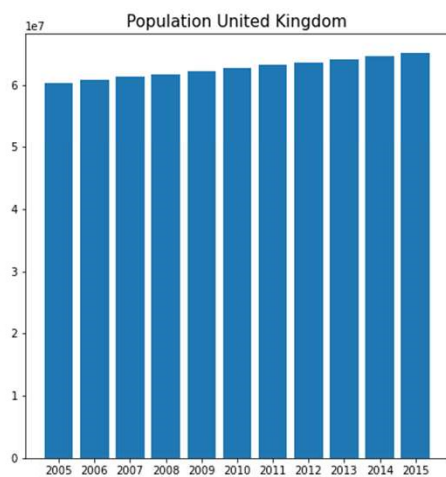
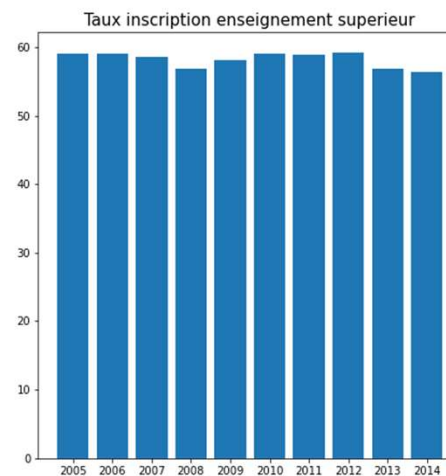
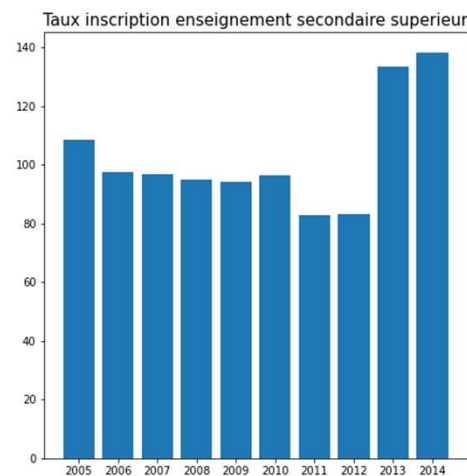
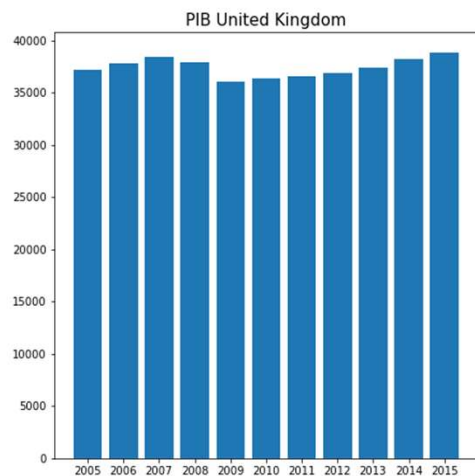




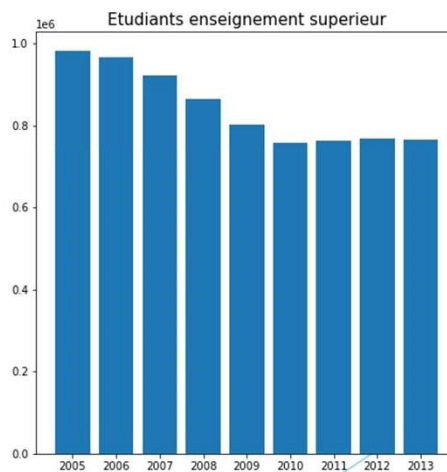
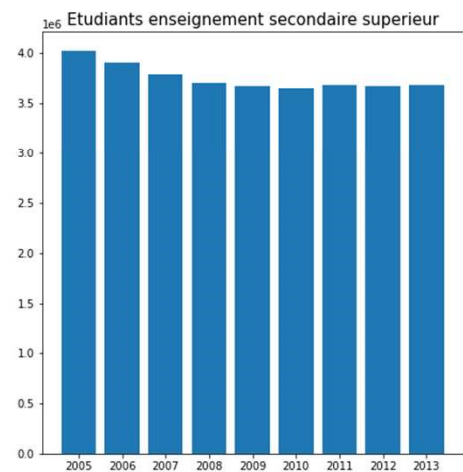
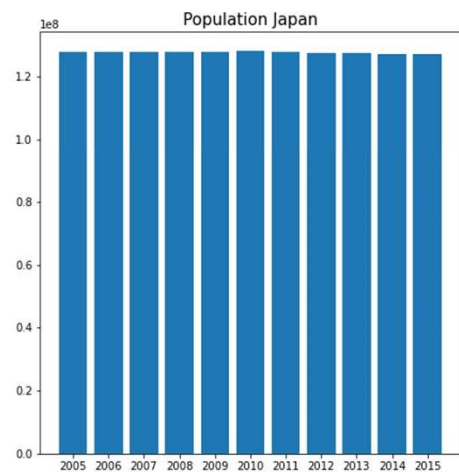
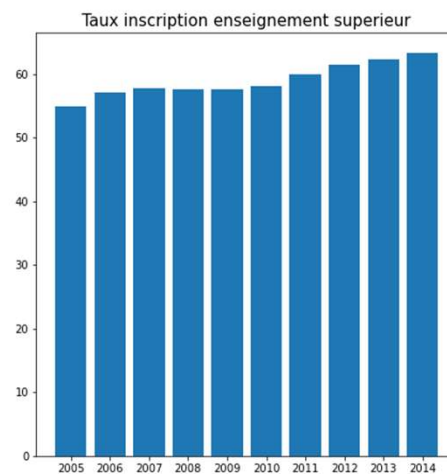
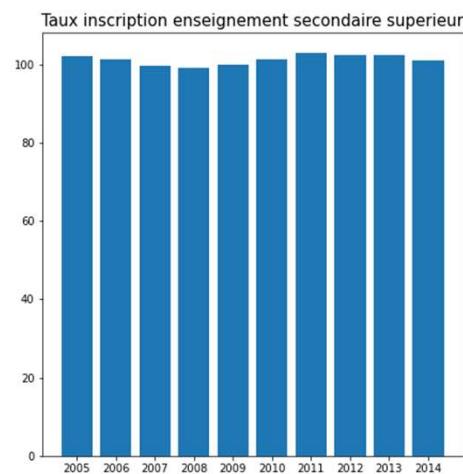
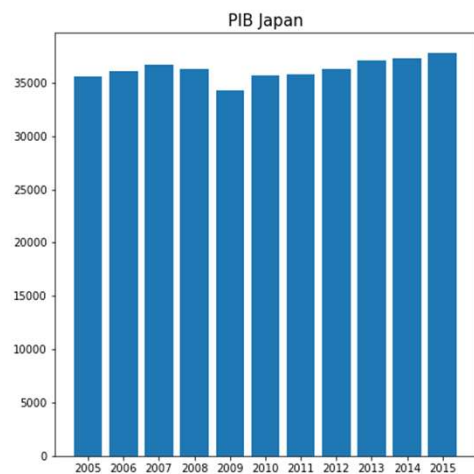
# Evolution indicateur AUSTRALIE



# Evolution indicateur Royaume Uni

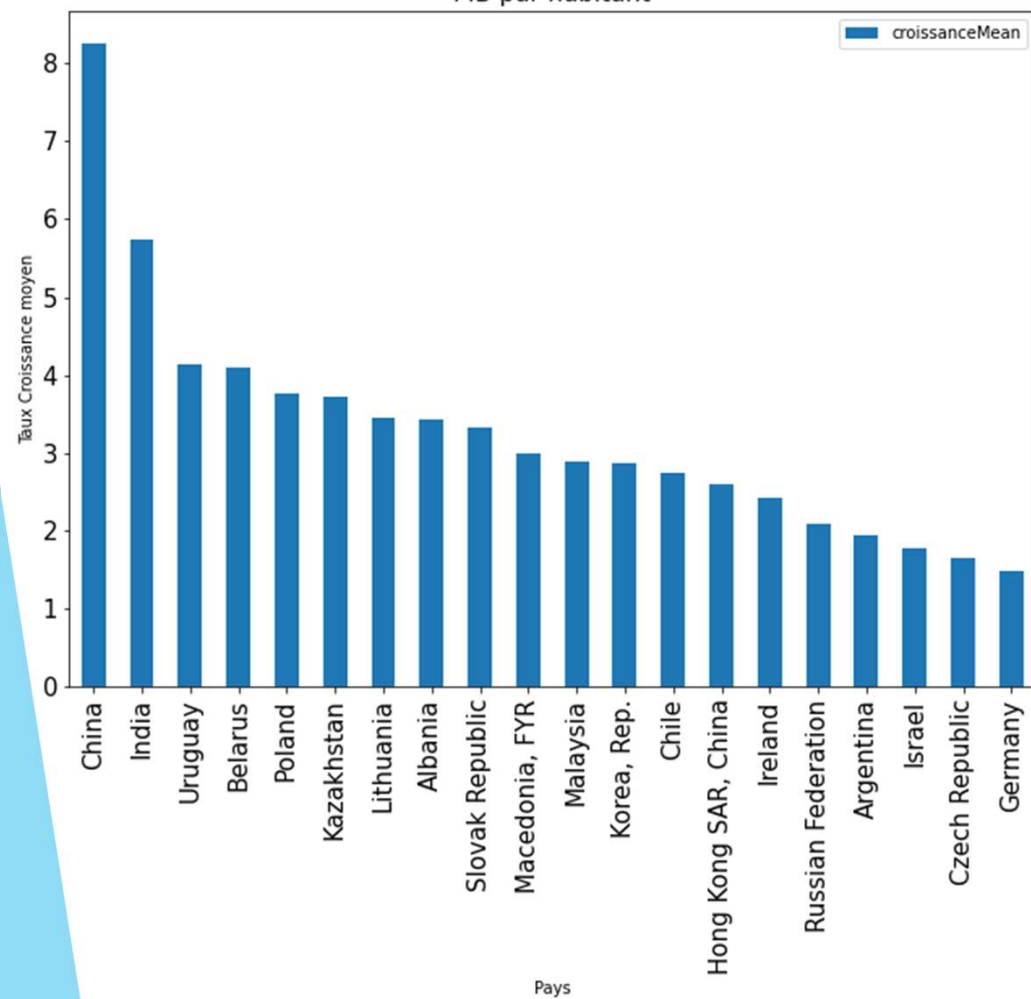


# Evolution indicateur au JAPON

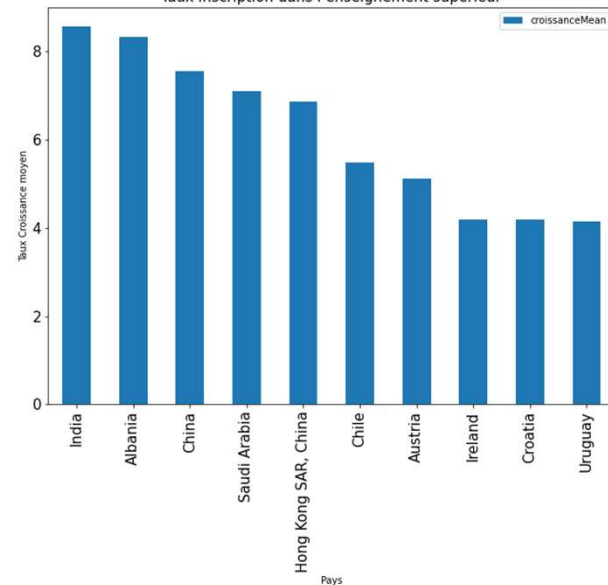


# Pays à forte croissance

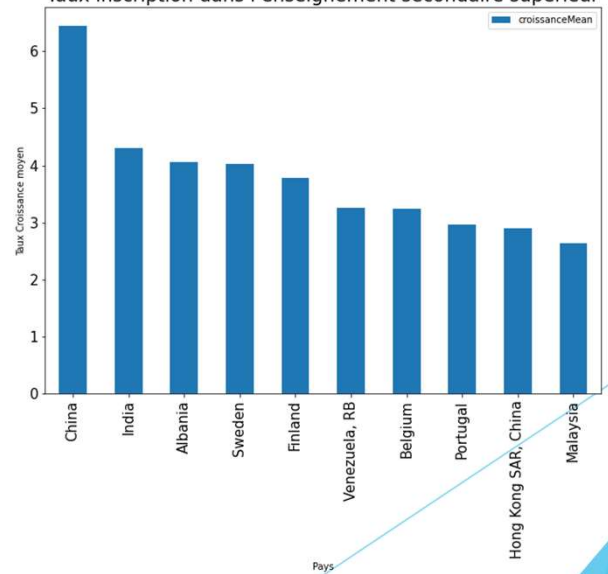
PIB par habitant



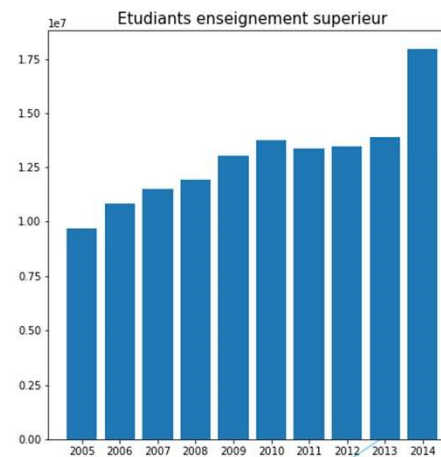
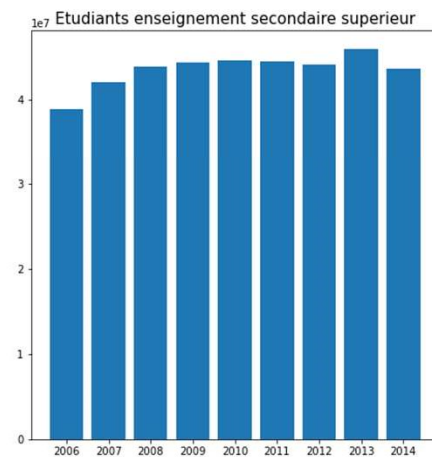
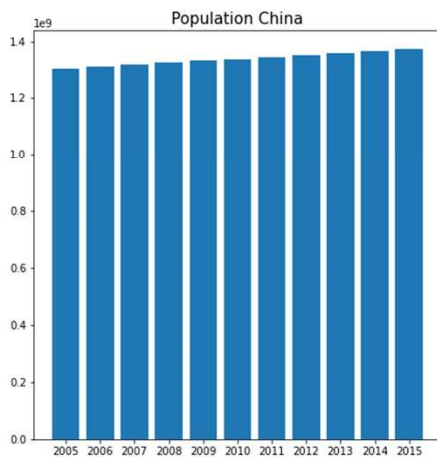
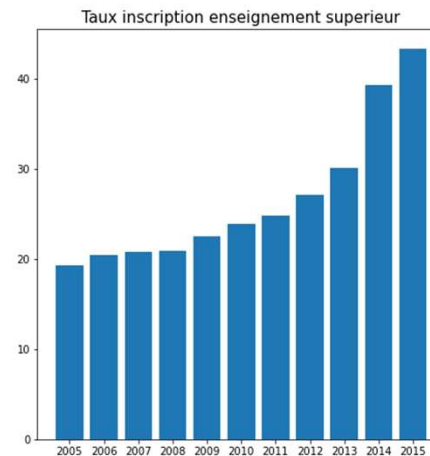
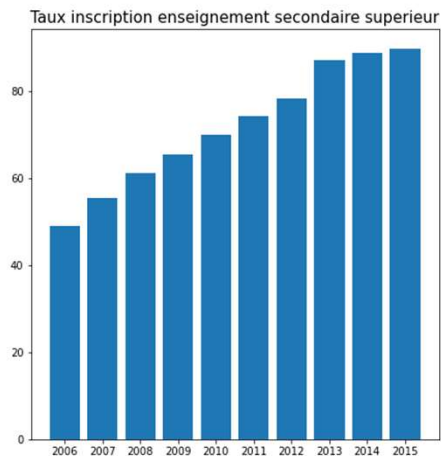
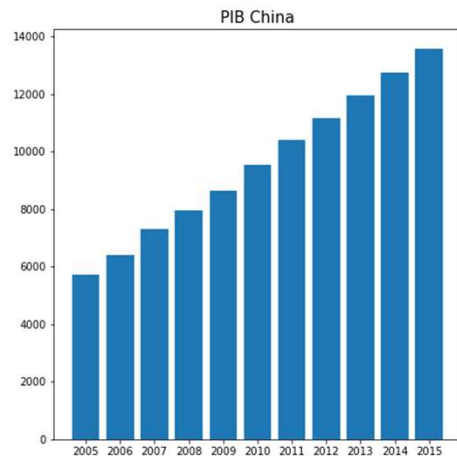
Taux inscription dans l'enseignement supérieur



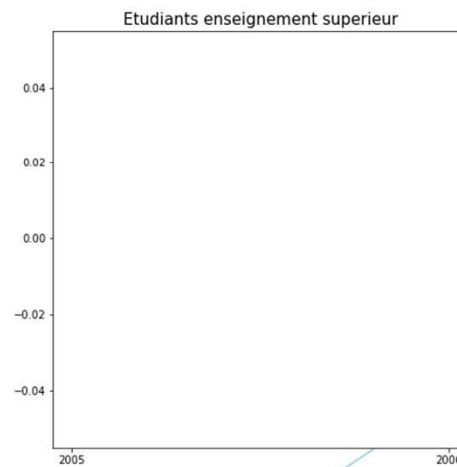
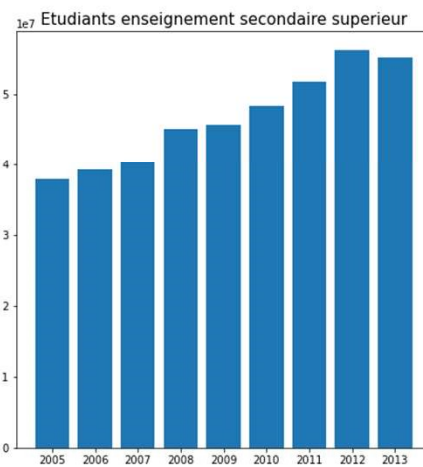
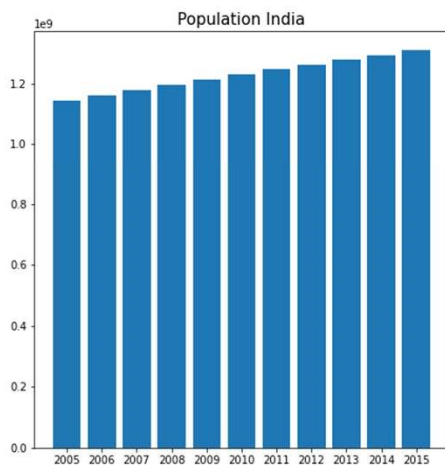
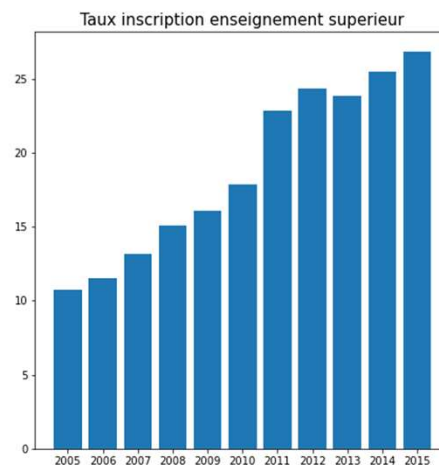
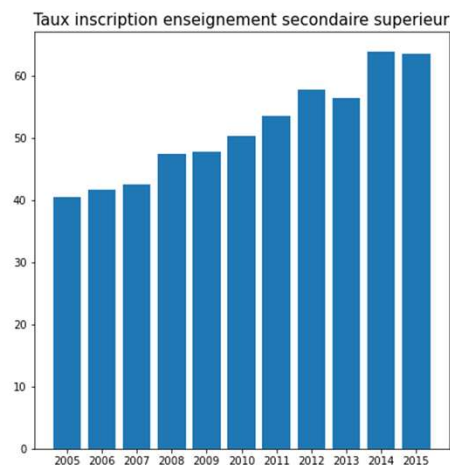
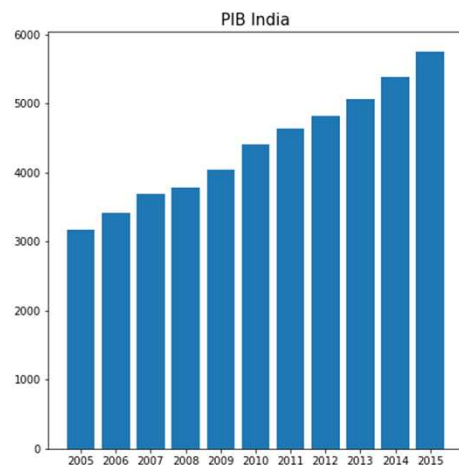
Taux inscription dans l'enseignement secondaire supérieur



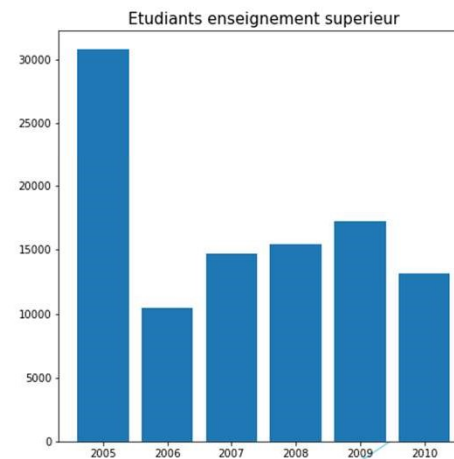
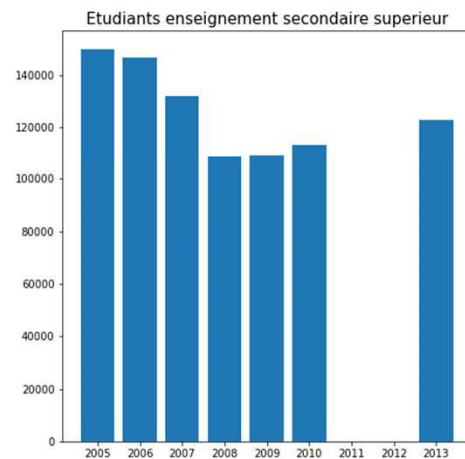
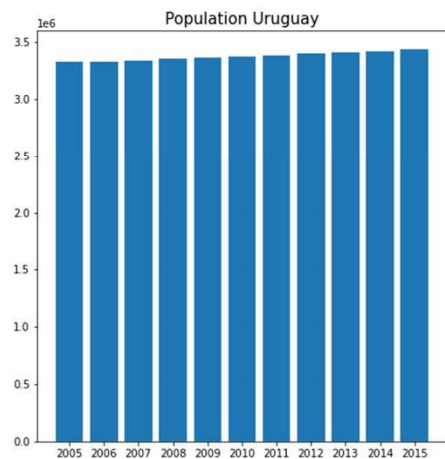
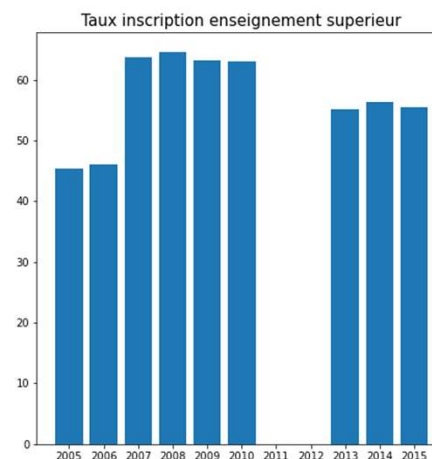
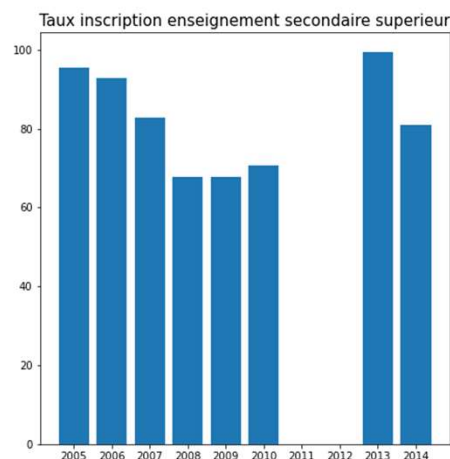
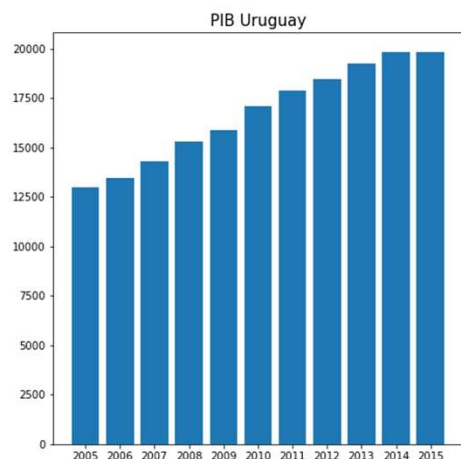
# Evolution indicateur CHINE



# Evolution indicateur Inde

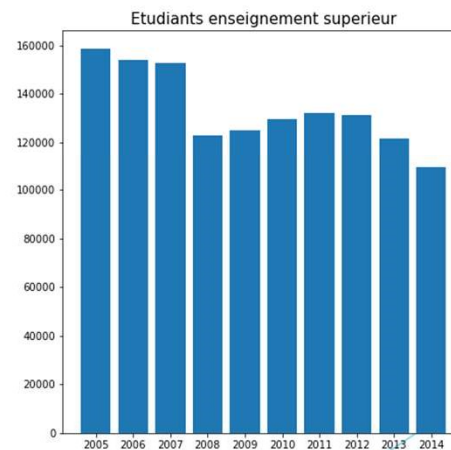
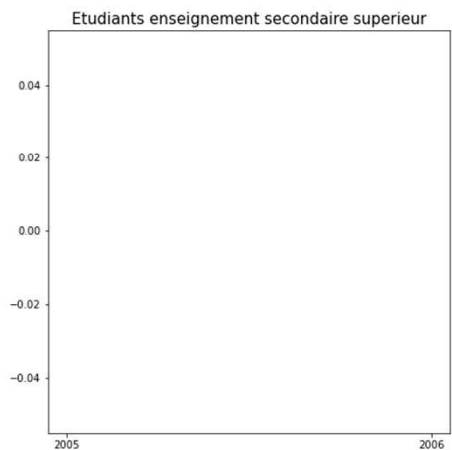
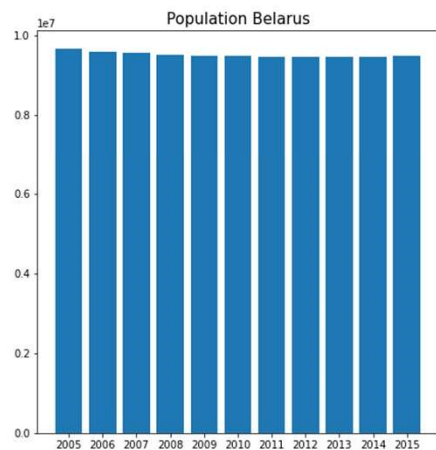
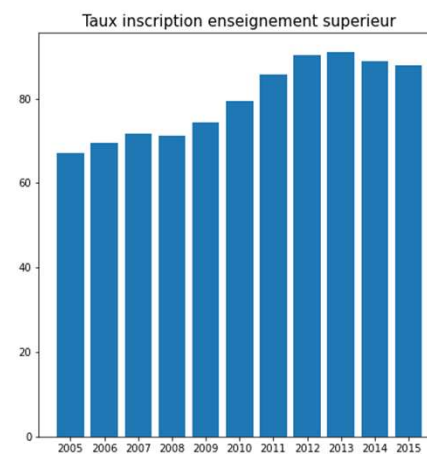
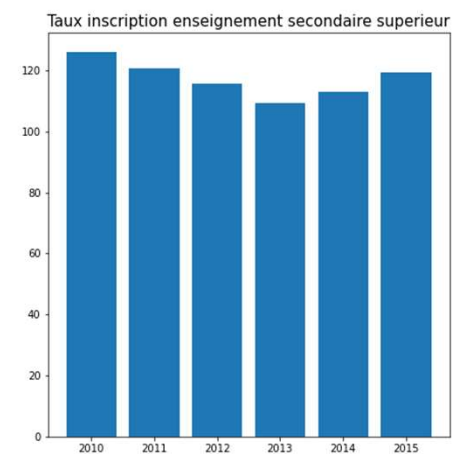
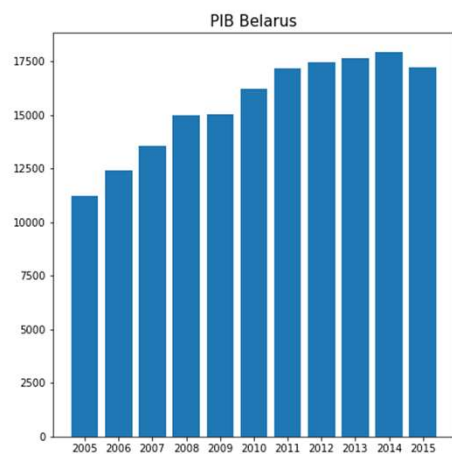


# Evolution indicateur URUGUAY





# Evolution indicateur Bielorussie



# Conclusion

- ▶ Pays à étudier en priorité :
  - Etats-Unis d’Amerique
  - Australie
  - Royaume Uni, Irlande
  - Finlande, Danemark, Norvège, Suède
  - Japon, Corée
- ▶ Pays a fort potentiel crossance importante
  - Chine
  - Inde
  - Pologne, Bielarussie
  - Uruguay

