

OpenClassrooms

Projet P2: Analysez des données de systèmes éducatifs

N. Alves

Projet P2: Analysez des données de systèmes éducatifs

Sommaire

- Contexte
- Présentation des informations disponibles
- Étude du jeu de données
- Analyse
- Conclusion et perspectives

Projet P2: Analysez des données de systèmes éducatifs

- Contexte
- Présentation des informations disponibles
- Étude du jeu de données
- Analyse
- Conclusion et perspectives

Projet P2: Analysez des données de systèmes éducatifs

Contexte

La société Academy propose des contenus en ligne à destination d'un public de niveau lycée et université. Dans le cadre d'un développement commercial à l'international, la société souhaite savoir si les données mises à disposition par la banque mondiale sont susceptibles de renseigner cette problématique : l'ouverture vers de nouveaux pays.

Projet P2: Analysez des données de systèmes éducatifs

- Contexte
- **Présentation des informations disponibles**
- Étude du jeu de données
- Analyse
- Conclusion et perspectives

Projet P2: Analysez des données de systèmes éducatifs

Présentation des informations disponibles

Le jeu de données est issu de la banque mondiale. Il comporte des informations :

- Par pays et régions du monde (217 pays et 25 régions)
- Pour des indicateurs multiples (3665 au total)
- Avec une évolution dans le temps de ces indicateurs (de 1970 à 2100)

Les indicateurs sont organisés:

- par catégorie sur le site de la banque mondiale (19 catégories)
- par catégorie et sous-catégories dans les fichiers (37 catégories et sous-catégories)

Une recherche par mot-clé est aussi disponible.

Projet P2: Analysez des données de systèmes éducatifs

- Contexte
- Présentation des informations disponibles
- **Étude du jeu de données**
- Analyse
- Conclusion et perspectives

Projet P2: Analysez des données de systèmes éducatifs

Étude du jeu de données

Face au volume d'informations disponibles, il convient d'en filtrer la partie la plus pertinente pour l'étude. Une première sélection d'indicateurs sera effectuée.

Les principaux thèmes retenus sont :

- les aspects socio-économiques
- les dépenses en matières d'éducation des gouvernements.

Une recherche par mot-clés (« internet », « electricity ») permet de déterminer des indicateurs possibles pour représenter l'accès à un moyen de visioconférence.

En vérifiant la présence de tous ces indicateurs dans le jeu de données, une première sélection est faite.

Projet P2: Analysez des données de systèmes éducatifs

Étude du jeu de données

Sélection d'indicateurs :

Les indicateurs macroéconomiques retenus relèvent de thématiques comme le contexte socio-économique, les dépenses de l'état et les infrastructures. Il s'agit d'indicateurs groupés selon les éléments suivants:

- la population,
- le revenu par habitant,
- l'accès à internet,
- les dépenses gouvernementales en matière d'éducation.

Projet P2: Analysez des données de systèmes éducatifs

Étude du jeu de données

Quelques pays et les indicateurs associés :

Pays	Revenu par habitant, PPA (\$ international courant, 2016)	Utilisateur d'internet (pour 100 personne, 2016)	Population, total (2015)	Dépense gouvernementale en éducation en % du PNB (%)	Dépense gouvernementale en éducation (Année)
Corée du sud	36 000	93	51 200 000	5.1	2015
France	42 000	86	66 900 000	5.5	2013
Finlande	44 000	88	5 500 000	7.2	2014
U.S.A.	59 000	76	323 100 000	5.4	2014
Singapour	85 000	81	5 600 000	2.9	2013

Projet P2: Analysez des données de systèmes éducatifs

Étude du jeu de données

Qualité des données :

Pour limiter les données manquantes, on utilise deux stratégies :

- On ne conserve que la valeur de la dernière année renseignée de chaque indicateur. On remarque qu'il n'y a pas d'information après 2016.
- On restreint l'étude aux pays dont les populations sont supérieures à 100 000 individus.

⇒ **L'étude concerne 156 pays et 4 indicateurs.**

Projet P2: Analysez des données de systèmes éducatifs

- Contexte
- Présentation des informations disponibles
- Étude du jeu de données
- **Analyse**
- Conclusion et perspectives

Projet P2: Analysez des données de systèmes éducatifs

Analyse

Remarque sur les indicateurs, le revenu par habitant

Le revenu par habitant est l'indicateur le plus dispersé.

On peut observer que la situation est déséquilibrée dans le monde. En effet, 80 % des pays ont un revenu par habitant inférieur à 30 000 \$ (parité de pouvoir d'achat).



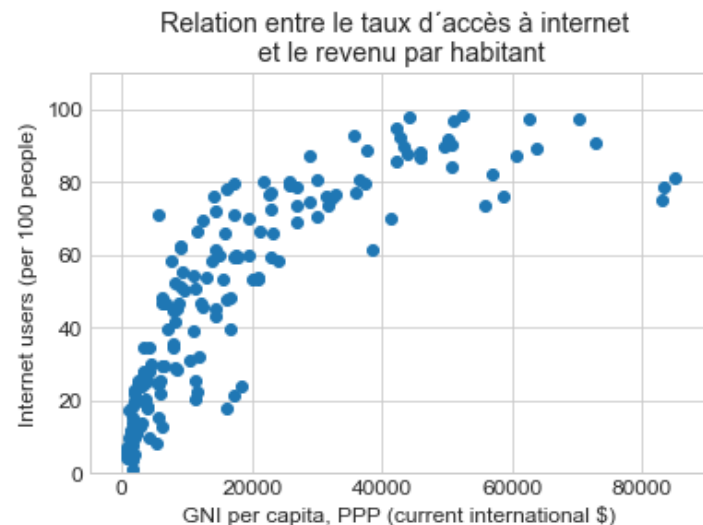
Projet P2: Analysez des données de systèmes éducatifs

Analyse

Remarque sur les indicateurs, le taux d'accès à internet

Cet indicateur est lié au revenu par habitant.

On observe une situation différente entre le groupe des revenus faibles et intermédiaires et celui des hauts revenus.



Projet P2: Analysez des données de systèmes éducatifs

Analyse

Construction d'un score d'attractivité :

Les indicateurs retenus permettent de construire un nouvel indicateur : le score d'attractivité. Il définit le potentiel de client dans un pays, basé sur une moyenne géométrique des indicateurs précédents.

- Les indicateurs sont normalisés pour pouvoir être utilisés sur une échelle comparable.
- Le revenu par habitant étant dispersé, on utilise le logarithme du revenu par habitant pour être moins discriminant vis à vis des autres indicateurs.

$$I_{\text{revenu par habitant}} = \ln(\text{revenu par habitant})$$

- Le score d'attractivité est construit avec la moyenne géométrique des indicateurs.

$$\text{Score} = \sqrt[4]{I_{\text{revenu}} \times I_{\text{population}} \times I_{\text{internet}} \times I_{\text{dépense}}}$$

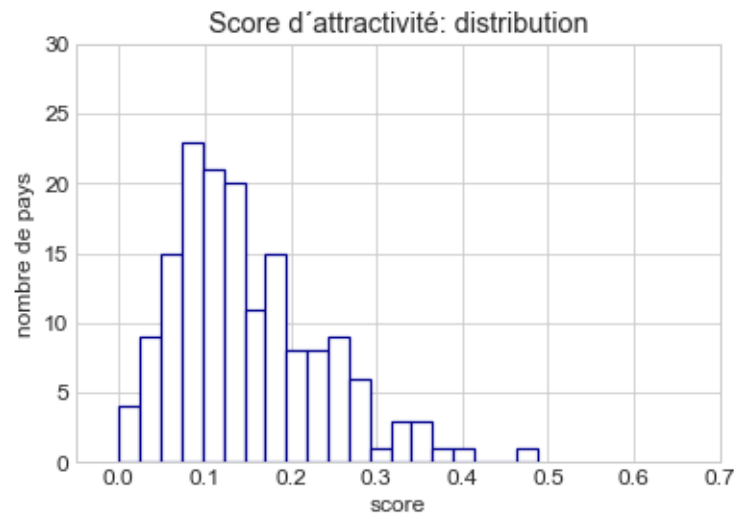
Cette construction s'inspire de la méthode de calcul de [l'indice de développement humain](#).

Projet P2: Analysez des données de systèmes éducatifs

Analyse

Statistique du score d'attractivité

- Moyenne : 0.149
Il s'agit de la moyenne mondiale.
- Médiane : 0.129
50 % des pays sont au-dessus et 50 % des pays sont au-dessous de cette valeur.
- Premier décile : 0.059
10 % des pays sont au-dessous de cette valeur.
- Dernier décile : 0.276
10 % des pays sont au-dessus de cette valeur.



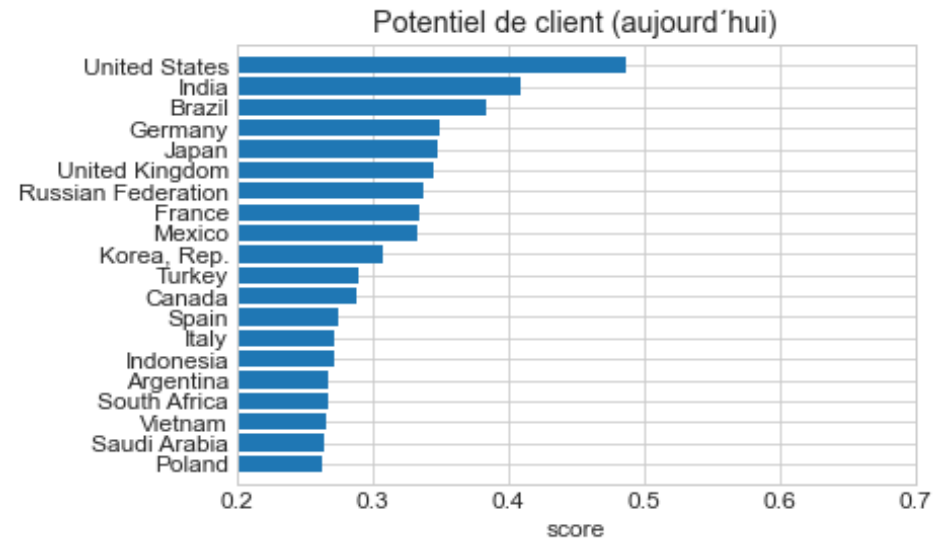
Projet P2: Analysez des données de systèmes éducatifs

Analyse

Les pays ayants le plus grand potentiel de clients.

Une liste des 20 pays les plus intéressants pour l'entreprise est calculée sur la base du score d'attractivité.

Il s'agit d'une étude conçue avec des indicateurs macroéconomiques. Cet étude ne prend pas en compte les disparités au sein de chaque pays.



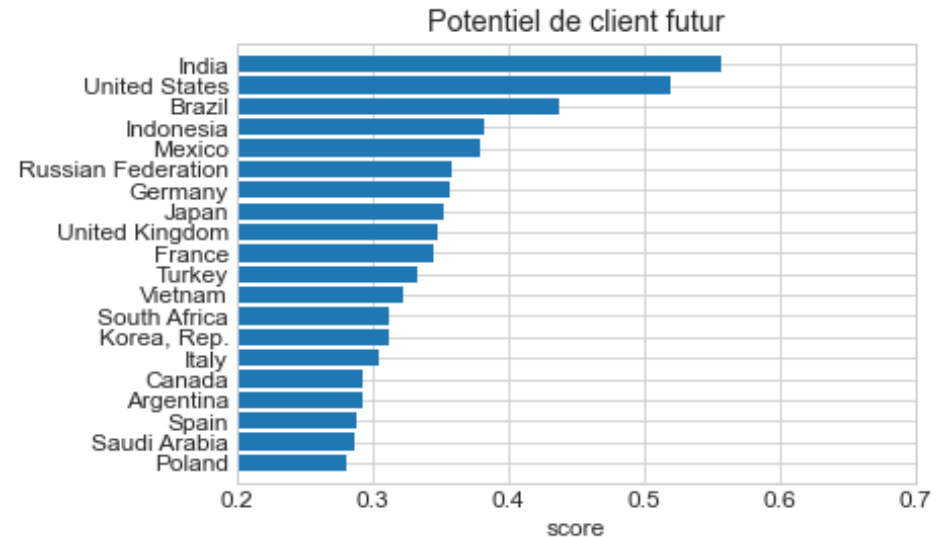
Projet P2: Analysez des données de systèmes éducatifs

Analyse

Étude du potentiel client dans le futur:

Pour calculer cette évolution, on prend en compte deux éléments supplémentaires :

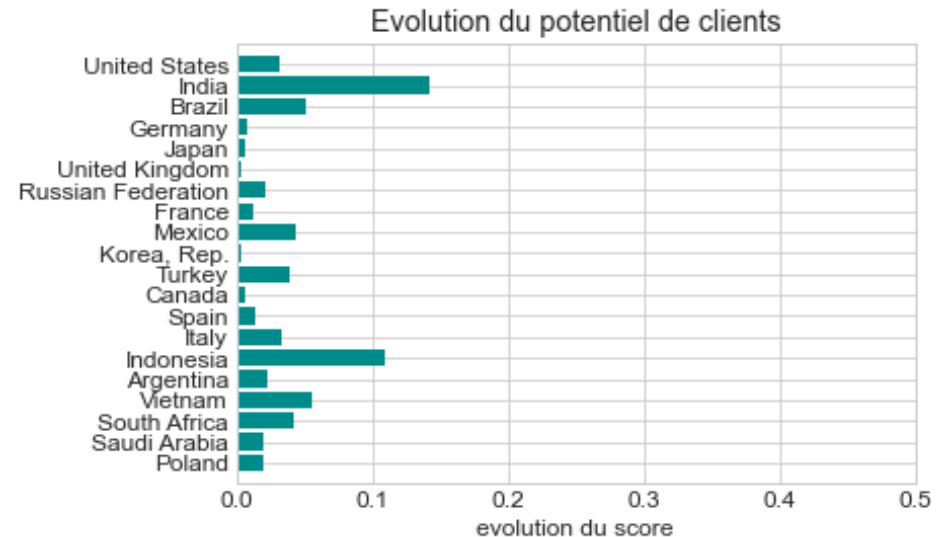
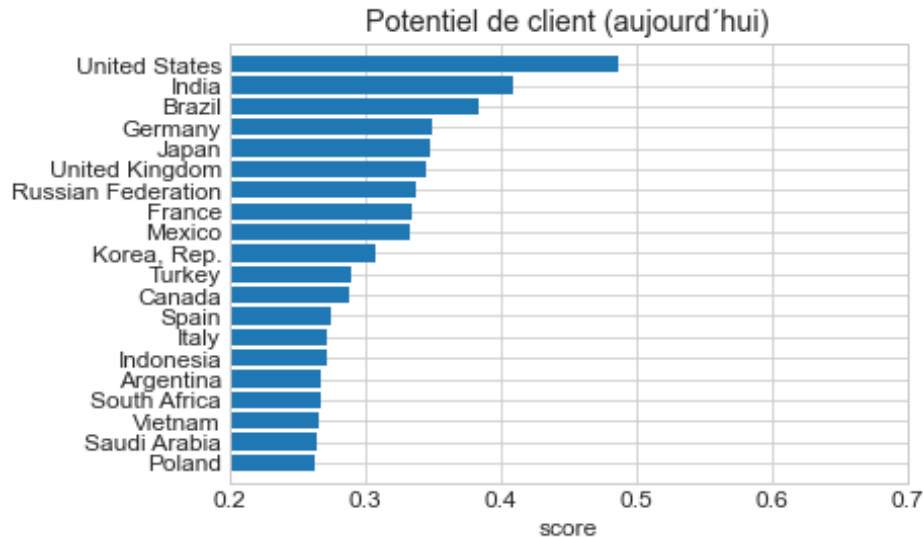
- Dans l'étude, on inclut les enfants âgés de 10 à 14 ans. Ce sont des élèves susceptibles d'être intéressés par les cours en ligne niveau lycée/université dans le futur.
- On suppose que dans un avenir très proche, la grande majorité des individus a accès à internet.



Projet P2: Analysez des données de systèmes éducatifs

Analyse

L'évolution du potentiel est positif pour l'ensemble des pays sélectionnés



Projet P2: Analysez des données de systèmes éducatifs

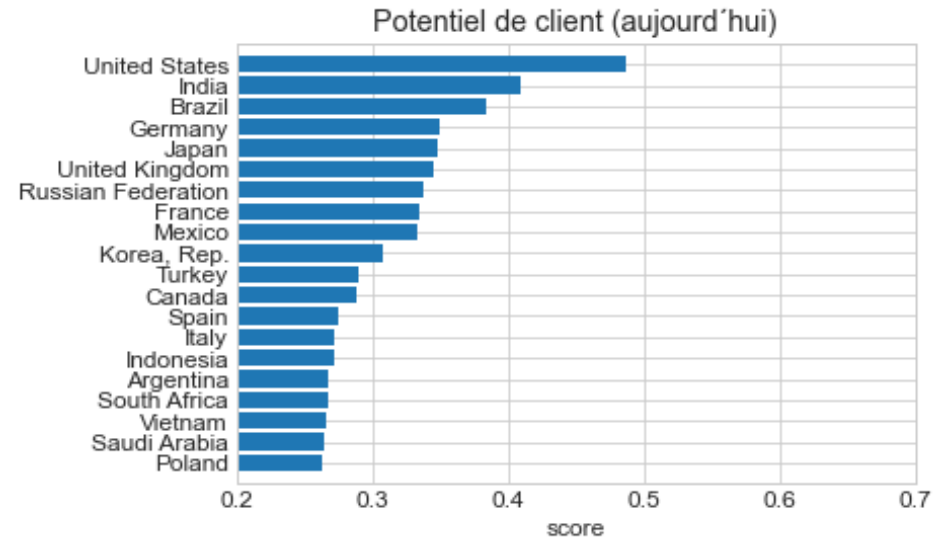
- Contexte
- Présentation des informations disponibles
- Étude Pré-exploratoire : description des données
- Analyse
- Conclusion et perspectives

Projet P2: Analysez des données de systèmes éducatifs

Conclusion et perspectives

Les données fournies par la banque mondiale permettent de déterminer une première liste de pays potentiellement intéressants pour l'entreprise.

Cette étude ne tient pas compte des disparités au sein de chaque pays, elle est à mettre en relation avec une étude de marché locale. Cela permettra de déterminer les possibilités de positionnement de l'offre éducative (barrières, tarif, contenu pédagogique).





Avez vous des questions ?