ANALYSE DE DONNÉES ÉDUCATIVES

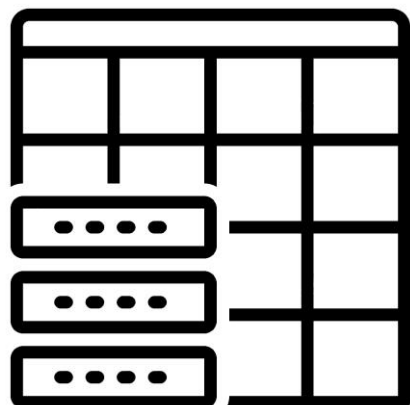academy

RAPPEL DES OBJECTIFS

DEMARCHE

240 PAYS

886 930 L / 70 C

- **EdStatsCountry**..**> Groupe revenu**
- **EdStatsCountry-Series**.......................**> Ø**
- **EdStatsData**...**> 3500 Indicateurs**
- **EdStatsFootNote**..............................**> Ø**
- **EdStatsSeries**.....................................**> Ø**
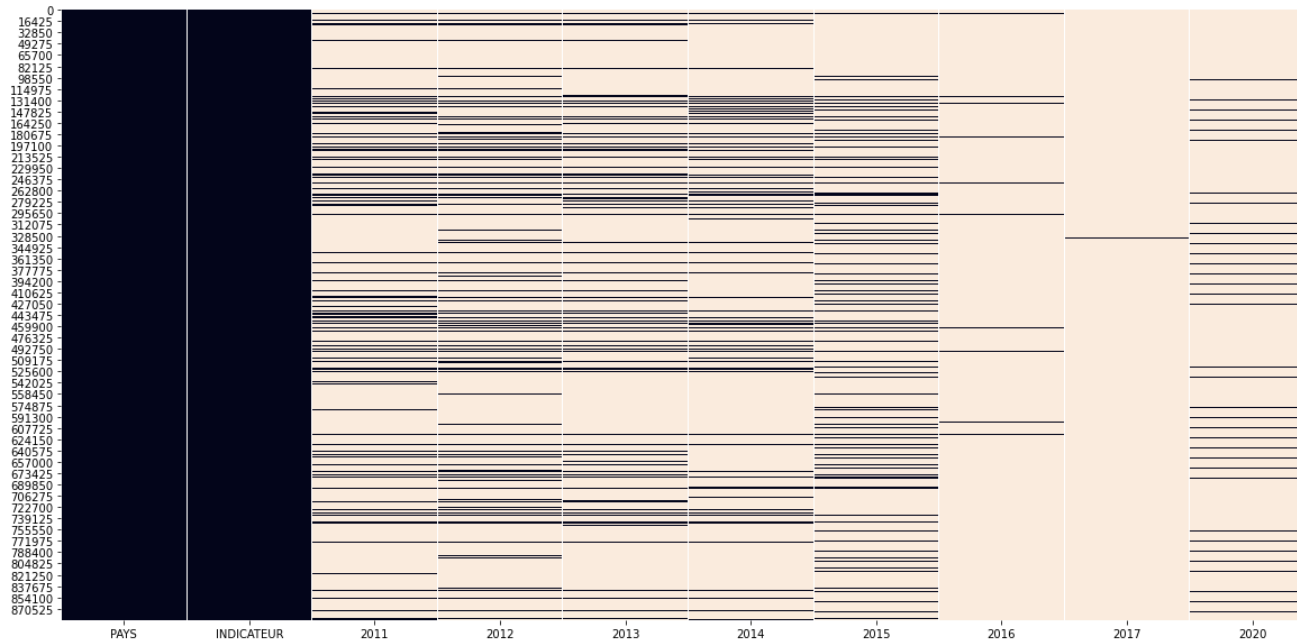
DEMARCHE

_INDICATEURS

3500 INDICATEURS

83%-100% Données manquantes

>

4 INCATEURS

f(NaN, Interpolate, Pertinence)

Valeurs existantes
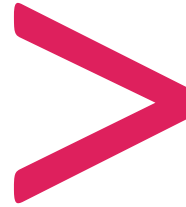
academy

DEMARCHE

4 INDICATEURS

Nombre d'étudiants

Population 15-24 ans

Utilisateurs d'internet

Richesse/habitant

>

SCORING VALEURS

98

SCORING CROISSANCE

98

RÉSULTATS VALEURS

Top 15

75

50

32

| academy

CROISSANCE

academy
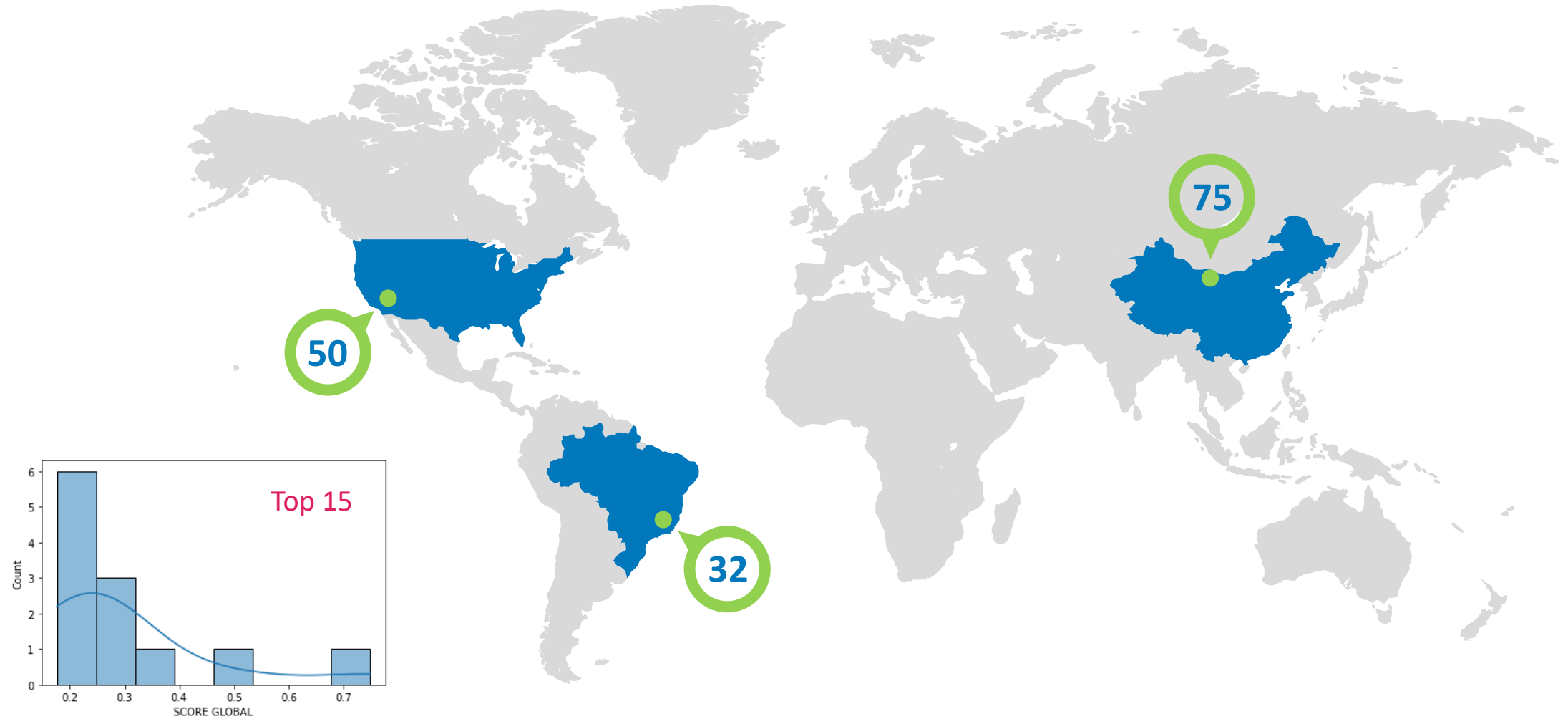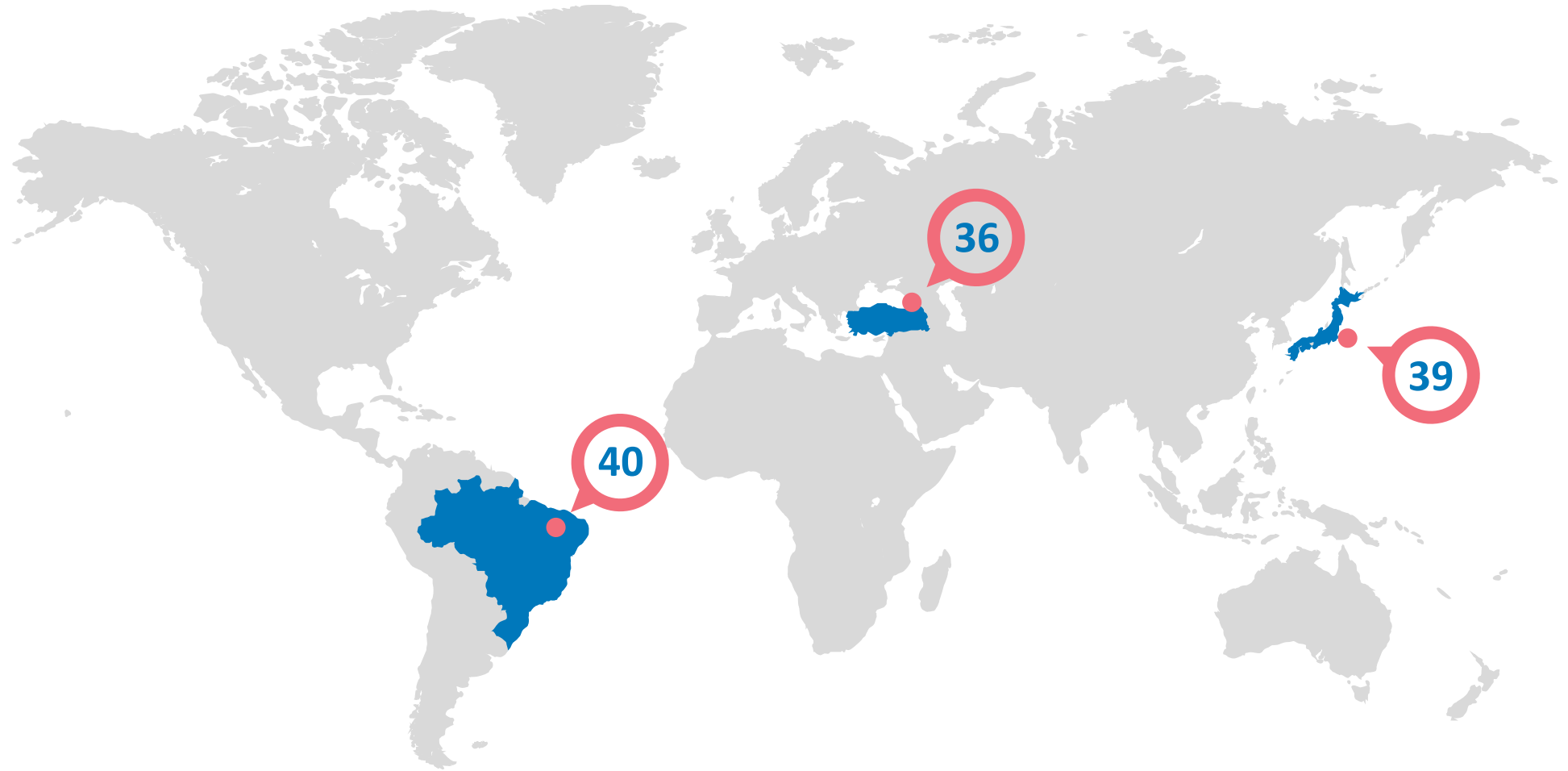
RÉSULTATS

75

24M Lycée / 43M Université

197M

15k < Moyenne (22k)

academy

RÉSULTATS

50

12M Lycée / 24M Université

45M

9 $ **59k**

RÉSULTATS

32

🎓 9M Lycée / 8M Université

🧍 34M

⚠️ $ 15k < Moyenne (22k)

RECOMMANDATION
_VALEUR
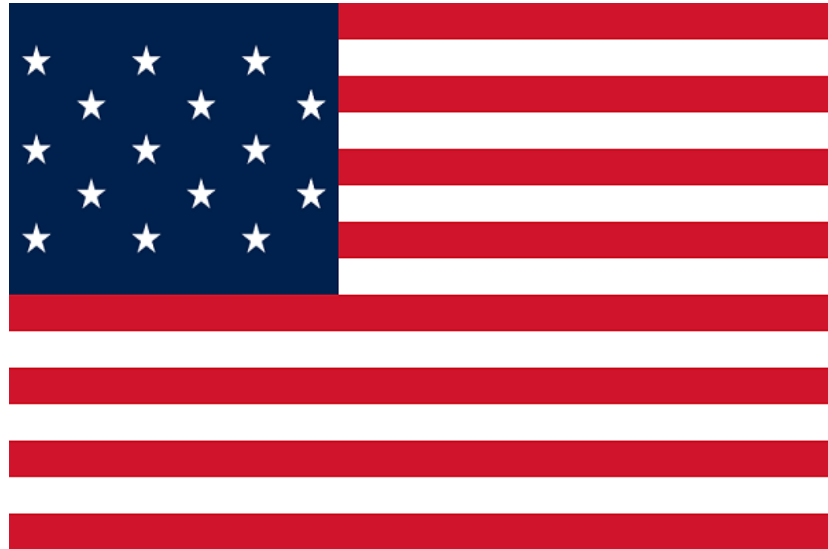
🥈 🎓 **Nombre d'étudiants**

🥈 🚶 **Population cible (15-24 ans)**

✅ 🌐 **Utilisation d'internet**

✅ $ **Pouvoir d'achat x4**

✅ 📈 **Prix des universités (15k-100k) x10**

academy

RECOMMANDATION
_CROISSANCE

🥇 🎓 **Croissance du nombre d'étudiants au Lycée**

4 🌐 **Croissance d'utilisation d'internet**

5 $ **Croissance du pouvoir d'achat**

academy

ANNEXES
_FONCTION

```python
from sklearn import linear_model # importation de linear_model

# Calcul du coeff. directeur
def calc_coeff(dataframe):
    X = pd.DataFrame([0,1,2,3], columns=list('X'), index=['2011', '2012', '2013', '2014'])
    for l in range(5):
        y = dataframe.iloc[l]
        model = linear_model.LinearRegression()
        results = model.fit(X, y)
        print(results.coef_)
```

academy

ANNEXES
_COMMANDES UTILISÉES

- **df = pd.read_csv()** # importation de fichier csv
- **df.columns** # nom colonnes
- **df.dtypes** # types données
- **df.shape** # taille
- **df.sort_values()** # trie par ordre
- **df.isna().mean()** # taux valeurs manquantes
- **sns.heatmap / sns.histplot()** # graphiques sns
- **df.boxplot()** # graphique python
- **df.loc[]** # sélection []

- **df.isin()** # présence élt dans dataframe (boolean)
- **df.drop() / df.dropna()** # Supprime () / Supprimer Na ()
- **df.interpolate()** # méthode pour remplir NaN
- **df.pivot()** # restructure dataframe
- **df.merge()** # combine dataframes
- **df.describe()** # statistiques dataframes
- **df.head** # affiche 15 premières lignes
- **df[df[]>x]** # mask de sélection

academy

ANNEXES
_DONNÉES UTILISÉES

Fournis en zip (origine WorldBank)

- **EdStatsCountry**
- **EdStatsCountry-Series3- EdStatsData**
- **EdStatsFootNote**
- **EdStatsSeries**

+

Infos coûts scolarité (source)
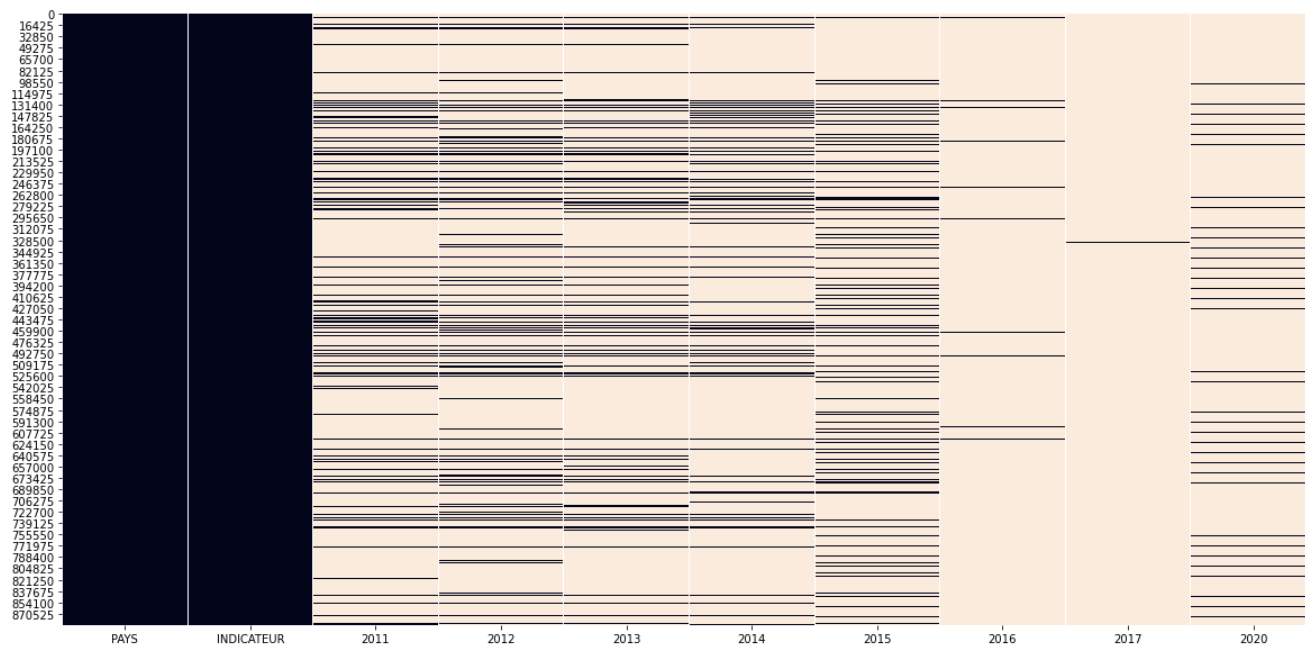
_NETTOYAGE

Après nettoyage

PAYS	0.000000
INDICATEUR	0.000000
2011	0.835374
2012	0.833962
2013	0.844961
2014	0.871705
2015	0.852234
2016	0.981442
2017	0.999839
2020	0.942007
dtype: float64	

PAYS	0.000000
INDICATEUR	0.000000
2011	0.193388
2012	0.198347
2013	0.213223
2014	0.244628
2015	0.395868
2016	0.645455
2017	1.000000
2020	1.000000
dtype: float64	

|	pays	nbre_inscrits_univ	nbre_inscrits_lycee	GNI-PPP par pers	utilisateur d'internet / 100 pers.	pop 15-24 ans	revenu	croissance(%)
count	129	1.290000e+02	1.290000e+02	129.000000	129.000000	1.290000e+02	129	129.000000
unique	129	NaN	NaN	NaN	NaN	NaN	5	NaN
top	Cabo Verde	NaN	NaN	NaN	NaN	NaN	Upper middle income	NaN
freq	1	NaN	NaN	NaN	NaN	NaN	34	NaN
mean	NaN	1.167340e+06	8.987429e+05	21676.744186	53.369339	5.986951e+06	NaN	1.193207
std	NaN	4.276535e+06	2.628682e+06	22089.615743	28.253221	1.877913e+07	NaN	1.116807
min	NaN	2.788000e+03	3.829000e+03	700.000000	4.000000	2.353600e+04	NaN	-1.720565
25%	NaN	7.453700e+04	7.677100e+04	5540.000000	25.570687	5.562690e+05	NaN	0.350911
50%	NaN	2.164680e+05	1.642670e+05	14420.000000	58.271236	1.493722e+06	NaN	1.073767
75%	NaN	6.323770e+05	5.462140e+05	31420.000000	77.289395	4.842803e+06	NaN	1.960334
max	NaN	4.336739e+07	2.447537e+07	125000.000000	98.240016	1.970268e+08	NaN	4.124170

ANNEXES

_INDICATEURS: NOMBRE D'ÉTUDIANTS

	pays	nbre_inscrits_lycee
37	China	24475374.0
187	United States	11736315.0
24	Brazil	9139433.0
77	Indonesia	5703003.0
13	Bangladesh	5355156.5
...	...	...
14	Barbados	8410.0
17	Belize	7680.0
162	St. Lucia	5391.0
105	Maldives	4552.0
66	Grenada	3829.0

	pays	nbre_inscrits_univ
37	China	43367392.0
187	United States	19531728.0
24	Brazil	8285475.0
179	Turkey	6062886.0
77	Indonesia	5107999.0
...	...	...
167	Swaziland	8057.0
39	Comoros	6499.0
105	Maldives	6089.0
100	Luxembourg	6085.0
162	St. Lucia	2788.0

academy
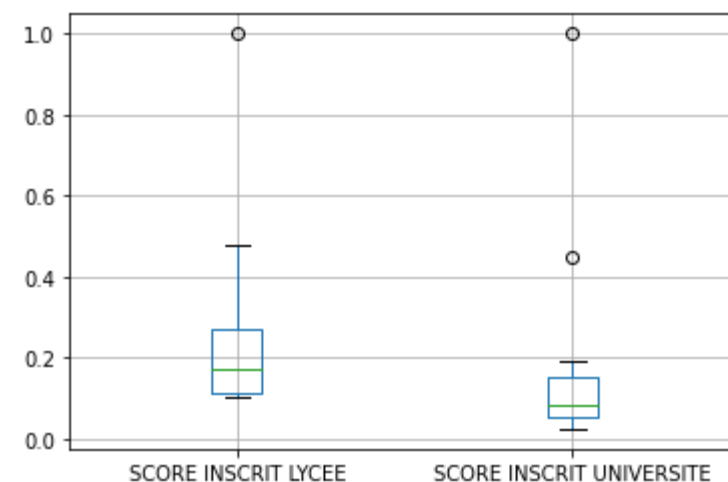
ANNEXES

	pays	pop 15-24 ans
37	China	197026759.0
187	United States	45147517.0
77	Indonesia	41819264.0
131	Pakistan	39080811.0
13	Bangladesh	34400594.0
...	...	...
107	Malta	51504.0
76	Iceland	44570.0
14	Barbados	37585.0
162	St. Lucia	29869.0
66	Grenada	23536.0

academy
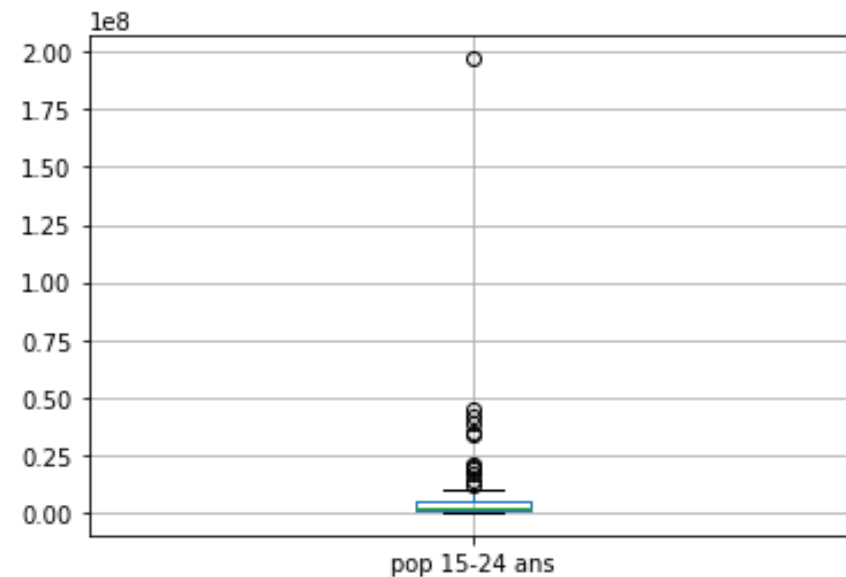
_INDICATEURS: RICHESSE

	pays	GNI-PPP par pers
141	Qatar	125000.0
101	Macao SAR, China	98650.0
90	Kuwait	83310.0
185	United Arab Emirates	72980.0
100	Luxembourg	70430.0
...	...	...
103	Malawi	1140.0
126	Niger	970.0
27	Burundi	770.0
96	Liberia	710.0
33	Central African Republic	700.0
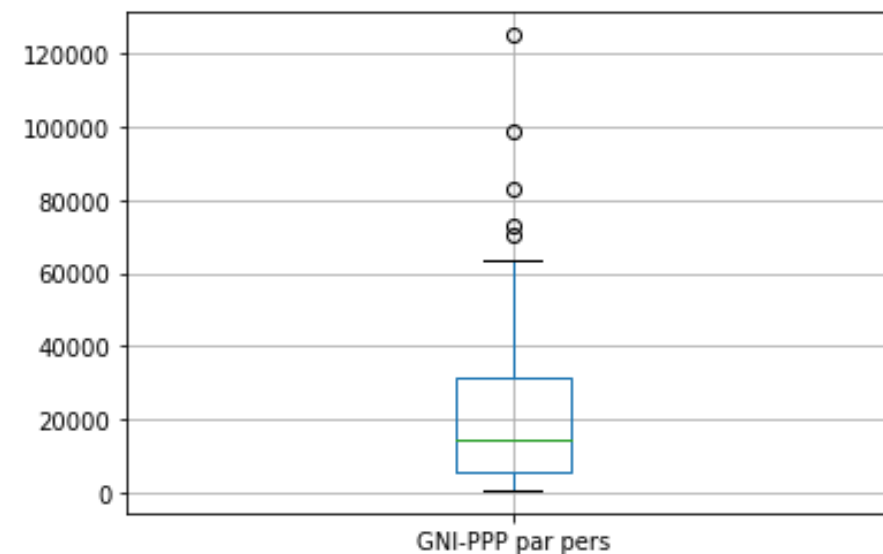
_INDICATEURS: CROISSANCE

	pays	croissance(%)
107	Malta	4.124170
126	Niger	3.771351
12	Bahrain	3.615325
183	Uganda	3.269713
4	Angola	3.218530
...	...	...
0	Albania	-0.577942
22	Bosnia and Herzegovina	-0.613298
93	Latvia	-0.643400
25	Bulgaria	-0.697303
112	Moldova	-1.720565

academy
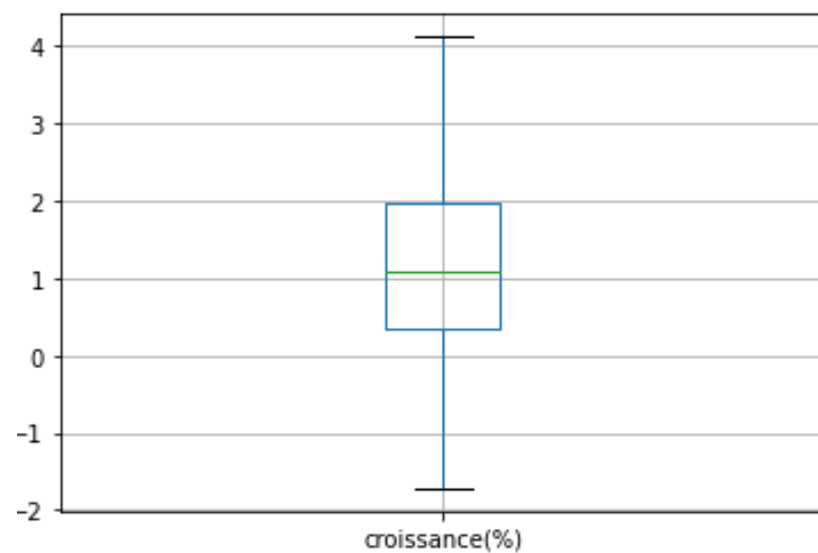
_SCORING

	pays	nbre_inscrits_univ	nbre_inscrits_lycee	GNI-PPP par pers	utilisateur d'internet / 100 pers.	pop 15-24 ans	revenu	croissance(%)	SCORE INSCRIT LYCEE	SCORE INSCRIT UNIVERSITE	SCORE INTERNET	SCORE POPULATION	SCORE GNI	CROISSANCE	SCORE GLOBAL
37	China	4.336739e+07	24475374.0	15500.0	53.200000	197026759.0	Upper middle income	0.314091	1.000000	1.000000	0.541531	1.000000	0.12400	0.076159	0.748338
187	United States	1.953173e+07	11736315.0	58700.0	76.176737	45147517.0	High income: OECD	0.350911	0.479515	0.450378	0.775415	0.229144	0.46960	0.085086	0.497828
141	Qatar	2.786600e+04	40403.0	125000.0	94.291118	129378.0	High income: nonOECD	1.715004	0.001651	0.000643	0.959804	0.000657	1.00000	0.415842	0.475719
12	Bahrain	3.890100e+04	37008.0	44250.0	97.999981	132609.0	High income: nonOECD	3.615325	0.001512	0.000897	0.997557	0.000673	0.35400	0.876619	0.446252
107	Malta	1.321600e+04	15110.0	35840.0	77.289395	51504.0	High income: nonOECD	4.124170	0.000617	0.000305	0.786740	0.000261	0.28672	1.000000	0.414929
100	Luxembourg	6.085000e+03	9797.0	70430.0	97.493930	65335.0	High income: OECD	1.960334	0.000400	0.000140	0.992405	0.000332	0.56344	0.475328	0.406409
101	Macao SAR, China	3.077100e+04	15239.0	98650.0	81.642985	62558.0	High income: nonOECD	1.379473	0.000623	0.000710	0.831056	0.000318	0.78920	0.334485	0.391278
90	Kuwait	7.178600e+04	108735.0	83310.0	78.367384	463706.0	High income: nonOECD	1.497756	0.004443	0.001655	0.797713	0.002354	0.66648	0.363166	0.367162
185	United Arab Emirates	1.566130e+05	151777.0	72980.0	90.600007	743608.0	High income: nonOECD	1.219429	0.006201	0.003611	0.922231	0.003774	0.58384	0.295679	0.363067
76	Iceland	1.909700e+04	15191.0	52490.0	98.240016	44570.0	High income: OECD	1.612716	0.000621	0.000440	1.000000	0.000226	0.41992	0.391040	0.362449
124	New Zealand	2.700742e+05	164749.0	37660.0	88.470186	615099.0	High income: OECD	2.086804	0.006731	0.006228	0.900551	0.003122	0.30128	0.505994	0.344781
146	Saudi Arabia	1.527769e+06	1592426.0	55860.0	73.750904	5298036.0	High income: nonOECD	1.578837	0.065062	0.035229	0.750722	0.026890	0.44688	0.382826	0.341522

academy

_SCORING FILTRÉ PAR SCORE ÉTUDIANTS

	pays	nbre_inscrits_univ	nbre_inscrits_lycee	GNI-PPP par pers	utilisateur d'internet / 100 pers.	pop 15-24 ans	revenu	croissance(%)	SCORE INSCRIT LYCEE	SCORE INSCRIT UNIVERSITE	SCORE INTERNET	SCORE POPULATION	SCORE GNI	CROISSANCE	SCORE GLOBAL
37	China	43367392.0	24475374.0	15500.0	53.200000	197026759.0	Upper middle income	0.314091	1.000000	1.000000	0.541531	1.000000	0.12400	0.076159	0.748338
187	United States	19531728.0	11736315.0	58700.0	76.176737	45147517.0	High income: OECD	0.350911	0.479515	0.450378	0.775415	0.229144	0.46960	0.085086	0.497828
24	Brazil	8285475.0	9139433.0	14840.0	59.682747	33595574.0	Upper middle income	0.712873	0.373413	0.191053	0.607520	0.170513	0.11872	0.172852	0.326814
84	Japan	3862460.0	2835697.0	42790.0	92.000000	12157426.0	High income: OECD	-0.340269	0.115859	0.089064	0.936482	0.061704	0.34232	-0.082506	0.292585
179	Turkey	6062886.0	2725972.0	24160.0	58.347734	13930222.0	Upper middle income	1.084194	0.111376	0.139803	0.593930	0.070702	0.19328	0.262888	0.274396
111	Mexico	3419391.0	3950126.0	17760.0	59.540446	20336734.0	Upper middle income	1.058240	0.161392	0.078847	0.606071	0.103218	0.14208	0.256595	0.269641
77	Indonesia	5107999.0	5703003.0	11240.0	25.366301	41819264.0	Lower middle income	1.065179	0.233010	0.117784	0.258207	0.212252	0.08992	0.258277	0.233890
157	South Africa	1018543.0	2609033.0	12880.0	54.000000	9998431.0	Upper middle income	1.273356	0.106598	0.023486	0.549674	0.050747	0.10304	0.308755	0.228460
131	Pakistan	1871575.0	4458868.0	5570.0	15.514558	39080811.0	Lower middle income	1.978320	0.182178	0.043156	0.157925	0.198353	0.04456	0.479689	0.221172
191	Vietnam	2466643.0	2532465.0	6170.0	46.500000	16939813.0	Lower middle income	0.904518	0.103470	0.056878	0.473331	0.085977	0.04936	0.219321	0.197667
13	Bangladesh	2068355.0	5355156.5	3790.0	18.246938	34400594.0	Low income	1.002774	0.218798	0.047694	0.185738	0.174599	0.03032	0.243146	0.180059
173	Thailand	2235450.0	2721167.0	16100.0	47.504966	9149718.0	Upper middle income	0.250165	0.111180	0.051547	0.483560	0.046439	0.12880	0.060658	0.176437

ANNEXES
_SCORING CROISSANCE

PAYS	coeff_inscrits_univ	coeff_inscrits_lycee	coeff GNI-PPP	coeff internet	coeff pop	SCORE COEFF LYCEE	SCORE COEFF UNIV	SCORE COEFF GNI	SCORE COEFF INTERNET	SCORE COEFF POP	SCORE COEFF GLOBAL
Brazil	305894.0	580070.0	600.0	2.675000	-64271.0	1.000000	0.091715	0.342857	0.583722	-0.013610	0.400937
Japan	-8897.5	4096.5	1745.0	4.582658	-100056.5	0.007062	-0.002668	0.997143	1.000000	-0.021187	0.396070
Turkey	579302.0	297725.0	1310.0	1.592145	75408.0	0.513257	0.173689	0.748571	0.347428	0.015968	0.359783
China	3335279.4	85373.6	1046.0	3.229988	-4722495.3	0.147178	1.000000	0.597714	0.704829	-1.000000	0.289944
Mexico	179882.0	138653.0	460.0	2.573705	133409.0	0.239028	0.053933	0.262857	0.561618	0.028250	0.229137
South Africa	6411.0	39685.0	280.0	4.000000	42083.5	0.068414	0.001922	0.160000	0.872856	0.008911	0.222421
Indonesia	348645.9	212458.7	529.0	1.500000	-78996.2	0.366264	0.104533	0.302286	0.327321	-0.016728	0.216735
United States	-496920.6	-32104.9	1750.0	0.651162	23156.7	-0.055347	-0.148989	1.000000	0.142093	0.004903	0.188532
Pakistan	117610.3	191272.8	195.0	0.994000	28167.8	0.329741	0.035263	0.111429	0.216905	0.005965	0.139860
Vietnam	137671.6	-98757.0	291.0	1.949000	-329207.3	-0.170250	0.041277	0.166286	0.425299	-0.069710	0.078580

academy