
PROJET 2 : ANALYSER DES DONNÉES DE SYSTÈMES ÉDUCATIFS

OBJECTIFS

Définir les critères

Nettoyage des données

Recherche des meilleurs indicateurs

Définition d'un score d'attractivité

DÉFINIR LES CRITÈRES

- **Taux d'éducation**
 - **Taux d'équipement et accès internet**
 - **Richesse par individu**
 - **Taille marché/difficulté accès au marché**
-

NETTOYAGE DES DONNÉES

	Country Name	Country Code	Indicator Name	Indicator Code	1970	1971	1972	1973	1974	1975	...	2060	2065
0	Arab World	ARB	Adjusted net enrolment rate, lower secondary, ...	UIS.NERA.2	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN
1	Arab World	ARB	Adjusted net enrolment rate, lower secondary, ...	UIS.NERA.2.F	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN
2	Arab World	ARB	Adjusted net enrolment rate, lower secondary, ...	UIS.NERA.2.GPI	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN

886 930 lignes * 70 colonnes = 62 millions cases

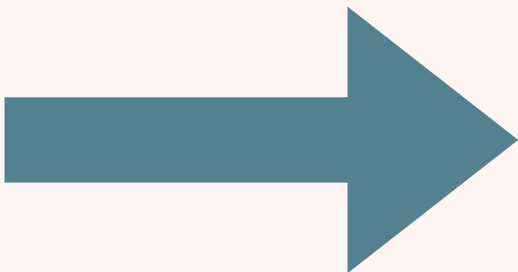
3 666 indicateurs pour 242 pays

NETTOYAGE DES DONNÉES

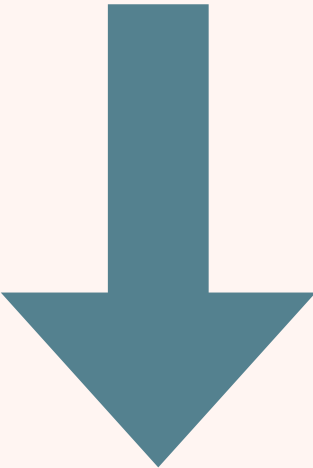
	Country Name	Country Code	Indicator Name	Indicator Code	1970	1971	1972	1973	1974	1975	...	2060	2065
0	Arab World	ARB	Adjusted net enrolment rate, lower secondary, ...	UIS.NERA.2	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN
1	Arab World	ARB	Adjusted net enrolment rate, lower secondary, ...	UIS.NERA.2.F	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN
2	Arab World	ARB	Adjusted net enrolment rate, lower secondary, ...	UIS.NERA.2.GPI	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN

886 930 lignes * 70 colonnes = 62 millions cases

3 666 indicateurs pour 242 pays



Récupère la dernière valeur (entre 2000 et 2020)



	Country Code	Country Name	Indicator Code	Indicator Name	last_element_year	last_element_value
4	ARB	Arab World	SE.PRM.TENR	Adjusted net enrolment rate, primary, both sex...	2014.0	85.320152
5	ARB	Arab World	SE.PRM.TENR.FE	Adjusted net enrolment rate, primary, female (%)	2014.0	83.820831
6	ARB	Arab World	UIS.NERA.1.GPI	Adjusted net enrolment rate, primary, gender p...	2014.0	0.966200
7	ARB	Arab World	SE.PRM.TENR.MA	Adjusted net enrolment rate, primary, male (%)	2014.0	86.753387
12	ARB	Arab World	UIS.NIRA.1	Adjusted net intake rate to Grade 1 of primary...	2014.0	76.620567

348 828 lignes * 6 colonnes = 2,1 millions cases

RECHERCHE MEILLEURS INDICATEURS

- Critère : EDUCATION
- Recherche du mot « éducation » dans les indicateurs (apparaît dans 1335 indicateurs)

RECHERCHE MEILLEURS INDICATEURS

- Critère : EDUCATION
- Recherche du mot « éducation » dans les indicateurs (apparaît dans 1335 indicateurs)
- Classement indicateurs par nombre d'apparitions

Indicateur	last_element_value
SP.POP.GROW : Population growth (annual %)	240
SP.POP.TOTL : Population, total	240
NY.GDP.MKTP.CD : GDP at market prices (current US\$)	234
NY.GDP.PCAP.CD : GDP per capita (current US\$)	234
IT.NET.USER.P2 : Internet users (per 100 people)	230
SE.PRM.ENRL : Enrolment in primary education, both sexes (number)	229
SE.PRM.ENRL.FE : Enrolment in primary education, female (number)	229
SE.PRM.ENRL.FE.ZS : Percentage of students in primary education who are female (%)	229
SE.SEC.ENRL.GC : Enrolment in secondary general, both sexes (number)	228
SE.SEC.ENRL.GC.FE : Enrolment in secondary general, female (number)	228
SE.SEC.ENRL.GC.FE.ZS : Percentage of students in secondary general education who are female (%)	228
UIS.E.2 : Enrolment in lower secondary education, both sexes (number)	228
NY.GDP.MKTP.KD : GDP at market prices (constant 2005 US\$)	227
NY.GDP.PCAP.KD : GDP per capita (constant 2005 US\$)	227
UIS.E.2.F : Enrolment in lower secondary education, female (number)	227
UIS.E.2.GPV : Enrolment in lower secondary general, both sexes (number)	227
UIS.E.3.GPV : Enrolment in upper secondary general, both sexes (number)	227
UIS.FEP.2 : Percentage of students in lower secondary education who are female (%)	227
UIS.GTVP.2.GPV : Percentage of students in lower secondary education enrolled in general programmes, both sexes (%)	227
NY.GNP.MKTP.CD : GNI (current US\$)	226
UIS.E.2.GPV.F : Enrolment in lower secondary general, female (number)	226
UIS.E.3.GPV.F : Enrolment in upper secondary general, female (number)	226
UIS.FEP.3.GPV : Percentage of students in upper secondary general education who are female (%)	226
UIS.GTVP.2.GPV.F : Percentage of female students in lower secondary education enrolled in general programmes, female (%)	226

RECHERCHE MEILLEURS INDICATEURS

RECHERCHE MEILLEURS INDICATEURS

➤ **Taux d'éducation**



- **SE.SEC.NENR : Net enrolment rate, secondary, both sexes (%)**
- **SE.TER.ENRR : Gross enrolment ratio, tertiary, both sexes (%)**

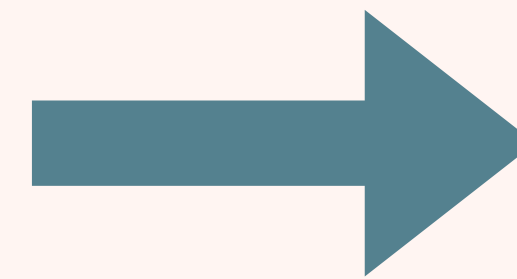
RECHERCHE MEILLEURS INDICATEURS

➤ Taux d'éducation



- SE.SEC.NENR : Net enrolment rate, secondary, both sexes (%)
- SE.TER.ENRR : Gross enrolment ratio, tertiary, both sexes (%)

➤ Taux d'équipement et accès internet



- IT.NET.USER.P2 : Internet users (per 100 people)
- IT.CMP.PCMP.P2 : Personal computers (per 100 people)

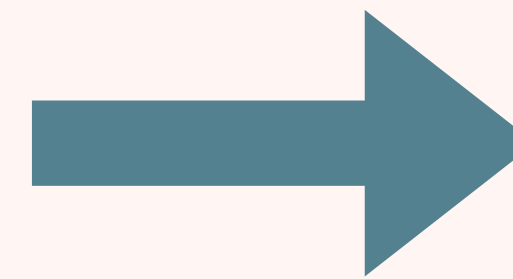
RECHERCHE MEILLEURS INDICATEURS

➤ Taux d'éducation



- SE.SEC.NENR : Net enrolment rate, secondary, both sexes (%)
- SE.TER.ENRR : Gross enrolment ratio, tertiary, both sexes (%)

➤ Taux d'équipement et accès internet



- IT.NET.USER.P2 : Internet users (per 100 people)
- IT.CMP.PCMP.P2 : Personal computers (per 100 people)

➤ Richesse par individu



- NY.GDP.PCAP.CD : GDP per capita (current US\$)

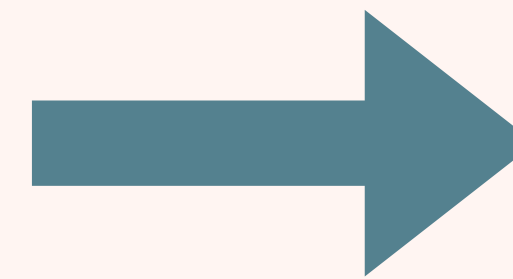
RECHERCHE MEILLEURS INDICATEURS

➤ **Taux d'éducation**



- **SE.SEC.NENR : Net enrolment rate, secondary, both sexes (%)**
- **SE.TER.ENRR : Gross enrolment ratio, tertiary, both sexes (%)**

➤ **Taux d'équipement et accès internet**



- **IT.NET.USER.P2 : Internet users (per 100 people)**
- **IT.CMP.PCMP.P2 : Personal computers (per 100 people)**

➤ **Richesse par individu**










- **NY.GDP.PCAP.CD : GDP per capita (current US\$)**

➤ **Taille marché/difficulté accès au marché**



TAILLE MARCHÉ / DIFFICULTÉ ACCÈS MARCHÉ

➤ Type de gouvernement

Name	Constitutional form	Head of state	Basis of executive legitimacy
 Afghanistan	Provisional	n/a	No constitutionally-defined basis to current regime ^[note 1]
 Albania	Republic	Ceremonial	Ministry is subject to parliamentary confidence
 Algeria	Republic	Executive	Presidency independent of legislature; ministry is subject to parliamentary confidence
 Andorra	Constitutional monarchy	Ceremonial	Ministry is subject to parliamentary confidence
 Angola	Republic	Executive	Presidency is independent of legislature
 Antigua and Barbuda	Constitutional monarchy	Ceremonial	Ministry is subject to parliamentary confidence
 Argentina	Republic	Executive	Presidency is independent of legislature

➤ Pas corrélé

TAILLE MARCHÉ / DIFFICULTÉ ACCÈS MARCHÉ

- Liste languages par pays :
- https://en.wikipedia.org/wiki/List_of_official_languages_by_country_and_territory

Country ↕	Official language ↕	Regional language ↕	Minority language ↕	National language ↕	W ↕
 Abkhazia ^[a]	<ul style="list-style-type: none">AbkhazRussian		Georgian	Abkhaz	
 Afghanistan ^[1]	<ul style="list-style-type: none">PashtoDari	<ul style="list-style-type: none">Uzbek^[b]Turkmen^[b]Pashai^[b]Nuristani^[b]Balochi^[b]Pamiri^[b]		Pashto	
 Albania ^[2]	Albanian		<ul style="list-style-type: none">GreekMacedonianAromanian		Italian
 Algeria ^[3]	<ul style="list-style-type: none">ArabicTamazight			<ul style="list-style-type: none">ArabicTamazight	French

- Language par nombre de locuteurs :
- https://en.wikipedia.org/wiki/List_of_languages_by_total_number_of_speakers

Rank ↕	Language ↕	Family ↕	Branch ↕	First language (L1) speakers ↕	Second language (L2) speakers ↕	Total speakers (L1+L2) ↕
1	English	Indo-European	Germanic	369.9 million	978.2 million ^[5]	1.348 billion
2	Mandarin Chinese (incl. Standard Chinese, but excl. other varieties)	Sino-Tibetan	Sinitic	921.2 million	198.7 million ^[6]	1.120 billion
3	Hindi (excl. Urdu)	Indo-European	Indo-Aryan	342.2 million	258.3 million ^[7]	600 million
4	Spanish	Indo-European	Romance	471.4 million	71.5 million ^[8]	543 million
5	Standard Arabic (excl. dialects)	Afro-Asiatic	Semitic	0 ^[a]	274 million ^[9]	274 million

TAILLE MARCHÉ / DIFFICULTÉ ACCÈS MARCHÉ

- Liste languages par pays :
- https://en.wikipedia.org/wiki/List_of_official_languages_by_country_and_territory

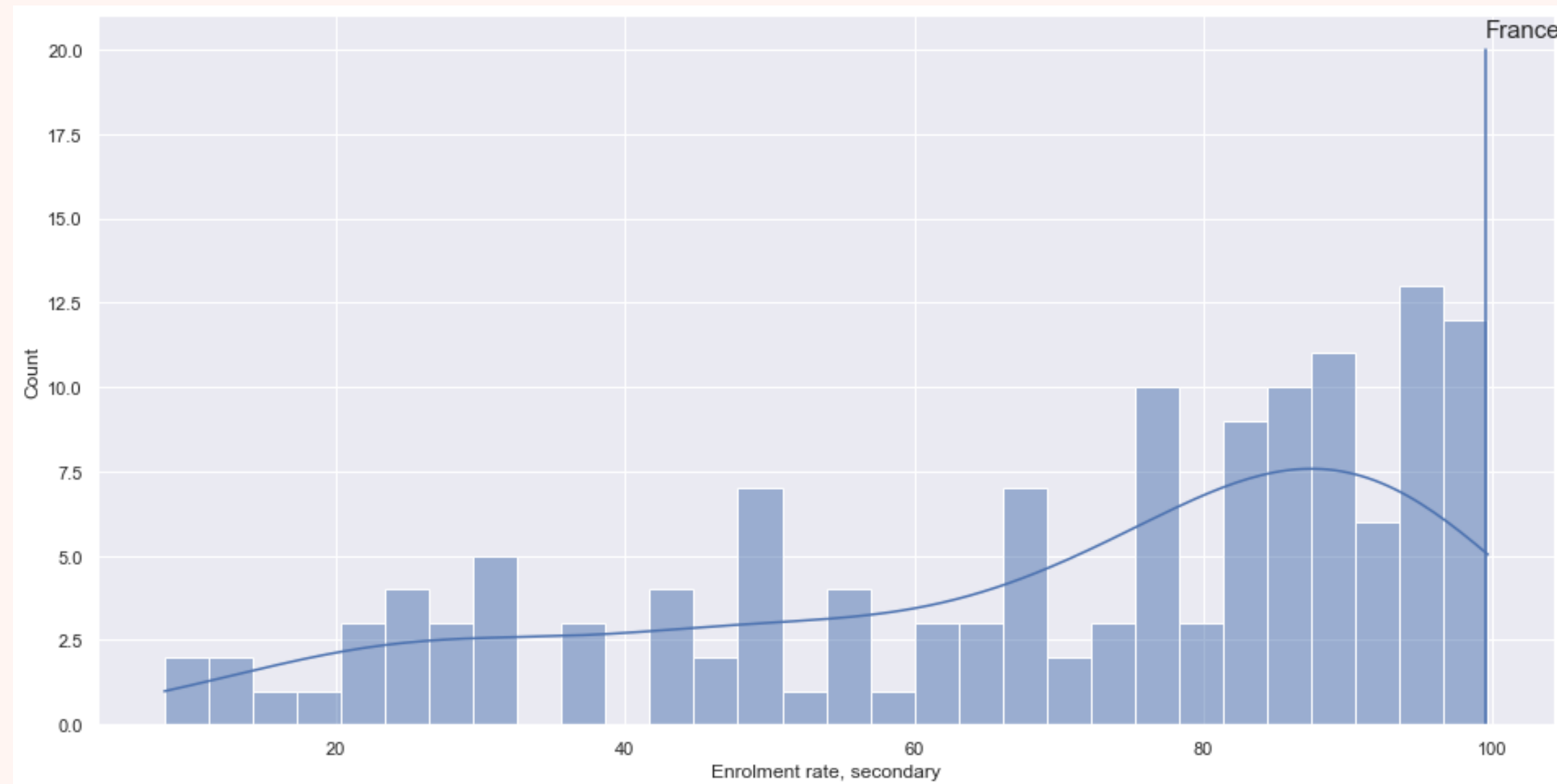
Country ↕	Official language ↕	Regional language ↕	Minority language ↕	National language ↕	W ↕
 Abkhazia ^[a]	<ul style="list-style-type: none">AbkhazRussian		Georgian	Abkhaz	
 Afghanistan ^[1]	<ul style="list-style-type: none">PashtoDari	<ul style="list-style-type: none">Uzbek^[b]Turkmen^[b]Pashai^[b]Nuristani^[b]Balochi^[b]Pamiri^[b]		Pashto	
 Albania ^[2]	Albanian		<ul style="list-style-type: none">GreekMacedonianAromanian		Italian
 Algeria ^[3]	<ul style="list-style-type: none">ArabicTamazight			<ul style="list-style-type: none">ArabicTamazight	French

- Language par nombre de locuteurs :
- https://en.wikipedia.org/wiki/List_of_languages_by_total_number_of_speakers

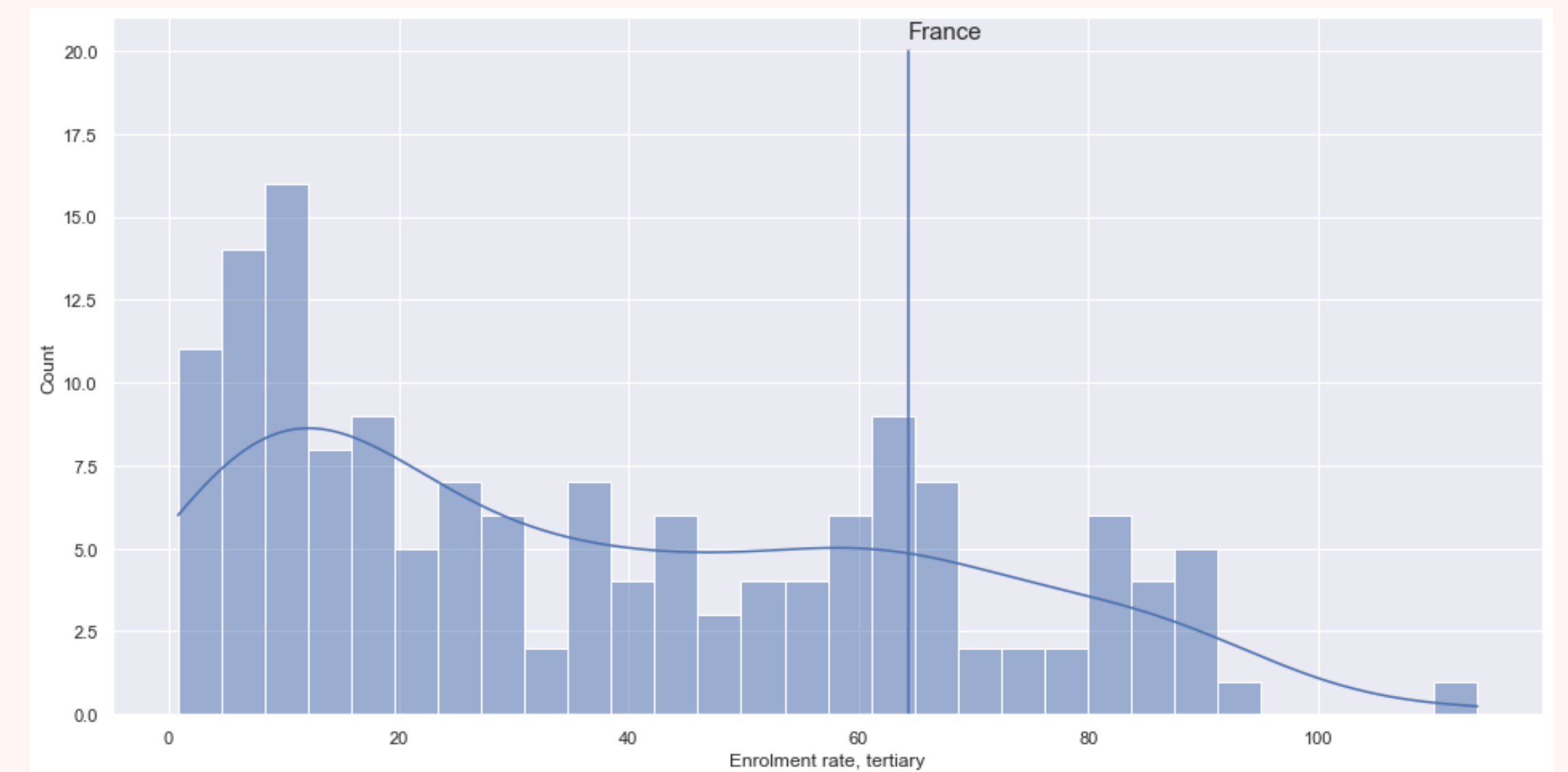
Rank ↕	Language ↕	Family ↕	Branch ↕	First language (L1) speakers ↕	Second language (L2) speakers ↕	Total speakers (L1+L2) ↕
1	English	Indo-European	Germanic	369.9 million	978.2 million ^[5]	1.348 billion
2	Mandarin Chinese (incl. Standard Chinese, but excl. other varieties)	Sino-Tibetan	Sinitic	921.2 million	198.7 million ^[6]	1.120 billion
3	Hindi (excl. Urdu)	Indo-European	Indo-Aryan	342.2 million	258.3 million ^[7]	600 million
4	Spanish	Indo-European	Romance	471.4 million	71.5 million ^[8]	543 million
5	Standard Arabic (excl. dialects)	Afro-Asiatic	Semitic	0 ^[a]	274 million ^[9]	274 million

➤ Plus attractif si pays utilise un des 30 languages les plus parlés

TAUX D'ÉDUCATION

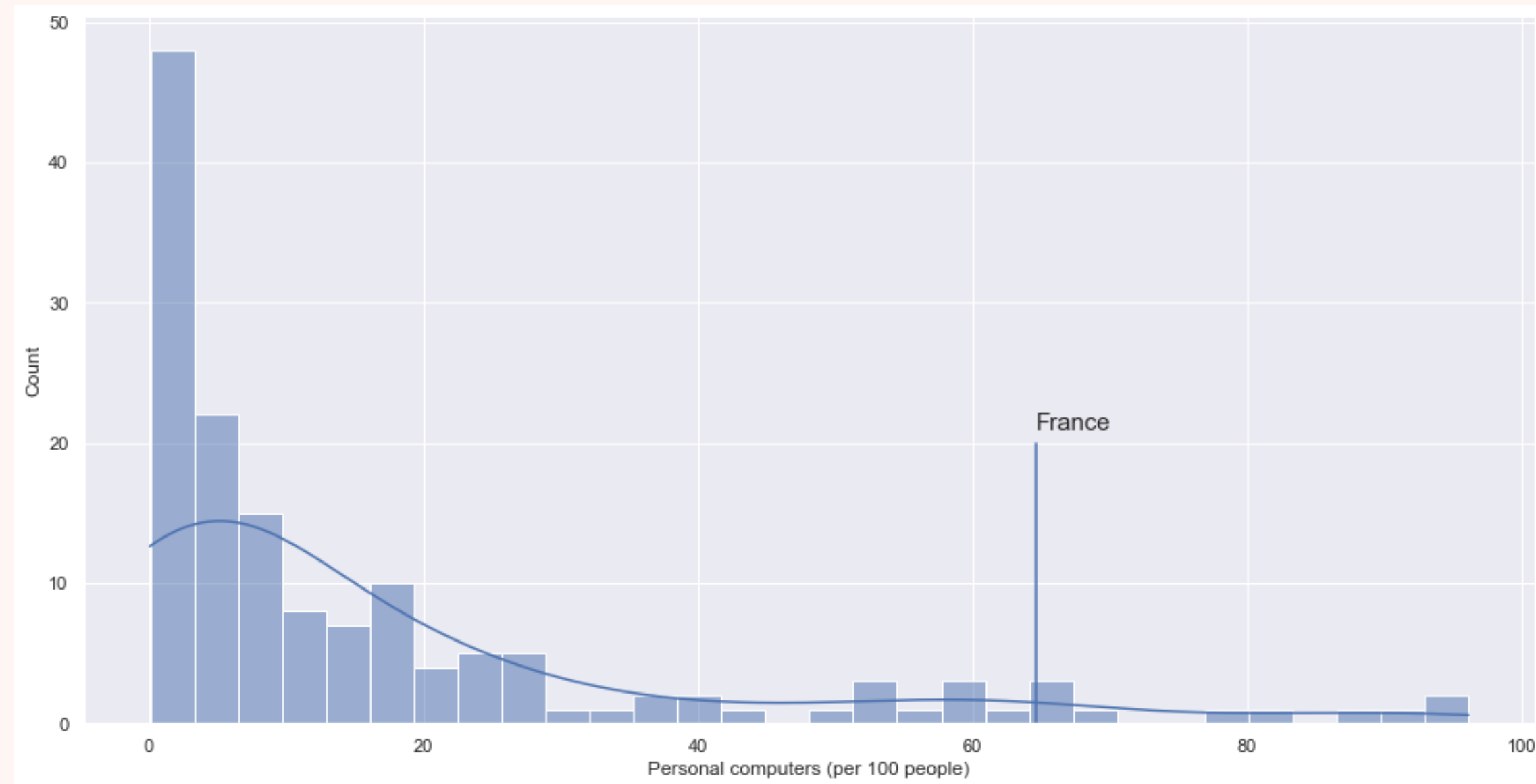


➤ **Enrôlement dans le secondaire**

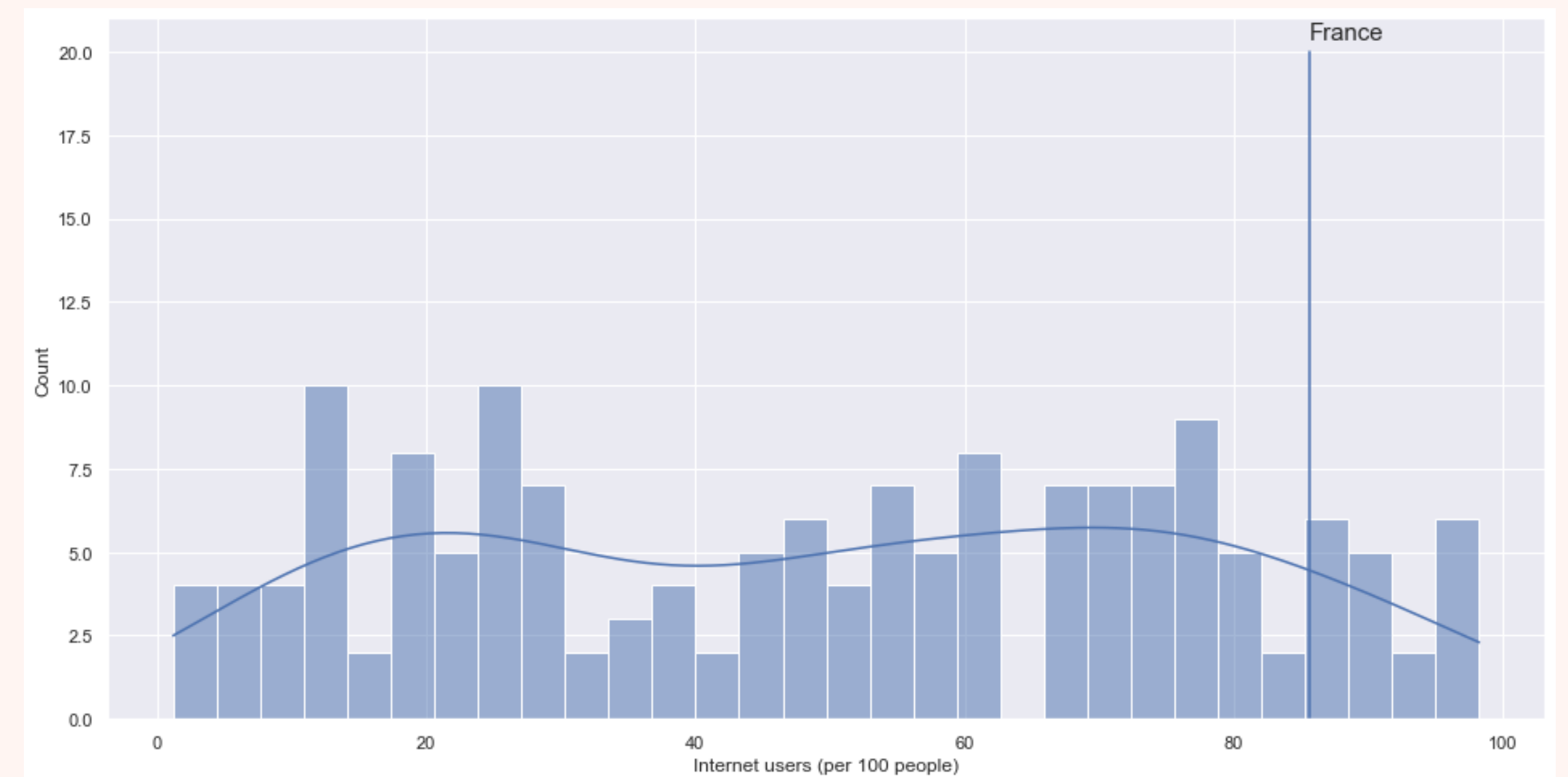


➤ **Enrôlement dans le tertiaire**

TAUX D'ÉQUIPEMENT

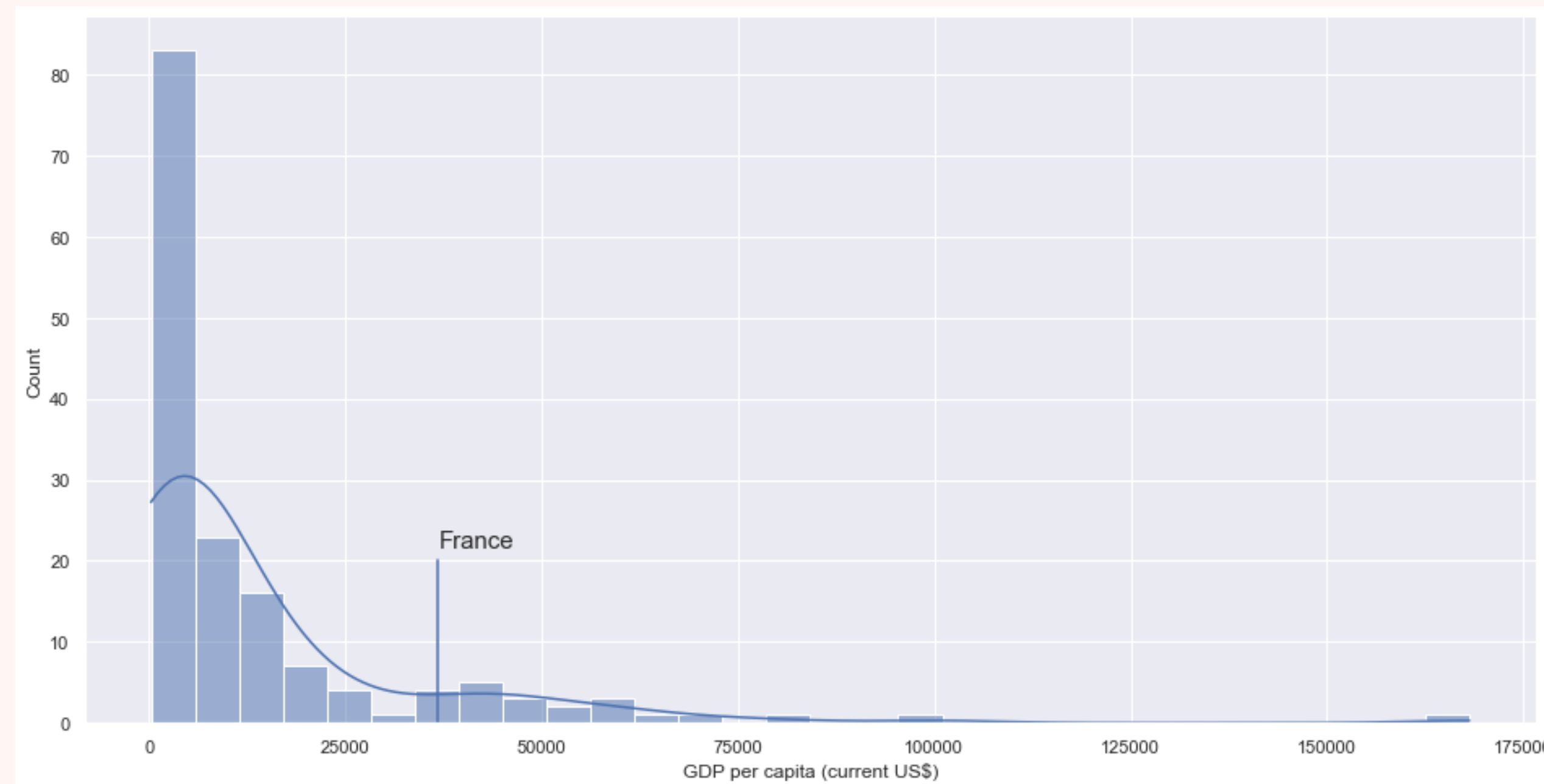


➤ **Nombre d'ordinateurs / 100 pers**



➤ **Utilisateurs d'internet / 100 pers**

RICHESSSE PAR INDIVIDU



➤ PIB / habitant

RESUMÉ DONNÉES SÉLECTIONNÉES

- **Taux d'éducation : enrôlement au second et troisième cycle**
- **Taux d'équipement et accès internet : ordinateur personnel et utilisateurs d'internet/100 habitants**
- **Richesse par individu : PIB/hab**
- **Taille marché/difficulté accès au marché : Vrai si un des langues parlés est dans les 30 langues les plus parlés au monde**



SCORE D'ATTRACTIVITÉ

	Enrolment rate, secondary	Enrolment rate, tertiary	Personal computers (per 100 people)	Internet users (per 100 people)	GDP per capita (current US\$)	Language premium
Afghanistan	48.8789291381836	8.66279983520508	0.390148257118001	10.59572642	561.77874627093	False
Albania	85.914451599121	58.1099510192871	4.59353838043875	66.3634447	4124.98239042866	False
Algeria		36.9222793579102	1.03059686016005	42.94552688	3916.88157088626	True
Angola	12.3750600814819	9.30801963806152	0.646018910072759	13.0	3308.70023298492	True
Antigua and Barbuda	81.6105194091797	23.4862403869629	20.9663699426121	73.0	14462.1762794658	True
Argentina	88.2460021972656	82.9173889160156	9.0561302575817	70.15076431	12440.3209818055	True
Armenia	89.1005783081055	44.3095016479492	9.85101745752868	62.0	3614.68835699811	False
Australia	87.5728530883789	90.3065032958984	60.1409674205855	88.23865805	49927.8195094751	True



Normalisation des indicateurs



$$Z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

SCORE D'ATTRACTIVITÉ

	Enrolment rate, secondary	Enrolment rate, tertiary	Personal computers (per 100 people)	Internet users (per 100 people)	GDP per capita (current US\$)	Language premium
Afghanistan	0.44520115431053237	0.0695568014976082	0.0032707030971145306	0.09703612744753587	0.0016445301492089154	0.0
Albania	0.849842519944298	0.5068556130933852	0.04701281869719139	0.67158850194301	0.022871728613055812	0.0
Algeria		0.3194768983215556	0.009935460830341789	0.4303231121404979	0.02163200227741442	1.0
Angola	0.0463685081233648	0.07526297154483257	0.00593339281942183	0.12180639126557591	0.018008862190316947	1.0
Antigua and Barbuda	0.8028187740049815	0.20065177424764827	0.21739488409714947	0.7399622623134965	0.08445385754964806	1.0
Argentina	0.8753164907536077	0.7262466771122094	0.09345228875326533	0.7106077328172836	0.07240898783276145	1.0
Armenia	0.8846533873747631	0.3848077303527409	0.10172419470750813	0.626633685954711	0.019831736009683937	0.0
Australia	0.8679618226783199	0.7915942423346971	0.625061011007643	0.8969600279884824	0.29573458205321507	1.0

➤ Attribution d'un poids (en fonction de l'importance estimée de l'indicateur)



➤ Weight = (1, 1, 1, 1, 1, 1)

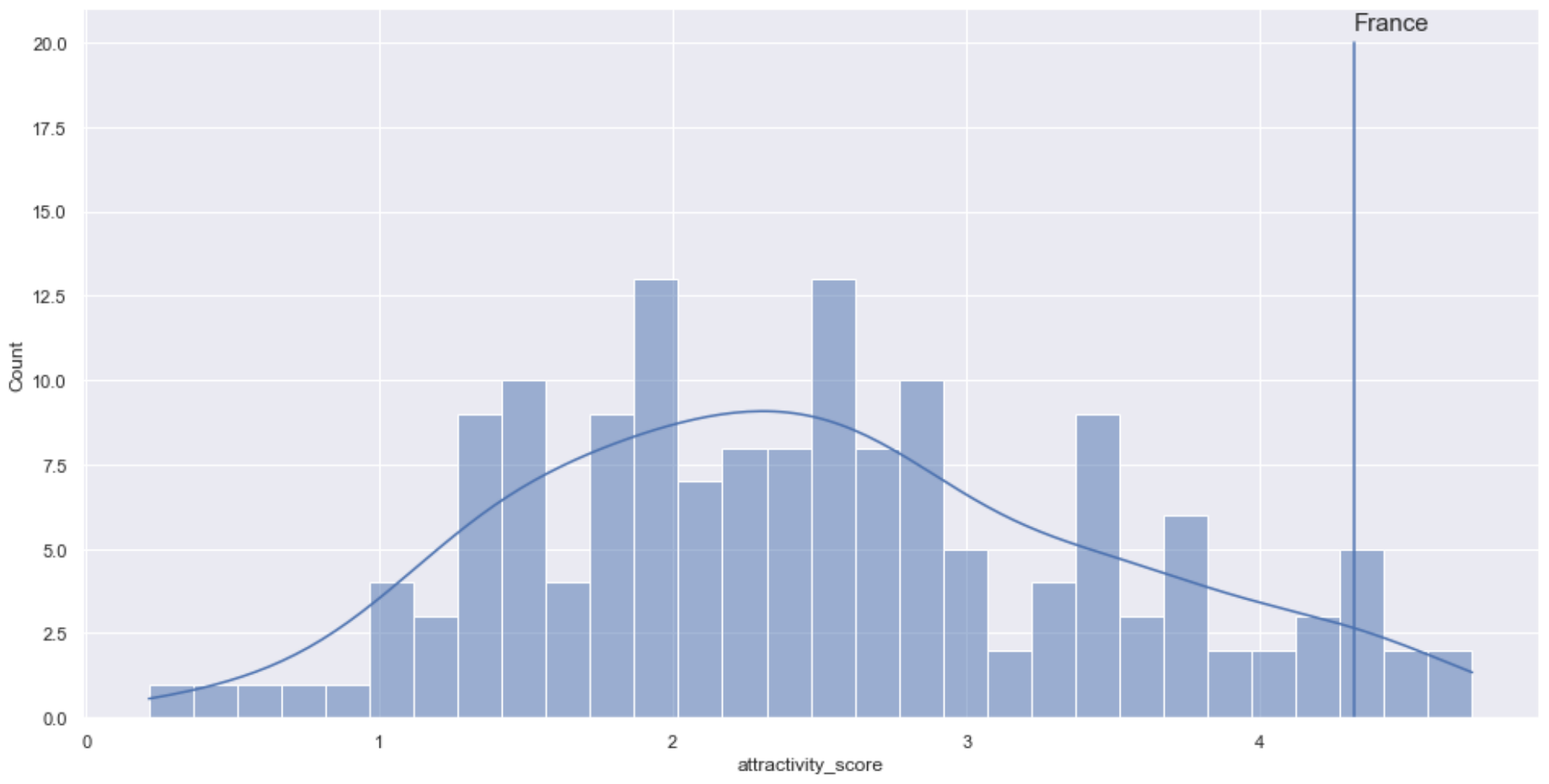
SCORE D'ATTRACTIVITÉ

	Enrolment rate, secondary	Enrolment rate, tertiary	Personal computers (per 100 people)	Internet users (per 100 people)	GDP per capita (current US\$)	Language premium	attractivity_score
Afghanistan	0.44520115431053237	0.0695568014976082	0.0032707030971145306	0.09703612744753587	0.0016445301492089154	0.0	0.6167093165019999
Albania	0.849842519944298	0.5068556130933852	0.04701281869719139	0.67158850194301	0.022871728613055812	0.0	2.098171182290941
Algeria	0.6635007226485137	0.3194768983215556	0.009935460830341789	0.4303231121404979	0.02163200227741442	1.0	2.4448681962183234
Angola	0.0463685081233648	0.07526297154483257	0.00593339281942183	0.12180639126557591	0.018008862190316947	1.0	1.267380125943512
Antigua and Barbuda	0.8028187740049815	0.20065177424764827	0.21739488409714947	0.7399622623134965	0.08445385754964806	1.0	3.045281552212924
Argentina	0.8753164907536077	0.7262466771122094	0.09345228875326533	0.7106077328172836	0.07240898783276145	1.0	3.4780321772691276
Armenia	0.8846533873747631	0.3848077303527409	0.10172419470750813	0.626633685954711	0.019831736009683937	0.0	2.017650734399407
Australia	0.8679618226783199	0.7915942423346971	0.625061011007643	0.8969600279884824	0.29573458205321507	1.0	4.477311686062357

➤ $\text{Score(pays)} = \text{Indicateurs(pays)} * \text{Weight}$

SCORE D'ATTRACTIVITÉ

	attractivity_score	
1		attractivity_score
2	Switzerland	4.7251378749020905
3	United States	4.596434494456767
4	Ireland	4.55599870772758
5	Australia	4.477311686062357
6	New Zealand	4.383306111053217
7	Liechtenstein	4.375225447986494
8	Canada	4.322162853650772
9	France	4.320195442599024
10	Luxembourg	4.29888563973414
11	Japan	4.139334130165732
12	Spain	4.135135144046202
13	Austria	4.128910513882273
14	Belgium	4.12284485001493
15	Germany	4.090829054544024
16	Saudi Arabia	3.964893072686532
17	Bahrain	3.950594435210752
18	San Marino	3.8058135184011195
19	Netherlands	3.7598583397343437
20	Cyprus	3.6974885187384188
21	Sweden	3.697423025227493
22	Norway	3.688944821077721
23	Italy	3.6884234932836586



POUR ALLER PLUS LOIN

- **Améliorer les normalisations en supprimant les min/max trop éloignés**
- **Ajouter d'autres indicateurs**