

cwurData Classement des universités dans le monde de la première à la dernière avec les notes données pour chaque critère

education_expenditure_supplementary_data Donne une note pour les différents types de niveaux pour un certain nombre de pays

educational_achievement_supplementary_data Donne des informations sur l'éducation dans tous les pays du monde ainsi que la répartition de la population selon les âges et le sexe

school_and_country_table Liste toutes les universités ainsi que le pays dans lequel elles se trouvent

shanghaiData Classement des universités dans le monde de la première à la dernière avec les notes données pour chaque critère

timesData Classement des universités dans le monde de la première à la dernière avec les notes données pour chaque critère

In [1]:

```
import pandas as pd
```

In [2]:

```
datacwur = pd.read_csv("/Users/valentinvermes/Desktop/Python_THP/19octobre/universites/cwurData.csv", encoding='latin-1', sep = ',')
```

datacwur

Out[2]:

	world_rank	institution	country	national_rank	quality_of_education	alumni_employ
0	1	Harvard University	USA	1	7	
1	2	Massachusetts Institute of Technology	USA	2	9	
2	3	Stanford University	USA	3	17	
3	4	University of Cambridge	United Kingdom	1	10	
4	5	California Institute of Technology	USA	4	2	
...
2195	996	University of the Algarve	Portugal	7	367	
2196	997	Alexandria University	Egypt	4	236	
2197	998	Federal University of CearÃ	Brazil	18	367	
2198	999	University of A CoruÃ	Spain	40	367	
2199	1000	China Pharmaceutical University	China	83	367	

2200 rows × 14 columns

In [3]:

```
datacwur.year.drop_duplicates()
```

Out[3]:

```
0      2012
100     2013
200     2014
1200    2015
Name: year, dtype: int64
```

In [6]:

```
datashan = pd.read_csv("/Users/valentinvermes/Desktop/Python_THP/19octobre/universites/shanghaiData.csv",encoding='latin-1', sep = ',')

datashan
```

Out[6]:

	world_rank	university_name	national_rank	total_score	alumni	award	hici	ns	
0	1	Harvard University	1	100.0	100.0	100.0	100.0	100.0	100.0
1	2	University of Cambridge	1	73.6	99.8	93.4	53.3	56.6	73.6
2	3	Stanford University	2	73.4	41.1	72.2	88.5	70.9	73.4
3	4	University of California, Berkeley	3	72.8	71.8	76.0	69.4	73.9	72.8
4	5	Massachusetts Institute of Technology (MIT)	4	70.1	74.0	80.6	66.7	65.8	70.1
...
4892	401-500	University of Trieste	11-20	NaN	0.0	0.0	5.0	10.9	0.0
4893	401-500	University of Zaragoza	9-13	NaN	0.0	0.0	7.6	5.1	0.0
4894	401-500	Utah State University	126-146	NaN	13.6	0.0	3.6	10.8	0.0
4895	401-500	Vienna University of Technology	4-6	NaN	0.0	0.0	0.0	12.2	0.0
4896	401-500	Wake Forest University	126-146	NaN	0.0	0.0	14.9	7.5	0.0

4897 rows × 11 columns

In [7]:

```
datashan.year.drop_duplicates()
```

Out[7]:

```
0      2005
500     2006
1000    2007
1510    2008
2013    2009
2514    2010
3014    2011
3514    2012
3798    2013
3897    2014
4397    2015
Name: year, dtype: int64
```

In [8]:

```
datatimes = pd.read_csv("/Users/valentinvermes/Desktop/Python_THP/19octobre/universites/timesData.csv", encoding='latin-1', sep = ',')
```

datatimes

Out[8]:

	world_rank	university_name	country	teaching	international	research	citations	incon
0	1	Harvard University	United States of America	99.7	72.4	98.7	98.8	34
1	2	California Institute of Technology	United States of America	97.7	54.6	98.0	99.9	83
2	3	Massachusetts Institute of Technology	United States of America	97.8	82.3	91.4	99.9	87
3	4	Stanford University	United States of America	98.3	29.5	98.1	99.2	64
4	5	Princeton University	United States of America	90.9	70.3	95.4	99.9	
...	
2598	601-800	Yeungnam University	South Korea	18.6	24.3	10.9	26.5	35
2599	601-800	Yıldırım Technical University	Turkey	14.5	14.9	7.6	19.3	44
2600	601-800	Yokohama City University	Japan	24.0	16.1	10.2	36.4	37
2601	601-800	Yokohama National University	Japan	20.1	23.3	16.0	13.5	40
2602	601-800	Yuan Ze University	Taiwan	16.2	17.7	18.3	28.6	39

2603 rows × 14 columns

In [9]:

```
datatimes.year.drop_duplicates()
```

Out[9]:

```
0      2011
200    2012
602    2013
1002   2014
1402   2015
1803   2016
Name: year, dtype: int64
```

In [14]:

```
print(
    'Datacwur :', datacwur['year'].value_counts().mean(),
    'DataShan :', datashan['year'].value_counts().mean(),
    'Datatimes :', datatimes['year'].value_counts().mean()
)
```

```
Datacwur : 550.0 DataShan : 445.1818181818182 Datatimes : 433.833333
3333333
```

In [16]:

```
dataeduc = pd.read_csv("/Users/valentinvermes/Desktop/Python_THP/19octobre/universites/educational_attainment_supplementary_data.csv",encoding='latin-1', sep = ',')

dataeduc
```

Out[16]:

	country_name	series_name	1985	1986	1987	1990	1991	1992	1993	1995	...	200
0	Afghanistan	Barro-Lee: Average years of primary schooling,...	0.33	NaN	NaN	0.44	NaN	NaN	NaN	0.57	...	0.8
1	Afghanistan	Barro-Lee: Average years of primary schooling,...	1.03	NaN	NaN	1.26	NaN	NaN	NaN	1.54	...	2.1
2	Afghanistan	Barro-Lee: Average years of primary schooling,...	0.83	NaN	NaN	0.95	NaN	NaN	NaN	1.26	...	1.0
3	Afghanistan	Barro-Lee: Average years of primary schooling,...	2.34	NaN	NaN	2.22	NaN	NaN	NaN	2.37	...	2.2
4	Afghanistan	Barro-Lee: Average years of primary schooling,...	0.54	NaN	NaN	0.92	NaN	NaN	NaN	0.94	...	2.0
...
79050	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN
79051	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN
79052	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN
79053	Data from database: Education Statistics: Educ...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN
79054	Last Updated: 10/20/2015	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN

79055 rows × 29 columns



In [20]:

```
datastrasbourg = pd.concat([datacwur, datashan, datatimes])
datastrasbourg
```

Out[20]:

	world_rank	institution	country	national_rank	quality_of_education	alumni_employi
0	1	Harvard University	USA	1	7.0	
1	2	Massachusetts Institute of Technology	USA	2	9.0	
2	3	Stanford University	USA	3	17.0	
3	4	University of Cambridge	United Kingdom	1	10.0	
4	5	California Institute of Technology	USA	4	2.0	
...
2598	601-800	NaN	South Korea	NaN	NaN	
2599	601-800	NaN	Turkey	NaN	NaN	
2600	601-800	NaN	Japan	NaN	NaN	
2601	601-800	NaN	Japan	NaN	NaN	
2602	601-800	NaN	Taiwan	NaN	NaN	

9700 rows × 30 columns

In [40]:

```
rang_strasbourg = datastrasbourg['institution'].str.contains('University of Strasbourg', na = False)
```

In [41]:

```
rang_strasbourg
```

Out[41]:

```
0      False
1      False
2      False
3      False
4      False
...
2598   False
2599   False
2600   False
2601   False
2602   False
Name: institution, Length: 9700, dtype: bool
```

In [42]:

```
datastrasbourg.loc[rang_strasbourg,:][['world_rank','university_name','year']]
```

Out[42]:

	world_rank	university_name	year
375	176	NaN	2014
1381	182	NaN	2015

In [43]:

```
datatimes.head(20)
```

Out[43]:

	world_rank	university_name	country	teaching	international	research	citations	incoi
0	1	Harvard University	United States of America	99.7	72.4	98.7	98.8	3
1	2	California Institute of Technology	United States of America	97.7	54.6	98.0	99.9	8
2	3	Massachusetts Institute of Technology	United States of America	97.8	82.3	91.4	99.9	8
3	4	Stanford University	United States of America	98.3	29.5	98.1	99.2	6
4	5	Princeton University	United States of America	90.9	70.3	95.4	99.9	
5	6	University of Cambridge	United Kingdom	90.5	77.7	94.1	94.0	5
6	6	University of Oxford	United Kingdom	88.2	77.2	93.9	95.1	7
7	8	University of California, Berkeley	United States of America	84.2	39.6	99.3	97.8	
8	9	Imperial College London	United Kingdom	89.2	90.0	94.5	88.3	9
9	10	Yale University	United States of America	92.1	59.2	89.7	91.5	
10	11	University of California, Los Angeles	United States of America	83.0	48.1	92.9	93.2	
11	12	University of Chicago	United States of America	79.1	62.8	87.9	96.9	
12	13	Johns Hopkins University	United States of America	80.9	58.5	89.2	92.3	10
13	14	Cornell University	United States of America	82.2	62.4	88.8	88.1	3
14	15	ETH Zurich "Swiss Federal Institute of Tech...	Switzerland	77.5	93.7	87.8	83.1	
15	15	University of Michigan	United States of America	83.9	53.3	89.1	84.1	5
16	17	University of Toronto	Canada	75.8	-	87.9	82.2	
17	18	Columbia University	United States of America	73.8	90.9	73.8	92.6	

	world_rank	university_name	country	teaching	international	research	citations	income
18	19	University of Pennsylvania	United States of America	71.8	32.9	82.7	93.6	4
19	20	Carnegie Mellon University	United States of America	70.3	39.1	79.3	95.7	5

In [46]:

```
datatimes2016 = datatimes.loc[datatimes['year'] == 2016]
```

In [47]:

```
datatimes2016
```

Out[47]:

	world_rank	university_name	country	teaching	international	research	citations	income
1803	1	California Institute of Technology	United States of America	95.6	64.0	97.6	99.8	9
1804	2	University of Oxford	United Kingdom	86.5	94.4	98.9	98.8	7
1805	3	Stanford University	United States of America	92.5	76.3	96.2	99.9	6
1806	4	University of Cambridge	United Kingdom	88.2	91.5	96.7	97.0	5
1807	5	Massachusetts Institute of Technology	United States of America	89.4	84.0	88.6	99.7	9
...
2598	601-800	Yeungnam University	South Korea	18.6	24.3	10.9	26.5	3
2599	601-800	Yildiz Technical University	Turkey	14.5	14.9	7.6	19.3	4
2600	601-800	Yokohama City University	Japan	24.0	16.1	10.2	36.4	3
2601	601-800	Yokohama National University	Japan	20.1	23.3	16.0	13.5	4
2602	601-800	Yuan Ze University	Taiwan	16.2	17.7	18.3	28.6	3

800 rows × 14 columns

In [48]:

```
datatimes2016.head(20)
```

Out[48]:

	world_rank	university_name	country	teaching	international	research	citations	inc
1803	1	California Institute of Technology	United States of America	95.6	64.0	97.6	99.8	
1804	2	University of Oxford	United Kingdom	86.5	94.4	98.9	98.8	
1805	3	Stanford University	United States of America	92.5	76.3	96.2	99.9	
1806	4	University of Cambridge	United Kingdom	88.2	91.5	96.7	97.0	
1807	5	Massachusetts Institute of Technology	United States of America	89.4	84.0	88.6	99.7	
1808	6	Harvard University	United States of America	83.6	77.2	99.0	99.8	
1809	7	Princeton University	United States of America	85.1	78.5	91.9	99.3	
1810	8	Imperial College London	United Kingdom	83.3	96.0	88.5	96.7	
1811	9	ETH Zurich "Swiss Federal Institute of Tech..."	Switzerland	77.0	97.9	95.0	91.1	
1812	10	University of Chicago	United States of America	85.7	65.0	88.9	99.2	
1813	11	Johns Hopkins University	United States of America	77.6	70.0	90.4	98.2	
1814	12	Yale University	United States of America	86.5	64.3	87.8	97.2	
1815	13	University of California, Berkeley	United States of America	80.4	61.9	91.1	99.7	
1816	14	University College London	United Kingdom	78.1	94.4	91.0	94.2	
1817	15	Columbia University	United States of America	85.9	73.5	82.2	98.1	
1818	16	University of California, Los Angeles	United States of America	80.8	56.4	88.6	98.5	
1819	17	University of Pennsylvania	United States of America	82.0	49.5	86.9	98.6	
1820	18	Cornell University	United States of America	77.9	63.9	86.1	97.2	
1821	19	University of Toronto	Canada	75.9	77.8	89.3	90.9	

	world_rank	university_name	country	teaching	international	research	citations	inc
1822	20	Duke University	United States of America	76.0	56.5	78.0	99.0	

In [49]:

```
datatimes2016 = datatimes2016.head(20)
```

In [50]:

```
datatimes2016
```

Out[50]:

	world_rank	university_name	country	teaching	international	research	citations	inc
1803	1	California Institute of Technology	United States of America	95.6	64.0	97.6	99.8	
1804	2	University of Oxford	United Kingdom	86.5	94.4	98.9	98.8	
1805	3	Stanford University	United States of America	92.5	76.3	96.2	99.9	
1806	4	University of Cambridge	United Kingdom	88.2	91.5	96.7	97.0	
1807	5	Massachusetts Institute of Technology	United States of America	89.4	84.0	88.6	99.7	
1808	6	Harvard University	United States of America	83.6	77.2	99.0	99.8	
1809	7	Princeton University	United States of America	85.1	78.5	91.9	99.3	
1810	8	Imperial College London	United Kingdom	83.3	96.0	88.5	96.7	
1811	9	ETH Zurich "Swiss Federal Institute of Tech..."	Switzerland	77.0	97.9	95.0	91.1	
1812	10	University of Chicago	United States of America	85.7	65.0	88.9	99.2	
1813	11	Johns Hopkins University	United States of America	77.6	70.0	90.4	98.2	
1814	12	Yale University	United States of America	86.5	64.3	87.8	97.2	
1815	13	University of California, Berkeley	United States of America	80.4	61.9	91.1	99.7	
1816	14	University College London	United Kingdom	78.1	94.4	91.0	94.2	
1817	15	Columbia University	United States of America	85.9	73.5	82.2	98.1	
1818	16	University of California, Los Angeles	United States of America	80.8	56.4	88.6	98.5	
1819	17	University of Pennsylvania	United States of America	82.0	49.5	86.9	98.6	
1820	18	Cornell University	United States of America	77.9	63.9	86.1	97.2	
1821	19	University of Toronto	Canada	75.9	77.8	89.3	90.9	

	world_rank	university_name	country	teaching	international	research	citations	inc
1822	20	Duke University	United States of America	76.0	56.5	78.0	99.0	

In [51]:

```
datatimes2016.loc[datatimes['country'] == 'United States of America']
```

Out[51]:

	world_rank	university_name	country	teaching	international	research	citations	incon
1803	1	California Institute of Technology	United States of America	95.6	64.0	97.6	99.8	97
1805	3	Stanford University	United States of America	92.5	76.3	96.2	99.9	63
1807	5	Massachusetts Institute of Technology	United States of America	89.4	84.0	88.6	99.7	95
1808	6	Harvard University	United States of America	83.6	77.2	99.0	99.8	45
1809	7	Princeton University	United States of America	85.1	78.5	91.9	99.3	52
1812	10	University of Chicago	United States of America	85.7	65.0	88.9	99.2	36
1813	11	Johns Hopkins University	United States of America	77.6	70.0	90.4	98.2	100
1814	12	Yale University	United States of America	86.5	64.3	87.8	97.2	43
1815	13	University of California, Berkeley	United States of America	80.4	61.9	91.1	99.7	47
1817	15	Columbia University	United States of America	85.9	73.5	82.2	98.1	
1818	16	University of California, Los Angeles	United States of America	80.8	56.4	88.6	98.5	47
1819	17	University of Pennsylvania	United States of America	82.0	49.5	86.9	98.6	47
1820	18	Cornell University	United States of America	77.9	63.9	86.1	97.2	33
1822	20	Duke University	United States of America	76.0	56.5	78.0	99.0	100

In [52]:

```
dataschoolcoun = pd.read_csv("/Users/valentinvermes/Desktop/Python_THP/19octobre/universites/school_and_country_table.csv", encoding='latin-1', sep = ',')
```

dataschoolcoun

Out[52]:

	school_name	country
0	Harvard University	United States of America
1	California Institute of Technology	United States of America
2	Massachusetts Institute of Technology	United States of America
3	Stanford University	United States of America
4	Princeton University	United States of America
...
813	Xidian University	China
814	Yeungnam University	South Korea
815	Yıldırım Technical University	Turkey
816	Yokohama City University	Japan
817	Yokohama National University	Japan

818 rows × 2 columns

In [53]:

```
datashan2015 = pd.merge(datashan, dataschoolcoun, left_on="university_name", right_on="school_name")
datashan2015 = datashan2015.loc[datashan2015['year'] == 2015]
```

In [55]:

```
datashan2015['country'].value_counts().sort_values(ascending=False)
```

Out[55]:

United States of America	90
United Kingdom	26
China	22
Canada	16
Italy	15
Germany	14
Japan	14
Australia	13
Spain	12
Netherlands	10
South Korea	9
Sweden	8
Taiwan	6
Finland	5
Switzerland	5
Austria	5
Denmark	5
France	5
South Africa	4
Brazil	3
Israel	3
Portugal	3
Belgium	3
Republic of Ireland	3
Singapore	2
Greece	2
Poland	2
Norway	2
Iran	2
Saudi Arabia	2
Mexico	1
New Zealand	1
Slovenia	1
India	1
Hungary	1
Egypt	1
Chile	1
Turkey	1
Hong Kong	1
Russian Federation	1
Czech Republic	1
Serbia	1

Name: country, dtype: int64

In [61]:

```
datacwur.loc[datacwur['country'] == 'France'].institution.unique()
```

Out[61]:

```
array(['University of Paris-Sud', 'École normale supérieure - Paris',
      'École Polytechnique', 'Pierre-and-Marie-Curie University',
      'Mines ParisTech', 'Paris Diderot University - Paris 7',
      'University of Strasbourg', 'Joseph Fourier University',
      'Aix-Marseille University', 'Claude Bernard University Lyon 1',
      'University of Montpellier 2',
      'Paul Sabatier University - Toulouse III',
      'Paris Descartes University', 'University of Lorraine',
      'University of Bordeaux II',
      'Versailles Saint-Quentin-en-Yvelines University',
      'University of Bordeaux I',
      'École normale supérieure de Lyon', 'University of Rennes 1',
      'Grenoble Institute of Technology', 'École Centrale Paris',
      'University of Montpellier 1',
      'University of Nice Sophia Antipolis', 'University of Savoy',
      'University of Nantes',
      'University Lille 1: Sciences and Technologies',
      'National Polytechnic Institute of Toulouse',
      'Paris-Est Créteil Val-de-Marne University',
      'University of Poitiers', 'University of Burgundy',
      'Lille 2 University of Health and Law',
      'University François Rabelais, Tours', 'University of Paris 13',
      'University of Évry Val d'Essonne',
      'Paul Valéry University, Montpellier III',
      'University of Caen Lower Normandy',
      'Institut national des sciences Appliquées de Lyon',
      'University of Perpignan Via Domitia', 'Blaise Pascal University',
      'University of Rouen', 'University of Picardie Jules Verne',
      'University of Franche-Comté', 'University of Angers',
      'Jean Monnet University', 'University of Orléans',
      'University of Western Brittany',
      'University of Reims Champagne-Ardenne',
      'University of Valenciennes and Hainaut-Cambresis',
      'University of Pau and Pays de l'Adour',
      'University of Maine (Le Mans)', 'University of Montpellier',
      'University of Bordeaux',
      'École normale supérieure de Cachan',
      'École centrale de Lyon'], dtype=object)
```

In []: