

University Rankings

PROJECT BRIEF

JUDITH IFARINDE LAY

Executive Summary

Universities are located all around the world. Due to the number of options available, determining the best university can take time and effort. Universities are ranked according to many factors, such as research, teaching practices, alums, and employer reputation, to name a few. There are numerous national and international ranking systems. These ranking systems often have differing views on which universities are the best.

In this analysis, we will examine three global university rankings: the Times Higher Education World University Ranking (THE), the Academic Ranking of World Universities (AWU), and the Center for World University Rankings (CWUR). We will explore the geographical trends and how these ranking systems compare to each other, the education expenditures, and the educational attainment.

Data Source

The data for this project is an open-source data download from <https://www.kaggle.com/mylesoneill/world-university-rankings>

Data Profile

Data Cleaning

- Completed Data Consistency Checklist for all datasets:
 - Check the Descriptive Statistics
 - Check for Mixed Datatypes
 - Adjusted Datatypes, if necessary
 - Check for Missing Values
 - Removed unnecessary columns
 - Check for Duplicates
- Exported Checked Dataframes (datasets)

Limitation and Ethics

- In the Center for World University Rankings (CWUR), the 'broad_impact' column only has data available for 2014 and 2015.
- The supplementary datasets, education attainment, and educational expenditure contain information from 1985 through 2010. They only have three years in common: 1995, 200, 2005, and 2010.
- There are quite a few missing values in the datasets, which may limit accurate and concise analysis of the datasets.
- In the education expenditures dataset, for the 'direct_expenditure_type', private and total data are only available for 2011.

Data Quality

Center for World University Rankings (CWUR)

Dimensions: Pre-Clean was 2200 rows x 14 columns; Post-Clean is 2200 rows x 14 columns.

Column	Description	Datatype	Time Variant
world_rank	world rank for university	Qualitative / Ordinal	yes
institution	name of university	Quantitative / Nominal	no
country	country of each university	Quantitative / Nominal	no
national_rank	rank of university within its country	Qualitative / Ordinal	yes
quality_of_education	rank for quality of education	Qualitative / Ordinal	yes
alumni_employment	rank for alumni employment	Qualitative / Ordinal	yes
quality_of_faculty	rank for quality of faculty	Qualitative / Ordinal	yes
publications	rank for publications	Qualitative / Ordinal	yes
influence	rank for influence	Qualitative / Ordinal	yes
citations	number of students at the university	Qualitative / Ordinal	yes
broad_impact	rank for broad impact	Qualitative / Ordinal	yes
patents	rank for patents	Qualitative / Ordinal	yes
score	total score, used for world rank	Qualitative / Ordinal	yes
year	year of ranking	Quantitative / Discrete	no

Times Higher Education World University Ranking (THE)

Dimensions: Pre-Clean was 2603 rows x 14 columns; Post-Clean is 2603 rows x 14 columns.

Column	Description	Datatype	Time Variant
world_rank	world rank for the university	Qualitative / Ordinal	no
university_name	name of the university	Quantitative / Nominal	no
country	country of each university	Quantitative / Nominal	no
teaching	university score for teaching (the learning environment)	Qualitative / Ordinal	yes
international	university score international outlook (staff, students, research)	Qualitative / Ordinal	yes
research	university score for research (volume, income, and reputation)	Qualitative / Ordinal	yes
citations	university score for citations (research influence)	Qualitative / Ordinal	yes
income	university score for industry income (knowledge transfer)	Qualitative / Ordinal	yes
total_score	total score for the university used to determine rank	Qualitative / Ordinal	yes
num_students	number of students at the university	Qualitative / Ordinal	yes
student_staff_ratio	number of students divided by the number of staff	Qualitative / Ordinal	yes
international_students	percentage of international students	Qualitative / Ordinal	yes
female_male_ratio	Female student to Male student ratio	Qualitative / Ordinal	yes
year	year of the ranking	Quantitative / Discrete	no

Academic Ranking of World Universities (AWU)

Dimensions: Pre-Clean was 4897 rows x 11 columns; Post-Clean is 4896 rows x 11 columns.

Column	Description	Datatype	Time Variant
world_rank	world rank of university	Qualitative / Ordinal	no
university_name	name of university	Qualitative / Ordinal	no
national_rank	rank of university within its country	Qualitative / Ordinal	yes
total_score	total score, use to determine the rank	Qualitative / Ordinal	yes
alumni	Alumni Score, based on the number of alumni of an institution winning Nobel prizes and fields medals	Qualitative / Ordinal	yes
award	Award Score, based on the number of staff of an institution winning Nobel Prizes in Physics, Chemistry, Medicine, and Economics and Fields Medals in Mathematics	Qualitative / Ordinal	yes
hici	HiCi Score, based on the number of Highly Cited Researchers selected by Thomson Reuters	Qualitative / Ordinal	yes
ns	N&S Score, based on the number of papers published in Nature and Science	Qualitative / Ordinal	yes
pub	PUB Score, based on the total number of papers indexed in the Science Citation Index-Expanded and Social Science Citation Index	Qualitative / Ordinal	yes
pcp	PCP Score, the weighted scores of the above five indicators divided by the number of full-time academic staff	Qualitative / Ordinal	yes
year	year of ranking	Quantitative / Discrete	no

School and Country

Dimensions: Pre-Clean was 818 rows x 2 columns; Post-Clean is 818 rows x 2 columns.

Column	Description	Datatype	Time Variant
school_name	university name	Quantitative / Nominal	no
country	location of university	Quantitative / Nominal	no

Educational Attainment

Dimensions: Pre-Clean was 79055 rows x 29 columns; Post-Clean is 79053 rows x 8 columns.

Column	Description	Datatype	Time Variant
country_name	name of country	Quantitative / Nominal	no
series_name	Age grouping type	Qualitative / Ordinal	no
1985	Year of collection	Quantitative / Discrete	no
1990	Year of collection	Quantitative / Discrete	no
1995	Year of collection	Quantitative / Discrete	no
2000	Year of collection	Quantitative / Discrete	no
2005	Year of collection	Quantitative / Discrete	no
2010	Year of collection	Quantitative / Discrete	no

Educational Expenditures

Dimensions: Pre-Clean was 333 rows x 9 columns; Post-Clean is 333 rows x 9 columns.

Column	Description	Datatype	Time Variant
country	A set of OECD countries, plus Brazil and the Russian Federation, plus an OECD average.	Quantitative / Nominal	no
institute_type	All Institutes (including preprimary education and subsidies to households, not separately shown), Elementary and	Qualitative / Ordinal	no
direct_expenditure_type	Public direct expenditure, private direct expenditure, or total (public + private) direct expenditure.	Qualitative / Ordinal	yes
1995	Year of collection	Quantitative / Discrete	no
2000	Year of collection	Quantitative / Discrete	no
2005	Year of collection	Quantitative / Discrete	no
2009	Year of collection	Quantitative / Discrete	no
2010	Year of collection	Quantitative / Discrete	no
2011	Year of collection	Quantitative / Discrete	no

Questions to Explore

- What are the top five universities when considering all three ranking systems?
- Which countries have the highest-ranked universities?
- How does national educational attainment relate to the quality of each nation's universities?
- Does spending more on education lead to better university rankings?