WORLD UNIVERSITY RANKINGS 2026

Analysis to explore global leading
institutions ranked as per
6 ranking indicators
by QS World University Rankings

BY
OLGA GAFFAROVA

Table of Content

	Page
Why I chose this data set?	3
QS World University Rankings 2026: Questions to explore	3
What QS rankings is about?	4
Data Source	4
Data Copyright	4
How data is collected?	4
How does QS evaluate universities' performance?	4
Data biases	4
Lens and indicators	4
Lens and indicators description	5
QS Institution Classifications	6
Size Classifications	6
Subject Range Classifications	6
Research Intensity & Focus Classifications	6
Faculty & Subject Area Classification	7
How does QS calculate Scores and Ranks?	8
Data Limitations and Ethical Concideration	9
Links to project files	10

Why did I choose this data set?

Would like to understand what is behind these ratings and how valuable and relevant all these data could be for future students when they make their choises, since I have a daughter and aim to get some insights so I can guide her in the future if needed.

QS World University Rankings 2026: Questions to explore

Ranking & Score Analysis

- Which factors correlate most strongly with the Overall Score? (e.g., Is Academic Reputation or Employer Reputation more important?)
- Do universities with higher Overall Scores also rank high across all sub-scores, or are there trade-offs?
- What is the distribution of Overall Scores across universities? (Are most clustered around the mean, or are there only a few top performers?)

Regional & Country Comparisons

- Which regions (Asia, Europe, Americas, etc.) dominate the top 100 universities?
- How do average Overall Scores differ across countries or regions?
- Which countries have the largest number of ranked universities?
- Are certain regions stronger in specific dimensions (e.g., research intensity in North America vs. teaching in Europe)?

Institutional Characteristics

- Does university Size (S, M, L, XL) correlate with better rankings?
- Do Public vs. Private institutions show differences in Overall Score or sub-scores?
- Does the type of Focus (Comprehensive vs. Focused) impact performance?

Score Dimension Insights

- Is there a relationship between Faculty-Student Ratio (FSR SCORE) and Academic Reputation (AR SCORE)?
- Do universities with strong International Faculty Ratio (IFR SCORE) also perform well in International Student Ratio (ISR SCORE)?
- Which dimension (e.g., Employer Reputation vs. International Research Network) tends to pull universities up or down in rank?

Year-to-Year Changes

- Which universities improved the most compared to their Previous Rank?
- Which universities dropped the most in rank?
- Do certain countries or regions show more upward mobility year to year?

Outliers & Patterns

- Are there universities that perform exceptionally well in one score (e.g., Sustainability) but poorly in others?
- Do any mid-ranked universities outperform top universities in specific metrics?
- Which universities are consistent across all dimensions, and which are more specialized?

QS World University Rankings 2026: Top global universities

What this rankings is about?

Over 1,500 of the world's top universities are included in the 2026 edition of the QS World University Rankings, with over 100 locations represented around the world.

Data Source - External / Open data
QS Quacquarelli Symonds (www.topuniversities.com).

Data Copyright — QS Quacquarelli Symonds rankings data is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License
Creative Commons License

Data Reporting Period: Academic Year — All the cycles start in November of each year - so, when the cycle begins in November 2025, QS teams are collecting data for the 2024 academic year.

How data is collected?

They claim to "collect data from 16.4 million academic papers, and gather insights from over 151,000 academics and 100,000 employers", and focus on "things like student experience, global partnerships, research opportunities, and the quality of teaching".
The reason why they elaborate universities' rankings is to offer students and applicants "the clearest picture of the universities that can help them achieve their ambitions."

How do they evaluate universities' performance?

They use four different sets of measurements:

1. "Lens: A broad theme made up of several related indicators. For example, a lens could focus on employability or research.

2. Indicator: A specific measurement of an institution's performance in one area, like academic reputation or student-to-faculty ratio.

3. Metric: A more precise calculation used within an indicator to determine how universities compare.

4. Each ranking has its own mix of lenses, indicators, and metrics. The final scores and rankings are determined by the weightings we assign to each of these areas, ensuring a fair and comprehensive view of every institution."

Lens and indicators

Below you can find the weightings for each performance lens and indicator in this ranking. Weightings are reviewed on an annual basis.

A metric used to calculate the final (overall) score for each ranking. A ranking can have multiple indicators. The weights of each indicator add up to 100% of the overall score. An example would be 'Citations per Faculty', an indicator used in our World University ranking.

Lens	Weighting	Indicator	Weighting
Research and Discovery	50 %	Academic Reputation	30 %
		Citations per Faculty	20 %
Employability and Outcomes	20 %	Employer Reputation	15 %
		Employment Outcomes	5 %
Learning Experience	10 %	Faculty Student Ratio	10 %
Global Engagement	15 %	International Faculty Ratio	5 %
		International Research Network	5 %
		International Student Diversity	0 %
		International Student Ratio	5 %
Sustainability	5 %	Sustainability	5 %

Lens and indicators description

Lens		Indicator	
Research and Discovery	The Research and Discovery lens measures both an institution's research quality, volume, and its reputation in the academic community. Institutions who score highly in this lens are likely to be producing high volumes of research output, have research which is widely cited by other academics, and will have built a reputation in the academic community to reflect this.	Academic Reputation	Academic Reputation The Academic Reputation (AR) indicator measures the reputation of institutions and their programmes by asking academic experts to nominate universities based on their subject area of expertise. The indicator not only illuminates the quality of an institution's research, but also their approach to academic partnerships, their strategic impact, their educational innovativeness and the impact they have made on education and society at large.
		Citations per Faculty	The Citations per Faculty (CPF) indicator is a measure of the relative intensity and volume of research being done at an institution. It highlights the volume of citations being achieved on average by an institution's academic staff. The citation count is divided by the number of individuals in the faculty in order to take into account different sizes of institution.
Employability and Outcomes	The Employability and Outcomes lens measures how well an institution prepares its graduates for employment, as well as the links it has to industry and its reputation outside of academia. Institutions who score highly in this lens are likely to have a strong reputation among employers, have a track record of graduates who go on to success in their field, have good rates of graduate employability, and established links with industry partners.	Employer Reputation	The Employer Reputation (ER) indicator measures the reputation of institutions and their programmes among employers. We remain the only major ranking to focus on this vital aspect of a student's educational journey. To measure this we collect the views of employers from around the world via our Employer Survey.
		Employer Reputation (Alumni Impact, Graduate Employment Index)	The Employment Outcomes (EO) indicator measures to what degree institutions can ensure a high level of employability for their graduates, and their record in producing graduates that have gone on to make a meaningful impact on society. Alumni Impact The Alumni Impact metric aims to measure an institution's record in producing graduates that have gone on to make a meaningful impact on society. Graduate Employment Index The Graduate Employment Index metric is based on graduate employment rate, which is defined as the percentage of graduates who go on to paid (non-voluntary) full or part time work within 15 months of finishing their degree.
Learning Experience	The Learning Experience lens aims to reflect the overall learning environment provided by a higher education institution to its students, through the level of support it offers all its students regardless of socioeconomic background. Institutions who score highly in this lens are likely to have more academic staff resources available to students and to hire high quality research staff.	Faculty Student Ratio	The Faculty-Student Ratio (FSR) indicator is a measure of the number of academic staff that an institution has to teach its students. The indicator is calculated by dividing the number of faculty staff by the number of students. A brief description of the indicator highlighting what it measures and why we use it.
Global Engagement	The Global Engagement lens aims to reflect the overall internationalization of higher education institutions, measuring an institution's international outlook in terms of foreign students, staff and its research links outside its own location. Institutions who score highly in this lens are likely to provide a more culturally diverse experience to students and staff, and be more connected to international networks of academic mobility and research.	International Faculty Ratio	The International Faculty Ratio (IFR) indicator looks at the ratio of international faculty staff to overall staff. Using the overall faculty staff numbers and the number of international faculty staff, we calculate what proportion of the overall faculty is international.
		International Research Network	International Research Network (IRN) is a measure of an institution's success in creating and sustaining research partnerships with institutions in other locations. The indicator measures how diverse and rich an institution's research network is by looking at the number of different countries represented, and whether these relationships are renewed and repeated.
		International Student Diversity	The International Student Diversity (ISD) indicator looks at the ratio of international students to overall students as well as the diversity of nationalities that those students are from. The indicator is an evolution of the International Student Ratio (ISR) indicator, with the aim of measuring not only the size of an institution's international student body, but also how successful the institution is at attracting students from a wide range of different countries and backgrounds.
		International Student Ratio	The International Student Ratio (ISR) indicator looks at the ratio of international students to overall students. Using the overall undergraduate and postgraduate student numbers and the number of international students, we calculate what proportion of the overall student body is international.
Sustainability	Sustainability provides students with a unique lens on which institutions are demonstrating a commitment to a more sustainable existence.	Sustainability	The Sustainability (SUS) indicator highlights which institutions are demonstrating a commitment to a more sustainable existence, and encompasses a variety of factors across environmental, social and governance (ESG). This includes everything from environmental projects on campus, through to diversity initiatives, institutional governance, and the impact of academic research across the UN's 17 sustainable development goals (SDGs). It evaluates the social and environmental impact of universities as not only centres of education and research, but also as major employers. Sustainability is an increasingly important issue for students when picking a study destination and QS is proud to be the first major university ranking provider to include it as an indicator in our core rankings.

QS Institution Classifications

QS classify institutions across four dimensions for the purposes of rankings comparison. They are listed below.

Size Classification

Based on the (full time equivalent) size of the degree-seeking student body. Where an FTE number is not provided or available, one will be estimated based on common characteristics of other institutions in the country or region in question.

	Size	Students
XL	Extra Large	More than 30,000
L	Large	>=12,000
M	Medium	>=5,000
S	Small	Fewer than 5,000

Subject Range Classification

Four categories based on the institution's provision of programs in the five broad faculty areas used in the university rankings. Due to radically different publication habits and patterns in medicine, an additional category is added based on whether the subject institution has a medical school.

Focus	Faculty Area	
FC	Full comprehensive	All 5 faculty areas + medical school
CO	Comprehensive	All 5 faculty areas
FO	Focused	3 or 4 faculty areas
SP	Specialist	2 or fewer faculty areas

Research Intensity & Focus Classification

Four levels of research activity evaluated based on the number of documents retrievable from Scopus in the five-year period preceding the application of the classification. The thresholds required to reach the different levels are different depending on the institution's pre-classification on aspects 1 and 2.

	RESEARCH INTENSITY
VH	Very High
HI	High
MD	Medium
LO	Low

RESEARCH	FOCUS	XL	L	M	S
VH	FC	13000	10000	5000	2500
HI	FC	4000	3000	1500	750
MD	FC	750	500	250	100
LO	FC	0	0	0	0
VH	CO	7000	5000	2500	1250
HI	CO	2000	1500	750	400
MD	CO	400	250	100	50
LO	CO	0	0	0	0
VH	FO	3500	2500	1250	650
HI	FO	1000	750	400	200
MD	FO	150	100	50	50
LO	FO	0	0	0	0
VH	SP	2 x mean for relevant broad subject areas	2 x mean for relevant broad subject areas	2 x mean for relevant broad subject areas	2 x mean for relevant broad subject areas
HI	SP	1 x mean for relevant broad subject areas	1 x mean for relevant broad subject areas	1 x mean for relevant broad subject areas	1 x mean for relevant broad subject areas
MD	SP	0.5 x mean for relevant broad subject areas	0.5 x mean for relevant broad subject areas	0.5 x mean for relevant broad subject areas	0.5 x mean for relevant broad subject areas
LO	SP	0	0	0	0

Faculty & Subject Area Classification

QS evaluates institutions across 5 broad faculty areas and 55 narrow subject areas. These classifications underpin many of our rankings, from the nominations received in our surveys, to the ways in which we use paper and citation data.

Arts and Humanities	Archaeology
	Architecture / Built Environment
	Art & Design
	Classics & Ancient History
	English Language & Literature
	History
	History of Art (New 2023)
	Modern Languages
	Music (New 2024)
	Linguistics
	Performing Arts
	Philosophy
	Theology, Divinity & Religious Studies

Engineering & Technology	Engineering - Chemical
	Engineering - Civil & Structural
	Engineering - General
	Computer Science & Information Systems
	Data Science (New 2023)
	Engineering - Electrical & Electronic
	Engineering - Mechanical, Aeronautical & Manufacturing
	Engineering - Mineral & Mining
	Engineering - Petroleum

Life Sciences & Medicine	Agriculture & Forestry
	Anatomy & Physiology
	Biological Sciences
	Dentistry
	Medicine
	Nursing
	Pharmacy & Pharmacology
	Psychology
	Veterinary Science

Natural Sciences	Chemistry
	Earth & Marine Sciences
	Environmental Sciences
	Geography
	Geology
	Geophysics
	Materials Science
	Mathematics
	Physics & Astronomy

Social Sciences & Management	Accounting and Finance
	Anthropology
	Business & Management Studies
	Communication & Media Studies
	Development Studies
	Economics & Econometrics
	Education
	Hospitality & Leisure Management
	Law
	Library & Information Management
	Marketing (New 2023)
	Politics & International Studies
	Social Policy & Administration
	Social Work
	Sociology
	Sports-related Subjects
	Statistics & Operational Research

How dooes QS calculate Scores and Ranks?

Sample size
QS calculate scores on the subset of all analysed institutions, as opposed to the subset of finally published (ranked) institutions. This means that a distribution of indicator or overall scores may not be from 0 to 100 in the publicly available results table.
Another consequence of that for rankings with increased sample sizes of analysed institutions, score distribution may change (stretch), so that same score may lead to a higher a rank in one edition, but to a lower rank in another one.

Calculation:
QS use the raw ratio or index as the original input.
They apply normalization to all institutions' ratios/indexes to standardize the input, generating a Z-Score for each institution;
They scale the Z Score for all the institutions from 0-100.
Ratio/Index – Input – Z Score – Scaled Score

Z-Score Normalization
The calculation of indicator scores process:
Once the underlying values (ratios or indices) for each indicator are calculated and, if needed, adjusted, the next thing to do is to rescale each indicator's value into a common range from 1 to 100, so that they are compatible with each other. This process is widely known as data normalization.
Firstly, the widely used z-score normalization (or standardization) is applied. A z-value shows the number of standard deviations a given data point lies above (positive z-value), below (negative z-value) or exactly on the mean (zero z-value)*.
Once z-scores are calculated, their position on the normal curve is plotted resulting in the scaled value from 0 to 1 for each indicator, showing the probability that an indicator value of a random institution from the population is less than or equal to the given value. For example, if an institution has Academic Reputation at this step of calculations equal to 0.9, this will indicate that it performs in the top 10% of institutions by this indicator.
The resulting scores are finally linearly scaled between 1 and 100 for each indicator using min-max normalization.

The calculation of overall scores process
After each indicator's data is compatible with each other, QS can combine the data reliably and apply weightings fairly to the calculation of the overall score.
The final overall score is scaled again using min-max normalization, where 100 goes to the maximum overall score, and 1 goes to the minimum overall score. For MBA and Business Masters Rankings, it was scaled from 20-100.
*As the number of institutions in a given ranking grows, this can have an inconsistent effect on how z-score normalization applies to different indicators. Typically, new institutions added will tend to have weaker performances in reputation and research indicators, but may have strengths in faculty student ratio or the international measures. The effect of this has been to bring down the means used for indicators more closely correlated with overall performance at a faster rate than those with indicators less strongly correlated. From 2016, we have locked the mean and standard deviation used for the standardization calculations in the QS World University Rankings to the top X in any given indicator (e.g. X=700 for Citations per Faculty), not including capped values. An impact from this is to space out the institutions above the mean a little more and another is that it is typical for an institution in the same rank position to have a lower score than previously in any given indicator.

Publication of Scores and Ranks:
Every institution has a published rank. It is either unique, joint, or in a band.
All indicator ranks are published and are specific up to a certain point and then banded thereafter. All indicator ranks are based on the underlying ratios or indices behind the corresponding scores, not the scores themselves. This doesn't affect the rankings results, but this does allow institutions more opportunity to showcase their performance, and rankings users to analyze their performance at a more granular level.
All indicator scores are displayed for all ranked institutions.
Overall scores are displayed to a certain point (e.g. Rank 500 in QS World University Ranking) and hidden thereafter.

Scores and Rankings in subregional rankings
An overall rank in the overall rankings is transformed into an overall rank in the corresponding sub-regional rankings, using same overall scores as in the overall rankings. Indicator ranks in the sub-regional rankings are not available.

Data Limitations

Your dataset provides useful insights into global university performance, but it is limited by missing data, bias in survey and research metrics, and comparability issues. Ethically, the rankings risk reinforcing inequalities, lack transparency, and may influence institutional behavior in ways that do not always align with educational values.

Data completeness

- Some columns have many missing values (e.g., not all universities are scored in all dimensions).
- Smaller or newer institutions may lack data in Scopus (research output) or survey responses.

Representation bias

- The dataset includes only ranked universities, so it does not represent all higher education institutions worldwide.
- Countries with fewer resources for international visibility may be underrepresented.

Subjectivity in measures

- Scores such as Academic Reputation (AR) and Employer Reputation (ER) are based on surveys, which may reflect perception rather than objective quality.

Comparability issues

- Different regions have varying higher education systems (size, funding models, language of publication), making direct comparison difficult.
- Research output in Scopus favors English-language publications.

Time-bound nature

Data reflects one snapshot (e.g., rankings for 2026). University performance may change significantly year-to-year.

Ethical Considerations

Equity & bias

- Rankings can reinforce existing hierarchies: well-known universities may get higher scores simply due to reputation, disadvantaging less visible institutions.
- Reliance on English-language journals and international collaborations can disadvantage Global South universities.
- The data set doesn't include the whole list of universities in the world
- QS evaluates institutions only across 5 broad faculty areas and 55 narrow subject areas
- Data can be incomplete or biased due to manual imput and reporting practices.
- Z-score calculation in ranking for new institutions will be low

Transparency

- The weighting and methodology of QS rankings are complex and not always fully transparent, making it difficult for stakeholders to interpret results fairly.

Impact on institutions

- Universities may make strategic decisions (resource allocation, hiring, admissions policies) to improve ranking positions rather than focusing on broader educational goals.
- Pressure to perform in rankings may disadvantage disciplines with less international visibility (e.g., local humanities research).

Use of survey data

- Ethical concerns around informed consent and representation: who is surveyed, how responses are weighted, and whether results reflect global diversity.

Potential misuse

- Policymakers, students, and the public may interpret rankings as absolute measures of "quality," which can be misleading without considering context.

Additional Notes & Links to project files

Rank

Usually, universities outside the top ~500 are not given a precise rank. QS wants to highlight that below a certain point, differences are not statistically significant.
Instead, they are grouped into ranges, for example: “Jeonbuk National University” => [701–710], This means it is somewhere in that range, but QS does not release the exact position.

Symbols meaning:

There are symbols (such as +, -, =) in some rank values. These symbols indicate movement compared to the previous year. = => No change in rank

- + => (ex: 701+) means that the institution is ranked in 701+ band, meaning beyond 700, but the exact rank is not published.

Overall Score

In the overall score, we have, for all the ranks that are > 706, an overall score value “-“ — this means no overall score available, so no calculated rank. Also, the Rank and previous Rank for all these values are not precise and are a range of ranks

Links to project files

Project Folder (FULL) Link

- **Data set (Excel) file Link**
- **Data profiling (Excel file) Link**
- **Cleaned Data Set (Excle file) Link**
- **Python Script (Jupiter file) Link**

Project Overview

Motivation

It is easy to assume that the best universities are simply those with the highest rankings at the very top of the QS list.
However, not everyone can realistically enroll in the top 10 due to high competition and cost. This project is motivated by the idea that high-quality education can also be found beyond the top tier, since there are many excellent universities worldwide that could be strong alternatives.

Objective

According to the QS methodology, the Discovery and Research lens carries the greatest weight (50%), with Academic Reputation (AR) as its key component (30%).
The goal of this analysis is to explore which other scores contribute to success in AR, helping us better understand what drives rankings outside of the top tier.

Scope

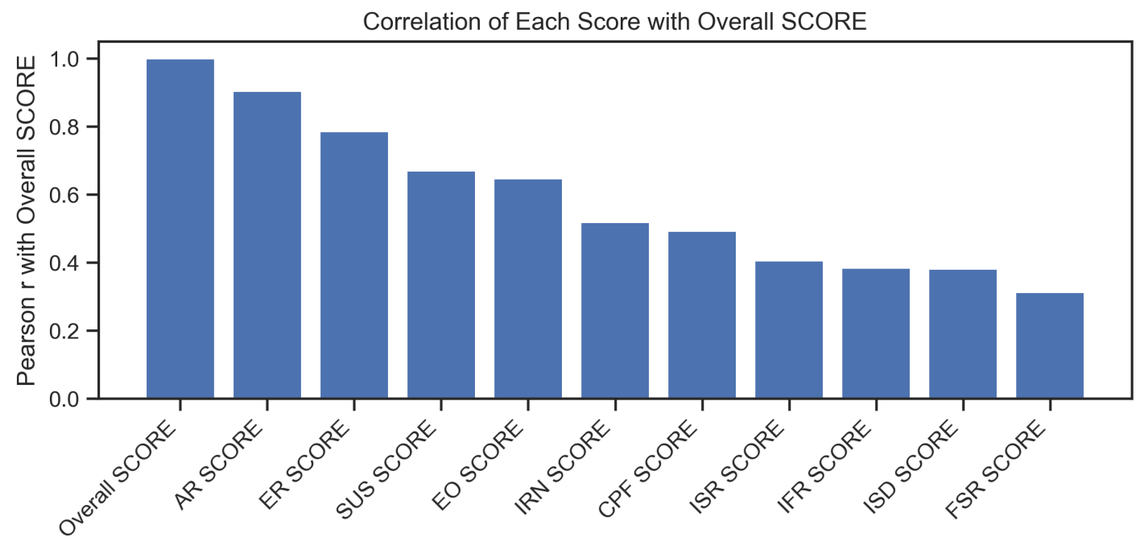
This analysis explores:
Global distribution of top universities (map) and descriptive statistics
Differences across quartiles of universities
Correlation patterns between indicators in each quartile
Which scores most strongly drive Academic Reputation (AR) in the middle quartiles
A few suggestions for undergraduates based on our findings

QS World University Rankings 2026: Questions to explore

Ranking & Score Analysis

Which factors correlate most strongly with the Overall Score?

The factor which correlates the most is Academic Reputation Score and second is Employer Reputation.



Do universities with higher Overall Scores also rank high across all sub-scores, or are there trade-offs?

Universities with higher Overall Scores also rank high across all sub-scores.

On radarplot we took 4 randon universities from 4 different quartiles - 25,50,75,95. Monash (almost uniformly high), but there are trade-offs: e.g., Khalifa (international > reputation), Technion (reputation > internationalization), Boston College (graduate outcomes > research).

Monash University (blue):

- Very strong overall — close to the outer ring in most metrics (AR, ER, FSR, ISD, IRN, SUS).
- Especially impressive in Internationalization scores (IFR, ISR) and Research (CPF).
- Looks like an “all-rounder” with only small dips.

Khalifa University of Science and Technology (orange):

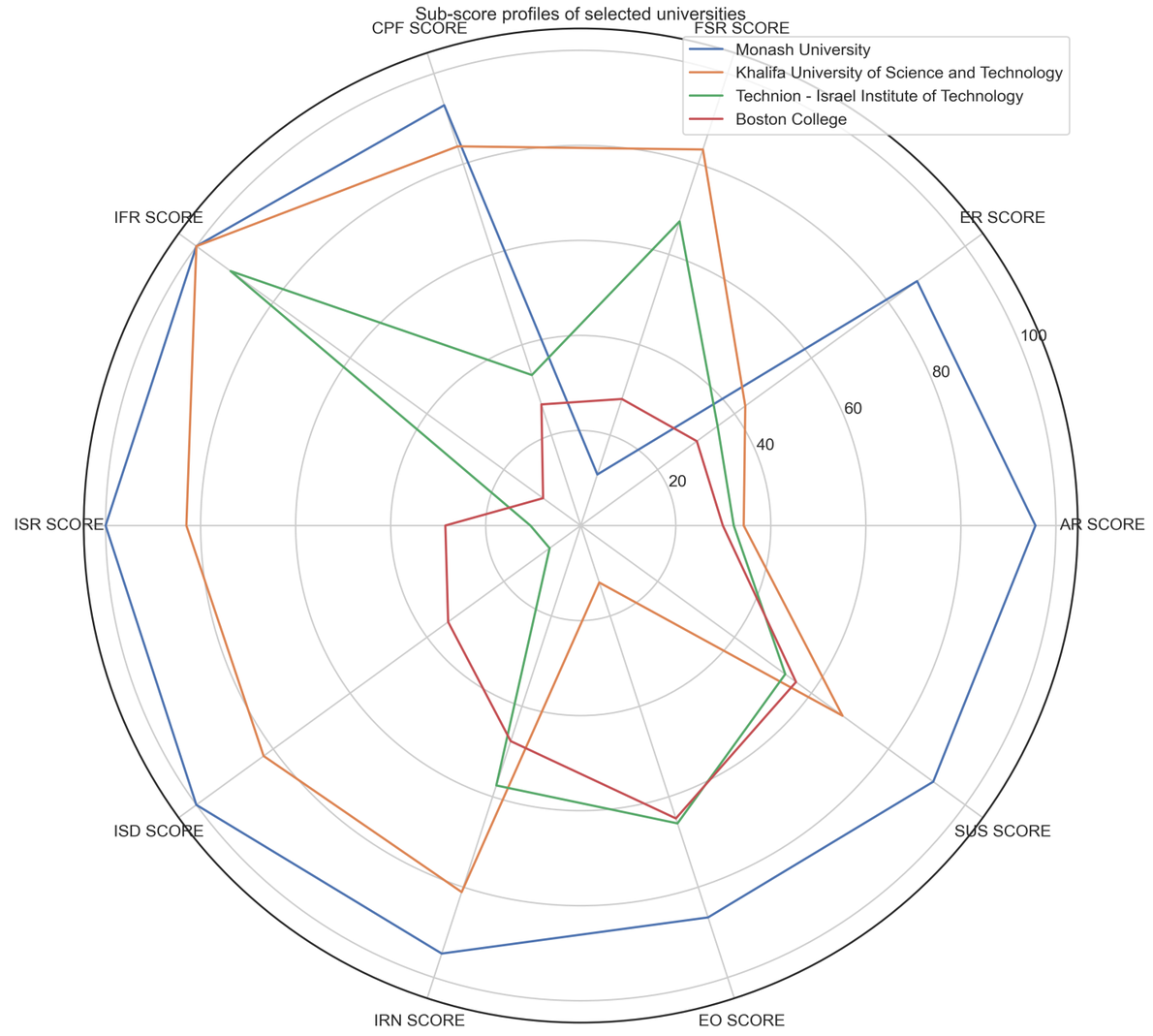
- Also strong in IFR, ISR, ISD (international profile).
- Lower in AR and ER, meaning reputation is weaker compared to Monash.
- Balanced but not top-tier in Sustainability.

Technion – Israel Institute of Technology (green):

- Strength in ER (Employer Reputation) and some research dimensions.
- Weaker in IFR/ISR (international ratios).
- The shape is spikier — some strong points, some weak points → a trade-off profile.

Boston College (red):

- Generally lower across most dimensions.
- Strongest in EO (Employment Outcomes) and moderate in Sustainability.
- Much weaker in research-related (CPF, ISD, IRN) and internationalization scores.



What is the distribution of Overall Scores across universities (by region)?

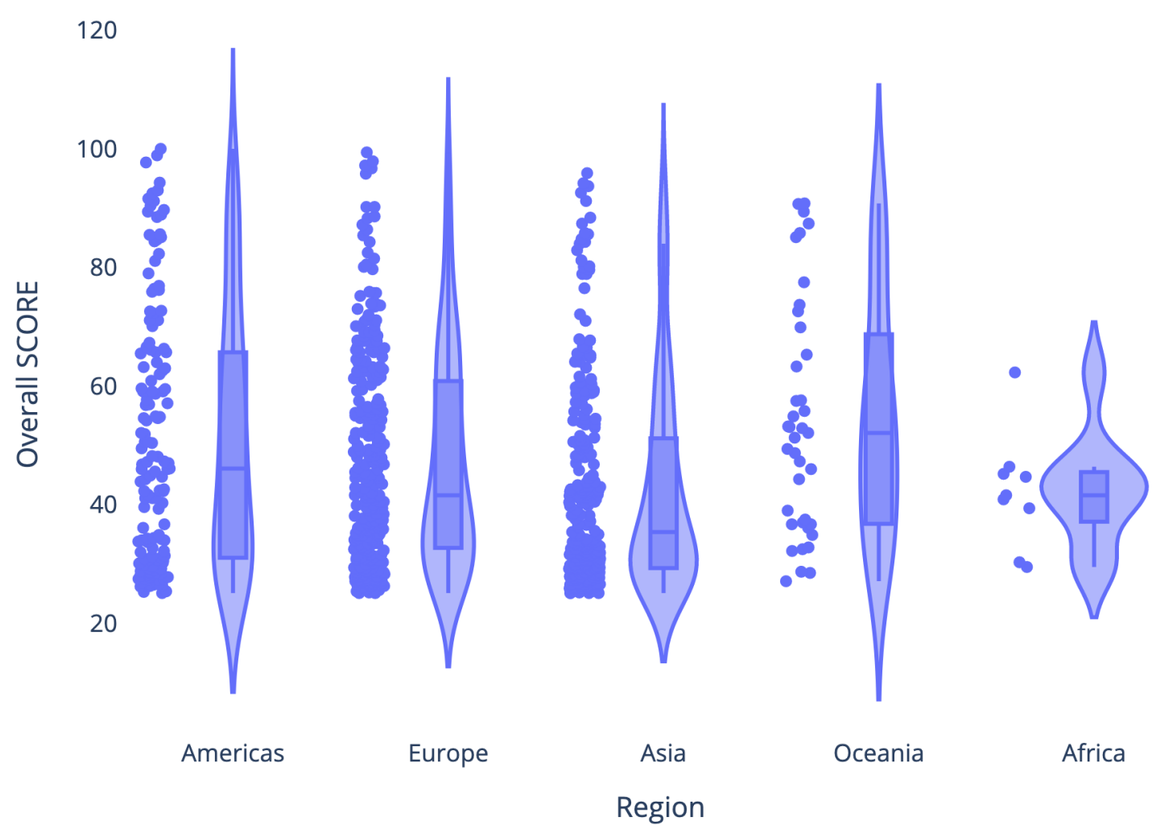
AR = Academic Reputation
ER = Employer Reputation
FSR = Faculty–Student Ratio
CPF = Citations per Faculty
IFR = International Faculty Ratio
ISR = International Student Ratio
ISD = International Research Network
IRN = International Research Network (check: IRN and ISD are often related, but different in QS)
EO = Employment Outcomes
SUS = Sustainability

Most universities cluster in the 30–60 score range across all regions. The violins are widest in this middle zone, meaning that’s where the majority of universities lie.

High performers are rare: Only a small fraction of universities achieve scores above 80. These appear as thinner “tails” at the top of the violins (a few top performers stretching the distribution upward).

The violin plot confirms that global rankings are driven by a relatively small group of elite universities, while the majority occupy mid-range positions.

- Europe & Asia: Very wide violins in the middle ranges, suggesting a large number of mid-ranked institutions, plus a few elite ones above 90.
- Americas: Similar clustering, but with slightly higher median and more spread toward the top.
- Oceania: Fewer institutions overall, but also concentrated around the mid-scores (40–60) with a handful high-scoring universities.
- Africa: Very small sample, mostly clustered around 40–50, with no very high-scoring outliers.



Average Overall by Region:

Oceania	54.238462
Americas	50.913793
Europe	47.099634
Asia	42.637238
Africa	42.255556

QS World University Rankings 2026: Questions to explore

Regional & Country Comparisons

Which regions (Asia, Europe, Americas, etc.) dominate the top 100 universities?
Top-100 count by Region:

Europe	37
Americas	31
Asia	22
Oceania	10

Top 15 Countries by Avg Overall SCORE:

Hong Kong SAR, China	72.600000
Singapore	63.125000
Netherlands	61.669231
Switzerland	61.200000
Mexico	60.550000
Sweden	59.962500
Qatar	58.700000
Denmark	58.040000
Australia	55.077419
United Kingdom	54.060000
Japan	53.362500
United States of America	52.871429
Ireland	51.966667
Belgium	51.450000
New Zealand	50.987500

Top 15 Countries by Median Overall SCORE:

Hong Kong SAR, China	81.20
Denmark	64.30
Singapore	62.65
Netherlands	62.50
Sweden	61.35
Switzerland	61.30
Mexico	60.55
Qatar	58.70
Japan	57.70
Australia	53.10
New Zealand	50.35
Belgium	50.25
United Kingdom	49.75
Colombia	49.50
Ireland	48.90

Distribution of Ranked Universities by Country (Top 20)

United States of America	192
United Kingdom	90
China (Mainland)	72
India	54
Germany	49
Japan	47
Italy	43
Republic of Korea	43
Russian Federation	40
Spain	38
Australia	36
France	35
Malaysia	32
Canada	29
Taiwan	28
Indonesia	26
Türkiye	26
Brazil	24
Mexico	22
Saudi Arabia	22

QS World University Rankings 2026: Questions to explore

Regional & Country Comparisons

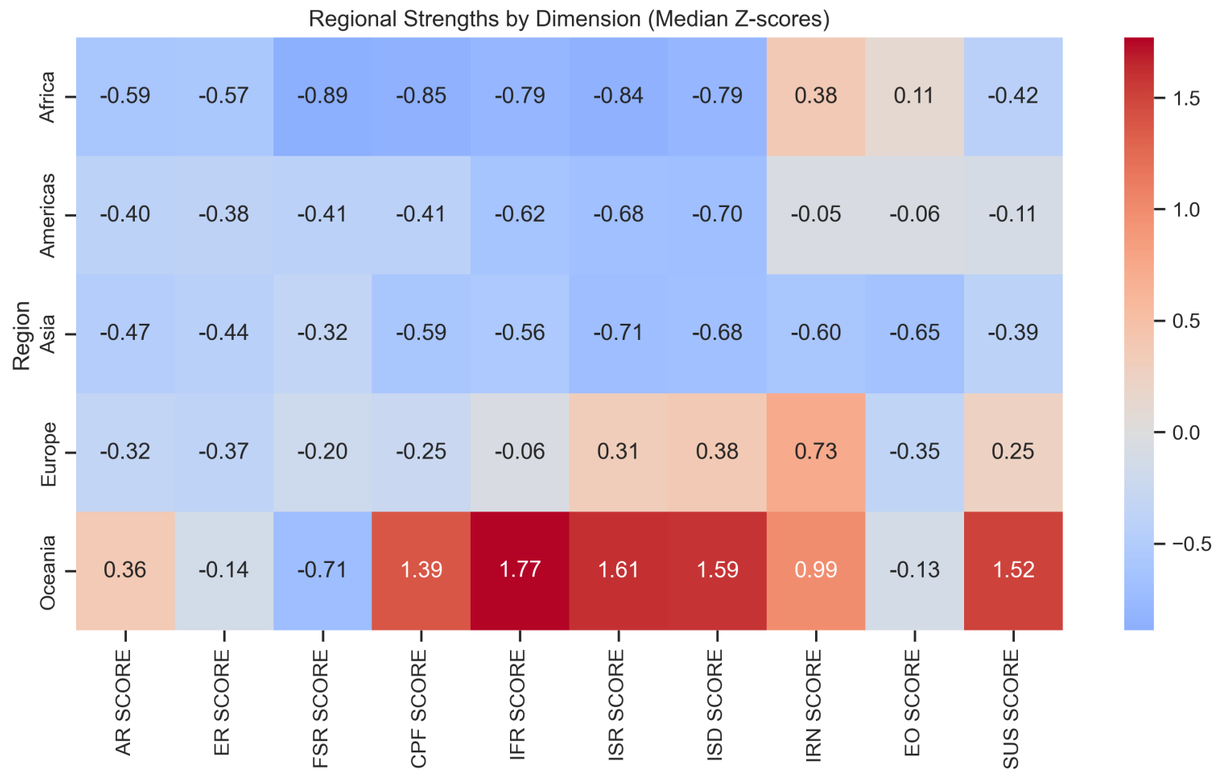
Are certain regions stronger in specific dimensions (e.g., research intensity in North America vs. teaching in Europe)?

Median Z-scores
Robust to outliers — reflects the "typical" institution in the region.

Oceania is still strong (CPF +1.39, IFR +1.77, ISD +1.59, SUS +1.52), but the values are even higher — meaning that not just a few, but the majority of Oceania's universities are well above average in these areas.

Europe shows modest strengths (IRN +0.73), but many metrics drop closer to 0 or negative compared to the mean. This suggests Europe has a few very strong outliers pulling up the mean.

Africa, Americas, Asia → median Z-scores are more negative than mean Z-scores. This means these regions have a few top performers improving the average, but the typical university is weaker.

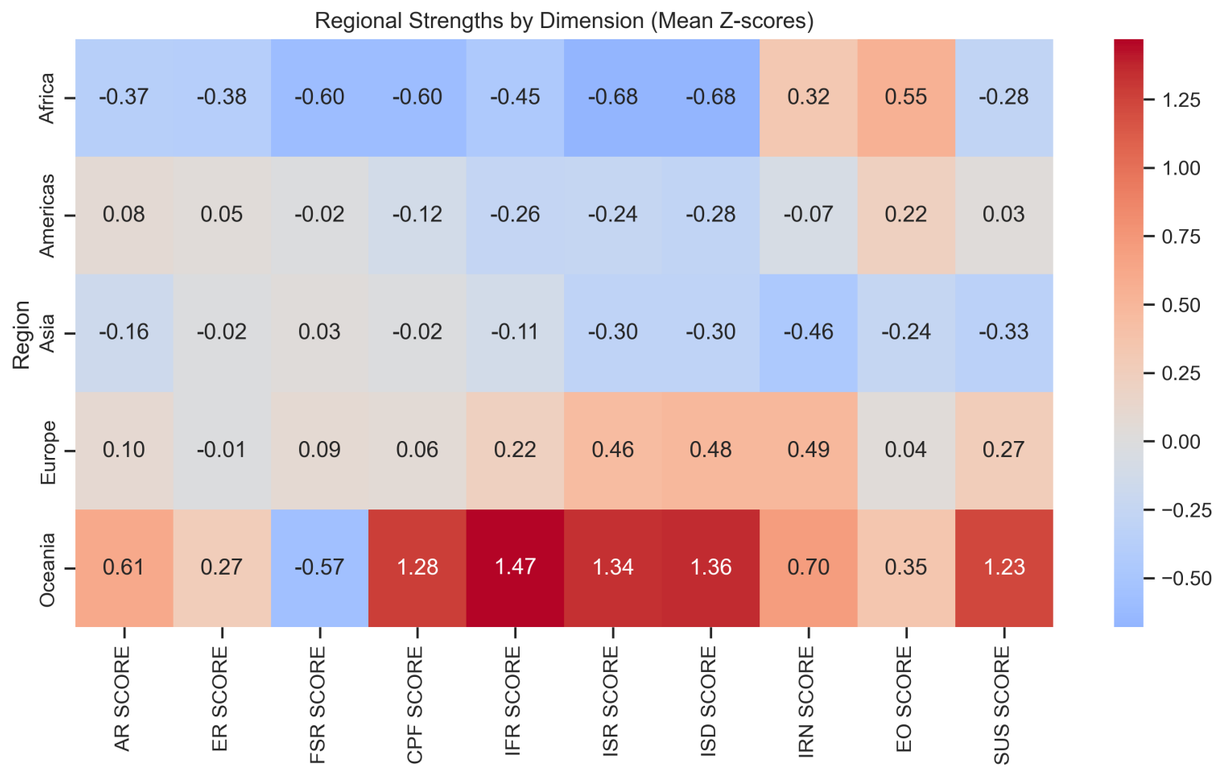


Mean Z-scores
Sensitive to outliers (a few very high or low universities can pull the mean up or down).

Oceania looks very strong across almost all internationalization metrics.

Europe shows balanced small positives (e.g. ISD +0.48, IRN +0.49).

Americas hover around 0, slightly positive in EO (+0.22).

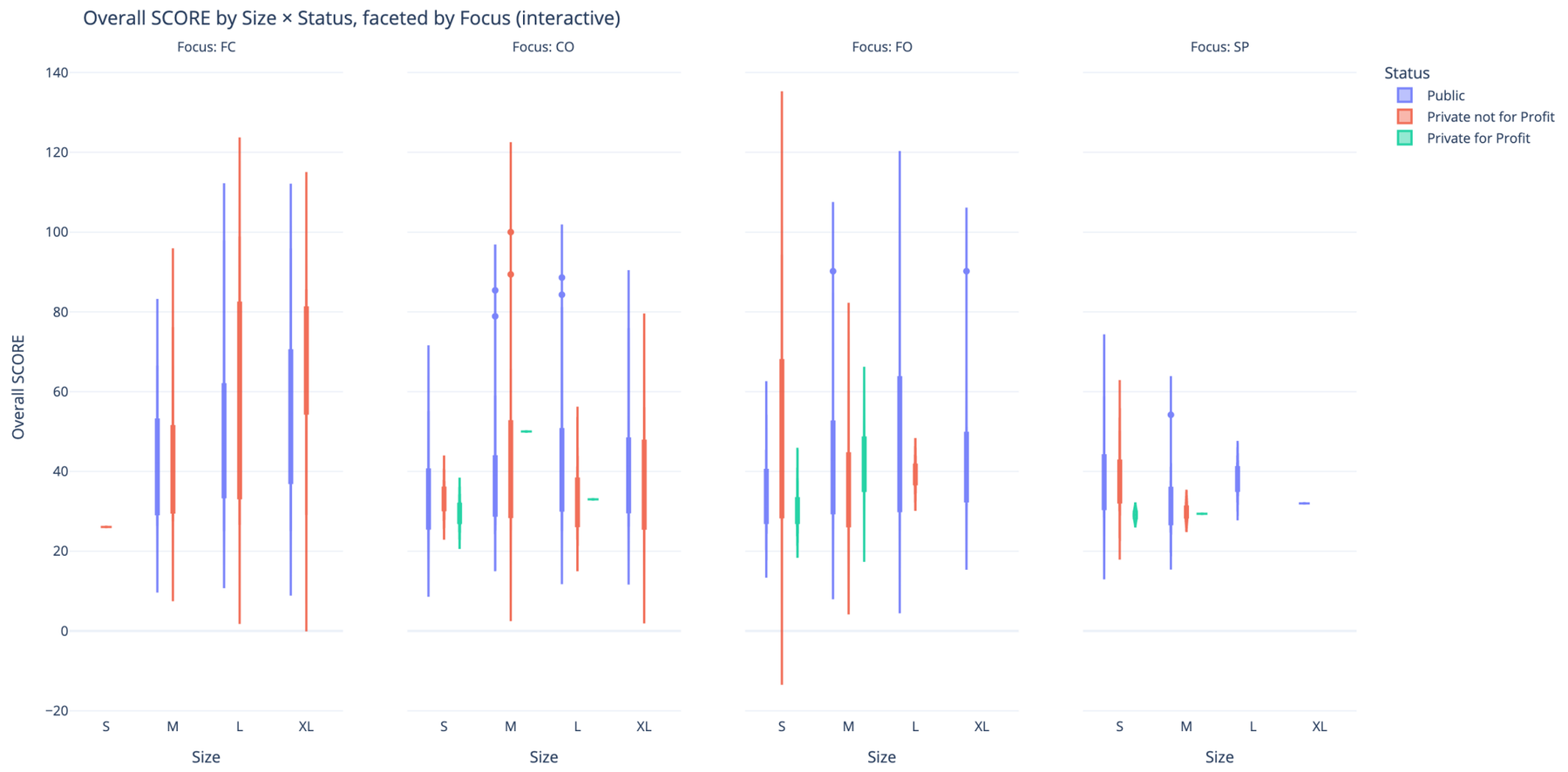


Means highlight elite outliers (e.g., top U.S. or European universities boost averages).
Medians show the typical university performance — and here it's clear that Oceania's strength is not just outliers, but consistent across the board.
For Africa, Asia, and Americas, the gap between mean and median suggests that strong outliers (top-ranked universities) mask weaker performance of the majority.

QS World University Rankings 2026: Questions to explore

Institutional Characteristics

Does university Size (S, M, L, XL) correlate with better rankings? Do Public vs. Private institutions show differences in Overall Score or sub-scores? Does the type of Focus (Comprehensive vs. Focused) impact performance?



1. Size of University

Trend: As university size increases (S → XL), performance improves.

Mean Rank improves dramatically:

Small (805) → Medium (874) → Large (725) → XL (631).

Lower rank = better performance.

Mean Overall Score rises steadily:

Small (35.7) → Medium (40.7) → Large (47.7) → XL (51.3).

Insight: Larger universities (L, XL) tend to achieve better rankings and higher overall scores than smaller ones. Possibly due to more resources, broader research, and stronger global presence.

Size	Count	Mean_Rank	Median_Rank	Mean_Overall	Median_Overall
S	97	805.07	825.5	35.71	31.3
M	382	874.11	925.5	40.69	34.9
L	677	724.93	735.5	47.74	42.9
XL	347	630.51	542.0	51.33	48.0

2. Status (Public vs. Private)

Private for Profit:

Very weak performance: Mean Rank ~1034, Mean Overall ~33.4.

These institutions perform significantly worse.

Private not for Profit:

Much better: Mean Rank ~813, Mean Overall ~47.3.

Comparable or slightly higher than public universities in scores.

Public Universities:

Strong performance: Mean Rank ~718, Mean Overall ~46.9.

They dominate the dataset in number (1185 vs. 263 private non-profit).

Insight: For-profit universities are consistently weaker.

Public and non-profit private universities perform comparably, with public universities having stronger rank positions.

Status	Count	Mean_Rank	Median_Rank	Mean_Overall	Median_Overall
Private for Profit	55	1033.94	1100.5	33.38	29.9
Private not for Profit	263	813.29	875.5	47.34	39.3
Public	1185	717.98	705.5	46.89	41.8

3. Focus Type

FC (Focused) universities do the best:

Mean Rank ~585, Median Rank ~518.

Mean Overall ~52, Median Overall ~49.

CO (Comprehensive) and FO (Focused Other?) are mid-range:

Ranks ~835–852, Overall scores ~41–43.

SP (Specialized) universities perform the weakest:

Rank ~913, Score ~34.7.

Insight: Focused universities (FC) perform best overall, with strong scores and ranks. Specialized (SP) universities are underperforming.

Focus	Count	Mean_Rank	Median_Rank	Mean_Overall	Median_Overall
CO	494	835.43	875.5	40.60	35.10
FC	584	584.71	518.0	51.98	48.95
FO	341	852.42	875.5	43.15	36.50
SP	84	913.30	900.5	34.71	31.95

Focus	Faculty Area	
FC	Full comprehensive	All 5 faculty areas + medical school
CO	Comprehensive	All 5 faculty areas
FO	Focused	3 or 4 faculty areas
SP	Specialist	2 or fewer faculty areas

Multiple Regression Model

This step is about understanding the explanatory power of the predictors. Once I know which predictors matter and how much they matter, I can move forward with confidence to the predictive stage — testing whether these predictors can accurately forecast Overall Scores for universities not used in model training.

In this step, I build a baseline multiple regression model to understand how well three key predictors — Academic Reputation (AR), Employer Reputation (ER), and Sustainability Score (SS) — explain variation in the Overall Score of universities.

Why OLS regression?

Ordinary Least Squares (OLS) is the standard method to test how multiple predictors jointly influence an outcome.

It provides coefficients that show the marginal effect of each predictor while holding the others constant.

By including all three predictors at once, I can see whether sustainability (SS) still matters after controlling for reputation (AR, ER).

Why robust standard errors (HC3)?

University scores are diverse and may not have equal variance across all observations (heteroskedasticity).

Using robust errors makes the p-values and confidence intervals more reliable, ensuring the results are not biased by unequal spread.

What does this achieve?

It gives me a clear picture of which factors are statistically significant and how large their contribution is.

The R^2 value (~0.85) shows how much of the overall ranking variation is already explained by these three scores.

The coefficients tell me the direction and relative strength of each factor's influence.

AR is the strongest driver.

SS has a surprisingly strong effect, even more than ER.

Why is this useful before prediction?

Before making regression-based predictions, it's important to check whether the chosen predictors are meaningful and stable.

This OLS step validates that AR, ER, and SS are not only correlated with Overall Score, but are also independent contributors.

With this evidence, I can justify using them as predictors when I move to the next stage: training regression models for prediction (on test data, or using cross-validation).

OLS Regression Results

Dep. Variable:	overall	R-squared:	0.846
Model:	OLS	Adj. R-squared:	0.845
Method:	Least Squares	F-statistic:	1162.
Date:	Mon, 08 Sep 2025	Prob (F-statistic):	4.36e-271
Time:	12:59:48	Log-Likelihood:	-2403.8
No. Observations:	703	AIC:	4816.
Df Residuals:	699	BIC:	4834.
Df Model:	3		
Covariance Type:	HC3		

	coef	std err	z	P>|z|	[0.025	0.975]
const	9.9479	1.132	8.785	0.000	7.728	12.167
ar	0.4454	0.022	20.011	0.000	0.402	0.489
er	0.1261	0.018	6.847	0.000	0.090	0.162
ss	0.1904	0.020	9.733	0.000	0.152	0.229

Omnibus:	0.613	Durbin-Watson:	1.628
Prob(Omnibus):	0.736	Jarque-Bera (JB):	0.708
Skew:	0.035	Prob(JB):	0.702
Kurtosis:	2.861	Cond. No.	398.

Notes:
[1] Standard Errors are heteroscedasticity robust (HC3)

Here, we test whether AR, ER, and SS contribute to explaining Overall Score.

All three coefficients are significant ($p < 0.001$), meaning each predictor has an independent effect.

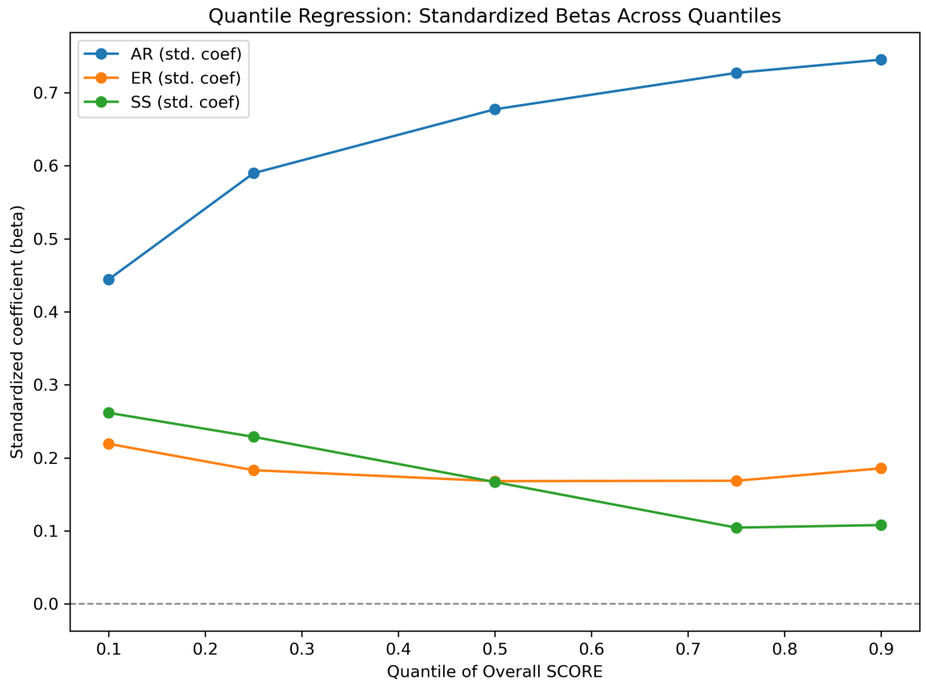
AR has the largest effect (coef ≈ 0.445) → it is the strongest driver.

SS (coef ≈ 0.19) has more influence than ER (≈ 0.13), which is an interesting finding: sustainability contributes meaningfully to overall performance.

The R^2 = 0.846 shows that ~85% of variation in Overall Score is explained by these three predictors, which indicates an excellent model fit.

Quantile Regression with AR, ER, and SS

In the previous OLS regression, I saw that Academic Reputation (AR), Employer Reputation (ER), and Sustainability (SS) all had effects on the Overall Score. However, OLS only explains the average effect across all universities. This hides an important question: do these predictors matter equally for low-ranked and high-ranked universities?
To answer this, I applied quantile regression. Unlike OLS, quantile regression estimates the relationship between predictors and the outcome at different points of the distribution (e.g., the 10th, 25th, 50th, 75th, and 90th percentiles of Overall Score). This allows me to see whether Sustainability, Academic, and Employer Reputation play the same role for weaker universities as they do for elite institutions. By including all three predictors in the model, I can directly compare their relative strength across the spectrum. This approach reveals differentiators for mid- and low-ranked universities versus dominant drivers at the very top.



Quantile regression (original scale):

	quantile	Intercept	ar_coef	ar_p	er_coef	er_p	ss_coef	ss_p
0	0.10	-0.883997	0.309280	2.486191e-18	0.146781	9.790282e-07	0.280427	5.276378e-29
1	0.25	2.849170	0.410537	7.760408e-34	0.122603	1.178625e-05	0.245186	2.270069e-18
2	0.50	10.140886	0.471497	4.288508e-60	0.112559	3.157356e-07	0.178828	3.593049e-11
3	0.75	18.354671	0.506203	3.127586e-69	0.112882	3.081595e-08	0.111835	1.897013e-04
4	0.90	21.295943	0.518844	1.954159e-75	0.124313	2.314062e-10	0.115681	7.356104e-05

	quantile	ar_coef	er_coef	ar_p	er_p	stronger_predictor
0	0.10	0.612411	0.144764	3.762519e-23	2.019574e-02	AR stronger
1	0.25	0.746637	0.130331	7.731867e-51	5.962677e-03	AR stronger
2	0.50	0.814503	0.121591	1.134130e-93	3.540206e-04	AR stronger
3	0.75	0.843431	0.136510	1.038669e-124	2.350194e-06	AR stronger
4	0.90	0.844650	0.153956	2.038563e-113	8.642825e-07	AR stronger

1) Quantile 0.10 (lower-performing universities)
AR = 0.31 ($p < 0.001$, highly significant) → a 1-point increase in AR raises Overall by ~0.31.
ER = 0.15 ($p < 0.001$) → weaker but still relevant.
SS = 0.28 ($p < 0.001$) → almost as strong as AR.
Among low-scoring universities, Sustainability (SS) matters almost as much as Academic Reputation.

2) Quantile 0.25 (lower quartile)
AR = 0.41 → influence becomes stronger.
ER = 0.12 → stays weaker.
SS = 0.25 → still relevant but below AR.
In the lower quartile, AR takes the lead, though SS remains an important factor.

3) Quantile 0.50 (median)
AR = 0.47 → very strong effect.
ER = 0.11 → weak but significant.
SS = 0.18 → moderate effect.
Around the median, AR clearly dominates, while SS plays a secondary role.

4) Quantile 0.75 (upper quartile)
AR = 0.51 → effect peaks.
ER = 0.11 → still weak but consistent.
SS = 0.11 → declines sharply.
For higher-scoring universities, AR is decisive, SS loses its impact.

5) Quantile 0.90 (top 10%)
AR = 0.52 → strongest influence overall.
ER = 0.12 → steady but minor.
SS = 0.12 → small and less relevant.
At the very top, Overall Score is almost entirely driven by AR, while SS and ER play only minor roles.

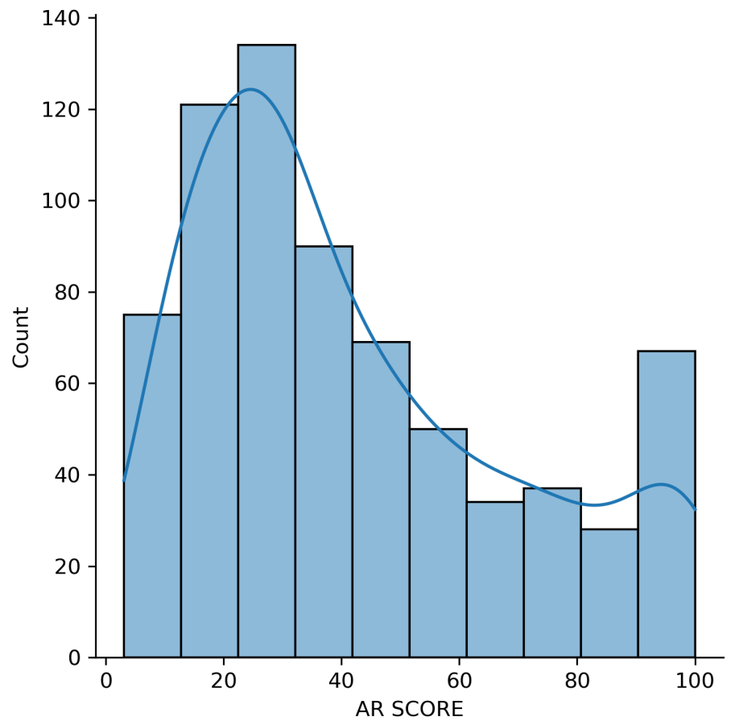
Overall:
AR (Academic Reputation) grows in strength across the distribution and dominates at the top.
ER (Employer Reputation) remains consistently weak but significant.
SS (Sustainability) is a differentiator for lower- and mid-tier universities, but its role diminishes in elite institutions.

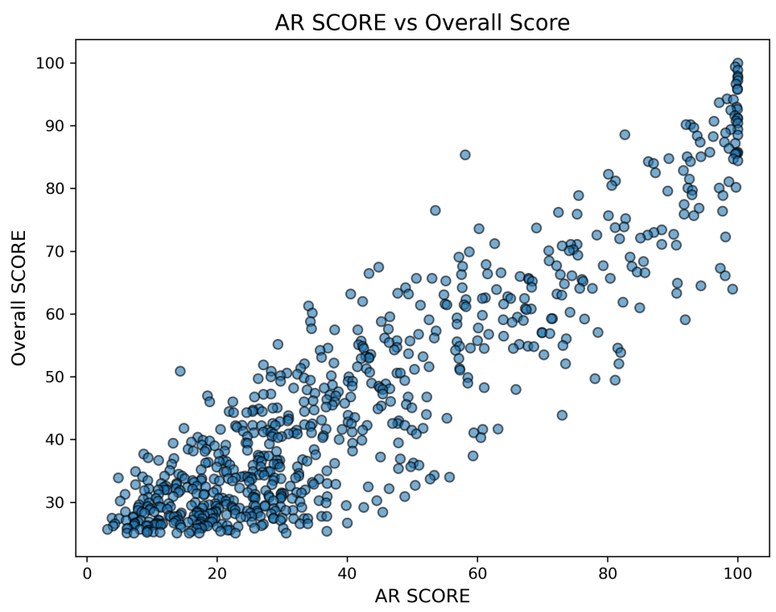
Interpretation of Quantile Regression with AR, ER, SS (standardized betas)
Quantile 0.10 (bottom 10%)
AR = 0.44 → strongest predictor.
ER = 0.22, SS = 0.26 → moderate influence.
For weaker universities, AR already dominates, but SS still has some differentiating role.
Quantile 0.25 (lower quartile)
AR grows to 0.59 → influence strengthens.
ER = 0.18, SS = 0.23 → both weaker than AR.
AR takes a clear lead, SS remains relevant, ER less so.
Quantile 0.50 (median)
AR = 0.68 → very strong effect.
ER = 0.17, SS = 0.17 → both much smaller.
Around the middle, AR fully dominates.
Quantile 0.75 (upper quartile)
AR = 0.73 → peaks.
ER = 0.17, SS = 0.10 → both declining.
At higher scores, AR is decisive, SS weakens.
Quantile 0.90 (top 10%)
AR = 0.75 → strongest across all quantiles.
ER = 0.19, SS = 0.11 → minor roles.
Among elite universities, Overall Score is almost entirely explained by AR.

Overall
AR (Academic Reputation) strengthens consistently across quantiles and is the decisive driver of Overall Score, especially for top universities.
ER (Employer Reputation) remains stable but weak across all quantiles.
SS (Sustainability) has moderate importance for lower-tier universities but loses influence at the top.
Winner: AR is the strongest predictor at every quantile.

Regression Analysis

Overview of how Academic Reputation (AR) is distributed among the 705 universities with an Overall Score





Model Validation (30% Test Set)
I evaluated the regression model predicting Overall Score from AR Score using a 70/30 train–test split.
Validation metrics:

R^2 = 0.804

→ About 80% of the variation in Overall Score is explained by AR Score alone.

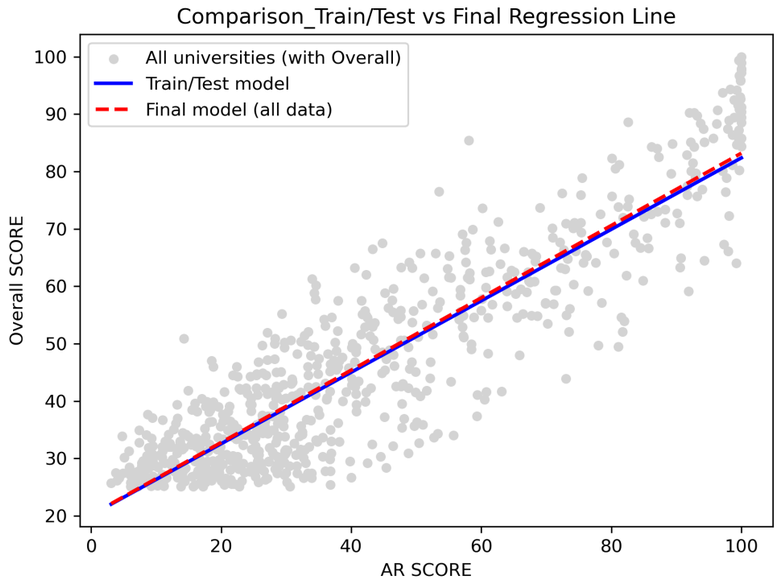
MAE = 6.57

→ On average, predictions are off by ~6.6 points from the true Overall Score.

RMSE = 8.30

→ Larger errors (outliers) are typically within ~8.3 points.

The model shows strong explanatory power (high R^2).
Prediction errors are moderate (MAE ~6.6 on a 0–100 scale).
RMSE > MAE indicates that a few universities have bigger prediction gaps, but overall the model performs consistently.




```
Train/Test model coefficients:
Intercept: [20.1155207]
Slope (AR → Overall): [[0.62252737]]

Final model coefficients:
Intercept: 20.18327470102191
Slope (AR → Overall): [0.63001484]
```

The slope (how much Overall Score rises with +1 AR point) changed only from 0.623 → 0.630 (Δ ≈ 0.007).

The intercept shifted only from 20.12 → 20.18 (Δ ≈ 0.06).

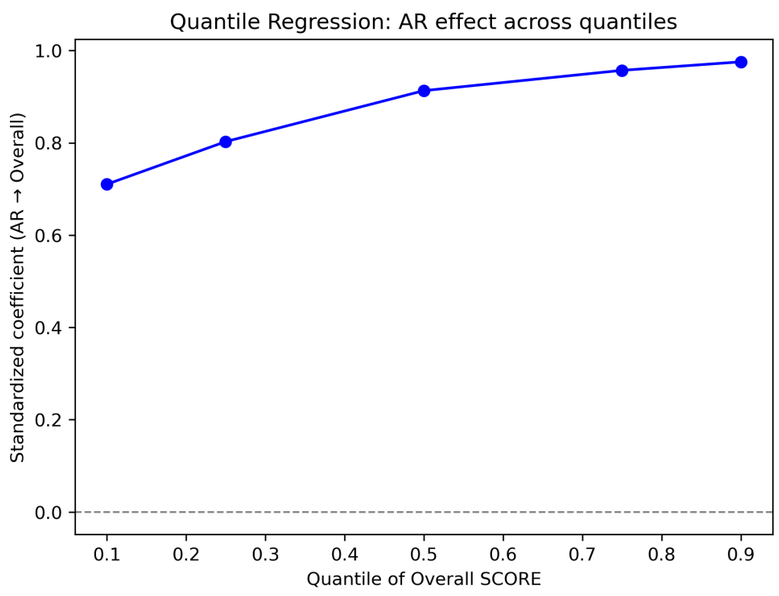
These differences are negligible → the relationship between AR and Overall is very stable, regardless of how the data is split.

"Regression is stable: AR consistently predicts Overall Score, with nearly identical coefficients in both validation and final models."

Does AR matter equally for low-ranked and high-ranked universities, or does its importance change across the distribution?

To answer this, I applied Quantile Regression. Unlike OLS, quantile regression estimates the effect of AR at different parts of the distribution (10th, 25th, 50th, 75th, and 90th percentiles). This allows me to capture how AR influences weaker, median, and elite universities differently.

	quantile	Intercept	ar_coef	ar_p	test_pinball_loss
0	0.10	-0.575931	0.710158	6.713336e-150	0.076608
1	0.25	-0.288988	0.802615	3.322297e-128	0.137989
2	0.50	0.023848	0.913063	2.819374e-145	0.167013
3	0.75	0.302807	0.957078	5.732185e-143	0.127478
4	0.90	0.513335	0.975516	1.020916e-143	0.070375



At the 10th percentile (low-ranked universities), AR has a moderate but significant effect (0.71 SD).
At the median, the effect strengthens (0.91 SD).
At the 90th percentile (top universities), AR nearly explains Overall one-to-one (0.98 SD).
AR is statistically significant at all quantiles (p < 1e-100).
Pinball loss shows the model performs reliably across quantiles, with especially good fit at the tails.

AR matters everywhere, but its impact grows as universities move up the ranking scale.
For elite institutions, Overall Score is almost entirely driven by AR.
This confirms that Academic Reputation is the decisive driver of success at the top of the distribution, while still being important for lower-ranked universities.

Clustering

Two clusters identified:

cluster_0 = universities with higher scores across almost all metrics especially international presence and reputation. Strong/top universities with high academic and employer reputation, strong international presence, and good sustainability.

cluster_1 = medium to lower-performing universities across most metrics. Medium/lower-performing universities with lower scores and higher variability.

Standard Deviation (std)
cluster_0 generally has lower std, meaning universities are more homogeneous in their performance.
cluster_1 has higher variation, meaning some universities are closer to top, while others are much lower.

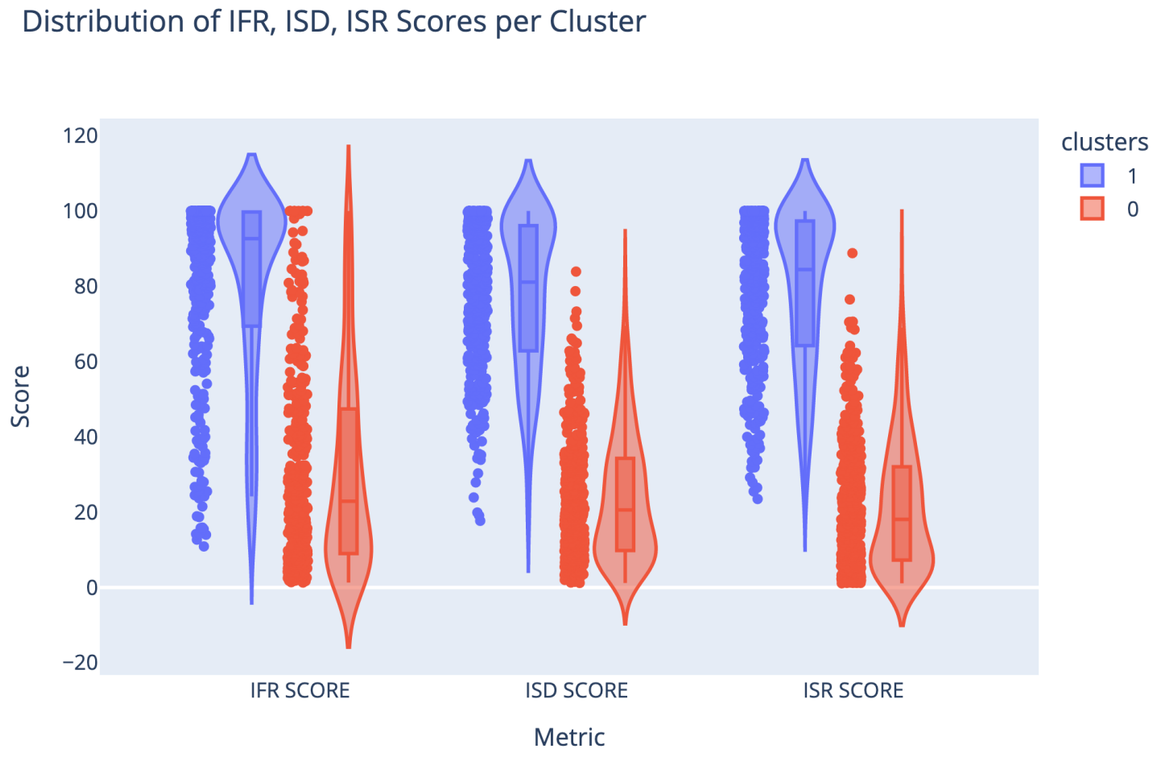
Min and Max values
Cluster 0's minimum scores are higher across most metrics, indicating fewer low-performing universities.
Maximum scores are similar in both clusters (close to 100), so top universities exist in both, but cluster 1 has many lower performers.

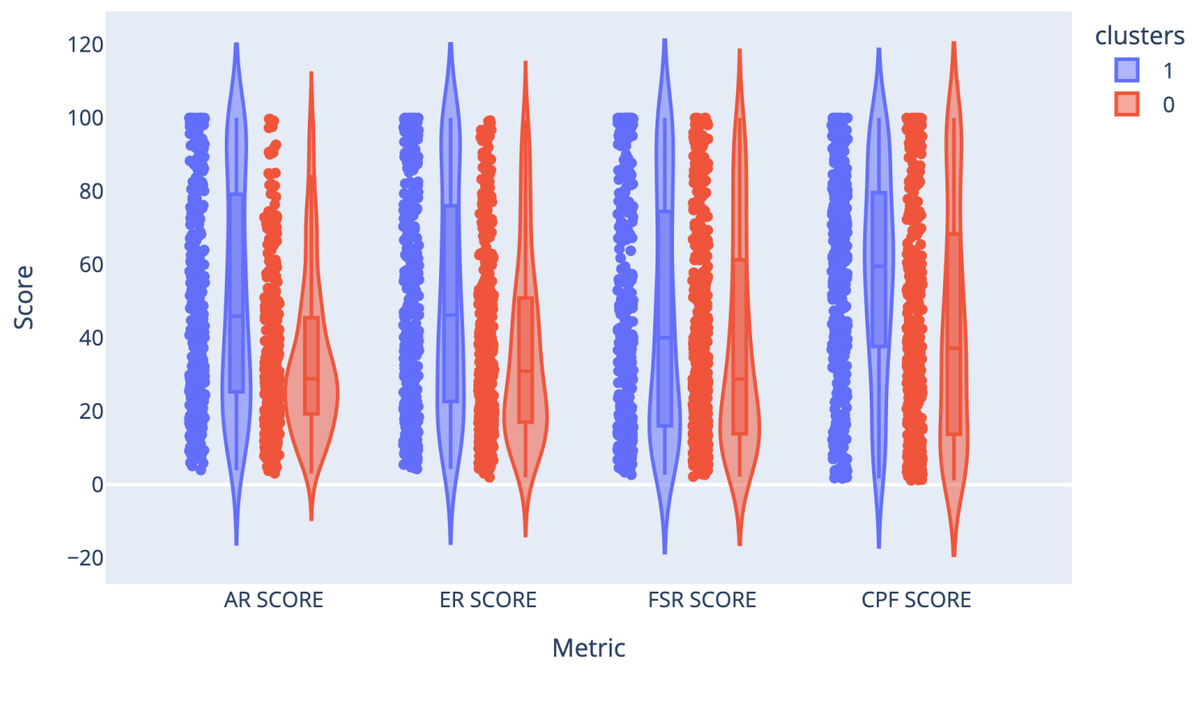
Internationalization drives cluster separation, despite its lower weight in overall rankings

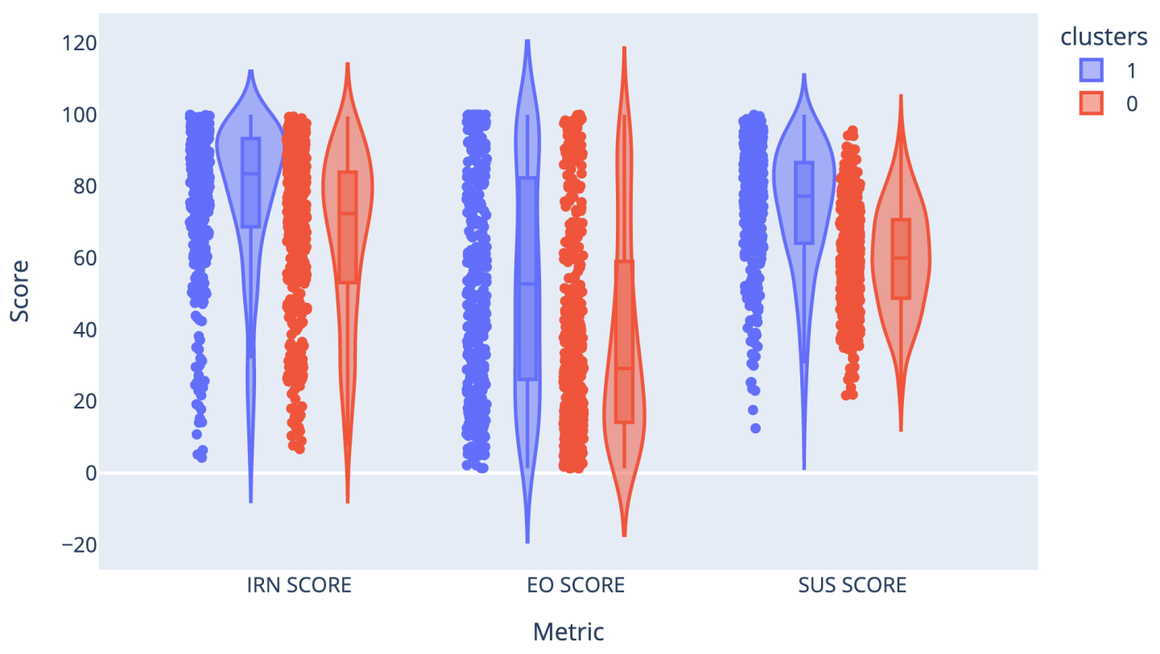
Biggest differences between clusters are in:
- IFR (International Faculty Ratio)
- ISD (International Student Diversity)
- ISR (International Student Ratio)

All three metrics are part of the Global Engagement lens, which has only 15% weight in the overall ranking.
Other lenses (Research & Discovery, Teaching, Sustainability) show smaller differences between clusters, meaning the clusters are not extremely different in the highest-weighted lenses.

Even though Global Engagement has lower weight, it still creates visible separation in the data.
This means that overall rankings might not fully reflect the large differences in internationalization, because this lens contributes less to the total score.
"The most visible cluster differences are in international engagement, even though it counts only 15% of the total score."
"Universities in Cluster 0 are far more globally connected (faculty and students), but in other high-weighted metrics like research, the differences are smaller."
"This suggests that clustering can reveal hidden patterns not immediately obvious from overall ranking scores."







4 Quartiles

To better understand differences between universities, first we splited them into four quartiles:

Q1 (Top 25%) – the highest-ranked universities.
Q2 (25–50%) – upper-middle universities.
Q3 (50–75%) – lower-middle universities.
Q4 (Bottom 25%) – the lowest-ranked universities.

Then, we created correlation heatmaps to see how different QS scores (Academic Reputation, Employer Reputation, Sustainability, etc.) relate to each other within each quartile to see how strongly one score tends to rise or fall with another score – and whether this changes depending on whether a university is top-ranked, middle, or lower-ranked.

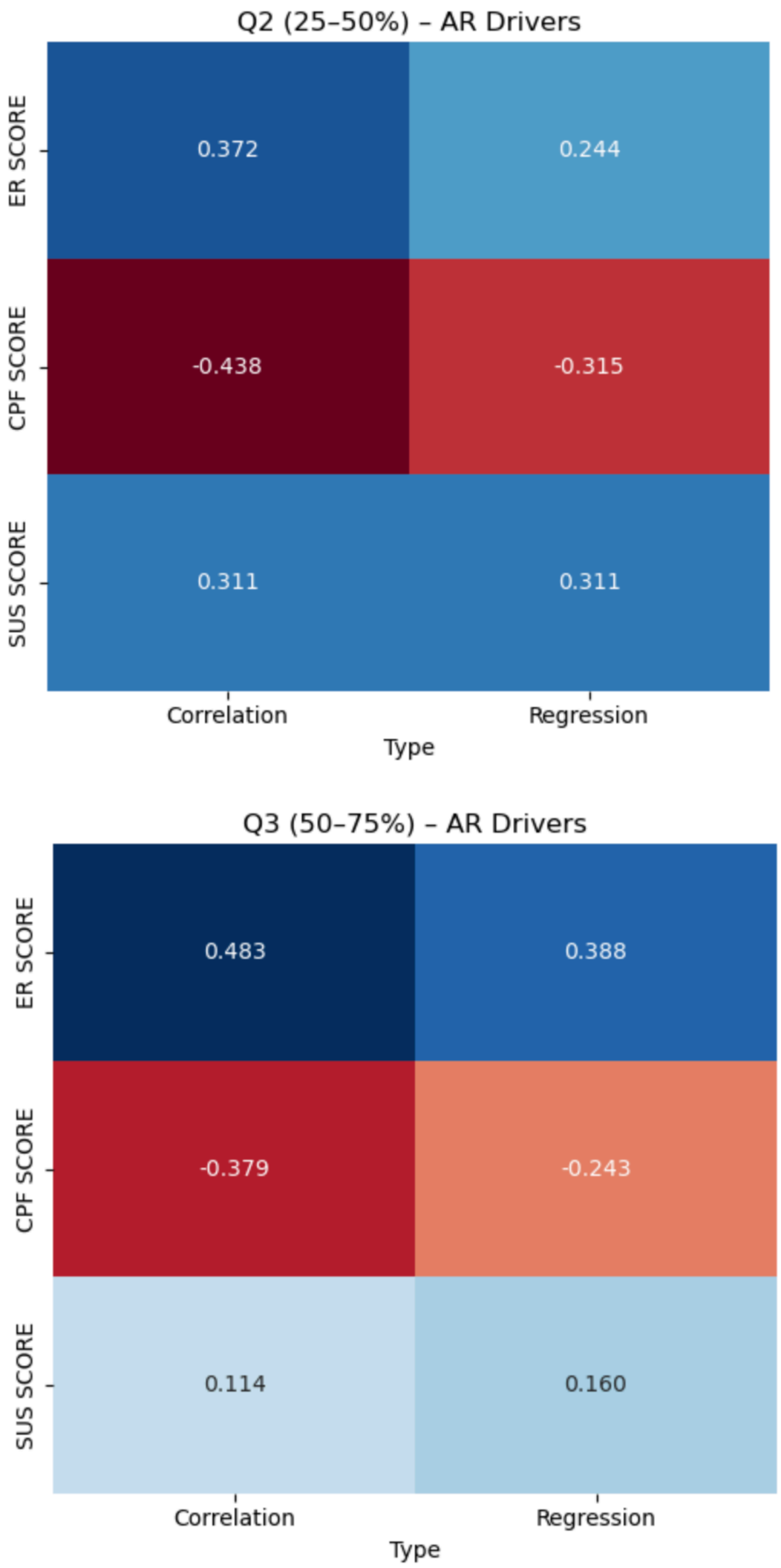
What drives AR beyond the top tier?

The analysis shows three indicators stand out for non-elite universities (Q2–Q3):

Sustainability (SUS)
In top universities (Q1), SUS aligns with prestige, but in Q2–Q4 it becomes more independent.
→ For students, strong SUS programs might signal a forward-thinking, socially responsible choice, even outside the global top 100.

Employer Reputation (ER)
Strongly tied to AR in Q1, but more variable in Q2–Q4.
→ For students, ER highlights graduate employability, industry links, and practical reputation.

Citations per Faculty (CPF)
Reflects research visibility, but in Q2–Q3 shows a weaker or even negative link to AR.
→ Suggests that research output alone doesn't guarantee academic reputation outside the elite tier.



Multiple Linear Regression allows us to test all predictors together.
It answers: "When universities have the same SUS and CPF, does ER still predict AR?"
This helps identify the independent contribution of each score to Academic Reputation.

What the Regression Says for Q2–Q3?

Employer Reputation (ER) → remains a strong positive driver of AR, especially in Q3.
Sustainability (SUS) → positive influence, strongest in Q2.
Citations per Faculty (CPF) → weak or negative, meaning research output alone does not raise AR outside top-tier schools.
R^2 ~ 0.30–0.35 → these 3 factors explain about one-third of the variation in AR.

Regression confirms what we saw in correlations: in mid-tier universities, Academic Reputation is shaped more by employer reputation and sustainability than by research citations. These effects are moderate — not rules — but they give useful hints about what contributes to reputation outside the elite group.

Main thoughts based on our findings

Since ER is 15% of the weight, doing well on employability and employer reputation gives non-elite schools (Q2–Q3) a clear path to increasing their AR.

Although CPF (20%) matters, outside elite tiers its influence on AR is weaker — so students should not assume that high research output always equals high reputation.

SUS is only 5% of the weighting, but because it's newly introduced and less saturated, schools that are proactive on sustainability may gain more relative benefit. Also, students increasingly care about ESG and sustainability, so SUS has growing signaling value.