Australian University Performance Analysis

Tony Nguyen

2022-12-12

##Loading required packages  
library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──  
## ✔ ggplot2 3.4.0 ✔ purrr 0.3.5   
## ✔ tibble 3.1.8 ✔ dplyr 1.0.10  
## ✔ tidyr 1.2.1 ✔ stringr 1.5.0   
## ✔ readr 2.1.3 ✔ forcats 0.5.2   
## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(scales)

##   
## Attaching package: 'scales'  
##   
## The following object is masked from 'package:purrr':  
##   
## discard  
##   
## The following object is masked from 'package:readr':  
##   
## col\_factor

##Importing datasets  
setwd('D:\\Comp\_Sci\_Projects\\Case Study')  
uni\_ranking <- read.csv("Data\\universities\_ranking\_v2.csv")  
uni\_score <- read.csv("Data\\universities\_scores\_v2.csv")  
  
uni\_data <- uni\_ranking %>% inner\_join(uni\_score,   
 by = c("ranking", "title", "location"))  
uni\_data\_v2 <- read.csv("Data\\uni\_data\_v2.csv")  
  
#Summarising necessary data for analysis  
overall\_score\_avg <- uni\_data %>%  
 group\_by(group = ifelse(location == "Australia", "Australia", "Global")) %>%  
 summarise(overall\_score = mean(overall.score))  
  
teaching\_score\_avg <- uni\_data %>%  
 group\_by(group = ifelse(location == "Australia", "Australia", "Global")) %>%  
 summarise(teaching\_score = mean(teaching.score))  
  
research\_score\_avg <- uni\_data %>%  
 group\_by(group = ifelse(location == "Australia", "Australia", "Global")) %>%  
 summarise(research\_score = mean(research.score))  
  
citations\_score\_avg <- uni\_data %>%  
 group\_by(group = ifelse(location == "Australia", "Australia", "Global")) %>%  
 summarise(citations\_score = mean(citations.score))  
  
industry\_income\_avg <- uni\_data %>%  
 group\_by(group = ifelse(location == "Australia", "Australia", "Global")) %>%  
 summarise(industry\_income\_score = mean(industry.income.score))  
  
intl\_outlook\_avg <- uni\_data %>%  
 group\_by(group = ifelse(location == "Australia", "Australia", "Global")) %>%  
 summarise(intl\_outlook\_score = mean(intl.outlook.score))  
  
  
##Joining all category score data into one dataframe  
aus\_vs\_global\_score\_avg <- list(overall\_score\_avg, teaching\_score\_avg, research\_score\_avg,  
 citations\_score\_avg, industry\_income\_avg, intl\_outlook\_avg) %>%  
 reduce(full\_join, by = "group")  
  
## Reordering x-axis categories on the column chart prep work  
order <- c("overall\_score", "teaching\_score", "research\_score", "citations\_score", "industry\_income\_score",   
 "intl\_outlook\_score")  
  
##Hide warnings  
knitr::opts\_chunk$set(warning = FALSE, message = FALSE)

## Introduction

This document will outline the analysis of the Times Higher Education (THE) University Rankings of both Australian and global universities and compare them. It will include analysis to determine whether there is any relationships between the student:staff ratio & total student enrollment numbers to teaching & research quality. There are also comparisons to Australian university performance with the global average using standardised metrics.

## Dataset

The dataset used for this analysis is the Times Higher Education World University Rankings 2021. It was accessed through [Kaggle](https://www.kaggle.com/datasets/matheusgratz/world-university-rankings-2021?select=universities_ranking.csv), and both the author of the dataset and THE allow the use of the data.

The Ranking ranks universities through several categories including:

* Teaching - assessing the learning environment
* Research - assessing reputation of the university’s research
* Citations - how often a university’s research is cited
* International Outlook - how many international students are enrolled and international staff are employed
* Industry Income - How much research income is earned from industry

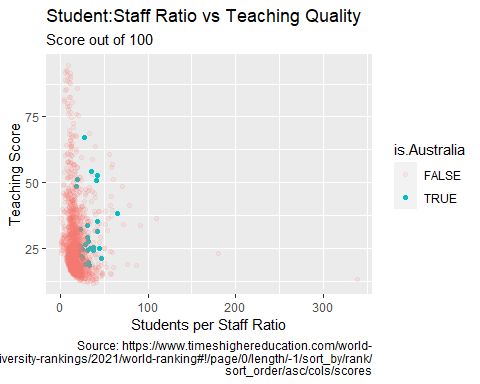
The source for the ranking’s methodology is provided [here](https://www.timeshighereducation.com/world-university-rankings/world-university-rankings-2021-methodology).

## Limitations

While 2022 Ranking data is already available, the latest dataset that includes both university demographics as well as their scores is 2021. In addition, THE does not provide the exact overall score for universities of rank 203 and below, so instead an average of the score band (e.g.band 50.6 to 54.2 into 52.4) is used instead. This shouldn’t affect the analysis as this only applies to the overall score, not the separate category scores that the analysis relies on.

## Relationship between Student:Staff Ratio and Teaching

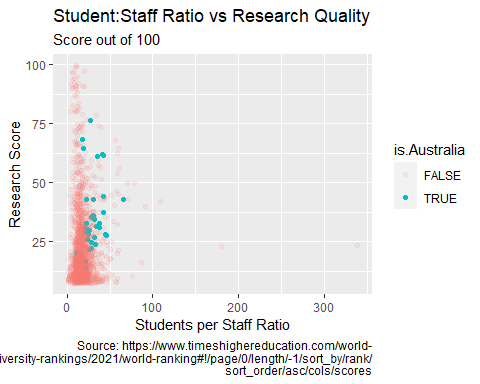
ggplot(data = uni\_data\_v2) +  
 geom\_point(mapping = aes(x = students.staff.ratio, y = teaching.score,  
 color = is.Australia, alpha = is.Australia))+  
 labs( x = "Students per Staff Ratio", y = "Teaching Score",  
 title = "Student:Staff Ratio vs Teaching Quality",  
 subtitle = "Score out of 100",  
 caption = "Source: https://www.timeshighereducation.com/world-  
 university-rankings/2021/world-ranking#!/page/0/length/-1/sort\_by/rank/  
 sort\_order/asc/cols/scores")



Looking at the plots for global universities, **there’s a visible correlation between lower student:staff ratio and higher teaching scores**, though that trend is less pronounced when looking at Australian universities only.

## Relationship between Student:Staff Ratio and Research

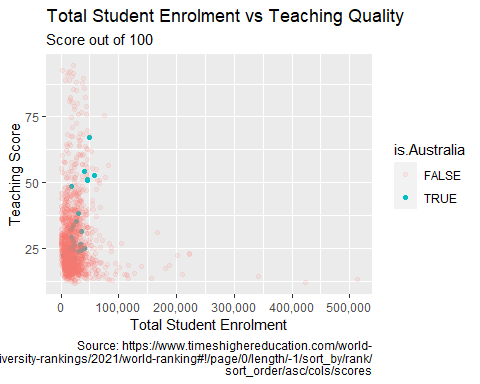
ggplot(data = uni\_data\_v2) +  
 geom\_point(mapping = aes(x = students.staff.ratio, y = research.score,  
 color = is.Australia, alpha = is.Australia))+  
 scale\_fill\_discrete(name = "Australia")+  
 labs( x = "Students per Staff Ratio", y = "Research Score",  
 title = "Student:Staff Ratio vs Research Quality",  
 subtitle = "Score out of 100",  
 caption = "Source: https://www.timeshighereducation.com/world-  
 university-rankings/2021/world-ranking#!/page/0/length/-1/sort\_by/rank/  
 sort\_order/asc/cols/scores")



Likewise, **universities with higher research scores tend to have lower student:staff ratios**, though again the correlation is less pronounced when looking at only Australian universities.

## Relationship between Total Student Enrollment and Teaching

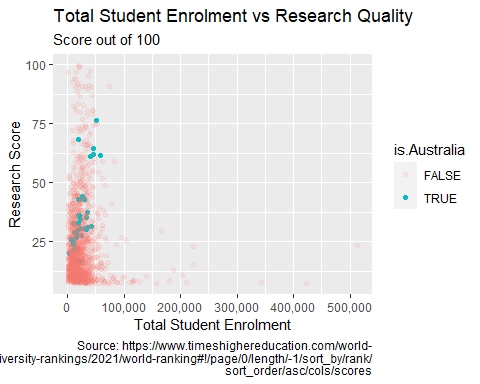
ggplot(data = uni\_data\_v2) +  
 geom\_point(mapping = aes(x = number.students, y = teaching.score,  
 color = is.Australia, alpha = is.Australia))+  
 scale\_fill\_discrete(name = "Australia")+  
 labs( x = "Total Student Enrolment", y = "Teaching Score",  
 title = "Total Student Enrolment vs Teaching Quality",  
 subtitle = "Score out of 100",  
 caption = "Source: https://www.timeshighereducation.com/world-  
 university-rankings/2021/world-ranking#!/page/0/length/-1/sort\_by/rank/  
 sort\_order/asc/cols/scores")+   
 scale\_x\_continuous(labels = label\_comma())



When looking at global trends, universities with lower total student enrollment tend to have higher teaching scores. However, this relationship is reversed when looking at only Australian universities, where **Australian universities with higher student numbers tend to receive higher teaching scores**.

## Relationship between Total Student Enrollment and Research

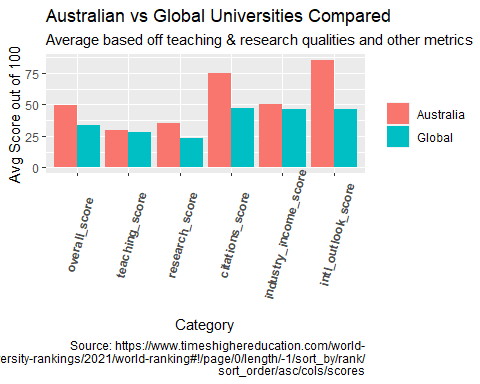
ggplot(data = uni\_data\_v2) +  
 geom\_point(mapping = aes(x = number.students, y = research.score,  
 color = is.Australia, alpha = is.Australia))+  
 scale\_fill\_discrete(name = "Australia")+  
 labs( x = "Total Student Enrolment", y = "Research Score",  
 title = "Total Student Enrolment vs Research Quality",  
 subtitle = "Score out of 100",  
 caption = "Source: https://www.timeshighereducation.com/world-  
 university-rankings/2021/world-ranking#!/page/0/length/-1/sort\_by/rank/  
 sort\_order/asc/cols/scores")+   
 scale\_x\_continuous(labels = label\_comma())



Likewise, global universities tend to have higher research scores with lower total student enrollment, whereas **Australian universities tend to have higher research scores with greater total students**.

## Comparing Category Scores of Australian Universities vs. Global Average

aus\_vs\_global\_score\_avg %>% pivot\_longer(cols=c('overall\_score',  
 'teaching\_score',  
 'research\_score',  
 'citations\_score',  
 'industry\_income\_score',  
 'intl\_outlook\_score'),  
 names\_to = "category",  
 values\_to = "score") %>%  
 ggplot() + geom\_col(mapping = aes( x = category, y = score, fill = group), position = "dodge") +  
 scale\_x\_discrete(limits = order) +  
 labs( x = "Category", y = "Avg Score out of 100",  
 title = "Australian vs Global Universities Compared",  
 subtitle = "Average based off teaching & research qualities and other metrics",  
 caption = "Source: https://www.timeshighereducation.com/world-  
 university-rankings/2021/world-ranking#!/page/0/length/-1/sort\_by/rank/  
 sort\_order/asc/cols/scores") +  
 theme(axis.text.x = element\_text(angle = 75 , vjust = 0.6, face = "bold"))+  
 scale\_fill\_discrete(name = "")



When assessing the category scores of Australian universities against the global average, **whilst research, citation and international outlook scores are well above average, teaching and industry income scores are about on-average.**

##Insights

Naturally, there’s a **negative correlation seen between student:staff ratio and teaching & research quality**. Being responsible for fewer students can allow professors to provide more individualized teaching and allow less time-pressured research.

However, **whilst global universities see a negative correlation between total enrollment numbers and teaching & research quality, Australian universities actually see a positive correlation**.

One explanation for this is that Australian universities receive public funding for enrollment numbers, and the more students that are enrolled into a course the more government funding they are allowed to receive which leads to an incentive to enroll as many students as possible.

When looking at individual categories, **Australian universities’ overall teaching quality as well as its connections to domestic industry are areas of improvement.**

## Recommendations

Australian universities can stand to improve their performance with additional funding. This is enable them to hire more staff to cater to their student populations whilst enabling them to hire more renowned staff from overseas.

In addition, changing funding allocation away from student enrollments will disincentivise universities away from enrolling as many students as possible. It is recommended that funding to be tied more closely to teaching outcomes to ensure students are receiving a higher standard of education.

Also, the federal government should focus on programs that better connect university research with industry outcomes and encourage more private funding from industry. This flow-on effect will enable the Australian economy to become more internationally competitive.

## Summary

In conclusion, analyzing the rankings for Australian universities reveals that they stand to benefit from increased funding that targets improved teaching performance and industry connections. This will enable them to improve their overall standing on the global rankings from average in the above mentioned categories to above the world average.