

Analysis of MOOC Platforms

Arqam Patel, Sanat Goel, Divya Gupta, Sudesh Kumari

2022-11-16

Introduction to MOOCs

MOOCs (Massive Open Online Courses) are internet based correspondence courses. These are one of the most revolutionary forms of knowledge transmission enabled by the internet, allowing large numbers of people access to top-notch educational resources at low prices. These have an increasing importance today, being commonly used by students and professionals to learn new skills and upgrade existing ones.

For the present project, we have limited the scope of our analysis to courses (as opposed to specializations and certifications) offered by conventional universities (as opposed to non-university non profit and corporate entities).

Major MOOC Platforms

Globally, Coursera (<https://coursera.org>) and edX (<https://edx.org>) are the two largest English language MOOC providers by number of users. These have a wide variety of courses and certifications offered by renowned universities and companies, and attract learners from all over the world.

NPTEL (<https://nptel>) stands for the National Programme on Technology Enhanced Learning. As the name suggests, it is an Indian government-funded initiative to make courses from top Indian technical universities available to students throughout the country.

Data

We looked at three datasets of course catalogues of the premier MOOC platforms edX and Coursera and NPTEL, as well as the QS world rankings.

NPTEL

This dataset was obtained by scraping the NPTEL course catalogue website (<https://nptel.ac.in/courses>) using the library rvest. It has 2568 observations of 3 variables, namely Course Name, Institute offering Course, Discipline.

Coursera

This dataset was obtained through the network API call (https://lua9b20g37-dsn.algolia.net/1/indexes/test_products?x-algolia-application-id=LUA9B20G37&x-algolia-api-key=dcc55281ffd7ba6f24c3a9b18288499b&hitsPerPage=1000&page=) from Coursera's search function. It was cleaned and filtered to leave only courses (as opposed to specializations and certifications) offered by universities (rather than companies and non profits). It had 3512 observations of 7 relevant fields.

QS World University Rankings

edX

Biases and shortcomings

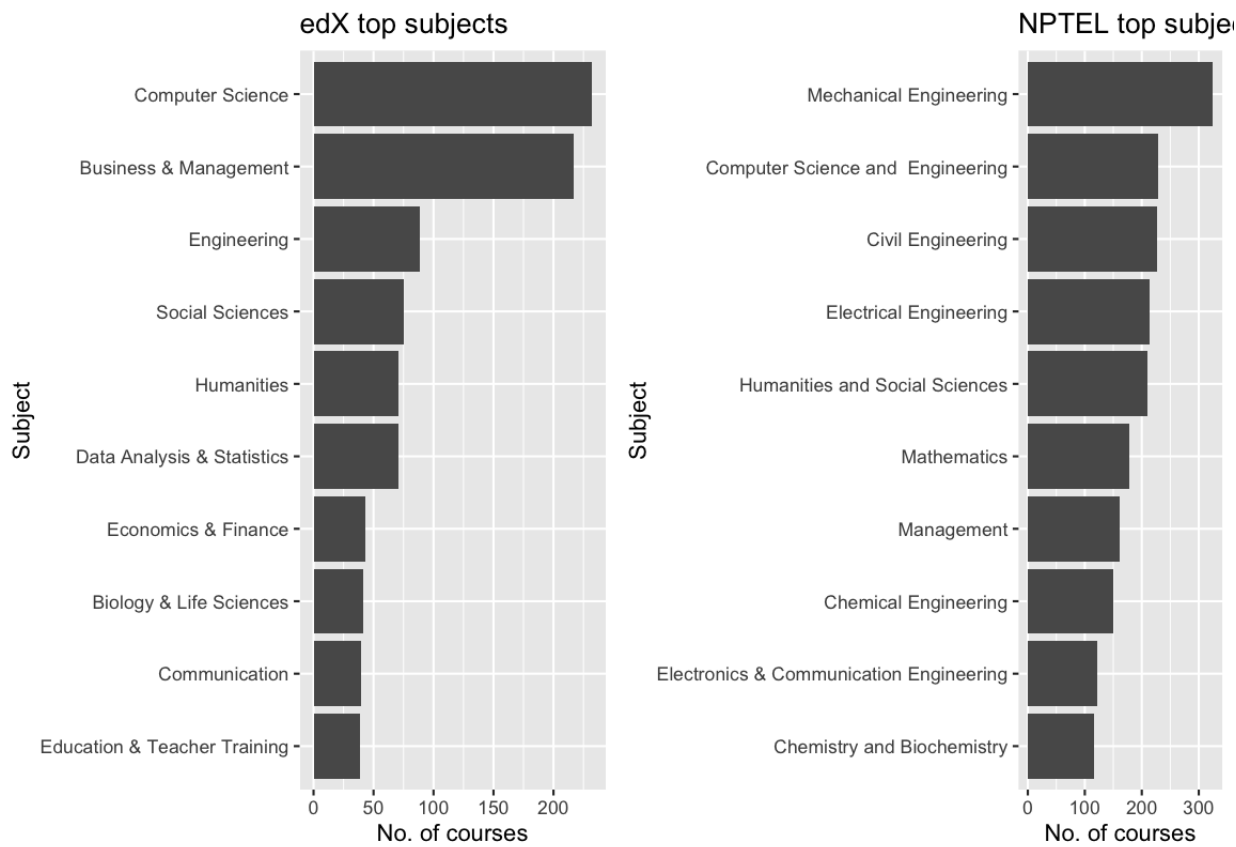
- ## Questions

- ## Plots

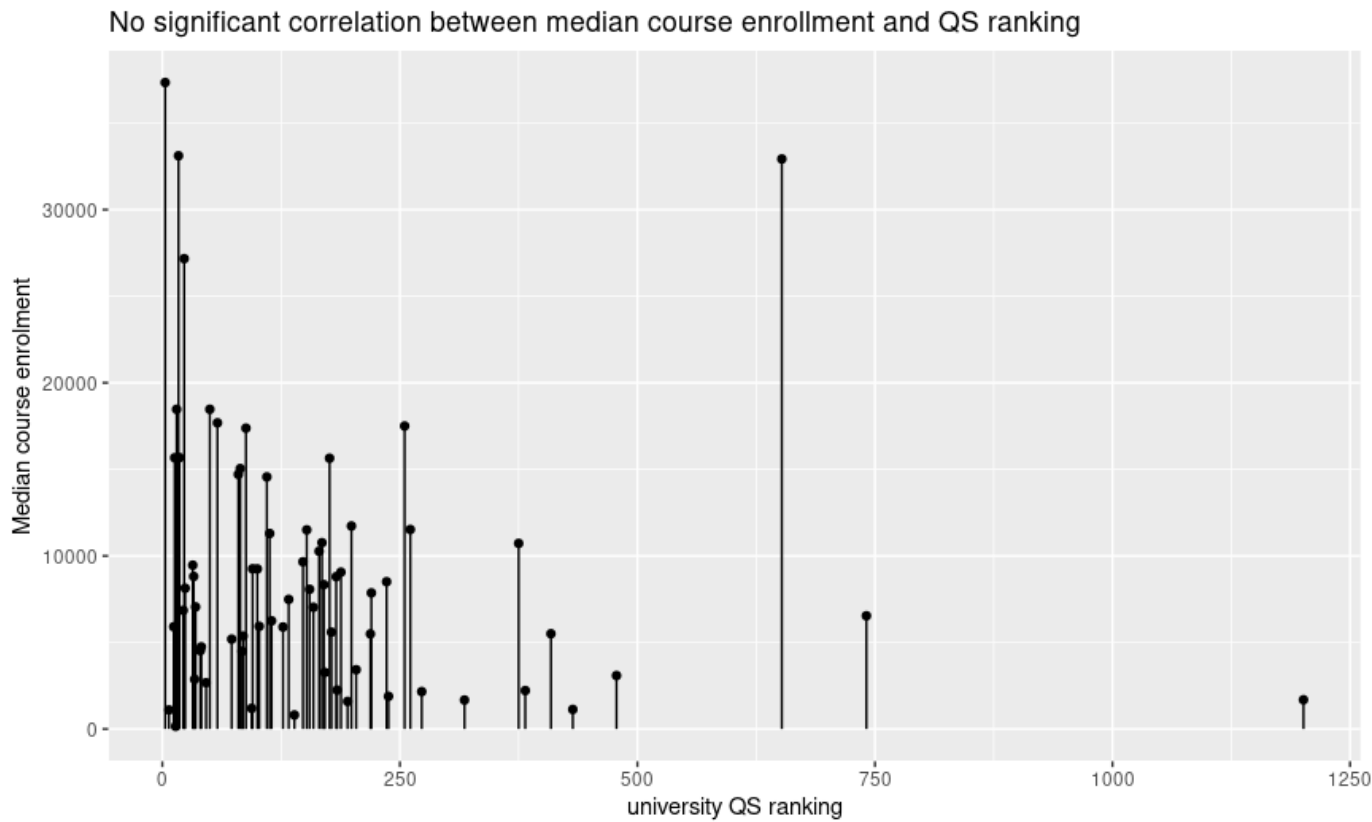
1. Word clouds



2. Subjects

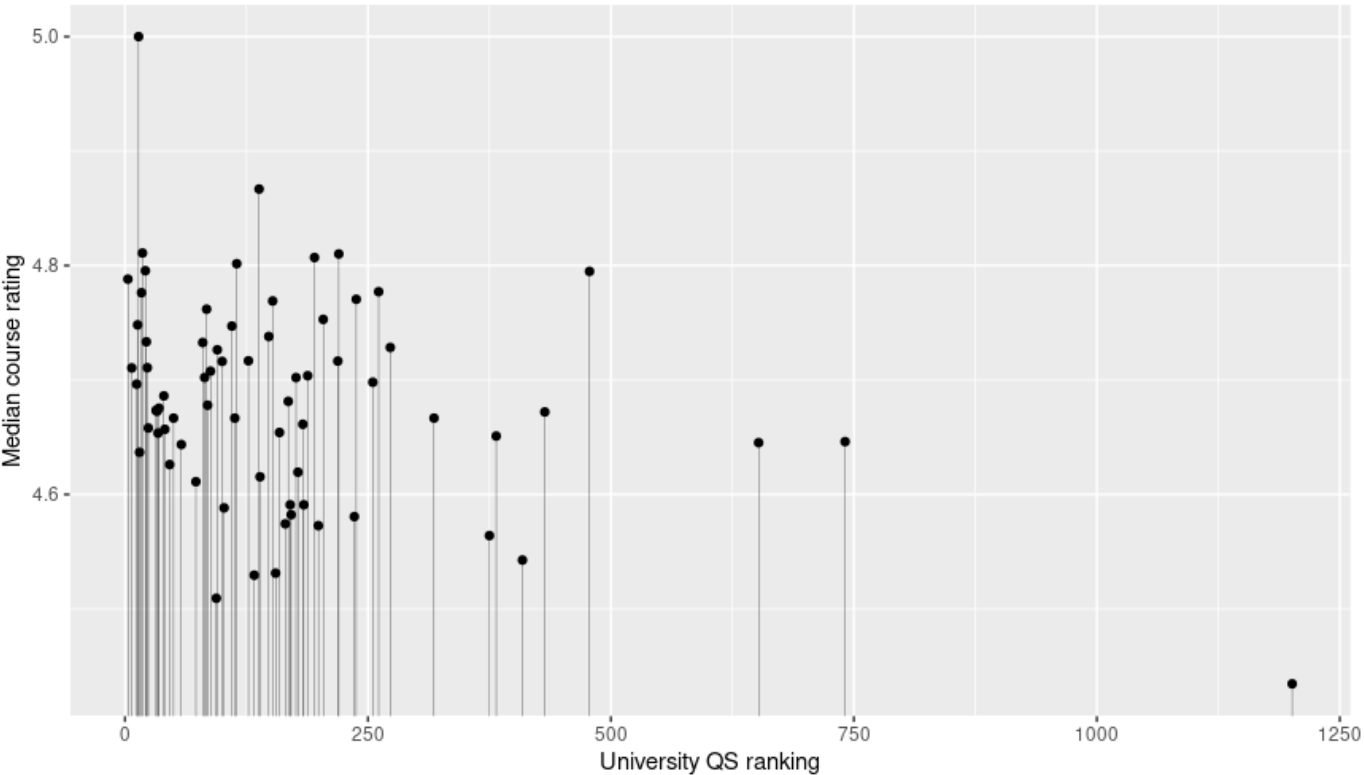


3. QS vs enrolment graph

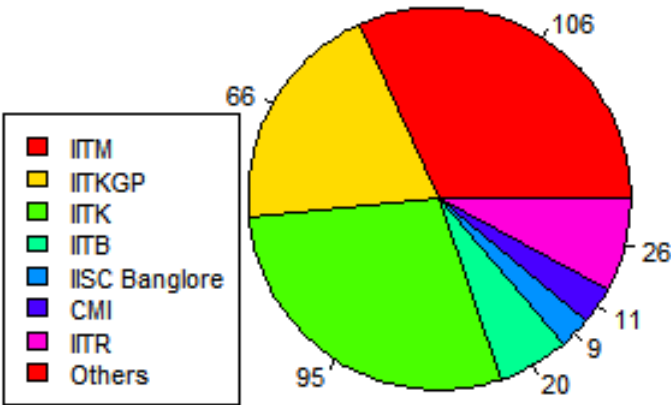


4. QS vs course rating

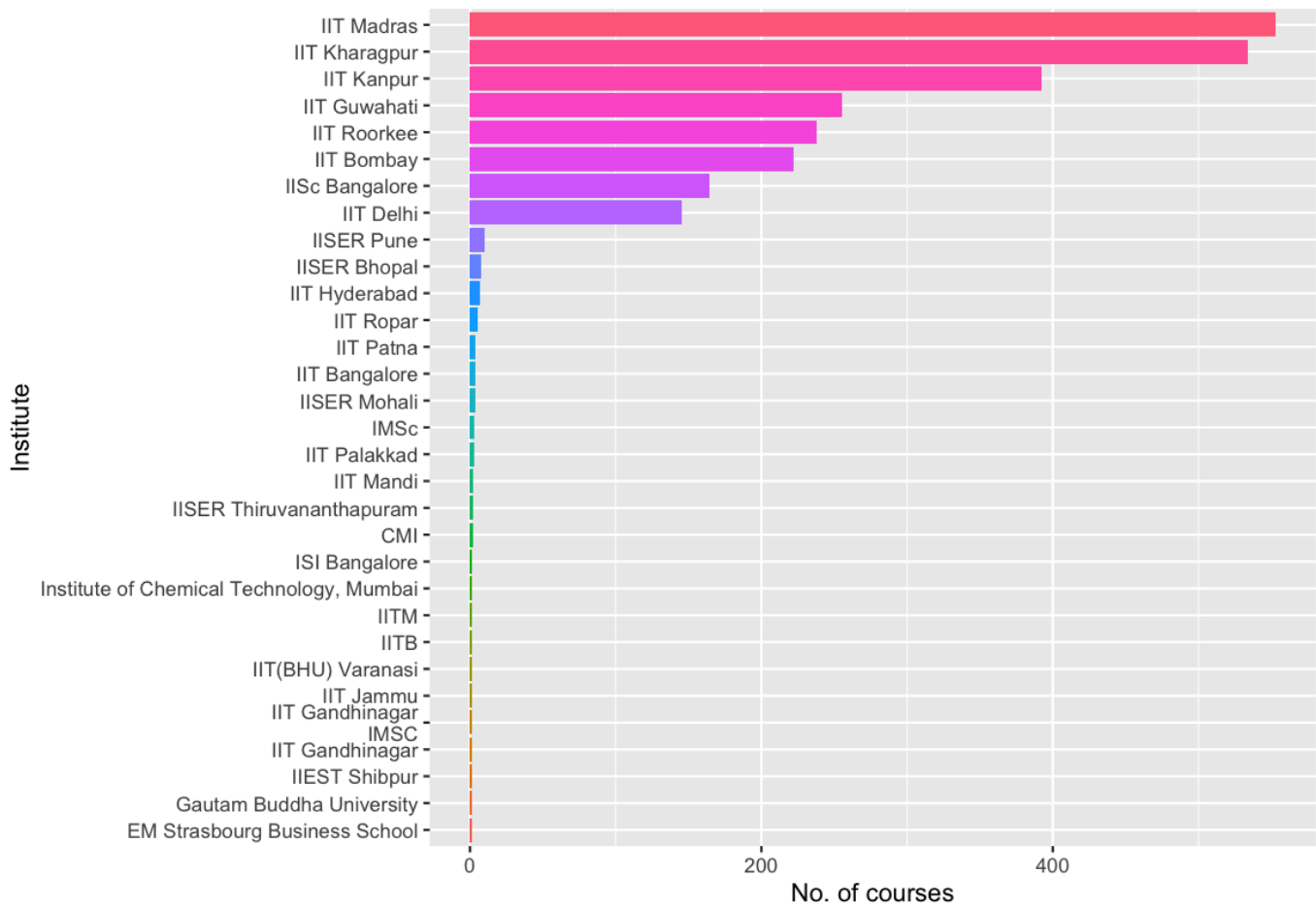
No significant correlation between median course rating and QS ranking



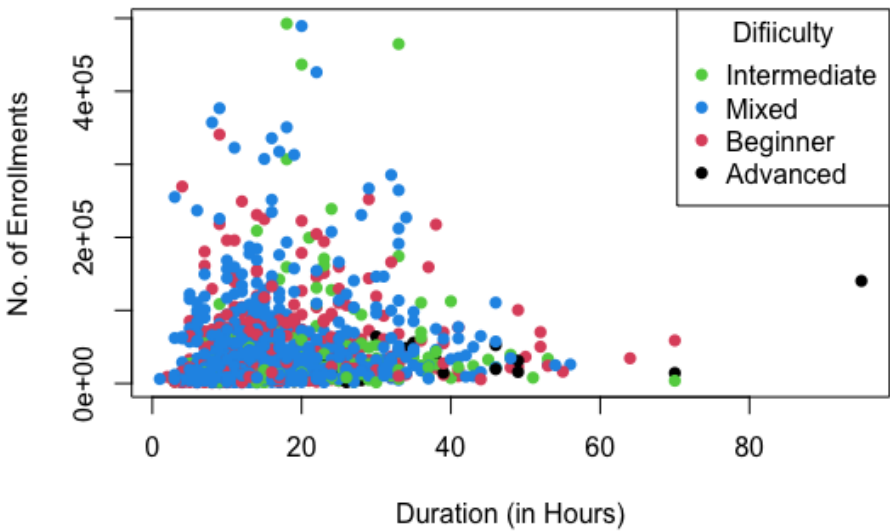
5. NPTEL piechart



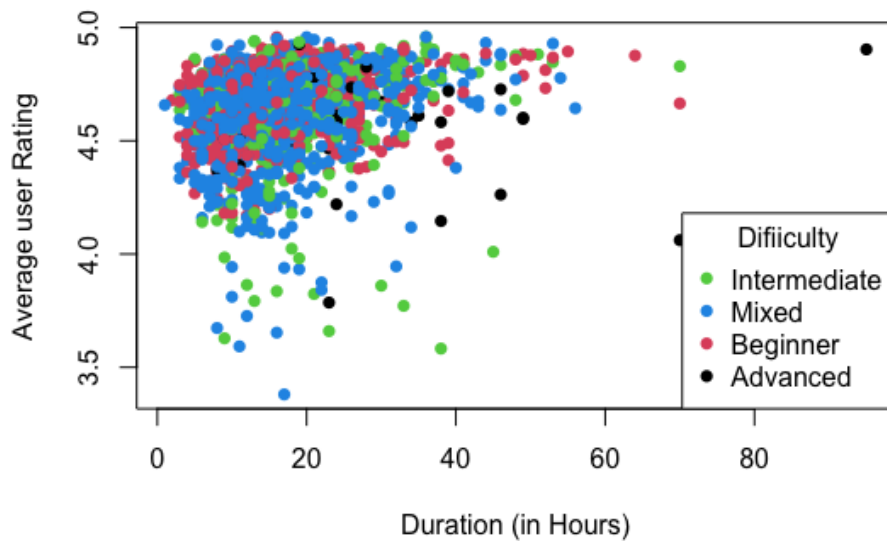
6. NPTEL institutes



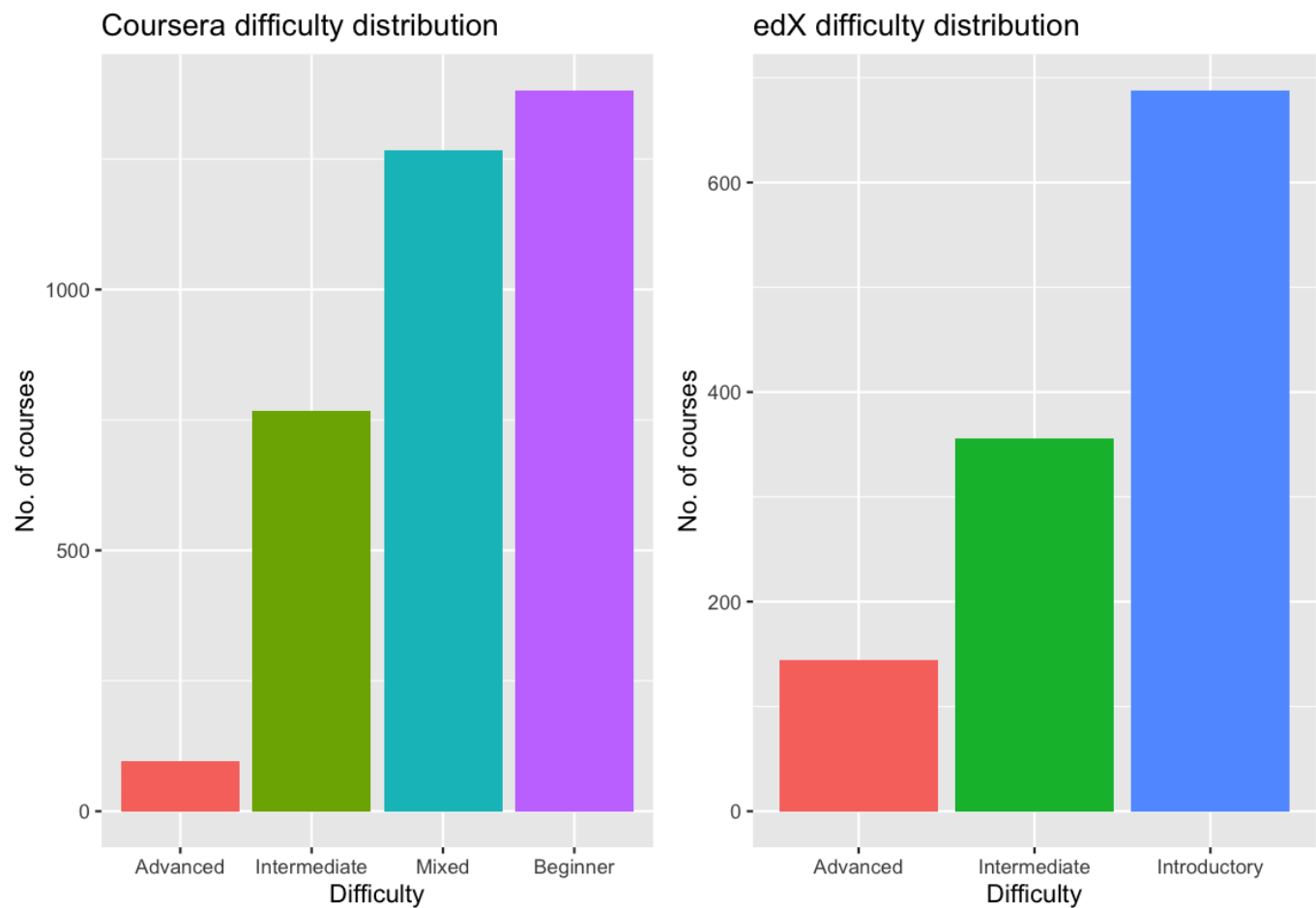
7. Duration vs Enrollments (Coursera)



8. Rating vs Duration (Coursera)



9. Difficulty of Moocs



Conclusions

1. There are relatively few MOOCs on both edX and Coursera for learning Advanced/Difficult topics compared to Beginner/Intermediate level ones. This might be because MOOCs are not a good medium to acquire such skills or because the core user base of these websites consists of introductory and intermediate learners.

2. NPTEL is more core engineering based (top subjects include Mechanical Engineering, Civil Engineering, Electrical Engineering, apart from Computer Science) and can also be said to be fairly representative of the general course catalogue at IITs, whereas edX and Coursera focus more heavily on computer science and management. This suggests that NPTEL is designed/used more as a supplementary academic tool for curricular study.
3. There are significant similarities in the word clouds of Coursera and edX as well as in the difficulty distribution structure. This suggests that both of these cater to similar needs, i.e. introductory/intermediate level learners interested in developing skills in tech, management and business. This is in contrast to NPTEL, which suggests a difference in the audience targeted by NPTEL and edX/Coursera.
4. The vast majority (2503 out of 2568, or 97.47%) of courses on NPTEL are provided by 8 institutes (out of 31 total). A report (<https://factordaily.com/nptel-india-edtech-mooc/>) from 2017 has a breakdown of MOOCs offered on the platform at that time by university. This revealed that IIT Madras, IIT Kanpur and IIT Kharagpur offered the majority, with the rest of the IITs not significantly ahead of the other universities like CMI. Juxtaposed with the current distribution of courses, we can see that IIT Madras, IIT Kanpur and IIT Kharagpur continue to be the top 3 institutes in terms of number of courses offered, although other institutes have also contributed significantly since then.
5. Median MOOC enrolment does not correlate very well with university QS ranking (correlation coefficient = -0.15), suggesting that other course specific variables (course content, quality, reputation, etc.) may matter more for enrolling users than the quality of institute offering the course.
6. Similarly, Median course rating has a weak correlation (correlation coefficient = -0.35) with university QS ranking and similar correlation with QS score (~0.36). This shows that MOOC quality does not solely depend on the quality of the institute, but that university quality does influence course quality to some extent.
7. A large majority (56 out of 68) of universities (listed in QS rankings) having more than 5 courses on Coursera fall in the top 250 QS ranks. Thus, top universities are overrepresented amongst MOOC organisers.
8. No single variable among rating (correlation coefficient = 0.05), price (c.c. = 0), university QS rank or course duration (c.c. = 0.10 for Coursera, 0.16 for edX) has a significant correlation (on its own) with Enrolments. This suggests that no single variable can be used as a good predictor for course success. This is valid because decisions to choose MOOCs are complex choices influenced simultaneously by multiple variables, including some which may not have been considered here (like course SEO, number of mentions on the web).

References

1. MOOC Data Insights through Scraping edX (<https://nycdatascience.com/blog/student-works/mooc-insights-scraping-edx/>)
2. What makes NPTEL India's EdTech hero — without newsy fundings and valuations (<https://factordaily.com/nptel-india-edtech-mooc/>)

3. Data Scraping Coursera to Find the Perfect Course
(<https://www.jibbyjames.com/post/coursera-data-scraper/>)
4. NPTEL (<https://nptel.ac.in>)
5. Coursera (<https://coursera.org>)
6. edX (<https://edx.org>)