**A panoramic map of MOOCs:**

**A comparison on MOOCs between the South and the North**

Tsinghua University

Fu Jie, Huang Zhenzhong

**Abstract**

*MOOCs was expected to decrease the educational inequality all over the world, while it is still uncertain that if MOOCs is a channel for dialogue and communication or just a one-way pipe from the academic core to the periphery. This study aims to find out whether there is knowledge production gap between South and North by using data mining techniques to analyze the course information of MOOCs offered by different type of countries. This study concentrates on the MOOCs supplying patterns in three levels, country level, institute level and discipline level. And it turns out that there is a huge gap between different division of countries in all kinds of sectors. Besides, even countries in the same division of South and North have various MOOCs supplying patterns due to the diverse economic and educational conditions.*

**Introduction**

MOOCs was expected to decrease the educational inequality all over the world. However, stating that “The world is not flat,” Rivard (2013) posited that three issues may limit the portability of Western MOOCs to other countries and cultures: language barriers; varying cultural expectations about pedagogies and learning environments; and limited access to the internet. On the other hand, once these barriers were broken through, there would be another awkward situation that MOOCs may represent or even increase the inequality between Southern and Northern countries. Wallerstein’s World System Theory divides the world into core countries, semi-periphery countries, and the periphery countries. Core countries focus on higher skill, capital-intensive production, and the rest of the world focuses on low-skill, labor-intensive production and extraction of raw materials. Those theories mostly concentrate on business field.

So, what is the situation when it comes to knowledge production? Normally, books or academic papers may be perfect to represent the knowledge production. However, it is not a good idea to analyze that large quantity of data. Thankfully, we have MOOCs nowadays, which is a fine knowledge carrier and a good case to study since it is open accessible and relatively small in data size.

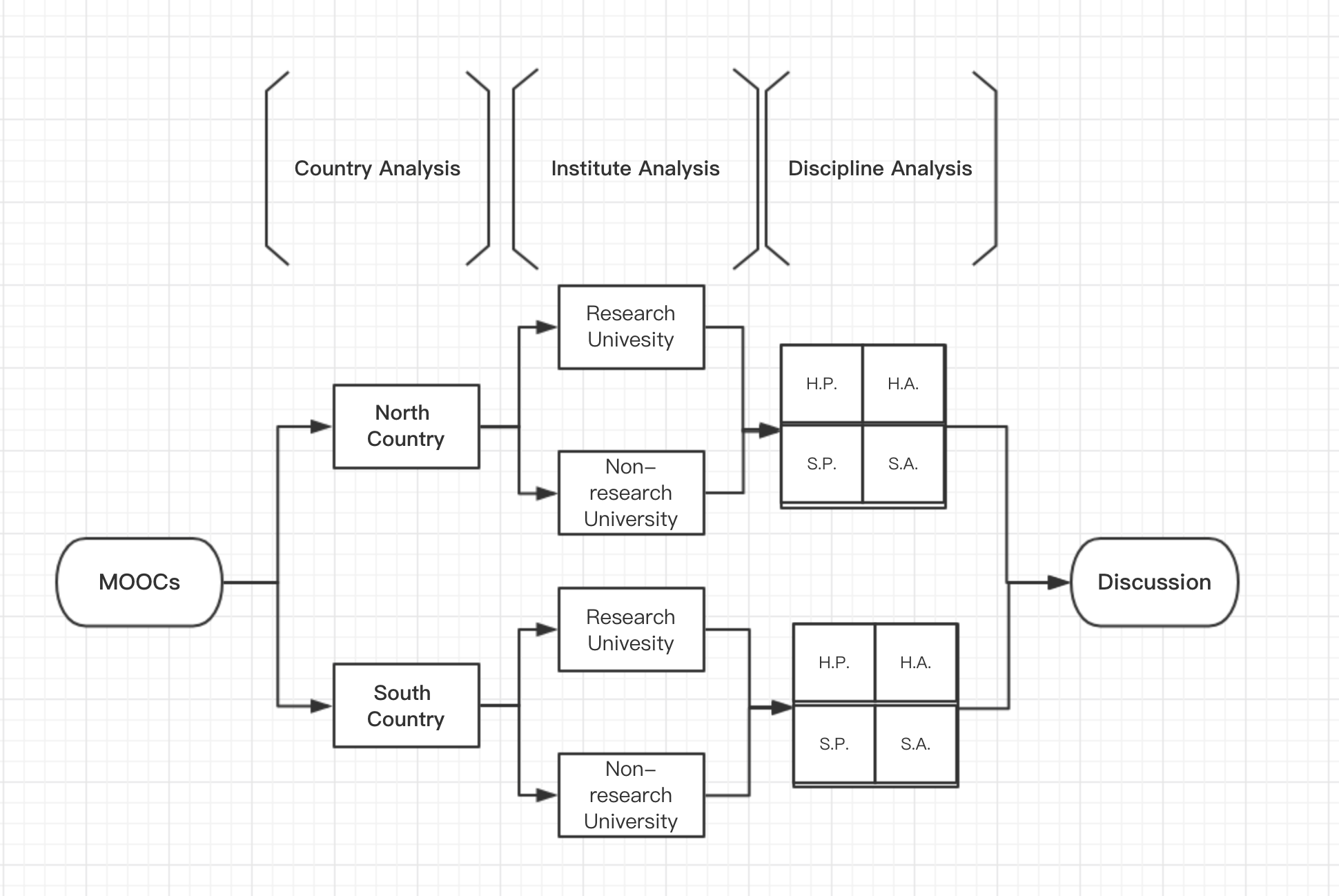
**Research Questions**

What are the differences between MOOCs offered by South and North countries? What do those differences mean? Why there are such differences existing?

**Research design and Framework**

**Research design**

This study will answer the research question “What are the differences between MOOCs offered by South and North countries? ” in three levels. Firstly, country level. By using country’s GDP per capita as a more specific index of country type, we can analyze correlation between country type and MOOCs offering pattern. Secondly, institute level. Introducing research and non-research university division, . Thirdly, encoding all the courses into four categories and then to find out what differences are between MOOCs’ disciplines of the North and South.



Pic 1. Analysis flow chart

**North South divide**

The Global South / North delineations are from the Strategic Plan: "For the purposes of our discussion, the Global North includes Australia, Canada, Israel, Hong Kong, Macau, New Zealand, Japan, Singapore, South Korea, Taiwan, the United States and all of Europe (including Russia). The Global South includes Asia (with the exception of Japan, Hong Kong, Macau, Singapore, South Korea and Taiwan), Central America, South America, Mexico, Africa, and the Middle East (with the exception of Israel)."

**Tony Becher disciplines theory**

Tony Becher divides all knowledge into four groups based on nature of knowledge and disciplinary culture:

Pure Sciences (e.g. Physics) : “hard pure”. The nature of hard pure knowledge is cumulative; atomistic; concerned with universals, quantities, simplifications; resulting in discovery. The nature of disciplinary culture is competitive; gregarious; politically well-organized; high publication rate.

Humanities (e.g. History) and pure social sciences (e.g. Anthropology): “soft pure”. The nature of soft pure knowledge is reiterative; holistic; concerned with particulars, qualities, complication; resulting in understanding. The nature of soft pure is individualistic, plurastic; loosely structured; low publication rate, etc.

Technologies (e.g. Engineering): “hard applied”. The nature of hard applied knowledge is purposive; pragmatic; concerned with mastery of physical environment; resulting in products.

The nature of disciplinary culture is entrepreneurial, cosmopolitan; dominated by professional values; patents substitutable for publication, etc.

Applied Social Sciences (e.g. Education): “soft applied”. The nature of soft applied knowledge is functional; ulitarian; concerned with enhancement of professional practice; resulting protocols. The nature of disciplinary culture is outward-looking; uncertain in status; dominated by intellectual fashions, etc.

Pic 2. Tony Becher’s disciplinary grouping theory

**GDP per capita**

Gross domestic product (at purchasing power parity) per capita, the purchasing power parity (PPP) value of all final goods and services produced within a country in a given year, divided by the average (or mid-year) population for the same year. GDP per capita is often considered an indicator of a country's standard of living; although this is problematic because GDP per capita is not a measure of personal income.

This paper uses GDP per capita as an explanatory variable, and also for clustering.

**Research university and non-research university**

Academic Ranking of World Universities (ARWU), also known as Shanghai Ranking, is an annual publication of university rankings by Shanghai Ranking Consultancy. The league table was originally compiled and issued by Shanghai Jiao Tong University in 2003, the first global ranking with multifarious indicators, after which a board of international advisories was established to provide suggestions

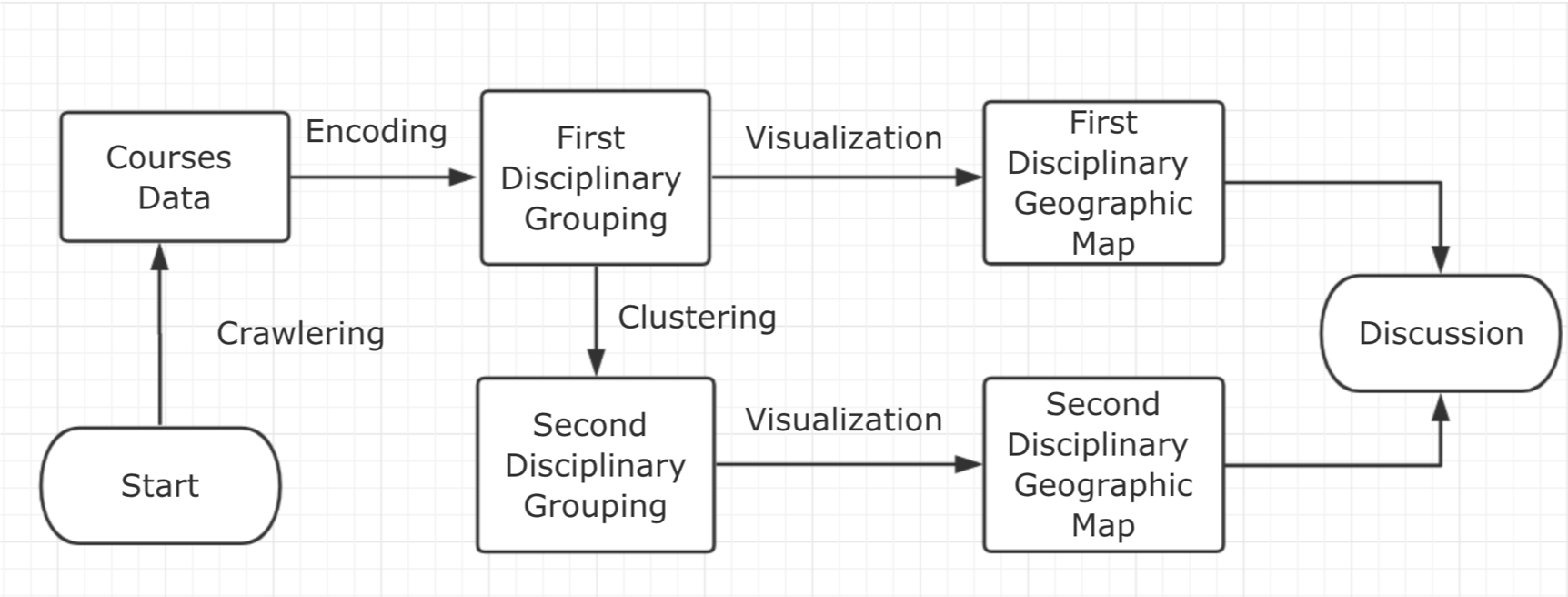
This paper uses research/non-research as an explanatory dummy variable, and also for clustering. We define institutes ranking top 200 in ARWU2017 as research institutes, otherwise non-research institutes.

**Data source**

This study analyzes and compares the basic information of MOOCs supplied by Southern and Northern countries, especially when it comes to the discipline types of the courses and what kind of knowledge they supply. The sample courses are from main MOOCs platforms like Coursera, edX, FutureLearn, and integrated MOOCs website Class-Central.

**Procedure**

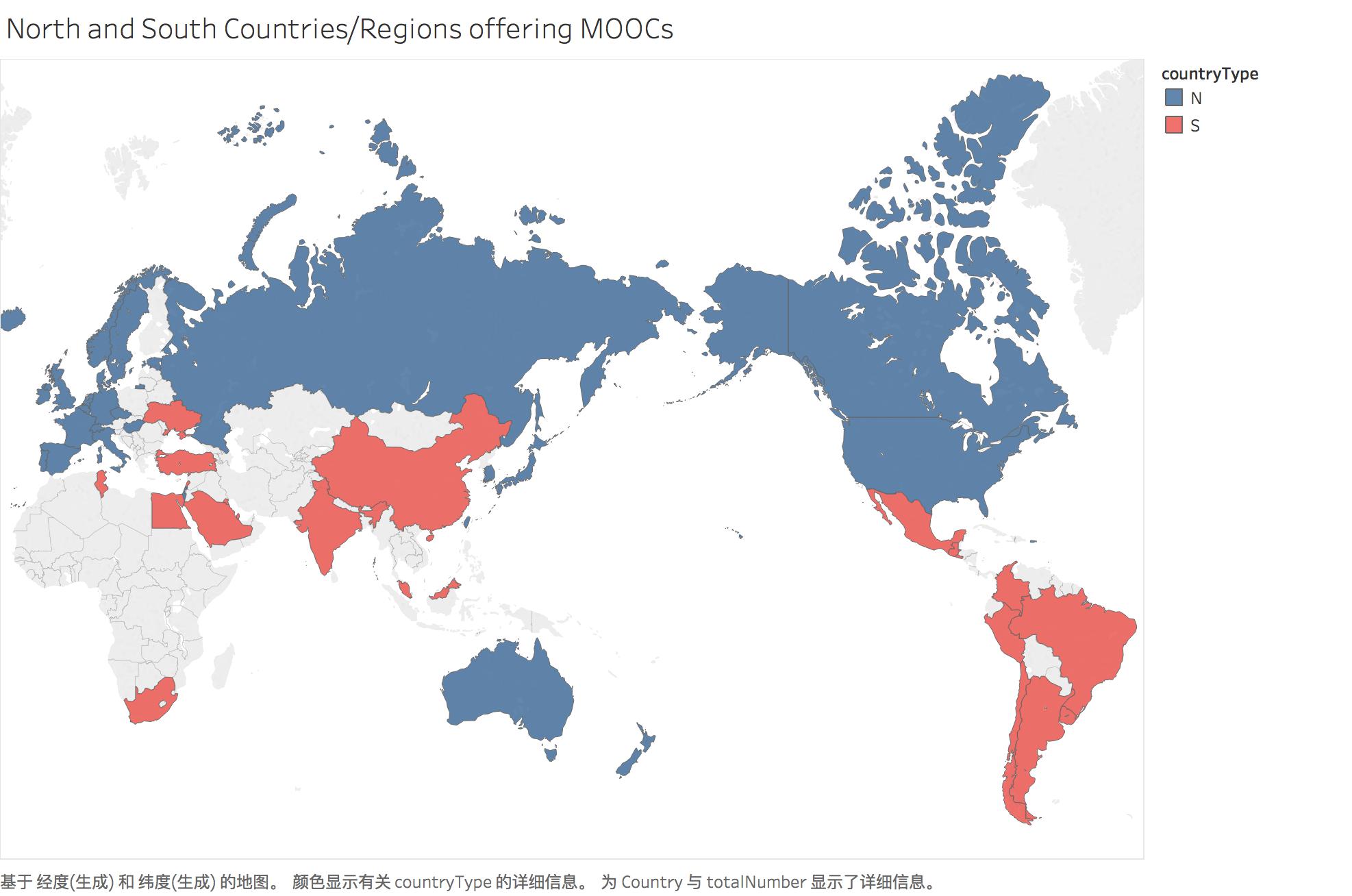
Firstly, using crawler technology to fetch courses information from these websites. Courses information includes but not limited to university, platform, country, language, subject, certificate, interested students (if the course is included in Class-Central), cycle and so on. Secondly, using machine learning technique to classify MOOCs types of different countries. Considering Mode 2 theory and Disciplinary Grouping theories. Gibbons and colleagues defined a new form of knowledge production called "mode 2", which is context-driven, problem-focused and interdisciplinary. He and his colleagues distinguished this from traditional research, labelled "mode 1", which is academic, investigator-initiated and discipline-based knowledge production. Meanwhile, according to knowledge and culture, Tony Becher (1987) classified all disciplines into four groups, Pure science(hard-pure), Humanities (pure social sciences), Technologies(hard-applied), and Applied social science('soft-applied'). Based on these two theoretical frameworks, all the courses will be coded into different categories. Finally, using data visualization methods to present the status quo of the MOOCs curriculum worldwide.



Pic 3.Task flow

**Results**

**Country analysis**



Pic 4. North and South countries publishing MOOCs

Tab 1. North and South countries publishing MOOCs

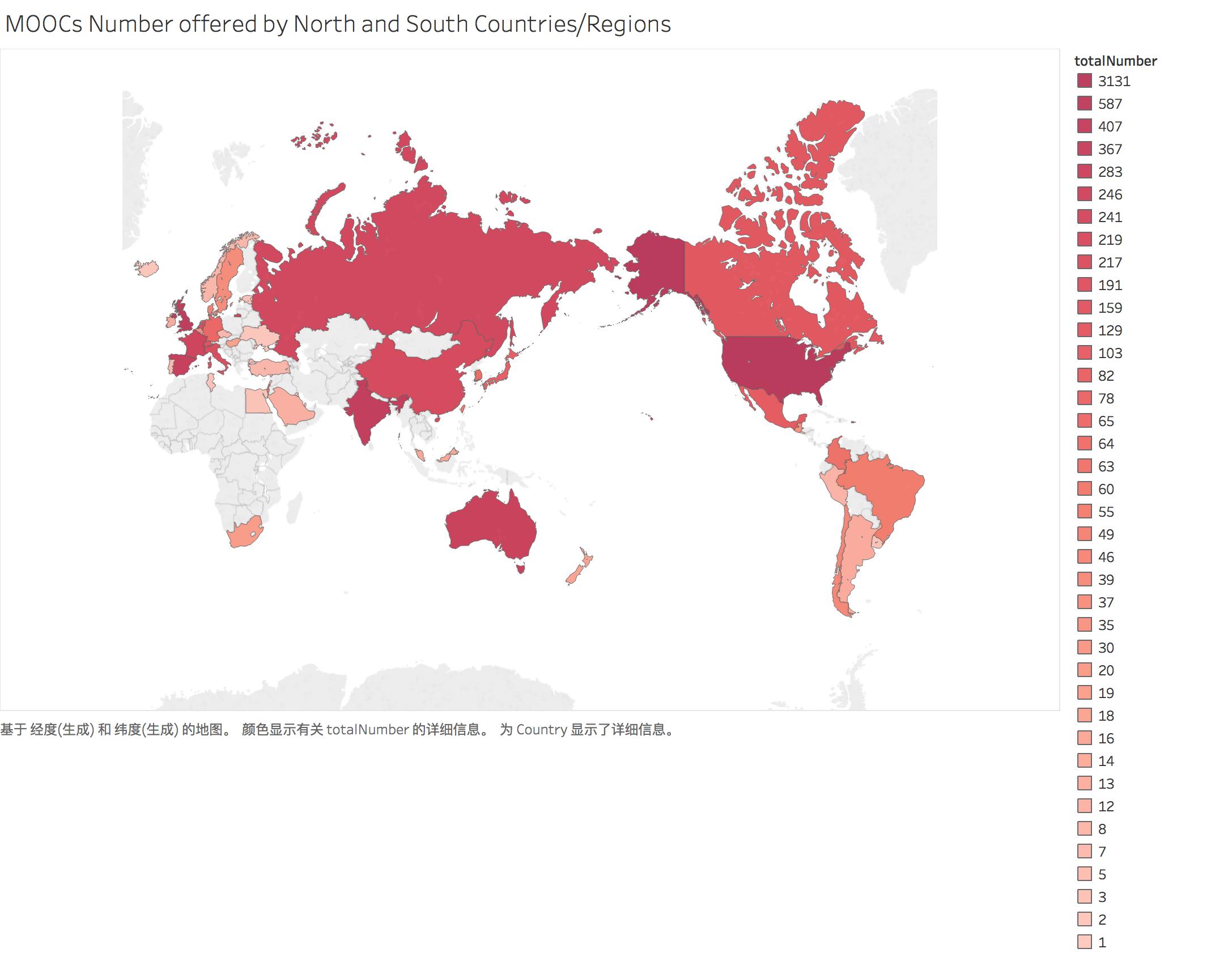
|  |  |  |  |
| --- | --- | --- | --- |
| Country Type | Number | MOOCs Number | MOOCs Number/Country |
| North Country | 28 | 6156 | 220 |
| South Country | 22 | 1005 | 46 |

There are 50 countries offering MOOCs in Class-Central, 28 of them are Global South countries, and 22 of them are Global North countries. Even though approximately the same number of countries in both sides have published MOOCs, there is huge difference in courses scales. The courses from North are six times larger than courses from South.

Among Top 15 offering MOOCs country/region, there are only 3 of them (India, China, Mexico) from south side.

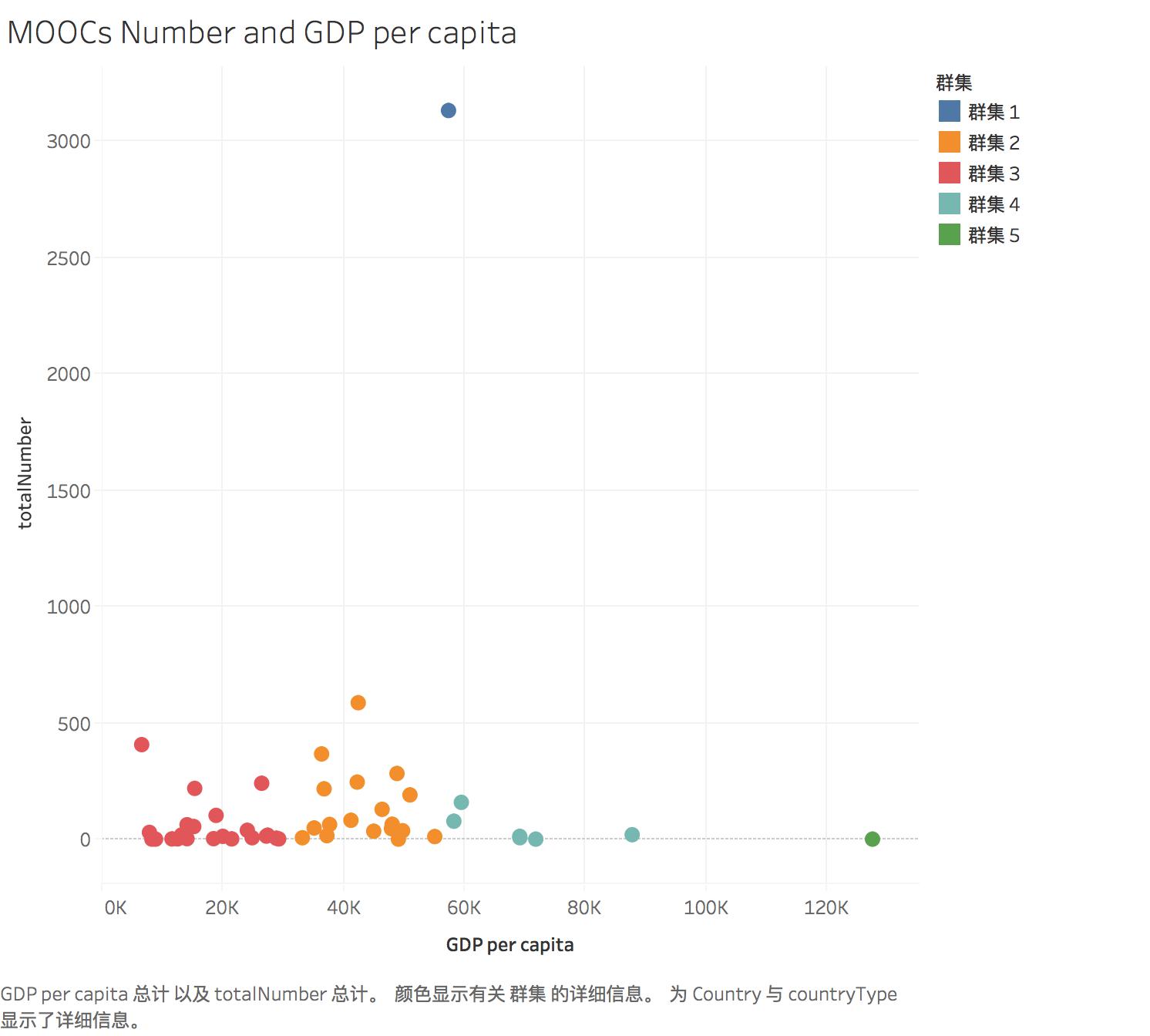
Tab 2. Top 15 publishing MOOCs

|  |  |  |
| --- | --- | --- |
| Country/Region | Courses Number | Country Type |
| USA | 3131 | N |
| UK | 587 | N |
| India | 407 | S |
| Spain | 367 | N |
| Australia | 283 | N |
| France | 246 | N |
| Russia | 241 | N |
| China | 219 | S |
| Italy | 217 | N |
| Netherlands | 191 | N |
| Switzerland | 159 | N |
| Canada | 129 | N |
| Mexico | 103 | S |
| Japan | 82 | N |
| Hong Kong China | 78 | N |

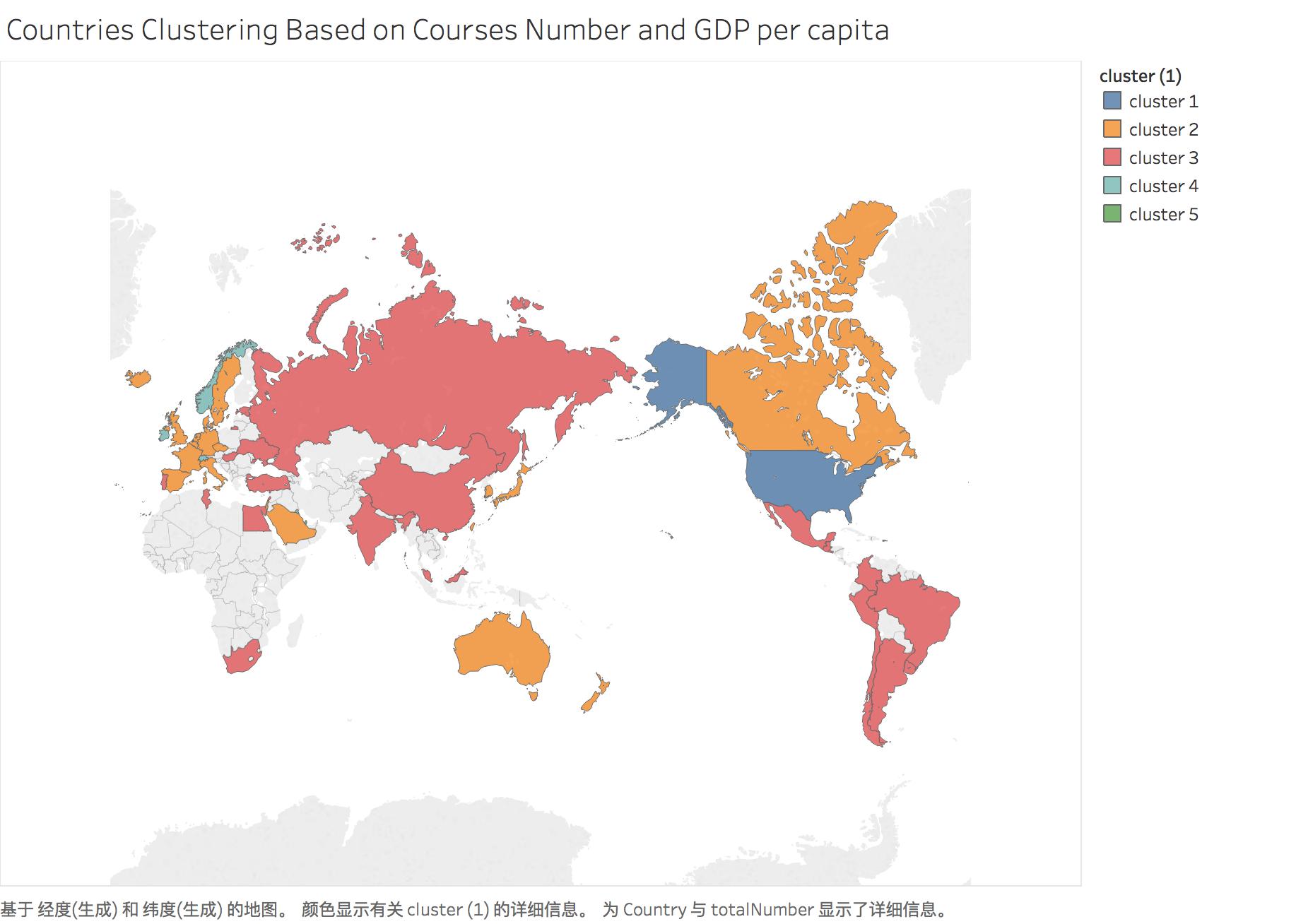


Pic 5. Different countries’ MOOCs number

Introducing GDP per capita as a new variable, along with country’s MOOCs number, and running clustering algorithm, we divide all the countries into five clusters automatically. Apparently, cluster 1 and cluster 5 are both outliers. America offering almost half of the MOOCs in Class-Central, while Qatar has highest GDP per capita in the world. Countries in cluster 2 have very high income, very large number of MOOCs and mainly from the North (except for Saudi Arabia). Countries in cluster 3 are mostly South countries, having middle income, supplying medium number of MOOCs. Countries in cluster 4 have very high income, offering relatively less courses.

****

Pic 6. Countries clustering based on MOOCs number and GDP per capita

****

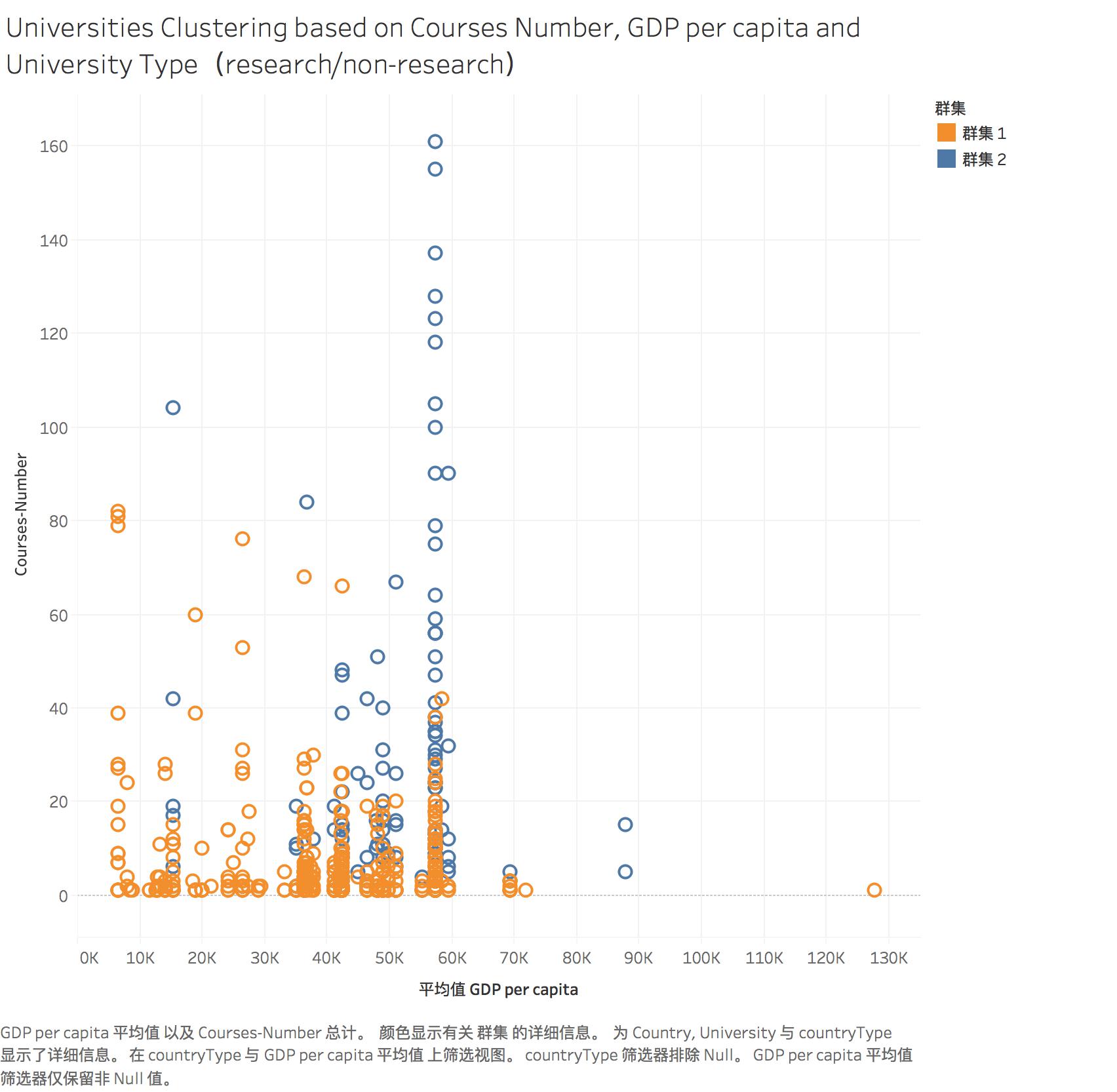
Pic 7. Countries clustering based on MOOCs number and GDP per capita

Tab 3. Countries clustering based on MOOCs number and GDP per capita

|  |  |  |
| --- | --- | --- |
| Cluster | Feature | Country/Region |
| Cluster 1 | High income, extremely large number of MOOCs | USA |
| Cluster 2 | Very high income, very large number of MOOCs | UK, Spain, Australia, France, Italy, Netherlands, Canada, Japan, Germany, Taiwan China, Israel, Denmark, Sweden, Belgium, New Zealand, Saudi Arabia, Czech  Iceland, South Korea |
| Cluster 3 | Mainly South countries, middle income, medium number of MOOCs | China, India, Russia, Mexico, Colombia, Brazil, Chile, Guatemala, South Africa, Hungary, Malaysia, Argentina, Peru, Turkey, Portugal, Lebanon., Grenada, Estonia, Uruguay, Egypt, Tunisia, Salvador, Ukraine |
| Cluster 4 | Very high income, relatively less in course | Switzerland, Hong Kong China, Singapore, Ireland, Norway, Kuwait |
| Cluster 5 | Extremely high income, very small number of MOOCs | Qatar |

**Institute analysis**

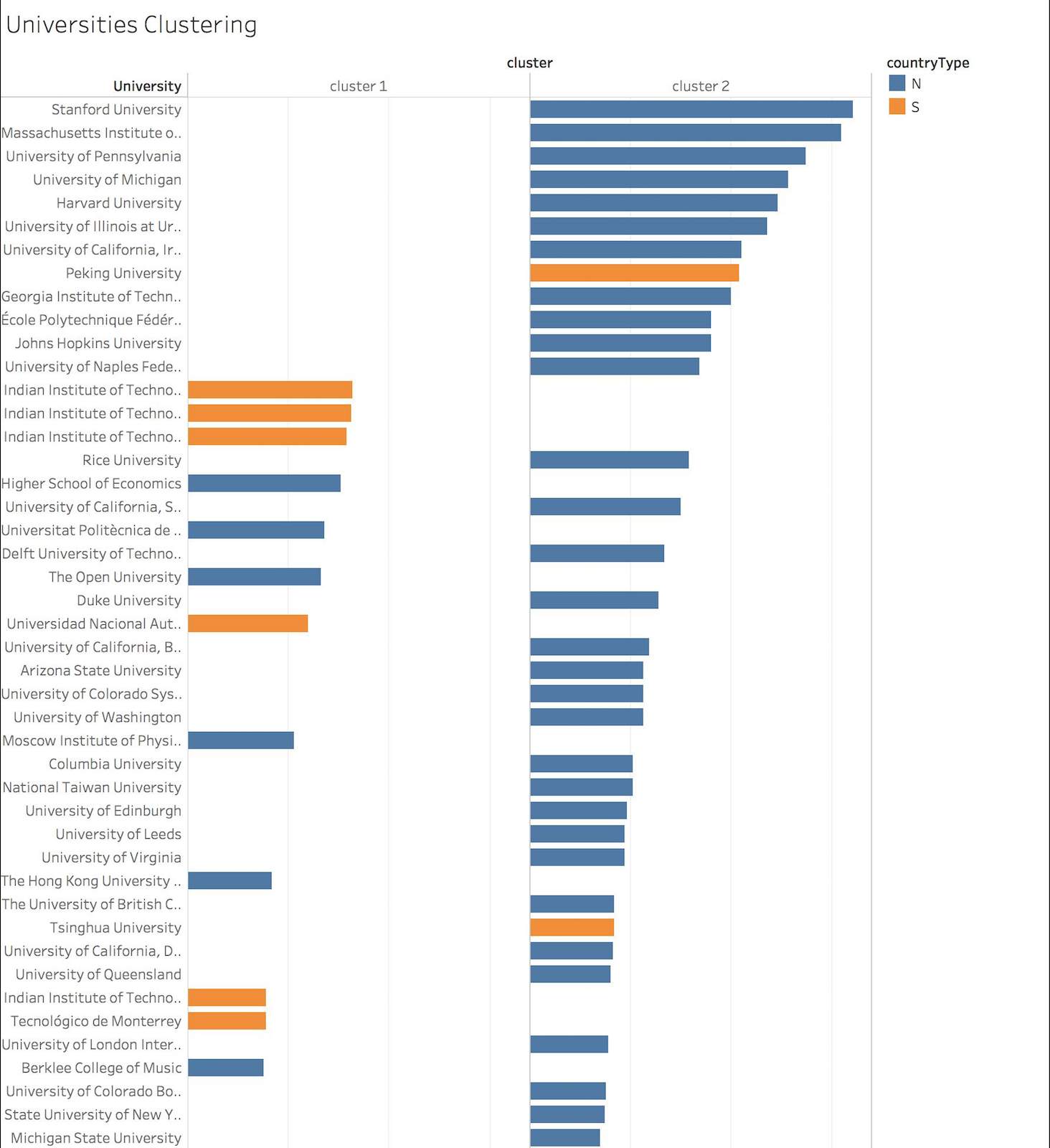
Adding boolean variable, research/non-research, along with GDP per capita, courses number, we have a new clustering of all the universities. and this clustering includes two clusters. Cluster 1 is mainly non-research type, small scales in MOOCs, and covering more South countries. There are 596 institutes in cluster 1. Cluster 2 is research type, supplying more courses. Notably, Tsinghua and Peking University from China are in cluster 2 along with Stanford, Harvard, MIT and so on, while institutes like Higher School of Economics are in cluster 1, even though these universities are from North countries. It is hardly surprising consequence considering the economic aggregate and economic growth speed of the two countries. Countries in the same division of South and North have different MOOCs supplying patterns due to the diverse economic and educational conditions.



Pic 8. Countries clustering based on MOOCs number, research/non-research and GDP per capita

Tab 4. Countries clustering based on MOOCs number, research/non-research and GDP per capita

|  |  |  |
| --- | --- | --- |
| Cluster | Feature | Universities |
| Cluster 1 | Non research institute; small number of MOOCs; more likely South countries | Indian Institute of Technology, Kharagpur;  Higher School of Economics;  Yonsei University;  Lancaster University;  Indian Institute of Technology Bombay;  Yeshiva University,  University of West Florida,  Norwich University of the Arts .etc  Total: 596 |
| Cluster 2 | Research institute;  Large number of MOOCs;  More likely North countries | Stanford University;  Massachusetts Institute of Technology;  University of Pennsylvania;  University of Michigan;  Harvard University;  Peking University;  Tsinghua University;  University of Queensland. etc  Total: 169 |



Pic 9. Countries clustering based on MOOCs number, research/non-research and GDP per capita

**Discipline analysis**

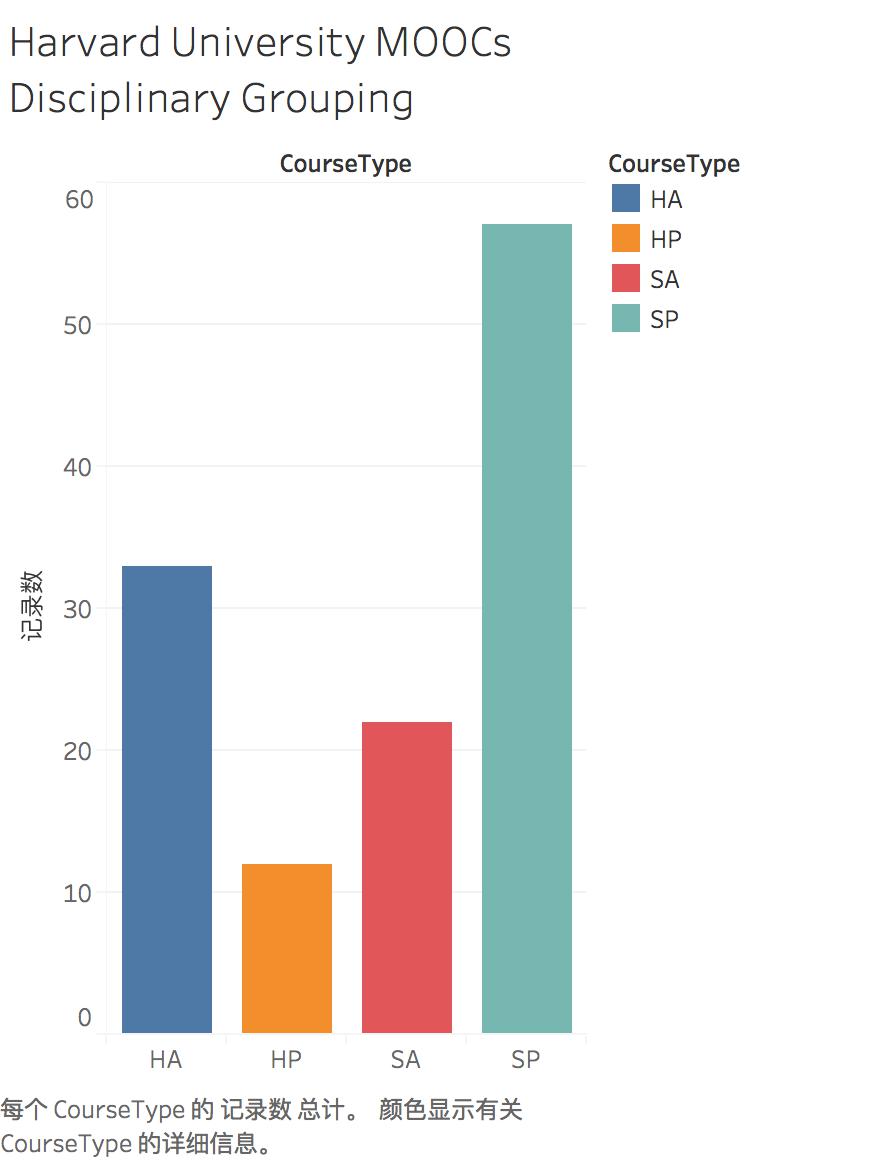
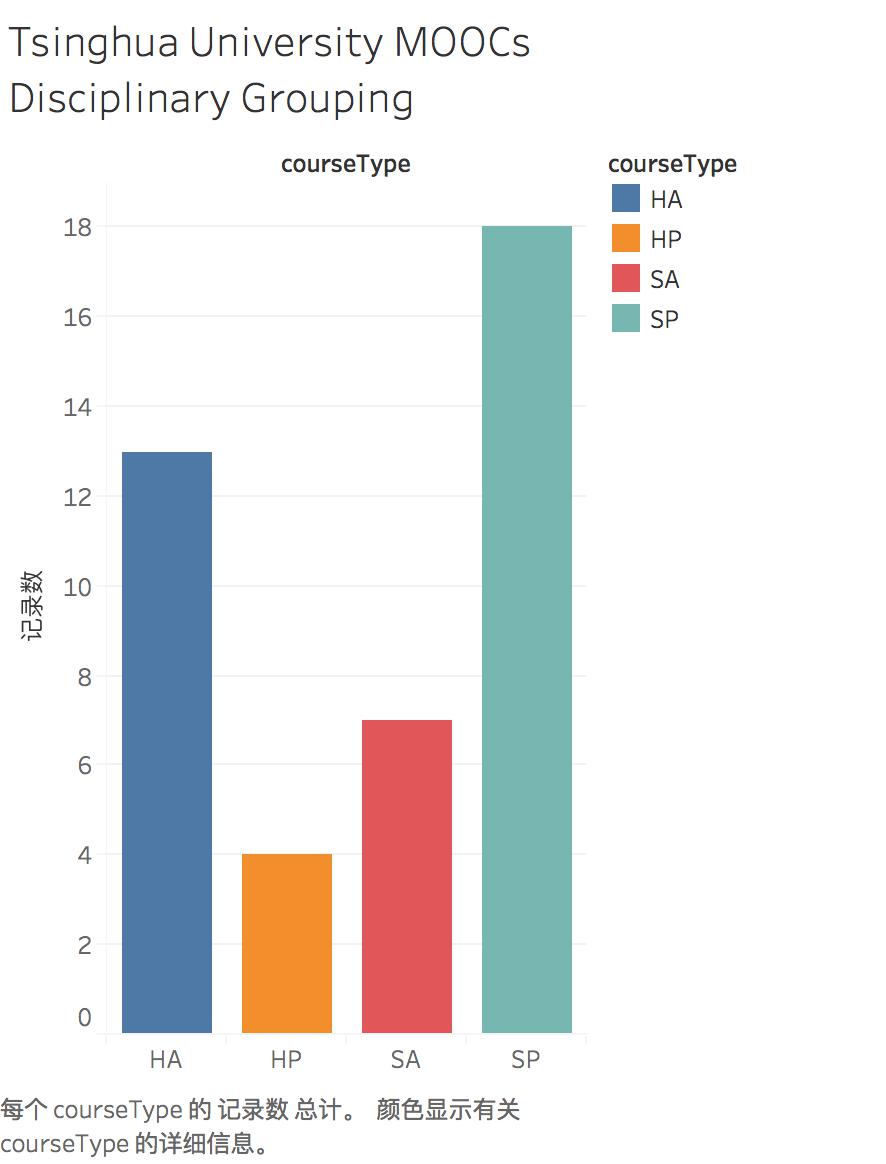
Since we do not have all the courses data right now, it is a compromise to carry on case study. Selecting four typical univerties (Harvard University, Tsinghua University, Melbourne University and Indian Institute of Technology, Kharagpur) from north and south sides, both research and non-research. Encoding all courses of the four institutes using Tony Becher’s disciplinary grouping theory.

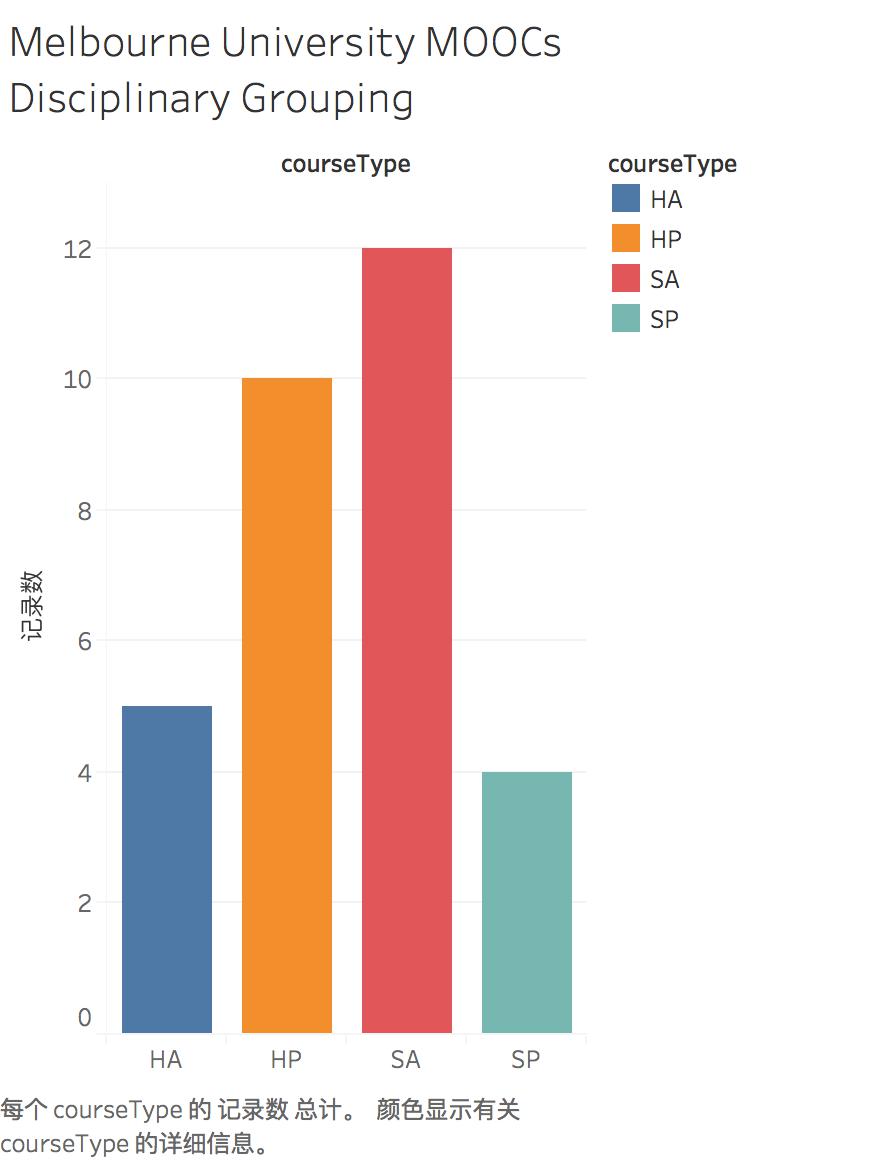
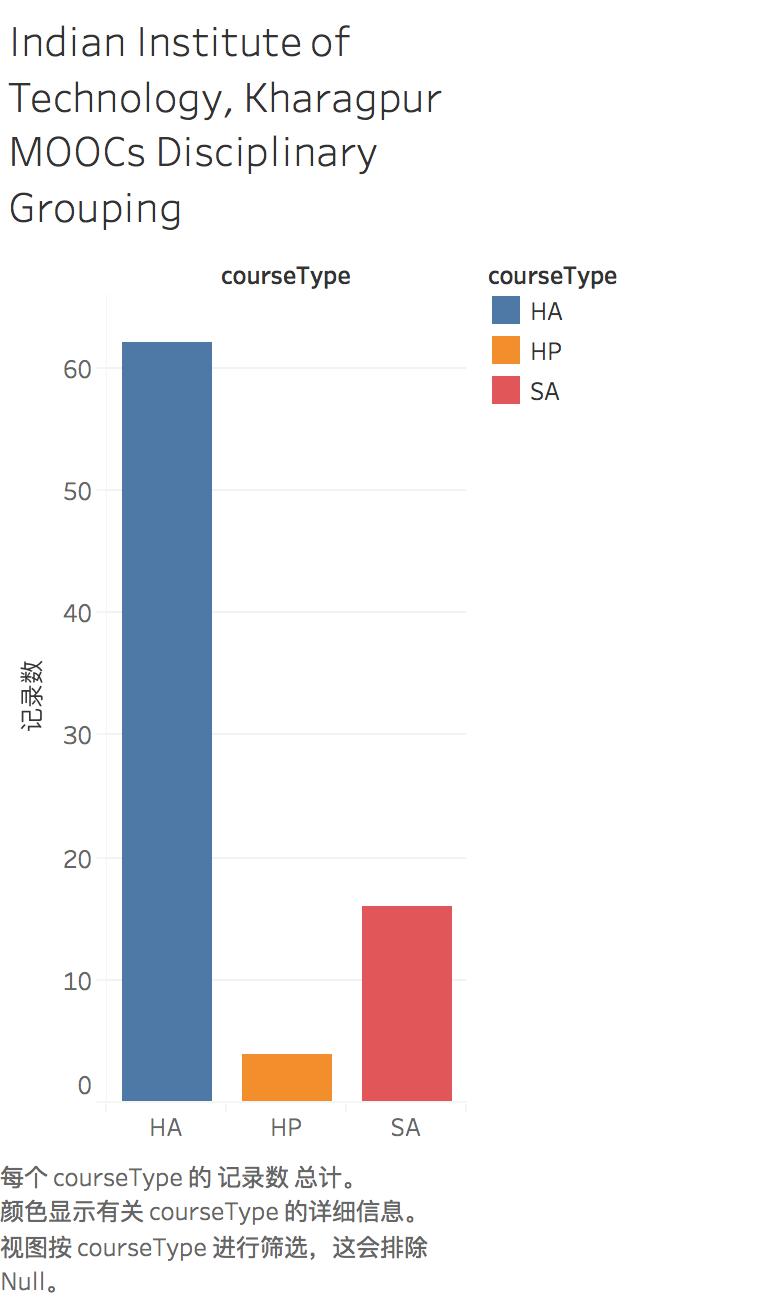
It is easy to find out that there are quite different discipline patterns among these four institutes. South universities more likely to offer hard applied courses comparing to other disciplines. North universities seem to have a higher percentage in soft disciplines. Besides, MOOCs disciplinary patterns between Harvard University and Tsinghua University are surprisely similar, even though Tsinghua is regarded as a university of science and techonology, while Harvard is a comprehensive university, which indicates that universities in different type of countries may have very similar disciplinary patterns.

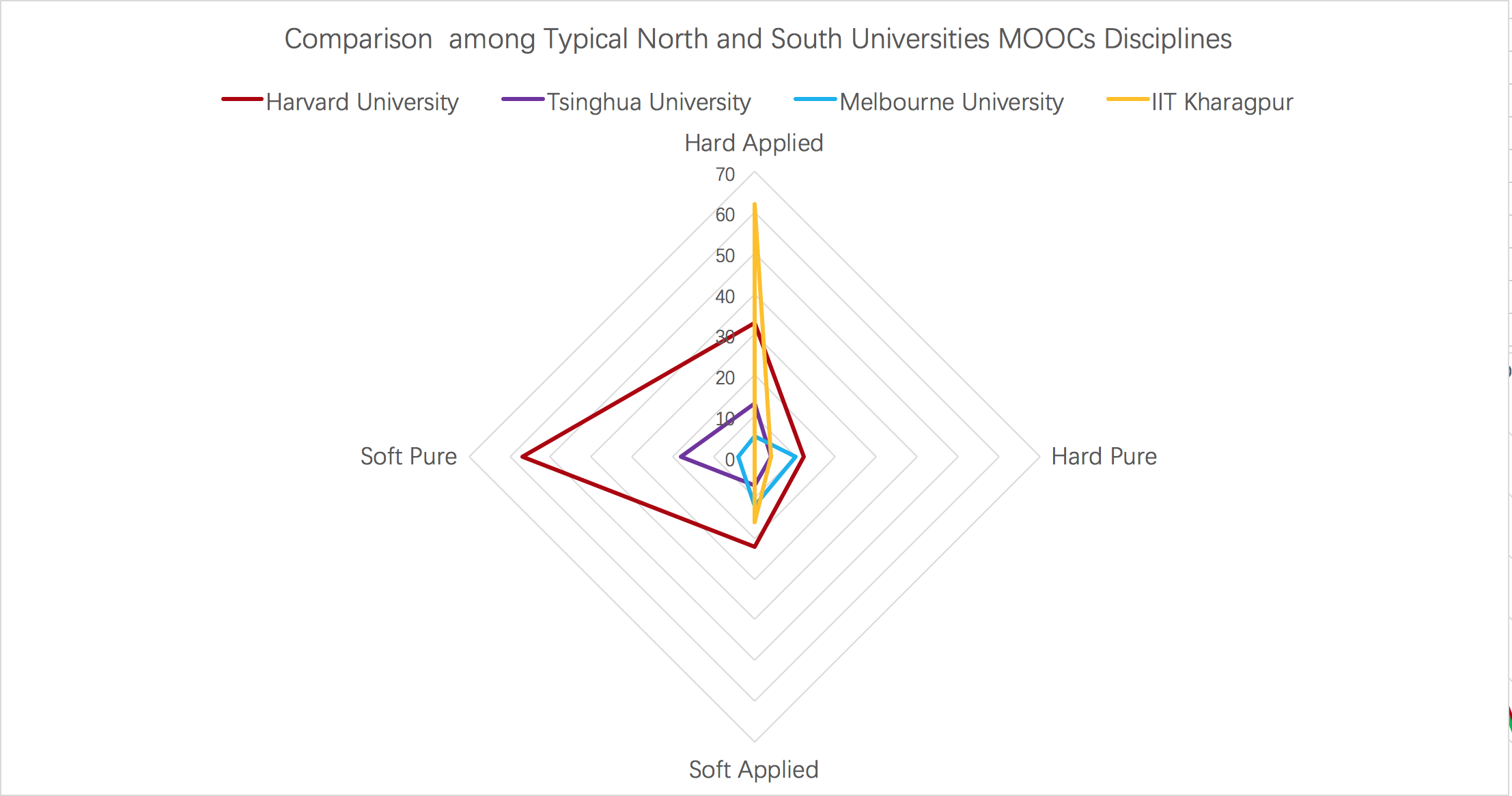
It should be noted there might be bias since what we discussed above is based on a small sample.

Tab 5. Comparison among typical North and South Universities MOOCs’ Disciplines

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| University | H.A. | H.P. | S.A. | S.P. | Courses Number | Country Type | Institue Type |
| Harvard University | 33 | 12 | 22 | 57 | 124 | North | Research |
| Tsinghua University | 13 | 4 | 7 | 18 | 42 | South | Research |
| Melbourne University | 5 | 10 | 12 | 4 | 31 | North | Research |
| Indian Institute of Technology, Kharagpur | 62 | 4 | 16 | 0 | 82 | South | Non-research |
| World | 3115 | 1165 | 2983 | 174 | 7437 | -- | -- |



Pic 10. Comparison among typical North and South Universities MOOCs’ Disciplines

**Conclusion and Discussion**

This study aims to sketch a panoramic map of global MOOCs, and concentrate on the differences of the country type, institute type and discipline between the southern countries and northern countries. It is a unique way to show the gap of knowledge production between the South and North.

This study tries to classify countries and universities into specific division by using machine learning technology, represent the differences by data visualization techniques, and to update the method of discipline classification by adopting data mining technology. The methods of this study are relatively new to represent and analyze this global education issue.

**Main findings**

There is a huge gap between South and North countries in terms of MOOCs, not only the courses number, but also the quality of the MOOCs which the institutes offering the courses can indicate, that the Global South countries have obvious disadvantages comparing to Global North countries.

Countries in the same division of South and North have different MOOCs supplying patterns due to the diverse economic and educational conditions. Therefore, some North countries who are not so promising in economy, for instance, Russia, may act like a South country when it offers online courses. However, China, the largest Global South country and its economy also has been regarded as the greatest miracle last four decades is more likely a North country when it comes to MOOCs.

The MOOCs’ disciplines are quite different between the South and North. From limited case study, it can tell that South countries show greater preference in Hard Applied courses such as Engineering, et cetera, than any other types of courses. North countries are interested in “Hard Applied” courses too, meanwhile, they also publish lots of “Soft Pure”, “Soft Applied” and “Hard Pure” courses, which are considered to be the core and foundation of knowledge production. An easy answer for the Engineering Hot in South countries is the demand of market, since most of the developing countries have an enormous infrastructure construction market, which means a huge demand of engineering talents. Consequently, institutes in those countries have to offer larger portion of Hard Applied courses, and naturally there are more “Hard Applied” courses online. On the contrary, North countries may do not have as huge as the demand in engineering like developing countries. At the same time, North countries are in the upstream of the global market, and they concentrate more in management, research and development, which can explain the MOOCs disciplinary patterns of these countries.

**Limitation and outlook**

Data problem. Courses data in Central-Class are only small part of the MOOCs worldwide, and there is small sample bias inevitably.

Deep comprehensive analysis on courses needed. This paper did only small number of case study about course list of North countries and South countries. To make points be more concrete, deeper comprehensive research is needed, for instance, pay more attention to courses name, language, enrolled students number and so on.