

'Converging Paths in Divergent Systems: A Comparative Analysis of Data Science Education Strategies in China and the United States

Duo Li

Department of Big Data Management
and Application
Shenyang Institute of Technology
Fushun, China
liduo@situ.edu.cn

Elizabeth Milonas

Computer Systems Technology
New York City College of Technology
(CUNY)
Brooklyn, USA
emilonas@citytech.cuny.edu

Qiping Zhang

Palmer School of Library and
Information Science
Long Island University
Brookville, USA
Qiping.Zhang@liu.edu

Abstract - This research-to-practice full paper compares data science education strategies in China and the United States, exploring whether different approaches can achieve similar educational outcomes. In the U.S., data science programs are typically developed by individual schools to meet labor market demands, integrating data science as an interdisciplinary subject. In contrast, Chinese universities follow a uniform plan set by the education department, resulting in specialized fields like 'Big Data Management and Application' and 'Data Science and Big Data Technology.' This research collected data on course offerings and university rankings in China and the U.S., analyzing curriculum content and program characteristics. Major findings indicate that in the U.S., Data Science and Business Analytics programs focus on technical skills and are primarily found in top-ranking universities, while in China, Data Science and Big Data Management programs emphasize management and are more widely distributed across different ranking groups. Despite these differences, both countries show similar geographic concentration patterns in data science programs. The study concludes that China and the U.S. adopt different educational strategies but achieve comparable effectiveness in data science education.

Keywords—Data Science Education, Comparative Analysis, Curriculum Development, University Rankings, Big Data Management

I. INTRODUCTION

Both China and the United States are developing specialties in data science, but they are following different paths. According to a research report from 2023, the development of data science in the United States is primarily led by individual schools, which tailor their programs based on the demands of the labor market for positions, or by integrating data science as an interdisciplinary subject with the schools' existing strengths [1]. In contrast, Chinese universities' specialties in data science or big data-related fields are uniformly planned by the education department. Currently, there are specialties like 'Big Data Management and Application' developing in the field of management, and 'Data Science and Big Data Technology' in engineering and science [2][3]. This setup has led to confusion in the job market between 'Big Data Management and Application' and 'Data Science and Big Data Technology', resulting in a basic requirement for big data professionals to understand programming languages, data analysis capabilities, and system development skills [4][5]. This has also created a gap

and disconnect between the setup of specialties and the job market.

II. LITERATURE REVIEW

In the article by Zhang et al. (2022) [6], the initial components of data science are discussed, highlighting how the field has expanded from data-centered disciplines and domains, such as statistics, computing, and informatics, to the social sciences and management [2]. The United States has been at the forefront of data science exploration due to its semiconductor and information technology industries. As a downstream player in these industries, China has also increasingly focused on data science in its universities. However, due to the structured nature of the Chinese higher education system, the overall professional layout and guidance are still set uniformly by the education department. Currently, there is significant uncertainty about which disciplines data science should belong to. As a result, China has established two closely related fields: Data Science and Big Data Technology, and Big Data Management and Application, reflecting the interdisciplinary nature of data science.

Data science (DS) is recognized as multi- and interdisciplinary in nature [2][7]. Many disciplines and academic units have begun to reflect on, discuss, and establish specific disciplinary identities within the DS research and education landscape. To meet the increasing demand for the DS workforce, various disciplines have developed and offered data-related courses and programs. A recent review of graduate-level DS education programs shows that Mathematics and Statistics, Computer Science, Business, and Library and Information Science (LIS) are the leading disciplines offering such programs [8]. Increasingly, business research institutions and business schools are introducing data science education in the fields of business and economics [9]. To further tailor data science education to different types of universities, a comparative analysis of the development of this field in China and the U.S. can help create a more systematic framework that can be used as a reference and applied by schools at various levels.

III. METHOD

This study primarily collected data on the number of course offerings in undergraduate Big Data majors at Chinese

universities and their rankings. It examines whether China and the U.S. approach the establishment and structuring of Data Science programs from different angles yet achieve the same effect, by comparing the relationship between university rankings and the offering of Data Science in both countries. In other words, it explores whether they 'reach the same destination by different paths'.

A. Research Question

RQ1: How do the approaches to developing Data Science programs in China and the U.S. differ?

RQ2: Does the correlation between university rankings and the availability of Data Science and Big Data courses in China and the U.S. indicate a convergence in educational effectiveness despite differing educational strategies?

B. Data Sources

1) Chinese Data Sources

In China, the Ministry of Education currently approves two undergraduate majors related to data science: 1) "Big Data Management and Application" (major code 120108T), primarily offered within management disciplines and granting a management degree. 2) "Data Science and Big Data Technology" (major code 080910T), mainly granting engineering or science degrees.

Ranking data for these majors can be found on the Shanghai Ranking Consultancy's website:

- Big Data Management and Application [10]
- Data Science and Big Data Technology [11]

2) U.S. Data Sources

The screening criteria focuses on academic programs that provide degrees in Data Science. The 171 college programs utilized in this study are derived from an earlier study [12, 13] in which the following websites [14-17] were crawled and college program data was recorded.

College ranking data used in this study was derived from the U.S. News college ranking data available in the U.S. News website [18].

C. Datasets

1) Chinese Data

Table 1 and 2 list the frequencies of all data science programs in the two above majors in China for the years 2021-2023.

Given relatively small sample sizes, all program data in China were collected for our study.

TABLE 1: DATA SCIENCE PROGRAMS IN BIG DATA MANAGEMENT AND APPLICATION IN CHINA (2021-2023)

| Ranking | Big Data Management & Applications | | |
|--------------|------------------------------------|------------|------------|
| | 2021 | 2022 | 2023 |
| A+ | 12 | 14 | 14 |
| A | 53 | 56 | 60 |
| B+ | 57 | 72 | 146 |
| B | 187 | 210 | 150 |
| Total | 309 | 352 | 370 |

TABLE 2: DATA SCIENCE PROGRAM IN DATA SCIENCE AND BIG DATA TECHNOLOGY IN CHINA (2021-2023)

| Ranking | Data Science & Big Data Technology | | |
|--------------|------------------------------------|-----------|------------|
| | 2021 | 2022 | 2023 |
| A+ | 2 | 3 | 4 |
| A | 6 | 16 | 19 |
| B+ | 8 | 19 | 43 |
| B | 24 | 53 | 44 |
| Total | 40 | 91 | 110 |

2) U.S. Data

A total of 171 undergraduate Data Science degree programs were initially crawled, but only 102 colleges were chosen for this study. Sixty-nine undergraduate programs were eliminated from the analysis due to the following 1) no Data Science courses were evident in the list of core requirements for the program, 2) only a graduate degree was offered, 3) only an associate degree was offered, 4) only a minor in Data Science was offered [12].

These 102 colleges are divided into two ranking categories according to U.S. News & World Report. 1) national ranking, 2) regional ranking. Each category was further divided into sub-groups that were roughly equal in size, resulting in 5 rank groups for national rankings and 3 rank groups for regional rankings (Table 3).

TABLE 3: U.S. NEWS RANKING CATEGORIES, THE NUMBER OF COLLEGES IN EACH CATEGORY, RANKING GROUPS AND FREQUENCIES FOR EACH CATEGORY (2021)

| Ranking Category | U.S. News & World Report | | |
|------------------|--------------------------|----|-------|
| | Ranking Group | N | Total |
| National | NG1: 1-75 | 24 | 66 |
| | NG2: 76-150 | 20 | |
| | NG3: 151-225 | 12 | |
| | NG4: 226-297 | 5 | |
| | NG5: 298-389 | 5 | |
| Regional | RG1: 1-16 | 13 | 36 |
| | RG2: 17-75 | 13 | |
| | RG3: 76-150 | 10 | |
| Total | | | 102 |

The degree name and course offerings for each data science program were reviewed as presented on the institution's and department's websites. Program competencies were evaluated based on the ACM competencies [19] and the competency clusters identified in our previous study [12].

Due to the unavailability of U.S. News data for 2022 and 2023, our U.S. dataset is limited to the year 2021.

D. Comparative Data Analysis

The comparative analysis aims to evaluate the effectiveness of Data Science education strategies between China and the U.S. by assessing how different educational approaches contribute to student outcomes and preparedness for the job market.

This involves three types of analyses: program degree majors, program geographic distribution, Key metrics for comparison include graduate employability, research output, and alignment with industry standards. This analysis will

determine if diverse educational frameworks in China and the U.S. converge towards similar quality outcomes or if distinct regional adaptations offer unique advantages.

1) Analysis of Program Degree Distribution

First, we collected ranking data for all undergraduate data science programs both in China and U.S.. Next, the tier distribution of program frequency by different degrees was calculated.

In China, only two degrees are offered for undergraduate data science program: Data Science and Big Data Technology, Big Data Management and Application.

In the U.S., programs with B.S. degree in Data Science were included for our analyses. Different types of degree majors include Data Science, Data Analytics, Big Data, Information Science and Technology (IST), Computer Science, and Math and Statistics.

2) Analysis of Program Geographic Distribution

Fig. 1 is a regional map of China, divided into seven regions: Central China (blue), East China (orange), North China (red), Northeast (teal), Northwest (green), South China (yellow), and Southwest (purple), as referenced in [20].

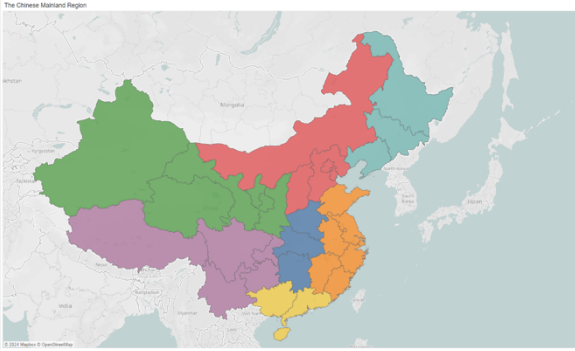


Fig. 1. Regional Map for Chinese Data

Fig. 2 is a regional map of the U.S., showing five regions according to the National Geographic Society classification [21]: Midwest (red), Northeast (teal), Southeast (blue), Southwest (purple), and West (green).

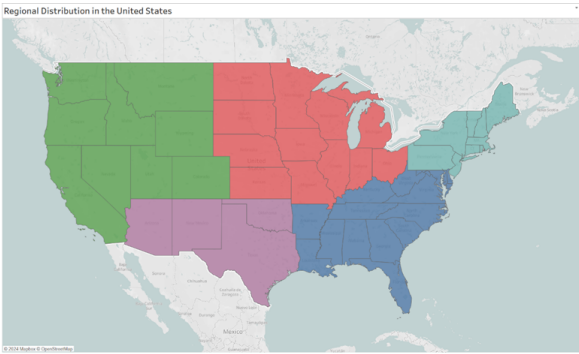


Fig. 2. Regional Map for U.S. Data

The tier distribution of program frequency by regions was analyzed for both China and U.S. data.

3) Analysis of Program Ranking & Competency

For Chinese data, a trend analysis of program ranking data was conducted for the years 2021 to 2023.

For the U.S. data, the program competencies were reported by ranking tiers; however, trend analysis could not be performed due to the absence of data for 2022 and 2023.

IV. RESULTS

A. Chinese Results

1) Results of Program Degree Distribution

Table 4 shows that the Chinese data spans three years, from 2021 to 2023, and can be filtered by year using the website's interface. By accessing the 2021, 2022, and 2023 data from Shanghai Ranking, we obtained ranking information for both majors over these three years. Six data tables correspond to "Data Science and Big Data Technology" and "Big Data Management and Application" for each year from 2021 to 2023. The column structure of each table includes Ranking (Major), Name, University, Grade, 211 Project, 985 Project, Province, Province (English), City, Region, and Region (English). Below is a summary of some statistical findings from this data.

TABLE 4: 2021-2023 STATISTICS OUTPUT OF MAJOR RANKING FROM CHINA UNIVERSITIES

| | 2021 | | 2022 | | 2023 | |
|----------------------------|--------------------------------------|-------------------------------------|--------------------------------------|-------------------------------------|--------------------------------------|-------------------------------------|
| | Data Science and Big Data Technology | Big Data Management and Application | Data Science and Big Data Technology | Big Data Management and Application | Data Science and Big Data Technology | Big Data Management and Application |
| Ranking Statistics: | | | | | | |
| Average R | 154.23 | 20.5 | 175.73 | 45.84 | 184.69 | 55.29 |
| Std Dev. | 89.02 | 11.69 | 101.31 | 26.41 | 106.56 | 31.87 |
| Highest R | 1 | 140 | 1 | 1 | 1 | 1 |
| Lowest R | 305 | | 351 | 90 | 367 | 110 |
| Tier Distribution: | | | | | | |
| A+ | 12 | 2 | 14 | 3 | 14 | 4 |
| A | 53 | 6 | 56 | 16 | 60 | 19 |
| B+ | 57 | 8 | 72 | 19 | 146 | 43 |
| B | 87 | 24 | 210 | 53 | 150 | 44 |

The annual data tables indicate that the number of higher-tier universities is relatively small, with B-level universities being the most numerous. While there are fluctuations in rankings across different years, universities in A+ and A tiers generally occupy higher ranking positions. The number of universities offering big data-related majors has increased, particularly in the B+ and B tiers.

The annual data tables reveal a relatively small number of higher-tier universities, with the majority being B-level institutions. While rankings fluctuate somewhat from year to year, universities in A+ and A tier generally maintain higher positions. The number of universities offering big data-related majors has increased, particularly in the B+ and B tiers.

We've organized the data by year and major, categorizing universities by tiers. Fig. 3 showed the key trends.

Data Science and Big Data Technology Major

From 2021 to 2023, tier A+ increased from 12 to 14; tier A increased from 53 to 60; tier B+ increased from 57 to 146; tier B decreased from 187 to 150.

Big Data Management and Application Major

From 2021 to 2023, tier A+ increased from 2 to 4; tier A increased from 6 to 19; tier B+ increased from 8 to 43; tier B increased from 24 to 44.



Fig. 3. Trends of University Grades in Data Science domain

Both majors saw an increase in the number of A+ and A tier universities from 2021 to 2023, indicating greater emphasis and development of these majors in top universities.

Despite the increase in B+ and A tier universities, the number of B tier universities offering Data Science and Big Data Technology decreased, suggesting some universities have moved up to higher tiers.

The growth may be related to policy support, regional economic development, and the allocation of higher education resources, particularly in areas with significant investment in the 211 and 985 projects.

2) Results of Program Geographic Distribution

Fig. 4 and Fig. 5 showed the geographic distribution of two data science programs in China from 2021 to 2023. East China and North China were top regions for both data science programs.

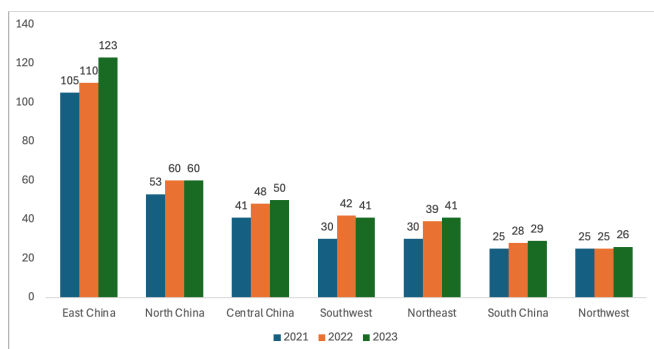


Fig. 4. Geographic Distribution of Data Science and Big Data Technology Programs (2021-2023)

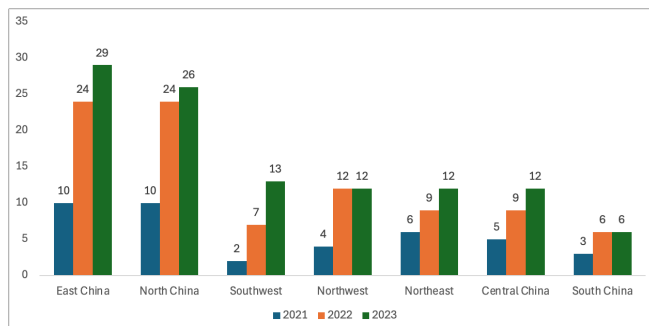


Fig. 5. Geographic Distribution of Big Data Management and Application Programs (2021-2023)

3) Results of Program Ranking Analyses

Correlation analyses were conducted to further explore trends and relationships in the data.

- whether there was a statistically significant correlation between university tier (e.g., A+, A, B+) and its ranking.
- whether the number of universities in different tiers shows a significant increase or decrease over time.

Table 5 revealed that only the correlation between 2021 and 2022 was positively significant ($p < .001$), indicating a very consistent change in the number of universities in these two years.

TABLE 5: THE CORRELATION MATRIX BETWEEN YEARS

| Year | Correlation Matrix | | |
|-----------|-------------------------|-------------------------|---|
| | Correlation Coefficient | Significance (2-tailed) | n |
| 2021-2022 | 0.99 | <0.001 | 3 |
| 2021-2023 | 0.73 | 0.10 | 3 |
| 2022-2023 | 0.77 | 0.09 | 3 |

Table 6 showed the result of Linear Regression analysis of program counts by year for Tier A+ universities. The regression equation is: Program Count = $2.0 \times \text{Year} - 4038$. However, the regression result was not significant ($p > .05$). Although the descriptive data showed an increasing trend in the number of Tier 'A+' universities each year, this trend was not statistically significant.

TABLE 6: HYPOTHESIS TESTING FOR LINEAR TREND OUTPUT

| Variable | Hypothesis Testing | | | | |
|----------------|--------------------|------------|---------------------------------|---------|---------|
| | Coefficient (B) | Std. Error | Beta (Standardized Coefficient) | T value | p value |
| Intercept | -4038 | 965.7 | N/A | -4.18 | 0.08 |
| Year (2021-23) | 2.0 | 0.93 | 0.96 | 2.15 | 0.18 |

B. U.S. Results

1) Results of Program Degree Distribution

Table 7 & 8 display the majors offered by nationally and regionally ranked colleges that have Data Science degree programs.

For national ranking programs, data science was a common major offered by all five ranking groups. However, the top two groups (NG1 and NG2) consistently offered all five data science majors. In contrast, the lower ranking groups (NG3 and NG4) primarily offered the data science major with the NG5 ranking groups also offering data analytics and IST majors.

TABLE 7: NATIONAL RANKINGS AND MAJORS (2021)

| Major Clusters | National Ranking | | | | | |
|---------------------|------------------|-----------|-----------|----------|----------|-----------|
| | NG1 | NG2 | NG3 | NG4 | NG5 | Total N |
| Data Science | 25% | 80% | 83% | 100% | 40% | 49 |
| Math and Statistics | 13% | 5% | 0% | 0% | 0% | 4 |
| Data Analytics | 8% | 5% | 8% | 0% | 40% | 6 |
| IST | 8% | 5% | 8% | 0% | 20% | 5 |
| Computer Science | 4% | 5% | 0% | 0% | 0% | 2 |
| Total | 24 | 20 | 12 | 5 | 5 | 66 |

For regionally ranked colleges, regardless of their ranking, the Data Science degree program is mainly housed in the Data Science and Data Analytics majors (Table 8). In addition, the lowest regionally ranked colleges (RG3), have a relatively high (30%) concentration of Data Science degree programs in their Math and Statistics major as well.

TABLE 8: REGIONAL RANKINGS AND MAJORS (2021)

| Major Clusters | Regional Ranking | | | |
|---------------------|------------------|-----------|-----------|-----------|
| | RG1 | RG2 | RG3 | Total N |
| Data Science | 85% | 77% | 40% | 25 |
| Data Analytics | 8% | 15% | 20% | 5 |
| Math and Statistics | 8% | 0% | 30% | 4 |
| Big Data | 0% | 8% | 10% | 2 |
| Total | 13 | 13 | 10 | 36 |

2) Results of Program Geographic Distribution

Table 9 & 10 display the regions in which nationally ranked (Table 9) and regionally ranked (Table 10) colleges offering Data Science degree programs are located.

On both the national and regional levels, the Northeast and Midwest regions have the largest number of data science programs across all ranking groups (NG1-NG5, RG1-RG3). The West region has more higher-ranked program (NG1-NG3, RG1-RG2), while the Southeast and Southwest regions primarily have lower-ranked programs (NG4-NG5, RG3).

TABLE 9: NATIONAL RANKINGS AND REGIONS (2021)

| Region | National Ranking | | | | | |
|--------------|------------------|-----------|-----------|----------|----------|-----------|
| | NG1 | NG2 | NG3 | NG4 | NG5 | Total N |
| NE | 38% | 25% | 42% | 20% | 40% | 22 |
| MW | 33% | 45% | 17% | 40% | 40% | 23 |
| W | 17% | 15% | 25% | 0% | 0% | 10 |
| SE | 13% | 10% | 8% | 20% | 0% | 7 |
| SW | 0% | 5% | 8% | 20% | 20% | 4 |
| Total | 24 | 20 | 12 | 5 | 5 | 66 |

TABLE 10: REGIONAL RANKINGS AND REGIONS (2021)

| Region | Regional Ranking | | | |
|--------------|------------------|-----------|-----------|-----------|
| | RG1 | RG2 | RG3 | Total N |
| MW | 50% | 17% | 30% | 11 |
| NE | 33% | 67% | 40% | 16 |
| W | 17% | 25% | 0% | 5 |
| SE | 8% | 0% | 10% | 2 |
| SW | 0% | 0% | 20% | 2 |
| Total | 13 | 13 | 10 | 36 |

3) Results of Program Ranking & Competency Analyses

Tables 11 and 12 present the ranking groups of data science programs and their corresponding competency frequencies for both national and regional rankings.

It is noteworthy that all ACM-recommended Data Science competencies are available at every college analyzed in this study. The top three competencies (Math and Statistics, Computing Fundamentals, and Data Management/Governance/Privacy) are consistently offered, regardless of whether the colleges are ranked nationally or regionally.

Nationally Ranking Result (Table 11)

Two competencies worth a closer examination in nationally ranked programs.

- *Math and Statistics*: This competency is emphasized across all ranking groups, with 37% of NG1 programs and a consistent presence in lower ranking groups. This highlights the foundational role of mathematical and statistical skills in Data Science education across top U.S. universities.
- *Machine Learning*: Particularly prominent in NG1 universities, where it is included in 50% of programs. This shows that the highest-ranked institutions are more likely to incorporate advanced and cutting-edge topics like Machine Learning into their curricula.

TABLE 11: U.S. DATA SCIENCE PROGRAM: NATIONAL RANKING AND PROGRAM COMPETENCIES (2021)

| Competencies | National Ranking | | | | | |
|-------------------------------------|------------------|-----------|-----------|-----------|-----------|------------|
| | NG1 | NG2 | NG3 | NG4 | NG5 | Total N |
| Math & Stats | 37% | 31% | 17% | 8% | 8% | 65 |
| Computing Fundamentals | 38% | 29% | 18% | 6% | 8% | 61 |
| Data Management/Governance/ Privacy | 38% | 28% | 16% | 8% | 10% | 50 |
| Data Mining/ Big Data | 38% | 25% | 19% | 6% | 11% | 47 |
| Data Science in Context | 36% | 32% | 18% | 7% | 7% | 44 |
| Machine Learning | 50% | 24% | 19% | 2% | 5% | 42 |
| Data Visualization | 33% | 30% | 22% | 8% | 5% | 36 |
| Total | 133 | 99 | 63 | 23 | 27 | 345 |

Regional Ranking Results (Table 12)

Two competencies worth a closer examination in regionally ranked programs.

- **Math and Statistics:** Again, this competency stands out, with 36% of programs across all regional ranking groups emphasizing it, reinforcing its fundamental importance in the curriculum of Data Science programs, regardless of the institution's ranking.
- **Data Visualization:** This competency is more strongly emphasized in RG2 programs (48%) compared to RG1 and RG3, indicating that certain regional institutions may place a particular focus on visual data representation, which is crucial for applied Data Science skills.

TABLE12: U.S DATA SCIENCE PROGRAM - REGIONAL RANKING AND PROGRAM COMPETENCIES (2021)

| Competencies | Regional Ranking | | | |
|-------------------------------------|------------------|-----------|-----------|-----------|
| | RG1 | RG2 | RG3 | Total N |
| Math & Stats | 36% | 36% | 28% | 36 |
| Computing Fundamentals | 36% | 33% | 30% | 33 |
| Data Management /Governance/Privacy | 28% | 34% | 27% | 29 |
| Data Science in Context | 36% | 39% | 25% | 28 |
| Data Visualization | 33% | 48% | 19% | 21 |
| Machine Learning | 21% | 43% | 28% | 21 |
| Data Mining/Big Data | 33% | 28% | 39% | 18 |
| Total | 13 | 13 | 10 | 36 |

V. DISCUSSIONS AND CONCLUSION

A. Key Findings

- **Popular Majors and Focus Areas:**

U.S.: The Data Science and Data Analytics majors are highly favored, with an emphasis primarily on cultivating advanced technical skills that encompass statistical analysis, machine learning, predictive modeling, data privacy and governance.

China: The popular majors are Data Science and Big Data Management, where the focus is more on management aspects intertwined with a strong technical foundation, preparing graduates to handle large-scale data operations and decision-making processes in business contexts.

- **Geographic Distribution of Programs:**

China and the U.S.: Both countries exhibit a similar pattern in the geographic distribution of data science programs. There is a notable concentration of these programs in the Northern and Eastern regions, which are major economic and educational hubs in both countries. This distribution contrasts with the relatively fewer programs located in the Southern and Western parts, indicating regional disparities in educational opportunities and industrial demand.

- **Ranking and Resource Allocation:**

U.S.: Data science programs are predominantly situated in universities ranked within the top 100 by U.S. News & World Report, indicating a significant allocation of resources and a high level of institutional support for these disciplines. This concentration in top-tier universities underscores the strategic importance placed on developing cutting-edge analytical capabilities and research output.

China: In contrast, the distribution of data science programs spans a wider range of university rankings. While top-tier universities (often part of the 985 and 211 projects) do offer robust data science programs, many mid-tier institutions also offer these majors, demonstrating a broad governmental push to enhance data literacy across varied educational levels.

B. Conclusion

Converging trends are evident in the types of data science degrees offered and their geographic concentration in key economic hubs, despite the divergent national strategies employed by China and the U.S. This suggests a shared understanding of the economic importance of data science education.

Despite these commonalities, there are stark differences in how curriculum is designed, how programs align with industry, and how resources are distributed. The U.S. model favors concentrating top-tier data science programs in elite universities, bolstering their global standing and capacity for innovation. Conversely, China's approach, with a more even distribution across university rankings, suggests a national focus on making data science education accessible to a wider range of students and preparing a larger workforce for this field.

To fully assess the effectiveness of these differing approaches, future studies should analyze the competencies developed by these programs. Such analyses should examine not only the technical and analytical skills imparted but also how these programs foster innovation, adaptability, and interdisciplinary integration. By comparing the outcomes of these educational strategies, researchers can better understand how different educational models influence the preparedness of graduates to meet the challenges of the global data economy.

C. Future Research:

Although this study primarily compared data science education strategies in China and the U.S., it established a framework to assess data science program competencies that can be used for future analysis of data science programs in other countries. Additionally, we plan to further obtain specific data science curriculum content from various universities in both the U.S. and China to provide more insights on aligning higher education data science programs with job market demands.

REFERENCES

- [1] Zhang, Y., Wu, D., Hagen, L., Song, I. Y., Mostafa, J., Oh, S., ... & Saltz, J. S. (2023). "Data science curriculum in the iField." *Journal of the Association for Information Science and Technology*, 74(6), 641-662.
- [2] Cao, L. (2017). "Data science: a comprehensive overview." *ACM Computing Surveys (CSUR)*, 50(3), 1-42.
- [3] Jahani, H., Jain, R., & Ivanov, D. (2023). "Data science and big data analytics: A systematic review of methodologies used in the supply chain and logistics research." *Annals of Operations Research*, 1-58.
- [4] Provost, F., & Fawcett, T. (2013). "Data science and its relationship to big data and data-driven decision making." *Big data*, 1(1), 51-59.
- [5] Li, Duo, Elizabeth Milonas, and Qiping Zhang. "Content Analysis of Data Science Graduate Programs in the US." In 2021 ASEE Virtual Annual Conference Content Access. 2021.
- [6] Zhang, Y., Wu, D., Hagen, L., Song, I. Y., Mostafa, J., Oh, S., ... & Saltz, J. S. (2023). "Data science curriculum in the iField." *Journal of the Association for Information Science and Technology*, 74(6), 641-662.

- [7] Raban, D. R., & Gordon, A. (2020). "The evolution of data science and big data research: A bibliometric analysis." *Scientometrics*, 122(3), 1563-1581.
- [8] Wu, D. (2019). "Research on data science curriculum [Conference presentation]." In Association for Information Science & Technology Annual Meeting, Melbourne, Australia.
- [9] Miah, S. J., Solomonides, I., & Gammack, J. G. (2020). "A design-based research approach for developing data-focussed business curricula." *Education and Information Technologies*, 25(1), 553-581.
- [10] "2021 中国大学专业排名：大数据管理与应用" <https://www.shanghairanking.cn/rankings/bcmr/2021/120108T> (accessed on Oct. 1, 2023)
- [11] "2021 中国大学专业排名：数据科学与大数据技术" <https://www.shanghairanking.cn/rankings/bcmr/2021/080910T> (accessed on Oct. 1, 2023)
- [12] Milonas, Elizabeth, Duo Li, Qiping Zhang (2021). "Content Analysis of Two-year & Four-year Data Science Programs in the United States." In proceedings of the 128th ASEE Annual Conference, July 26-29, 2021.
- [13] Milonas, E., & Zhang, Q., & Li, D. (2022). "Do Undergraduate Data Science Program Competencies Vary by College Rankings?" In Proceedings of the 2022 ASEE Annual Conference & Exposition, Minneapolis, MN. 10.18260/1-2--41153).
- [14] "60 Bachelor Degrees in Data Science" <https://www.mastersindatascience.org/specialties/bachelor-degrees-in-data-science/> (accessed on Feb 1, 2021)
- [15] Swanstrom, R. (2021) Data Science Colleges and Universities <https://ryanswanstrom.com/colleges/> (accessed on Feb 1, 2021)
- [16] "Top 20 most affordable data science bachelor degree programs" <https://www.datasciencedegreeprograms.net/rankings/affordable-bachelors/> (accessed on Feb. 1, 2021)
- [17] Mathur, S. (2019) "The List of All U.S. Colleges With a Data Science Major" <https://blog.collegevine.com/the-list-of-all-u-s-colleges-with-a-data-science-major/> (accessed on Feb. 1, 2021)
- [18] U.S. News "Best Undergraduate Data Science Programs" <https://www.usnews.com/best-colleges/rankings/computer-science/data-analytics-science> (accessed on Feb. 1, 2021)
- [19] A. Danyluk and P. Leidig, "Computing competencies for undergraduate data science curricula," Association of Computing Machinery (ACM), January 2021. [Online]. Available: https://dstf.acm.org/DSTF_Final_Report.pdf
- [20] "中国七大地理分区" [Online]. Available: https://www.douban.com/note/512456305/?_i=4337069pe4FXGP,43372485U7I12g (accessed on Oct. 1, 2023)
- [21] National Geographic Society, "United States Regions," [Online]. Available: <https://media.nationalgeographic.org/assets/file/us-regions-map.pdf> (accessed on Feb. 1, 2021).