An Analysis of ARL Research University Library Funding: The 2017-2018 ARL Annual Statistical Report as a Case Study

Authors: Su Li, Su Li

Date: 2023-06-23T00:00:00+00:00

Abstract

Abstract: Based on data from the Association of Research Libraries (ARL) 2017-2018 Annual Statistics Report, this study conducts correlational statistical analysis on selected data from 116 research university libraries, including total budget, total collection procurement budget, total number of collection volumes/issues, and their ranking positions, together with the corresponding universities' rankings in the 2018 Times Higher Education World University Rankings. Analyses employing Pearson, Spearman, and Kendall correlation coefficients all demonstrate positive correlations with strong relationships. Three recommendations are proposed based on these research findings.

Full Text

Preamble

A Discussion on ARL Research University Library Funding: A Case Study of the ARL 2017-2018 Annual Statistical Report

Su Li
(Nanjing Normal University Library, Nanjing 210023)

Abstract

Based on data from the Association of Research Libraries (ARL) 2017-2018 Statistics, this study conducts a correlation analysis of 116 research university libraries, examining total library expenditures, total collection procurement funding, total volumes/issues/serials expenditures, and their rankings in relation to their universities' positions in the 2018 Times Higher Education (THE) World University Rankings. Using Pearson, Spearman, and Kendall correlation coefficients, all analyses reveal a positive and strong correlation. Based on these findings, three recommendations are proposed.

Keywords: ARL; research university library; Times Higher Education

Introduction

The United States represents one of the most developed higher education systems globally, and North America stands as one of the most advanced regions for higher education worldwide. In the 2023 Times Higher Education World University Rankings, North American institutions occupy 20 of the top 30 positions, with 19 from the United States and one from Canada. The remaining positions are held by the United Kingdom (5), Mainland China (2), Switzerland (1), Germany (1), and Singapore (1). Similar patterns held true from 2017 to 2022, underscoring the leading position of North American higher education.

1.1 ARL Member Libraries

The Association of Research Libraries (ARL), established in 1932, is the largest research library organization in North America, dedicated to advancing research, learning, and scholarly communication while influencing public policy. Since 1969, ARL has published annual statistical reports, with collection-related statistics forming their most critical component. Annual expenditures on collection resources reach $1 billion. According to the 2017-2018 report, ARL comprised 124 research libraries from the United States and Canada as of January 2018, organized as a non-profit entity. Of these, 116 are university libraries, with the remaining eight comprising public, governmental, and non-profit research libraries. By country, there are 16 Canadian and 108 U.S. institutions; among academic libraries specifically, Canada contributes 16 and the United States 100, representing 93.6% of the total membership. This composition largely reflects the current state and development level of research university libraries in North America.

1.2 ARL Research University Libraries

All 116 university libraries within ARL are research libraries, and their parent institutions consistently rank within the top 500 of the THE World University Rankings, qualifying them as research universities. Since university libraries drive institutional advancement, ARL member statistics largely represent the sophisticated educational standards of North American research universities.

What relationship exists between the information embedded in ARL research university library annual statistics and THE World University Rankings? Current domestic research on data analytics in this domain remains limited. Overall, data analytics applications in libraries are still nascent in China, whereas North American library professionals not only recognize their importance but have initiated national-level advocacy and action. Relevant practices at Chinese university libraries have yet to achieve scale, and institutional support remains inadequate.

1.3 Pearson, Spearman, and Kendall Correlation Coefficients and Their Applicability

Pearson, Kendall, and Spearman correlation coefficients all measure relationships and strength between variables. The calculated coefficient determines the nature of correlation: a coefficient of 0 indicates no relationship; a value between 0 and 1 signifies positive correlation (the dependent variable increases as the independent variable increases); a value between -1 and 0 indicates negative correlation (the dependent variable decreases as the independent variable increases). All three coefficients range from -1 to 1, with absolute values indicating correlation strength—larger absolute values denote stronger relationships, with values approaching 1 representing stronger correlations and values near 0 representing weaker ones. Changes in independent variables exert greater influence on dependent variables when correlations are stronger. Correlation strength can be further classified into five ranges: 0.8-1.0 (very strong), 0.6-0.8 (strong), 0.4-0.6 (moderate), 0.2-0.4 (weak), and 0.0-0.2 (very weak or no correlation).

Pearson correlation coefficient applies when: (1) two variables have a linear relationship, (2) both are continuous data, (3) both populations follow normal distributions or approximately normal unimodal distributions, and (4) observations are paired and independent.

Spearman correlation coefficient applies when variables are paired rank-order data or when continuous data has been converted to ranks, regardless of population distribution shape or sample size. For skewed distributions, Spearman provides more authentic results than Pearson.

When data are continuous, normally distributed, and linearly related, Pearson is most appropriate, though Spearman may also be used with lower efficiency. Spearman is also used for two ordinal variables rather than Pearson.

Kendall rank correlation coefficient, also known as Kendall's tau, applies to paired rank-order data or rank-converted continuous variable observations.

2 Correlation Data Analysis

The ARL Statistics 2017-2018 report was published in 2019, with data collected from July 2017 to June 2018. This study analyzes correlations between raw data from the ARL Statistics 2017-2018 report and corresponding universities' 2018 THE World University Rankings. Using Excel and SPSS 27, we examine research university libraries' total expenditures, total collection procurement funding, total volumes/issues/serials expenditures, and top 30 rankings against their universities' THE rankings. Since the raw data on these expenditures exhibit linear relationships, approximate normal unimodal distributions (with some skewed distributions), while THE ranking data represent rank-order or rank-converted continuous data, we employ Pearson, Spearman, and Kendall coefficients for cross-validation.

2.1 ARL Research University Library Total Expenditures (Original: Total Library Expenditures)

Total library expenditures (translated as "Total Library Expenditures") ([1]P51) are critical to sustainable library development. Among 116 ARL research university libraries, institutions ranking in the top 30 for total expenditures consistently place within the top 300 of THE World University Rankings. Accordingly, we analyze correlations between the top 30 ARL research university libraries' expenditure rankings and their 2018 THE rankings.

First, we correlate the top 30 libraries' expenditure rankings with their universities' 2018 THE rankings. The data and analysis results follow:

```
[1]P51t
Mean SD world world rank
Pearson Correlation 1 Sig. (2-tailed) .539**
world Pearson Correlation .539** Sig. (2-tailed) .002
**. Correlation is significant at the 0.01 level (2-tailed).
Sig. (2-tailed) 95% Confidence Interval (2-tailed) a Pearson Correlation rank - world a. Est

Kendall's tau_b rank correlation coefficient .423** world Sig. (2-tailed) . world correlatio
Spearman's Rho rank correlation coefficient .585** Sig. (2-tailed) . world correlation coeff
**. Correlation is significant at the 0.01 level (2-tailed).

Confidence Interval for Kendall's tau_b
Sig. 95% Confidence Interval (2-tailed) a tau_b (2-tailed) rank - world a. Estimation is bas

Confidence Interval for Spearman's rho
Sig. 95% Confidence Interval (2-tailed) a,b Spearman's Rho (2-tailed) rank - world a. Estima
```

Results: Statistically significant correlations exist between the top 30 ARL research university libraries' expenditure rankings and their 2018 THE rankings (Spearman r=0.585, P<0.001; Kendall r=0.423, P=0.001; Pearson r=0.539, P=0.002). (Due to skewed distributions, Spearman results are more authentic.)

Findings: The top 30 ARL research university libraries' expenditure rankings show significant positive correlation with 2018 THE rankings, demonstrating moderate correlation strength. This indicates that greater investment in research university library total expenditures corresponds to higher THE rankings.

Second, for further validation, we analyze the relationship between the top 30 libraries' actual expenditure amounts and THE rankings. During analysis, we transformed the data: THE rankings were reverse-coded using (Max-X)/(Max-Min), while library expenditures were forward-coded using (Y-Min)/(Max-Min), where X represents university ranking data and Y represents expenditure data. Results remained largely consistent:

```
[1]P51 zxhV Pearson Correlation 1 Sig. (2-tailed) .493** nxhW Pearson Correlation .493** Sig
```

```
**. Correlation is significant at the 0.01 level (2-tailed).

Nonparametric Correlations
Kendall's tau_b zxhV correlation coefficient .423** Sig. (2-tailed) . nxhW correlation coeff
Spearman's Rho zxhV correlation coefficient .585** Sig. (2-tailed) . nxhW correlation coeffi
**. Correlation is significant at the 0.01 level (2-tailed).
```

Results: Statistically significant correlations exist between the top 30 libraries' expenditure amounts and THE rankings (Spearman r=0.585, P=0.001; Kendall r=0.423, P=0.001; Pearson r=0.496, P=0.006). (Due to skewed distributions, Spearman results are more authentic.)

Findings: The top 30 ARL research university libraries' expenditure amounts show significant positive correlation with 2018 THE rankings, demonstrating moderate correlation strength. This further confirms that greater investment in research university library total expenditures corresponds to higher THE rankings.

Data on ARL research university library total expenditures, top 30 rankings, and corresponding 2018 THE rankings:

Rank Order Table 6 : Total Library Expenditures

University	Value	2018world
TORONTO	96,957,697	22
MICHIGAN	75,322,139	20
COLUMBIA	71,249,875	14
NEW YORK	70,837,125	27
PRINCETON	67,381,347	7
CALIFORNIA, LOS ANGELES	60,224,261	15
CALIFORNIA, BERKELEY	59,966,729	18
PENNSYLVANIA STATE	58,789,073	96
PENNSYLVANIA	57,471,947	10
TEXAS	56,174,948	49
CORNELL	54,691,065	19
TEXAS A&M	54,231,581	159
ILLINOIS, URBANA	53,326,289	37
SOUTHERN CALIFORNIA	52,646,302	66
OHIO STATE	49,228,384	70
EMORY	48,201,137	84
WASHINGTON	48,035,687	25
JOHNS HOPKINS	47,389,161	13
VIRGINIA	47,239,047	113
MINNESOTA	47,113,006	56
RUTGERS	46,209,907	176
NORTH CAROLINA	43,361,552	56
INDIANA	41,362,397	117

University	Value	2018world
NORTHWESTERN	40,660,526	20
MICHIGAN STATE	40,135,018	141
ALBERTA	40,106,398	119
NORTH CAROLINA STATE	38,375,475	251
WISCONSIN	37,820,507	43
	36,047,384	43
	35,453,700	

2.2 ARL Research University Library Total Collection Procurement Expenditures (Original: Total Library Materials Expenditures)

Total collection procurement expenditures (translated as "Total Library Materials Expenditures") ([1]P48) determine collection richness and accessibility. These expenditures account for 46.05% of total library budgets ([1]P1), representing nearly half of all funding. Among 116 ARL research university libraries, the top 30 in collection procurement expenditures consistently rank within the top 250 of THE World University Rankings. Accordingly, we analyze correlations between the top 30 libraries' collection procurement expenditure rankings and their 2018 THE rankings.

First, we correlate the top 30 libraries' collection procurement expenditure rankings with their universities' 2018 THE rankings:

```
[1]P48m
Mean SD world world rank
Pearson Correlation 1 Sig. (2-tailed) world
world Pearson Correlation .319 Sig. (2-tailed) .086
Pearson Correlation Sig. (2-tailed) 95% Confidence Interval (2-tailed) a rank - world a. Est

Nonparametric Correlations
Kendall's tau_b rank correlation coefficient .301* world Sig. (2-tailed) . world correlation
Spearman's Rho rank correlation coefficient .415* Sig. (2-tailed) . world correlation coeffi
*. Correlation is significant at the 0.05 level (2-tailed).

Confidence Interval for Kendall's tau_b
Sig. 95% Confidence Interval (2-tailed) a tau_b (2-tailed) rank - world a. Estimation is bas

Confidence Interval for Spearman's rho
Sig. 95% Confidence Interval (2-tailed) a,b Spearman's Rho (2-tailed) rank - world a. Estima
```

Results: Statistically significant correlations exist between the top 30 libraries' collection procurement expenditure rankings and THE rankings (Spearman r=0.415, P=0.023; Kendall r=0.301, P=0.019). (Due to skewed distributions, Spearman is more authentic; Pearson requires bivariate normal distribution, and its results are less accurate for non-normal data, so Pearson r=0.319,

P=0.086 can be disregarded here.)

Second, for further validation, we analyze the relationship between the top 30 libraries' actual collection procurement expenditure amounts and THE rankings, applying the same data transformations:

```
nxhworld zxhvalues
nxhworld Pearson Correlation 1 Sig. (2-tailed)
zxhvalues Pearson Correlation .334 Sig. (2-tailed) .071
Kendall's tau_b nxhworld correlation coefficient .301* Sig. (2-tailed) . nxhworld zxhvalues
Spearman's Rho nxhworld correlation coefficient .415* Sig. (2-tailed) . zxhvalues correlatio
*. Correlation is significant at the 0.05 level (2-tailed).
```

Results: Statistically significant correlations exist between the top 30 libraries' collection procurement expenditure amounts and THE rankings (Spearman r=0.415, P=0.023; Kendall r=0.301, P=0.019; Pearson r=0.334, P=0.071). (Due to skewed distributions, Spearman is more authentic.)

Findings: The top 30 ARL research university libraries' collection procurement expenditure amounts show significant positive correlation with THE rankings, demonstrating moderate correlation strength. This indicates that greater investment in collection procurement corresponds to higher THE rankings.

Data on ARL research university library total collection procurement expenditures, top 30 rankings, and corresponding 2018 THE rankings:

Rank Order Table 3 : Total Library Materials Expenditures

University	Value	2018world
HARVARD	49,064,256	6
COLUMBIA	42,897,682	14
PRINCETON	33,436,481	7
NEW YORK	32,728,697	27
TORONTO	32,397,006	22
MICHIGAN	31,074,257	20
TEXAS A&M	27,934,438	159
TEXAS	27,252,769	49
PENNSYLVANIA STATE	24,381,968	96
PENNSYLVANIA	23,211,152	10
JOHNS HOPKINS	23,064,668	13
CALIFORNIA, BERKELEY	22,327,045	18
EMORY	22,303,279	84
SOUTHERN CALIFORNIA	22,217,728	66
CORNELL	21,364,254	19
MINNESOTA	21,314,825	56
ILLINOIS, URBANA	20,259,941	37
MICHIGAN STATE	20,067,226	141

University	Value	2018world
CALIFORNIA, LOS ANGELES	19,831,535	15
ALBERTA	19,740,337	119
MCGILL	19,714,669	42
OHIO STATE	19,293,476	70
NORTHWESTERN	19,258,945	20
WASHINGTON	19,189,962	25
INDIANA	18,930,795	117
CHICAGO	18,687,658	9
PITTSBURGH	18,370,609	82
	18,234,849	
	17,826,259	
	17,558,257	

2.3 ARL Research University Library Total Volumes/Issues/Serials Expenditures (Original: Volumes in Library)

Total volumes/issues/serial expenditures (translated as "Volumes in Library") ([1]P46) constitute a major collection component. These expenditures represent 40.4% of total collection procurement funding ([1]P8, Mean 40.4%, Totals 40.4%) and approximately 20% of total library budgets (calculated from [1]P2-7, P46, P48, P51), representing a substantial portion of library spending. Among 116 ARL research university libraries, the top 30 in volumes/issues/serials expenditures consistently rank within the top 250 of THE World University Rankings. Accordingly, we analyze correlations between the top 30 libraries' volumes/issues/serials expenditure rankings and their 2018 THE rankings.

First, we correlate the top 30 libraries' volumes/issues/serials expenditure rankings with their universities' 2018 THE rankings:

```
[1]P46v
Mean SD world world rank
Pearson Correlation 1 Sig. (2-tailed) world
world Pearson Correlation .430* Sig. (2-tailed) .018
.430* *. Correlation is significant at the 0.05 level (2-tailed).
Sig. (2-tailed) 95% Confidence Interval (2-tailed) a Pearson Correlation rank - world a. Est

Nonparametric Correlations
Kendall's tau_b rank correlation coefficient .368** world Sig. (2-tailed) . world correlatio
Spearman's Rho rank correlation coefficient .515** Sig. (2-tailed) . world correlation coefi
**. Correlation is significant at the 0.01 level (2-tailed).

Confidence Interval for Kendall's tau_b
Sig. 95% Confidence Interval (2-tailed) a tau_b (2-tailed) rank - world a. Estimation is bas

Confidence Interval for Spearman's rho
```

Sig. 95% Confidence Interval (2-tailed) a,b Spearman's Rho (2-tailed) rank - world a. Estima

Results: Statistically significant correlations exist between the top 30 libraries' volumes/issues/serials expenditure rankings and THE rankings (Spearman r=0.515, P=0.004; Kendall r=0.368, P=0.004; Pearson r=0.430, P=0.018). (Due to skewed distributions, Spearman is more authentic.)

Second, for further validation, we analyze the relationship between the top 30 libraries' actual volumes/issues/serials expenditure amounts and THE rankings, applying the same data transformations:

nxhworld zzhvalues
nxhworld Pearson Correlation 1 Sig. (2-tailed)
zzhvalues Pearson Correlation .436* Sig. (2-tailed) .016
.436* *. Correlation is significant at the 0.05 level (2-tailed).
Kendall's tau_b nxhworld correlation coefficient .368** Sig. (2-tailed) . nxhworld zzhvalues
Spearman's Rho nxhworld correlation coefficient .515** Sig. (2-tailed) . zzhvalues correlati
**. Correlation is significant at the 0.01 level (2-tailed).

Results: Statistically significant correlations exist between the top 30 libraries' volumes/issues/serials expenditure amounts and THE rankings (Spearman r=0.515, P=0.004; Kendall r=0.368, P=0.004; Pearson r=0.436, P=0.016). (Due to skewed distributions, Spearman is more authentic.)

Findings: The top 30 ARL research university libraries' volumes/issues/serials expenditure amounts show significant positive correlation with 2018 THE rankings, demonstrating moderate correlation strength. This indicates that greater investment in volumes/issues/serials corresponds to higher THE rankings.

Data on ARL research university library total volumes/issues/serials expenditures, top 30 rankings, and corresponding 2018 THE rankings:

Rank Order Table 1 : Volumes in Library

University	Value	2018world
HARVARD	21,273,455	6
MICHIGAN	15,669,216	20
TORONTO	15,092,789	22
ILLINOIS, URBANA	15,033,508	37
COLUMBIA	14,431,600	14
CALIFORNIA, LOS ANGELES	14,374,971	15
CALIFORNIA, BERKELEY	13,047,310	18
CHICAGO	12,955,237	9
TEXAS	11,906,082	49
WISCONSIN	11,694,468	43
INDIANA	11,369,391	117
PRINCETON	10,737,321	7

University	Value	2018world
CORNELL	10,224,241	19
WASHINGTON	9,940,375	25
NORTH CAROLINA	9,247,789	56
COLORADO	9,030,377	129
OHIO STATE	8,969,582	70
PENNSYLVANIA STATE	8,762,717	96
PENNSYLVANIA	8,647,496	10
NEW YORK	8,496,644	27
ALBERTA	8,278,842	119
MICHIGAN STATE	7,970,760	141
MINNESOTA	7,777,787	56
ARIZONA	7,765,996	161
NORTHWESTERN	7,669,601	20
BRITISH COLUMBIA	7,663,181	34
BROWN	7,584,660	53
	7,515,156	
	7,191,341	
	6,851,411	

3.1 Conclusions

Through correlation analysis, this study finds that ARL research university libraries' total expenditures, total collection procurement funding, and total volumes/issues/serials expenditures all positively correlate with THE World University Rankings, demonstrating moderate correlation strength. Given that ARL research university libraries represent advanced educational standards in North America—the world's most developed educational region—these data reflect sophisticated educational levels. Therefore, increasing research university library total expenditures, collection procurement funding, and volumes/issues/serials expenditures enhances support services for university education, thereby improving educational quality and THE World University Rankings.

3.2 Recommendations

Due to funding constraints, many Chinese university libraries currently cannot match ARL standards for total expenditures, collection procurement funding, or volumes/issues/serials expenditures. Chinese libraries should therefore learn from ARL members to diversify funding sources and increase total investment. ARL libraries receive funding through multiple channels beyond regular institutional budgets, including research grants, special project funds, donations, and service fees. Most Chinese university libraries rely primarily on institutional budget allocations, representing a single source. Therefore, the first priority is

increasing total library expenditures to fundamentally improve service quality and level.

Second, diversifying funding sources will increase total expenditures. As total funding grows, collection procurement and other expenditures can increase accordingly. Third, optimizing fund allocation ensures efficient resource utilization. With increased expenditures, Chinese research university libraries can enhance support services for university education, improve educational quality, and elevate THE World University Rankings. These recommendations may offer reference points for the future development of Chinese research university libraries.

References

[1] Association of Research Libraries. ARL Statistics 2017-2018. [M]. Washington: Association of Research Libraries, 2019.

[2] 2018 Times Higher Education World University Rankings. [EB/OL]. [2020-09-04]. http://www.qianmu.org/.

[3] Wuyun Gaowa. Interpretation and Implications of ARL's *2018 Library Assessment Conference Proceedings* [J]. Library and Information Service, 2019(05): 79-84.

[4] Wan Xiaogang, Hong Qiulan. Research on Outcome Assessment of American Academic Libraries and Its Implications: An Analysis Based on ARL Library Assessment Conferences 2010-2018 [J]. Library Science Research, 2020(19): 88-94.

[5] Hu Hongxiao, Liu Junying, Wu Xu, Zhang Xiaoping. Analysis of Literature Resource Construction in Research Universities [J]. Jiangxi Library Science Journal, 2003(4).

[6] Peng Tao, Sun Huiming. Analysis of Literature Resource Construction in American ARL University Libraries in the Past Decade [P]. Journal of Academic Libraries, 2012(6): 62-65.

[7] Jantti M, Heath J. What role for libraries in learning analytics? [J]. Performance Measurement and Metrics, 2016, 17(2): 203-210.

[8] Zhu Lei, Yuan Xiaoping. Interpretation of the *2020 Top Trends in Academic Libraries* Report [J]. Inner Mongolia Science, Technology and Economy, 2021(7): 134-137.

[9] David J, Sheskin. Handbook of Parametric and Nonparametric Statistical Procedures [M]. New York: CRC, 2004.

[10] https://blog.csdn.net/weixin_{43876625}/article/details/123921224?spm=1001.2101.3001.6650.3&utm_{
task-blog-2$_{\text{default}}$CTRLIST$_{\text{Rate-3-123921224-blog-82455285.235\%5Ev32\%5Epc_\{\{\{relevant\}\}\{\{default\}\}\}\{base3\}\&depth_1-utm_}}$
3-123921224-blog-82455285.235%5Ev32%5Epc_{{{relevant}}}*{{default}}*{base3}&utm_{{{relevant}}}_{{ind

Author Biography: Su Li (born 1981), female, holds a graduate degree and serves as a librarian at Nanjing Normal University Library. Email: suli313@qq.com. Address: Nanjing Normal University, 1 Wenyuan Road, Qixia District, Nanjing, 210023. Contact: 13390901388.

Note: Figure translations are in progress. See original paper for figures.