JPPIPA 12(11) (2026)

Jurnal Penelitian Pendidikan IPA

http://jppipa.unram.ac.id/index.php/jppipa/index

Perceived Readiness for Statistics-Based Big Data Analytics in Higher Education: Evidence from Faryab University

Abdul Wahab Muradi[1], Naqibullah Rezayi[2*], Zaidullah Zaheed[3]

[1] Statistics and Econometrics Department, Economic Faculty, Faryab University, Afghanistan.
[2] Information Systems Department, Computer Science Faculty, Faryab University, Afghanistan.
[3] Mathematics Department, Education Faculty, Samangan University, Afghanistan.

Received: August 12, 2026
Revised: November 12, 2026
Accepted: November 25, 2026
Published: November 30, 2026

Corresponding Author*:
Naqibullah Rezayi
n.rezaee.of@gmail.com

DOI: 10.29303/jppipa.v12i11.16500

Open Access

© 2026 The Authors. This article is distributed under a (CC-BY License)

Abstract: The effective use of big data for institutional decision-making requires statistical competence, reliable infrastructure, and data quality. This study examined perceptions of the opportunities and challenges associated with statistics-based big data analytics at Faryab University, Afghanistan. A quantitative cross-sectional survey involved 300 respondents, comprising 224 students, 51 academic staff, and 25 administrative and information technology staff. Data were collected using a structured questionnaire covering seven dimensions, with Cronbach's alpha coefficients ranging from .88 to .92. Data were analysed using descriptive statistics, Welch's t-tests, one-way analysis of variance, chi-square tests, and multiple regression. The results revealed moderate or undecided perceptions across the measured constructs (M = 2.96–3.06). Respondents with formal statistical training reported higher statistical competence than those without training (d = 0.35). Perceived opportunities (β = .36) and statistical competence (β = .25) demonstrated the strongest positive associations with perceived analytical effectiveness. The six predictors collectively explained 45.5% of the outcome variance (R^2 = .455). These findings highlight the perceived potential of statistics-based big data analytics while identifying challenges related to institutional readiness and supporting conditions. The study underscores the importance of strengthening statistical competence and institutional infrastructure to support data-driven decision-making. However, the cross-sectional design and reliance on self-reported data limit causal interpretation and direct assessment of operational analytical capacity. Further research employing longitudinal designs, objective analytical assessments, and institutional datasets is recommended to evaluate actual analytical performance and intervention effectiveness.

Keywords: Afghanistan; Analytical readiness; Big data analytics; Higher education; Statistics

Introduction

Universities generate and maintain increasingly large volumes of digital information through student enrolment systems, examination records, learning management systems, library services, research activities, human resource administration, and financial operations. Although many of these records are initially collected for administrative purposes, their systematic analysis can provide useful insights into student progress, institutional performance, resource allocation, and academic decision-making. Big data analytics (BDA) offers opportunities to transform these records into meaningful information that can support planning and improve institutional processes. Nevertheless, the availability of large datasets does not automatically lead to better decisions. Their effective use depends on the quality of the underlying data, the analytical methods employed, and the ability of institutional members to interpret and apply the resulting information (Baig et al., 2020; Daniel, 2019; Stojanov & Daniel, 2024).

Statistics plays an essential role in this process because it provides the methods required to organise, analyse, interpret, and evaluate data. Statistical

How to Cite:
Muradi, A. W., Rezayi, N., & Zaheed, Z. (2026). Perceived Readiness for Statistics-Based Big Data Analytics in Higher Education: Evidence from Faryab University. *Jurnal Penelitian Pendidikan IPA*, *12*(11), 136–146. https://doi.org/10.29303/jppipa.v12i11.16500

competence enables individuals to select appropriate analytical techniques, identify patterns, assess uncertainty, and distinguish meaningful findings from misleading associations. Consequently, the value of big data analytics cannot be assessed solely through the availability of software, databases, or computing infrastructure. It also depends on the knowledge and skills of the people responsible for analysing and interpreting data. Previous research has similarly emphasised that organisational analytical capability involves a combination of technological resources, data management, employee competencies, and managerial support (Ghasemaghaei et al., 2018; Mikalef et al., 2018; Wang et al., 2018). Daniel (2019) further cautions that the use of large datasets does not eliminate methodological limitations; rather, inappropriate analytical choices may produce misleading conclusions even when substantial amounts of data are available.

Recent research has provided a broader understanding of big data analytics in higher education. Stojanov and Daniel (2024), in a systematic review of 75 studies, examined applications of big data and analytics in learning, teaching, and university administration. Their findings indicated that analytics has been used to support learning-related activities, while its contribution to teaching and administrative decision-making remains an important area for further development. Similarly, empirical research on learning analytics adoption has demonstrated that institutional implementation involves more than technical considerations. Alzahrani et al. (2023) identified interconnected challenges involving institutional culture, ethical considerations, analytical capabilities, leadership, and technology evaluation. Mukred et al. (2024) also found that perceived usefulness, management support, financial resources, and institutional support are relevant to the adoption of learning analytics tools in higher education institutions. More recently, Ajanaku and Ajiferuke (2025) examined big data analytics readiness and use in selected Canadian higher education institutions, identifying fragmented data environments, integration difficulties, and resource constraints as important institutional challenges. These studies demonstrate that readiness for data analytics is a multidimensional issue shaped by technological, organisational, and human factors.

Despite these developments, several challenges continue to limit the effective application of statistics-based big data analytics in universities. Institutional records may be incomplete, inconsistent, or distributed across systems that do not communicate effectively. In addition, limited infrastructure, inadequate analytical training, insufficient institutional coordination, and concerns about privacy and information security may restrict the use of available data. Sivarajah et al. (2017) classified big data challenges into issues associated with data, analytical processes, and management, while research on learning analytics has highlighted the importance of stakeholder engagement, ethical responsibility, data access, and institutional support (Alzahrani et al., 2023; Tsai et al., 2019). These challenges indicate that successful implementation requires not only access to analytical technologies but also appropriate institutional arrangements and the ability of users to understand and apply statistical information responsibly.

Although previous studies have examined big data applications, learning analytics adoption, and institutional readiness, important questions remain concerning how different groups within universities perceive their capacity to use statistics-based analytics. Three related gaps are particularly relevant to the present study. First, the methodological gap concerns the need to examine statistical competence, infrastructure support, data quality, perceived opportunities, perceived challenges, and analytical effectiveness within a common institutional survey framework. Second, the conceptual gap relates to the distinction between perceived readiness and actual analytical capability. Although respondents may consider their institution capable of using data analytics, questionnaire responses alone cannot establish whether the university possesses operational systems or produces accurate analytical results. Third, the contextual gap concerns the limited institution-specific evidence available from lower-resource higher education settings, particularly Afghanistan. Findings from universities with different technological and organisational conditions cannot automatically be generalised to institutions operating under different resource constraints.

Against this background, the present study examines perceptions of statistics-based big data analytics at Faryab University, Afghanistan. It focuses on students, academic staff, and administrative and information technology personnel because these groups may have different experiences, responsibilities, and levels of exposure to statistical analysis and institutional data systems. The study adopts a socio-technical perspective, treating perceived readiness as a combination of human capabilities, technological and data-related conditions, institutional opportunities and challenges, and perceived analytical effectiveness. Importantly, the study measures respondents' perceptions rather than directly auditing institutional databases or testing the performance of operational big data analytics systems. Its findings should therefore be interpreted as evidence of perceived readiness and perceived effectiveness, rather than verified institutional analytical capacity.

Accordingly, the study addresses the following four research questions:

RQ1: How do students and staff at Faryab University perceive their statistical competence, the opportunities associated with big data analytics, and the challenges affecting its use?

RQ2: How do respondents assess the university's perceived readiness in terms of infrastructure support, data quality, privacy and security, and analytical effectiveness?

RQ3: Do perceptions of statistics-based big data analytics differ according to respondents' institutional roles, experience, and formal statistical training?

RQ4: Which factors are associated with respondents' perceived analytical effectiveness in the context of statistics-based big data analytics?

The study contributes institution-specific empirical evidence concerning perceptions of statistical competence and big data analytics readiness in an Afghan higher education setting. By examining the views of students and different categories of university personnel, it provides a comparative understanding of perceived opportunities, institutional constraints, and factors associated with perceived analytical effectiveness. The findings may help identify areas requiring further investigation and inform the design of future institutional initiatives. However, recommendations concerning training, infrastructure development, and analytics implementation should be regarded as potential priorities for evaluation rather than interventions whose effectiveness has already been established.

Method

Research design and respondents

We used a quantitative, cross-sectional survey design and treated Faryab University as a single case. The data set contains 300 complete responses from three groups in the university community: students, academic staff, and administrative or IT staff. Respondents were aged 18 to 50 (M = 34.50, SD = 9.80). Table 1 gives the profile. Because about three in four respondents were students, the results lean toward the student viewpoint; the staff groups (51 academic and 25 administrative/IT respondents) are reported separately but are small.

Table 1. Profile of respondents (N = 300)

Characteristic	Category	*n*	%
Gender	Male	192	64.0
	Female	108	36.0
Role	Student	224	74.7
	Academic staff	51	17.0
	Administrative/IT staff	25	8.3
Highest education	Bachelor	97	32.3
	Master	134	44.7
	Doctorate	45	15.0
	Other/professional	24	8.0
Big data experience	None	95	31.7
	Less than 2 years	100	33.3
	2 to 5 years	78	26.0
	More than 5 years	27	9.0
Formal training in statistics/big data	Yes	126	42.0
	No	174	58.0

Instrument and measures

The questionnaire had eight background questions and 35 statements covering seven constructs, each answered on a five-point scale from 1 (strongly disagree) to 5 (strongly agree). Table 2 lists the constructs with an example statement. A construct score is the mean of its items. Internal consistency was checked with Cronbach's alpha (Equation 1), where k is the number of items, σ^2_i the variance of item i and σ^2_T the variance of the total score. All constructs were above the conventional .70 benchmark (Taber, 2018).

$$\alpha = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k}\sigma_i^2}{\sigma_T^2}\right) \quad (1)$$

Table 2. Constructs, number of items, example statements and reliability

Construct	Items	Example statement	α
Statistical competence (SC)	5	I can interpret statistical results generated from large datasets.	.91
Perceived opportunities (OP)	6	Statistics can help identify patterns in student performance and learning.	.91
Perceived challenges (CH)	6	Limited computing infrastructure makes big-data analysis difficult.	.92
Infrastructure support (INF)	5	Relevant statistical and data-analysis software is available.	.90
Data quality (DQ)	4	University data are generally accurate and consistent.	.88
Privacy/security concern (PV)	4	The university needs stronger data-security measures for big-data analytics.	.88
Analytical effectiveness (OUT)	5	Statistical tools improve the efficiency of institutional analysis.	.90
Readiness index (SC, INF, DQ, OUT)	19	Composite of the four constructs above	.92

Overall readiness for big data analytics (R) was defined as the average of statistical competence, infrastructure support, data quality and analytical effectiveness (Equation 2). Scores below 2.34 were classed as low readiness, 2.34 to 3.66 as moderate, and 3.67 or above as high, which splits the 1 to 5 scale into equal thirds. The same bands were used to read construct means: roughly, below 2.34 means disagreement, above 3.66 means agreement, and anything between is close to neutral.

$$R = \frac{SC+INF+DQ+OUT}{4} \quad (2)$$

Data analysis

Analyses were run in Python with pandas, SciPy and statsmodels. We first looked at means, standard deviations, skewness and kurtosis. Shapiro-Wilk tests were significant for several constructs, which is common with 300 cases, but skewness and kurtosis stayed within ±0.3, so we judged the scores adequate for parametric tests. Group differences used Welch t-tests (two groups) and one-way ANOVA (more than two groups), with Kruskal-Wallis tests as a check, and chi-square tests for readiness levels. Cohen's d and η^2 describe effect size. To see which factors go with analytical effectiveness, we fitted a multiple regression with the other six constructs as predictors and examined variance inflation factors (Hair, Black, Babin, & Anderson, 2019). The significance level was .05. Many comparisons were made, so results with p between .01 and .05 should be read as leads, not firm findings.

Result and Discussion

Overall pattern of responses

The most striking thing about the averages is how close together they are (Table 3, Figure 1). All seven construct means fall between 2.96 and 3.06, in other words right around the neutral midpoint. Perceived opportunities scored highest (M = 3.06) and data quality lowest (M = 2.96), but a gap of one tenth of a point is too small to build an argument on. The readiness index averaged 2.98 (SD = 0.56). Respondents, taken together, are neither enthusiastic nor sceptical about big data analytics; they are largely undecided.

There are a few reasons this might be so. Big data analytics is still a new idea on many campuses, and almost a third of respondents (31.7%) said they had no experience with big data tools at all. Three-quarters were students, who may see little of how institutional data are actually handled. It is also possible that the scale simply attracted many neutral answers. Whatever the cause, the practical reading is that opinion is still open, and what the university does next can shape it.

Table 3. Descriptive statistics of the study constructs (N = 300)

Construct	M	SD	Skewness	Kurtosis	Reading
Statistical competence	2.97	0.78	-0.01	-0.08	Neutral
Perceived opportunities	3.06	0.71	-0.27	-0.15	Neutral
Perceived challenges	3.01	0.75	-0.11	-0.28	Neutral
Infrastructure support	2.99	0.74	-0.04	-0.14	Neutral
Data quality	2.96	0.78	0.19	-0.15	Neutral
Privacy/security concern	2.98	0.78	0.01	-0.25	Neutral
Analytical effectiveness	3.02	0.73	0.06	-0.19	Neutral
Readiness index	2.98	0.56	0.11	-0.09	Neutral

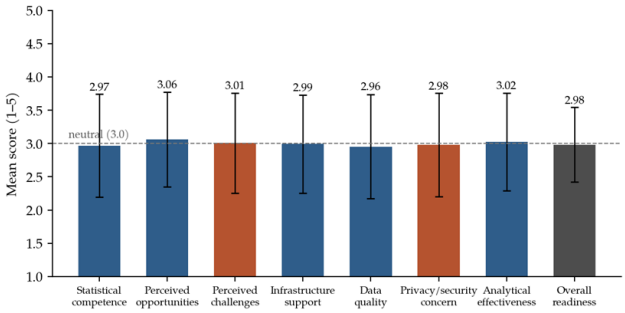


Figure 1. Mean score (±1 SD) for each construct. Blue bars are favourable constructs, orange bars are constraint constructs, grey is the composite index; the dashed line marks the neutral value of 3.

Opportunities and challenges, item by item

Looking at single statements gives a slightly sharper picture (Figure 2). On every one of the six opportunity items, more people agreed than disagreed: for example 31.3% agreed that analytics can improve evidence-based decisions against 23.0% who disagreed, and 31.0% against 25.7% for spotting patterns in student learning. The margins are small, but they all point the same way. The challenge items were more evenly split. Limited computing infrastructure (30.7% agree, 27.0% disagree) and privacy and security restrictions (29.3% vs. 25.7%) tilted slightly toward agreement, whereas limited access to quality datasets tilted the other way (27.7% vs. 30.0%). Neutral was the most common answer on every item, ranging from about 42% to 56%.

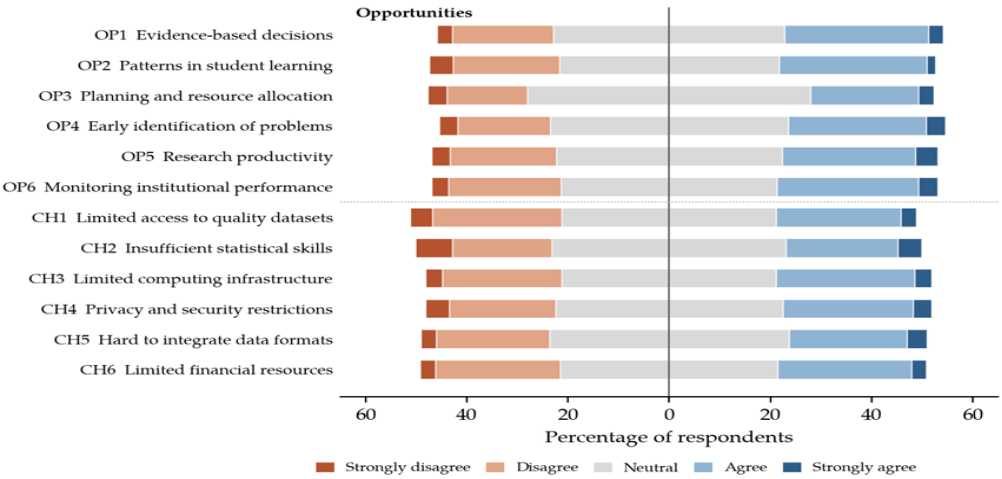


Figure 2. Distribution of responses to the opportunity items (OP1 to OP6, upper six bars) and challenge items (CH1 to CH6, lower six bars), centred on the neutral category.

This fits the mixed evidence in the literature. Respondents see some upside, which echoes the optimism about analytics in higher education (Baig et al., 2020), but they do not rank any single obstacle above the rest, unlike the sharper data, process and management distinctions drawn by Sivarajah et al. (2017). A community without much hands-on experience may simply not yet know which barrier will bite first.

Readiness levels

When the readiness index was split into bands, 226 respondents (75.3%) were moderately ready, 38 (12.7%) were low and 36 (12.0%) were high (Figure 3). The low and high groups are almost the same size, so the university is not tilted toward either extreme. It is a broad middle with two thin tails.

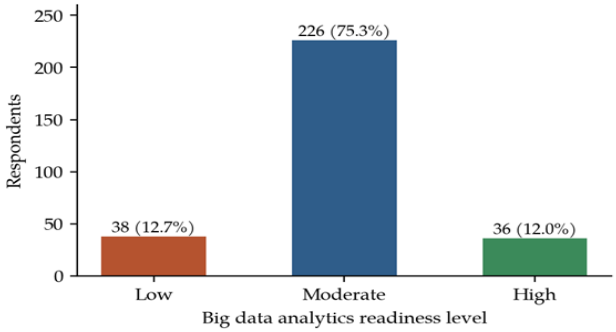


Figure 3. Number and share of respondents in each readiness band.

How the constructs relate

Figure 4 shows the correlations. The readiness index correlates strongly with its four components (r = .72 to .77), but that is built in by the way it is calculated. The other links are more informative. Perceived opportunities and analytical effectiveness go together (r = .57), as do statistical competence and effectiveness (r = .50). Perceived challenges run the opposite way, correlating negatively with effectiveness (r = −.27), infrastructure support (r = −.31) and data quality (r = −.29). Privacy and security concern is tied mainly to perceived challenges (r = .43) and hardly at all to the other constructs ($|r| \leq .19$). In other words, worry about privacy seems to be one part of a general sense of obstacles, not a separate brake on enthusiasm for analytics.

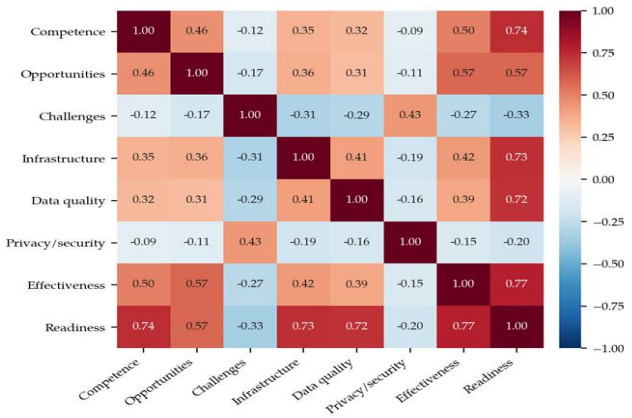


Figure 4. Pearson correlations among the study constructs (N = 300).

Differences between groups

Formal training made the clearest difference (Table 4). Respondents with training rated their statistical competence higher than those without (3.12 vs. 2.86; $t(281) = 2.97$, $p = .003$, $d = 0.35$), and scored somewhat higher on perceived opportunities ($d = 0.24$, $p = .044$) and analytical effectiveness ($d = 0.24$, $p = .037$). The overall readiness index was close to, but not past, the conventional cut-off ($p = .057$). Training had no bearing on challenges, infrastructure, data quality or privacy

concern, which makes sense: a course changes what a person can do, not what the university provides.

Table 4. Comparison of respondents with and without formal training (Welch *t*-test)

Construct	No training (n = 174) M (SD)	Training (n = 126) M (SD)	t	p	d
Statistical competence	2.86 (0.79)	3.12 (0.73)	2.97	.003	0.35
Perceived opportunities	2.99 (0.72)	3.16 (0.69)	2.02	.044	0.24
Perceived challenges	2.98 (0.74)	3.04 (0.77)	0.64	.524	0.07
Infrastructure support	3.02 (0.73)	2.96 (0.76)	-0.68	.494	-0.08
Data quality	2.91 (0.78)	3.02 (0.78)	1.30	.196	0.15
Privacy/security concern	2.95 (0.76)	3.01 (0.81)	0.63	.530	0.07
Analytical effectiveness	2.95 (0.77)	3.12 (0.68)	2.09	.037	0.24
Readiness index	2.93 (0.57)	3.06 (0.55)	1.91	.057	0.22

Note. d is positive when the trained group scored higher. No correction was applied for multiple comparisons.

The same pattern shows up in the readiness bands (Figure 5a): 16.7% of trained respondents reached the high band against 8.6% of untrained ones, while 7.1% of trained respondents were in the low band against 16.7% ($\chi^2(2) = 9.20$, $p = .010$). We cannot say training caused this. People who are already keen on statistics may be the ones who sign up for courses, so the data show an association and nothing more.

Experience with big data tools told a similar story for competence. Mean statistical competence rose from 2.81 for people with no experience to 2.95, 3.01 and 3.43 for those with less than two years, two to five years and more than five years ($F(3, 296) = 4.69$, $p = .003$, $\eta^2 = .045$). Privacy and security concern also differed somewhat across experience groups ($F(3, 296) = 2.84$, $p = .038$, $\eta^2 = .028$), with the two most experienced groups more concerned (means 3.15 and 3.13) than the least experienced (2.83). This is plausible: people who have worked with real data tend to notice the risks. Readiness bands were associated with experience too ($\chi^2(6) = 12.96$, $p = .044$). Gender, highest education and years of work experience made no meaningful difference to any construct. Role came closest, with perceived opportunities highest among administrative/IT staff ($M = 3.27$), then students (3.08) and academic staff (2.88; $F(2, 297) = 2.98$, $p = .052$). With only 25 administrative/IT respondents, we treat this as a lead for future work, not a result.

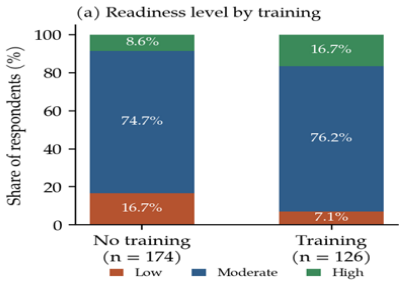
What goes with analytical effectiveness

The regression model was significant, $F(6, 293) = 40.80$, $p < .001$, and explained 45.5% of the variance in analytical effectiveness (adjusted $R^2 = .444$). Variance inflation factors were all at or below 1.39, so overlap among predictors was not a concern (Hair et al., 2019). Perceived opportunities were the strongest predictor ($\beta = .36$), followed by statistical competence ($\beta = .25$), infrastructure support ($\beta = .13$) and data quality ($\beta = .11$). Perceived challenges had a small negative effect ($\beta = -.11$), and privacy and security concern added essentially nothing once the others were accounted for ($\beta = .01$, $p = .890$). Figure 5b shows the standardised coefficients.

Table 5. Multiple regression predicting analytical effectiveness (N = 300)

Predictor	B	SE	β	t	p	VIF
Statistical competence	0.233	0.048	0.246	4.88	< .001	1.36
Perceived opportunities	0.371	0.052	0.359	7.11	< .001	1.37
Perceived challenges	-0.107	0.049	-0.109	-2.18	.030	1.36
Infrastructure support	0.128	0.051	0.129	2.53	.012	1.39
Data quality	0.106	0.046	0.113	2.28	.023	1.32
Privacy/security concern	0.006	0.045	0.007	0.14	.890	1.23

Note. $R^2 = .455$, adjusted $R^2 = .444$, $F(6, 293) = 40.80$, $p < .001$. B = unstandardised coefficient; β = standardised coefficient.



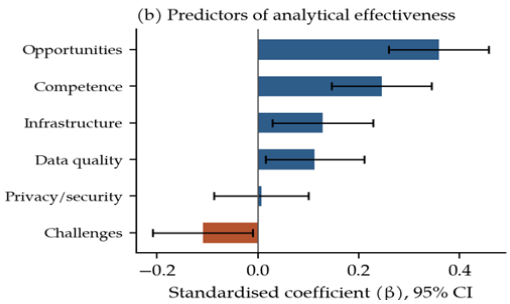


Figure 5. (a) Readiness band by training status; (b) standardised regression coefficients (β) with 95% confidence intervals for the prediction of analytical effectiveness.

Two points deserve care. First, the result that opportunities and competence matter most agrees with the argument that skills and positive expectations, not equipment alone, drive effective analytics (Ghasemaghaei et al., 2018; Mikalef et al., 2018). Second, opportunity and effectiveness were both reported by the same people in the same survey, and the effectiveness items overlap in content with the opportunity items, so part of this strong link may reflect a general positive attitude. The model describes how perceptions line up; it does not show that raising one will raise the other.

Implications and limitations

Several practical points follow. Training is the most obvious lever, since it was the one background factor tied to competence, opportunity ratings and readiness; short, applied courses in descriptive and inferential statistics for large datasets would be a reasonable start. Infrastructure and data quality matter, but their weights in the model were modest, and infrastructure ratings were neutral on average, so these are improvements to pursue alongside training and not before it. The privacy result is more nuanced. Concern was not related to perceived effectiveness, yet it was higher among more experienced respondents, which suggests that clear data-protection rules should be set up early, before heavier use brings the issue to the surface. Finally, with so many people undecided, one or two small, visible projects, such as tracking patterns in student performance or monitoring institutional indicators (the opportunity items with the most support), might do more to move opinion than general announcements.

The study has limits. Everything rests on self-reported perceptions, not on measured skills or audited data quality. It covers one university, at one time, with students making up most of the sample, so it cannot be generalised to other institutions and cannot show change over time. The readiness bands are conventional equal-thirds cut-offs and have not been validated against outside criteria. Finally, many tests were run without correction, which is why the smaller *p*-values are best treated as leads. Future work could add objective measures of skill and data quality, include more staff, and follow the university over several years.

Discussion

The findings indicate that respondents at Faryab University hold generally moderate and largely undecided perceptions regarding the institution's readiness for statistics-based big data analytics. The mean scores of the seven constructs ranged from 2.96 to 3.06, suggesting that respondents neither expressed strong agreement nor substantial disagreement about the university's analytical capabilities. Perceived opportunities recorded the highest mean (M = 3.06), whereas data quality received the lowest (M = 2.96). However, the difference between these values is relatively small and should not be interpreted as evidence of a meaningful practical distinction between the constructs.

The readiness index also produced a moderate mean score (M = 2.98, SD = 0.56). This finding suggests that respondents do not have a particularly positive or negative perception of the conditions required for implementing statistics-based big data analytics. Nevertheless, a moderate perception score should not be equated with verified institutional readiness. The study did not conduct a technical assessment of university databases, evaluate existing analytical systems, or test the accuracy of statistical outputs. Consequently, the results describe respondents' perceptions rather than the university's objectively measured operational capacity.

The relatively neutral responses may partly reflect differences in respondents' familiarity with big data analytics. Approximately 31.7% reported having no previous experience with big data tools, while students represented 74.7% of the total sample. These characteristics may have influenced how respondents evaluated institutional analytical opportunities and constraints. Students, for example, may have limited direct exposure to administrative databases, institutional reporting procedures, and analytical decision-making. Therefore, their responses may reflect general expectations rather than practical experience with university data systems.

These findings are broadly consistent with the multidimensional understanding of analytics readiness discussed in previous higher education research. Stojanov and Daniel (2024) emphasised that the application of analytics in universities extends across learning, teaching, and administrative activities, while Alzahrani et al. (2023) identified institutional, technological, ethical, and organisational considerations as interconnected factors affecting learning analytics adoption. Similarly, Mukred et al. (2024) highlighted the relevance of management support, financial resources, and institutional conditions. The present findings extend this discussion by providing institution-specific evidence from an Afghan university, although the single-institution design does not permit generalisation to all Afghan higher education institutions.

Perceived Opportunities and Institutional Challenges

The item-level findings reveal that respondents generally recognised potential opportunities associated with statistics-based big data analytics, particularly its possible contribution to evidence-based decision-making and the identification of patterns in student performance and learning. Nevertheless, the differences between agreement, disagreement, and neutral responses were relatively limited. This pattern suggests that the perceived value of analytics is acknowledged, but respondents may not have sufficient practical experience to assess its potential contribution confidently.

The findings are consistent with the broader argument that statistical and analytical methods can support the interpretation of institutional information and improve the availability of evidence for decision-making (Baig et al., 2020; Daniel, 2019). However, recognising these potential benefits does not establish that analytics has already improved institutional performance. The distinction is important because the present study measures perceived opportunities rather than documented changes in academic outcomes, administrative efficiency, or institutional decision quality.

The findings concerning institutional challenges were similarly mixed. Respondents expressed some concern about limited computing infrastructure and privacy and security restrictions, while their responses regarding access to quality datasets were comparatively divided. These results indicate that the challenges associated with analytics are not perceived as a single, clearly dominant problem. Instead, they may involve several interconnected conditions, including technological resources, data management practices, institutional coordination, and awareness of responsible data use.

This interpretation is consistent with Sivarajah et al. (2017), who discussed big data challenges in relation to data characteristics, analytical processes, and management considerations. Nevertheless, the present study does not independently establish the actual extent of infrastructure limitations or deficiencies in university databases. The findings should therefore be understood as evidence of perceived institutional constraints. A future technical assessment involving database audits, system interoperability checks, and data-quality indicators would be necessary to establish the operational significance of these concerns.

Statistical Training, Analytical Competence, and Experience

One of the more informative findings concerns the association between formal statistical training and perceived statistical competence. Respondents who reported previous training obtained a higher competence mean (M = 3.12) than those without formal training (M = 2.86), with a small-to-moderate effect size (d = 0.35). The difference was statistically significant (p = .003), indicating that training status was associated with respondents' assessments of their statistical capabilities.

The relationship between training and competence is theoretically plausible because formal instruction can provide opportunities to develop knowledge of statistical concepts, analytical procedures, and data interpretation. Previous research has also emphasised the importance of human analytical capabilities in organisational big data analytics (Ghasemaghaei et al., 2018; Mikalef et al., 2018). In the context of Faryab University, the finding suggests that statistical training may be an important area for institutional development.

However, the association should not be interpreted as evidence that training directly improved respondents' competence. Individuals with greater interest in statistics or previous analytical experience may be more likely to participate in training programmes. This possibility of self-selection means that the observed difference may partly reflect pre-existing differences between trained and untrained respondents. Furthermore, competence was measured through self-reported questionnaire items rather than an objective statistical knowledge examination.

The analysis also identified differences in perceived opportunities and analytical effectiveness according to training status, although their effect sizes were relatively small (d = 0.24 for both comparisons). The readiness index difference did not reach the conventional significance threshold (p = .057). These results provide a more cautious picture than a general claim that training is associated with every aspect of institutional readiness. In particular, training status was not significantly associated with perceived infrastructure support, data quality, challenges, or privacy and security concerns.

A similar pattern emerged when respondents were grouped according to their experience with big data tools. Statistical competence increased across the reported experience categories, from 2.81 among respondents without experience to 3.43 among those with more than five years of experience. The difference was statistically significant, although the effect size was relatively small (η^2 = .045). This finding suggests that familiarity with analytical tools is associated with greater perceived statistical competence. Nevertheless, because experience and competence were measured at the same point in time, it is not possible to determine whether experience contributes to competence or whether individuals with stronger existing skills are more likely to acquire practical experience.

Privacy and security concerns also differed across experience categories (η^2 = .028). Respondents with greater experience reported somewhat higher concern,

which may reflect increased awareness of the responsibilities associated with handling institutional information. However, this interpretation remains tentative because the study did not measure actual privacy incidents, security practices, or respondents' knowledge of data-protection procedures.

Relationships Among Statistical Competence, Opportunities, and Perceived Analytical Effectiveness

The correlation analysis revealed positive associations between perceived opportunities and analytical effectiveness ($r = .57$) and between statistical competence and analytical effectiveness ($r = .50$). In contrast, perceived challenges were negatively associated with analytical effectiveness ($r = -.27$), infrastructure support ($r = -.31$), and data quality ($r = -.29$). These relationships suggest that respondents who evaluated the opportunities and their own statistical competence more positively also tended to evaluate analytical effectiveness more positively.

The regression analysis provided further evidence of these relationships. The six predictors collectively explained 45.5% of the variance in perceived analytical effectiveness ($R^2 = .455$; adjusted $R^2 = .444$). Perceived opportunities ($\beta = .359$) and statistical competence ($\beta = .246$) had the largest positive standardised coefficients. Infrastructure support ($\beta = .129$) and data quality ($\beta = .113$) also showed positive associations, whereas perceived challenges had a negative coefficient ($\beta = -.109$). Privacy and security concerns did not demonstrate a statistically significant independent association with the outcome ($\beta = .007$, $p = .890$).

The relatively strong association between perceived opportunities and analytical effectiveness may indicate that respondents who recognise the potential value of analytics also tend to evaluate its usefulness more positively. Similarly, the relationship between statistical competence and perceived analytical effectiveness highlights the relevance of human capabilities in interpreting and using statistical information. These findings are compatible with the organisational capability perspective discussed by Ghasemaghaei et al. (2018) and Mikalef et al. (2018), which emphasises the interaction between analytical skills and organisational resources.

Nevertheless, the regression coefficients should be interpreted as associations with self-reported analytical effectiveness, not as evidence of actual improvements in institutional analytical performance. All constructs were measured using the same questionnaire and collected from the same respondents. Consequently, common-method variance, social desirability, and similarities in item wording may have contributed to some of the observed relationships. In particular, perceived opportunities and analytical effectiveness may overlap conceptually because both concern respondents' positive assessments of the usefulness of analytics. This measurement overlap could partly account for their relatively strong association.

The low variance inflation factor values (1.23–1.39) indicate that severe multicollinearity among the predictors was not evident according to this diagnostic. However, low multicollinearity does not eliminate the possibility of common-method bias or conceptual overlap between constructs. Other factors, including institutional role, previous training, analytical experience, and access to relevant systems, may also influence respondents' assessments. Therefore, the regression model provides an account of how the measured perceptions are statistically related, rather than establishing a causal pathway through which one construct produces changes in another.

Differences Across Institutional Roles and Respondent Characteristics

The subgroup findings require careful interpretation because the sample was not evenly distributed across institutional roles. Students accounted for 224 respondents, compared with 51 academic staff and 25 administrative or IT staff. Consequently, the overall descriptive results are more representative of student perceptions than of the views of university personnel responsible for teaching, administration, or information technology.

Perceived opportunities differed across institutional roles, with the highest mean reported by administrative and IT staff ($M = 3.27$), followed by students ($M = 3.08$) and academic staff ($M = 2.88$). However, the overall comparison did not reach conventional statistical significance ($p = .052$). The finding should therefore be treated as exploratory rather than as evidence of a confirmed difference between institutional groups. The relatively small administrative and IT subgroup also limits the precision of estimates for this category.

The results indicate that institutional role may be relevant to how individuals perceive the opportunities associated with analytics, but the present sample does not provide sufficient evidence to establish reliable differences between all respondent groups. Future research should recruit more balanced numbers of students, academic staff, and administrative or IT personnel. Such a design would provide a stronger basis for comparing perceptions and examining whether direct involvement in institutional data management is associated with different assessments of analytical readiness.

Implications for Higher Education Institutions

The findings provide several implications for universities considering the development of statistics-

based big data analytics. First, the association between formal training and perceived statistical competence suggests that statistical education and practical analytical training deserve institutional attention. Universities may consider offering introductory and intermediate training in descriptive statistics, inferential analysis, data visualisation, and responsible interpretation of analytical results. However, the effectiveness of such programmes should be evaluated through objective assessments before their benefits are assumed.

Second, the relatively neutral perceptions of infrastructure support and data quality indicate areas that warrant further institutional investigation. Rather than immediately investing in advanced analytical platforms, universities could begin by examining the availability of existing data, the consistency of administrative records, the compatibility of information systems, and the accessibility of statistical software. These preliminary assessments would help identify practical requirements and avoid investments that are not aligned with institutional needs.

Third, privacy and security considerations should be incorporated into any future analytics initiative. Although privacy concerns did not independently predict perceived analytical effectiveness in the regression model, this finding does not imply that data protection is unimportant. Universities should establish appropriate access controls, data-handling procedures, anonymisation practices, and institutional responsibilities before using student and staff information for analytical purposes.

A feasible approach would involve a small-scale pilot project using a clearly defined institutional dataset, such as anonymised student examination or enrolment records. The pilot could evaluate data completeness, analytical accuracy, processing time, and the usefulness of resulting reports for institutional decision-making. Baseline measurements should be established before implementation, followed by a subsequent evaluation using the same indicators. The responsible academic, statistical, and information technology units should also monitor data security and ensure that the analysis serves legitimate institutional purposes.

These proposals represent potential institutional priorities derived from the observed perceptions and relevant literature. They should not be interpreted as interventions whose effectiveness has already been demonstrated by the present study.

Limitations and Directions for Future Research

Several limitations should be considered when interpreting the findings. First, the study was conducted at a single university in Afghanistan. Although it provides institution-specific evidence, its findings cannot automatically be generalised to other universities with different institutional structures, financial resources, technological environments, and analytical capabilities.

Second, the cross-sectional design captures respondents' perceptions at one point in time. It therefore cannot establish the direction of relationships or determine whether statistical training, analytical competence, or perceived opportunities produce improvements in analytical effectiveness. The observed associations may also reflect self-selection into training and differences in respondents' previous experience.

Third, the study relied on self-reported questionnaire responses rather than objective assessments of statistical knowledge or institutional analytical performance. Consequently, the reported competence scores may not correspond directly to respondents' demonstrated analytical skills. Similarly, perceived data quality and infrastructure support do not replace technical assessments of databases, software availability, or information-system performance.

Fourth, the respondent distribution was uneven, with students accounting for approximately three-quarters of the sample. The smaller academic and administrative/IT subgroups limit the precision of comparisons across institutional roles. Future investigations should consider stratified sampling to obtain more balanced representation and should report subgroup estimates with appropriate confidence intervals.

Fifth, although the questionnaire demonstrated satisfactory internal consistency, Cronbach's alpha alone does not establish the validity or conceptual distinctness of the seven constructs. Additional evidence from expert assessment, exploratory or confirmatory factor analysis, and discriminant validity testing would strengthen the measurement framework. The readiness index also includes perceived analytical effectiveness as one of its components. Therefore, it should be interpreted as a broad composite of perceived analytical conditions and effectiveness rather than as an independent measure of operational readiness.

Finally, the use of a common questionnaire for all constructs creates the possibility of common-method variance and measurement overlap. The positive relationship between perceived opportunities and analytical effectiveness should therefore be interpreted with caution.

Future research could address these limitations through a preregistered longitudinal or intervention-based design involving multiple Afghan universities. Such studies could combine questionnaires with objective statistical knowledge tests, audited institutional datasets, and practical analytical tasks. A structured training intervention could be evaluated by

comparing participants' analytical performance before and after training, using predefined indicators such as accuracy, processing time, data completeness, and the quality of statistical interpretation. More balanced sampling across students, academic staff, and administrative or IT personnel would also provide a stronger basis for examining institutional differences.

Conclusion

Faryab University's students and staff are, on balance, undecided about big data analytics: they see some opportunity, recognise some obstacles, and most sit in the moderate readiness band (75.3%). What separates the more ready from the less ready is mostly skill and exposure. Formal training and hands-on experience were linked to higher statistical competence, and perceived opportunities and competence were the strongest predictors of analytical effectiveness in a model that explained almost half of its variance. Infrastructure and data quality contributed less but still significantly, and privacy concern was not a predictor. The most useful next step for the university is probably to invest in statistical training and a few well-chosen pilot projects, and to build data-protection practice while use is still modest.

Acknowledgments

The authors thank the students and staff of Faryab University who took part in the survey.

Author Contributions

Conceptualization, A.W.M. and N.R.; methodology, A.W.M.; software, N.R.; formal analysis, A.W.M.; writing—original draft preparation, A.W.M.; writing—review and editing, Z.Z.; supervision, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

References

Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2020). Big data in education: A state of the art, limitations, and future research directions. *International Journal of Educational Technology in Higher Education, 17*, Article 44. https://doi.org/10.1186/s41239-020-00223-0

Daniel, B. K. (2019). Big data and data science: A critical review of issues for educational research. *British Journal of Educational Technology, 50*(1), 101-113. https://doi.org/10.1111/bjet.12595

Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems, 27*(1), 101-113. https://doi.org/10.1016/j.jsis.2017.10.001

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Andover, UK: Cengage Learning.

Ifenthaler, D., & Yau, J. Y.-K. (2020). Utilising learning analytics to support study success in higher education: A systematic review. *Educational Technology Research and Development, 68*(4), 1961-1990. https://doi.org/10.1007/s11423-020-09788-z

Mikalef, P., Pappas, I. O., Krogstie, J., & Giannakos, M. (2018). Big data analytics capabilities: A systematic literature review and research agenda. *Information Systems and e-Business Management, 16*(3), 547-578. https://doi.org/10.1007/s10257-017-0362-y

Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research, 70*, 263-286. https://doi.org/10.1016/j.jbusres.2016.08.001

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Tsai, Y.-S., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2019). Complexity leadership in learning analytics: Drivers, challenges and opportunities. *British Journal of Educational Technology, 50*(6), 2839-2854. https://doi.org/10.1111/bjet.12846

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior, 89*, 98-110. https://doi.org/10.1016/j.chb.2018.07.027

Wang, Y., Kung, L., & Byrd, T. A. (2018). Big data analytics: Understanding its capabilities and potential benefits for healthcare organizations. *Technological Forecasting and Social Change, 126*, 3-13. https://doi.org/10.1016/j.techfore.2015.12.019