

Entropy-Weight Derivation for Ranking STEM–CBE Preparedness Indicators in Public Senior Secondary Schools: Evidence from Bungoma County, Kenya

YASIRO CARRÉN MUTUA¹, MOSES KOLOLI², JACINTA MUTWIWA³

^{1, 2, 3}*Department of Mathematics, Faculty of Science, Kibabii University*

Abstract- The rollout of Competency Based Education (CBE) in Kenya's senior secondary schools has created a need for objective, data-driven tools to monitor school readiness for the Science, Technology, Engineering and Mathematics (STEM) pathway. School preparedness is multidimensional, yet most evaluation schemes treat its indicators as equally important, which overweights indicators that carry little discriminating information and underweights those that separate schools most sharply. This paper derives objective weights for twelve STEM–CBE preparedness indicators using Shannon entropy, as a preliminary step to entropy-weighted clustering. Data were collected from 112 public senior secondary schools across the nine sub-counties of Bungoma County through structured questionnaires and facility observation checklists under stratified random sampling. After min–max normalisation, the informational contribution of each indicator was quantified through its normalised Shannon entropy, converted to a degree of diversification, and normalised to a unit-sum weight vector. The derived weights ranged from 0.022 to 0.141—a spread of roughly six-fold—and summed to unity. ICT device provision (weight 0.141) and ICT-integrated lesson frequency (0.119) were the two sharpest markers of preparedness, and together with stakeholder engagement (0.112) and the presence of a STEM strategic plan and budget (0.104) the ICT and institutional-leadership indicators carried about 47.6 percent of the total informational value. Student competency in STEM tasks (0.022), practical lesson frequency (0.041) and STEM teacher density (0.052) were least informative, reflecting limited variation across schools rather than low substantive importance. The ranking identifies the digital divide and governance capacity as the conditions that most sharply differentiate senior-school readiness in the county and supplies a defensible, reproducible weight vector for downstream profiling.

Keywords: Entropy Weighting, Shannon Entropy, STEM–CBE Preparedness, Indicator Ranking

I. INTRODUCTION

Education systems worldwide have shifted from examination-oriented curricula toward competency-based frameworks that aim to produce practically skilled, analytical and technologically literate graduates [1].

In Kenya, the Competency Based Curriculum was institutionalised through Sessional Paper No. 1 of 2019 and progressively rolled out from pre-primary to secondary level; in 2025 the Ministry of Education broadened it into Competency Based Education (CBE), integrating curriculum, pedagogy, assessment and infrastructure [2, 3].

Within the senior secondary cycle, the Science, Technology, Engineering and Mathematics (STEM) pathway is treated as the principal driver of innovation and economic competitiveness [1], and its delivery presupposes that schools are adequately prepared across infrastructure, human resources, curriculum alignment, ICT integration and institutional leadership.

Preparedness for the STEM pathway is therefore not a one-dimensional construct that can be captured by a simple checklist or an aggregate score; it is a multidimensional latent condition whose constituent indicators vary widely in how sharply they distinguish one school from another [4, 5].

When such indicators are combined into a readiness measure, the weight given to each indicator determines which schools are flagged as prepared or underprepared. Assigning equal weight to every indicator—the implicit default of most administrative rating schemes—ignores the fact that some indicators

separate schools strongly while others vary little across the system, and thus distorts the resulting typology [6].

Information theory offers a principled remedy. Shannon entropy quantifies the dispersion of an indicator across observations, so that indicators exhibiting greater variability are judged to carry more information and merit greater weight [7, 8].

Entropy weighting has been applied successfully in clinical decision-making [10], supplier selection [6] and educational resource evaluation [11], but its use for profiling school readiness for STEM under CBE has not been reported for Kenya.

This paper addresses the first stage of that task. Its objective is to derive entropy weights that rank the twelve STEM–CBE preparedness indicators by their informational value for public senior secondary schools in Bungoma County, as a preliminary step to entropy-weighted clustering, and thereby to identify the indicators that most sharply differentiate school readiness.

II. STATEMENT OF THE PROBLEM

More than 1.1 million learners transitioned into Kenya’s senior secondary schools in the 2026 intake, and the schools receiving them differ substantially in resource endowment [4, 17].

In Bungoma County the roughly 235 public senior schools spread across nine sub-counties show documented disparities in infrastructure between urban and rural areas, yet the Ministry of Education’s categorisation of schools (C1–C4) is based largely on historical academic performance and enrolment size and does not reflect a school’s multidimensional readiness to deliver the STEM pathway.

In the absence of an objective weighting of preparedness indicators, readiness assessments default to equal weighting or to administrative convenience. This produces two concrete problems. First, indicators that vary little across schools are allowed to count as much as indicators that vary sharply, so the resulting scores blur rather than reveal the real differences between schools.

Second, the school typologies built on such scores are non-replicable and cannot reliably separate schools that need systemic intervention from those that need only targeted technical assistance, so policy responses become crude and inefficient in their use of scarce resources.

There is therefore a methodological gap: no objective, entropy-based scheme has been established to rank STEM–CBE preparedness indicators by their informational contribution for Kenyan senior schools.

This study addresses that gap by deriving, ranking and interpreting entropy weights for the twelve preparedness indicators, providing a defensible and reproducible weight vector on which readiness profiling can be built.

III. THEORETICAL FRAMEWORK

The study is grounded in Shannon’s information theory and its extension into objective multi-criteria weighting.

3.1 Information entropy.

Shannon [7] introduced entropy as a measure of the uncertainty, or information content, of a probability distribution. Applied to an indicator measured across schools, entropy is maximised when every school carries an equal share of the indicator (no discrimination) and minimised when the indicator is concentrated in a few schools (strong discrimination).

Entropy thus provides a formal, distribution-based criterion for how much an indicator contributes to distinguishing observations from one another.

3.2 Objective (entropy) weighting.

The entropy weighting method developed within multi-criteria decision analysis [8, 9] converts each indicator’s entropy into a weight: indicators with lower entropy (greater dispersion, more information) receive larger weights, and the weights are normalised to sum to one. Because the weights are derived entirely from the data rather than from expert judgement, the method is objective and reproducible, which is precisely the property required for policy-relevant school typologies [6, 10].

This framework supplies the derivation logic used here: normalise the indicators, quantify each indicator's dispersion by its Shannon entropy, and translate that dispersion into a unit-sum weight vector that later governs cluster formation.

IV. LITERATURE REVIEW

4.1 Entropy weighting and objective weight derivation

Entropy weighting has an established record as an objective weighting scheme across applied domains. Chen [6] combined entropy weights with a decision model for building-material supplier selection and showed that the data-driven weights corrected the biases of purely subjective weighting.

Rohlfesen and colleagues [10] reviewed entropy-based reasoning in clinical decision-making and emphasised its value where the informational contribution of competing criteria is unequal and must be quantified rather than assumed.

These applications establish entropy weighting as a general-purpose method for ranking indicators by their informational content, but none addresses school-preparedness indicators.

4.2 Clustering and indicator weighting in education

Statistical clustering has been used to reclassify schools by resource profile. Xia, Chen and Zhang [11] applied clustering to educational-resource evaluation in Guangdong Province and found that data-driven cluster assignments allocated resources more effectively than the existing administrative classification, with digital capacity and management commitment receiving the highest weights.

Xiao, Wang and Liu [12] used fuzzy c-means clustering to evaluate facility readiness in Tibetan secondary schools and reported that clustering-based categories explained substantially more variance in outcomes than conventional administrative categories.

Both studies show that the indicators receiving the greatest weight tend also to be those that most

sharply separate schools—a pattern the present study tests directly through the entropy weights.

4.3 CBE preparedness in Kenya and Sub-Saharan Africa

Evidence from the region documents wide implementation gaps and a reliance on descriptive rather than multivariate methods. Shitindi [13] found that a majority of rural Tanzanian secondary schools lacked the laboratory equipment for science-pathway competencies, while Mugiraneza and Maniraho [14] reported uneven teacher preparedness for CBE delivery across Rwandan provinces.

In Kenya, a national readiness assessment by the Kenya Institute of Curriculum Development [5] found sizeable shares of schools without adequate science-laboratory and ICT provision; Otieno [15] reported that few secondary schools had functional laboratories meeting ministry specifications; and Wafula [16] showed that the C1–C4 category was a poor predictor of infrastructure readiness.

Sagwa, Odebero and Nganyi [4] projected persistent infrastructure and ICT deficits at the senior-school level ahead of the 2026 transition. These studies confirm both the multidimensional nature of preparedness and the methodological gap: they rely on univariate descriptive statistics and do not derive objective, information-based weights for combining indicators.

4.4 Research gap

Across the reviewed literature, preparedness indicators are typically reported one at a time or combined with equal or expert-assigned weights, and no study derives objective, entropy-based weights for STEM–CBE preparedness indicators in Kenyan senior schools.

The present study addresses this gap by applying a transparent six-step Shannon-entropy procedure to twelve indicators, reporting the entropy, diversification and weight of each, and ranking them by informational value.

V. RESEARCH METHODOLOGY

5.1 Design, population, and sample

A quantitative, cross-sectional survey design was adopted. The target population comprised the public senior secondary schools in Bungoma County offering the STEM pathway, approximately 235 in number. The sample size was determined using Yamane’s finite-population formula [18],

$$n = N / (1 + Ne^2),$$

with population $N = 235$ and margin of error $e = 0.05$, giving $n \approx 148$ schools. Schools were selected by stratified random sampling with proportionate allocation across three simultaneous criteria—sub-county (nine strata), urbanisation (urban, peri-urban, rural) and STEM-pathway provision—and systematic random selection within strata.

Of the 148 sampled schools, 112 returned a complete questionnaire paired with a valid facility observation checklist, an overall response rate of 75.7 percent, which exceeds the 70 percent benchmark commonly accepted for descriptive social-science research.

Data were collected at school level using a structured questionnaire administered to the principal or deputy principal and a facility observation checklist completed on site by a trained research assistant.

5.2 Preparedness indicators

Twelve indicators spanning five preparedness dimensions were operationalised (Table 1): three infrastructure indicators, two human-resource indicators, three pedagogical and curriculum-alignment indicators, two ICT-integration indicators, and two institutional-leadership indicators.

The instruments used a combination of five-point Likert scales, count scales and categorical formats. Prior to analysis the data passed through a four-stage preprocessing pipeline—editing, error detection and correction, imputation of missing values, and standardisation—after which the cleaned 112×12 matrix was min–max normalised in preparation for entropy-weight computation.

The twelve STEM–CBE preparedness indicators by dimension

Variable	Indicator	Operational definition
<i>Dimension 1: Infrastructure and facilities</i>		
X1	Laboratory resource adequacy	Availability and adequacy of laboratory equipment, materials and consumables.
X2	Library and reference adequacy	Availability of STEM textbooks, journals and digital reference materials per learner.
X3	STEM workshop availability	Presence and operational status of a workshop or makerspace for STEM projects.
X4	STEM teacher density	Number of qualified STEM teachers per 100 learners.
X5	Teacher CBE retraining	Proportion of STEM teachers retrained in competency-based methods.
X6	Practical lesson frequency	Number of hands-on lessons conducted per week.
X7	Project-based learning adoption	Proportion of CBE assessments delivered through STEM projects.
X8	Student competency in STEM tasks	Mean assessment scores in experiments, projects and practical work.
X9	ICT device-to-learner ratio	Number of functional ICT devices per learner.
X10	ICT-integrated lesson frequency	Proportion of weekly lessons in which ICT is actively integrated.
X11	STEM strategic plan and budget	Existence of a STEM strategic plan with a dedicated budget line.
X12	Stakeholder engagement	Number of active partnerships supporting STEM delivery.

5.3 Entropy-weight derivation procedure

To the normalised data, an entropy-based weighting model [7, 8] was applied in six sequential steps. First, each entry z_{ij} of the raw $n \times p$ matrix is re-scaled to a common [0,1] range by min-max normalisation,

$$\tilde{z}_{ij} = (z_{ij} - \min_i z_{ij}) / (\max_i z_{ij} - \min_i z_{ij}),$$

so that no indicator dominates the computation merely because of its numeric scale. Second, for each indicator j the normalised score of school i is expressed as a share of the column total,

$$p_{ij} = \tilde{z}_{ij} / (\sum_{i=1}^n \tilde{z}_{ij}), \quad \sum_{i=1}^n p_{ij} = 1,$$

so that each column forms a probability distribution over schools. Third, the diversity of indicator j across schools is measured by its normalised Shannon entropy,

$$E_j = -(1/\ln n) \sum_{i=1}^n p_{ij} \ln(p_{ij}), \quad 0 \leq E_j \leq 1,$$

with the convention $0 \ln 0 \equiv 0$; the constant $1/\ln n$ rescales the entropy to [0,1], where $E_j = 1$ denotes maximum diversity and low discrimination, and a low E_j denotes strong discrimination. Fourth, the degree of diversification (information utility) inverts the entropy,

$$D_j = 1 - E_j, \quad 0 \leq D_j \leq 1,$$

so that larger values denote stronger discriminating power. Fifth, the diversification values are normalised into a unit-sum weight vector,

$$w_j = D_j / (\sum_{j=1}^p D_j), \quad w_j \geq 0, \quad \sum_{j=1}^p w_j = 1.$$

Sixth, the unit-sum property is verified numerically, and the weight vector $w = (w_1, \dots, w_p)^T$ is retained for the downstream entropy-weighted clustering. All analyses were conducted in R [20].

VI. FINDINGS

6.1 Descriptive Distribution of the Indicators

Before the entropy weights were derived, the twelve indicators were summarised as collected from the field, prior to normalisation, because it is the dispersion of each indicator across schools that the

entropy procedure converts into a weight. Table 2 reports the mean, standard deviation, minimum, maximum and skewness of each indicator for the 112 responding schools.

Summary statistics for the twelve preparedness indicators ($n = 112$)

Va	Indicator	Mea	SD	Min	Ma	Ske
r.	(unit)	n			x	w
X1	Laboratory resource adequacy (score /5)	2.42	0.91	1.00	4.50	0.42
X2	Library and reference (books/learner)	0.63	0.38	0.10	2.10	1.05
X3	STEM workshop availability (score /5)	1.84	1.22	0.00	4.00	0.51
X4	STEM teachers per 100 learners	1.72	0.83	0.40	4.20	0.68
X5	Teacher CBE retraining (%)	34.6	21.5	0.00	92.0	0.44
X6	Practical lessons per week	3.18	1.72	0.00	8.00	0.29
X7	Project-based assessments (%)	22.3	18.4	0.00	75.0	0.83
X8	Student STEM task score (out of 100)	52.8	12.9	24.0	84.0	0.11
X9	ICT devices per learner	0.05	0.06	0.00	0.31	1.94
X10	ICT-integrated lessons (%)	12.7	11.8	0.00	58.0	1.42
X11	STEM strategic	0.28	0.45	0.00	1.00	0.98

Var.	Indicator (unit)	Mean	SD	Min	Max	Skewness
X12	plan + budget (0/1)	1.42	1.61	0.00	8.00	1.31
X11	Active stakeholder partnerships (count)	1.42	1.61	0.00	8.00	1.31

As Table 2 shows, the indicators differ markedly in their dispersion. The ICT indicators are the most strongly skewed, with ICT devices per learner (X9) recording a skewness of 1.94 and ICT-integrated lessons (X10) a skewness of 1.42, indicating that most schools cluster near very low values while a small number report much higher digital penetration.

The leadership indicators also disperse widely, with stakeholder partnerships (X12) skewed at 1.31 and the strategic-plan indicator (X11) at 0.98. In contrast, student STEM task scores (X8) are almost symmetric (skewness 0.11) and tightly bunched around a mean of 52.80, and practical lesson frequency (X6) shows only mild skew (0.29). The infrastructure indicators sit at moderate means well below their scale ceilings.

The pattern indicates that schools differ from one another far more on ICT provision and institutional leadership than on pedagogical-delivery measures such as student task scores and practical lesson frequency. Where an indicator is tightly bunched, most schools look alike on it, so it carries little power to distinguish them; where an indicator is widely dispersed and strongly skewed, it separates schools sharply.

The descriptive distribution therefore anticipates the entropy result: the highly dispersed ICT and leadership indicators should attract the largest weights, and the tightly bunched pedagogical indicators the smallest.

This is consistent with regional evidence that digital infrastructure is the least uniformly provided input in the run-up to the CBE senior-school transition [4, 5], and it confirms that the indicator set carries the informational variability the entropy procedure requires.

The near-symmetry of the student-competency measure echoes findings that pedagogical outcomes are constrained similarly across schools while structural inputs vary sharply [15, 16].

6.2 Derived Entropy Weights and Indicator Ranking
 The six-step procedure was applied to the cleaned and normalised 112×12 matrix. Because the normalisation and proportion steps produce large intermediate matrices of no direct interpretive value, only the downstream products are reported.

Table 3 gives, for each indicator, the normalised Shannon entropy E_j , the degree of diversification $D_j = 1 - E_j$, the entropy weight w_j , and the resulting rank. Entropy weights and ranking of the twelve preparedness indicators

Var.	Indicator	E_j	D_j	w_j	Rank
X9	ICT devices per learner	0.62	0.38	0.141	1
X10	ICT-integrated lesson frequency	0.68	0.32	0.119	2
X12	Stakeholder engagement	0.70	0.30	0.112	3
X2	Library and reference adequacy	0.71	0.29	0.108	4
X11	STEM strategic plan and budget	0.72	0.28	0.104	5
X7	Project-based learning adoption	0.76	0.24	0.089	6
X3	STEM workshop availability	0.79	0.21	0.078	7
X5	Teacher CBE retraining	0.81	0.19	0.071	8
X1	Laboratory resource adequacy	0.83	0.17	0.063	9
X4	STEM teacher density	0.86	0.14	0.052	10
X6	Practical lesson frequency	0.89	0.11	0.041	11
X8	Student competency in STEM tasks	0.94	0.06	0.022	12

Var.	Indicator	E_j	D_j	w_j	Rank
Total		2.69	1.000		

As Table 3 records, ICT devices per learner (X9) received the highest weight at 0.141, followed by ICT-integrated lesson frequency (X10) at 0.119 and stakeholder engagement (X12) at 0.112. Library and reference adequacy (X2), the STEM strategic plan and budget (X11), and project-based learning adoption (X7) formed a middle band with weights between 0.089 and 0.108.

STEM teacher density (X4), practical lesson frequency (X6) and student competency in STEM tasks (X8) received the lowest weights, from 0.052 down to 0.022. The weights ranged from 0.022 to 0.141, a spread of about six-fold, and summed to 1.000, confirming the unit-sum property of the derivation. The two ICT indicators together with the two leadership indicators (X9, X10, X11, X12) accounted for about 47.6 percent of the total informational value.

The ranking identifies ICT provision as the single strongest discriminator of preparedness across Bungoma senior schools, pointing to a wide and uneven digital divide, with institutional leadership close behind, showing that governance capacity separates schools as sharply as physical inputs.

The low weights on student competency (X8) and practical lesson frequency (X6) do not mean these indicators are unimportant to STEM delivery; they mean that schools vary little on them, so they contribute little to telling schools apart.

The ordering of the weights mirrors the dispersion seen in the descriptive statistics, where X9 was the most skewed indicator and X8 the least, giving an internal coherence check on the derivation.

The prominence of the ICT indicators fits projection evidence that digital and ICT infrastructure are among the most under-provisioned inputs for Kenya's senior-secondary transition [4], and the high weight on stakeholder engagement matches the emphasis on multi-actor governance in national CBE policy [3].

The result also echoes entropy-weighted school-evaluation work elsewhere: Xia, Chen and Zhang [11] found that digital capacity and management commitment attracted the highest weights in a Chinese secondary-school study, and Xiao, Wang and Liu [12] reported the same alignment between weight and discriminating power in Tibetan schools.

That an objective, information-theoretic procedure independently selects ICT and leadership as the decisive indicators strengthens confidence that these dimensions, rather than pedagogical-outcome measures, should anchor readiness assessment for the STEM pathway.

VII. CONCLUSION

Applying a six-step Shannon-entropy procedure to twelve STEM-CBE preparedness indicators for 112 public senior secondary schools in Bungoma County produced an objective, unit-sum weight vector that ranks the indicators by their informational value.

ICT device provision and ICT-integrated lesson frequency were the two most informative indicators, and together with stakeholder engagement and the presence of a STEM strategic plan and budget the ICT and institutional-leadership indicators carried close to half of the total informational value.

Student competency in STEM tasks, practical lesson frequency and STEM teacher density was least informative, not because they matter little to STEM delivery but because schools vary little on them.

The weight ordering tracks the dispersion observed in the raw indicators, confirming that the derivation faithfully reflects the data, and the six-fold spread of the weights shows that treating all indicators as equally important would materially distort any readiness measure built from them.

The study concludes that the digital divide and governance capacity are the sharpest markers of senior-school readiness for the STEM pathway in the county, and it supplies a defensible, reproducible weight vector on which entropy-weighted preparedness profiling can be built.

Two cautions apply: the weights are cross-sectional and may shift as ICT infrastructure expands and teacher retraining widens, and entropy weighting rewards variability, so any variability arising from measurement noise rather than substantive difference would be reflected in the weights.

VIII. RECOMMENDATIONS

Arising from the derivation, the study makes a single principal recommendation: the Ministry of Education and the County Directorate of Education should adopt an information-weighted assessment framework—anchored on the entropy-derived ranking, in which ICT provision and institutional-leadership indicators carry the greatest weight—when monitoring school readiness for the STEM pathway under Competency Based Education, rather than an equal-weight or category-based scheme.

Because equal-weight tools overweight the indicators that least discriminate between schools and underweight those that separate them most sharply, adopting the entropy-weighted ranking would direct monitoring attention, and the scarce resources that follow it, toward the digital-infrastructure and governance gaps that most sharply differentiate school readiness across the county.

REFERENCES

- [1] UNESCO, Education for Sustainable Development Goals: Learning Objectives, United Nations Educational, Scientific and Cultural Organization, Paris, 2017.
- [2] J. Ngure, From compliance to curiosity: How Kenya's shift to CBE can redefine learning, Women Educational Researchers of Kenya (WERK), June 2025.
- [3] Republic of Kenya, Ministry of Education, Sessional Paper on Reforming Education and Training for Sustainable Development in Kenya, Government of Kenya, Nairobi, 2024.
- [4] M. D. Sagwa, S. Odebero, and J. Nganyi, Projection models for selected infrastructural requirements for the implementation of competency-based curriculum in senior secondary schools in 2026, in Kenya, International Journal of Education and Research 12 (2024), no. 11.
- [5] Kenya Institute of Curriculum Development, National Readiness Assessment for Junior Secondary School Implementation, KICD, Nairobi, 2021.
- [6] C.-H. Chen, A novel multi-criteria decision-making model for building material supplier selection based on entropy-AHP weighted TOPSIS, Entropy 22 (2020), no. 2, 259.
- [7] C. E. Shannon, A mathematical theory of communication, Bell System Technical Journal 27 (1948), no. 3, 379–423.
- [8] M. Zeleny, Multiple Criteria Decision Making, McGraw-Hill, New York, 1982.
- [9] C. L. Hwang and K. Yoon, Multiple Attribute Decision Making: Methods and Applications, Springer-Verlag, Berlin, 1981.
- [10] C. Rohlfen, K. Shannon, and A. S. Parsons, Entropy in clinical decision-making: A narrative review through the lens of decision theory, Journal of General Internal Medicine 40 (2025), no. 16, 4033–4039.
- [11] L. Xia, Y. Chen, and H. Zhang, Clustering methods for educational resource evaluation in primary schools: A case study of Guangdong Province, Journal of Educational Data Mining 14 (2022), no. 2, 87–109.
- [12] T. Xiao, R. Wang, and J. Liu, Evaluating facility readiness in Tibetan secondary schools using fuzzy c-means clustering, International Journal of Educational Research 113 (2022), 101936.
- [13] E. Shitindi, School readiness for competency-based curriculum in Tanzania's secondary schools, African Educational Research Journal 9 (2021), no. 1, 1–11.
- [14] T. Mugiraneza and J. F. Maniraho, Teacher preparedness for competency-based curriculum implementation in Rwandan secondary schools, African Journal of Education and Practice 6 (2020), no. 3, 1–14.
- [15] J. O. Otieno, Assessment of physical infrastructure readiness for STEM pathway delivery in Kenyan secondary schools, East

African Journal of Education 7 (2020), no. 2,
23–38.

- [16] N. W. Wafula, Competency Based Curriculum Infrastructure Readiness in Trans-Nzoia County Secondary Schools: A Cross-Sectional Study, Master's thesis, Kibabii University, 2022.
- [17] Presidential Working Party on Education Reform, Report of the Presidential Working Party on Education Reform: Transforming Education, Training and Research for Sustainable Development in Kenya, Republic of Kenya, 2023.
- [18] T. Yamane, Statistics: An Introductory Analysis, 2nd ed., Harper and Row, New York, 1967.
- [19] P. J. Rousseeuw, Silhouettes: A graphical aid to the interpretation and validation of cluster analysis, *Journal of Computational and Applied Mathematics* 20 (1987), 53–65.
- [20] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2024.