# DIAGNOSIS OF THE QUALITY OF BRAZILIAN BASIC EDUCATION IN MUNICIPAL NETWORKS

## Pedro Henrique Alves Barros[1], Regina Serrão Lanzillotti[2].

**ABSTRACT**

Motivation punctuates Brazilian basic education in municipal networks, Saeb (2021). The objective is to verify the proficiency of the average values in Mathematics, Portuguese Language and Literacy Rate, between states in municipal and private schools, for the capital and interior. The mean, mode, and median were calculated to evaluate symmetry, kurtosis, and coefficient of variation, with the histogram for graphical visualization. Covariances for pairs of disciplines were obtained in relation to autarchies and regionalities both in the capital and in the interior. The test of equality of covariance measured the similarity in the performance of the disciplines. The scatter plot between Euclidean and Mahalanobis Distances for the Literacy Rate by state showed sensitivity that the Southeast and Midwest regions.

**Keywords:** Test of Covariance, Distance from Mahalanobis, Euclidean distance, Statistics, Performance.

[1] Master of Science in Mathematics
State University of Rio de Janeiro (UERJ)
Email: pedro.barros@pos.ime.uerj.br
[2] Doctor in Transportation Engineering
Coordination of Graduate and Engineering Research Programs (COPPE - UFRJ)
Email: reginalanzillotti@ime.uerj.br

## INTRODUCTION

Basic education is the first stage in the formation of the individual. This stage has divisions, each with its objectives and methodologies, which are always being evaluated and reevaluated to keep up with the changes experienced by the student generations. Education has a central role in the nation's project, because "[...] alone does not transform society, without it society does not change either" (Freire, 2000), but it also acts by constituting "[...] broader social relations, thus contributing, contradictorily, to the transformation and maintenance of these relations." Brazilian basic education needs to be constantly analyzed, so that there is a quality education offered. Interpreting the different levels of learning in Elementary School in countries of continental dimensions, Brazilian territory, becomes challenging. It is estimated that there are 178 thousand schools that aggregate more than 47 million enrollments, which is a complex but necessary task, as the results allow directing investments to regions of greater demand. One of the alternatives for diagnosing Brazilian basic education has been large-scale external evaluations. The evaluation model in question emerged in Brazil in the 1990s, aiming to assess the quality of teaching, being prepared and carried out by the Ministry of Education. This is one of the means used as a tool to map the positive and negative aspects, indicating the possible difficulties encountered, especially in literacy, thus measuring the results obtained by groups of students, providing inputs for constant monitoring, planning, guidance and intervention in the educational process. Luckesi cites that:

> Evaluation can be characterized as a way of judging the quality of the evaluated object, a factor that implies taking a position on it, to accept or transform it. Evaluation is a value judgment on relevant manifestations of reality, with a view to decision-making. (LUCKESI, 2000, p. 33)

From then on, other external evaluations emerged, such as the National Literacy Assessment (ANA), the National High School Exam (Enem) and the National Student Performance Exam (Enade).

One of the first external evaluations in education was the Basic Education Evaluation System (SAEB), which aimed to measure student proficiency in the initial grades in Portuguese Language and Mathematics. The exam is a valuable instrument in decision-making aimed at achieving better quality of education through these results. the National Institute for Research and Studies (Inep) cites that:

> The SAEB allows schools and municipal and state education networks to evaluate the quality of education offered to students. The result of the evaluation is an indication of the quality of Brazilian education and offers subsidies for the elaboration, monitoring and improvement of evidence-based educational policies. (INEP/MEC, 2019)

Since 1990, the SAEB has undergone methodological changes, as education has undergone pedagogical and administrative transformations, the most recent being the implementation of the National Common Curriculum Base (BNCC), one of the main documents that deals with Brazilian basic education. The SAEB census evaluation maps education in seven dimensions: school attendance, teaching and learning, investment, education professionals, management, equity, and human rights. In 2007, Inep determined the Basic Education Index that associates the grades of students in the SAEB, with the rates of approval, failure and dropout, from the School Census.

In constant improvements, especially from 2019 onwards, there is a restructuring of the exam to adapt to the BNCC, as the matrices of Portuguese Language and Mathematics were inserted in the 2nd year of Elementary School, while Human Sciences and Natural Sciences for the 9th year and this school segment, which as of 2023, these subjects were included from the 5th year. The 2021 exam was considered a pilot study, an exam with the most recent publication of results, available by Inep, which discriminate literacy, language and mathematics rates at the level of municipal, state, federal and private public schools, also by state and national aggregation. The grades of Mathematics and Portuguese Language and of the other subjects are determined by the pre-stipulated proficiency scales for the exam, so there is an association in relation to the performance scale in line with the skills and competencies present in the BNCC, considered acquired or developed. The literacy rate corresponds to the relative index of the proportion of students who obtained a satisfactory level of literacy according to the National Literacy Assessment (ANA), such an exam is carried out annually to assess the level of literacy and literacy in Portuguese and literacy in Mathematics of children regularly enrolled in the 3rd year of elementary school.

It is also worth adding that according to Creso Franco:

> [...]    Studies that test students only once have important limitations to explain the effect of schools on learning. The reason for this lies in the fact that the proficiency measure is a measure of student learning over many years. Therefore, this measure cannot be explained as a function of the students' recent past (variables of the school that the student attended in the SAEB year). It is for this reason that longitudinal studies are considered the "gold standard" of school and educational system evaluation studies. In longitudinal studies, the measure of prior proficiency is used to filter the effect of the student's prior learning. Thus, the second measured

proficiency (after filtering previous learning) represents learning throughout the year. Certainly, this learning throughout the year can be explained by the characteristics of the school attended. This is the main limitation of the SAEB for the explanation of the factors that influence learning. (Creso Franco, 2001)

Even so, it is believed that the global analysis of information can be revealing regarding student performance, as Alavarse indicates:

> Knowing and using the results of external evaluations in the classrooms and comparing them with internal evaluations means understanding them not as an end in itself, but rather as a possibility of associating them with the necessary transformations in order to strengthen the quality of the democratic public school, which is the one that is organized to guarantee the learning of all (ALAVARSE, 2013, p. 75).

In post-pandemic times, an analysis of the scenario revealed in 2021, the most recent, can be demonstrative and indicative for future policies to combat illiteracy, because, due to the need for schools to remain closed due to health issues, it was found that many students, even with the efforts of the respective educational authorities, were harmed in the teaching and learning process in that period, carrying a gap that will last in the following school years. This proposal aims to map the divergence of the realities of public and private education in terms of student performance in Elementary Education at the state level, using as a database the results provided by Inep of Saeb 2021, the year considered the most lethal of Covid-19 in Brazil. At that critical moment, when the country ended the year with 619,109 deaths due to the pandemic, the restrictive health measures determined teaching, initially remote, and later hybrid. Article 80 of the LDB ensures:

> [...] the Government shall encourage the development and dissemination of distance learning programs, at all levels and modalities of education, and of continuing education", whose paragraph 1 defines that "distance education, organized with special openness and regime, shall be offered by institutions specifically accredited by the Union (Brasil, 1996).

Distance Education (DE), currently regulated by Decree No. 9,057/2017, is applicable to both Higher Education and Basic Education courses. However, in 2021, during the pandemic, society showed little confidence in this teaching method for basic education, mainly due to the immense social inequality in the country, where a considerable portion of the population did not even have access to basic personal hygiene utensils. Cunha and Ferrari quote:

> [...] It is important to highlight that the pandemic only accentuated vulnerabilities that were already present in the Brazilian social fabric. At least since 2014, although with more emphasis from 2016 onwards, we have not only gone through the exhaustion of the expansive economic cycle (2004 to 2014) but, mainly, through the economic recession (2014 to 2016) and economic stagnation (2016 to the present day), already defined by some authors as "stagflation" (Cunha; Ferrari, 2021).

Due to this scenario, it can be seen that the pandemic exerted a disproportionate influence on the various social strata, exerting a greater impact, especially on the less favored classes, and those more dependent on public services, especially in the provision of public education. This study is, therefore, also a way to verify the effectiveness in complying with article 206 of the Federal Constitution of 1988, which guarantees Education a duty of the State and of the family, and a right of all.

## OBJECTIVES

The general objective of this work is to diagnose the level of proficiency of students in the geographic structure both intra and inter at the state level regarding the subjects of Mathematics, Portuguese Language and Literacy in public and private institutions through the results of Saeb 2021, the most recent released by Inep at the time of completion of this work.

The specificity is concentrated in the sole paragraph of article 211 of the Federal Constitution, which dictates that the City Hall of the municipalities will act primarily in Elementary Education and Early Childhood Education, while High Schools under the management of the States of the federation. This taxonomy of public education was confronted both for private schools in the capital and those in the interior. to that of private schools both in the capital and in the countryside. In this way, it will be possible to measure the differences in learning geographically and according to the autarchy, which reflects the degree of dissimilarity in the acquisition of acquired knowledge to consciously provide opportunities for the Ministry of Education's investments.

## THEORETICAL FRAMEWORK

Initially, a literature review was carried out with articles published from 2017 to 2024 in the electronic databases: *Scientific Electronic Library Online* - Scielo and Google Scholar, using the descriptors: Saeb. A search was added to the Inep database about the SAEB and its proposals, approaches, history of changes and published results. A consultation of previous evaluations of Mathematics and Portuguese Language was also carried out to

verify the evaluation model applied, in addition to a search for literacy evaluations and literacy criteria stipulated by ANA.

## METHODOLOGY

The most recent results of the SAEB carried out in 2021 are on the Inep website. At the moment, the data for 2023 has not yet been released, it will be in the preliminary phase. This Management Information System - GIS subsidized the descriptive and inferential statistical study for the quantitative years referring to the years of the global pandemic, a critical moment when many students were away from schools, despite remote learning having been offered in many municipal networks, but few students had access to the use of assistive technology and education.

On the Inep website there are the average proficiency values for private education in the capital, but the absence was observed for the state of Ceará. There was an attempt to contact Inep, but without success. Therefore, the state was disregarded in this study. The availability of average performances in Portuguese, Mathematics and Literacy were verified, but the competent bodies did not include measurement of the variability around this average.

Initially, the histograms of the available averages were created, which allowed a visualization of the students' profile in relation to learning. The measures of mean, median, and mode position were obtained to evaluate the occurrence or not of heterogeneity of the disclosed data, and in one distribution it can be considered symmetrical, if these measures are equal (Correa, 2002), but the asymmetric distribution is positive, if the value of the mean is higher than the mode, while in the negative, the mode is higher than the mean. The measure of dispersion, coefficient of variation (CV), can be obtained by the ratio between the standard deviation and the arithmetic mean, with the standard deviation being the square root of the variance.

$$CV = \frac{S}{\bar{x}}, \qquad\qquad (1)$$

where S = standard deviation
$\bar{x}$ = arithmetic mean.

Low dispersion is considered when the VC values are close to zero, because the increase in this measure infers in the expressiveness of the variation around the mean value. The percentile coefficient of kurtosis indicates the degree of flattening of the frequency polygon of the histogram, measured by the ratio between the mean distance of the distance of quartiles 3 and 1, in relation to the distance of the extreme percentiles, 90th percentile, and 10th percentile. (Naghettini; Pinto, 2007). The equation is expressed as follows:

$$K = \frac{(Q_3 - Q_1)/2}{(P_{90} - P_{10})}$$

where (2)

$$Q_3, 3^\circ\ quartil,\quad Q_1, 1^\circ\ quartil,\qquad P_{90}, percentil\ 90,\qquad P_{10}, percentil\ 10.$$

The Normal Probability Distribution has a percentile kurtosis coefficient corresponding to 0.263 when the mesocuritic distribution is classified. If the value is higher, it is called platicuric, if not, leptocurtic.

Cluster analysis is a multivariate technique that aims to provide one or more partitions in the mass of data, in groups, by some classification criterion, in such a way that there is homogeneity within and heterogeneity between groups (Sneath & Sokal, 1973; Mardia et al., 1997) The covariance test was performed to verify the equality in the covariances of the performances in Mathematics, Portuguese Language and Literacy Rate in the comparison of the education offered in the capital and in the countryside, both at the private and public levels, at the municipal level of each state of the federation. The methodology tests the null hypothesis H0 of the equality of covariances.

$$H_0:\ ,\ (3)\textstyle\sum_1 = \sum_2 = \ldots = \sum_k$$

where , represent the covariance matrices of a multinormal -dimensional population system. Let $\sum_1 \sum_2, \ldots, \sum_k kpS_i$ be a non-biased sample estimate of , based on $\sum_i n_i$ degrees of freedom, where $n_{i\ i\ for} = N - 1$ the case of a random sample with $N_i$ vectors observed in the population. If $i - \acute{e}sim H_0$ is true, we have:

$$S = \frac{1}{\sum n_i} \sum_{i=1}^{k} n_i S_i, \qquad (4)$$

where S is the pooled estimate of the covariance matrices. The test statistic:

$$M = \sum n_i \ln|S| - \sum_{i=1}^{k} n_i \ln S_i \qquad (5)$$

Box (1949) showed that if the scale factor

$$C^{-1} = 1 - \frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \left( \sum_{i=1}^{k} \frac{1}{n_i} - \frac{1}{\sum n_i} \right) \qquad (6)$$

is introduced, the quantity M presents a distribution that can be approximated to the Chi-square ($\chi$) with α = degrees of freedom, and as the value of , sample sizes of the groups increases, if all are equal to n, we have:$C^{-1} \, 1/2 \, (k - 1)p(p + 1)n_i n_i$

$$C^{-1} = 1 - \frac{(2p^2 + 3p - 1)(k+1)}{6(p+1)kn} \qquad (7)$$

If, $MC^{-1} < \chi_{\alpha}$(tabulated) the veracity of 0 is accepted$H$.

In a multivariate group for each Federation Unit, similarity measures were carried out with the objective of diagnosing the intensity of interstate discrepancies in student performance within the North, South, Northeast, Southeast and Midwest regions. In clusters it is possible to make an exploratory analysis of similarity through the calculation of distances. Based on this measure, similar individuals are grouped together and the others are placed in separate groups (Aaker et al., 2001). Different scales of magnitude measures can distort reality, and the most appropriate way to verify it is the use of standardized measures in units of standard deviation, which is explained below:

$$z_i = \frac{x_i - \bar{x}}{s}, \qquad (8)$$

where is the observed value of x $z_i$ i on the ratio scale, is the arithmetic mean, the sample standard deviation.$\bar{x} s$

Most methods of group analysis require the analysis of the relationships between them. According to Barroso & Artes (2003) there are two types of similarity: similarity measures (directly proportional to the similarity between objects) and dissimilarity measures (inversely proportional to the similarity between objects). Similarity can be defined as "the measure of correspondence, or similarity, between the objects to be grouped together." (RUN; PAULO and DIAS FILHO, 2007, p.333) Such measures are easily interrelated and transformable among themselves (Bussab et al., 1990). Usually, the measure of similarity or dissimilarity is expressed as a distance or metric function (Doni, 2004). According to Khattree and Naik (2000) and Cruz and Carneiro (2006), among the measures of dissimilarities, the Euclidean distance and the Mahalanobis distance are highlighted, due to their greater use.

The formulation of Euclidean distance (DE) is based purely on spatial points, and disregards the variability between data, is demonstrated by the application of the Pythagorean Theorem in relation to distance, as Euclidean space becomes a metric space.

$$\text{DE}(,) = \vec{x}\vec{y} \sqrt{\sum_{i=1}^{p}(x_i - y_i)^2} \qquad (9)$$

The adoption of the Mahalanobis Generalized Distance (Dm), a distance introduced by the Indian mathematician Prasanta Chandra Mahalanobis in 1936, is based on the variables expressed in units of the standard deviation (z) and the inverse of the covariance matrix S' to identify different patterns, with the formulation estimated according to (Rao, 1952) by:

$$\text{Dm}(z) = \sqrt{z^T . S^{-1} . z}, \qquad (10)$$

where $S^{-1}$ is the inverse of the covariance matrix.

The Mahalanobis distance considers, therefore, the variability within each sampling unit, and not only the central tendency, being a more acceptable measure for group analysis (Riboldi, 1986).

## RESULTS AND DISCUSSIONS

The analysis of student performance for the 2021 SAEB began with the average and median of literacy, Portuguese Language and Mathematics rates in public and private educational institutions in the capital and countryside, with the exception of Ceará, discriminating standard deviation and coefficient of variation. The results obtained for the Literacy Rate, Portuguese Language and Mathematics, table 1, indicated that the best results are found in the capital for private education, with averages between 76.63, 772.34 and 772.60 with respective coefficients of variation 0.26, 0.01 and 0.02. The least expressive average values were found in the capital in the municipal schools for the subjects under analysis, 28.03 for Literacy Rate, 704.04 for Portuguese Language and 724.78 for Mathematics. The highest coefficients of variation for the three subjects are recorded in the municipal public schools in the interior, which reduces the reliability of the average values.

Table 1: Mean, median, and coefficient of variation for Literacy, Portuguese Language, and Mathematics Rates for municipal schools in the capital and interior of Brazilian states, SAEB 2021. (more)

| Discrimination | Municipal School | | Private School | |
|---|---|---|---|---|
| | Capital | Interior | Capital | Interior |
| Literacy rate | | | | |
| Average | 28,03 | 28,84 | 76,63 | 74,16 |
| Median | 27,56 | 25,36 | 79,34 | 74,86 |
| Standard deviation | 10,10 | 13,38 | 19,73 | 10,04 |
| Coefficient of Variation | 0,36 | 0,46 | 0,26 | 0,14 |
| Portuguese language | | | | |
| Average | 704,04 | 705,37 | 772,34 | 766,56 |

Table 1: Mean, median, and coefficient of variation for Literacy, Portuguese Language, and Mathematics Rates for municipal schools in the capital and interior of Brazilian states, SAEB 2021. (conclusion)

| Discrimination | Municipal School | | Private School | |
|---|---|---|---|---|
| | Capital | Interior | Capital | Interior |
| Portuguese language | | | | |
| Median | 705,75 | 702,98 | 770,46 | 765,53 |
| Standard deviation | 16,46 | 20,11 | 15,12 | 12,25 |
| Coefficient of Variation | 0,02 | 0,03 | 0,01 | 0,02 |
| Mathematics | | | | |
| Average | 724,78 | 726,23 | 772,60 | 764,92 |
| Median | 722,41 | 724,29 | 777,08 | 761,76 |
| Standard deviation | 18,92 | 19,95 | 19,17 | 14,20 |

| Coefficient of Variation | 0,03 | 0,03 | 0,02 | 0,02 |
|---|---|---|---|---|

Source: The author.

An option to evaluate the dispersion around the mean values simultaneously corresponds to the calculation of the combined covariance of dimensionality four, constructed according to the covariances obtained for the taxonomy referring to the capital and interior both at the municipal and private public levels, for the disciplines Mathematics, Portuguese Language and Literacy Rate.

According to (Morrison, 1971), the value of the scale factor ($^{C-1}$) was 0.92, while for the test statistic (M), the value was 26.15. The product of these statistics resulted in 24.15, which compared to the value 31.14 associated with the Chi-square value for 20 degrees of freedom at the 5% significance level, configures the equality of covariances between the performances of Mathematics, Literacy Rate and Portuguese Language. This result represents that the grades, regardless of the disciplines, vary in a similar way regardless of the region and autarchy. In other words, the statistical analysis of Literacy Rates can represent the learning scenario in the other subjects as well. This finding can be questioned why literacy rates are below the assessment of Mathematics and Portuguese Language, but it should be noted that literacy rates require students to achieve the minimum expected literacy according to the ANA. However, this does not inform about the observed veracity of the evaluative discrepancies observed in Table 1.

It is necessary to obtain the covariances according to the pairs of disciplines, both for public and private and also at the regional level.  Similarly, in the previous procedure, the value of the scale factor ($^{C-1)\ was\ obtained}$, which was equal to 0.95, while the "M" Test Statistic, the value was 37.98. The product of these statistics resulted in 35.92, which compared to the value 28.87 associated with the Chi-square value for 18 degrees of freedom at the 5% significance level, configures the inequality of covariances between the performances of Mathematics, Literacy Rate and Portuguese Language. This result shows that depending on the region and autarchy, school performance becomes differentiated. It is interpreted that the first Combined Covariance test did not inform the reliability, but perhaps the perpetuation of the scenario, because if there is a deficiency in at least one of the disciplines, it may affect the others. The performance of post-literacy students does not evolve equally, even the adequate measurement of the literacy rate is a precursor to the expected proficiency of students.

The percentile coefficient of kurtosis is obtained according to the ratio between the mean distance of the quartiles in relation to the distance of the extremes, 90th percentile and 10th percentile. If this value tends to zero, it indicates that the variability around the median value is small, if not the higher the coefficient, the greater heterogeneity is observed in the observed values. The Literacy Rates of the municipal schools in the capital and in the interior presented values of this coefficient of 0.309 and 0.289, respectively, indicating platycurtic distributions, that is, there is no expressive variability around the median, which presented values of 27.56 and 25.36, respectively, not distancing from the average values for the capital evaluated at 28.03, but for the interior, The average is above this value, 28.84, which configures a tendency to asymmetry. In this case, the capital showed an asymmetry coefficient of 0.78, describing an asymmetric distribution to the right, which implies that most students are below the average.

On the other hand, the distribution of the Literacy Rate in private schools in the capital and in the interior showed percentile kurtosis of 0.247 and 0.207, respectively, not far from the kurtosis of the normal distribution, in the order of 0.263. Regarding the asymmetry, the results for the interior showed values close to the means, which can be inferred that these results adhere to the normal distribution. In the case of the capital, the value was -0.23, configuring an asymmetric distribution to the left, which shows that most students are above average.

The analysis at the state level allowed us to consider the state of Piauí as a discrepant in relation to the average for the Literacy Rates of the municipal schools in the capital (figure 1), and this state stood out at the national level, practically double the average. Despite the surprise, it becomes a relevant fact, which leads to the questioning of the literacy technique.

Figure 1: Boxplot graph in relation to the average for the Literacy Rates of the Capital.
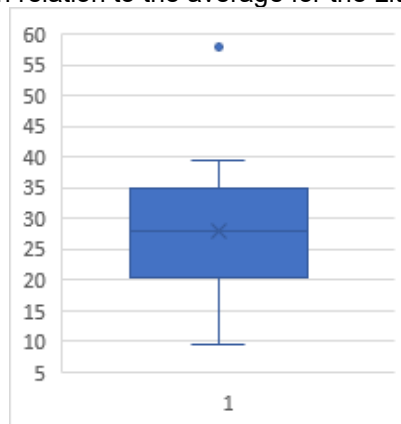


Source: The author.

One of the objectives is to evaluate the discrepancies in the results of Literacy at the state level, grouping the states at the regional level, North, Northeast, South, Southeast and Midwest, opting for the Euclidean Distance and Mahalanobis Distance that were plotted in the scatter plot for each region. It should be noted that the Euclidean Distance is obtained in relation to the nearest neighbor, while the Mahalanobis distance considers the variability according to the covariance matrix. Figure 2 shows the growth for the North region.

Figure 2: Scatter plot of the discrepancy in the Literacy Rate among states in the North Region.
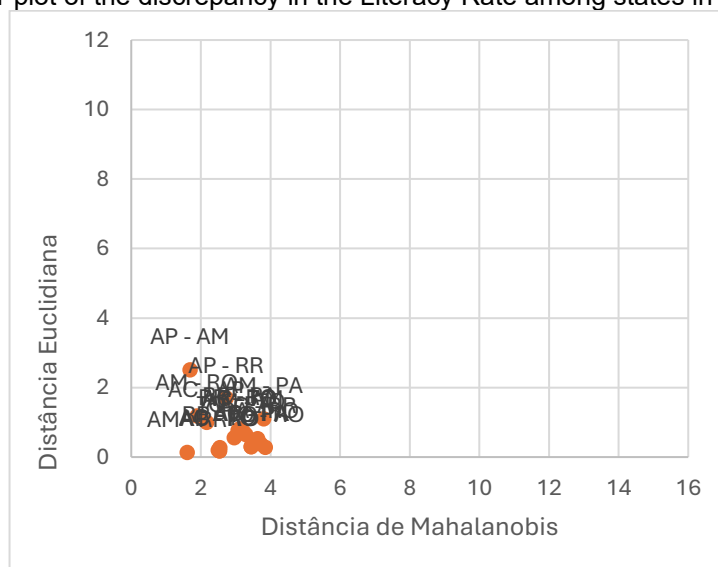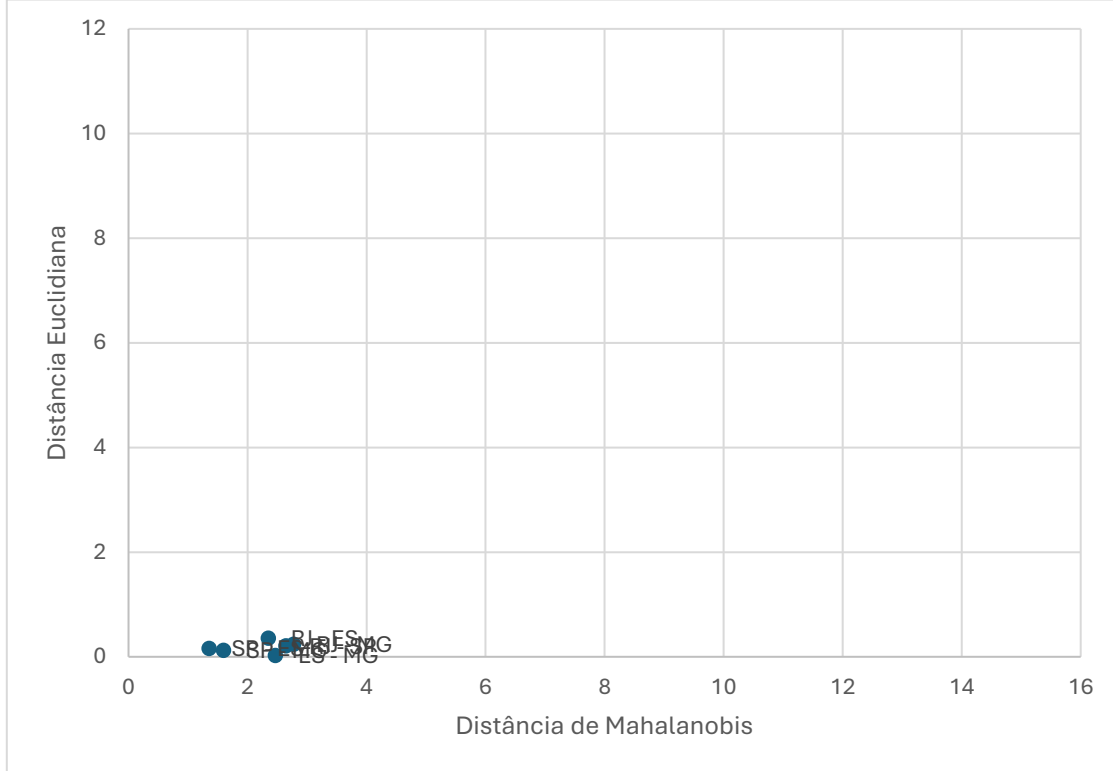


Source: The author.

The states of Amazonas and Pará should be separated from the other groupings, with Amazonas tending to aggregate to the states in the lower left part of the figure, and Pará tending to aggregate to the states in the lower right part of the graph, having smaller Euclidean Distances. From the observed, these are the ones that contained the greatest

discrepancy in the pairs ordered as a function of the Euclidean and Mahalanobis distances, and the Mahalanobis distance contemplates the influence of variability around the mean. Figures 3, 4 and 5 show the scatter plots for the states of the Southeast, South and Midwest regions, respectively.

Figure 3: Scatter plot of the discrepancy in the Literacy Rate among states in the Southeast Region.
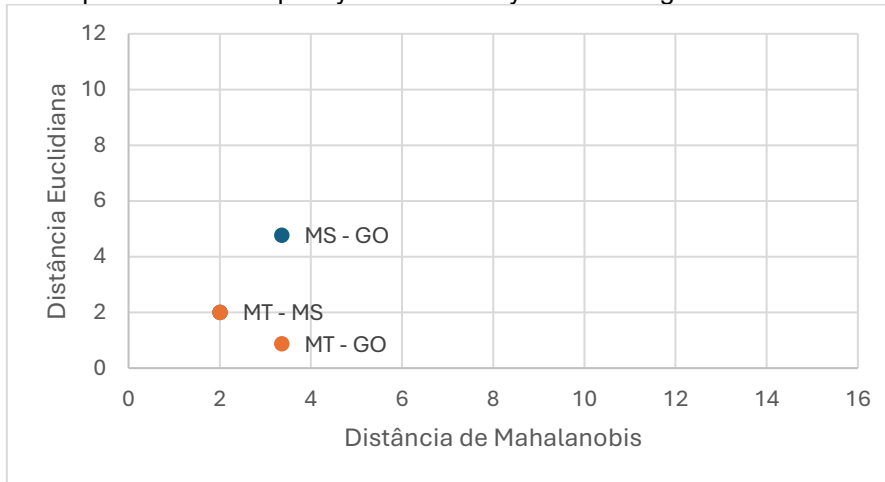


Source: The author.

Figure 4: Scatter plot of the discrepancy in the Literacy Rate among states in the South Region.
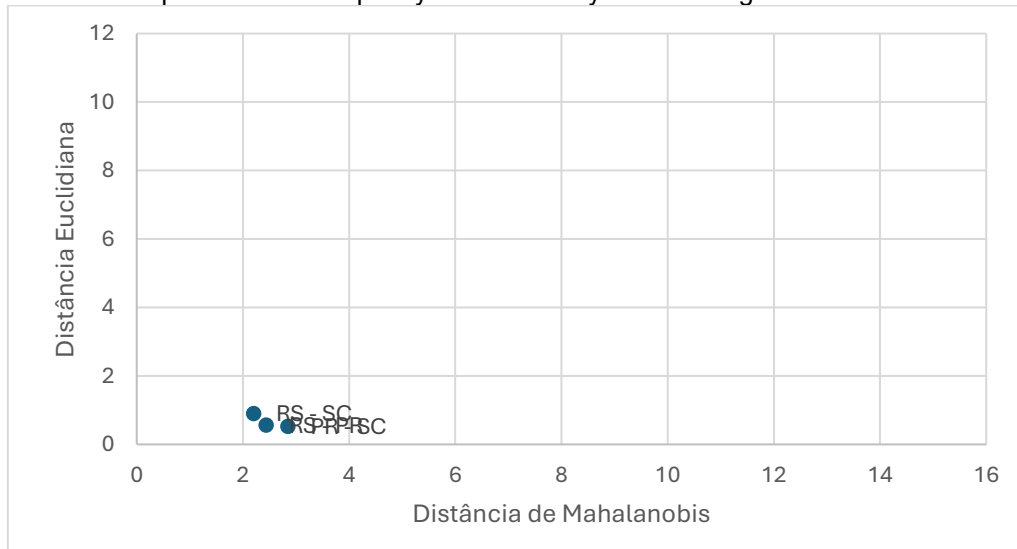


Source: The author.

Figure 5: Scatter plot of the discrepancy in the Literacy Rate among states in the Midwest Region.
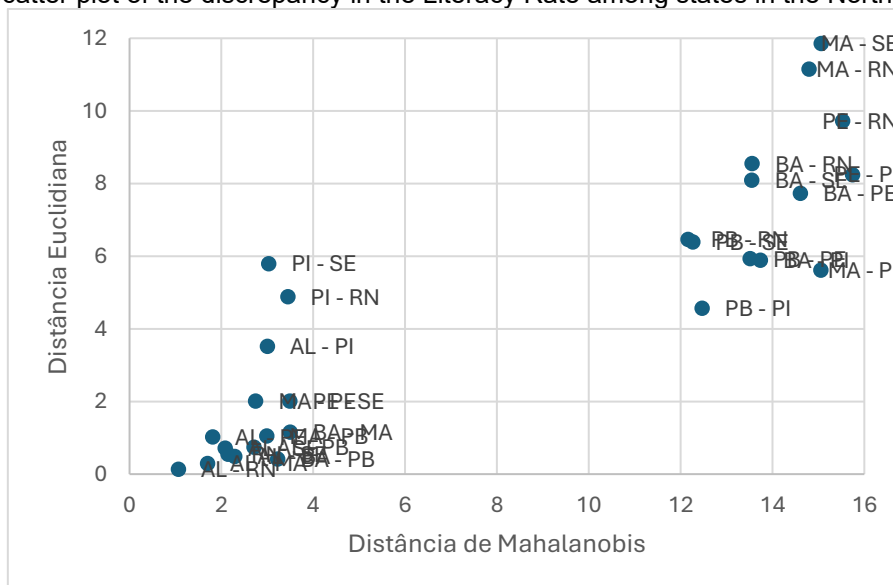


Source: The author.

It can be observed that in the South region, there was little expressiveness in the differences in the limit values obtained for the Euclidean Distance and Mahalanobis Distance, which contributed to the aggregation of Rio Grande do Sul and Santa Catarina.

The distance from Mahalanobis to the Midwest region signaled a prominence of Goiás in relation to Mato Grosso, when analyzed according to the Euclidean distance evidences the discrimination of Mato Grosso. Finally, the graph in figure 6 illustrates the Northeast region.

Figure 6: Scatter plot of the discrepancy in the Literacy Rate among states in the Northeast Region.



Source: The author.

In the Northeast region, two clusters were clearly observed, the one on the lower right under the strong influence of Piauí and Alagoas due to the frequency of the two

distances measured by ordered pairs. The grouping on the left with greater influence from Manaus and Bahia, much more sensitive to the effects of variability around the mean.

## CONCLUSION

Literacy rates in municipal schools in the capital, municipal schools in the interior, and in private schools in the capital and interior, are representative of the level of student proficiency measured by the national exam in 2021. This year was a critical year, schools were still adapting to the changes imposed by hybrid teaching. Even today, these effects are felt by education professionals throughout the country, who contributed to the minimum necessary to leverage the learning of the subjects analyzed. Therefore, it is coherent to adopt the literacy rate as a way to diagnose future scenarios of the school learning deficit, allowing debates among educators on ways to overcome such gaps and enable the student to exercise their citizenship in the future. The statistical treatment carried out allows us to perceive the equality in the variation of the achievements in relation to the average in the subjects of Mathematics, Portuguese Language and Literacy Rate, according to the test of covariance treated both in municipal and private schools, for the capital and the interior. It should be noted that the average values of 28.03 of the municipal schools in the capital and 76.63 of the private schools in the capital, represented the lowest and highest rates of Literacy Rate, respectively. The evaluation of kurtosis and symmetry allowed us to assess that there is a difference between the results of municipal and private schools, because in the latter the results of the students are higher than the average value, while in the municipal schools, there is a tendency for the values to be close to the averages. It was noticed that there is a similarity between the evaluation of Mathematics and Portuguese Language, but when compared with the Literacy rate, most are below 50%, showing that, perhaps, the way of building the evaluation instrument in the classroom is distant from the concepts used in the formation of this rate. In addition, at the time of the pandemic, the literacy process was interrupted. The need for supplementary classes of the Portuguese language is evidenced in order to overcome the deficiency of precarious literacy.

The Covariance test, when applied according to the location and type of study to be offered, proved that there is a difference in variability around the average of the results in performance in the subjects of Portuguese and Mathematics and in Literacy Rates, proving that, in the pandemic scenario, the offer of education was not the same, once it was realized that the use differed in relation to the autarchy and the regionality. This descriptive

statistical analysis indicated that in the pandemic scenario, private education, even in a time of crisis, managed to provide better education, despite the adoption of hybrid teaching.

Since the interest was to evaluate the degree of similarity in teaching, and that the Covariance test indicated the representativeness of the Literacy Rates, the calculation of the distances between the states of the federation was carried out. The analysis of Euclidean distances and Mahalanobis distances between the states in the interior of the South, Southeast, North, Northeast and Midwest regions, showed that in the Southeast region there is a strong influence of variability around the mean in the calculation of distances, which is therefore the one that contains the highest degree of dissimilarity in the level of proficiency measured in the 2021 SAEB. The Midwest region was the only one where the effects of variability around the mean reduced distances between the states in terms of Literacy Rate.

Thus, it is evident the need to use qualitative and quantitative methods to measure the impairment of the teaching and learning process, at the regional, state, and municipal levels, since it subsidizes the investments to be applied by the municipalities in the respective regionalities.

# REFERENCES

1. Aaker, D., Kumar, V., & Day, G. S. (2001). Pesquisa de marketing. São Paulo: Atlas.

2. Alavarce, O. M. (2009). A organização do ensino fundamental em ciclos: algumas questões. Revista Brasileira de Educação, 14(40), 35–50.

3. Barroso, L. P., & Artes, R. (2003). Análise multivariada. Lavras: UFLA.

4. Bonamino, A., & Franco, C. (1999). Avaliação e política educacional: o processo de institucionalização do Saeb. Cadernos de Pesquisa, 108, 101–132.

5. Brasil. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). (2020). Censo da Educação Básica 2019: resumo técnico. Brasília.

6. Brasil. (1996). Lei de Diretrizes e Bases da Educação Nacional, LDB 9394/1996.

7. Bussab, W. O., Miazaki, E. S., & Andrade, D. F. (1990). Introdução à análise de agrupamentos. In 9º Simpósio Brasileiro de Probabilidade e Estatística. São Paulo: IME – USP.

8. Corrar, L. J., Paulo, E., & Dias Filho, J. M. (Eds.). (2007). Análise multivariada: para os cursos de administração, ciências contábeis e economia (pp. 73–130). São Paulo: Atlas.

9. Correa, S. M. B. B. (2003). Probabilidade e estatística (2nd ed.). Belo Horizonte: PUC Minas Virtual.

10. Cruz, C. D., & Carneiro, P. C. S. (2006). Modelos biométricos aplicados ao melhoramento genético. Viçosa: UFV.

11. Franco, C., Fernandes, C., Soares, J. F., Beltrão, K., Barbosa, M. E., & Alves, M. T. G. (2003). O referencial teórico na construção dos questionários contextuais do SAEB 2001. Estudos em Avaliação Educacional, 28, 39–71.

12. Freire, P. (2000). Pedagogia da indignação: cartas pedagógicas e outros escritos. São Paulo: UNESP.

13. Khattree, R., & Naik, D. N. (2000). Multivariate data reduction and discrimination with SAS software. New York: BBU Press and John Wiley & Sons Inc.

14. Luckesi, J. C. (2000). Avaliação da aprendizagem escolar: estudos e proposições (10th ed.). São Paulo: Cortez.

15. Mardia, K. V., Kent, J. T., & Bibby, J. M. (1997). Cluster analysis. In Mathematical Statistics (pp. 360–393). Academic Press.

16. Morrison, D. F. (1971). Multivariate statistical methods. New York: McGraw Hill.

17. Naghettini, M., & Pinto, E. J. A. (2007). Hidrologia estatística. Belo Horizonte: CPRM.

18. Rao, C. R. (1952). An advanced statistical method in biometric research. New York: John Wiley & Sons.

19. Riboldi, J. (1986). Análise de agrupamento: Cluster analysis e suas aplicações. Piracicaba: ESALQ.

20. Sneath, P. H. A., & Sokal, R. R. (1973). Numeric taxonomy: the principles and practice of numerical classification (pp. 15–20). San Francisco: W. H. Freeman.