TIA- A smart psychometric analysis app for Albanian high school exams framework

Eralda Gjika
University of Tirana, Tirana, Albania
E-mail address: eraldagjika@gmail.com

Lule Basha
University of Tirana, Tirana, Albania
E-mail address: lule.hallaci@fshn.edu.al

Afërdita Alizoti
3Centre of Educational Services, Tirana, Albania
E-mail address aferdita.alizoti@gmail.com

Abstract
In support of experts working on high school examinations in Albania, there have been prior studies and research highlighting the importance of psychometric analysis in standard setting for these exams. The State Matura process, initiated in 2006, has undergone several modifications, and since 2016, these results play a crucial role in determining university entry for students in Albania. TIA (Test Item Analysis) Shiny App is an inovative application designed to facilitate psychometric analysis in the development of Albanian high school final exams. The application improves professional decision-making by providing quick insights into item and exam performance empowering psychometricians and subject experts to make data-driven decisions. By integrating item statistics like difficulty and discrimination with flexible cut-score setting, it enhances the efficiency of exam scoring and evaluation processes. In addition, the application uses advanced classification techniques to identify similarities between items. The framework i, s versatile, supporting exams composed of only MCQs, only SAQs, or a combination of bothmaking it a valuable tool for various assessment scenarios. In addition it also offeres detailed item disctractor reports for MCQ items for promoting continuous improvement of exam quality. With comprehensive numeric reports and dynamic visual insights, the TIA-Smart Solution empowers psychometricians and educators to make informed, data-driven decisions effectively.

Keywords: psychometrics, item statistics, exam performance, high school, shiny app.

1. Introduction
Equity and data-driven examinations in today's global evolving educational scenario become imperative to conduct accurate assessment of students' knowledge and capabilities.

Psychometric analysis has emerged as one of the strong pillars of modern assessment as a measure of cognitive and academic skills while ensuring a balanced and fair test development. Albania's State Matura exams, which are crucial for students' academic futures, now can benefit from a cutting-edge tool known as TIA (Test Item Analysis). This article explores into how TIA, a dedicated psychometric analysis tool, may enhance exam standards, helping educators achieve fairer, more accurate assessments, and transforming the educational experience for Albanian high school students. The integration of automated psychometric analytical tools will undoubtedly improve the speed and fairness of evaluations by reducing dependence on human judgement, thus minimizing errors and bias in the process.

In an early study by [1], the author explored student perceptions of e-assessment by conducting a survey with 130 undergraduates who had engaged in online assessments during the 2007–2008 academic year. The survey focused on six key areas: affective factors, validity, practical issues, reliability, security, and learning and teaching. The results of the survey highlighted varied student opinions, with a notable concern about the fairness of item banking. Despite this, students generally viewed e-assessment positively, particularly for its benefits to teaching and learning. The use of digital technologies has clearly improved creation, delivery, evaluation and feedback to assessments in real time.

According to [2] e-assessment has significantly advanced the learning process by offering more efficient ways to evaluate learner outcomes and provide immediate feedback. With the integration of ICT, e-assessment makes the entire process holistic and electronic, right from test design to feedback delivery, hence increasing the measures of student performance. In [3], authors describe a technology-enriched, web-enabled exam preparation and evaluation service aimed at enhancing academic performance. The service allows students to perform preparatory exercises and evaluation tests in a simulated examination environment, providing a realistic experience. This platform offers content personalization to address individual scholastic weaknesses and utilizes data mining techniques to ensure content effectiveness while proactively identifying the academic needs of different student segments. The solution provided was designed with a client-server architecture, utilizing Java technology and XML for information exchange over the internet, which enhances accessibility and user experience.

During and after the COVID-19 pandemic, the necessity of utilizing e-assessment became critical as educational institutions were forced to quickly adapt to online learning and evaluation methods. Institutions had to rely on the technological solutions available in the market to continue operations while ensuring academic integrity. In their study [4], authors address the challenges faced by higher education institutions (HEIs) in assessing students through online examinations. With the shift to online assessments, the risk of student malpractice became more prevalent, especially in technical courses where designing unique and novel questions posed a challenge. In light of this, safeguarding academic integrity has become a priority for HEIs. The authors demonstrate a detailed procedure for creating diverse question formats, such as assertion-reasoning, matching, and multiple-select types, with clear examples.

The challenges posed by the increasing volume of large-scale assessment data, particularly in terms of managing data and conducting psychometric analysis efficiently are also studied in [5] [6]. Traditional psychometric software often lacks the necessary functionality, resulting in high costs and inefficiencies. To solve these issues, the authors designed and implemented a modernized data pipeline that streamlines the management of data, psychometric analysis, report generation, and quality assurance. This pipeline scales well with large databases, reduces human error by minimizing manual processes, ensures repeatability of complex tasks, and improves output quality, ultimately reducing costs. The paper provides insights into the design and functionality of this modernized pipeline and discusses the development of non-technical, user-friendly interfaces for psychometric analysis in large-scale assessments. Challenges are also presented when considering how

such approaches align with learning goals and how various barriers might be reduced in order to establish ease of adoption. Yet well-designed e-assessment systems could also offer significant support for students and teachers concerning skill development and the achievement of long-term learning goals.

2. High School Exams Challenges

The State Matura exams represent an important assessment for Albanian high school students, determining university admissions and influencing academic path. Introduced in 2006 to standardize educational outcomes across Albania, these exams are designed to evaluate core subjects [7] [8].

One of the challenges addressed in this study is related to the lack of historical records on the psychometric performance of multiple-choice questions (MCQs) and written questions (SAQs) over time. This limitation currently prevents item developers and subject experts from effectively rekeying or reconstructing new questions based on previously used ones.

Without this historical context, it is difficult to establish relationships between new and old questions, and to pilot questions or predict outcomes effectively. Item banks, which are essential for developing e-assessments, provide a structured repository of validated questions, facilitating the piloting and refining of items to ensure accurate and reliable assessments [9] [10] [11]. To improve this process, we are developing a solution that offers a user-friendly Shiny app interface. This will enable thorough statistical and psychometric analysis, addressing the current challenges and enhancing the overall efficiency, accuracy, and reliability of the e-assessment development process.

Exam results often reflect students' knowledge inconsistently due to discrepancies in question quality, scoring methods, and the subjective interpretation of results. The TIA app aims to address these issues head-on by analyzing each exam item systematically, ultimately contributing to a more reliable and equitable Matura exam system.

Albania's State Matura exams are plagued by several challenges, which can skew the accuracy and fairness of student assessments [12] [13]. One of the primary issues is limited data availability. In Albania, comprehensive datasets on exam performance are scarce, limiting the insights educators can extract for exam improvement. Additionally, educator resistance to data-driven change is another hurdle; many teachers may be hesitant to adopt new, analytics-driven methods, especially those unfamiliar with psychometrics. Moreover, inadequate psychometric training among educators hampers the effective application of data insights. Together, these factors have contributed to the call for an advanced tool like TIA, designed to overcome these barriers.

3. Introduction to TIA (Test Item Analysis) Shiny app

TIA, short for Test Item Analysis, is an innovative Shiny app specifically crafted to address Albania's unique educational needs. Built on psychometric principles, TIA systematically evaluates the quality of test items, offering detailed feedback on individual questions and highlighting issues like question difficulty and discriminatory power. The app's primary objective is to improve the quality of exams, ensuring that assessments are both challenging

and fair. By applying advanced statistical techniques, TIA provides actionable insights that educators and administrators can use to continuously refine exam quality, making it a valuable tool for educational development in Albania. This app is constructed using foundational and advanced statistical and psychometric libraries in R, enabling robust analysis and precise assessment insights.

3.1. Core Features of TIA

TIA brings several advanced features that empower educators to conduct thorough psychometric analyses of exam items, resulting in fairer and more insightful assessments.

TIA Shiny app Layout

TIA layout provides a comprehensive set of tabs, each dedicated to a specific output focused on either overall exam performance or item-level analysis

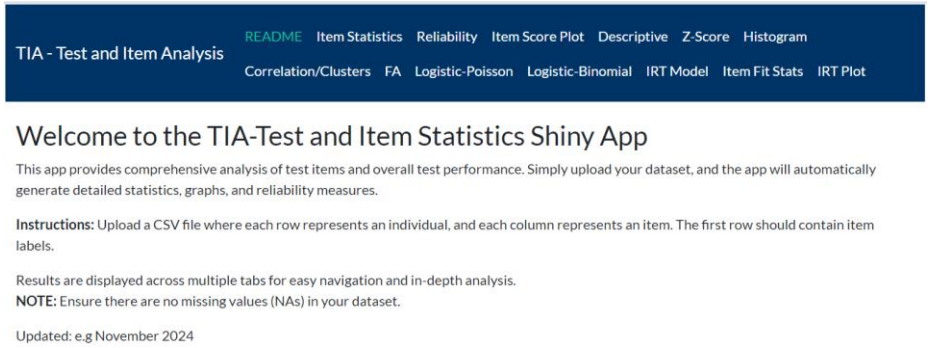


Fig. 1. Shiny app interface layout showing the main features and navigation structure
Source: Authors

The sidebar Menu has an upload tab which will allow the user to upload the response item matrix of the exam (e.g., CSV, Excel). The item matrix is a dataframe where columns represent individual items, rows represent students, and the cells contain the scores for each item, providing a detailed view of student performance across all items.

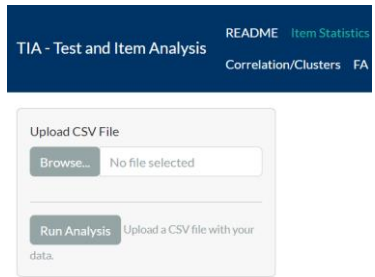


Fig. 2. Sidebar menu in the Shiny app showing the file upload options and action buttons.
Source: Authors

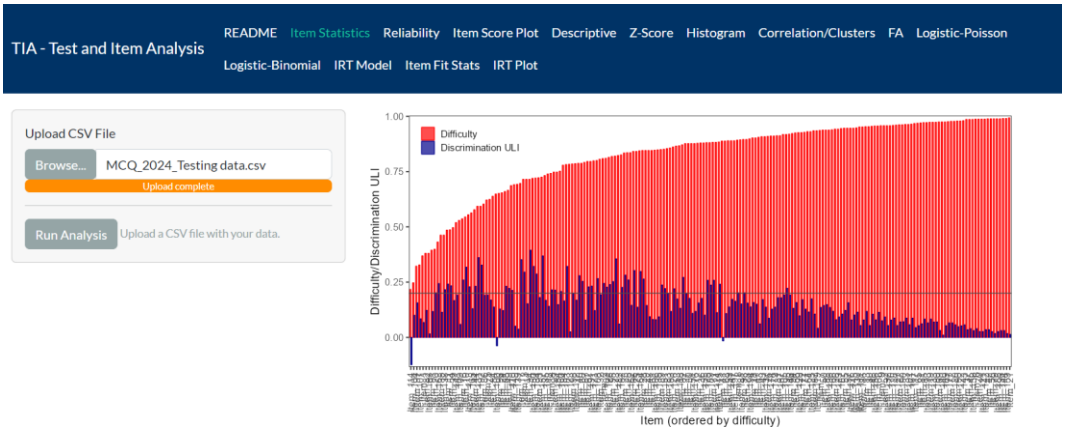
Main Menu Tabs: The main menu consist of multiple tabs, classified into 4 main categories:

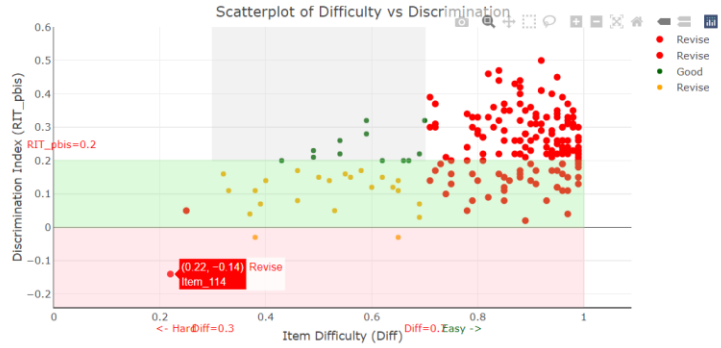
- Item Statistics
- Reliability
- Item performance
- Advanced Statistical and psychometric Models

3.2. Item Statistics

The Item Statistics category has four main tabs. Each of them offering informative statistics about items and overall test performance.

- Item Statistics: This section provides two key visualizations for assessing item characteristics: item difficulty and discrimination index [14] [15] [16] [17] [18]. The correlation scatterplot illustrates the relationship between item difficulty and discrimination index. This dynamic plot allows users to interact by selecting and deselecting specific items, helping to explore the behavior and characteristics of individual items. Users can further analyze the items by focusing on critical or safe regions, which highlight whether items fall within expected performance ranges for difficulty and discrimination. This feature provides an interactive way to assess item quality and take actions about potential improvements.





Loaded dataset with 705 rows and 197 columns.

Show 10 entries Search:

	Item_ID	Diff.	Avg.score	SD	Min.	Max.	Obs.min.	Obs.max.	Prop.max.	RIT_pbis	RIR_DI	Alpha_Drop	PBIS_Flag
1	Item_1	0.37	0.37	0.48	0	1	445	260	0.37	0.04	0.04	0.89	LOW <0.2
2	Item_2	0.81	0.81	0.39	0	1	135	570	0.81	0.22	0.19	0.89	
3	Item_3	0.71	0.71	0.45	0	1	201	504	0.71	0.3	0.27	0.89	

Fig. 3. Item statistics chart and scatterplot visualizing item performance.
Source: Authors

- Downloadable Summary Statistics Table: In addition to the visualizations, the system generates a downloadable table containing essential summary statistics for each assessment item [19] [20] [21] [22] [23] [24] [25]. This table includes the following key metrics: mean, standard deviation, difficulty index, discrimination index, flag column etc. This table serves as a comprehensive summary of the statistical properties of each item, providing a clear, at-a-glance overview of the assessment's performance. Users can download and examine this data for further analysis or reporting purposes.

Loaded dataset with 705 rows and 197 columns.

Show 10 entries Search:

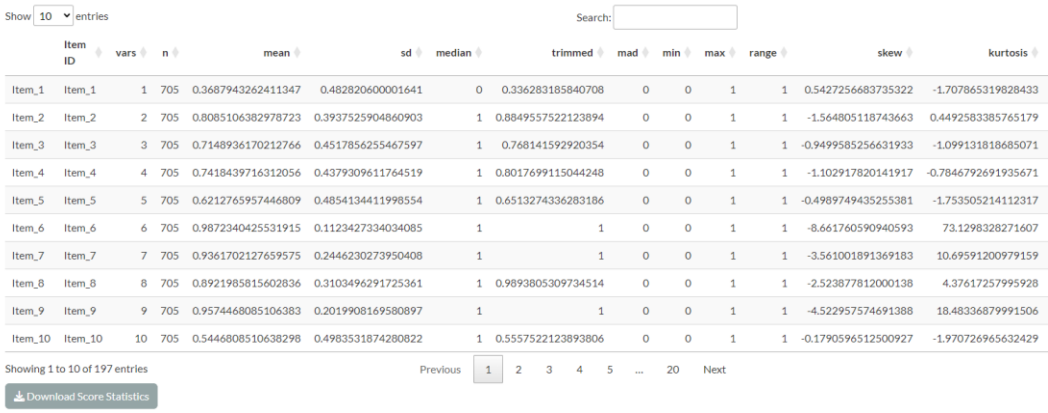
	Item_ID	Diff.	Avg.score	SD	Min.	Max.	Obs.min.	Obs.max.	Prop.max.	RIT_pbis	RIR_DI	Alpha_Drop	PBIS_Flag
1	Item_1	0.37	0.37	0.48	0	1	445	260	0.37	0.04	0.04	0.89	LOW <0.2
2	Item_2	0.81	0.81	0.39	0	1	135	570	0.81	0.22	0.19	0.89	
3	Item_3	0.71	0.71	0.45	0	1	201	504	0.71	0.3	0.27	0.89	
4	Item_4	0.74	0.74	0.44	0	1	182	523	0.74	0.21	0.18	0.89	
5	Item_5	0.62	0.62	0.49	0	1	267	438	0.62	0.2	0.16	0.89	LOW <0.2
6	Item_6	0.99	0.99	0.11	0	1	9	696	0.99	0.19	0.18	0.89	LOW <0.2
7	Item_7	0.94	0.94	0.24	0	1	45	660	0.94	0.28	0.26	0.89	
8	Item_8	0.89	0.89	0.31	0	1	76	629	0.89	0.35	0.33	0.89	
9	Item_9	0.96	0.96	0.2	0	1	30	675	0.96	0.26	0.25	0.89	
10	Item_10	0.54	0.54	0.5	0	1	321	384	0.54	0.26	0.23	0.89	

Showing 1 to 10 of 197 entries Previous 1 2 3 4 5 ... 20 Next

Download Statistics

Fig. 4. Summary item statistics table with downloadable option.
Source: Authors

- The Descriptive Tab provides detailed statistical summaries for each item in the assessment, offering a comprehensive overview of item-level performance. The statistics presented include: Minimum, Maximum, Mean, Median, Standard Error (S.E.), Skewness, Range, Kurtosis. This tab is designed to allow users to filter, search, and sort the statistical data for easier exploration and identification of trends. The data tables are fully interactive, enabling users to quickly isolate specific items based on their statistical characteristics, facilitating efficient analysis and decision-making.



Show 10 entries Search:

	Item ID	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis
Item_1	Item_1	1	705	0.3687943262411347	0.482820600001641	0	0.336283185840708	0	0	1	1	0.5427256683735322	-1.707865319828433
Item_2	Item_2	2	705	0.8085106382978723	0.3937525904860903	1	0.8849557522123894	0	0	1	1	-1.564805118743663	0.4492583385765179
Item_3	Item_3	3	705	0.7148936170212766	0.4517856255467597	1	0.768141592920354	0	0	1	1	-0.9499585256631933	-1.099131818685071
Item_4	Item_4	4	705	0.7418439716312056	0.4379309611764519	1	0.8017699115044248	0	0	1	1	-1.102917820141917	-0.7846792691935671
Item_5	Item_5	5	705	0.6212765957446809	0.4854134411998554	1	0.6513274336283186	0	0	1	1	-0.4989749435255381	-1.753505214112317
Item_6	Item_6	6	705	0.9872340425531915	0.1123427334034085	1	1	0	0	1	1	-8.661760590940593	73.1298328271607
Item_7	Item_7	7	705	0.9361702127659575	0.2446230273950408	1	1	0	0	1	1	-3.561001891369183	10.69591200979159
Item_8	Item_8	8	705	0.8921985815602836	0.3103496291725361	1	0.9893805309734514	0	0	1	1	-2.523877812000138	4.37617257995928
Item_9	Item_9	9	705	0.9574468085106383	0.2019908169580897	1	1	0	0	1	1	-4.522957574691388	18.48336879991506
Item_10	Item_10	10	705	0.5446808510638298	0.4983531874280822	1	0.5557522123893806	0	0	1	1	-0.1790596512500927	-1.970726965632429

Showing 1 to 10 of 197 entries Previous 1 2 3 4 5 ... 20 Next

Download Score Statistics

Fig. 5. Summary descriptive statistics table with downloadable option.
Source: Authors

- Item Score Plot: This feature provides a detailed visualization of the distribution of scores for each individual item on the assessment [26] [27] [28].Users can interact with the plot by scrolling through a slider to examine the performance on each specific item. The plot offers insights into how students are responding to each question, helping to identify potential issues such as questions that are too easy or too difficult based on the distribution of scores.

Fig. 6. Item score distribution bar chart.
Source: Authors

3.3. Descriptive

- Histogram: The histogram tab visualizes the distribution of total scores across all students for the entire examination. This tool allows users to explore the overall performance trends of the cohort. Additionally, the histogram provides flexibility by allowing users to adjust the cut score (passing percentage) and customize the number of bins in the visualization [29] [30]. Adjusting the number of bins helps to refine the granularity of the distribution, providing a clearer view of the performance distribution and offering a more informative and accurate representation of the score range.

Fig. 7. Histogram of total scores with adjustable cut score feature in the Shiny app.
Source: Authors

- Z-Score in Descriptive Section: The Z-score analysis in the descriptive section provides standardized scores for the entire exam, offering a way to assess how far each student's score deviates from the mean in terms of standard deviations [31] [32] [33] [34]. This feature allows for a better understanding of the overall exam performance, providing context for how individual scores compare to the collective results. It is especially useful for identifying outliers, understanding the performance distribution, and gaining insights into the overall difficulty of the exam.

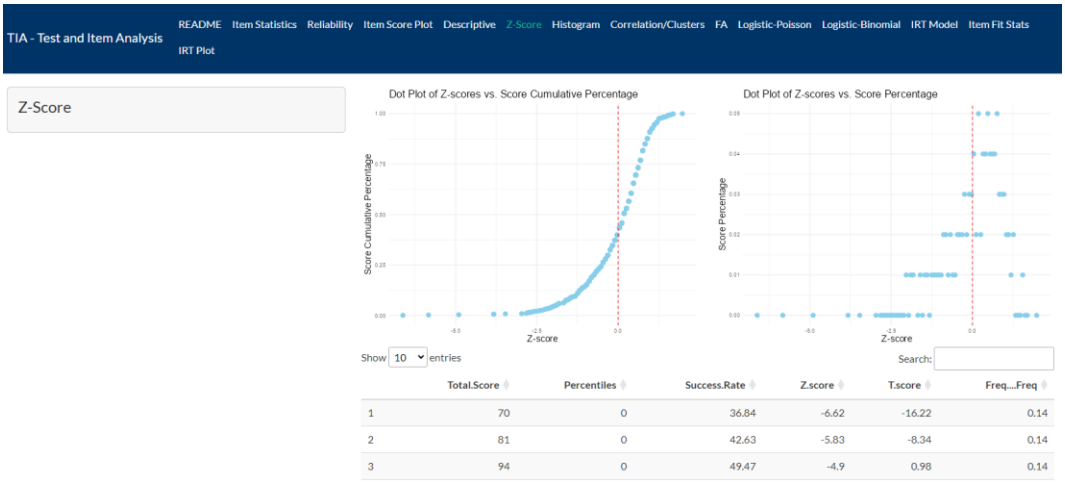


Fig. 8. Z-score table alongside cumulative score percentage plots.
Source: Authors

3.4. Reliability

Reliability: The reliability tab provides calculations of various reliability coefficients to assess the consistency and stability of the test results [35] [36] [37] [38]
These coefficients include:

- *Greatest Lower Bound (GLB)*: GLB as a conservative estimate of reliability, often considered a lower bound. It is especially useful in situations where item difficulty levels vary across the test. By providing the GLB along with other reliability coefficients, the app offers a more comprehensive and reliable evaluation of the test's internal consistency.
- *Cronbach's Alpha*: Measures the internal consistency of the test, indicating how well the items on the exam are correlated with each other [39]. A higher value suggests that the items are consistently measuring the same underlying construct.
- *Kuder-Richardson Formula 20 (KR-20):* Used for dichotomous items (e.g., true/false or multiple-choice questions), it assesses the internal consistency of the test, similar to Cronbach's Alpha but specifically for binary responses.

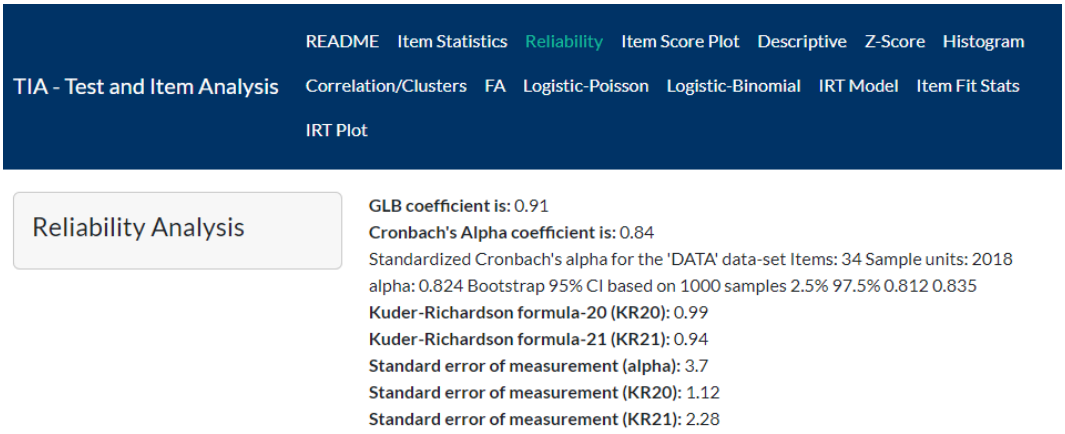


Fig. 9. Reliability coefficients and standard errors.
Source: Authors

3.5. Item Performance

The app generates a distance matrix and a correlation heatmap to visually represent the relationships between items based on their psychometric properties [40] [41]. The user can interactively select the number of clusters for the analysis, allowing for flexible segmentation of items based on their similarity or dissimilarity. In addition to the heatmap, a dendrogram is displayed, showing a hierarchical clustering analysis of the items. This dendrogram is directly associated with the correlation heatmap, providing deeper insights into how items group together based on their correlations. For users who want to conduct further analysis outside the app, the correlation matrix is available for download, enabling them to export the data for use in other statistical software or tools. This feature allows for more advanced analysis and further manipulation of the data, supporting a more comprehensive approach to item analysis and assessment design.

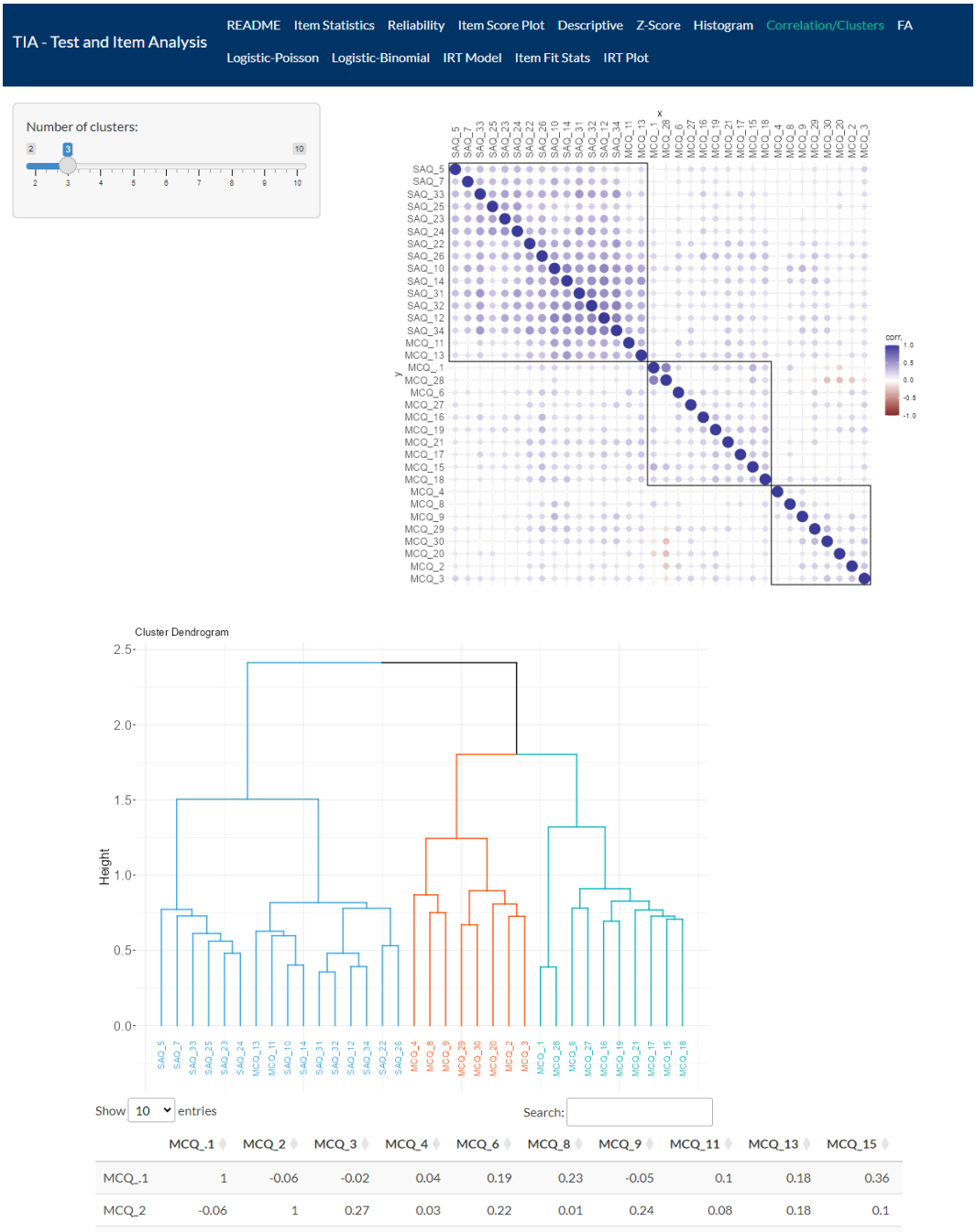


Fig. 10. Correlation heatmap and dendrogram visualizations showing clustering patterns based on the manually selected number of clusters.
Source: Authors

TIA leverages advanced classification methods to analyze and group exam questions with precision. This ensures that every question and response is evaluated within a framework that considers various scoring nuances and question types. Classification techniques are particularly valuable in exams with multiple-choice questions (MCQs), as they reveal patterns in students' answer choices and assess consistency across exam sections. TIA's algorithms support educators in refining test quality through systematic classification, supporting exams to become more insightful and representative of student competencies.

4. Conclusions

TIA was developed in response to the need for statistical and psychometric reports in national examinations in Albania. The lack of such reports and the absence of historical data on item and test performance over the years highlighted the necessity for a framework that could support test developers and subject specialists in Albania. This interface can also be beneficial for university-level examinations. TIA's user-friendly design makes it accessible and easy to use for educators at various levels. It can also be applied to exam analyses for professional certification tests in Albania such as: physician profession, engineer profession etc. Some psychometric reports, however, may require training to ensure accurate interpretation of the results. Overall, TIA offers visual and numeric reports that simplify decision-making for educators and test administrators. These reports display item performance through charts, graphs, and other visual aids, providing a clear, at-a-glance view of exam quality. Numeric data complements these visuals, enabling educators to track metrics like item difficulty, discrimination levels, and response frequencies. This interface can also contribute to the creation of an item bank, enhancing exam development and scoring practices over time by enhancing educational experience.

References

[1] J. Dermo, "e-Assessment and the student learning experience: A survey of student perceptions of e-assessment," *British Journal of Educational Technology,* vol. 40, no. 2, pp. 203-214, 2009.

[2] N. Alruwais, G. Wills and M. Wald, "Advantages and challenges of using e-assessment," *International Journal of Information and Education Technology,* vol. 8, no. 1, pp. 34-37, 2018.

[3] S. R. Abidi and A. Goh, "A Web-enabled exam preparation and evaluation service: providing real-time personalized tests for academic enhancement," in *Proceedings*

IEEE International Conference on Advanced Learning Technologies, Madison, WI, USA, 2001.

[4] S. Allamsetty, M. V. S. S. Chandra, N. Madugula and B. Nayak, "Improvement of the quality of question papers for online examinations toward simultaneous enhancement of students' learning," *IEEE Transactions on Learning Technologies,* vol. 17, pp. 135-142, 2024.

[5] R. Schwarz, H. C. Bulut and C. Anifowose, "A data pipeline for e-large-scale assessments: Better automation, quality assurance, and efficiency," *International Journal of Assessment Tools in Education,* vol. 10, no. Special Issue, pp. 116-131, 2023.

[6] R. Schwarz and E. Gjika, "Modernized psychometric analysis for digital assessments: An open-source approach for automation, quality assurance, and efficiency," in *Proc. 48th Int. Assoc. Educ. Assessment Annu. Conf.*, Kingston, Jamaica, 2023.

[7] E. Gjika, J. P. A. Lessa, A. Alizoti and L. Basha, "Modernizing high school assessments in Albania: Exploring item response theory in state matura exams," in *Proc. Int. Meeting of the Psychometric Society*, Prague, Czech Republic, 2024.

[8] E. Gjika, L. Basha and A. Alizoti, "Exploring the dynamics of assessments: A comprehensive review of MCQs, SAQs, and e-assessment in Albania," *Journal of Natural Sciences,* vol. 35, 2024.

[9] B. D. Wright and S. R. Bell, "Item banks: What, why, how," *Journal of Educational Measurement,* vol. 21, no. 4, p. 331–345, 1984.

[10] J. B. Bjorner, C.-H. Chang, D. Thissen and B. B. Reeve, "Developing tailored instruments: Item banking and computerized adaptive assessment," *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation,* vol. 16, no. 1, pp. 95-108, 2007.

[11] C. Ke, Y. W. Yingwei, Y. J. Yajun and L. R. Runhua, "Item-bank in language e-assessment: Issues and perspectives," in *5th Int. Conf. Computer Science & Education*, Hefei. China, 2010.

[12] E. Gjika, J. Lessa, A. Alizoti and L. Basha, "Modernizing high school assessments in Albania: Exploring item response theory in state matura exams," in *Proc. Int. Meeting of the Psychometric Society*, Prague, 2024.

[13] E. Gjika, L. Basha and A. Alizoti, "Exploring the dynamics of assessments: A comprehensive review of MCQs, SAQs, and e-assessment in Albania," *Journal of Natural Sciences,* vol. 35, 2024.

[14] W.Revelle, "Package 'psych'," 26 June 2024. [Online]. Available: https://cran.r-project.org/web/packages/psych/psych.pdf. [Accessed 9 November 2024].

[15] H. Wickham, "Elegant Graphics for Data Analysis," *Springer-Verlag,* 2016.

[16] C. Sievert, "Interactive web-based data visualization with R, plotly, and shiny," *Chapman & Hall,* 2019.

[17] P. Martinková and A. Hladká, "Computational Aspects of Psychometric Methods: With R," *Chapman and Hall/CRC,* 2023.

[18] P. Martinková and A. Drabinová, "ShinyItemAnalysis for teaching psychometrics and to enforce routine analysis of educational tests," *The R Journal,* vol. 10, no. 2, pp. 503-515, 2018.

[19] W.Revelle, "Package 'psych'," 26 June 2024. [Online]. Available: https://cran.r-project.org/web/packages/psych/psych.pdf. [Accessed 9 November 2024].

[20] P. Martinková and A. Hladká, "Computational Aspects of Psychometric Methods: With R," *Chapman and Hall/CRC,* 2023.

[21] P. Martinková and A. Drabinová, "ShinyItemAnalysis for teaching psychometrics and to enforce routine analysis of educational tests," *The R Journal,* vol. 10, no. 2, pp. 503-515, 2018.

[22] H. Wickham, J. Hester and R. Francois, "Package 'readr'," 1 November 2024. [Online]. Available: https://cran.r-project.org/web/packages/readr/readr.pdf. [Accessed 9 November 2024].

[23] H. Wickham and J. Bryan, "Package ‘readxl’," 7 July 2023. [Online]. Available: https://cran.r-project.org/web/packages/readxl/readxl.pdf. [Accessed 9 November 2024].

[24] D. Magis and J. T. Beland, "Package ‘difNLR’," 3 May 2023. [Online]. Available: https://cran.r-project.org/web/packages/difNLR/difNLR.pdf. [Accessed 9 November 2024].

[25] H. Wickham, "Package ‘tidyr’," 24 January 2024. [Online]. Available: https://cran.r-project.org/web/packages/tidyr/tidyr.pdf. [Accessed 9 November 2024].

[26] H. Wickham, "Elegant Graphics for Data Analysis," *Springer-Verlag,* 2016.

[27] H. Wickham, "Package ‘tidyr’," 1 January 2024. [Online]. Available: https://cran.r-project.org/web/packages/tidyr/tidyr.pdf. [Accessed 9 November 2024].

[28] H. Wickham, R. François, L. Henry and K. Müller, "Package ‘dplyr’," 16 November 2023. [Online]. Available: https://cran.r-project.org/web/packages/dplyr/dplyr.pdf. [Accessed 9 November 2024].

[29] H. Wickham, "Elegant Graphics for Data Analysis," *Springer-Verlag,* 2016.

[30] H. Wickham, R. François, L. Henry and K. Müller, "Package ‘dplyr’," 16 November 2023. [Online]. Available: https://cran.r-project.org/web/packages/dplyr/dplyr.pdf. [Accessed 9 November 2024].

[31] W. Revelle, "Package ‘psych’," 26 June 2024. [Online]. Available: https://cran.r-project.org/web/packages/psych/psych.pdf. [Accessed 9 November 2024].

[32] H. Wickham, "Elegant Graphics for Data Analysis," *Springer-Verlag,* 2026.

[33] P. Martinková and A. Hladká, "Computational Aspects of Psychometric Methods: With R," *Chapman and Hall/CRC,* 2023.

[34] P. Martinková and A. Drabinová, "ShinyItemAnalysis for teaching psychometrics and to enforce routine analysis of educational tests," *The R Journal,* vol. 10, no. 2, pp. 503-515, 2018.

[35] S. Zhang, "R Programming for Psychometrics," 13 November 2023. [Online]. Available: https://bookdown.org/sz_psyc490/r4psychometics. [Accessed 9 November 2024].

[36] W. Revelle and R. E. Zinbarg, "Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma," *Psychometrika,* vol. 74, no. 1, pp. 145-154, 2008.

[37] C. P. Malkewitz, P. Schwall, C. Meesters and J. Hardt, "Estimating reliability: A comparison of Cronbach's α, McDonald's ωt and the greatest lower bound," *Social Sciences & Humanities Open,* vol. 7, no. 1, 2023.

[38] D. Rizopoulos, "Package 'ltm'," 13 October 2022. [Online]. Available: https://cran.r-project.org/web/packages/ltm/ltm.pdf. [Accessed 9 November 2024].

[39] L. J. Cronbach, "Coefficient alpha and the internal structure of tests," *Biometrika,* vol. 16, pp. 297-335, 1951.

[40] D.Kassambara, "Package 'factoextra'," 13 October 2022. [Online]. Available: https://cran.r-project.org/web/packages/factoextra/factoextra.pdf . [Accessed 9 November 2024].

[41] T. Galili, "dendextend: an R package for visualizing, adjusting, and comparing trees of hierarchical clustering," *Bioinformatics,* vol. 31, no. 22, pp. 3718-3720, 2015.