Ahlussunnah: Journal of Islamic Education
https://ejournal.stitahlussunnah.ac.id/index.php/JIE

Integrating Advanced Data Analysis in Islamic Education Research

Soibatul Aslamiah Nasution[1], Zulfani Sesmiarni[2], Iswandi[1], Ramadhoni Aulia Gusli[2]
[1]IAI YAPTIP Chadijah Ismail Pasaman Barat, Indonesia
[2]Universitas Islam Negeri Sjech M. Djamil Djambek Bukittinggi, Indonesia
aslamiahs114@gmail.com*

Article Information:
Received October 21, 2025
Revised November 28, 2025
Accepted December 31, 2025

Keywords: *Data analysis, mixed methods, quantitative analysis, qualitative analysis, NVivo*

Abstract

This study aims to integrate various high-level data analysis techniques in quantitative and qualitative approaches to produce a comprehensive understanding of educational phenomena. A mixed methods approach is used to systematically combine the strengths of numerical and interpretive analysis. In the quantitative stage, data were analyzed using advanced statistical techniques such as regression analysis, multiple regression, principal component analysis (PCA), discriminant analysis, canonical analysis, path analysis, factor analysis, ANAVA, ANCOVA, MANOVA, and MANCOVA to test the significant relationships, differences, and effects between variables. Meanwhile, the qualitative stage uses taxonomy/domain analysis, constant comparative analysis, and symbolic analysis to explore the deeper meaning of field data inductively and contextually. The use of the NVivo application supports the qualitative analysis process through thematic coding, visualization of conceptual relationships, and evidence-based data validation. The integration of these two approaches strengthens the internal and external validity of the research, increases the reliability of the results, and broadens theoretical and practical understanding in the context of modern education. The results of this study indicate that the application of a combination of high-level data analysis methods and qualitative analysis technology support is capable of producing richer, more accurate, and more relevant interpretations for the development of data- and meaning-based educational science.

INTRODUCTION

The development of science and technology in the era of the 4.0 industrial revolution has brought fundamental changes to the scientific research paradigm. Researchers are no longer only required to describe phenomena, but also to analyze complex relationships between variables and interpret the meanings contained therein. In this context, a single methodological approach is often insufficient to capture the complexity of dynamic social and educational realities. Therefore, the integration of high-level data analysis in quantitative and qualitative research has become a strategic necessity in producing more holistic and credible findings

How to cite:	Nasution, S. A., Sesmiarni, Z., Iswandi, I., Gusli, R. A. (2025). Integrating Advanced Data Analysis in Islamic Education Research. *Ahlussunnah: Journal of Islamic Education*, *4*(3), 763-778.
E-ISSN:	2827-9573
Published by:	The Institute for Research and Community Service

(Tashakkori & Teddlie, 2021). The mixed methods approach has emerged as an alternative that combines the strengths of quantitative objectivity with qualitative interpretive depth. Through this integration, researchers can answer research questions that simultaneously test the relationships between variables and interpret the social and psychological contexts behind them.

In quantitative analysis, various advanced statistical techniques are used to analyze relationships, differences, and influences between variables. Methods such as regression analysis and multiple regression are useful for predicting dependent variables based on a number of predictors (Hair et al., 2019a). Meanwhile, principal component analysis (PCA) is used to reduce data dimensions while retaining the greatest variance, discriminant analysis (LDA) is used to classify objects into specific groups, and canonical analysis (CCA) is used to assess the relationship between two sets of variables.

Furthermore, techniques such as path analysis and factor analysis serve to identify causal relationships between latent variables and the indicators that compose them (Byrne, 2016). Meanwhile, ANAVA, ANCOVA, MANOVA, and MANCOVA are used to test differences in means between groups in the context of one or more dependent variables, taking into account covariates or interactions between variables (Tabachnick & Fidell, 2019). The use of these techniques allows researchers to gain a more accurate and measurable understanding of complex empirical phenomena.

Meanwhile, qualitative analysis plays an important role in exploring the meanings, symbols, and subjective experiences of participants. Approaches such as taxonomic/domain analysis are used to identify the structure of meaning categories based on hierarchical relationships in the data (Spradley, 2016). The Constant Comparative Method introduced by Glaser and Strauss in 1967 allows researchers to continuously compare data until coherent theoretical categories emerge. Meanwhile, symbolic analysis interprets the social meanings attached to actions, language, and symbols in the context of everyday life. Through this approach, researchers not only understand "what" happened, but also 'why' and "how" a phenomenon occurred in a particular socio-cultural context.

In the context of modern methodology, the use of qualitative data analysis technology such as NVivo has become an integral part of the research process. NVivo serves to assist in data coding, theme exploration, and visualization of relationships between concepts efficiently (Bazeley & Jackson, 2019). This software improves the accuracy and transparency of qualitative analysis, while enabling collaboration among research teams in managing large and complex data (Zamawe, 2015). The integration of NVivo with advanced statistical analysis also opens up opportunities to combine quantitative and qualitative results in a single integrated analytical framework, for example by linking qualitative coding results with descriptive or inferential statistical results.

Thus, this study aims to develop an integrative model in research data analysis that combines quantitative and qualitative methods in a complementary manner. Through the application of high-level data analysis techniques and the use of analysis software such as NVivo, this study is expected to contribute methodologically and practically to the development of science, particularly in the fields of education, social sciences, and management. This approach is also expected to strengthen the quality of research results by improving the validity, reliability, and depth of interpretation of the phenomena studied.

METHODS

This study uses a mixed methods approach that integrates advanced quantitative data analysis and interpretive qualitative analysis to gain a comprehensive understanding of the phenomenon under study (Engkizar et al., 2024; 2025;

Kasheem et al., 2025; Kassymova et al., 2025). This approach allows researchers to empirically explain the relationship between variables through quantitative data and deepen its meaning through qualitative exploration. Quantitative data were collected through questionnaires and analyzed using various techniques such as regression analysis, multiple regression, principal component analysis (PCA), discriminant analysis (LDA), canonical analysis (CCA), path analysis, factor analysis (EFA and CFA), ANOVA, ANCOVA, MANOVA, and MANCOVA with the help of statistical software such as SPSS, AMOS, and SmartPLS (Sugiyono, 2020).

Meanwhile, qualitative data was obtained through interviews, observations, and document analysis, then analyzed inductively using a taxonomy/domain, constant comparative, and symbolic analysis approach. In the qualitative analysis stage, this study utilized the NVivo application to assist in the coding process, theme grouping, thematic analysis, and visualization of relationships between categories, thereby increasing the efficiency and transparency of the analysis. Data validity was ensured through triangulation of sources, methods, and theories, as well as examination of the results by experts and member checking. The reliability of quantitative instruments was tested using Cronbach's Alpha, while the credibility of qualitative data was strengthened by peer debriefing and audit trails. The integration of these two approaches resulted in an in-depth, objective, and comprehensive analysis of the research phenomenon (Sugiyono, 2017; Engkizar et al., 2022; 2022; Htay et al., 2025; Fauzi et al., 2025; Ferdinan et al., 2024; Uyubah & Anawati, 2024; Feri, 2025; Asy'ari et al., 2023).

RESULT AND DISCUSSION

Advanced Data Analysis

Advanced data analysis refers to the application of complex statistical and computational techniques to understand multidimensional phenomena. In social and educational research, this approach includes various methods such as factor analysis, multivariate regression analysis, path analysis, and Structural Equation Modeling (SEM) (Sugiyono, 2014). This approach allows researchers to test theoretical models with latent variables, which is highly relevant in Islamic education because many concepts such as faith, morals, and spirituality are abstract and cannot be measured directly (Sugiyono, 2013). In Islamic education, data analysis is not only interpreted as a quantitative process, but also as an epistemological effort to understand the essence of education within the framework of tauhid values. According to Al-Attas (1980), Islamic education must be oriented towards the formation of insan kamil through the integration of knowledge and morals. Therefore, data analysis in this context must be able to reveal the spiritual and moral dimensions of students scientifically, without neglecting the aspect of values. In other words, modern methodologies such as SEM or confirmatory factor analysis (CFA) can be used to measure Islamic value constructs validly and reliably, as long as the instrument and model designs are constructed in accordance with the principles of Islamic education.

Quantitative Data Analysis Techniques (Regression Analysis, Multiple, Principal Component Analysis, Discriminant Analysis, Canonical Analysis, Path Analysis, Factor Analysis, Anava, Anacova, Manova, and Mancova

Regression Analysis

Regression analysis is one of the most fundamental statistical techniques in quantitative research used to model the functional relationship between one dependent variable and one or more independent variables. In simple regression, this relationship involves only one predictor, whereas in multiple regression there are several predictors that contribute simultaneously to the dependent variable. The main purpose of regression is to explain the variation in the dependent variable and predict unknown values based on a combination of predictor variables. Regression coefficients indicate the direction and magnitude of the influence of each predictor after controlling for the influence of other predictors. According to (Miles et al., 2014), regression helps researchers understand the strength of the relationship between variables, both in an explanatory and predictive context, while (Frost 2020) emphasizes that interpreting regression results requires an understanding of the theoretical context and data underlying the relationship.

To ensure the validity of regression results, a number of basic assumptions must be met. First, the relationship between the dependent and independent variables must be linear so that the model can capture the actual relationship pattern. Second, the residuals from the model must be independent, meaning that errors between observations must not be correlated with each other. Third, the residual variance must be constant (homoscedasticity) so that the significance test results remain accurate, and fourth, the residuals must be normally distributed to allow for valid parametric inference. In addition, there should be no high multicollinearity between predictors as this can confound coefficient estimates and reduce the reliability of the model. Diagnostic checks such as residual plots and Variance Inflation Factor (VIF) are needed to test for violations of these assumptions. Good regression reporting also includes R^2 values, standardized and unstandardized coefficients, confidence intervals, and significance test results (Sugiyono, 2015).

In practice, regression is often widely used in various fields such as education, psychology, economics, and health. However, its use often draws methodological criticism, especially in the use of automatic variable selection techniques such as stepwise regression. (Miles & Huberman, 2014) explain that this technique can produce models that are unstable and prone to overfitting, which ultimately reduces the model's ability to generalize to the population. Therefore, theory-driven modeling and cross-validation are more recommended. In addition, in observational studies, it is necessary to consider confounding variables that can affect causal interpretation. To overcome this, researchers can use advanced approaches such as multivariate regression, propensity score matching, or structural equation modeling (SEM) (Greenland, Pearl, & Robins, 1999). Thus, regression remains a powerful analytical tool in quantitative research, but its application must be accompanied by methodological caution to ensure that the results obtained are valid, reliable, and scientifically meaningful.

Multiple

Multiple regression analysis is an advanced statistical technique used to study the relationship between one continuous dependent variable and two or more independent variables or predictors. This approach allows researchers to understand how several factors simultaneously affect a particular outcome, as well as determine how much each predictor contributes to the total variation in the dependent variable. According to Frost (2020), multiple regression serves not only predictive purposes, but also tests theoretical models that explain causal relationships between variables. In the context of education or psychology, for example, multiple regression is often used to assess the influence of factors such as learning motivation, teaching quality, and school environment on student learning outcomes (Black & Wiliam, 2018).

Mathematically, a multiple regression model can be expressed as ($Y = \beta_0 + \beta_1X_1 + \beta_2X_2 + ... + \beta_nX_n + \varepsilon$), where ($Y$) is the dependent variable, ($X_1, X_2, ..., X_n$) are the predictor variables, (β_0) is the constant, and (ε) represents the error or residual. Each coefficient ((β_i)) indicates the average change in the dependent variable when one independent variable increases by one unit, while the other variables are held constant. The main assumptions in multiple regression include linearity of the relationship between variables, independence of residuals, homoscedasticity (constant residual variance), normality of residuals, and no high multicollinearity between predictors (Field, 2018).

Violations of these assumptions can cause bias in estimates and reduce the reliability of the model. Therefore, diagnostic tests such as residual plots, Durbin-Watson tests, and Variance Inflation Factor (VIF) are very important to ensure the validity of the model. In reporting results, researchers usually include the coefficient of determination (R^2) value to show the proportion of variation in the dependent variable explained by all predictors, as well as a significance test (F-test) to assess whether the model as a whole is significant. In addition, standardized coefficients are used to compare the relative strength of predictors in the model. However, common practices such as stepwise regression are often criticized for producing unstable models and potentially overfitting, where the model over-adjusts to the sample data and fails to generalize to a broader population (Frost, 2020).

A more recommended approach is theory-driven selection or the use of cross-validation to ensure that the model remains consistent and can be applied to new data. Furthermore, multiple regression also plays an important role in inferential analysis and causal testing, especially in observational studies. However, causal interpretation can only be done if important assumptions, such as the absence of confounding and omitted variable bias, are met. In educational or social research, multiple regression is often integrated with path analysis or structural equation modeling (SEM) to gain a deeper understanding of the relationships between variables. Thus, multiple regression serves not only as a descriptive statistical tool, but also as an analytical method that allows researchers to construct, test, and validate complex conceptual models.

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a statistical technique used to reduce the dimensions of data by transforming a set of correlated variables into a small number of new components that are orthogonal (uncorrelated with each other). The main purpose of PCA is to retain as much of the total variation in the original data as possible while simplifying the data structure to make it easier to explore and visualize. Thus, PCA helps researchers extract key information from large datasets without losing significant statistical meaning. In social, educational, and data science research, PCA is often used to identify latent patterns or hidden dimensions behind complex observed variables, while reducing the computational load before further analysis such as regression, clustering, or discriminant analysis (Jolliffe & Cadima, 2016).

Mathematically, PCA generates a linear combination of the original variables to form principal components, where the first component explains the largest proportion of the total variance, followed by subsequent components that explain the remaining variance in sequence. This process is based on the decomposition of the covariance or correlation matrix, which then produces eigenvalues and eigenvectors as the basis for component formation. The selection of the number of components usually refers to the Kaiser criterion (eigenvalue > 1), scree plot to see the inflection point (elbow), or cumulative variance explained up to a certain limit (usually 70–90%) (Sugiyono, 2017). To ensure that the PCA results are valid, it is important for researchers to standardize the variables first, generally through z-score transformation, especially if the variables have different measurement units. This is because PCA is sensitive to data scales, and variables with large variances will dominate the results if normalization is not performed.

In interpreting the results, the loading matrix is used to see the contribution of each variable to each component, while the biplot allows visualization of the relationship between variables and observations in the principal component space. However, it is important to emphasize that PCA is not a causal model; the resulting components cannot be interpreted as cause-and-effect relationships between variables, but only as representations of data variance patterns. Therefore, PCA is more appropriately used as an exploratory and pre-processing tool prior to inferential analysis. In the modern context, PCA also forms the basis for various machine learning algorithms such as Principal Component Regression (PCR) and Kernel PCA, which extend its use to non-linear and large-scale data.

Discriminant Analysis (LDA)

Discriminant Analysis is a multivariate statistical technique that aims to identify the most effective linear combination of predictor variables in distinguishing between two or more predefined categorical groups. In its classical form, Linear Discriminant Analysis (LDA) constructs a discriminant function that maximizes the ratio of between-group variance to within-group variance, thereby producing an optimal decision boundary for classification. This technique is not only used to classify new objects into specific groups, but also to understand the extent to which predictor variables contribute to the separation between groups. In social, educational, and psychological research, LDA is often used to analyze differences in characteristics between groups of students, learning styles, or academic performance based on multivariate indicators.

The main assumptions in classical discriminant analysis include the multivariate normality of predictor variables in each group and the homogeneity of the covariance matrix between groups (homoscedasticity). When these assumptions are met, LDA produces statistically optimal linear separation boundaries. However, violations of assumptions such as differences in covariance between groups or non-normal distributions can reduce classification accuracy and model stability. In such conditions, alternatives such as Quadratic Discriminant Analysis (QDA) are used, which allow for different covariances between groups. In addition, modern approaches such as Regularized Discriminant Analysis (RDA) and Penalized LDA improve performance on small or high-dimensional datasets through shrinkage and penalization techniques (Li et al., 2020).

In the context of big data and machine learning, LDA has evolved into part of probabilistic classification algorithms, where the discriminant function is viewed as a generative model that explicitly estimates the distribution of each class. LDA also plays an important role in feature extraction before being applied to advanced classification models such as Support Vector Machines (SVM) or Neural Networks (Li et al., 2020). However, for datasets with a small number of samples per class (small-sample problem), covariance estimation can become unstable or singular, requiring approaches such as shrinkage LDA or Bayesian LDA to avoid overfitting. Thus, although LDA is a classical method, it remains relevant in the modern data era due to its ability to be generalized through regularization approaches and integration with machine learning methods.

Canonical Analysis (Canonical Correlation Analysis – CCA)

Canonical Analysis (Canonical Correlation Analysis, CCA) is a multivariate statistical method used to explore and measure linear relationships between two sets of variables, for example between a set of independent variables (set X) and a set of dependent variables (set Y). This technique was first introduced by Hotelling (1936) with the aim of finding linear combinations of each set of variables called canonical variates that have maximum correlation with each other. CCA is very useful in the context of social, educational, and psychological research when researchers want to understand how two groups of complex constructs are simultaneously related to each other, such as the relationship between learning style indicators and academic outcomes, or between demographic variables and job satisfaction.

In its implementation, CCA assumes a linear relationship, multivariate normality, and no high multicollinearity within each set of variables. High multicollinearity can cause instability in canonical weight estimates and misinterpretation of canonical loadings. Therefore, diagnostic checks such as variance inflation factor (VIF) and inter-variable correlation matrices are essential before running the analysis (Hair et al., 2019). Interpretation of CCA results focuses on canonical correlations, loadings, and cross-loadings that describe how strongly individual variables contribute to the formed canonical functions. To avoid overinterpretation, only canonical pairs with significant correlations and eigenvalues above the threshold (e.g., $\lambda > 0.1$) are analyzed further (Tabachnick & Fidell, 2019).

In modern developments, CCA has been expanded to address the challenges of large data sets and high multicollinearity through the Regularized Canonical Correlation Analysis (rCCA) and Partial Least Squares (PLS) approaches. The rCCA approach adds a regularization penalty to stabilize estimates when the number of variables exceeds the number of samples, while PLS focuses on maximizing covariance rather than correlation, making it more robust to noise and non-normal data. Furthermore, the integration of CCA with machine learning has resulted in variants such as Kernel CCA (KCCA) and Deep CCA, which are capable of capturing non-linear relationships between sets of variables in high-dimensional feature spaces. Overall, Canonical Analysis remains an important technique in exploring complex relationships between multivariate constructs, especially in research involving multidimensional interactions between variables.

Path Analysis

Path analysis is an extension of multiple regression used to test causal relationship models between variables, both direct and indirect through mediators. This technique was first introduced by Sewall Wright (1934) as a method for describing and measuring causal effects in a system of interrelated variables. In practice, path analysis is used to verify whether the structure of relationships between variables in a theoretical model is consistent with empirical data. Path coefficients represent the strength and direction of relationships between variables and are often interpreted in the context of mediation or moderation models (Kline, 2016). Thus, path analysis serves as a bridge between multiple regression and Structural Equation Modeling (SEM), where SEM is an extension of path analysis that includes latent variables and more complex covariance models (Byrne, 2016).

The basic assumptions in path analysis include multivariate normality, model identifiability, and adequate sample size relative to the number of parameters estimated. A large sample size is necessary for parameter estimates to be stable and model results to be reliable. Parameter estimation is usually performed using the Maximum Likelihood Estimation (MLE) method, which is sensitive to violations of multivariate normality. If these assumptions are not met, alternative approaches such as Bootstrapping or Asymptotically Distribution-Free (ADF) estimation can be used to obtain more robust estimates (Bollen & Stine, 1992). In addition, model fit testing is performed using various goodness-of-fit indices such as the Comparative Fit Index (CFI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR), which provide an overview of how well the theoretical model fits the empirical data.

In the context of social and educational research, path analysis is often applied to understand the complex causal relationships between factors such as motivation, learning quality, and student learning outcomes. For example, the effect of motivation on academic achievement can occur directly, or indirectly through mediating variables such as learning engagement (Meyers et al., 2013). Furthermore, with the help of software such as AMOS, LISREL, or SmartPLS, researchers can construct path models that integrate latent variables, allowing for a richer interpretation of multidimensional educational phenomena. Thus, path analysis serves not only as a statistical tool, but also as a methodological approach to testing the validity of theoretical models and understanding the complex dynamics between constructs in scientific research.

Factor Analysis

Factor analysis is one of the most important multivariate statistical techniques in social, psychological, and educational research because it allows researchers to identify the underlying latent structure of a set of measured variables. There are two main approaches to factor analysis, namely Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA). EFA aims to explore patterns of relationships between variables and discover previously unknown dimensions or latent factors, while CFA is used to test whether a theoretically formulated factor structure is consistent with empirical data (Brown, 2015). In the context of educational and psychometric research, factor analysis is often used to validate instrument constructs such as attitude scales, learning motivation, or teacher competence, thereby ensuring that the indicators used actually measure the intended construct.

In conducting EFA, researchers must determine the appropriate factor extraction method, such as Principal Axis Factoring (PAF) or Maximum Likelihood (ML), as well as the rotation method to obtain a factor structure that is easier to interpret. Rotation can be orthogonal (e.g., Varimax), which assumes that factors are uncorrelated, or oblique (e.g., Promax), which allows for correlations between factors. Criteria for determining the number of factors can use eigenvalue ≥ 1 (Kaiser's rule), scree plot, or parallel analysis, which is more accurate in avoiding over factoring. In addition, sample size also plays an important role: various studies suggest a minimum of 5–10 respondents per item or a total sample of at least 200 for stable and reliable results.

Meanwhile, Confirmatory Factor Analysis (CFA), which is generally conducted within the framework of Structural Equation Modeling (SEM), aims to test the extent to which theoretical factor models fit the data obtained. In CFA, the model is evaluated using model fit indices such as Chi-Square (χ^2), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR). In addition, construct validity is examined through convergent validity (via Average Variance Extracted ≥ 0.5) and discriminant validity (For discriminant validity, correlations between factors < 0.85 or AVE > MSV). In modern research, CFA and EFA are often used sequentially: EFA for initial exploration and CFA for theoretical confirmation, ensuring that the instrument has a robust, reliable, and valid factor structure (Hair et al., 2019b).

ANOVA (Analysis of Variance)

Analysis of variance (ANOVA) is an inferential statistical technique used to test whether there are significant differences between the means of two or more groups. Unlike the t-test, which only compares two groups, ANOVA can analyze more than two groups at once while maintaining the error rate (Type I error) within reasonable limits (Field, 2018). ANOVA was developed by Ronald A. Fisher in the 1920s as part of experimental methodology, and to this day remains one of the most important methods in educational, social, and psychological research.

In principle, ANOVA compares the variability between groups (between group means) with the variability within groups (random error). If the variation between groups is much greater than the variation within groups, it can be concluded that there is a significant effect of the independent variable on the dependent variable (Howell, 2013). The simplest form is One-Way ANOVA, which involves one independent variable with two or more levels (groups). Furthermore, there is Two-Way ANOVA, which allows for the analysis of two factors simultaneously as well as the interaction between them. In educational research, ANOVA is often used to test the effectiveness of learning methods, differences in learning outcomes between classes, or the effect of training on teacher performance (Gravetter & Wallnau, 2017).

In its application, ANOVA has several key assumptions: (1) normally distributed data, (2) homogeneity of variances between groups, and (3) independence between observations. When the assumption of homogeneity of variances is violated, researchers can use Welch's ANOVA or the more robust Brown–Forsythe test. After the F-test shows a significant difference, further steps are taken through post-hoc analysis (e.g., Tukey HSD, Bonferroni, or Scheffé test) to determine which groups differ significantly from one another (Field, 2018).

In the context of modern social and educational research, ANOVA is often integrated into more complex models such as Analysis of Covariance (ANCOVA) to control for covariate variables, or Multivariate ANOVA (MANOVA) to handle more than one dependent variable simultaneously (Tabachnick & Fidell, 2019). Thus, ANOVA not only functions as a tool for testing mean differences, but also serves as the basis for the development of more advanced multivariate analysis techniques.

ANCOVA (Analysis of Covariance)

Analysis of Covariance (ANCOVA) is an extension of Analysis of Variance (ANOVA) that combines elements of linear regression and ANOVA to control for the influence of covariates, which are continuous variables that may affect the dependent variable but are not the main focus of the study (Field, 2018). The main purpose of ANCOVA is to reduce error variance and increase statistical power by adjusting group differences for covariates before testing the main treatment effect. Thus, ANCOVA allows researchers to obtain more accurate estimates of the effect of independent variables on dependent variables after the covariate effect has been statistically controlled.

Conceptually, ANCOVA calculates the difference in means between groups after adjusting for covariates, so that comparisons are made as if all groups had the same covariate values (Tabachnick & Fidell, 2019). For example, in educational research, ANCOVA is often used to compare the effectiveness of new learning methods on student learning outcomes while controlling for initial scores (pretests) as covariates. With this approach, the measured treatment effect truly reflects the influence of the learning method, rather than the result of variations in students' initial abilities.

In applying ANCOVA, there are several important assumptions that must be met: i) a linear relationship between the covariates and the dependent variable; ii) homogeneity of regression slopes between groups, meaning that the relationship between the covariates and the dependent variable must be the same for all groups; iii) normality of residuals; iv) homogeneity of variance between groups; and v) independence between observations. If the assumption of homogeneity of regression slopes is violated, then the ANCOVA model is no longer valid, and researchers are advised to use a covariate-factor interaction model or a nonparametric approach.

In modern practice, ANCOVA is also used in the framework of mixed models and structural equation modeling (SEM) to expand the analytical capabilities of hierarchical or longitudinal data. In addition, statistical software such as IBM SPSS, R, and Jamovi provide comprehensive facilities for calculating ANCOVA, including main effect tests, interactions, and assumption checks using Levene's Test and residual plots. In the context of education and social sciences, ANCOVA is an important tool for improving the internal validity of experimental and quasi-experimental research.

MANOVA (Multivariate Analysis of Variance)

Multivariate Analysis of Variance (MANOVA) is an extension of Analysis of Variance (ANOVA) that allows researchers to test the effect of one or more independent variables on more than one continuous dependent variable simultaneously (Tabachnick & Fidell, 2019). In other words, MANOVA considers the correlation between dependent variables, making it more efficient and informative than conducting a series of separate ANOVAs, which risk increasing the Type I error rate (Hair et al., 2021). This approach is particularly useful in social, psychological, and educational research when the measurement results consist of several interrelated aspects, such as academic achievement, motivation, and learning satisfaction.

Conceptually, MANOVA analyzes differences in multivariate mean vectors between groups. This model tests whether certain linear combinations of dependent variables differ significantly between treatment groups. The main test statistics in MANOVA include Wilks' Lambda, Pillai's Trace, Hotelling's Trace, and Roy's Largest Root, each of which represents a different approach to evaluating the similarity of covariance between groups. These statistical values are then tested using the F distribution to determine multivariate significance. If the test results show significant differences, further steps are taken with univariate analysis or post-hoc tests to identify which dependent variables contribute to these differences.

In its application, MANOVA has a number of statistical assumptions that must be met in order for the analysis results to be valid: i) multivariate normality of the dependent variables; ii) homogeneity of the covariance matrix between groups, which can be tested using Box's M test; iii) independence of observations; and iv) no high multicollinearity between dependent variables (Hair et al., 2021). If the assumptions of homogeneity or normality are violated, robust alternatives such as Pillai's Trace are more recommended due to their resistance to assumption violations. In the context of education, MANOVA is often used to evaluate the impact of new learning methods on several learning outcome indicators such as students' knowledge, skills, and attitudes simultaneously.

In the era of modern data analysis, MANOVA also forms the basis for advanced models such as MANCOVA (Multivariate Analysis of Covariance), which controls for the influence of covariates, and SEM (Structural Equation Modeling), which extends the analysis to latent relationships between variables (Warne, 2017). MANOVA can be implemented using software such as SPSS, R, SAS, or AMOS, which provide comprehensive analysis including assumption testing, main effects, interactions, and multivariate interpretation. Thus, MANOVA is a powerful analytical tool for understanding the effects of treatment on multiple interrelated outcomes comprehensively.

MANCOVA (Multivariate Analysis of Covariance)

Multivariate Analysis of Covariance (MANCOVA) is an extension of MANOVA (Multivariate Analysis of Variance) used to test differences in the means of several dependent variables simultaneously between groups, while controlling for the influence of one or more continuous covariates (Tabachnick & Fidell, 2019). Thus, MANCOVA combines the principles of ANOVA, linear regression, and covariance analysis, providing more accurate and efficient results in situations where dependent variables are correlated and influenced by additional factors that need to be controlled (Hair et al., 2021). In the context of educational and social research, MANCOVA is often used to evaluate the effect of treatment on multiple outcomes (e.g., academic achievement, motivation, and attitude) while controlling for initial differences between participants such as pretest scores or socioeconomic background.

Conceptually, MANCOVA adjusts the scores of the dependent variables based on covariates before comparing groups. This process produces adjusted means that represent the values of the dependent variables if all groups had the same covariate values (Grimm, Yarnold, & Peng, 2016). After adjustment, the model tests for significant differences in linear combinations of dependent variables between groups using multivariate statistics such as Wilks' Lambda, Pillai's Trace, Hotelling's Trace, and Roy's Largest Root (Olson, 1974). Significant results indicate that there are at least differences in the linear combinations of dependent variables that cannot be explained by covariates alone.

For MANCOVA results to be valid, several important statistical assumptions must be met, namely: (1) a linear relationship between covariates and each dependent variable; (2) homogeneity of regression slopes between groups; (3) multivariate normality of the dependent variables; (4) homogeneity of covariance matrices between groups, which can be tested using Box's M test; and (5) independence of observations (Field, 2018; Stevens, 2009). If the assumption of homogeneity of slopes is not met, MANCOVA is not appropriate to use, and a covariate-factor interaction model should be considered. In addition, to prevent type I error inflation, it is recommended that the number of covariates be relatively small compared to the total sample size.

In modern practice, MANCOVA is widely applied in experimental and quasi-experimental research in the fields of education, psychology, and social sciences, due to its ability to improve internal validity and control for confounding variables (Keselman et al., 1998). Statistical software such as SPSS, R, SAS, and AMOS provide comprehensive facilities for performing MANCOVA, including tests for main effects, interaction effects, and diagnostic assumption checks. In addition, MANCOVA also forms the basis for advanced methods such as Multilevel Modeling and Structural Equation Modeling (SEM), which enable the analysis of relationships between latent variables with complex covariates (Byrne, 2016). Thus, MANCOVA not only expands the inferential capabilities of researchers but also improves the accuracy of interpreting multivariate results in multidimensional research.

Qualitative Research Data Analysis Techniques (Taxonomy/Domain, Constant Comparative, Symbolic), Utilization of the NVIV-O Application

Taxonomy/Domain Analysis

Domain analysis is one of the main stages in ethnographic qualitative data analysis developed by James P. Spradley (1979). The purpose of this analysis is to identify the structure of meaning in a culture or social context based on conceptual categories called domains. Each domain includes one cover term that encompasses a number of included terms, which are linked by semantic relationships such as "is a type of" or "is part of." Thus, domain analysis helps researchers understand how participants organize their knowledge and experiences within a particular conceptual system.

The next stage, taxonomic analysis, deepens the discovered domain by breaking it down into hierarchical subcategories. This analysis aims to understand the internal structure of a domain and the relationships between more specific elements. Taxonomy describes the relationships between categories based on hierarchy and levels of inclusion, for example, how a general concept has more specific derivative concepts in a particular social context (Fetterman, 2019). In educational research, for example, taxonomy can help identify conceptual patterns about "effective learning strategies" based on the perceptions of teachers and students, and then group these components into a meaningful conceptual hierarchical structure.

Domain and taxonomy analysis is usually conducted iteratively alongside the data collection process through interviews, observations, or documents. This process is inductive, whereby categories are developed gradually from the data until saturation is reached. According to (Given, 2016), this approach provides a basis for systematic and organized qualitative analysis, while maintaining the depth of contextual interpretation. Currently, software such as NVivo or ATLAS.ti is often used to facilitate coding, visualization of taxonomy hierarchies, and exploration of relationships between categories, which strengthens the validity of findings through transparent documentation.

Analysis Constant Comparative

Constant comparative analysis is the core of the grounded theory approach, first developed by Glaser and Strauss (1967). The primary goal of this method is to develop theory rooted directly in field data through a continuous process of comparison between emerging data units, categories, and concepts. This process allows researchers to discover patterns, relationships, and meanings that emerge inductively without being influenced by pre-existing theories.

In practice, the constant comparative technique is carried out in four main stages: (1)comparing incidents applicable to each category, namely comparing each incident or unit of data with existing categories; (2)integrating categories and their properties, namely combining categories that have similarities or thematic relationships; (3)delimiting the theory, namely narrowing the focus to the most relevant and meaningful categories; and (4)writing the theory, where the relationships between categories are integrated into a conceptual model or substantive theory (Corbin & Strauss, 2015).

This method is iterative and dynamic, meaning data collection and analysis are conducted simultaneously. As new categories emerge, researchers continually compare them with older data to ensure consistency and deepen conceptual understanding. This approach also increases the credibility and validity of the analysis because each finding is consistently tested against the entire body of data (Charmaz, 2014). In the context of educational research, for example, constant comparative analysis can be used to understand the process of developing student learning motivation or teachers' strategies for creating active learning by grouping conceptual themes that develop from their experiences.

In the digital age, the use of qualitative analysis software such as NVivo or ATLAS.ti is very helpful in implementing this method. This software allows researchers to conduct coding, memoing, and cross-comparison efficiently, and visualize the relationships between emerging categories (Paulus et al., 2019). This approach not only enhances the transparency of the analysis process but also supports the reproducibility of findings and deeper interpretation of theory.

Symbolic Analysis

Symbolic analysis is a qualitative approach rooted in the theory of symbolic interactionism (symbolic interactionism) developed by Herbert Blumer (1969). This approach emphasizes that social meaning is formed through human interactions with symbols, language, actions, and objects that have symbolic value in a particular social context. In qualitative research, symbolic analysis is used to interpret how individuals give meaning to experiences, events, and social phenomena through the symbols used in everyday life (Charmaz, 2014).

This approach is highly relevant in the study of education, culture, and communication because it helps researchers understand the hidden dimensions of meaning behind social actions and interactions. For example, in the context of Islamic education, symbols such as uniforms, greetings, or religious activities at school have different meanings for students, teachers, and the community. Symbolic analysis allows researchers to explore how these symbols reflect spiritual values, collective identities, and the social structures underlying educational practices (Ramadhoni Aulia Gusli & Hamdi Abdul Karim, 2024).

In practice, the symbolic analysis process involves identifying, categorizing, and interpreting symbols that appear in data from interviews, observations, and documents. Researchers attempt to relate these symbols to broader social contexts and systems of meaning. Techniques such as coding, memoing, and narrative mapping is often used to explore the relationship between symbols and their meanings. With the help of qualitative analysis applications such as NVivo or ATLAS.ti, researchers can systematically organize symbolic data, mark meaningful themes, and visualize conceptual relationships between symbols and social meanings (Paulus, Lester, & Dempster, 2014).

This approach produces not only thematic descriptions but also a deep interpretive understanding of how meaning is constructed and negotiated in social contexts. Thus, symbolic analysis makes an important contribution to qualitative research oriented toward the exploration of meaning and social representations, rather than simply the description of phenomena.

Utilization of NVivo Application in Qualitative Data Analysis

NVivo is a computer-based qualitative data analysis software (Computer-Assisted Qualitative Data Analysis Software/ CAQDAS) that is the most widely used to support researchers in the process of organizing, coding, and interpreting non-numerical data such as interviews, observations, transcripts, field notes, and documents (Bazeley & Jackson, 2019). This application allows researchers to conduct systematic and transparent analysis of large amounts of data, as well as facilitate the application of various analytical approaches such as grounded theory, thematic analysis, content analysis, discourse analysis, and case study. NVivo helps maintain audit trail research, strengthening the credibility and validity of qualitative findings (Woolf & Silver, 2018).

In analytical practice, NVivo serves as a tool for conducting the coding process, which is the core of qualitative research. Researchers can mark data segments and group them into categories or themes (nodes) that are relevant. Features such as query, word frequency, text search, and matrix coding allows exploration of relationships between themes and patterns of meaning that emerge from the data. NVivo also supports visual analysis through models and mind maps, which makes it easier for researchers to describe the relationship between concepts or phenomena. This function is very helpful at the research stage. Constant comparative analysis and symbolic interpretation, where researchers must continuously compare and relate between categories or symbols.

Furthermore, the use of NVivo increases the efficiency and transparency of qualitative analysis. With features coding comparison, researchers can evaluate inter-coder reliability (inter-coder reliability) to ensure consistency of analysis results. In addition, NVivo supports the integration of multimedia data such as audio, video, and images, which enables a more intuitive approach. Multi-modal qualitative research (Paulus & Lester, 2016). In the context of educational research, for example, NVivo can be used to analyze teacher reflections, student interactions, and digital learning artifacts to gain a deeper understanding of pedagogical practices or technology implementation (Zamawe, 2015). Thus, NVivo is not only a technical tool but also a conceptual tool that strengthens the rigor, validity, and depth of interpretation of qualitative research.

CONCLUSION

The integration of advanced data analysis in quantitative and qualitative research is a strategic step in addressing the complexity of social, educational, and managerial phenomena in the digital era. This approach enables researchers to combine the strength of the objectivity of numerical data with the interpretative depth of qualitative data, resulting in a more comprehensive, valid, and contextual understanding. Various quantitative techniques, such as regression, multiple regression, principal component analysis (PCA), discriminant analysis, canonical analysis, path analysis, factor analysis, ANOVA, ANCOVA, MANOVA, and MANCOVA, play a crucial role in explaining the relationships, differences, and influences between variables empirically and measurably. On the other hand, qualitative analyses such as taxonomy/domain analysis, constant comparative analysis, and symbolic analysis provide a deep interpretative dimension to the meanings, values, and social symbols emerging from field data. The use of NVivo as qualitative analysis software enhances the coding process, organizing themes, and visualizing conceptual relationships efficiently and transparently. Thus, the combination of advanced statistical techniques and technology-based interpretive analysis not only enhances the validity and reliability of research results but also enriches the theoretical and practical dimensions in the development of modern science. Overall, this integrative application represents a new paradigm in scientific research, where quantitative and qualitative data are no longer seen as two separate approaches, but rather as two complementary sides in answering research questions in a more in-depth, comprehensive, and relevant way to the needs of the times.

REFERENCES

Asy'ari, A. A., Makalao, D. A. M., & Irawan, I. (2023). Analisis Metode Penelitian Kuantitatif dalam Manajemen Pendidikan Islam. *Tadbir: Jurnal Manajemen Pendidikan Islam, 11*(2), 152-175. https://doi.org/10.30603/tjmpi.v11i2.3796

Bazeley, & Jackson. (2019). Qualitative Data Analysis with NVivo (3rd ed.). *SAGE Publications*.

Black, & Wiliam. (2018). Assessment and Classroom Learning. In J. Gardner (Ed.), Assessment and Learning (2nd ed.). *SAGE Publications*.

Brown. (2015). *Confirmatory Factor Analysis for Applied Research (2nd ed.)*. Guilford Press.

Byrne. (2016). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming (3rd ed.)*. https://doi.org/10.4324/9781315757421

Corbin, & Strauss. (2015). Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory (4th ed.). *SAGE Publications*. https://journals.sagepub.com/home/qrj

Engkizar, E., Jaafar, A., Hamzah, M. I., Fakhruddin, F. M., Oktavia, G., & Febriani, A. (2023). Changes in Students' Motivation to Memorize the Quran: A Study at Quranic Higher Education Institutions in Indonesia. *International Journal of Islamic Studies Higher Education, 2*(3), 240–258. https://doi.org/10.24036/insight.v2i3.240

Engkizar, E., Jaafar, A., Hamzah, M. I., Langputeh, S., Rahman, I., & Febriani, A. (2025). Analysis Problems of Quranic Education Teachers in Indonesia: Systematic Literature Review. *International Journal of Islamic Studies Higher Education, 4*(2), 92–108. https://insight.ppj.unp.ac.id/index.php/insight

Engkizar, E., Jaafar, A., Sarianto, D., Ayad, N., Rahman, A., Febriani, A., Oktavia, G., Puspita, R., & Rahman, I. (2024). Analysis of Quran Education Problems in Majority Muslim Countries. *International Journal of Islamic Studies Higher Education, 3*(1), 65–80. https://doi.org/10.24036/insight.v3i1.209

Engkizar, E., Sarianti, Y., Namira, S., Budiman, S., Susanti, H., & Albizar, A. (2022). Five Methods of Quran Memorization in Tahfidz House of Fastabiqul Khairat Indonesia. *International Journal of Islamic Studies Higher Education, 1*(1), 54–67. https://doi.org/10.24036/insight.v1i1.27

Fauzi, M. H., Salsabila, S., Diniyati, A. I. L., Pebriani, A. R., Fithriya, R. A. I., & Suresman, E. (2025). Integrasi Nilai Islam dan Inovasi Pembelajaran dalam Pendidikan Agama Islam di Perguruan Tinggi dalam Perspektif Akademik dan Keagamaan. *Reflection: Islamic Education Journal, 2*(2), 186-196. https://doi.org/10.61132/reflection.v2i2.771

Ferdinan, F., Rahman, A., & Pewangi, M. (2024). Integrasi Nilai-Nilai Islam pada Supervisi Pendidikan Kepala Sekolah dalam Meningkatkan Kinerja Guru. *Didaktika: Jurnal Kependidikan, 13*(3), 4031-4044. https://doi.org/10.58230/27454312.713

Feri, F. R. D. (2025). Integrasi Metode Kualitatif dan Kuantitatif dalam Penelitian Manajemen Pendidikan Islam. *Islamic Management: Jurnal Manajemen Pendidikan Islam, 1*(2), 28-32. https://doi.org/10.63097/f75r7p71

Fetterman. (2019). Ethnography: Step-by-Step (4th ed.). *SAGE Publications*. https://journals.sagepub.com

Field. (2018). *Discovering Statistics Using IBM SPSS Statistics (5th ed.)*. SAGE Publications.

Frost. (2020). *Regression Analysis: An Intuitive Guide for Using and Interpreting Linear Models*. Statistics by Jim Publishing.

Given. (2016). 100 Questions (and Answers) About Qualitative Research. *SAGE Publications*.

Gravetter, & Wallnau. (2017). Statistics for the Behavioral Sciences (10th ed.). *Boston, MA: Cengage Learning*.

Hair, Black, Babin, & Anderson. (2019a). Multivariate Data Analysis (8th ed.). *Pearson*.

Hair, Black, Babin, & Anderson. (2019b). Multivariate Data Analysis (8th ed.). *Pearson Education*.

Htay, S. S., Po, E. T. H., & Kaewkanlaya, P. (2025). Building Student Character through Worship in Elementary Schools. *Muaddib: Journal of Islamic Teaching and Learning*, *1*(2), 55-63. https://muaddib.intischolar.id/index.php/muaddib/article/view/11

Jolliffe, & Cadima. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Kasheem, M., Yahya, N., Shalghoum, N., Masuwd, M., Alriteemi, A., Abdullah, M., Alsaeh, F., & Alrumayh, S. (2025). Artificial Intelligence in Academic Research: Adoption, Opportunities, and Barriers among Faculty in Libya Higher Education. *Multidisciplinary Journal of Thought and Research*, *1*(3), 109-127. https://mujoter.intischolar.id/index.php/mujoter/article/view/20

Kassymova, G. K., Talgatov, Y. K., Arpentieva, M. R., Abishev, A. R., & Menshikov, P. V. (2025). Artificial Intelligence in the Development of the Theory and Practices of Self-Directed Learning. *Multidisciplinary Journal of Thought and Research*, *1*(3), 66-79. https://mujoter.intischolar.id/index.php/mujoter/article/view/19

Li, Cheng, Wang, Morstatter, Trevino, Tang, & Liu. (2020). Feature selection: A data perspective. *ACM Computing Surveys*, *50*(6), 1–45. https://doi.org/10.1145/3136625

Miles, & Huberman. (2014). *Analisis Data Kualitatif*. Universitas Indonesia Press.

Miles, Huberman, & Saldana. (2014). *Analisis data kualitattif: buku sumber tentang metode-metode* (3rd ed.). UIPress.

Paulus, Lester, & Dempster. (2019). Digital Tools for Qualitative Research. *SAGE Publications*, *4*(6), 22–34. https://doi.org/10.4135/9781473957671

Ramadhoni Aulia Gusli, & Hamdi Abdul Karim. (2024). Application of School Financial Management in Managing the Bos Fund in Sdn 09 V Koto Kampung Dalam. *ICMIE Proceedings*, *1*(20), 199–206. https://doi.org/10.30983/icmie.v1i.13

Spradley. (2016). Participant Observation. *Waveland Press*.

Sugiyono. (2013). *Meto Penelitian Kualitatif Kuantitatif dan R&D*. Alfabeta.

Sugiyono. (2014). *Metode Penelitian Pendidikan Penelitian Kuantitatif, Kualitatif, Dan R&D*. Alfabeta.

Sugiyono. (2015). *Metode Penelitian Kombinasi (Mixed Methods)*. Alfabeta.

Sugiyono. (2017). *Metode Penelitian Kuantitatif, Kualitatif, Dan R&D*. Alfabeta.

Sugiyono. (2020). *Metode Penelitian Kualitatif Bandung*. Alfabeta.

Tabachnick, & Fidell. (2019). Using Multivariate Statistics (7th ed.). *Pearson*.

Tashakkori, & Teddlie. (2021). Mixed Methods Research: Integrating Quantitative and Qualitative Approaches in the Social and Behavioral Sciences. *SAGE Publications*.

Uyubah, M., & Anawati, S. (2024). Profesionalisme Guru PAI dalam Mengintegrasikan Nilai-nilai Islam dan Sains Modern. Jurnal Ilmiah dan Penelitian, 2(1).

Warne. (2017). A primer on multivariate analysis of variance (MANOVA) for behavioral scientists. *Practical Assessment, Research, and Evaluation*, *19*(17), 1–10. https://doi.org/10.7275/sm63-7h70

Woolf, & Silver. (2018). Qualitative Analysis Using NVivo: The Five-Level QDA Method. *Routledge/Taylor & Francis*. https://doi.org/10.4324/9781351723135

Zamawe. (2015). The implication of using NVivo software in qualitative data analysis: Evidence-based reflections. *Malawi Medical Journal*, *27*(1), 13–15. https://doi.org/10.4314/mmj.v27i1.4