Journal of Research in Didactical Sciences
ISSN: 2997-5905 (Online)
VOL. 5, ISSUE 1, 2026
Pages: 77-91

RESEARCH ARTICLE

OPEN ACCESS

Authentic Learning of Correlation and Regression Through Gamified FIFA World Cup Ball Possession and Goal-Scoring Data

Tirivanhu Muchuweni [1*]

[1] Rhodes University, South Africa
* Corresponding author: tmuchuweni@hotmail.com

Abstract: Authentic datasets can make statistical concepts more meaningful by connecting classroom learning with real-world situations. However, limited attention has been given to translating authentic football match statistics into structured gamified activities for teaching correlation and regression in secondary mathematics. This paper examines 2026 FIFA World Cup ball-possession and goal-scoring data and uses the analysis to inform a proposed gamified activity for teaching correlation and regression in secondary mathematics. Guided by Authentic Learning Theory, the study employed a quantitative correlational design using publicly available 2026 FIFA World Cup Round of 32 statistics comprising 32 team observations. No students participated, no classroom intervention was conducted, and no student learning outcomes were measured. Pearson correlation and simple linear regression were used to analyse the relationship between ball-possession percentage and goals scored. The findings revealed a weak positive relationship between ball-possession percentage and goals scored that was not statistically significant ($r = .290, p = .108$), with the regression model explaining 8.4% of the variation in goals scored ($R^2 = .084$). These findings describe an association in the match data; they do not demonstrate that gamification improves student learning. The study contributes the proposed five-stage SCORE Instructional Model, which integrates authentic sports data, statistical reasoning and gamification into a structured approach for teaching correlation and regression. The model offers teachers a potential framework for designing meaningful statistical investigations, although its effects on student learning and engagement require classroom evaluation.

Keywords: Authentic learning; Correlation and regression; Gamification; Mathematics education; Statistical reasoning

INTRODUCTION

Helping learners understand correlation and regression is an important goal of secondary mathematics because these concepts enable students to interpret relationships between variables, analyse data, and make evidence-based decisions in real-world situations (Engel & Sedlmeier, 2011; Garrett et al., 2016; Tsarev et al., 2025). Many learners find these statistical concepts abstract when they are presented through textbook exercises that lack meaningful contexts, making it difficult to connect mathematical ideas with everyday experiences (Dierdorp, 2013; Garrett et al., 2016). Authentic learning addresses this challenge by engaging learners with realistic problems and real-world data that encourage active exploration, critical thinking, and meaningful application of mathematical concepts (Herrington, 2006; Lombardi & Oblinger, 2007; Mims, 2003; Rule, 2006).

© 20236 The Author(s). This article is distributed under the Creative Commons Attribution 4.0 International License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction, as long as the original author(s) and source are credited and any modifications are indicated. The data and materials are also dedicated to the public domain under the CC0 1.0 waiver, unless otherwise stated.

Sport provides one of the richest sources of authentic data for mathematics classrooms because it naturally generates numerical information that learners encounter and discuss beyond school (Garrett et al., 2016; Herrington et al., 2013). In this paper, the term football refers to association football (soccer) to distinguish it from American football. Football, in particular, offers a wide range of statistics, including ball possession, passes, shots, and goals, which can be analysed using statistical methods (Collet, 2013; Lago & Martín, 2007; Wang et al., 2022). Previous FIFA World Cup Football tournaments have generated extensive datasets that have been used to examine performance indicators, tactical patterns, and goal-scoring behaviours, demonstrating the educational value of authentic sports data for quantitative analysis (Almeida et al., 2025; Degrenne & Carling, 2024; Kubayi, 2020; Soroka et al., 2023).

At the same time, gamification has emerged as an effective instructional approach for increasing learners' motivation, participation, and engagement in mathematics (Alt, 2023; Dichev & Dicheva, 2017; Hamari et al., 2014; Landers, 2014). Gamification refers to the use of game design elements in non-game learning environments to enhance learner motivation, engagement, and participation (Deterding et al., 2011; Landers, 2014). Within mathematics education, this approach can make statistical concepts more interactive by encouraging learners to investigate patterns, test ideas, and interpret evidence from authentic datasets (Alt, 2023; Hamari et al., 2014). Recent studies (Jack et al., 2025; Lee & Lai, 2024; Muchuweni & Jojo, 2026a, 2026b; Muchuweni et al., 2025) have also reported positive effects of gamification on mathematics achievement, engagement, and statistical learning, suggesting that authentic data and game-based learning can complement one another in mathematics classrooms.

Despite growing interest in gamified mathematics instruction and the increasing availability of authentic sports data, limited attention has been given to using 2026 FIFA World Cup ball-possession and goal-scoring data to design activities for teaching correlation and regression. Most football studies have focused on analysing team performance or identifying factors associated with match outcomes while giving little attention to demonstrating how these datasets can support the teaching of statistical concepts in secondary mathematics (Jones et al., 2004; Kapsalis et al., 2023; Lago-Peñas et al., 2010; Taha & Ali, 2023). This points to a need for instructional examples showing how authentic match data can be used to teach statistical concepts; whether such activities improve student learning requires classroom evaluation. Against this background, this paper analyses 2026 FIFA World Cup Round of 32 ball-possession and goal-scoring data and uses the analysis to inform a proposed gamified activity for teaching correlation and regression in secondary mathematics. Specifically, the paper seeks to answer the following research question: *How can analysis of 2026 FIFA World Cup Round of 32 ball-possession and goal-scoring data inform a proposed gamified activity for teaching correlation and regression in secondary mathematics*?

LITERATURE REVIEW

Authentic learning encourages learners to develop mathematical understanding by working with meaningful tasks that resemble real-world situations as opposed to isolated classroom exercises (Garrett et al., 2016; Herrington, 2006; Lombardi & Oblinger, 2007; Rule, 2006). Meaningful learning experiences become more effective when students engage with genuine datasets that require them to organise, analyse, and interpret information while constructing their own understanding of mathematical relationships (Dierdorp, 2013; Garrett et al., 2016). Mims (2003) further explains that authentic learning provides opportunities for learners to apply classroom knowledge to realistic situations, helping them appreciate the practical value of mathematics beyond examinations.

Gamification provides an instructional approach that can make authentic mathematical tasks more engaging by incorporating game elements such as challenges, feedback, competition, and achievement into learning activities (Deterding et al., 2011; Hamari et al., 2014; Landers, 2014). Dichev and Dicheva (2017) argue that gamification has considerable educational potential when it supports meaningful learning in contrast to simply increasing enjoyment. Similarly, Alt (2023), Diaz and Estoque-Loñez (2024), and Li et al. (2024) report that well-designed gamified activities can improve learner motivation, engagement, and participation in mathematics classrooms. Recent work has also demonstrated that gamified learning environments support mathematical argumentation, classroom interaction, and statistical learning, creating opportunities for learners to investigate mathematical ideas through active participation (Jack et al., 2025; Lee & Lai, 2024; Lo & Hew, 2020). This evidence is reinforced by recent mathematics education studies highlighting the instructional value of digital game-based platforms for promoting learner engagement and meaningful mathematical participation (Muchuweni & Jojo, 2026a, 2026b; Muchuweni et al., 2025).

Correlation and regression are most meaningful when learners analyse authentic datasets that contain observable relationships between variables, unlike artificial examples created solely for instruction (Dierdorp, 2013; Engel & Sedlmeier, 2011; Tsarev et al., 2025). Asuero et al. (2006), Franzese and Iuliano (2018), and Zou et al. (2003) explain that these statistical techniques enable learners to measure the strength of relationships, identify patterns, and make informed predictions from data. Developing these competencies is very important as data interpretation and evidence-based reasoning continue to play a central role across many disciplines, including sport.

Football offers an appropriate context for learning these concepts because match statistics naturally generate paired numerical variables that can be explored using correlation and regression (Dierdorp, 2013; Engel & Sedlmeier, 2011; Wang et al., 2022). Previous studies have examined relationships among ball possession, passing, shots, and goal scoring to understand team performance and match success (Collet, 2013; Hughes & Franks, 2005; Jones et al., 2004; Wang et al., 2022). Analyses of FIFA World Cup competitions have similarly investigated tactical patterns, possession behaviour, and goal-scoring

opportunities across different tournaments (Almeida et al., 2025; Degrenne & Carling, 2024; Kubayi, 2020; Soroka et al., 2023). While these studies provide valuable insights into football performance, their primary focus remains on sports analysis, with little attention given to mathematics teaching.

The reviewed literature therefore indicates an opportunity to connect authentic learning, gamification, and statistical reasoning within a single classroom activity. Although authentic learning and gamification have both demonstrated value in mathematics education, and football research has produced rich datasets suitable for statistical analysis, limited attention has been given to using FIFA World Cup ball possession and goal-scoring data as a gamified context for teaching correlation and regression in secondary mathematics. Addressing this gap may provide learners with a more meaningful way of developing statistical understanding while demonstrating how mathematics can be applied to real-world situations.

THEORETICAL FRAMEWORK

This study is guided by Authentic Learning Theory, proposed by Herrington and Oliver (2000), which explains that meaningful learning occurs when learners engage with realistic tasks that resemble situations encountered outside the classroom. Instead of learning concepts through isolated examples, learners develop deeper understanding by investigating real-life problems, analysing genuine data, and applying knowledge to practical situations (Garrett et al., 2016; Lombardi & Oblinger, 2007; Rule, 2006). In this paper, 2026 FIFA World Cup ball-possession and goal-scoring data provide an authentic context for a proposed activity on correlation and regression. The activity was not implemented with students.

Authentic Learning Theory also emphasises active learner participation, collaboration, reflection, and the construction of knowledge through inquiry (Herrington & Oliver, 2000; Herrington et al., 2013). The proposed activity uses gamification through team challenges, points, feedback, and collaboration (Deterding et al., 2011; Hamari et al., 2014; Landers, 2014). Learners would use the match data to create a scatter plot, calculate and interpret correlation and regression, and explain what the results show.

METHODOLOGY

This study used a quantitative correlational design to examine the association between 2026 FIFA World Cup ball-possession percentage and goals scored. The analysis informed a proposed gamified activity for teaching correlation and regression in secondary mathematics; the study did not evaluate an implemented activity. A correlational design was considered appropriate because it allows the relationship between variables in authentic data to be examined (Asuero et al., 2006; Franzese & Iuliano, 2018; Zou et al., 2003).

The dataset consisted of 16 Round of 32 matches, representing 32 participating teams and 32 team observations. Each observation included two variables: ball possession percentage and goals scored. The

Round of 32 was purposively selected because it provided a manageable dataset of 32 team observations for a proposed classroom activity while retaining the competitive nature of knockout-stage football. Using the entire group stage would have produced a considerably larger dataset that was less practical for the proposed activity, whereas the Round of 16 would have resulted in fewer observations, limiting opportunities for learners to explore data patterns and relationships within that activity.

The study used publicly available 2026 FIFA World Cup match statistics retrieved from the Quantasys Flow Match Reports website (*https://wc26.com.au/matchreports.html*). No human participants, personal information, or confidential information were involved. The researchers organized the data, constructed scatter plots, calculated the Pearson correlation coefficient, developed a simple linear regression model, and interpreted the relationship between ball possession and goal scoring. These analyses informed the proposed classroom activity; learners did not conduct them as part of this study.

The collected data were analysed using descriptive statistics, scatter plots, Pearson correlation analysis, and simple linear regression. Descriptive statistics summarised the characteristics of the dataset, while the scatter plot provided a visual representation of the relationship between the variables. Pearson correlation measured the strength and direction of the relationship between ball possession and goals scored, and simple linear regression examined the extent to which ball possession predicted goal-scoring performance. Regression diagnostics were examined using the Shapiro–Wilk and Breusch–Pagan tests. Match-clustered standard errors were used to account for paired team observations, while Poisson regression was conducted as a sensitivity analysis because goals scored is a discrete count variable.

RESULTS AND DISCUSSION

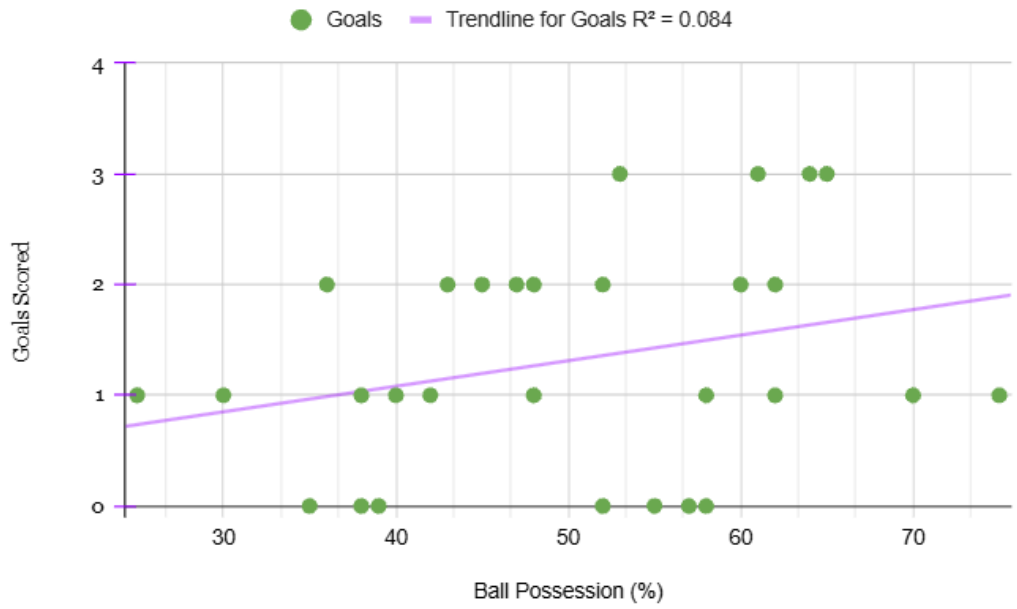
Data from 32 team observations obtained from the 2026 FIFA World Cup Round of 32 were analysed to examine the relationship between ball possession percentage and goals scored. Ball possession percentage was treated as the independent variable, while goals scored represented the dependent variable. Pearson correlation and simple linear regression were performed to determine the relationship between the two variables. Table 1 presents the descriptive statistics.

Table 1. Descriptive Statistics for Ball Possession and Goals Scored ($N = 32$)

Statistic	Ball possession (%)	Goals scored
Mean	50	1.31
Standard deviation	12.11	0.97
Minimum	25	0
Maximum	75	3

Figure 1 presents the scatterplot of ball possession percentage and goals scored for the 32 team observations, together with the fitted regression line. The plot suggests a weak positive relationship between the variables.

Figure 1. Ball possession and goals scored in the 2026 FIFA World Cup Round of 32 (*Author's own elaboration*)



Pearson correlation analysis revealed a weak positive relationship between ball possession and goals scored, $r(30) = .290, p = .108$. The relationship was not statistically significant at the .05 level. Because opposing teams belonged to the same matches, this result was interpreted cautiously. Simple linear regression was then performed to estimate the relationship between the two variables.

$$\hat{Y} = \beta_0 + \beta_1 X$$

where:

$$\hat{Y} = \ predicted\ goals\ scored,$$
$$\beta_0 \ = \ regression\ intercept,$$
$$\beta_1 \ = \ regression\ slope, and$$
$$X \ = \ ball\ possession\ percentage.$$

The fitted regression model obtained from the analysis was:

$$\hat{Y} = 0.1581 + 0.0231X$$

The regression slope indicates that each additional percentage point of ball possession was associated with an estimated increase of 0.0231 goals scored. The model explained 8.4% of the variation in goals scored ($R^2 = .084$), indicating limited predictive power. This association does not establish causation, as other factors, including team strength, shooting accuracy and defensive performance, may also influence goal scoring.

Because the 32 observations represented opposing teams in 16 matches, the assumption of independence required examination. Opposing teams had complementary possession percentages, meaning their observations were not fully independent. The regression was therefore re-examined using match-clustered standard errors. The positive association was not statistically significant ($p = .126$). Diagnostic tests showed no statistically significant departure from residual normality (Shapiro–Wilk, $p = .260$) but indicated evidence of heteroscedasticity (Breusch–Pagan,$p = .045$). However, these results do not establish that all regression assumptions were satisfied.

Goals scored is also a discrete count variable, which presents limitations for ordinary least squares regression. Although linear regression provides an accessible instructional approach for examining statistical relationships, it can produce unrealistic negative predictions and may not fully represent the distribution of goal counts. A Poisson regression was therefore conducted as a sensitivity analysis and indicated a positive but statistically nonsignificant association between ball possession and goals scored ($\beta = 0.0177, p = .090$). Ordinary least squares was retained as the primary instructional model because the proposed activity focuses on introductory correlation and linear regression.

The weak positive correlation ($r = .290$) indicates that teams with higher ball possession tended to score slightly more goals, although the relationship was not statistically significant. The regression model explained 8.4% of the variation in goals scored, leaving considerable variation unexplained. These findings provide an opportunity for learners to examine the strength of statistical relationships, interpret predictions and distinguish association from causation. Using FIFA World Cup data for these investigations reflects Authentic Learning Theory by connecting statistical reasoning to realistic problems (Garrett et al., 2016; Herrington & Oliver, 2000; Herrington et al., 2013).

The proposed activity would involve learners in collecting, organising, representing and interpreting match statistics through team challenges, collaboration and feedback. These tasks reflect inquiry and knowledge construction within authentic learning environments (Herrington & Oliver, 2000; Rule, 2006), while incorporating gamification elements associated with learner engagement and participation in previous research (Alt, 2023; Deterding et al., 2011; Hamari et al., 2014; Landers, 2014). However, the present study did not investigate whether these activities improve student engagement or mathematical understanding.

Previous football studies have examined possession and goal-scoring statistics primarily to understand team performance and match outcomes (Collet, 2013; Hughes & Franks, 2005; Wang et al., 2022). This study proposes an instructional application of such data to secondary mathematics. Teachers could use the observed relationship to develop questions about prediction accuracy, unexplained variation and the limitations of statistical models. Although this approach offers a possible alternative to textbook datasets, its educational effectiveness requires classroom implementation and evaluation.

To translate the statistical analysis into classroom practice, the proposed SCORE Instructional Model was developed. The model provides a structured approach to teaching correlation and regression through statistical investigation, collaboration, feedback and assessment. The following section presents its five stages and illustrates their potential classroom application.

DEVELOPMENT OF THE PROPOSED SCORE INSTRUCTIONAL MODEL

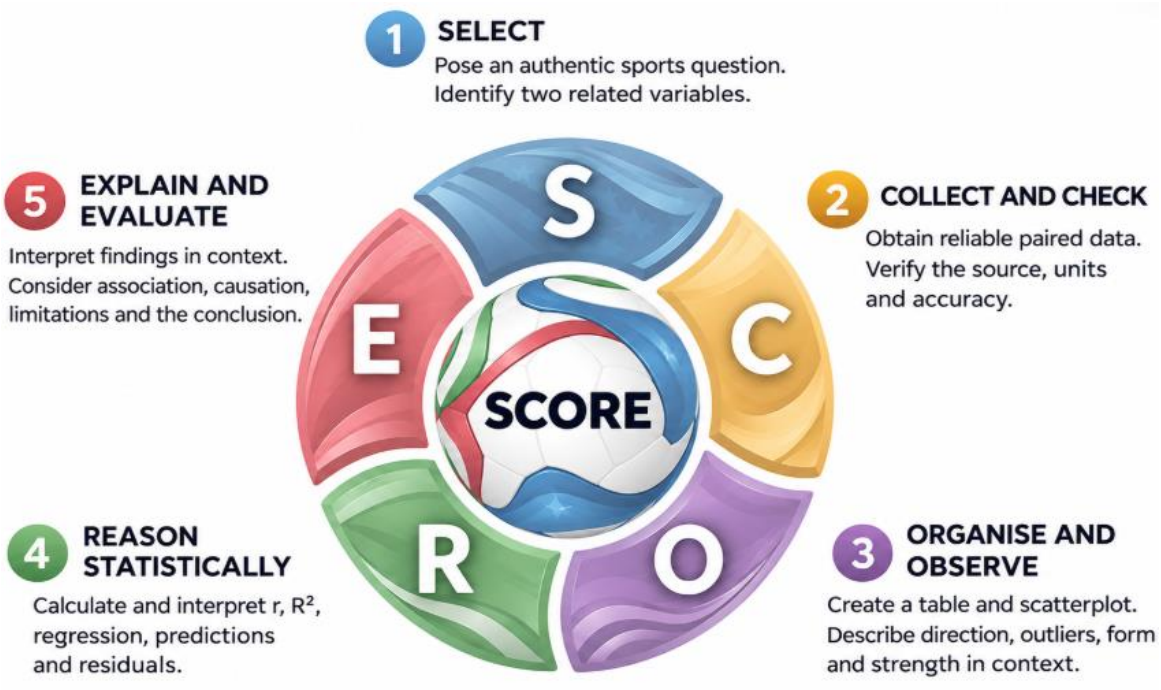
The SCORE Instructional Model was developed to show how authentic sports data can be translated into a structured activity for teaching correlation and regression. The model is informed by Authentic Learning Theory, which emphasises meaningful contexts, realistic tasks, inquiry, collaboration, reflection and the application of knowledge (Herrington & Oliver, 2000; Herrington et al., 2013). It also incorporates gamification through team challenges, points, feedback and the presentation of findings. The model was developed conceptually from the theoretical framework, reviewed literature and statistical analysis presented in this paper. It has not yet been implemented or empirically validated. Unlike activities that focus primarily on analysing authentic datasets, the SCORE Model explicitly connects data selection, verification, visual exploration, statistical reasoning and evidence-based communication within a five-stage gamified structure. Its contribution lies in organising these processes into a flexible instructional sequence that incorporates team challenges, feedback and assessment.

SCORE represents five connected stages: Select, Collect and Check, Organise and Observe, Reason Statistically, and Explain and Evaluate. In the Select stage, the teacher introduces an authentic sporting context, and learners identify a meaningful question and two related variables. In the Collect and Check stage, learners obtain paired data from a reliable source, confirm the units of measurement and check the data for completeness and accuracy. In the Organise and Observe stage, learners arrange the observations in a table, construct a scatterplot and describe the direction, outliers, form and strength of the relationship.

During the Reason Statistically stage, learners calculate and interpret the correlation coefficient, coefficient of determination, regression equation and residuals. They use the regression model to make predictions within the observed range of the data. In the final Explain and Evaluate stage, learners interpret their findings in the sporting context, distinguish association from causation, identify limitations and communicate an evidence-based conclusion.

The model incorporates gamification across the five stages. Learners work in teams, earn points for accurate analysis and clear explanations, and receive teacher feedback as they complete each stage. Groups may present their findings through a poster or short presentation. Assessment focuses on data accuracy, statistical reasoning, contextual interpretation, collaboration and the clarity of the final conclusion. Figure 2 presents the proposed SCORE Instructional Model.

Figure 2. The proposed SCORE Instructional Model for teaching correlation and regression through authentic sports data (*Author's own elaboration*)



Illustrative Application of the SCORE Model

The 2026 FIFA World Cup investigation provides an illustrative application of the SCORE Model. The activity aims to enable learners to organise paired data, construct and interpret scatterplots, calculate and interpret correlation and regression statistics, make predictions, and evaluate statistical relationships within an authentic sporting context. In the Select stage, learners investigate the question: *Can ball possession predict goals*? They identify ball-possession percentage as the explanatory variable and goals scored as the response variable. During Collect and Check, groups use the 32 team observations from the 16 Round of 32 matches and verify the data using the stated source.

In Organise and Observe, learners construct a scatterplot and describe the relationship in context. During Reason Statistically, they calculate and interpret the correlation coefficient (r), coefficient of determination (R^2), regression equation and selected residuals. In Explain and Evaluate, groups determine whether ball possession provides useful predictive information, explain why the relationship does not establish causation and identify other factors that may influence goal scoring. Teams earn points for accurate calculations, clear interpretations and effective use of evidence. Each group produces a poster or short presentation, which the teacher assesses using a rubric covering data accuracy, scatterplot construction, statistical calculations, interpretation of findings, collaboration and clarity of presentation. Feedback is provided throughout the five stages to support learners' statistical reasoning and understanding. To support reproducibility and flexibility, teachers can implement the SCORE Model

using suitable authentic sports datasets, a graphing tool and a structured worksheet aligned with its five stages.

To address common difficulties in understanding correlation and regression, the activity includes guiding questions such as: *Does higher ball possession always result in more goals*? *What does the correlation coefficient indicate about the relationship*? *Can the regression equation accurately predict goals scored?* Learners would be expected to recognise that the weak positive correlation was not statistically significant, that higher possession does not necessarily result in more goals, that predictions are estimates rather than exact outcomes, and that correlation does not establish causation. These questions are intended to help learners distinguish statistical relationships from causal explanations and interpret regression results within the limitations of the data.

Although the FIFA World Cup data provide the illustrative context in this paper, the SCORE Model can also be adapted to other paired sports data, such as field-goal percentage and points scored in basketball, attack efficiency and sets won in volleyball, or first-serve percentage and service points won in tennis. These applications remain proposals for future classroom implementation and evaluation.

CONCLUSION

This paper analysed FIFA World Cup ball-possession and goal-scoring data to examine their statistical relationship and inform a proposed gamified activity for teaching correlation and regression in secondary mathematics. The analysis revealed a weak positive relationship that was not statistically significant. By using authentic match statistics, the proposed activity could give learners an opportunity to explore statistical relationships using data that are meaningful, relevant, and connected to real-world events. Because no classroom intervention was conducted and student learning was not measured, the findings do not demonstrate that the activity improves students' understanding or that gamification is effective.

The analysis addresses the research question by showing how FIFA World Cup data can inform the design of a proposed instructional activity. Beyond providing a context for teaching statistical concepts, this approach illustrates how real-world datasets may connect mathematics with learners' everyday experiences. The study therefore contributes the proposed five-stage SCORE Instructional Model, which integrates authentic learning, gamification and statistical reasoning into a structured approach for teaching correlation and regression in secondary mathematics. Its effects on student learning and engagement require classroom evaluation.

Implications for Classroom Instruction and Recommendations

The findings suggest that mathematics teachers can use authentic datasets from major sporting events to support the teaching of statistical concepts in meaningful contexts. Using FIFA World Cup ball-

possession and goal-scoring data could provide learners with opportunities to apply correlation and regression to real situations, helping them connect classroom mathematics with everyday experiences. Incorporating gamified learning activities alongside real-life datasets may also encourage greater learner participation, curiosity, and statistical reasoning.

Future studies should implement this learning approach in secondary mathematics classrooms to examine its effects on learners' understanding of correlation and regression, engagement, and achievement. Researchers may also extend this approach by using other authentic datasets from sport, science, economics, or environmental studies to investigate how real-world contexts can support the teaching and learning of statistics.

Limitation of the study

This study was limited to 32 team observations from the FIFA World Cup Round of 32, which means the findings should not be generalised to all football matches or all mathematics classrooms. The study was therefore positioned as a focused statistical investigation informing a proposed classroom activity rather than an evaluation of instructional effectiveness. Furthermore, the 32 observations came from 16 matches involving opposing teams whose possession percentages were complementary. This dependence was considered through match-clustered standard errors, although the small number of matches limits the strength of statistical inference. The study used only two variables: ball possession percentage and goals scored. Other match factors, such as shots on target, passing accuracy, team strength, and defensive strategies, were not included. The analysis was deliberately restricted to these two variables to maintain a manageable focus on correlation and regression. The study also used publicly available match statistics and did not test the activity directly with learners. Consequently, no conclusions can be drawn about actual student learning, learner engagement, or the effectiveness of gamification. The proposed SCORE Instructional Model requires classroom implementation and empirical evaluation before its educational effectiveness can be established.

REFERENCES

Almeida, C. H., Paixão, P., Jorge, A., Vargas, P., Gonçalves, R., & Batalau, R. (2025). Impact of situational variables on goal-scoring offensive sequences in the 2022 FIFA World Cup. *Journal of Human Kinetics, 100*, 237. https://doi.org/10.5114/jhk/202563

Alt, D. (2023). Assessing the benefits of gamification in mathematics for student gameful experience and gaming motivation. *Computers & Education, 205*, 104806. https://doi.org/10.1016/j.compedu.2023.104806

Asuero, A. G., Sayago, A., & González, A. G. (2006). The correlation coefficient: An overview. *Critical Reviews in Analytical Chemistry, 36*(1), 41-59. https://doi.org/10.1080/10408340500526766

Collet, C. (2013). *The possession game? A comparative analysis of ball retention and team success in European and international football, 2007–2010*. *Journal of Sports Sciences, 31*(2), 123-136. https://doi.org/10.1080/02640414.2012.727455

Degrenne, O., & Carling, C. (2024). *Comparison of goalscoring patterns between the 2018 and 2022 FIFA World Cups*. *Frontiers in Sports and Active Living, 6*, 1394621. https://doi.org/10.3389/fspor.2024.1394621

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. *Proceedings of the 15th International Academic MindTrek Conference*, 9–15. https://doi.org/10.1145/2181037.2181040

Dichev, C., & Dicheva, D. (2017). *Gamifying education: What is known, what is believed and what remains uncertain: A critical review*. *International Journal of Educational Technology in Higher Education, 14*(1), 1–36. https://doi.org/10.1186/s41239-017-0042-5

Diaz, A. F., & Estoque-Loñez, H. (2024). A meta-analysis on the effectiveness of gamification on student learning achievement. *International Journal of Education in Mathematics, Science and Technology, 12*(4), 1095–1116. https://doi.org/10.46328/ijemst.4185

Dierdorp, A. (2013). *Learning correlation and regression within authentic sciences* (Doctoral dissertation, Utrecht University).

Engel, J., & Sedlmeier, P. (2011). Correlation and regression in the training of teachers. In *Teaching Statistics in School Mathematics-Challenges for Teaching and Teacher Education: A Joint ICMI/IASE Study: The 18th ICMI Study* (pp. 247-258). Dordrecht: Springer Netherlands. https://doi.org/10.1007/978-94-007-1131-0_25

Franzese, M., & Iuliano, A. (2018). Correlation analysis. In *Encyclopedia of bioinformatics and computational biology: ABC of bioinformatics* (Vol. 1, pp. 706-721). Elsevier. https://doi.org/10.1016/b978-0-12-809633-8.20358-0

Garrett, L., Huang, L., & Charleton, M. C. (2016). A framework for authenticity in the mathematics and statistics classroom. *The Mathematics Educator, 25*(1). https://doi.org/10.63301/tme.v25i1.2012

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). https://doi.org/10.1109/hicss.2014.377

Herrington, J. (2006, October). Authentic e-learning in higher education: Design principles for authentic learning environments and tasks. In *eLearn: World Conference on EdTech* (pp. 3164-3173). Association for the Advancement of Computing in Education (AACE).

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. *Educational technology research and development, 48*(3), 23-48. https://doi.org/10.1007/bf02319856

Herrington, J., Reeves, T. C., & Oliver, R. (2013). Authentic learning environments. *Handbook of research on educational communications and technology*, 401-412. https://doi.org/10.1007/978-1-4614-3185-5_32

Hughes, M., & Franks, I. (2005). Analysis of passing sequences, shots and goals in soccer. *Journal of Sports Sciences, 23*(5), 509-514. https://doi.org/10.1080/02640410410001716779

Jack, E., Alexander, C., & Jones, E. M. (2025). Exploring the impact of gamification on engagement in a statistics classroom. *Teaching Mathematics and Its Applications, 44*(2), 156–172. https://doi.org/10.1093/teamat/hrae009

Jones, P. D., James, N., & Mellalieu, S. D. (2004). Possession as a performance indicator in soccer. *International Journal of Performance Analysis in Sport, 4*(1), 98-102. https://doi.org/10.1080/24748668.2004.11868295

Kapsalis, M., Plakias, S., Kyranoudis, A., Zarkadoula, A., Lathoura, A., & Tsatalas, T. (2023). Exploring the impact of possession-based performance indicators on goal scoring in elite football leagues. *Journal of Physical Education and Sport, 23*(8), 2004-2015.

Kubayi, A. (2020). Analysis of goal scoring patterns in the 2018 FIFA World Cup. *Journal of Human Kinetics, 71*, 205. https://doi.org/10.2478/hukin-2019-0084

Lago, C., & Martín, R. (2007). Determinants of possession of the ball in soccer. *Journal of Sports Sciences, 25*(9), 969-974. https://doi.org/10.1080/02640410600944626

Lago-Peñas, C., Lago-Ballesteros, J., Dellal, A., & Gómez, M. (2010). *Game-related statistics that discriminated winning, drawing and losing teams from the Spanish soccer league*. *Journal of Sports Science & Medicine, 9*(2), 288.

Landers, R. N. (2014). Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming, 45*(6), 752–768. https://doi.org/10.1177/1046878114563660

Lee, W.-C., & Lai, C.-L. (2024). Facilitating mathematical argumentation by gamification: A gamified mobile collaborative learning approach for math courses. *International Journal of Science and Mathematics Education, 22*(1), 211–232. https://doi.org/10.1007/s10763-024-10462-6

Li, L., Hew, K. F., & Du, J. (2024). *Gamification enhances student intrinsic motivation, perceptions of autonomy and relatedness, but minimal impact on competency: A meta-analysis and systematic review*. *Educational Technology Research and Development, 72*(2), 1561–1590. https://doi.org/10.1007/s11423-023-10337-7

Lo, C. K., & Hew, K. F. (2020). A comparison of flipped learning with gamification, traditional learning, and online independent study: The effects on students' mathematics achievement and cognitive engagement. *Interactive Learning Environments, 28*(4), 464–481. https://doi.org/10.1080/10494820.2018.1541910

Lombardi, M. M., & Oblinger, D. G. (2007). Authentic learning for the 21st century: An overview. *Educause learning initiative, 1*(2007), 1-12.

Mims, C. (2003). Authentic learning: A practical introduction & guide for implementation. *Meridian: A Middle School Computer Technologies Journal, 6*(1), 1-3.

Muchuweni, T., & Jojo, Z. (2026a). The culture of play: How Kahoot suggests a reconfiguration of participation and relationships in secondary mathematics classrooms. *International Electronic Journal of Mathematics Education, 21*(3), Article em0887. https://doi.org/10.29333/iejme/18852

Muchuweni, T., & Jojo, Z. (2026b). Gimkit's instructional value in secondary mathematics education: An integrative review. *International Journal of Learning, Teaching and Educational Research, 25*(3), 1051–1074. https://doi.org/10.26803/ijlter.25.3.47

Muchuweni, T., Jojo, Z., & Kariyana, I. (2025). Enhancing mathematics instruction through Quizizz: A systematic literature review. *International Journal of Learning, Teaching and Educational Research, 24*(10), 106–124. https://doi.org/10.26803/ijlter.24.10.5

Rule, A. C. (2006). Editorial: The components of authentic learning. *Journal of Authentic Learning, 3(1), 1–10.*

Soroka, A., Duda, H., Stuła, A., Ambroży, T., Kromke, C., & Te Poel, H. D. (2023). *Ball possession as an indicator identifying differences in the efficient operation of football teams during the world cup-Qatar 2022*. *Journal of Kinesiology and Exercise Sciences*, *33*(102), 9-20. https://doi.org/10.5604/01.3001.0053.5968

Taha, T., & Ali, A. Y. (2023). Greater numbers of passes and shorter possession durations result in increased likelihood of goals in 2010 to 2018 World Cup Champions. *PLoS One, 18*(1), e0280030. https://doi.org/10.1371/journal.pone.0280030

Tsarev, R., Senapati, B., García, D. S., Islamova, T., Nikulushkin, A., & Muracova, N. (2025, October). Correlation and Regression analysis of student grades and predicting academic performance. In *Software Engineering: Emerging Trends and Practices in System Development: Proceedings of 14th*

Computer Science On-line Conference 2025 (Vol. 3, p. 250). Springer Nature. https://doi.org/10.1007/978-3-032-00239-6_17

Wang, S. H., Qin, Y., Jia, Y., & Igor, K. E. (2022). A systematic review about the performance indicators related to ball possession. *PLoS One, 17*(3), e0265540. https://doi.org/10.1371/journal.pone.0265540

Zou, K. H., Tuncali, K., & Silverman, S. G. (2003). Correlation and simple linear regression. *Radiology, 227*(3), 617-628. https://doi.org/10.1148/radiol.2273011499

Please Cite: Muchuweni, T. (2026). Authentic Learning of Correlation and Regression Through Gamified FIFA World Cup Ball Possession and Goal-Scoring Data. *Journal of Research in Didactical Sciences,* 5(1), 77-91. https//doi.org/10.51853/jorids/515

Received: 07.02.2026 ▪ Accepted: 17.11.2026