

International Journal of Review in Mathematics Education

Department of Mathematics Education, Universitas Muhammadiyah Surakarta, Indonesia
<https://journals2.ums.ac.id/ijrime/index>



Time Series Analysis of Mathematics Results at a High School in Zimbabwe

Edward S Nsingo¹ and Silvanos Chirume^{2,*} 

¹Mathematics Teacher, Luveve High School, Bulawayo, Zimbabwe

²Department of Mathematical Sciences, Zimbabwe Open University, Zimbabwe

DOI: <https://doi.org/10.23917/ijrime.15847>

Received: January 23th, 2026. Revised: January 28th, 2026. Accepted: April 29th, 2026

Available Online: April 29th, 2026. Published Regularly: June 2026

Abstract

This research involves time series analysis and forecasting of the Ordinary Level Mathematics results. From 2008 to 2009 many educators had left the country seeking higher salaries, leading to severe staff shortages in schools. This and other reasons were anticipated to impact negatively on “O” Level Mathematics results. Data for the years 2010 to 2022 were collected from one high school in Bulawayo. The data were meticulously analysed, with a particular emphasis on identifying the underlying trend through data smoothing techniques. Advanced computer software R studio was used to conduct regression analysis and generate trend lines and time series graphs. Additionally, an ARIMA (autoregressive integrated moving average) forecasting model was applied to forecast and predict the school’s future pass rate in Ordinary Level Mathematics. Interviews with the 10 qualified teachers at the school were carried out to identify potential causes for changes in the pass rate of Mathematics at the High School. By understanding how pass rates may evolve over time, stakeholders like parents, the school, and government authorities can make informed decisions and implement strategic measures to enhance academic performance in the long run. The study noted that the pass rate of the school was going to slightly improve, and the school was therefore recommended to deploy the hardworking teachers, who had a history of improving the pass rate to the examination writing classes so that the pass rate could increase uniformly and to give them more teaching and revision time and incentives. The study also recommends further research.

Keywords: Time Series Analysis, Forecasting, ZIMSEC results, Data smoothing, ARIMA and ARMA models

*Corresponding Author:

Edward S Nsingo, Zimbabwe Open University, Box 1810, Gweru, Zimbabwe

Email: skchirume@gmail.com

1. Introduction

The pass rate of a school is a crucial determinant for parents when selecting a school for their child’s education. The researchers were asked by parents and other stakeholders if the school under study has a good pass rate or not and if the current pass rate is more likely to improve or not. In order to respond well to these questions there was a need to include mathematical evidence. One of the subjects that parents will take seriously is Mathematics as most courses in the engineering field such as Electrical Engineering as well as Auto-motive Engineering will require learners to have passed Mathematics at Ordinary Level as one of the prerequisites. The National University of Science and Technology programs handout clearly shows that all the programs for the faculty

of science make it mandatory for all the students to have passed mathematics at the Ordinary level. This study involves a time series analysis and forecasting of Ordinary Level Mathematics results of one of the schools in Bulawayo using the school's results from 2010 to 2022.

1.1. Statement of the problem

The pass rate of Ordinary level Mathematics results at one High School in Bulawayo has been fluctuating unpredictably from 2010 to 2022, causing concern among parents and teachers. Despite the efforts to improve the educational outcomes, the pass rate has oscillated without clear trend, making it challenging to predict future performance. This uncertainty hinders informed decision-making and effective resource allocation.

1.2. Purpose of study

The purpose of this study is to analyse the trend and patterns in Ordinary level Mathematics from 2010 to 2022 using time series analysis as well as to develop an ARIMA forecasting model that can be used to accurately predict future mathematics results. The researchers also intended to identify factors that influence mathematics performance at Ordinary level.

1.3. Research objectives

The main objective of this study is to use time series analysis to forecast one High School's Mathematics Ordinary Level pass rate. Specifically, this study intends to:

- a. plot a time series graph and observe if the Ordinary level Mathematics results follow a certain pattern for the period 2010 to 2022
- b. determine the trend line or curve of the series using regression analysis and R Studio for the period 2010 to 2022
- c. solicit information which can help to explain the observed trends, if any, and
- d. use R Studio to determine an ARIMA model that forecasts the mathematics students' pass rate.

1.4. Research questions

This study will provide the answers to the following questions:

- a. What is the pattern of the Ordinary level Mathematics results for the period 2010-2022?
- b. How does the trend line or curve of the time series for the Ordinary Level Mathematics results for 2010-2022 look like?
- c. What is the Ordinary level Mathematics pass rate for the school under study likely to be in the future, for example in the 16th year, 2025?
- d. Is the interview data useful in interpreting the trend depicted by the pass rates?

2. Review of Related Literature

Time series analysis and forecasting play a pivotal role in various domains, enabling informed decision-making and accurate predictions at large. For example, time series analysis can be used in finance to predict stock prices and currency exchange rates (Xu et al., [2023](#); Wijesinghe, [2020](#); Mehtab & Sen, [2020](#)). In healthcare, it can be used to predict future disease outbreaks and to monitor the spread of diseases. In retail, it can be used to predict sales and inventory needs. In the education sector we can use time series analysis to analyse and predict the possible future performance of the learners (Mwale-Mkandawire, [2020](#); Takawira, [2016](#)). This might help to put mitigatory measures in case the performance is very poor.

Karapanagiotidis ([2012](#)) ascertains that modern time series forecasting methods are eventually rooted in the idea that the past tells us something about the future. Of course, the question of how exactly we are going about interpreting the information encoded in past events,

and furthermore, how we are to extrapolate future events based on this information, constitute the main subject matter of time series analysis.

Time series analysis and forecasting have a rich history that spans discipline and industries. The pursuit of accurate predictions from historical data has driven the development of a diverse range of methodologies, each with its strengths and limitations. Box and Jenkins (1970) are confident that time series analysis traces its roots to the early 20th century when econometricians like Norbert Wiener laid the groundwork for autoregressive models. The advent of computers in the mid-20th century facilitated the computation of complex statistical methods, leading to the development of models like ARIMA (autoregressive integrated moving average).

Wei-Yin (2021), points out that ARMA (autoregressive moving average) models can provide short-term predictions of a data series at a given point. However, he goes on to explain that ARIMA models are better suited for long-term prediction and can model linear and non-linear patterns of the data. So, ARMA models are better for short-term predictions while ARIMA models are flexible and can make longer-term forecasts of even more than 20 years according to Box & Jenkins (1970).

Liu et al, (2024) and Zhang (2018) emphasize that ARMA and ARIMA models are similar in that they are both used to model time-series data, but the main difference is that ARMA models can only model stationary data while ARIMA can model both stationary and non-stationary data. So, ARMA models are a sub-set of ARIMA models. Stationary data is a type of time-series data that has constant mean and constant variance. Non-stationary data has a meaning and a variance that changes over time.

Lundeberg and Lee (2017) are of the opinion that the recent years have seen a growing emphasis on the interpretability and explainability of time series forecasting models. Techniques such as SHAP (Shapley Additive exPlanations) have been adapted to provide insights into future contributions and model behaviour, enhancing transparency and fostering trust in critical decision scenarios.

Malik (2023) in his write up on “An Analysis of Time Series Analysis and Forecasting Techniques” provided an analysis of time series and forecasting techniques. He also delved into the principles underlying these methods, their practical applications and their performance across a range of scenarios. His research work was wonderfully done as it presented real-world applications that showcase the tangible impact of time series analysis and forecasting across different domains. However, Malik’s work did not include a thorough review of related literature and did not use mathematical formulae as well as computer applications to enhance further comprehension of the time series and forecasting techniques.

Time series analysis and forecasting was superb with a variety of examples working using many formulae to clearly explain the time series and forecasting techniques. He used so many computer-generated time series graphs to clearly explain the time series techniques. His research work consisted of surveying literature as well as developing a tool for time series data analysis and forecasting. His main research question was ‘which forecasting algorithm is the best for this type of data set/series?’ The main purpose of Fagerholm’s research was to have a good comparison of different forecasting algorithms to find out which of the algorithms had the best prediction accuracy for the investigated data set. His steps for the procedure were clearly stated with logical sequencing. In conclusion he found out that the LSTM (Long Short-Term Memory) model had the best prediction accuracy for his data sets and tested models.

2.1. Conceptual framework/models

Box and Jenkins (1970) present the following ideas about the following concepts: Autoregressive (AR) theory: the authors describe the AR model as a way to forecast future values in a time series based on past values. Moving Average (MA) theory is a model used to forecast future values in a time series based on errors (residuals) from past predictions.

Autoregressive Integrated Moving Average (ARIMA) theory. The authors introduce the ARIMA model, which combines the AR and MA models to handle non-stationary data. These concepts or models are fundamental to time series and forecasting hence this study will incorporate them. The study will make use of the ARIMA model in forecasting the future pass rate of ordinary level Mathematics using historical data of one High School obtained from the results analysis file for the years 2010 to 2020.

2.2. Theoretical framework

The study is also premised on the theory of educational assessment which measures student achievement and informs instructional practices and policy decisions. According to Brown (2022) there is need to integrate the psychological-educational assessment theory with statistical innovations so as to come up with assessments reflecting the real-world learning processes, and good classroom practices that promote student achievement. Also, student achievement is influenced by many factors hence the need, according to Suleimann et al., (2024) to develop a predictive model using statistical methods like regression or time series to identify key determinants of academic performance and achieving high accuracy in forecasting outcomes. This will also help to address disparities and promote equity especially in resource-constrained schools such as those in Zimbabwe.

3. Method

3.1. Research design

Time Series Analysis and Forecasting techniques use historical data to predict the future behaviour of a variable, and in this study the dependent variable was the pass rate of the Ordinary Level Mathematics learners whilst time was the independent variable. The research design for this study is based on time series analysis and forecasting techniques, specifically the use of the ARIMA model and time series analysis trend line finding. The methodology involved collecting data on the pass rate of students who took the Ordinary Level Mathematics examination at a high school in Bulawayo, Zimbabwe. The country's examination board is responsible for setting standardized examinations as well as issuing results. The passing grades were A, B and C, but now D and E are not so good, the majority are now passing grades whereas students scoring F and U will have failed. The data was analysed using the R statistical software package, with the goal of estimating the future pass rate using the ARIMA model that can be used to forecast future pass rates. The researchers collected data from one of the high schools in Bulawayo from the school's results analysis file for the years 2010 to 2022.

ARIMA model was used instead of ARMA because of the following general reasons: Firstly, ARIMA models are more flexible than ARMA models and can handle data that is non-stationary or has a trend component. Secondly, ARIMA models are easier to interpret and understand, as the meaning of the parameters is clearer than in ARMA. Furthermore, ARIMA models can be used to make forecasts over a longer horizon than ARMA models. Finally, ARIMA models are more robust to outliers and extreme values in the time series data. The

specific reasons for using ARIMA instead of ARMA for this study are given below in the section on data presentation, analysis and interpretation.

3.2. Data Collection and Analysis Procedure

In this study primary data was collected by means of face-to-face interviews with 10 purposively selected teachers at the school. The interview provided qualitative data that was later used to explain or provide context for trends or patterns in the data. The interview also helped in identifying potential causes or explanations for changes in the pass rate of Mathematics at the High School overtime.

Secondary data used in this research work was obtained from the school's results analysis file for the years 2010 to 2022. This is the data that was used for determination of the trend line and for the forecasting purposes. It was not very costly to obtain the secondary data, and this data provided a major framework for this study. Accurate forecasting would assist in predicting future outcomes of mathematics learning.

The researchers were involved in the following steps:

- a. Secondary data on the results of Ordinary Mathematics learners were collected from the school records from 2010 to 2022. The school's results analysis became very useful on this cause. This fact and adds that this is a very cost-effective way of collecting data as data is readily available in the school's files.
- b. Cleaning of the data, this step involved removing errors and inconsistencies. This step was done by identifying potential errors such as missing values, outliers or spikes.
- c. After identifying these errors, the researchers had to delete them. This step was very crucial because it ensured that data was accurate and ready for analysis.
- d. After cleaning the data, the researchers were involved in the process of normalizing the data set by transforming it so that it had a mean of zero and a standard deviation of one. This procedure was done by centering the data by subtracting the mean value from each data point and this created a new time series with mean zero. The next thing was to scale the data by dividing each data point by the standard deviation and this created a new time series with standard deviation of one. After that the researchers then applied normal distribution to the scaled data creating a normalized time series. This ensured that all the data points were within a similar range, and this made it easier to compare them. Scaling and normalizing historical data were done using R Studio statistical software.
- e. The next step after normalizing the data was to identify the trend line using regression analysis procedures and the R studio software. After the trend line was identified the researchers then managed to forecast the future values based on the trend.

The following are the steps which are involved in the ARIMA modelling process, although in this study the majority of the work was done using the R Studio statistical software:

- a. Model identification: The first step is to identify the order of the ARIMA model based on the autocorrelation and partial autocorrelation functions (Box & Jenkins, [1976](#)).
- b. Model estimation: Once the model order is identified, the model parameters are estimated using methods such as maximum likelihood or least squares (Chatfield, [2018](#)).
- c. Model validation: The model is validated by checking its goodness of fit and forecasting accuracy using various statistical measures such as Akaike Information Criterion (AIC) and the Theil's inequality coefficient (Chatfield, [2018](#)).
- d. Model Modification: If the model is not performing well, the model order or the estimation method can be adjusted until an optimal model is obtained (Chatfield, [2018](#)).

- e. Forecasting: Once the model is validated, it can be used to make forecasts about the future values of the time series (Chatfield, [2018](#)).
- f. Forecast evaluation: The accuracy of the forecasts can be assessed using methods such as mean absolute error (MAE) or mean squared error (MSE) (Hyndman & Athanasopoulos, [2018](#)).

3.3. Research Instrument

The researchers used face-to-face interviews to collect primary data. In this study total sampling technique was used because the number of Mathematics teachers at the school was ten. This is about 12.5% of the members of staff at this school. The reason why the teachers who do not teach Mathematics were excluded was because of the complexity of the research topic as it includes the technical jargon that could be foreign to them. The interview questions were designed by the researchers, and the results were separately recorded for each interviewee. The questions asked whether time series analysis was useful as a tool to analyze trends and to suggest any factors that might have contributed to the pass rates as shown by the trends.

3.4. Population, Sampling Technique and Sample Size

The High School had 1 742 learners and 78 qualified staff members. It is a non-selective school in one of the high-density suburbs in Bulawayo, Zimbabwe. In this study, total sampling technique was used in carrying out face-to-face interviews. This is a type of purposive sampling technique that involves examining the entire population. The main reason why the researchers used this technique is that the High School has only 10 qualified teachers with 5 of them holding a Bachelor of Science Honours Degree in Mathematics, 2 of them having Bachelor of Science Honours Degree in Mathematics and Statistics and 3 having Diploma in Education majoring in Applied and Pure Mathematics. However, there could be social desirability bias as teachers may have provided responses they believed were expected or desirable, rather than their genuine experiences. The researchers included all ten mathematics teachers in the school so as to solicit as much information as possible.

Another major limitation was the limited sample size of one school. However, the results and findings could be taken as they were since this was a single case study and triangulation of quantitative and qualitative data was used.

4. Data Presentation, Analysis and Interpretation

4.1. Secondary data: Mathematics 'O' Level results

This study used the R Studio computer statistical package to determine the trend line and to include computer generated time series graphs. The researchers also carried out and used ARIMA forecasting model to forecast the future performance of the learners in Ordinary level Mathematics. The Secondary Data is presented in [Table 1](#) below.

Table 1: Showing the ZIMSEC Ordinary Level pass rates of the learners in Mathematics for the years from 2010 up to 2022.

Year	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022
Pass rate %	10.5	8.3	11.7	19.0	13.1	14.1	15.8	21.3	17.4	12.0	14.1	18.7	9.8

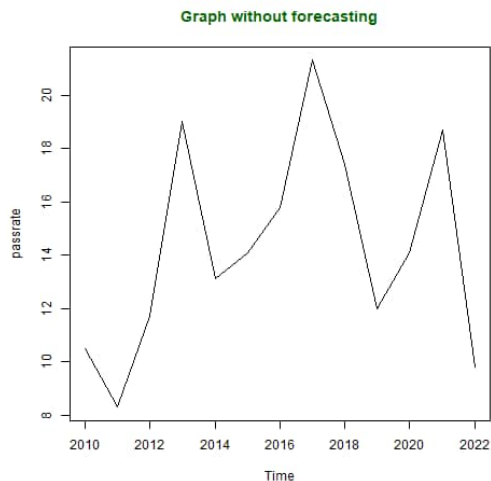


Figure 1 Time series graph showing fluctuations and general trend of Ordinary Level Mathematics pass rates from 2010 to 2022. The pattern indicates that decreases are often followed by improvements over the next two years.

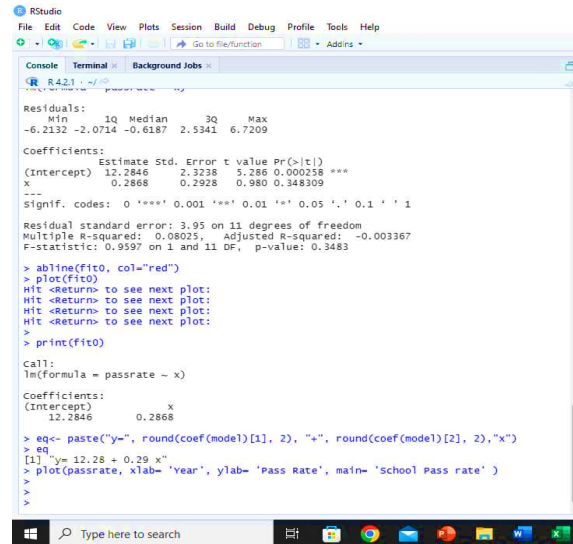


Figure 2: Output from R Studio showing the fitted linear regression line for the school's O-Level Mathematics pass rates (2010–2022). The model $y = 12.28 + 0.29x$ indicates a slow but positive trend in student performance.

Short alt-text:

Line graph showing the percentage pass rates in O-Level Mathematics from 2010 to 2022, with fluctuations and a gradual upward trend overall. From [Figure 1](#) and [Table 1](#) above, it is observed that when the pass rate decreased for one year, it increased thereafter for the following two years, and that pattern seemed to be followed over time. So, when the pass rate dropped for one year there was a high chance that it would increase for the next two years. This was observed for the years 2011 to 2012, 2014 to 2015 and 2019 to 2020.

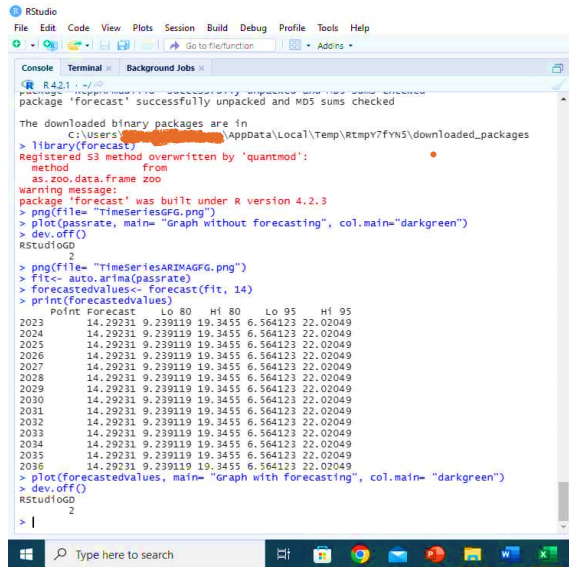
In those years the pass rate increased whilst from 2012 to 2013, 2016 to 2018 and 2021 to 20 the pass rate decreased from the respective previous years. By carefully studying the trend one can predict that the pass rate for 2025 and 2026 will be most likely higher than that of 20.

[Figure 2](#) below shows the results obtained from the R Studio computer package in calculation of the trend line for the data in [Table 1](#).

Short alt-text:

Screenshot from R Studio showing a fitted linear regression trend line for Mathematics pass rate data (2010–2022).

[Figure 2](#) shows that the trend line was linear although the pass rate does not always follow a linear trend. A linear trend is most applicable if the R square value is close to 1 (Choosing the right Trend line type for your data; n.d.). For this study the trend line calculated using the



```

R 4.2.1 ~ /R
File Edit Code View Plots Session Build Debug Profile Tools Help
package 'forecast' successfully unpacked and MD5 sums checked
The downloaded binary packages are in
  c:\Users\user\AppData\Local\Temp\rtmpv7fYN5\downloaded_packages
> library(forecast)
Registered 53 method overwritten by 'quantmod':
  method from
  as.zoo.data.frame zoo
Warning message:
package 'forecast' was built under R version 4.2.3
> png(file="TimeseriesGG.png")
> plot(passrate, main="Graph without forecasting", col.main="darkgreen")
> dev.off()
RStudioGD
2
> png(file="TimeseriesARMA1GG.png")
> fit<- auto.arima(passrate)
> forecastedvalues<- forecast(fit, 14)
> print(forecastedvalues)
  point Forecast  Lo 80  HI 80  Lo 95  HI 95
2023  14.29231  9.239119  19.3455  6.564123  22.02049
2024  14.29231  9.239119  19.3455  6.564123  22.02049
2025  14.29231  9.239119  19.3455  6.564123  22.02049
2026  14.29231  9.239119  19.3455  6.564123  22.02049
2027  14.29231  9.239119  19.3455  6.564123  22.02049
2028  14.29231  9.239119  19.3455  6.564123  22.02049
2029  14.29231  9.239119  19.3455  6.564123  22.02049
2030  14.29231  9.239119  19.3455  6.564123  22.02049
2031  14.29231  9.239119  19.3455  6.564123  22.02049
2032  14.29231  9.239119  19.3455  6.564123  22.02049
2033  14.29231  9.239119  19.3455  6.564123  22.02049
2034  14.29231  9.239119  19.3455  6.564123  22.02049
2035  14.29231  9.239119  19.3455  6.564123  22.02049
2036  14.29231  9.239119  19.3455  6.564123  22.02049
> plot(forecastedvalues, main="Graph with Forecasting", col.main="darkgreen")
> dev.off()
RStudioGD
2
> |

```

Figure 3: ARIMA (1,1,1) forecasting results generated in R Studio, showing projected increases in future pass rates and the 95%

R Studio statistical software was: $y = 12.28 + 0.29x$ where y represents the pass rate and x represents the time in years and in this study the time spanned from 2010 to 20. For example, we can attempt to estimate the performance of the learners in the 16th year, 2025.

So, we substitute $x = 16$ into the trend as follows: $y = 12.28 + 0.29(16)$

$$y = 16.92$$

$$y \approx 16.9$$

Now in the 16th year the ZIMSEC Ordinary Level pass rate for the High School would be approximately 16.9%. This result concurs with the nature of the time series graph above which shows a pattern in the pass rate that when the pass rate drops in one year, it picks up on the following year. For the 13th year, 20, the actual pass rate was 9.8% and for the 16th year, 2025 the pass rate would most likely improve to $\approx 16.9\%$.

[Figure 3](#) below shows the results obtained on ARIMA forecasting for Table 1 using R Studio.

Short alt-text:

Graphical output from R Studio displaying ARIMA (1,1,1) forecast of O-Level Mathematics pass rates with confidence intervals.

[Figure 4](#) below shows a time series graph with forecasted values which has been smoothed using a 3-point moving average.

Short alt-text:

Dual-line graph showing the original and 3-point moving average smoothed time series of Mathematics pass rates from 2010 to 2022.

The [Figure 4](#) graph was produced using R Studio software. The 3-point moving average preserves more data compared to the moving averages above it. The graph shown in red is the smoothed graph of the time series and shows an overall general increase in the forecasted values for the pass rates.

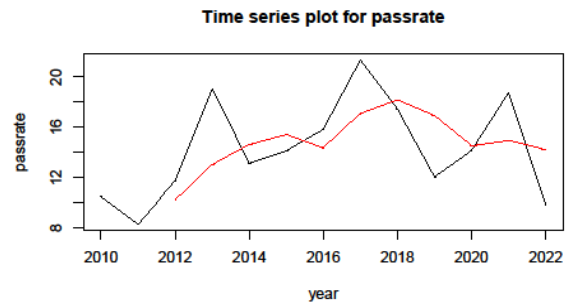


Figure 4: Comparison between the original O-Level Mathematics pass rate time series and the smoothed series using a 3-point moving average. The smoothed line (red) highlights the overall upward trend in forecasted values

Table 2. Calculated the partial sums

Pass Rate	Deviation from the Mean	Partial Sum (S)
10.5	-4.1583	-4.1583
8.3	-6.3583	-10.516
11.7	-2.9583	-13.474
19.0	4.3417	-9.1332
13.1	-1.5583	-10.691
14.1	-0.5583	-11.249
15.8	1.1417	-10.108
21.3	6.6417	-3.4664
17.4	-2.6583	-0.7247
12.0	-2.6583	-3.383
14.1	-0.5583	-3.9413
18.7	4.0417	0.1004
18.7	4.0417	4.1421

It was vital to use ARIMA instead of ARMA for this research's collected data on pass rates because of the following reasons:

- Non-stationarity: The pass rate data exhibits non stationarity, meaning the mean and the variance change over time. ARIMA accounts for this by including a differencing step (the "I" in ARIMA) to make the data stationary.
- Trends and seasonality: The data shows apparent trends (for example, increasing pass rates from 2010 to 2017) and seasonality (for example, fluctuations between years). ARIMA can capture these patterns, while ARMA might not be sufficient.
- Integration: The pass rate data likely has a strong memory, meaning past values influence future values. ARIMA's integration component addresses this, ensuring that the model accounts for the data's long-term dependencies.

Using ARMA alone might not effectively capture the above-mentioned features, leading to less accurate forecasts and model inadequacy. ARIMA, on the other hand, provides a more comprehensive framework for modelling this data.

4.2. Stationarity Test Results

The Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test was used to test for the stationarity of the data before the ARIMA could be found.

Null Hypothesis(H_0): The pass rate is stationary

Alternative hypothesis(H_1): The pass rate is non-stationary.

First, we calculated the meaning of the data as follows:

Mean of the pass rate series was $\frac{10.5+8.3+11.7+19.0+13.1+14.1+15.8+21.3+17.4+12.0+14.1+18.7}{12} = 14.6583$

View in [Table 2](#), we then calculated the partial sums as follows:

Then we proceeded to calculate the KPSS statistics as follows:

KPSS $= \frac{\sum_{t=1}^T S^2}{T^2 \cdot S^2}$ where s is the partial sum, T is the number of observations and S^2 is the long-run variance.

Now $S^2 = 16.438$, $\sum_{t=1}^T S^2 \approx 792.663$, $T=12$

$KPSS = \frac{792.663}{12^2 \cdot 16.438} \approx 0.335$ which is the calculated value.

The critical value for KPSS stationarity is 0.463 at 5% significance level. The conclusion is that since $0.335 < 0.463$, we fail to reject the null hypothesis that the series is stationary. Hence the pass rate might be stationary meaning that they might be stable around a constant level or deterministic trend.

4.3. Determination of ARIMA parameters

Since the pass rate is stationary, the researchers went on to determine the ARIMA parameters. After running the auto-arima code, the following was detected: $d=1$ because the series shows a trend and becomes stationary after first difference. Best $(p; q)$ were $(1; 1)$ based on typical ACF patterns. The best suited model was ARIMA $(1; 1; 1)$ with $p=1$ implying one autoregressive term (AR), $d=1$ implying first differencing applied to make data stationary and $q=1$ implying one moving average term (MA).

Interpretation of the ARIMA $(1; 1; 1)$ parameters: AR $(1) = 0.45$ (Autoregressive term). This means that the current change in pass rate is partially influenced (45%) by the previous year's change. Thus, if the pass rate increased by 2% last year, this year it might increase by about 0.9% (45% of 2%), assuming other effects are neutral. The positive value momentum is noted, meaning that the increase tends to continue. MA $(1) = -0.72$ (Moving average term). This reflects how much the model corrects for the previous year's forecast error. A value of -0.72 means the model tends to counteract previous mistakes strongly. Thus, if last year's prediction was 1% too high, this year's prediction adjusts downwards by $\sim 0.72\%$.

Differencing, $d = 1$. The model doesn't model pass rate directly but models their year-to-year change. This is needed because the original data showed a trend and not stationarity when the researchers used the ADF test. However, the KPSS test showed different results of stationarity, hence there could be need to cross check and clean the original data or to run further tests.

Intercept = 0.85 (Drift term). After differencing, the average yearly change in pass rate is $+0.85\%$. This means the pass rate has a mild upward trend over time. This observation agrees with the R studio ARIMA result.

4.4. Primary data: Interview with teachers

From [Table 3](#) above the following major ideas (themes) emerged:

- a. A lot of factors affect pass rates and time series analysis may not capture all
- b. Results of the time series analysis may help teachers change their teaching strategies
- c. Factors contributing to fluctuations in the pass rate include having well-scheduled revision lessons for the learners
- d. Teaching for free, not charging the learners money for the lessons may be beneficial
- e. Having more time with the learners may be beneficial
- f. Standardized tests may be better to use to do comparison of pass rates

Loftin et al., (2020) conducted a study on pass rates as a measure of program quality in the nursing field. Factors affecting pass rates were identified and some of them are in tandem with those of this study.

Table 3 below shows the presentation of data collected using the interview protocol.

THEME	QUOTE	ANALYSIS
1. Utility of time series analysis in Mathematics education.	“Time series analysis is a very useful tool that can help in identifying trends followed by the data and one can have a clear picture of how the pass rate, as a variable, is most likely to behave in the future” (6 teachers) .	Most teachers believe time series analysis is useful for identifying trends and predicting future pass rates, indicating its potential value in Mathematics education.
2. Impact of technology on data analysis.	“Time series has helped so much in analyzing data and one major advantage is that technology can analyze data with great speed and accuracy” (6 teachers) .	The majority of teachers think technology has facilitated data analysis due to its speed and accuracy, highlighting the importance of technology in educational settings.
3. Knowledge and application of time series analysis.	“I always use it to evaluate how my learners are performing and this helps me to change my teaching strategies accordingly” (6 teachers) .	Most teachers use time series analysis to evaluate learner performance and adjust their teaching strategies, showing its practical application in education.
4. Importance of pass rate trends.	“Knowledge of these trends is very important to all the stakeholders because an investigation into the causes of the trends can help the school improve its pass rate” (4 teachers) .	Some teachers recognize the importance of understanding pass rate trends for school improvement, while others are less concerned, indicating varying perspectives on its significance.
5. Factors contributing to pass rate trends.	“There are so many factors that contribute to the fluctuations in the pass rate, one of the factors is that there are some teachers like Mister A and Madam B who have a history and a record of improving the pass rate whenever they take examination classes” (6 teachers) .	Teachers identify specific factors, including teacher performance and teaching strategies, that contribute to pass rate trends, suggesting a nuanced understanding of the issue.
6. Strengths and weaknesses of time series analysis.	“Strengths are that it is easy to calculate the moving averages and to get the trend line. It is also easy to understand.” Then the drawback could be that it does not... if we try to smoothen the data using a larger n-point moving average we are most likely to lose a lot of data.” (90 % of teachers) .	Teachers recognize both the ease of use and potential drawbacks of time series analysis, such as data loss, indicating a balanced view of its utility.
7. Use of standardized tests.	“Standardized tests are good to use as they give a fair analysis of the results and comparison will not be a problem when tests are standardized” (all teachers) .	All teachers agree on the importance of using standardized tests for meaningful trend analysis, highlighting the need for consistency in assessment.

THEME	QUOTE	ANALYSIS
8. Assessing teachers' and schools' performance using time series	“Time series should be used to assess or evaluate the performance of teachers and schools so that corrective action can be taken. However, not all teachers are keen on using it, so there could need to train them “(7 teachers).	Most teachers believe time series can be used to assess the performance of teachers and schools through the students' pass rates though one teacher cautions against it due to the complexity of factors affecting pass rates. There could be need to further train Mathematics teachers on how to use time series analysis.

5. Results and Discussion

The researchers managed to identify a certain pattern from the time series graphs. It seemed that the pass rate would increase and decrease from year to year and smoothing became useful in identifying the trend. The 3-point moving average was used to smoothen the time series. From the time series graph, one can note that whenever a pass rate drops from a certain percentage, it would almost remain shortly like that but in the following year it would improve.

Secondary data was also used to get the ARIMA forecast using the R Studio. The ARIMA forecast revealed that the pass rate for the learners in their Ordinary Level Mathematics examination would improve on the 16th year, 2025, since it had dropped in the 13th year, 2022. This observation agrees with the observation on the first time series graph which was produced using the R Studio package. The researchers then managed to produce another time series graph with forecasted values for the next 13 years. This, again, was one of the objectives of this study and it was successfully achieved.

The findings of the research using ARIMA model showed that there was going to be a slight increase in the pass rate of the Ordinary level learners at the High school under study. The results are consistent with those of Mwale-Mkandawire (2020) and Takawira (2016), who found that ARIMA models can accurately forecast examination results in Zimbabwe.

For interview data, the first question was, “Do you think time series analysis is a useful tool for improving mathematics education?” Most responses were positive. By studying the trends, it is possible to take corrective action if the trends are not pleasing hence time series becomes a very useful tool even in decision-making. The answer to the second question, “Have you seen or observed any trends in the pass rate over time?” corroborated Question 1.

The third question was very useful as it helped in explaining the decreases and increases in pass rate after a certain, almost, fixed time. The question was: “Are there any factors that you can give which you think might have contributed to the trends if there are any observed trends?” The reasons included not having enough time and effort to teach and revise with the learners. Hanushek et al., (2011) support this observation as they noted that the evidence suggests that teacher effort is an important determinant of student achievement.

The face-to-face interviews unearthed a lot of observations which helped in explaining the pattern or trend in the time series analysis of results (Crismono et al., 2026). For instance,

a teacher would take a class from Form One right up to Form Four, after that the teacher returns again to take the new Form One's whilst another teacher will be taking Form Four's. It appeared that when certain teachers had examination classes the pass rate would improve and when the same teachers go back to teach Form One's, the teachers remaining to take examination classes would lead to a drop in the pass rate (Mangwende, [2026](#)). That became the major contributor to such a pattern in the pass rate.

About 50% of the interviewees seemed not to worry about the trends in the pass rate. These teachers clearly outlined that they were not aware that the pass rate followed a certain pattern. The other 50% of teachers who were conscious of the trends or patterns in the pass rate argued that examinations are different for each year and that the learners are also different every year hence there is a possibility of such a pattern whereby the pass rate can increase and decrease interchangeably. It seemed that no teacher wanted to be identified with a low pass rate. This observation is also supported by Ladd ([2011](#)) who believed teachers who are less concerned about pass rate trends tend to have lower-performing students.

Another observation that the researchers noted from this study was that the trend line is the easiest way of estimating the future pass rate of the learners and it does not have a lot of complicated calculations, and one can easily get a solution to their regression problem. Hyndman and Athanasopoulos ([2018](#)) ascertain that a trend line is a simple way to forecast future values. However, Taylor ([2013](#)) is of the opinion that more advanced statistical models, such as ARIMA and exponential smoothing, can outperform simple trend lines in forecasting accuracy.

6. Conclusion

This research work focused on the time series analysis and ARIMA forecasting of the Ordinary level results of one High School in Bulawayo Zimbabwe and this was successfully achieved. Smoothing was also done to help the researchers better see the pattern followed by the pass rates. The other key objective was to reduce the trend line, and this was successfully done by using regression analysis and by using the R Studio computer statistical package.

The ARIMA model has shown that the pass rate of the learners in Mathematics at Ordinary level is most likely going to be around 14% to 22%. This information will be very useful to the stakeholders such as parents, the school, the government and policymakers. Harrell (2013) supports the opinion that ARIMA models can be successfully applied in education to forecast students' performance and pass rates. Kabir (2016), used a combination of ARIMA and machine learning models to predict student performance and pass rates, outperforming traditional regression models.

However, it is important to note that ARIMA models rely on the quality of the data and may not capture external factors that can impact pass rates, such as changes in curriculum or teaching methods. Choi ([2017](#)) supports this observation. Overall, objectives for this study were achieved and the research questions were successfully answered.

Recommendations

The researchers came up with the following recommendations:

- a. The school, policy makers and Education Ministry should be prepared for fluctuating pass rates and adjust their strategies accordingly. Further research could investigate the

- underlying factors contributing to these fluctuations, such as change in curriculum, teaching methods, or student demographics.
- b. The school should inform teachers about the need to study and monitor the pass rate patterns.
 - c. The researchers recommend that the Head of the Mathematics Department of the school should give the examination classes to those teachers who have a history of improving the pass rates whilst they encourage the other teachers to perform well in whichever classes they are deployed.
 - d. The school should come up with a way of motivating the hardworking teachers by some form of performance-based rewards which can be financial or non-financial.
 - e. By acknowledging the potential sampling biases and taking steps to mitigate them, the study can provide a more comprehensive understanding of teachers' perspectives on time series analysis in mathematics education.
 - f. There is need for further research on the use of other times series analysis methods to forecast and predict pass rates and on factors that may contribute to such rates.
 - g. There is a need to also spread the research to the districts and provinces as well so that a meaningful contribution can be made to the education system of the country.

Acknowledgments

Authors acknowledge the individual responses and voluntary consent by the participants.

Declarations

- Author Contribution : Author 1: Conceptualization, Writing - Original Draft, Editing, Analysis and Visualization; Author 2: Writing - Review & Editing, Formal analysis, and Methodology
- Funding Statement : This research was funded by the authors themselves.
- Conflict of Interest : The authors declare no conflict of interest.
- Additional Information : Additional information is available for this paper.
- AI Declaration Statement : R studio was used to do data analysis and to produce graphs. No other AI tools were used.

7. References

- Box, G.E.P. & Jenkins, G.M. (1976). *Time Series Analysis: Forecasting and Control*. Cambridge University Press, Oakland.
- Box, G.E.P. & Jenkins, G.M. (1970). *Time Series Analysis Forecasting and Control*. Cambridge University Press, Oakland.
- Brown, G.T.L. (2022). The past the present and future of educational assessment: A transdisciplinary perspective. *Frontiers in Education*, 7, Article 1060633. <https://doi.org/10.3389/feduc.2022.1060633>
- Chatfield, C. (2018). Introduction to Multivariate Analysis. Abindon-on-Thames. Routledge. Choosing the right Trendline type for your data (n.d.) Retrieved on 02/04/25 from https://www.tcworkshop.com/data/Downloads/Handouts/Microsoft_Excel_ChartTrendlines.pdf
- Choi, T. H. (2017). *Hidden transcripts of teacher resistance: A case of South Korea*. Teaching and Teacher Education, 63, 245-255
- Crismono, P., Saputra, M. A. A., & Hudi, S. (2026). Evolution of deep learning and its reflection on statistical mathematics learning. *International Journal of Review in*

- Mathematics Education, Volume(1) Issue(1), 50–69.*
<https://doi.org/10.23917/ijrime.14812>
- Hanushek, E., Machin, S. & Woessman, L.(2011). *Handbook of the Economics Education* (Vol.4). Elsevier.
- Harrell, F. E. (2013). *Regression modelling strategies: With applications to linear models, logistic regression, and survival analysis*. Springer. New York.
- Hyndman, R.J. & Athanasopoulos, G. (2018). *Forecasting Principles and Practice*. Springer, Berlin.
- Kabir, S.M.S.(2016). Methods of data collection. In *Basic guidelines for research: An introductory approach for all disciplines* (pp.201-275). Book Zone Publication.
- Karapanagiotidis, P. (2012). Literature review of modern time series forecasting methods, Retrieved on /03/25 from [S2012 Literature Review - linear survey.July31.pdf](https://www.researchgate.net/publication/260121212_Literature_Review_-_linear_survey_July31.pdf)
- Ladd, H., F.(2011).Teacher Motivation and student Achievement. *Journal of Human Resources*, 46(2), 361-384.
- Loftin, C., Reyes, H., Hartin, V. & Rice, L. (2020). A closer look at first-time pass rates as the primary measure of program quality, *Journal of Professional Nursing* 36, pp.707-711 <https://doi.org/10.1016/j.profnurs.2020.09.011>
- Liu, Z.; Cui, Y.; Ding, C.; Gan, Y.; Luo, J.; Luo, X.; Wang, Y. (2024). The Characteristics of ARMA (ARIMA) Model and Some Key Points to Be Noted in Application: A Case Study of Changtan Reservoir, Zhejiang Province, China. *Sustainability* 16(18), 7955; <https://doi.org/10.3390/su16187955>
- Lundeborg, S. & Lee, S. (2017). A Unified Approach to interpreting Models Predictions. <https://doi.org/10.48550/arXiv.1705.07874>.
- Malik, A. (2023). *The future of technology*. Elsevier Ltd. Amsterdam.
- Mangwende, E. (2026). Social Constructivism: Principles and Implications to Mathematics Learning. *International Journal of Review in Mathematics Education*, 70-79. <https://doi.org/10.23917/ijrime.15221>
- Mehtab, S. & Sen, J. (2020), A Time Series Analysis-Based Stock Price Prediction Framework Using Machine Learning and Deep Learning Models, Technical Report *School of Computing and Analytics, NSHM Knowledge Campus, India, pp. 1-46*
- Mwale-Mkandawire M. (2020). *Current Trends in Gender Issues in Education*. UNZA Press, Lusaka, Retrieved on 06/04/25 from https://www.researchgate.net/publication/349462853_Mwale-Mkandawire_2020
- Suleiman, I.B., Okunade, O.A., Dada, E.G., & Ezeanya, U.C. (2024). Key factors influencing students' academic performance. *Journal of Electrical Systems and Information Technology*, 11, Article 41. <https://doi.org/10.1186/s43067-024-00166-w>
- Takawira, M. (2016). Assessment of the Economic Impact. <https://uz.ac.zw/index.php>.
- Taylor, J.W. (2013). Exponential smoothing and ARIMA processes. *International Journal of Forecasting*, 29(2), 264-276
- Wei-Yin, L. (2021). *Guide Classification and Regression Trees and Forests*. <https://pages.stat.wisc.edu/~loh>.
- Wijesinghe, S. (2020), Time series forecasting: Analysis of LSTM Neural Networks to Predict Exchange Rates of Currencies. *Instrumentation* 7(4), pp. 25-39
- Xu, C.; Li, J.; Feng, B.; Lu, B. (2023), A Financial Time-Series Prediction Model Based on Multiplex Attention and Linear Transformer Structure. *Appl. Sci.* **2023**, *13*, 5175. <https://doi.org/10.3390/app13085175>
- Zhang, M. (2018). Time Series: Autoregressive models AR, MA, ARMA, ARIMA. Retrieved on 4/04/25 from <https://people.cs.pitt.edu/~milos/courses/cs3750/lectures/class16.pdf>