International Journal of Economics, Commerce and Management
United Kingdom ISSN 2348 0386 Vol. 14, Issue 10, Oct 2026

https://ijecm.co.uk/

DATA DRIVEN ANALYTICS FOR MAINTENANCE STRATEGY OPTIMIZATION IN POWER GENERATION: A CASE STUDY OF KENGEN PLC, KENYA

Joseph Guchu Ndungu
Master of Business Administration (Management Information Science)
School of Business and Management Science,
University of Nairobi, Kenya
guchunjoseph@gmail.com

Anthony Wainaina Ndungu
Senior Lecturer, School of Business and Management Science,
University of Nairobi, Kenya
awndungu@uonbi.ac.ke

Roba Adan Abdi Boyante
Jomo Kenyatta University of Agriculture and Technology, Kenya
aroba99@gmail.com

Abstract

This study examined how data driven analytics can be leveraged within Plant Information Systems to optimize maintenance strategies at Kenya Electricity Generating Company PLC (KenGen). The study adopted a descriptive and explanatory research design and utilized both primary and secondary data. Primary data was collected through a structured questionnaire administered to employees involved in plant operations, maintenance management, and Plant Information Systems, while secondary data was obtained from SAP Plant Maintenance records and organizational reports. A total of 162 questionnaires were completed and returned, representing an 81% response rate. The study examined the influence of Data Quality and Availability, Analytics Tools and Technologies, Staff Analytical Competency, and System

Licensed under Creative Common

Integration on Optimized Maintenance Strategies. Multiple regression analysis showed that the model was statistically significant (F = 19.315, p < 0.001) and explained 33.0% of the variation in optimized maintenance strategies. Staff Analytical Competency was the strongest statistically significant predictor (B = 0.736, β = 0.392, p < 0.001), followed by Analytics Tools and Technologies (B = 0.260, β = 0.211, p = 0.012). Data Quality and Availability and System Integration had positive but statistically insignificant effects when all variables were considered simultaneously. The findings indicate that effective maintenance optimization depends not only on data availability and integrated systems, but also on staff analytical capability and effective utilization of analytics technologies.

Keywords: Data driven analytics; Maintenance optimization; Plant Information Systems; Predictive maintenance; Analytical competency; Power generation

INTRODUCTION

Maintenance management is fundamental to asset intensive industries such as power generation, where equipment reliability directly influences operational continuity, safety, service delivery, and organizational performance (Alvarez Alvarado et al., 2022). Advances in digital technologies have transformed maintenance from a reactive function into a strategic process that improves asset performance and operational efficiency.

Data driven analytics has become a key enabler of modern maintenance by converting operational data into actionable insights for decision making. Through statistical analysis, machine learning, and predictive models, organizations can detect equipment deterioration, anticipate failures, and schedule maintenance proactively, resulting in improved reliability, reduced downtime, extended asset life, and lower maintenance costs (Carvalho et al., 2019; Papacharalampopoulos et al., 2022).

These capabilities are particularly important in the energy sector, where electricity generation depends on highly capital intensive assets operating under demanding conditions. Equipment failures may lead to production losses, increased operating costs, environmental risks, and interruptions to electricity supply. Consequently, electricity utilities increasingly integrate Plant Information Systems (PIS), Supervisory Control and Data Acquisition (SCADA) systems, industrial sensors, and Enterprise Resource Planning (ERP) platforms to support continuous equipment monitoring and evidence based maintenance decisions.

Data driven analytics refers to the systematic collection, processing, analysis, and interpretation of data to support informed decision making. In maintenance management, it

Licensed under Creative Common

transforms operational data generated by Plant Information Systems into actionable insights for maintenance planning and asset management. Continuous equipment monitoring enables early fault detection, while machine learning improves predictive maintenance by identifying failure patterns that conventional approaches may overlook (Alvarez Alvarado et al., 2022; Carvalho et al., 2019).

Despite growing adoption in developed economies, implementation in many developing countries remains constrained by poor data quality, fragmented information systems, limited analytical skills, and inadequate system integration. These challenges underscore the need to examine how organizations can better leverage operational data to improve maintenance performance within the energy sector.

Maintenance strategies comprise the planned approaches adopted to ensure that physical assets continue to perform their intended functions efficiently and reliably throughout their operational life. These strategies include reactive, preventive, predictive, and condition based maintenance, each differing in planning, execution, and dependence on operational data (Gholipour et al., 2025). While reactive maintenance addresses failures after they occur and preventive maintenance follows predetermined schedules, predictive and condition based maintenance utilize real time equipment information to determine the optimal timing of maintenance interventions.

Maintenance performance is commonly assessed using indicators such as Mean Time Between Failures (MTBF), Mean Time to Repair (MTTR), Overall Equipment Effectiveness (OEE), equipment availability, and maintenance cost. Improvements in these indicators reflect enhanced reliability, operational efficiency, and resource utilization. Consequently, organizations are increasingly adopting predictive and condition based maintenance to maximize asset availability while minimizing maintenance costs.

The integration of data driven analytics within Plant Information Systems (PIS) has significantly transformed maintenance management by enabling organizations to make timely, evidence based maintenance decisions. Analytics supports continuous equipment monitoring, early identification of abnormal operating conditions, optimization of maintenance schedules, and prioritization of maintenance activities according to asset criticality. These capabilities facilitate the transition from reactive maintenance towards predictive and condition based maintenance, resulting in improved reliability, reduced downtime, and enhanced operational efficiency (Carvalho et al., 2019).

In Kenya, the adoption of advanced maintenance analytics within electricity utilities remains at an early stage despite increasing investment in digital technologies. KenGen PLC, which operates Geothermal, Hydro, Thermal, and Wind power plants, has implemented PIS,

Licensed under Creative Common

including SCADA and SAP Plant Maintenance (SAP PM), to support plant operations and maintenance. However, maintenance activities remain predominantly preventive and reactive, contributing to recurring equipment failures, unplanned downtime, and increasing maintenance costs (EPRA, 2023).

Furthermore, limited empirical evidence exists on how data quality and availability, analytics tools and technologies, staff analytical competency, and system integration collectively influence maintenance strategy optimization within Kenyan electricity utilities. This study therefore sought to address this contextual and empirical gap by examining how data driven analytics could be leveraged within Plant Information Systems to optimize maintenance strategies at KenGen PLC.

Statement of the Problem

Effective and efficient maintenance management is fundamental to asset intensive industries such as power generation, where equipment reliability directly influences operational performance, service delivery, and organizational sustainability. Although the integration of data driven analytics into maintenance management has enabled organizations to optimize maintenance strategies and improve asset reliability, its adoption and utilization remain uneven across electricity utilities, particularly in developing countries.

Within Kenya's energy sector, maintenance management continues to face significant challenges. Although KenGen PLC has implemented Plant Information Systems (PIS) to support plant operations, the organization has yet to fully leverage data driven analytics to optimize maintenance strategies. Consequently, maintenance remains largely reactive or time based, resulting in frequent equipment failures, prolonged downtime, and increasing operational costs (KenGen PLC, 2023). Given that KenGen PLC generates more than 60% of Kenya's electricity, these maintenance inefficiencies have implications for operational efficiency and the reliability of the national power supply.

While studies conducted in China, the Middle East, and Europe have demonstrated the effectiveness of predictive maintenance and data driven analytics in improving maintenance performance (Carvalho et al., 2019), empirical evidence within Kenya's energy sector remains limited. Specifically, little is known about how data quality and availability, analytics tools and technologies, staff analytical competency, and system integration influence the optimization of maintenance strategies within Plant Information Systems. This study therefore sought to address this gap by examining the influence of data driven analytics on optimizing maintenance strategies at KenGen PLC.

Licensed under Creative Common

Research Objectives

The general objective of the study was to examine how data driven analytics could be leveraged within Plant Information Systems to optimize maintenance strategies at KenGen PLC. The specific objectives were to:

i) Assess the influence of data driven analytics on KenGen PLC's predictive, condition based and preventive maintenance strategies.
ii) Evaluate the effectiveness of data driven analytics in improving asset reliability, reducing failure rates and downtime, and enhancing operational efficiency.
iii) Examine the challenges and enablers affecting the integration of data driven analytics into maintenance decision making processes at KenGen PLC.

THEORETICAL REVIEW

This study is anchored on three complementary theories: Reliability Centered Maintenance (RCM), Systems Theory, and Data Driven Decision Making (DDDM). Individually, these theories explain different aspects of maintenance optimization. Collectively, they provide a comprehensive framework for understanding how data driven analytics can improve maintenance strategies within PIS.

Reliability Centered Maintenance (RCM)

Reliability Centered Maintenance (RCM), developed in the late 1960s and early 1970s by U.S. commercial airlines and the FAA (Nowlan & Heap, 1978), optimizes maintenance by analyzing asset functions, failure modes, and their impacts. Formalized by the Society of Automotive Engineers (SAE), RCM is widely adopted in industries like power generation, oil and gas, and manufacturing. Recent advancements integrate IoT, AI, and machine learning to enable predictive and prescriptive maintenance (Benhanifia et al., 2025). These technologies enhance RCM's ability to anticipate failures and tailor strategies. Its structured methodology ensures alignment with operational and organizational objectives.

RCM provides a theoretical foundation for KenGen PLC to optimize maintenance through data driven analytics. By integrating real time monitoring tools like IoT sensors and predictive algorithms, RCM enables a shift from reactive and preventive to predictive and prescriptive maintenance. This reduces unplanned outages, extends asset lifespans, and enhances cost efficiency in Kenya's power sector. High quality data, advanced analytical tools, skilled staff, and integrated systems are essential for effective implementation. RCM supports regulatory compliance and boosts stakeholder confidence by improving plant reliability.

Licensed under Creative Common

RCM is criticized for being resource intensive, requiring extensive time, expertise, and data for failure mode analysis (Alvarez Alvarado et al., 2022). Its aviation origins may not fully adapt to less regulated, complex environments like power generation in developing countries. Poor data quality or fragmented systems can undermine RCM's reliability. Rigid application may fail to address real time decision making needs in dynamic energy utilities. These challenges highlight the need for flexibility and robust data governance.

Systems Theory

Systems Theory, introduced by Ludwig von Bertalanffy (1968), frames complex entities as interdependent systems rather than isolated parts. In engineering, it has evolved into applications like the “system of systems” (SoS) framework and digital twins, modelling interactions in complex infrastructures (Adesanya et al., 2025). It emphasizes technical, organizational, and human factors converging to enhance performance in asset intensive industries. The theory’s adaptability supports integrated approaches to operational efficiency. Its holistic perspective is widely applied in energy, manufacturing, and other sectors.

Systems Theory frames KenGen’s power plants as interconnected subsystems, including equipment, control systems, operators, and environmental factors. Plant Information Systems (PIS) integrate data from operational technologies (e.g., SCADA, sensors) and IT platforms (e.g., SAP PM, analytics dashboards). This enables feedback loops and coordinates decision making for optimized maintenance strategies. Variables like data quality, analytic tools, staff competencies, and system integration collectively influence outcomes. The theory supports predictive maintenance and operational reliability at KenGen PLC.

Systems Theory is criticized for its abstract nature and limited prescriptive guidance (Jackson, 2019). Its holistic approach can lead to analytical complexity, complicating causal analysis. Implementation requires significant data and resources, which may be challenging in developing countries. The theory may overlook conflicts and institutional constraints affecting system performance. These limitations necessitate practical tools to operationalize the framework effectively.

Data Driven Decision Making (DDDM)

Data Driven Decision Making (DDDM) emerged from evidence based management frameworks in the late 20th century, popularized by Davenport and Harris (2007) as 'competing on analytics.' Advances in Industrial Internet of Things (IIoT), machine learning, and simulation modelling have extended DDDM into operational and maintenance domains (Ledmaoui et al., 2025). In the energy sector, DDDM improves transparency, accountability, and predictive

Licensed under Creative Common

accuracy in maintenance planning (Carvalho et al., 2019). It leverages data to optimize decision making processes.

DDDM transforms raw data from KenGen's Plant Information Systems (PIS), SCADA, and IIoT sensors into actionable maintenance insights. It enables predictive and prescriptive maintenance, improving decision accuracy and resource allocation. By anticipating equipment failures, DDDM enhances reliability, safety, and efficiency. High quality data, analytic tools, staff competencies, and system integration are critical for its success.

DDDM's reliance on high quality data means incomplete or inconsistent data can lead to flawed decisions (Chen et al., 2025). Overemphasis on quantitative evidence may undervalue managerial intuition or experiential knowledge (Szukits, 2022). It requires significant investment in digital infrastructure and skilled personnel, posing challenges in developing countries. Data privacy, cybersecurity, and governance risks are notable concerns.

EMPIRICAL REVIEW

This section reviews empirical literature related to the dependent and independent variables of the study. It examines previous findings on data driven analytics and maintenance strategy optimization and identifies the conceptual, methodological, contextual, and empirical gaps that justify the current study.

Data Quality and Availability

The effectiveness of analytics in optimizing maintenance strategies is fundamentally dependent on the quality and accessibility of data. High quality data must be accurate, complete, timely, and relevant to maintenance objectives (Carvalho et al., 2019). In energy utilities, poor data quality arising from sensor malfunctions, incomplete logging, or inconsistent reporting can undermine predictive models and lead to unreliable maintenance decisions (Rustamova & Rustamov, 2024).

Data availability encompasses the ability to access and retrieve relevant operational information across Supervisory Control and Data Acquisition (SCADA) systems, historian databases, and enterprise platforms such as SAP Plant Maintenance (SAP PM). Where utilities achieve seamless data flow across operational and enterprise systems, predictive models perform more reliably, resulting in improved Mean Time Between Failures (MTBF) and reduced downtime (IoT Analytics, 2023).

Consistency and standardization of data formats across different plant systems are essential for interoperability and effective analytics implementation. Without standardized data models, analytics tools may generate inconsistent results, weakening maintenance decision

Licensed under Creative Common

making (Sahal et al., 2020). Standardization enhances integration with dashboards and reporting systems used in maintenance planning.

Data governance and cybersecurity are also critical dimensions of data quality. As utilities adopt cloud based analytics, digital twins, and remote monitoring technologies, maintaining data integrity and security becomes increasingly important (Chen et al., 2025). Weak governance frameworks may distort predictive insights and reduce organizational confidence in analytics outcomes.

In the Kenyan context, challenges such as fragmented databases, inconsistent logging practices, and limited integration between SCADA and SAP PM systems may constrain predictive maintenance adoption. Strengthening data quality controls and governance mechanisms within KenGen's Plant Information Systems is therefore central to achieving maintenance optimization.

Analytics Tools and Technologies

Analytics tools and technologies provide the infrastructure necessary to capture, process, and interpret operational data for maintenance decision making. In energy utilities, these tools typically include SCADA systems, historian databases, enterprise resource planning modules such as SAP Plant Maintenance (SAP PM), and advanced machine learning algorithms for anomaly detection and predictive modelling (Sahal et al., 2020).

Empirical studies indicate that integration of predictive analytics within maintenance systems significantly improves performance outcomes. Carvalho et al. (2019) demonstrate reductions in unplanned outages and maintenance costs following adoption of machine learning based predictive maintenance, with models capable of identifying subtle deviations in equipment performance that traditional threshold based monitoring systems may fail to detect.

In Kenya, organizations such as KenGen PLC have invested in foundational digital infrastructure including SCADA systems and SAP PM modules. However, advanced predictive analytics tools such as AI enabled prognostics, digital twins, and real time predictive dashboards remain underutilized. This suggests a gap between available technological infrastructure and full analytics driven maintenance implementation.

Effective deployment of analytics tools requires not only acquisition of technology but also integration with existing systems and alignment with operational processes. Where tools are deployed without integration or adequate utilization, expected maintenance performance improvements may not be fully realized.

Licensed under Creative Common

Staff Analytical Competency

Staff analytical competency refers to the knowledge, skills, and organizational readiness required to interpret data and apply analytics insights to maintenance decision making (Mehta et al., 2025). Technical staff must possess competencies in predictive modelling, statistical analysis, and visualization tools to translate operational data into actionable maintenance strategies.

In Sub Saharan Africa, limited availability of advanced analytics skills remains a barrier to predictive maintenance adoption (Rustamova & Rustamov, 2024). Even where digital infrastructure exists, inadequate training and limited cross functional collaboration may restrict effective use of analytics tools.

Organizations that invest in structured training programs and continuous professional development demonstrate measurable improvements in maintenance key performance indicators such as Mean Time to Repair (MTTR), Overall Equipment Effectiveness (OEE), and plant availability (Sensemore, 2023–2024). Analytical competency also involves the ability to communicate insights across departments, ensuring coordination between maintenance engineers, ICT professionals, and operations managers.

Within KenGen PLC, strengthening analytical competencies among engineers and ICT personnel is essential for transitioning from reactive and time based maintenance to predictive and condition based approaches. Without sufficient technical capacity, even well integrated systems may fail to generate meaningful performance improvements.

System Integration

System integration refers to the interoperability between plant level operational systems and enterprise platforms that support maintenance decision making. In energy utilities, integration typically involves linking SCADA systems, historian databases, and ERP modules such as SAP PM to enable seamless real time data exchange (Rustamova & Rustamov, 2024). Integration challenges frequently arise where legacy systems lack compatibility with modern IoT and analytics platforms. Fragmented systems may delay data transmission, limit predictive accuracy, and weaken coordination between maintenance planning and execution processes (Sahal et al., 2020).

Empirical evidence suggests that utilities with integrated systems report improved equipment reliability, optimized spare parts management, and enhanced predictive maintenance capabilities (Sahal et al., 2020). Real time synchronization of operational and enterprise data enables faster fault response and more accurate maintenance scheduling.

Licensed under Creative Common

For KenGen PLC, harmonizing SCADA, SAP PM, and analytics platforms is critical for building an integrated maintenance ecosystem. Effective system integration supports predictive analytics, enhances decision transparency, and strengthens the transition from reactive to proactive maintenance strategies.

Optimized Maintenance Strategies

Reduced downtime is a key outcome of optimized maintenance strategies. In traditional maintenance setups, downtime often results from unanticipated failures or rigid preventive schedules that do not reflect the actual condition of equipment. By leveraging real time data and predictive analytics, utilities can anticipate potential equipment failures, prioritize interventions, and execute timely repairs, thereby minimizing loss of revenue (Carvalho et al., 2019).

Improved asset reliability is another central aspect of optimization. Reliability is the likelihood that plant assets perform their intended function without failure for a designated period (Gholipour et al., 2025). Data driven maintenance strategies such as Reliability Centered Maintenance (RCM) and condition based monitoring focus on aligning interventions with actual equipment conditions. Increased reliability translates into higher capacity utilization and improved operational stability (Alvarez Alvarado et al., 2022).

Cost efficiency is also critical. Maintenance activities constitute a significant proportion of operational expenditure in electricity utilities. Ineffective strategies may lead to over maintenance or under maintenance. Optimized strategies apply predictive models and analytics to guide decisions on resource allocation, spare parts planning, and labor deployment (Ledmaoui et al., 2025).

Summary of Literature Review and Knowledge Gaps

The reviewed literature demonstrates that data driven analytics is a key enabler of maintenance strategy optimization in asset intensive industries. Previous studies have shown that predictive and condition based maintenance reduce equipment downtime, improve asset reliability, and enhance cost efficiency through the effective use of analytics (Carvalho et al., 2019; Alvarez Alvarado et al., 2022).

Despite these advances, the literature reveals several conceptual, methodological, contextual, and empirical gaps. Most existing studies have been undertaken in developed countries and focus on specific technologies such as Artificial Intelligence (AI), Internet of Things (IoT), or machine learning, with limited attention given to the combined influence of data quality and availability, analytics tools and technologies, staff analytical competency, and

Licensed under Creative Common

system integration on maintenance strategy optimization. Furthermore, few empirical studies have examined these relationships within the Kenyan energy sector. This study addresses these gaps by investigating how these four analytics capabilities influence maintenance strategy optimization using KenGen PLC as a case study.

RESEARCH METHODOLOGY

The study adopted a descriptive and explanatory research design (Creswell & Creswell, 2022). The descriptive design was used to assess the existing state of data driven analytics and maintenance practices within KenGen PLC, while the explanatory design was used to examine the relationship between data driven analytics and optimized maintenance strategies.

The target population comprised employees of Kenya Electricity Generating Company PLC. KenGen had 2,425 employees as of 30 June 2025, but the study focused on employees involved in plant operations, maintenance management, asset reliability, and the use of Plant Information Systems, including SCADA, SAP Plant Maintenance, and historian databases. The accessible population comprised 400 employees drawn from Operations, Geothermal Development, ICT, and other relevant functions involved in maintenance planning, execution, and decision making.

The study applied stratified random sampling to ensure representation of the different power generation technologies. The accessible population was stratified into Geothermal, Hydro, Thermal, and Wind. The sample size was determined using Yamane's (1967) formula:

$$n = \frac{N}{1 + N(e^2)}$$

Where:

n = sample size

N = accessible population

e = margin of error

Using an accessible population of 400 and a margin of error of 5%:

$$n = \frac{400}{1 + 400(0.05^2)}$$

$$n = \frac{400}{2}$$

$$n = 200$$

The study therefore targeted 200 respondents. The sample was proportionately allocated across the four strata as shown below in Table 1. Within each stratum, respondents were selected using simple random sampling.

Licensed under Creative Common

Table 1: Proportionate Sample Allocation by Plant Type

Plant Type	Population (Nh)	Weight (Nh/N)	Sample (nh)
Geothermal	180	0.45	90
Hydro	140	0.35	70
Thermal	40	0.10	20
Wind	40	0.10	20
Total	400	1.00	200

Primary data were collected using a structured questionnaire administered electronically through Google Forms. The questionnaire captured demographic information and measured Data Quality and Availability, Analytics Tools and Technologies, Staff Analytical Competency, System Integration, and Optimized Maintenance Strategies. Responses were measured using a five point Likert scale ranging from 1, Strongly Disagree, to 5, Strongly Agree. Open ended questions were also included to obtain respondents' views on the challenges and enablers affecting the integration of data driven analytics into maintenance decision making.

The study also used secondary data obtained from SAP Plant Maintenance records and organizational reports to assess maintenance strategy distribution, downtime trends, and plant operation and maintenance expenditure.

Data was analyzed using IBM SPSS Statistics, applying descriptive and inferential statistical techniques. Reliability of the research instrument was assessed using Cronbach's Alpha, while multicollinearity was tested using the Variance Inflation Factor (VIF). Pearson's correlation analysis was used to determine the strength and direction of relationships among the study variables. Multiple linear regression analysis was used to examine the influence of the four independent variables on optimized maintenance strategies using the model:

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \varepsilon$$

Where:

Y = Optimized Maintenance Strategies

X1 = Data Quality and Availability

X2 = Analytics Tools and Technologies

X3 = Staff Analytical Competency

X4 = System Integration

α = Constant

β1–β4 = Regression coefficients

ε = Error term

The significance of the overall regression model was assessed using the F test, while the significance of the individual regression coefficients was assessed using the t test. The

Licensed under Creative Common

coefficient of determination, R^2, was used to determine the proportion of variation in optimized maintenance strategies explained by the independent variables. Statistical significance was assessed at the 5% level.

Ethical principles were observed throughout the study. Participation was voluntary, respondents provided informed consent, and confidentiality and anonymity were maintained. Approval to conduct the study was obtained from the University of Nairobi, while authorization to collect data within KenGen PLC was obtained from the relevant management.

ANALYSIS, FINDINGS AND DISCUSSION

The data collected was analyzed using descriptive and inferential statistical techniques, and the findings were interpreted in relation to the study objectives. Primary data was obtained through the questionnaire survey, while secondary data was obtained from SAP Plant Maintenance records and organizational reports.

Response Rate

The study targeted 200 employees involved in Plant Information Systems (PIS) and maintenance management at KenGen PLC. A total of 162 questionnaires were completed and returned, representing a response rate of 81%, while 38 questionnaires, representing 19%, were not returned as shown in Table 2.

Table 2: Questionnaire Response Rate

Response	Frequency	Percentage
Responded	162	81%
Not Returned	38	19%
Total	200	100%

According to Mugenda and Mugenda (2003), a response rate of 50% is adequate, 60% is good, while a response rate of 70% and above is considered very good. The response rate of 81% was therefore considered adequate for statistical analysis and interpretation of the study findings.

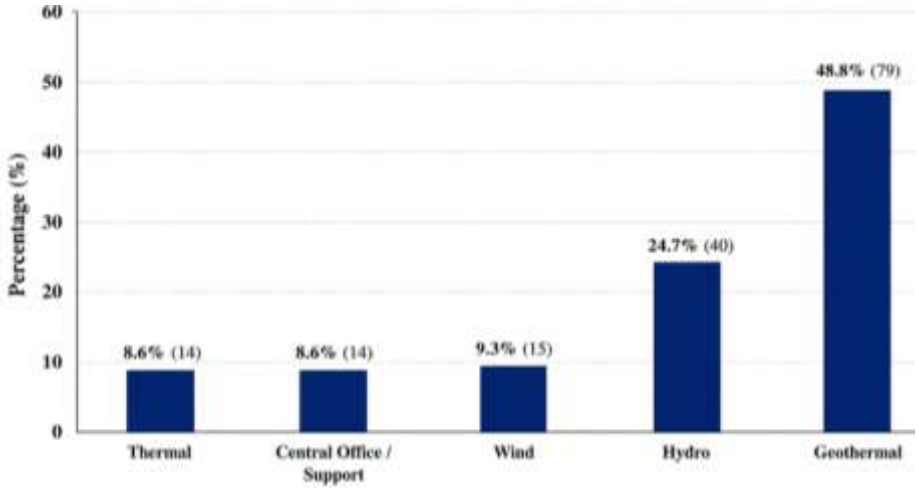
Respondent Profile

The respondent profile was examined to establish the operational and professional background of the participants. The distributions by plant, department, experience, education level, and primary role are presented in Figures 1 to 5.

Licensed under Creative Common

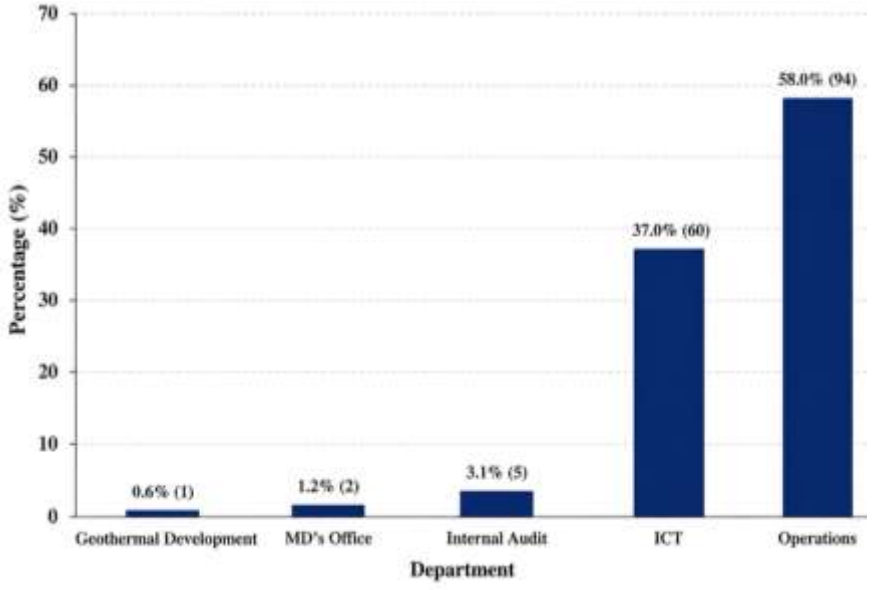
Geothermal contributed the largest share of respondents at 48.8%, followed by Hydro at 24.7%. Wind accounted for 9.3%, while Thermal and Central Office and support functions each accounted for 8.6% as shown in Figure 1.

Figure 1: Respondents by Plant Type



Operations and ICT accounted for most respondents, reflecting the study's focus on maintenance operations, Plant Information Systems, and analytics as shown in Figure 2.

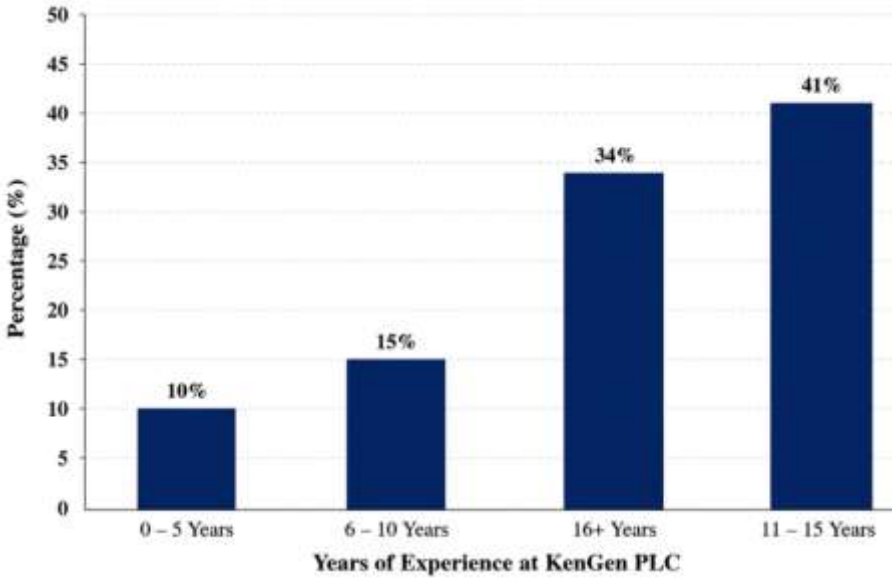
Figure 2: Respondents by Department



Most respondents had substantial institutional experience, with 75.0% having worked at KenGen PLC for more than ten years as shown in Figure 3.

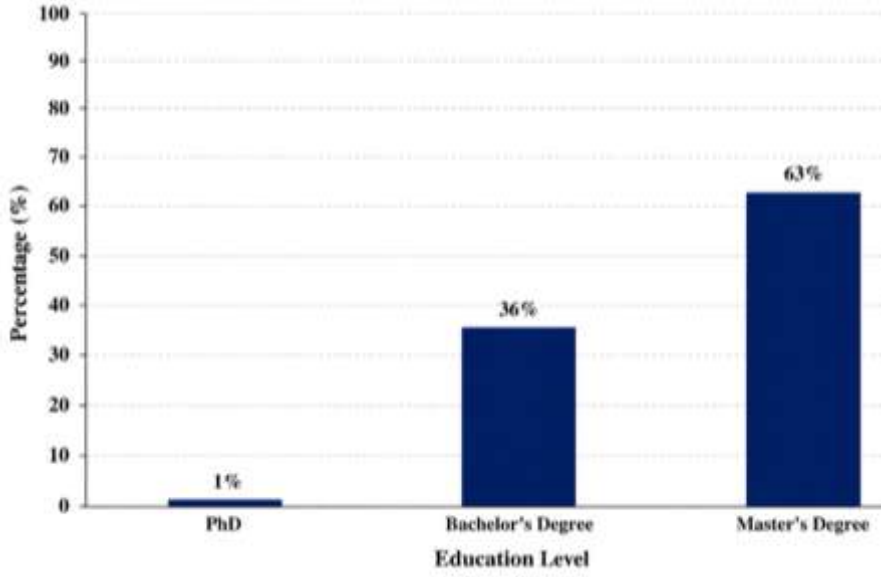
Licensed under Creative Common

Figure 3: Respondents by Years of Experience at KenGen PLC



The majority of respondents held postgraduate qualifications, with master's degree holders accounting for 63.0% of the respondents as shown in Figure 4.

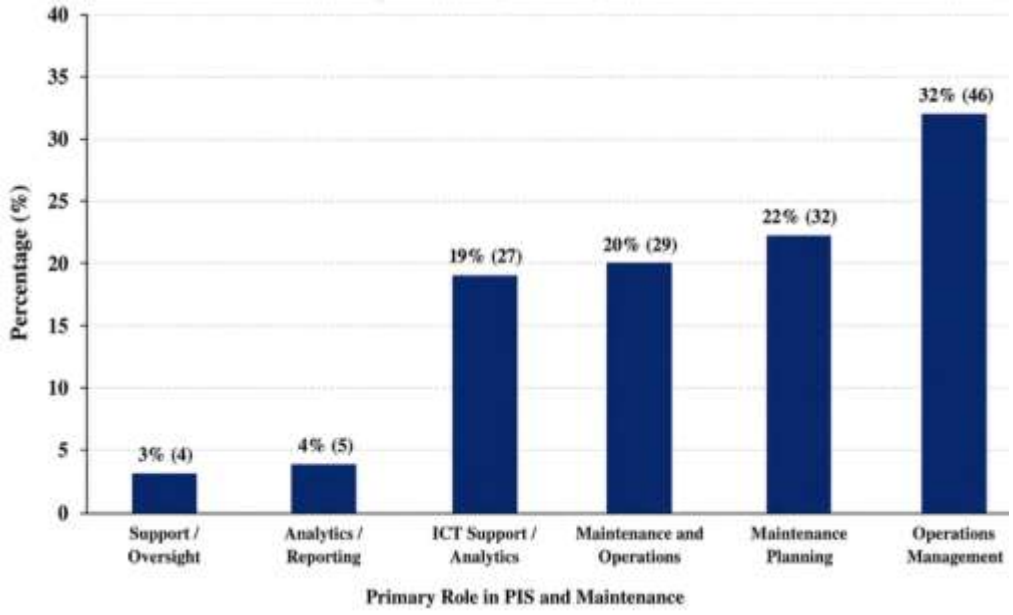
Figure 4: Respondents by Education Level of Respondents



Operations Management, Maintenance Planning, Maintenance and Operations, and ICT Support and Analytics formed the main functional groups represented in the study as shown in Figure 5.

Licensed under Creative Common

Figure 5: Respondents by Primary Role in PIS and Maintenance



Reliability and Diagnostic Tests

Reliability was assessed using Cronbach's Alpha. According to Mugenda and Mugenda (2003), a coefficient of 0.70 and above indicates acceptable reliability. Cronbach's Alpha was established separately for each construct as shown in Table 3.

Table 3: Reliability Test Results

Variable	Cronbach's Alpha	Interpretation
Data Quality and Availability	0.607	Moderate Reliability
Analytics Tools and Technologies	0.862	Good Reliability
Staff Analytical Competency	0.743	Acceptable Reliability
System Integration	0.812	Good Reliability
Optimized Maintenance Strategies	0.981	Excellent Reliability

Four of the five study constructs recorded coefficients above 0.70. Data Quality and Availability recorded 0.607 and was retained because a minimum reliability level of 0.60 may be acceptable in exploratory and social science research (Hair et al., 2022).

Multicollinearity was assessed using the Variance Inflation Factor (VIF). According to Hair et al. (2022), VIF values below 10 indicate that multicollinearity is not a significant concern as shown in Table 4.

Licensed under Creative Common

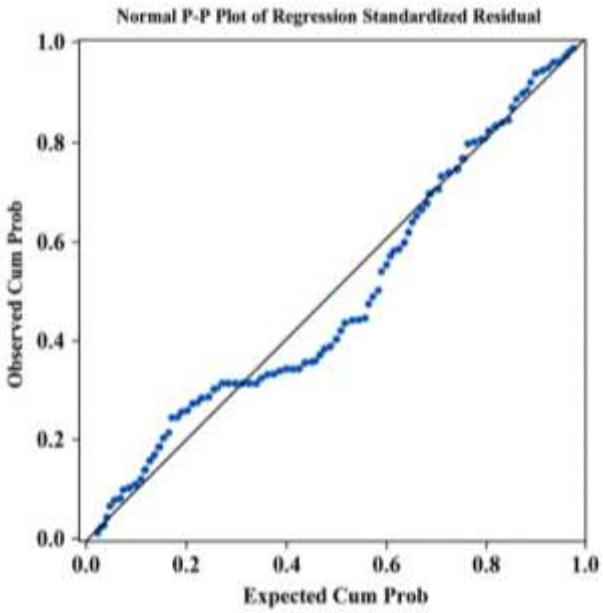
Table 4: Multicollinearity Test

Variable	VIF	Interpretation
Data Quality and Availability	1.490	No Multicollinearity
Analytics Tools and Technologies	1.659	No Multicollinearity
Staff Analytical Competency	1.219	No Multicollinearity
System Integration	1.974	No Multicollinearity

Normality Test

Normality was assessed using the Normal Probability to Probability Plot of the standardized regression residuals. Hair et al. (2022) indicate that normality is supported when the observed cumulative probabilities closely follow the expected cumulative probabilities along the diagonal reference line as shown in Figure 6.

Figure 6: Normality Test



The observed cumulative probabilities closely followed the diagonal reference line, with only minor deviations in the central region. The residuals were therefore approximately normally distributed, and the normality assumption was adequately satisfied.

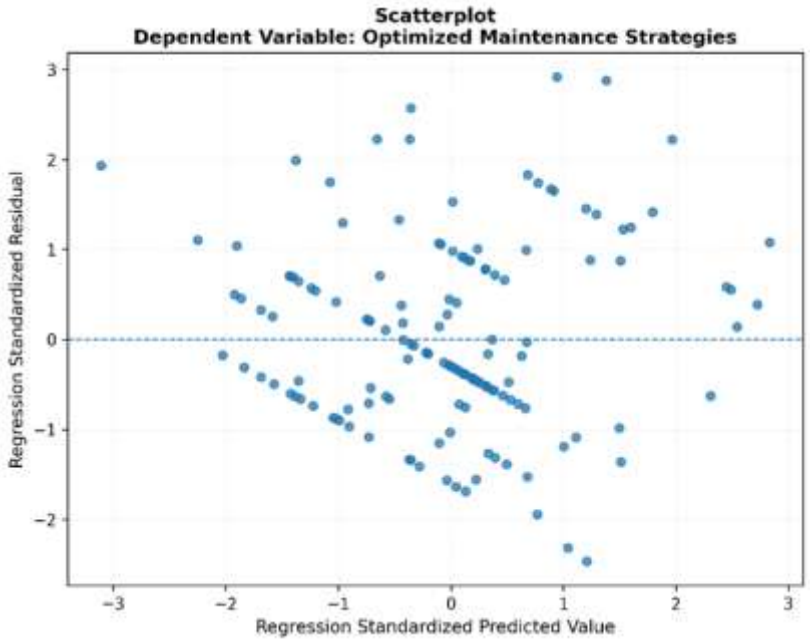
Homoscedasticity Test

Homoscedasticity was assessed using a scatterplot of standardized residuals against standardized predicted values. Hair et al. (2022) indicate that the assumption is satisfied when

Licensed under Creative Common

residuals are randomly dispersed around the horizontal axis without forming a distinct pattern as shown in Figure 7.

Figure 7: Homoscedasticity Test



The residuals were distributed above and below the zero line, but the scatterplot showed some clustering and variation in the spread of residuals across predicted values. The Breusch Pagan test was statistically significant, indicating that the variance of the residuals was not entirely constant and suggesting the presence of some heteroscedasticity.

Linearity Test

Linearity was assessed from the scatterplot of standardized residuals against standardized predicted values. The residuals were generally dispersed around the zero line without a pronounced curved or systematic pattern. The relationship between the independent variables and Optimized Maintenance Strategies was therefore considered reasonably linear, although some variation in the spread of residuals was observed.

Independence of Errors

Independence of errors was assessed using the Durbin Watson statistic. According to Field (2018), the value between 1.5 and 2.5 indicates that the assumption is satisfied. The model produced a Durbin Watson statistic of 1.644, indicating that the residuals were independent and that there was no evidence of significant autocorrelation.

Licensed under Creative Common

Descriptive Statistics

Descriptive statistics were computed to summarize respondents' perceptions of the four independent variables and Optimized Maintenance Strategies as shown in Table 5.

Table 5: Descriptive Statistics

Variable	Mean	Standard Deviation
Data Quality and Availability	3.027	0.425
Analytics Tools and Technologies	2.582	0.768
Staff Analytical Competency	2.809	0.503
System Integration	3.046	0.586
Optimized Maintenance Strategies	2.231	0.943

System Integration recorded the highest mean score (M = 3.046), followed by Data Quality and Availability (M = 3.027). Staff Analytical Competency (M = 2.809) and Analytics Tools and Technologies (M = 2.582) recorded lower scores. Optimized Maintenance Strategies had the lowest mean (M = 2.231), indicating that the expected maintenance optimization outcomes were not yet strongly perceived by respondents.

Correlation Analysis

Pearson correlation analysis was conducted to determine the strength and direction of the relationships between the study variables as shown in Table 6.

Table 6: Pearson Correlation Matrix

Variable	DQA	ATT	SAC	SI	OMS
Data Quality and Availability	1	.384**	.275**	.540**	.325**
Analytics Tools and Technologies	.384**	1	.299**	.609**	.393**
Staff Analytical Competency	.275**	.299**	1	.374**	.500**
System Integration	.540**	.609**	.374**	1	.371**
Optimized Maintenance Strategies	.325**	.393**	.500**	.371**	1

*** Correlation is significant at the 0.01 level.*

Licensed under Creative Common

All four independent variables were positively associated with Optimized Maintenance Strategies. Staff Analytical Competency had the strongest relationship with Optimized Maintenance Strategies (r = 0.500), followed by Analytics Tools and Technologies (r = 0.393), System Integration (r = 0.371), and Data Quality and Availability (r = 0.325).

Regression Analysis

Multiple linear regression analysis was conducted to determine the combined influence of Data Quality and Availability, Analytics Tools and Technologies, Staff Analytical Competency, and System Integration on Optimized Maintenance Strategies.

Table 7: Model Summary

Statistic	Value
R	0.574
R Square	0.330
Adjusted R Square	0.313
Standard Error of Estimate	0.781

The model produced an R Square of 0.330, indicating that the four independent variables jointly explained 33.0% of the variation in Optimized Maintenance Strategies as shown in Table 7.

ANOVA Results

The regression model was statistically significant ($F = 19.315$, $p < 0.001$), indicating that the independent variables jointly had a significant influence on Optimized Maintenance Strategies as shown in Table 8.

Table 8: ANOVA Results

Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	47.141	4	11.785	19.315	0.000
Residual	96.176	157	0.613		
Total	143.317	161			

Regression Coefficients

Staff Analytical Competency was the strongest statistically significant predictor of Optimized Maintenance Strategies ($B = 0.736$, $\beta = 0.392$, $p < 0.001$), followed by Analytics

Licensed under Creative Common

Tools and Technologies (B = 0.260, β = 0.211, p = 0.012). Data Quality and Availability and System Integration were positive but were not statistically significant when the variables were considered simultaneously as shown in Table 9.

$Y = -1.458 + 0.263X_1 + 0.260X_2 + 0.736X_3 + 0.051X_4$

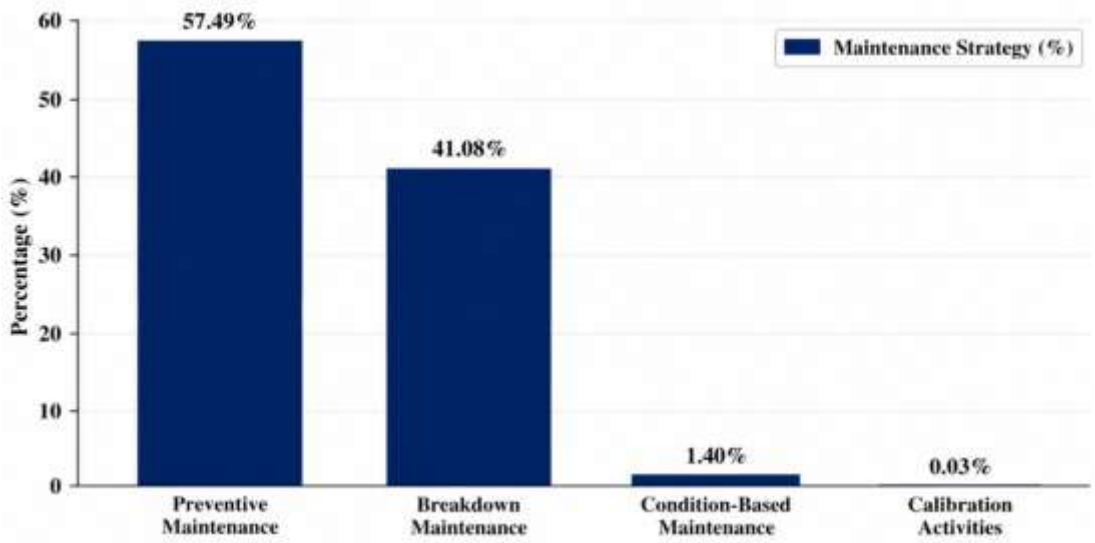
Table 9: Regression Coefficients

Variable	B	Beta	t	Sig.
Constant	-1.458		-2.901	0.004
Data Quality and Availability	0.263	0.119	1.517	0.131
Analytics Tools and Technologies	0.260	0.211	2.548	0.012
Staff Analytical Competency	0.736	0.392	5.519	<0.001
System Integration	0.051	0.032	0.344	0.732

Findings by Research Objectives

The first objective assessed the influence of data driven analytics on predictive, condition based, and preventive maintenance strategies. The survey findings showed positive relationships between all four analytics capabilities and Optimized Maintenance Strategies, with Staff Analytical Competency recording the strongest relationship. The secondary SAP Plant Maintenance data provided further evidence of the maintenance strategy mix as shown in Figure 8.

Figure 8: Distribution of Maintenance Strategies (2021-2025)



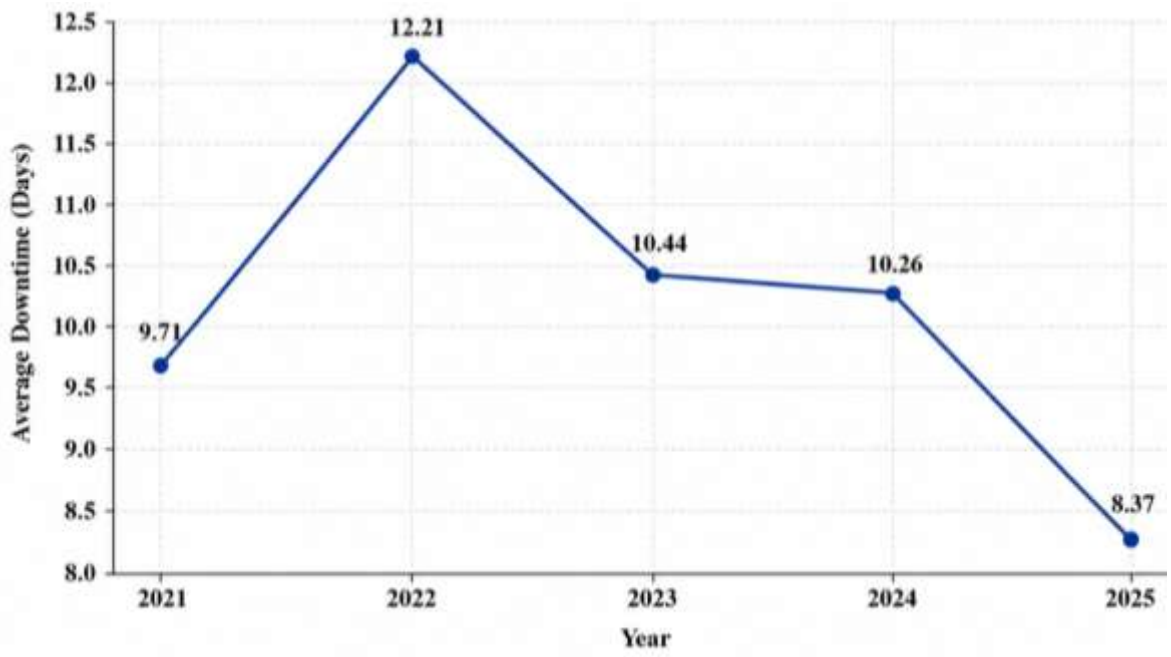
Licensed under Creative Common

Preventive Maintenance accounted for 57.49% of maintenance activities, followed by Breakdown Maintenance at 41.08%. Condition Based Maintenance represented only 1.40%, while Calibration Activities accounted for 0.03%. The results show that maintenance remained predominantly preventive, with limited adoption of condition based maintenance as shown in Figure 8.

The second objective evaluated the effectiveness of data driven analytics through downtime reduction, asset reliability, and cost efficiency. Staff Analytical Competency and Analytics Tools and Technologies were the statistically significant predictors of Optimized Maintenance Strategies.

Average downtime generally declined over the study period, from 9.73 days in 2021 to 8.37 days in 2025, despite an increase in 2022. This points to gradual improvement in maintenance planning and restoration processes as shown in Figure 9.

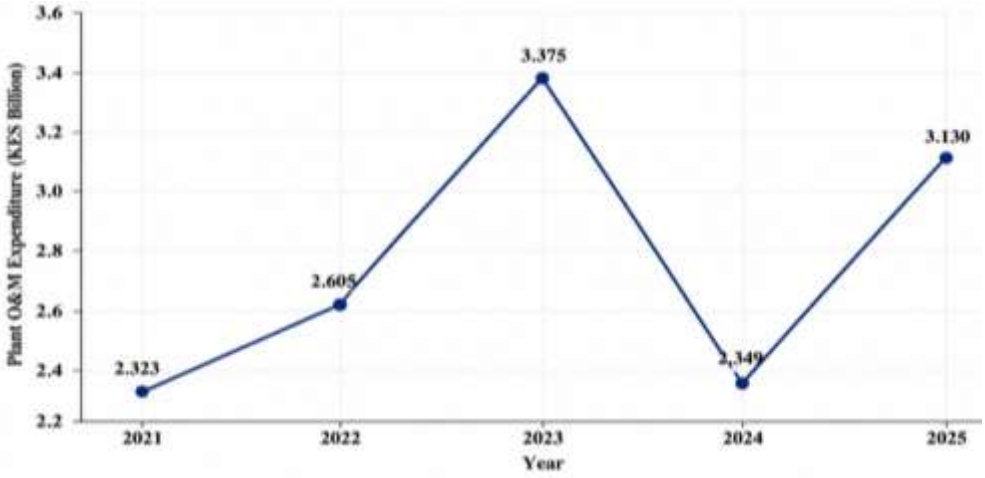
Figure 9: Average Downtime Trend (2021-2025)



Plant operation and maintenance expenditure increased from KES 2.323 billion in 2021 to KES 3.130 billion in 2025, despite a temporary decline in 2024. The expenditure pattern indicates that the improvement in downtime had not yet translated into sustained maintenance cost reductions as shown in Figure 10.

Licensed under Creative Common

Figure 10: Plant Operation and Maintenance Expenditure Trend (2021-2025)



The third objective examined the challenges and enablers affecting analytics integration into maintenance decision making. Of the 162 respondents, 50 (30.9%) provided substantive responses to at least one open ended question, generating 129 coded responses. The main challenges were skills, training and analytical knowledge gaps, system integration and legacy system incompatibility, data quality and accessibility issues, and low adoption or underutilization of analytics tools. Improved planning, scheduling and downtime reduction, together with management support for digital transformation, were identified as key enablers as shown in Table 10.

Table 10: Qualitative Themes for Open Ended Responses

(n = 129 coded responses from 50 respondents)

Theme	Frequency (n)	% of Responses
Challenges		
Skills, training, and analytical knowledge gaps	38	29.5%
System integration and legacy system incompatibility	32	24.8%
Data quality, availability, and accessibility issues	21	16.3%
Low adoption / underutilization of analytics tools	19	14.7%
Resistance to change and organizational culture	12	9.3%
Network connectivity and real time monitoring infrastructure	9	7.0%
Budget and cost constraints	6	4.7%
Cybersecurity concerns	6	4.7%
Enablers		
Improved planning, scheduling, and downtime reduction	8	6.2%
Management/leadership support for digital transformation	3	2.3%

Licensed under Creative Common

Discussion of Findings

The findings indicate that data driven analytics is positively related to Optimized Maintenance Strategies at KenGen PLC, although the low mean score for the dependent variable shows that the expected benefits have not yet been fully realized. Staff Analytical Competency emerged as the strongest predictor, followed by Analytics Tools and Technologies. This is consistent with the Data Driven Decision Making perspective, which emphasizes that value from data depends on the ability to analyze and convert it into operational decisions.

Data Quality and Availability and System Integration recorded relatively higher descriptive scores but were not statistically significant predictors in the regression model. This suggests that data and interconnected systems alone are insufficient to optimize maintenance unless they are supported by adequate analytical capability and effective use of analytics tools.

The secondary data showed that preventive maintenance remained dominant and condition based maintenance was minimally adopted. Although average downtime declined between 2021 and 2025, maintenance expenditure continued to fluctuate. The qualitative findings reinforce the importance of analytical skills, analytics tools, and system integration in extending the use of predictive and condition based maintenance.

SUMMARY

The study examined how data driven analytics could be leveraged within Plant Information Systems to optimize maintenance strategies at KenGen PLC. Specifically, the study assessed the influence of data driven analytics on predictive, condition based and preventive maintenance strategies, evaluated its effectiveness in improving asset reliability, reducing downtime and enhancing operational efficiency, and examined the challenges and enablers affecting the integration of analytics into maintenance decision making.

The regression model was statistically significant, indicating that the study variables jointly influenced Optimized Maintenance Strategies. Staff Analytical Competency emerged as the strongest statistically significant predictor, followed by Analytics Tools and Technologies. Data Quality and Availability and System Integration were positively associated with Optimized Maintenance Strategies but were not statistically significant predictors when the four variables were considered simultaneously.

The first objective established that data driven analytics had a positive influence on maintenance strategies at KenGen PLC. However, secondary SAP Plant Maintenance data showed that Preventive Maintenance remained the dominant maintenance approach, accounting for 57.49% of maintenance activities, followed by Breakdown Maintenance at

Licensed under Creative Common

41.08%. Condition Based Maintenance accounted for only 1.40%, indicating limited adoption of analytics driven maintenance approaches.

The second objective established that data driven analytics contributed to improved maintenance planning, asset reliability and operational decision making. Average equipment downtime generally declined from 9.73 days in 2021 to 8.37 days in 2025. However, plant operation and maintenance expenditure increased from KES 2.323 billion in 2021 to KES 3.130 billion in 2025, despite a temporary decline in 2024. This indicates that improvements in maintenance performance had not yet translated into sustained cost efficiency.

The third objective identified analytical skills gaps, system integration and legacy system incompatibility, data quality and accessibility issues, and low utilization of analytics tools as the main challenges affecting analytics integration. Improved maintenance planning, reduced downtime and management support for digital transformation were identified as key enablers.

CONCLUSIONS

The study concludes that data driven analytics has a positive influence on optimized maintenance strategies at KenGen PLC. Staff Analytical Competency and Analytics Tools and Technologies were the only statistically significant predictors of Optimized Maintenance Strategies, demonstrating that the availability of skilled personnel and effective use of analytics technologies are central to achieving maintenance optimization.

Data Quality and Availability and System Integration recorded relatively favourable descriptive results but did not remain statistically significant predictors after controlling for the other variables. The findings therefore indicate that having reliable data and integrated information systems is not sufficient on its own to improve maintenance outcomes. These capabilities must be supported by analytical competence and effective utilization of appropriate analytics tools.

The study further concludes that data driven analytics has contributed to improved maintenance planning, asset reliability and operational decision making. The reduction in average downtime between 2021 and 2025 indicates gradual improvement in maintenance performance. However, the continued dominance of preventive and breakdown maintenance and the minimal adoption of condition based maintenance show that the full potential of analytics driven maintenance has not yet been realized.

Maintenance expenditure also continued to fluctuate over the study period, indicating that improvements in maintenance performance had not yet resulted in sustained cost efficiency. Overall, KenGen PLC has established a foundation for data driven maintenance through its Plant Information Systems and maintenance data infrastructure, but greater

Licensed under Creative Common

analytical capability, advanced analytics tools and effective system integration are required to strengthen predictive and condition based maintenance.

RECOMMENDATIONS FOR POLICY AND PRACTICE

The recommendations of the study are presented as short, medium and long term actions based on the findings.

Short Term: Within One Year

The study established that Staff Analytical Competency was the strongest predictor of Optimized Maintenance Strategies. KenGen PLC should therefore prioritize continuous training and capacity development for maintenance, operations, engineering and ICT personnel. The training should strengthen staff capability to interpret maintenance data, use available analytics tools and apply analytical insights in maintenance planning and decision making. The organization should also promote more effective use of existing analytics tools and dashboards to support data driven maintenance activities.

Medium Term: One to Three Years

The findings showed that the utilization of advanced analytics tools and technologies remained relatively low despite the availability of maintenance data. KenGen PLC should therefore invest in advanced analytics technologies, including predictive analytics, machine learning applications, real time monitoring tools and maintenance dashboards, to support predictive and condition based maintenance. The organization should also strengthen data governance practices to improve the accuracy, consistency, completeness and accessibility of maintenance data used in maintenance decision making.

Long Term: Beyond Three Years

The study established that System Integration and Data Quality and Availability, although rated relatively favorably by respondents, had not translated into significant improvements in maintenance optimization. KenGen PLC should therefore accelerate the integration of Plant Information Systems, SAP Plant Maintenance, SCADA systems, historians and other operational platforms to enable seamless data sharing and real time maintenance decision making. In addition, the organization should develop an enterprise wide analytics strategy supported by a performance measurement framework to monitor the contribution of analytics initiatives to maintenance performance, asset reliability, downtime reduction and cost efficiency, while supporting continuous improvement.

Licensed under Creative Common

LIMITATIONS OF THE STUDY

The study was conducted within KenGen PLC, and the findings may therefore not be directly generalized to organizations operating in different industries or institutional contexts. The study also relied partly on questionnaire responses based on respondents' perceptions and experiences, and the possibility of response bias could not be eliminated.

The scope of the study was limited to Data Quality and Availability, Analytics Tools and Technologies, Staff Analytical Competency and System Integration. Other organizational, technological and environmental factors that may influence maintenance performance were not examined.

The sample achieved also differed slightly from the planned stratified sample. Geothermal respondents were over represented, Hydro respondents were under represented, and respondents from Central Office and support functions were included although they were not part of the original four stratum sampling frame. This may have affected the representativeness of the intended population. Finally, the cross sectional research design limited the ability to establish causal relationships among the study variables.

FURTHER RESEARCH

Future studies may examine the application of data driven analytics in maintenance management across multiple power generation companies and other asset intensive industries to improve the generalizability of findings beyond a single organization. Longitudinal studies may also be undertaken to examine how adoption of data driven analytics influences asset reliability, equipment downtime, maintenance costs and operational efficiency over time.

Further studies may investigate other organizational, technological and environmental factors that were beyond the scope of this study, including organizational culture, leadership support, change management, digital maturity, data governance, cybersecurity and resource availability. Future research may also examine the contribution of emerging technologies such as Artificial Intelligence, Machine Learning, Digital Twins and the Industrial Internet of Things to predictive maintenance and maintenance optimization.

REFERENCES

Adesanya, F., Castro Silva, K., Graciano Neto, V. V., & David, I. (2025). Systems of twinned systems: A systematic literature review. arXiv. https://doi.org/10.48550/arXiv.2505.19916

Alvarez Alvarado, M. S., Donaldson, D. L., Recalde, A. A., Noriega, H. H., Khan, Z. A., Velasquez, W., & Rodriguez Gallegos, C. D. (2022). Power system reliability and maintenance evolution: A critical review and future perspectives. IEEE Access, 10, 51922–51950. https://doi.org/10.1109/ACCESS.2022.3172697

Licensed under Creative Common

Benhanifia, A., Ben Cheikh, Z., Moura Oliveira, P., Valente, A., & Lima, J. (2025). Systematic review of predictive maintenance practices in the manufacturing sector. Intelligent Systems with Applications, 26, 200501. https://doi.org/10.1016/j.iswa.2025.200501

Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. d. P., Basto, J. P., & Alcalá, S. G. S. (2019). A systematic literature review of machine learning methods applied to predictive maintenance. Computers & Industrial Engineering, 137, 106024. https://doi.org/10.1016/j.cie.2019.106024

Chen, S., Turanoglu Bekar, E., Bokrantz, J., & Skoogh, A. (2025). AI enhanced digital twins in maintenance: Systematic review, industrial challenges, and bridging research–practice gaps. Journal of Manufacturing Systems, 82, 678–699. https://doi.org/10.1016/j.jmsy.2025.07.006

Creswell, J. W., & Creswell, J. D. (2022). Research design: Qualitative, quantitative, and mixed methods approach (6th ed.). SAGE Publications.

Davenport, T. H., & Harris, J. G. (2007). Competing on analytics: The new science of winning. Harvard Business School Press.

Energy and Petroleum Regulatory Authority (EPRA). (2023). Annual report and financial statements 2022/2023. Nairobi: EPRA. https://www.epra.go.ke

Field, A. (2018). Discovering statistics using IBM SPSS Statistics (5th ed.). Sage Publications.

Gholipour, Y., Zare, M., Sereshk, M. V., & Gholipour, Y. (2025). A comprehensive review of maintenance strategies: From reactive to proactive approaches. Central Asia and the Caucasus, 26(1), 70–83.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2022). Multivariate data analysis (9th ed.). Cengage Learning.

IoT Analytics. (2023). *Predictive maintenance market: 5 highlights for 2024 and beyond*. https://iot-analytics.com/predictive-maintenance-market/

Jackson, M. C. (2019). Critical systems thinking and the management of complexity. Wiley.

Kenya Electricity Generating Company PLC (KenGen). (2023). Integrated annual report and financial statements 2022/2023. Nairobi: KenGen PLC.

Ledmaoui, Y., El Maghraoui, A., El Aroussi, M., & Saadane, R. (2025). Review of recent advances in predictive maintenance and cybersecurity for solar plants. Sensors, 25(1), 206. https://doi.org/10.3390/s25010206

Mehta, K., Lwakatare, B., Zörner, W., & Ehrenwirth, M. (2025). Mini grid performance in Sub Saharan Africa: Case studies from Tsumkwe and Gam, Namibia. Sustainable Energy Research, 12, 30. https://doi.org/10.1186/s40807-025-00174-y

Mugenda, O. M., & Mugenda, A. G. (2003). Research methods: Quantitative and qualitative approaches. Acts Press.

Nowlan, F. S., & Heap, H. F. (1978). Reliability centered maintenance. United States Department of Defense.

Papacharalampopoulos, A., Stavropoulos, P., & Salonitis, K. (2022). Data driven predictive maintenance in manufacturing: Current status and future trends. Procedia CIRP, 106, 112–118. https://doi.org/10.1016/j.procir.2022.05.018

Rustamova, S., & Rustamov, F. (2024). Problems of integrating artificial intelligence with SCADA systems. *ScienceRise, 2*, 106–117. https://doi.org/10.21303/2313-8416.2024.003691

Sahal, R., Breslin, J. G., & Ali, M. I. (2020). Big data and stream processing platforms for Industry 4.0 requirements mapping for a predictive maintenance use case. Journal of Manufacturing Systems, 54, 138–151. https://doi.org/10.1016/j.jmsy.2019.11.004

Sensemore. (2023–2024). Essential KPIs and metrics for maintenance management. Sensemore Insights. https://sensemore.io

Szukits, Á. (2022). The illusion of data driven decision making – The mediating effect of digital orientation and controllers' added value in explaining organizational implications of advanced analytics. Journal of Management Control, 33(3), 403–446. https://doi.org/10.1007/s00187-022-00343-w

von Bertalanffy, L. (1968). General system theory: Foundations, development, applications. George Braziller.

Yamane, T. (1967). Statistics: An introductory analysis (2nd ed.). Harper & Row.

Licensed under Creative Common