Communication in Physical Sciences, 2023, 9(4): 1247-1263

A Data-Driven Framework for Detecting Procurement Anomalies and Financial Control Risks in Service Organisations Using Machine Learning and Risk-Based Analytics

Mary Jatta*, Sarah Cobbinah, Alimatou Jallow, Veronica Ibe, Toluwalase Ayeni, and Quadri Salami

Received: 06 July 2023/Accepted: 05 September 2023/Published: 19 September 2023

Abstract: *The fact that purchasing activities are a major source of financial exposure in service organizations arises because procurement transactions cover a large number of vendors, involve various invoices, purchase orders, approvals, and contract-related decisions. Control procedures that rely mainly on periodic sampling are unlikely to pick up complex, scattered, or quickly changing anomalies. In this study, a data-driven approach is presented for detecting purchasing anomalies and financial control risks by combining statistical red-flag analysis, anomaly detection, machine learning, and risk-based audit prioritization. The methodology was created through a secondary analysis of actual procurement cases, comprising about 216,000 procurement transactions from a large government organization in Singapore, over 50,000 procurement observations from SF Express, and around 3.5 million electronic invoices with more than 11 million line items from public bodies in Brazil. The framework in question brings together transaction-value outliers, unusual purchasing frequencies, duplicate or near-duplicate invoices, vendor concentration, abnormal price deviations, temporal patterns, approval exceptions, and Benford-law-based indicators to form a composite purchasing-risk score. Existing evidence shows the practical usefulness of this method: an explainable procurement anomaly-detection system that used 48 fraud indicators reached 67.1% precision in identifying suspicious transactions in the Singapore example, while a machine-learning study based on more than 50,000 procurement records proved the possibility of automatically classifying procurement fraud. Further analysis of the Brazilian electronic invoices revealed that price alerts of high intensity amounted to more than R$560 million in invoice-item value. The results show that data analytics can shift purchasing controls from being mainly based on retrospective sampling to continuous, risk-based monitoring. Nevertheless, the anomaly scores should be seen as indicators that need to be investigated by humans and not as proof of fraud. The study concludes that by combining machine learning, statistical analysis, explainability, and well-established internal-control principles, a practical foundation is provided for improving financial control in service organizations.*

Keywords: *purchasing anomalies; procurement fraud; financial controls; machine learning; anomaly detection; forensic analytics; service organisations.*

Mary Jatta*
School of Business, San Francisco Bay University, California, USA
Email: Maryjatta1996@gmail.com

Sarah Cobbinah
School of Business, San Francisco Bay University, California, USA
Email: sarahcobbinah15@gmail.com

Alimatou Jallow
American International University West Africa, Serekunda, Gambia
Email: jallowalima50@gmail.com

Veronica Ibe
Department of Microbiology, Lagos State University, Lagos, Nigeria
Email: Veronicaibe95@gmail.com

Toluwalase Ayeni
Department of Economics, University of Ilorin, Ilorin, Nigeria
Email: toluwalase09@gmail.com

Quadri Salami
Department of Electrical and Electronics Engineering, University of Lagos, Nigeria
Email: boboabolajie@gmail.com

1.0 Introduction

The purchasing and procurement functions are essential to the financial activities of service organisations (Ali *et al*., 2023; Okpolosa, 2022). Hotels, hospitals, universities, financial institutions, consulting firms, telecommunications companies, government agencies, professional-service organisations and other service-oriented bodies usually handle large numbers of purchase orders, invoices, vendor records, expense claims and payments (Obiebi & Otemu, 2023; World Bank, 2020). Because these transactions are decentralised there are opportunities for errors, duplicate payments, unauthorized purchasing, overpricing, fictitious vendors, conflicts of interest, split purchases and other kinds of control failure (Fang *et* al., 2022; Hysaj, 2023; Nkunda *et al*., 2023). The fact that procurement is becoming more digital has at the same time provided an opportunity to improve financial controls (Harju *et al*., 2023; Duguay et al., 2023; Karttunen *et al*., 2023). Since electronic procurement systems produce transaction-level data which includes dates, suppliers, quantities, prices, approval records, payment details and the purchasing categories, this type of information can be analysed using statistical and machine-learning techniques in order to spot instances that differ from normal purchasing behaviour (Cocciolo *et* al., 2023; Madzimure *et al*., 2023). It has long been known in fraud-detection research that statistical methods and machine learning can be used to identify unusual financial behaviour which cannot easily be picked up by traditional manual procedures (Bello *et* al., 2023; Bolton & Hand, 2002).

Data mining has likewise become an important element of financial fraud detection. As Ngai et al. (2011) thoroughly reviewed, fraud-detection methods have been divided into classification, regression, clustering, prediction, outlier detection and visualization, showing how widely analytical techniques can be applied to financial-control issues. Procurement transactions are particularly suitable for anomaly detection because routine purchasing activities generally follow identifiable and predictable patterns. For instance, suppliers often maintain relatively consistent transaction frequencies, operate within established price ranges, provide specific categories of goods or services, and conduct transactions at fairly regular intervals (Westerski et al., 2021). Consequently, deviations from these established patterns may indicate potential procurement risks or fraudulent activities. Such deviations include purchases made at prices substantially higher than historical levels, the issuance of multiple invoices with similar amounts within a short period, transactions occurring outside normal working hours, and sudden increases in the proportion of an organisation's expenditure allocated to a particular supplier. These irregularities can serve as warning indicators for further investigation.

It should be pointed out that an anomaly does not always indicate fraud; unusual transactions can be the result of legitimate business activities (Hilal *et al*., 2022). For example, emergency purchases, inflation, the signing of new contracts, changes in the demand for services, the introduction of new suppliers, and one-off capital expenditures might all appear statistically unusual even though they do not involve any wrongdoing. For this reason, a sound financial control system should make the distinction between detecting anomalies and determining whether there is fraud. The analytical systems should flag the transactions that need to be examined, and properly

authorized staff should then decide whether the exception is due to an error, legitimate business activity, a breach of policy, or suspected fraud.
Recent studies into procurement provide evidence in favour of this method. Westerski et al. (2021) created 48 fraud indicators and used them on about 216,000 procurement transactions which had taken place over a period of four years in a big government organisation. The explainable anomaly-detection system they developed was able to achieve a precision of 67.1% when identifying suspicious transactions.
Bai & Qiu (2023) looked at the problem of automatically detecting procurement fraud by using a database provided by SF Express. In their study, they represented each procurement event with nine features and applied neural-network models to more than 50,000 procurement records. The results showed that it was feasible to use machine-learning methods to identify and classify suspicious procurement events.
A large-scale study carried out by Interian et al. (2023) looked at about 3.5 million Brazilian electronic invoices involving more than 11 million invoice items. In order to examine abnormal quantities and prices, the researchers applied data-cleaning procedures and used clustering and outlier-detection algorithms. As a result of their analysis, they found price alerts of a high intensity concerning invoice items with a total value exceeding R$560 million. The studies show that procurement analytics can be used on a much larger scale than traditional manual audit procedures. Yet a great deal of reliance is still placed by many service organizations on periodic reconciliations, management approvals and audit procedures based on samples. While such methods are still important, they can be supplemented by continuous analytical monitoring.
A further useful method is Benford's Law. This involves checking whether the distribution of the first digits in naturally occurring numerical data sets differs significantly from the distribution that is expected. Nigrini (2012) gives examples of the application of Benford's Law in forensic accounting, auditing and fraud detection, such as in the analysis of vendors and financial data. Yet Benford's Law should not be regarded as a standalone test for fraud since a number of legitimate data sets do not meet the assumptions required for the law to be applied, and Nigrini himself warns that some financial statement data sets are too small to allow for valid conclusions.
The importance of the financial control aspect of procurement risk should not be overlooked. A transaction can be considered anomalous if an approval has been bypassed, if the segregation of duties has been violated, if the purchase has gone beyond the authorization limit, or if payment has been made without the relevant supporting documentation. It therefore follows that purchasing analytics should not concentrate only on numerical outliers; it should also look into whether transactions comply with the organisation's control rules. The general framework set out in ISO 31000:2018 includes the steps for identifying, analysing, assessing, addressing, monitoring and communicating risks. The standard stresses the need to incorporate risk management into the way organizations make decisions and carry out their operations. The principles involved are compatible with a data-driven framework for monitoring procurement in which the results of analysis are used in risk assessment, investigation and the improvement of control measures.

1.1 Problem statement

Traditional purchasing controls usually function after the transactions have taken place, with internal auditors periodically examining samples of those transactions and the finance departments spotting errors when they carry out reconciliations. The issue is that a small sample might not properly represent the entire population of procurement transactions, and in addition sophisticated control failures can involve a number of individual transactions which are each ordinary but together form an unusual pattern. The research problem

addressed in this study is the development of a systematic, explainable, and data-driven method for detecting purchasing anomalies and financial-control risks in service organisations.

1.2 The aim of the study

This study is aimed at developing and assessing a data-driven framework which can be used for detecting purchasing anomalies and financial-control risks in service organisations by means of statistical analysis, anomaly detection, machine learning and risk-based prioritisation.

1.3 Research objectives

The current study aims at

(i) establishing prominent features in purchasing which have high financial control risk potential,

(ii) formulating a system of analysis parameters in detecting purchasing anomalies,

(iii) analyzing the applicability of various machine learning/statistics methods to the process of identifying procurement risks,

(iv) analyzing the evidential power of documented studies on large-scale procurement,

(v) developing a purchasing risk score for audit prioritization,

(vi) establishing the practical control implications for service organizations.

2.0 Materials and Method

2.1 Research design

The study used a secondary-data analytical research design; instead of creating fictional organizational transactions or reporting simulated results as if they were empirical observations, it relied on the findings from actual documented procurement analytics investigations to develop and validate a general framework.

Three major empirical evidence sources were examined, including:

(i) a Singapore procurement study involving approximately 216,000 transactions; (ii) an SF Express procurement study involving more than 50,000 procurement observations; and (iii) a study of Brazilian electronic invoices which included about 3.5 million invoices and over 11 million invoice items. These sources were chosen because they include real procurement observations and explicitly address anomaly- and fraud-detection problems.

2.2 Data sources

2.2.1 Singapore procurement transactions

Westerski and others in 2021 described a practical system for detecting procurement fraud which was based on about 216,000 transactions from four years of procurement activity in a large government organisation in Singapore; the system used 48 fraud indicators and generated an interpretable risk measure at the transaction level.

The methodology of the study is especially useful to our present research since it uses a number of different indicators instead of depending on a single statistical test.

2.2.2 Data on SF Express procurements

Bai and Qiu (2023) looked at procurement data supplied by SF Express relating to procurement activities from 2015 to 2017 and examined more than 50,000 samples on the basis of nine features connected with procurement. They built neural-network models in order to detect suspicious procurement events and to classify the different types of fraud.

2.2.3 Brazilian electronic invoices

Interian and others analysed nearly 3.5 million electronic invoices which had been issued in relation to purchases made by public organisations in Mato Grosso, Brazil, between 2016 and 2019. The database included over 11 million separate invoice items, and their method involved data cleaning, clustering and the detection of outliers, with a special focus on abnormal quantities and unusual prices.

2.3 Purchasing-risk variables

Fig. 1 presents the proposed data-driven framework for detecting purchasing anomalies and financial-control risks in service organizations, showing the progression from procurement data collection and variable identification through analytical assessment, risk scoring, reporting, follow-up actions and continuous control improvement.

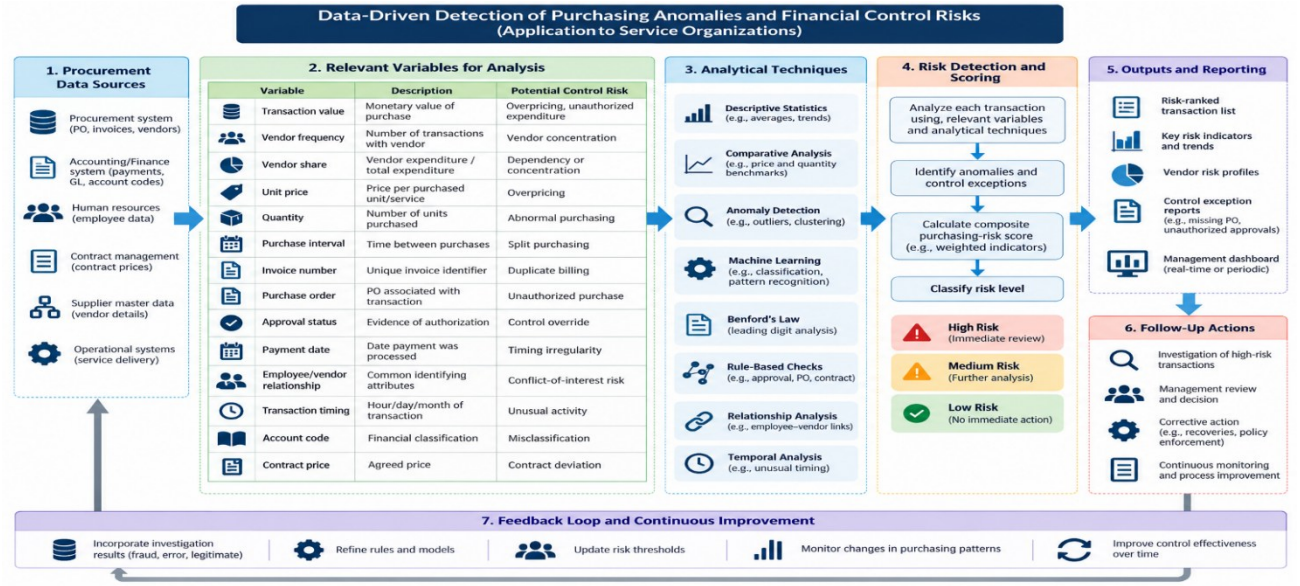


Fig. 1: Data-Driven Framework for Detecting Purchasing Anomalies and Financial-Control Risks in Service Organizations

2.4 Data preprocessing

When putting together a practical implementation, the first step should be to carry out an assessment of data quality. Before any work on the model begins, investigation should be made into the presence of duplicate records, missing supplier identifiers, inconsistent date formats, negative quantities, abnormal currency values and invalid invoice numbers. It is necessary to standardise continuous variables when this is required. Standardisation for a variable x can be written as equation 1

$$Z_i = \frac{x_i - \mu}{\sigma} \qquad (1)$$

where Z_i is the standardized value, x_i is the observation, μ is the mean and σ is the standard deviation. For strongly skewed financial values, logarithmic transformation (equation 2) can be applied

$$x_i' = \ln(1 + x_i) \qquad (2)$$

It diminishes the impact of the very large observations while still keeping their relative order.

2.5 Statistical anomaly detection

A basic statistical anomaly score was constructed using the standardized deviation as shown in equation 3

$$A_i = \left|\frac{x_i - \mu}{\sigma}\right| \qquad (4)$$

Compared with the reference population, a transaction involving a large A_i is statistically unusual. Concerning procurement prices, the adopted approach was to compare each transaction with the distribution of prices for the same supplier, category, location or service type, including equation 5

$$PDR_i = \frac{R_i - \tilde{P}_c}{\tilde{P}_c} \qquad (5)$$

where R_i denotes the transaction price and $\tilde{P}_c$ is the median of the prices from comparable transactions in category c. A large positive price-deviation ratio should be taken as a warning sign rather than as direct proof of overbilling.

2.6 Benford-law analysis

The expected probability of a first digit d under Benford's Law is as expressed in equation 6

$$\mathrm{P(D = d)} = \log_{10}\left(1 + \frac{1}{\mathrm{d}}\right) \qquad (7)$$

Equation 6 is based on the assumption that d = 1…..9

The expected distribution includes about 30.1% for the digit 1, 17.6% for the digit 2 and 12.5% for the digit 3, with the percentages becoming lower as the leading digits increase. The chi-square statistic (equation 8) was used to compare observed and expected frequencies:

$$\chi^2 = \frac{(O_d - E_d)^2}{E_d} \qquad (8)$$

where O_d stands for the observed frequency

and E_d for the expected frequency. Since Benford's law depends on the features of the dataset in question, the test should be carried out only on suitable numerical populations and should be used in conjunction with other indicators. Nigrini (2012) stresses that an analysis based on Benford's law is only an auxiliary tool and not conclusive proof of financial wrongdoing.

2.7 Machine-learning anomaly detection

The proposed framework uses both supervised and unsupervised methods.

Isolation Forest

Isolation Forest is the suitable choice when fraud labels are incomplete or not available; it achieves this by means of recursive partitioning, the observations that are more easily isolated being given higher anomaly scores.

The general anomaly score may be represented as equation 9

$$s(x,n) = 2^{-\frac{E(h)(x)}{c(n)}} \quad (9)$$

Here, E(h(x)) is the expected path length for the observation x, and c(n) is the normalization factor, which is based on the sample size.

Random Forest

Wherever information is available from historical audits, Random Forest is able to categorise transactions as being of low, medium or high risk. The technique involves the use of several decision trees and is fairly robust with regard to nonlinear relationships and interactions.

Neural networks. Neural-network methods are especially relevant when dealing with high-dimensional procurement datasets, as Bai and Qiu (2023) showed that neural-network models can be used for identifying procurement fraud based on more than 50,000 procurement observations.

2.8 Composite purchasing-risk score

To facilitate practical audit prioritization, individual indicators can be normalized and combined:

$$R_i = \sum_{j=1}^{m} w_j X_{ij} \quad (10)$$

where R_i = overall risk score for transaction i, X_{ij} = normalized risk indicator j, w_j = weight assigned to indicator j and m stands for the number of indicators.

2.9 Evaluation measures

For labeled transactions, model performance should be evaluated using equation 11

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

$$Recall = \frac{TP}{TP+FN} \quad (12)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (13)$$

where TP, FP, FN and TN stand for true positives, false positives, false negatives and true negatives, respectively. Precision is especially important when carrying out audit triage since too many false positives can overwhelm the internal audit teams.

3.0 Results and Discussion

3.1 The evidence base for data-driven procurement analytics

Table 1 presents the empirical findings from three published studies on procurement analytics and shows how data-driven methods can be applied to large-scale purchasing transactions. Westerski et al. (2021) looked at about 216,000 procurement transactions relating to four years of activity in a large government organization in Singapore. Their method included 48 fraud indicators and combined the evidence from these different indicators to produce an interpretable transaction-level score for procurement audit officers. The final assessment reached a precision of 67.1% in identifying suspicious transactions. Bai and Qiu (2023) looked at over 50,000 procurement records from SF Express and represented each procurement event by means of nine features before using neural-network models to detect suspicious transactions and to classify the various types of procurement fraud.

Table 1. Empirical evidence used in the study

Study	Organization/context	Data volume	Analytical approach	Reported result
Westerski et al. (2021)	Large government organization, Singapore	Approximately 216,000 transactions	48 fraud indicators, explainable anomaly detection	Precision = 67.1%
Bai & Qiu (2023)	SF Express	Greater than 50,000 procurement observations	9 observed features; neural-network models	The results indicated feasibility of automated procurement-fraud identification
Interian et al. (2023)	Brazilian public-entity electronic invoices	Approximately 3.5 million invoices; >11 million items	Data cleaning, clustering and outlier detection	>R$560 million associated with high-intensity price alerts

Source: Compiled from published studies

Their results showed that machine-learning models could offer useful support for automated procurement-fraud detection, although the authors also noted room for further improvement. Interian et al. (2023) analysed nearly 3.5 million Brazilian electronic invoices containing more than 11 million invoice items from purchases made by public entities in Mato Grosso between 2016 and 2019. By applying data-cleaning procedures, clustering and outlier-detection algorithms, the study examined abnormal quantities and prices and found high-intensity price alerts linked to invoice items worth more than R$560 million.

The study by Westerski et al. (2021) offers direct evidence showing that it is better to combine several procurement-risk indicators than to depend on a single anomaly criterion. The system they developed used 48 indicators, each representing a different type of procurement-fraud risk, and then combined the evidence from these to produce an interpretable score at the transaction level. When it was applied to about 216,000 procurement transactions, the system obtained a precision of 67.1% in identifying suspicious transactions. This shows that a multi-indicator analytical approach can serve as a practical way of reducing a large number of procurement transactions to those that need further audit examination.

The figures in Table 1 show that procurement analytics can be successfully used with transaction sets varying from about tens of thousands to several million records. The sample of approximately 216,000 transactions examined by Westerski et al. (2021) highlights the challenges of depending solely on a manual, transaction-by-transaction review, especially when audit resources and the time available for review are limited. By using 48 indicators and an interpretable scoring system, a large number of transactions could be reduced to a smaller group of observations that were prioritised on the basis of risk. In a similar way, Bai and Qiu's (2023) application of neural networks to more than 50,000 procurement records shows the ability of machine-learning techniques to detect patterns spanning a number of transaction attributes. The Brazilian study also offers additional evidence that automated analytics can be applied to millions of invoices and to individual invoice items. Taken together, these findings indicate that data analytics should be used as a complementary



approach to increasing both the scope and efficiency of procurement controls.

3.2 The effectiveness of multi-indicator anomaly detection

The precision mentioned should nevertheless be understood with care. A precision figure of 67.1% indicates that about 67 out of every 100 transactions identified as suspicious by the system in question were indeed cases that had been classified as suspicious according to the evaluation method used in the study. It does not mean that 67.1% of all procurement transactions were fraudulent, nor does it suggest that those transactions which had not been flagged by the system were necessarily free from fraud. The finding thus shows that the system is useful as a means of screening and prioritising transactions rather than as a method for independently verifying fraud.

The fact that interpretability is important can also be seen in the way the study was designed. Instead of just producing an unexplained anomaly label, the system combined information from various indicators in order to give procurement audit officers an interpretable score. This aspect is especially important for service organisations since procurement officers and internal auditors need to have clear evidence before carrying out further reviews or taking management actions. For this reason, the current study backs a multi-indicator approach in which transaction value, supplier behaviour, price deviations, invoice characteristics, approval status and other control variables all contribute together to the assessment of purchasing risk.

From the point of view of financial control, the result is important since procurement departments usually have only a limited amount of audit resources compared to the number of transactions they process. A multi-indicator system is able to rank the transactions according to the level of concern they represent, which enables auditors to focus their documentary testing and investigation on those cases showing a number of concurrent risk indicators. This method can enhance audit efficiency while still relying on human judgment as the last step in deciding whether an unusual transaction is genuine business activity, an accounting error, a breach of policy or possible fraud.

3.3 The Use of Machine Learning to Detect Anomalies in Procurement

This result matters because machine-learning models can assess combinations of variables rather than relying only on individual thresholds (Onyijen et al., 2023). For instance, in a service organisation a transaction might not be unusually large if looked at on its own, but it could become more unusual when its value is taken together with a supplier who is rarely used, a very short time span since the last purchase, a price that is considerably different from that of similar transactions, the lack of a standard purchase order, or an odd relationship between the employee carrying out the transaction and the supplier. Multivariate relationships of this kind are difficult to spot using separate rule-based tests but can possibly be picked up by machine-learning models.

Bai and Qiu (2023) provide additional evidence on using machine learning to detect procurement risk. Based on more than 50,000 procurement observations acquired from SF Express, the researchers used nine features to represent each procurement event and set up neural-network models in order to identify suspicious procurement transactions and to classify the various types of fraud. The study found that the models were useful for detecting procurement fraud, even though further improvements could still be made.

The findings reported by Interian et al. (2023) provide further evidence of the value of transaction-level analytics for identifying purchasing price risks. The researchers analyzed nearly 3.5 million Brazilian electronic invoices containing more than 11 million invoice items generated from public-entity purchases in Mato Grosso between 2016 and 2019. Their analytical procedure incorporated data cleaning, clustering and outlier-detection algorithms to investigate two principal risk



patterns: abnormal quantities and anomalous prices. The study identified high-intensity price alerts associated with invoice items having an aggregate value exceeding R$560 million.

Even so, the study shows that proper model governance is important. The effectiveness of machine-learning models relies on the quality, representativeness and completeness of the training data; models trained on historical audit decisions might end up reproducing features of earlier decisions, while unsupervised algorithms might detect legitimate transactions simply because they differ from the usual purchasing pattern. It therefore follows that the outputs of machine-learning systems should be regarded as analytical risk indicators which require professional interpretation rather than as conclusive proof of fraud. Since Bai and Qiu (2023) reached this conclusion, their findings support combining machine learning with explainable indicators and with existing procurement-control procedures rather than replacing traditional audit judgment.

3.4 Large-scale anomalies in invoice prices

The fact shows that total expenditure by itself might not contain enough information to evaluate procurement efficiency or to assess financial-control risk. It is possible for an organization to show a seemingly satisfactory level of overall spending while the prices paid for individual items or services are considerably different from those of similar transactions. Transaction-level price analysis can thus offer further information by comparing individual purchases with appropriate reference groups using criteria such as supplier, product or service category, quantity, location, contract terms, and historical purchasing patterns.

For service organisations, this method can be applied not only to physical products but also to a range of professional services, information-technology contracts, equipment maintenance, transportation, healthcare supplies, facility management, consulting services, accommodation and other activities which are outsourced. The main aim of the analysis is not merely to spot transactions that have a high monetary value but to assess whether the price of a transaction is unusually high in comparison with an appropriate group of comparable cases. It is important to make this distinction since a high-value transaction can still be entirely legitimate if it is backed by a valid contract and proper authorisation, while a transaction of moderate value may pose a greater risk if its price is greatly at variance from that of similar purchases.

The Brazilian evidence also shows that automated anomaly detection is scalable. It would take a great deal of time and resources to examine more than 11 million invoice items manually, whereas computational methods are able to systematically compare large groups of transactions and spot observations that deserve further investigation. It follows that automated price-risk analysis should be used as part of continuous procurement monitoring. Yet, as is the case with the other analytical methods examined in this study, a price alert should be regarded as a warning signal, not as proof of overpricing or fraud since valid differences can occur due to quality, urgency, contractual conditions, market conditions, location and other factors.

3.5 Purchasing Control-Risk Indicators

The framework suggested classifies purchasing anomalies according to six main dimensions: price risk, transaction risk, vendor risk, approval risk, temporal risk and relationship risk. The indicators associated with these dimensions and their financial-control implications are given in Table 2. These dimensions have been chosen since they reflect different aspects of procurement behaviour and enable both transaction-level and supplier-level risks to be examined within a single analytical framework.



Table 2. Suggested purchasing-risk indicators

Risk dimension	Examples of analytical indicators	Control implication
Price risk	Unit-price deviation, price percentile, supplier-price variance	Possible overpricing or contract deviation
Transaction risk	Duplicate invoices, unusual quantities, unusual values	Error, duplicate payment or manipulation
Vendor risk	New supplier, concentration, inactive supplier reactivation	Vendor-management weakness
Approval risk	Missing PO, approval after purchase, threshold override	Authorization failure
Temporal risk	Weekend/holiday purchases, unusual timing, clustered transactions	Possible circumvention or unusual activity
Relationship risk	Shared addresses, bank details, contact information	Potential conflict-of-interest indicator

For instance, the fact that a new supplier has appeared is not in itself suspicious since it is common for organisations to enter into agreements with new vendors. The level of risk, however, could go up if the new supplier is linked to a high-value transaction, involves a large price deviation, lacks competitive documentation, has had recent changes to its bank account details or has triggered an approval exception. The framework therefore focuses on the combination of various risk indicators rather than considering each anomaly as separate evidence of misconduct. This method is in line with the multi-indicator procurement analytics approach described by Westerski et al. (2021), in which several indicators were combined to produce an interpretable risk assessment at the transaction level.

Table 2 shows that price risk involves checking whether the monetary features of a purchase differ considerably from those of appropriate reference transactions, while transaction risk looks at any irregularities in the transaction itself. Vendor risk broadens the analysis to include supplier behaviour, such as supplier concentration and unusual supplier activity, and approval risk assesses compliance with the existing procurement authorization procedures. Temporal and relationship risks add further dimensions that might indicate unusual purchasing patterns or possible conflicts of interest. It is important to combine these different aspects since no one indicator is necessarily enough on its own to prove there is a control problem.

3.6 Duplicate and Nearly Duplicate Purchasing Transactions

Duplicate-payment detection is an important application of procurement analytics since duplicate transactions can lead to actual financial losses even if each invoice seems legitimate. An exact duplicate can be found by matching combinations of the supplier's identification, the invoice number, the invoice date and the transaction amount. For example, a simple duplicate key can be expressed as

$D_i = (\text{Vendor}_i, \text{Invoice Number}_i, \text{Amount}_i)$

while an alternative matching structure can incorporate invoice date:

$D_i = (\text{Vendor}_i, \text{Invoice Date}_i, \text{Amount}_i)$

Such procedures are able to detect transactions which have identical or very similar attributes. Yet exact matching will fail in cases where duplicate invoices include minor alterations to the invoice numbers, the dates, the descriptions or the amounts. As a result, fuzzy-matching procedures can be used to compare the invoice descriptions, the supplier identifiers, the

transaction values and the relevant purchase-order information.

For instance, if a series of transactions involving Vendor A with invoice numbers 10345, 10345A and 10346—all amounting to ₦2,450,000—relate to the same service period or purchase order, they should be subject to analytical review. The significance of this pattern increases when the transactions take place over a short time frame and seem to reflect a single underlying expenditure. This type of analysis may also help in identifying split purchasing, which occurs when a large expenditure is divided into a number of smaller transactions, possibly in an attempt to avoid procurement or authorization thresholds.

The importance of carrying out duplicate and near-duplicate analysis lies in the fact that it looks at the relationships between transactions rather than assessing each transaction on its own. This approach is in agreement with the wider fraud-detection literature, where the usefulness of statistical and analytical methods for spotting unusual patterns in groups of financial transactions is acknowledged (Bolton & Hand, 2002; Ngai et al., 2011). Yet an analytical match should be treated as a signal for further review since legitimate recurring invoices can also have similar features regarding the supplier, the amount and the date.

3.7 Vendor Concentration

Vendor concentration offers a supplementary measure at the level of suppliers with regard to procurement risk. The fraction of the total amount spent on procurement that relates to a specific supplier can be expressed as

$$VC_j = \frac{Purchase\ from\ vendor, j}{Total\ procurement\ expenditure} \quad (14)$$

The figure (VC_j) is the proportion of the vendor's expenditure. The higher this value, the greater the proportion of the organization's procurement expenditure that is concentrated with that supplier. The concentration measure can be complemented by the Herfindahl–Hirschman Index (HHI), expressed as

$$HHI = \sum_{j=1}^{N} s_j^2 \quad (15)$$

It is significant to note that vendor concentration by itself is not proof of fraud or improper procurement. A service organization may properly rely heavily on a specialized supplier due to technical requirements, the existence of long-term contracts, geographical restrictions, or because there are no suitable alternatives available. Yet, very high levels of concentration can also warrant examining competitive bidding processes, contract renewal practices, supplier dependence, and controls related to conflicts of interest. The value of this indicator is therefore in highlighting the procurement arrangements that deserve further investigation rather than in automatically considering concentrated procurement as being a problem.

In which s_j denotes the fraction of the total procurement expenditure that belongs to supplier (j) and N is the total number of suppliers; the HHI gives an overall indication of the degree of concentration among suppliers, higher values meaning a greater level of concentration.

3.8 Control Exceptions

One of the special features of the proposed framework is the combination of financial-control exceptions with statistical and behavioural anomalies. A transaction can have a normal monetary value and yet indicate a serious control failure if the necessary authorization procedures have been avoided. On the other hand, a high-value transaction can be legitimate if it is supported by a valid contract, appropriate authorization and documented operational requirements.

The presence of a required approval can therefore be represented by the binary control-exception variable

$$CE_i = \begin{cases} 1, if\ required\ approval\ is\ absent \\ 0. otherwise \end{cases} \quad (16)$$

Similarly, a purchase-order exception can be represented as

$$PO_i = \begin{cases} 1, if\ the\ transaction\ lacks\ a\ valid\ purchase\ order \\ 0. otherwise \end{cases} \quad (17)$$

while an authorization-threshold exception can be expressed as

$$PO_i = \begin{cases} 1, if\ the\ transaction\ exceeds\ the\ applicable\ authorization\ threshold \\ 0. otherwise \end{cases} \quad (18)$$

The binary indicators can then be included in the overall purchasing-risk score. Doing so enhances the framework since it enables analytical detection to take into account not just whether a transaction is unusual but also whether it adheres to the organization's established controls. This difference is especially significant in the case of service organizations, where procurement authorization, segregation of duties and documentation requirements are vital elements of financial governance.

3.9 Benford-Law Indicators

Benford's Law offers another statistical method for identifying appropriate transaction populations. The probability that the first digit (d) will occur in a naturally occurring numerical data set is given by

$$P(D = d) = log_2\left(1 = \frac{1}{D}\right), D = 1, 2 \ldots .9$$

The expected first-digit distribution that results is given in Table 3.

Table 3. The expected first-digit distribution according to Benford's Law

First digit	Expected proportion (%)
1	30.10
2	17.61
3	12.49
4	9.69
5	7.92
6	6.69
7	5.80
8	5.12
9	4.58

Table 3 shows that the smaller digits are likely to occur more often than the larger ones, the digit 1 as the first digit making up about 30.10% of the observations as against 4.58% for the digit 9. A large discrepancy between the observed and the expected distributions might thus serve as a signal for further investigation. Nigrini (2012) speaks of the use of Benford's analysis in forensic accounting, auditing and fraud detection.

Benford's analysis must, however, be approached with caution since a departure from the expected distribution does not prove that fraud has taken place, as it is possible for legitimate datasets to fail to meet the conditions required for Benford's Law to apply. Things such as product prices, regulated tariffs, contractual amounts, identification numbers and invoice values with very tight constraints may naturally show non-Benford distributions. It follows that Benford analysis should be used as a supplementary indicator as part of the overall purchasing-risk assessment and not regarded as a standalone fraud test.

3.10 Proposed Risk Classification

The individual indicators mentioned above can be combined together to form a composite purchasing-risk score in order to aid in the prioritization of audits. For instance, the indicators might be adjusted so that they all fall on a common 0–1 scale and then be given weights based on their relative importance within the organization's control environment. A general representation of the suggested score is

$$R_r = \sum_{I=K}^{k} w_K X_{IK} \quad (19)$$

The overall risk score for transaction (i) is represented by (R_i), the normalized value of risk indicator (k) by (X_{ik}), the weight given to that indicator by (w_k), and K stands for the total number of indicators. The weights should

be determined by referring to the organization's risk priorities, historical audit results and, where appropriate, by the performance of the empirical model. The proposed operational classification is presented in Table 4.

Table 4. Proposed Purchasing-Risk Classification

Risk score	Category	Recommended action
0.00–0.24	Low	Routine monitoring
0.25–0.49	Moderate	Periodic review
0.50–0.74	High	Targeted audit testing
0.75–1.00	Critical	Prompt investigation and control review

Table 4 shows that as the risk scores increase, the review requirements become stronger. Transactions that are classified as low risk will still be subject to routine monitoring, while those that are classified as high risk or critical risk will receive a more detailed analytical or documentary examination. This arrangement enables the limited audit resources to be focused on transactions which have one or more severe risk characteristics.

The operational thresholds listed in Table 4 are suggested rather than having been empirically proven as universal cutoff values; they must therefore not be taken to mean that a transaction with a score of 0.75 is inherently fraudulent or that one with a score of 0.74 is necessarily acceptable. Each organization should adjust the thresholds on the basis of its historical transaction data, confirmed audit results, records of control failures, and its own risk tolerance. This method is in line with the risk-management principles set out in ISO 31000:2018, which stress the importance of identifying, assessing, treating and monitoring risk in a way that is specific to the particular context.

3.11 Discussion

The results and the empirical evidence examined in this study show that data-driven procurement analytics can in several important respects supplement traditional financial-control procedures. Firstly, analytical methods offer considerably wider population coverage since the whole transaction database can be checked before only a smaller number of cases are chosen for in-depth examination. The large datasets studied by Westerski et al. (2021), Bai and Qiu (2023) and Interian et al. (2023) prove that it is possible to apply computational techniques to procurement populations consisting of anywhere from tens of thousands to millions of observations.

Secondly, the suggested framework allows the relationships between variables to be looked at rather than considering each control feature on its own. A transaction might have a normal value yet become more significant when it is combined with an unusual supplier, a large price deviation, a missing purchase order, an approval exception, or an unusual transaction interval. This approach that takes multiple variables into account is in line with the wider financial-fraud literature, in which classification, clustering, prediction and outlier-detection techniques are regarded as useful methods for detecting unusual financial behaviour (Bolton & Hand, 2002; Ngai et al., 2011). Thirdly, the framework enhances the process of prioritising audits by transforming a number of analytical signals into a structured risk assessment. Westerski et al. (2021) illustrated this approach by using 48 procurement fraud indicators together with an interpretable score at the transaction level and achieved a precision of 67.1% in their study. While this finding supports the application of analytical triage, it should not be taken as representing the overall performance level for all organisations. Instead, it shows that combining several indicators can help auditors to reduce large groups of transactions to those that need further examination.

The results also indicate that explainable analytical systems should be used. In order to



start a further investigation, procurement officers and internal auditors need to know the reason why a transaction has been flagged. It is therefore more suitable for practical implementation to adopt a hybrid architecture which combines deterministic control rules, statistical tests, anomaly detection, machine learning and human review than to depend on a single algorithm. With this approach, explicit control violations can be detected by means of rules while statistical and machine-learning methods are able to identify less obvious relationships and deviations.

The findings of Interian et al. (2023) also support the idea that outlier detection can be scaled up for use in the analysis of procurement prices. By looking at millions of electronic invoice items, their study shows that automated methods are able to detect possible important anomalies in both price and quantity within groups of transactions that could not easily be examined in full using manual methods. It is necessary, though, to provide a contextual interpretation of the alerts obtained, since price differences can be due to legitimate reasons such as quality, urgency, location, contract terms, quantity and market conditions.

The results indicate that the main benefit of data-driven procurement analytics is not in taking the place of auditors or in automatically declaring transactions as fraudulent, but rather in enhancing the coverage, consistency, explainability, and prioritization of financial-control activities; this conclusion is in line with the difference between anomaly detection and fraud determination as stated in the financial-fraud literature (Bolton & Hand, 2002; Ngai et al., 2011).

3.12 The implications for financial control

The fact that the proposed framework has significant implications for financial control in service organizations means that its capacity to screen transactions continuously can help complement periodic reconciliation and audit activities by enabling potential anomalies to be spotted nearer to the time when they take place. As a result, there will be less need for retrospective detection and management will be given more up-to-date information regarding emerging procurement risks.

The framework also enables risk-based internal auditing since it allows auditors to prioritize transactions on the basis of their analytical risk features rather than depending solely on random or judgmental sampling. By means of supplier-level indicators, further assistance can be given to vendor management through the identification of unusual concentration, odd price behaviour and changes in supplier activity. In a similar way, approval indicators can be used to detect transactions which do not comply with purchase-order and authorization requirements.

From the point of view of an investigation, high-risk transactions can be put together with the relevant supplier, employee, purchase order, approval, invoice and historical transaction data. Doing so allows investigators to obtain a more comprehensive evidence trail than if they looked at each invoice separately. At the management level, procurement-risk information can also be summed up by department, supplier, expenditure category and reporting period in order to assist with financial-control oversight.

The applications are in line with ISO 31000:2018, a standard that stresses the need for a systematic approach to the identification, analysis, assessment, management, monitoring and communication of organizational risks. It follows that the framework in question offers a means by which procurement analytics can be integrated into a wider risk-management and internal-control system rather than remaining just a standalone fraud-detection tool.

3.13 Limitations

When interpreting the results, several limitations of the study should be taken into account. The framework is mainly based on empirical evidence that has already been published and on methodological analysis rather than on a proprietary transaction dataset

obtained from one particular service organization. As a result, while the study shows the usefulness and conceptual value of data-driven procurement analytics, it does not prove what the universal performance level would be for all service organizations.

Second, the empirical studies looked at different organizational settings, procurement procedures, transaction structures and analytical methods. For instance, the 67.1% precision cited by Westerski et al. (2021) refers to the performance achieved in the particular evaluation context of their study and cannot be taken as the precision that could be expected in other organizations. In the same way, the findings given by Bai and Qiu (2023) and Interian et al. (2023) pertain to their own datasets and methodological circumstances.

Thirdly, anomaly detection is able to identify observations or patterns that differ from expected behaviour without thereby proving that there has been intentional misconduct. As a result, legitimate transactions can be flagged due to emergency procurement, the establishment of new contracts, unusual operational requirements, changes in market conditions or other factors. Human investigation together with relevant documentary evidence is still necessary before any claim of fraud or misconduct can be made.

Fourthly, supervised machine-learning techniques rely on the quality of the labels employed when developing the model, since historical audit decisions might have inconsistencies or contextual biases which a trained model could then replicate. Unsupervised methods, on the other hand, have a different kind of limitation in that they may treat legitimate but unusual transactions as anomalies. These limitations highlight the need for model validation, explainability and human oversight.

In the end, implementation calls for proper data governance, privacy protection, cybersecurity measures and access controls since procurement analytics can involve sensitive information about employees, suppliers, finance and contracts. For successful deployment it is necessary not only to use appropriate analytical techniques but also to have reliable data, clearly established governance responsibilities, secure information systems and well-defined procedures for examining and settling analytical alerts.

4.0 Conclusion

A data-driven framework for the detection of purchasing anomalies and financial-control risks in service organisations was developed in this study. The framework combines statistical outlier detection, Benford's law screening, duplicate detection, vendor analysis, price-deviation analysis, control-exception rules, machine learning and explainable risk scoring. Evidence from actual procurement studies shows the practical potential of this method. In the case of a procurement system in Singapore which used 48 fraud indicators, about 216,000 transactions were analysed and a precision rate of 67.1% was achieved in spotting suspicious transactions. Another procurement fraud study, based on more than 50,000 observations of SF Express, showed that neural-network-based procurement analysis is feasible. Moreover, a large-scale analysis of around 3.5 million Brazilian electronic invoices and over 11 million invoice items also proved that outlier-detection techniques can detect abnormal prices and quantities, with high-intensity price alerts relating to more than R$560 million in invoice-item value.

The main point is that financial control can be improved by changing from mostly periodic and sample-based reviews to continuous, risk-based and data-driven monitoring, since such a system is able to look at the whole purchasing population, detect any unusual transactions, rank the level of risk and give auditors evidence for specific investigations.

It by no means follows that analytical flags should be automatically regarded as indicating fraud. The proper function of data analytics is to pinpoint transactions, vendors, employees, purchasing categories and control processes



which then need to be looked into more closely. Final judgments must still be based on documentary evidence, organizational policies, professional audit procedures and suitable human judgment.

For service organizations, the recommended architecture is therefore a hybrid model in which business rules provide control assurance, statistical methods identify unusual observations, machine learning detects complex patterns, explainable analytics supports investigation, and human auditors make final determinations. Such integration provides a scalable foundation for improving purchasing transparency, reducing control failures and strengthening financial-risk management.

5.0 References

Ali, F. A., Hassan, H. B., & Kadir, M. R. A. (2023). The effect of budgetary control on organisational performance in small and medium enterprises (SMEs) in Somalia. *International Journal of Professional Business Review*, *8*, 12, Article e04119. https://doi.org/10.26668/businessreview/2023.v8i12.4119

Bai, J., & Qiu, T. (2023). *Automatic procurement fraud detection with machine learning* [Preprint]. arXiv. https://arxiv.org/abs/2304.10105

Bello, O. A., Folorunso, A., Onwuchekwa, J., Ejiofor, O. E., Budale, F. Z., & Egwuonwu, M. N. (2023). Analysing the impact of advanced analytics on fraud detection: A machine learning perspective. *European Journal of Computer Science and Information Technology*, *11*, 6, pp. 103–126. https://doi.org/10.37745/ejcsit.2013/vol11n6103126

Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, *17*, 3, pp. 235–255. https://doi.org/10.1214/ss/1042727940

Cocciolo, S., Samaddar, S., & Fazekas, M. (2023). Government analytics using procurement data. In *The government analytics handbook: Leveraging data to strengthen public administration*. World Bank.

Duguay, R., Rauter, T., & Samuels, D. (2023). The impact of open data on public procurement. *Journal of Accounting Research*. https://doi.org/10.1111/1475-679X.12479

Fang, H., Fang, F., Hu, Q., & Wan, Y. (2022). Supply chain management: A review and bibliometric analysis. *Processes*, *10*, 9, Article 1681. https://doi.org/10.3390/pr10091681

Harju, A., Hallikas, J., Immonen, M., & Lintukangas, K. (2023). The impact of procurement digitalization on supply chain resilience: Empirical evidence from Finland. *Supply Chain Management: An International Journal*, *28*, 7, pp. 62–76. https://doi.org/10.1108/SCM-08-2022-0312

Hilal, W., Gadsden, S. A., & Yawney, J. (2022). Financial fraud: A review of anomaly detection techniques and recent advances. *Expert Systems with Applications*, *193*, Article 116429. https://doi.org/10.1016/j.eswa.2021.116429

Hysaj, A. (2023). Decentralized finance. *International Journal of Advanced Natural Sciences and Engineering Researches*, *7*, 10, pp. 158–162.

International Organization for Standardization. (2018). *ISO 31000:2018: Risk management—Guidelines*. International Organization for Standardization.

Interian, R., Carpanese, I., Mello, B., & Ribeiro, C. C. (2023). Red flag algorithms for Brazilian electronic invoices: Outlier detection and price risk classification. *International Transactions in Operational Research*, *31*, 3, pp. 1533–1553. https://doi.org/10.1111/itor.13323

Karttunen, E., Lintukangas, K., & Hallikas, J. (2023). Digital transformation of the purchasing and supply management process. *International Journal of Physical Distribution & Logistics Management*, *53*, 5–6, pp. 685–706.

https://doi.org/10.1108/IJPDLM-06-2022-0199

Madzimure, J., Mafini, C., & Dhurup, M. (2020). E-procurement, supplier integration and supply chain performance in small and medium enterprises in South Africa. *South African Journal of Business Management*, *51*, 1, Article a1838. https://doi.org/10.4102/sajbm.v51i1.1838

Ngai, E. W. T., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. (2011). The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature. *Decision Support Systems*, *50*, 3, pp. 559–569. https://doi.org/10.1016/j.dss.2010.08.006

Nkunda, R. M., Kazungu, I., & Changalima, I. A. (2023). Collaborative procurement practices in public organizations: A review of forms, benefits and challenges. *Ghana Journal of Development Studies*, *20*, 1. https://doi.org/10.4314/gjds.v20i1.3

Nigrini, M. J. (2012). *Benford's law: Applications for forensic accounting, auditing, and fraud detection*. John Wiley & Sons. https://doi.org/10.1002/9781119203094

Obiebi, O. S., & Otemu, E. (2023). Audit fees and auditor's tenureship: Perception of selected audit firms in Nigeria. *Research Journal of Management Practice*, *3*, 11, pp. 41–55.

Okpolosa, M. O. (2022). Budgetary control system and financial performance of listed manufacturing firms in Nigeria. *BW Academic Journal*, *1*, 1, p. 9. https://bwjournal.org/index.php/bsjournal/article/view/703

Onyijen, O. H., Olaitan, E. O., Olayinka, T. C., & Oyelola, S. (2023). Data-driven machine learning techniques for the prediction of cholera outbreak in West Africa. *International Journal of Applied and Natural Sciences*. https://bluemarkpublishers.com/index.php/IJANS

Westerski, A., Kanagasabai, R., Shaham, E., Narayanan, A., Wong, J., & Singh, M. (2021). Explainable anomaly detection for procurement fraud identification—Lessons from practical deployments. *International Transactions in Operational Research*, *28*, 6, pp. 3276–3302. https://doi.org/10.1111/itor.12968

World Bank. (2020). *Egypt digital economy country assessment*. World Bank Group. https://www.worldbank.org/

Compliance with Ethical Standards

Declarations:

The authors declare that they have no conflict of interest.

Data availability

All data used in this study will be readily available to the public.

Consent for publication

Not Applicable.

Ethical Consideration

This has been declared in Section 3 of the text

Availability of data and materials: The publisher has the right to make the data public.

Competing interests

The authors declared no conflict of interest.

Funding

The authors declared no source of funding

Authors' Contributions

Mary Jatta conceptualised the study, developed the methodology and coordinated the research. Sarah Cobbinah and Alimatou Jallow contributed to literature review and data analysis. Veronica Ibe and Toluwalase Ayeni participated in evidence synthesis, interpretation and framework development. Quadri Salami contributed to machine-learning analysis, technical evaluation and manuscript review. All authors contributed to the preparation, critical revision and approval of the final manuscript