Design of a decision-making system for monitoring and early identification of underperformance risks in companies with multiple subsidiaries *

Abdoul Kamal ASSOUMA[1,*], Abdoul-Matine OUSMANE[1], Tahirou DJARA[1], Maria Sandra DEMAHOU[1], Houévo Pascal Anicet FADOHOUN[1]

[1] *Université d'Abomey-Calavi*

Abstract

In this study, we designed a decision support system for monitoring performance and detecting risks of underperformance within companies and organizations. The developed architecture relies on a star-schema data warehouse, with an ETL process implemented using Talend for data integration into PostgreSQL. The system is based on a hybrid data-mining approach, combining K-Means clustering with business rules to detect underperformance risks. The results show a concordance rate of more than 80% between the clusters and the business rules, validating the robustness of the identification model. The interactive dashboard developed in Power BI enables real-time comparative monitoring of performance indicators. This solution provides decision-makers with a dynamic and scalable strategic management tool, with proven capability in early detection of units at risk of underperformance.

Keywords

Decision support system, data warehouse, ETL, data mining, K-Means, performance.

1. Introduction

Digital transformation places data valorization at the heart of organizational performance. Organizations generate significant volumes of data daily, creating both an opportunity and a challenge for strategic decision-making. However, many institutions struggle to detect signs of underperformance promptly, due to poor utilization of their data.

Several case studies of companies worldwide well illustrate the consequences of the underperformance. For example, NOKIA[1] lost its dominant position due to its slow pace of innovation and poor strategic choices, resulting in a drastic decline in its market share. Similarly, KODAK[2] was weakened by its inability to adapt to digital technologies and by ineffective technology management, leading to its bankruptcy in 2012. In Africa, Air Afrique[3] and NIPOST [4] also suffered from poor operational performance and poor governance, resulting in financial losses, the closure of units, and a decline in customers. These examples demonstrate that underperformance, whether operational, strategic, or technological, can have serious consequences for an organization's survival and competitiveness.

Recent studies confirm that organizational underperformance stems less from a lack of data than from an inability to leverage it strategically. In light of this, this research aims to design a decision-

*CITA 2026 – Land, Smart Cities, and Sustainable Development, 25-26 June 2026, Cotonou, Benin

1,*Corresponding author.

abdoulk1@hotmail.com (A. K. ASSOUMA); matines1@yahoo.fr (A. Matine OUSMANE); csm.djara@gmail.com (T. DJARA) ; demahoumarias@gmail.com (M. S. DEMAHOU); hfadohoun@gouv.bj (H. P. A. FADOHOUN).

0000-0002-8591-6610 (T. DJARA)

making system that enables performance monitoring and the early detection of underperformance risks.
The central issue is how to implement a decision-making system capable of transforming operational data into actionable indicators for anticipating underperformance within companies/organizations.
This paper proposes the design of a decision-making system combining a star-shaped data warehouse modeled in PostgreSQL, an ETL process with Talend, and an interactive dashboard using Power BI. The solution includes a hybrid data mining approach combining K-Means clustering with business rules for detecting underperformance risks, thus offering a dynamic and scalable strategic management tool.

2. Literary review

2.1. Decision-making systems: concepts and architecture

A decision support system (DSS) encompasses all the IT tools used to analyze operational data to inform strategic decision-making. Unlike transactional systems (OLTPs) focused on processing routine operations, DSSs rely on an analytical approach (OLAP) that enables complex processing of large amounts of data. Their architecture is generally structured around a data warehouse that consolidates and stores information from heterogeneous sources (Excel files, databases, applications).
According to Inmon [5], the data warehouse forms the foundation of the DSS by providing an integrated and consistent view of the data, while Kimball and Ross [6] favor a user-oriented approach through dimensional modeling. These systems facilitate multidimensional analysis, trend identification, and predictive modeling, thus becoming essential levers for organizational performance.

2.2. The Data Warehouse: Principles and Modeling

A data warehouse is defined as a subject-oriented, integrated, historical, and non-volatile collection of data (Inmon)[5]. Its classic architecture comprises several layers: heterogeneous data sources, ETL (Extract, Transform, Load) processes for data cleaning and integration, the central data warehouse, thematic data marts, and reporting tools.

Multidimensional modeling relies on three key elements: facts (measurable events), dimensions (contextual axes of analysis), and measures (numerical indicators).

Several schemas structure this modeling: the star schema (simple and efficient), the snowflake schema (normalized dimensions), and the constellation schema (sharing dimensions among several facts). This structure enables OLAP operations (slice and dice, drill-down, pivot) essential for decision analysis.

2.3. Organizational performance and indicators

According to Lebas [7], performance is "the result of collective action directed towards an objective, in a given context." It is therefore measured by both effectiveness (achievement of results) and efficiency (optimal use of resources).

Organizational performance encompasses an organization's ability to achieve its objectives by optimizing its resources, based on three fundamental dimensions: relevance (suitability of objectives), effectiveness (achievement of results), and efficiency (optimization of resources).

Kaplan and Norton's Balanced Scorecard [8] structured this approach by integrating multiple perspectives (financial, customer, internal processes, learning).

A performance indicator is a quantitative or qualitative variable that provides information on the degree to which a set objective has been achieved. Key performance indicators (KPIs) fall into three categories: effectiveness, efficiency, and quality indicators. Early detection of underperformance risks, whether internal (poor resource allocation) or external (economic changes), is a major strategic challenge for organizations and their performance.

2.4. BI and Data Mining Tools

Business Intelligence (BI) encompasses the technologies that transform raw data into actionable insights for decision-making. Major solutions include Power BI (intuitive interface, Microsoft ecosystem), QlikView/Sense (associative model), SAP BI (ERP integration), and Oracle BI (high performance with large data volumes).

Data mining complements BI by applying statistical and algorithmic methods to uncover hidden patterns within data. Data mining is an integral part of the overall Knowledge Discovery in Databases (KDD) data extraction process, which includes data selection, preprocessing, transformation, mining, and interpretation.

Techniques such as clustering (K-Means), classification, and regression help identify abnormal trends and anticipate risks of underperformance, thus providing essential predictive insights for strategic management.

2.5. Related works

The study of organizational performance has garnered increasing interest in the literature, particularly since the emergence of Decision Support Systems (DSS) in the 1990s. These systems have evolved over time from simple technological tools to genuine levers for improving organizational performance. The seminal research of Inmon [5] and Kimball & Ross [6] established the principles of data warehouse structuring and multidimensional modeling, highlighting that design quality directly influences the speed and reliability of analyses.

In 2009, Golfarelli and Rizzi [9] subsequently demonstrated the crucial importance of ETL processes and data governance in the overall performance of DSS. This approach was complemented in 2017 by the work of Osunrinde and Tiamiyu [10], which highlighted the correlation between information quality and organizational performance, particularly in African contexts where a lack of integrated analytical tools persists.

The specific field of employment has benefited from the contributions of the ILO [11] and the research of Al-Radaideh & Al-Nagi [12], validating the usefulness of performance indicators and data mining techniques for predictive analysis. In 2023, Koch-Rogge [13] further enriched this approach with Data Envelopment Analysis, enabling the identification of underperforming entities.

The shift towards Business Intelligence and Data Mining marked a decisive turning point, enabling the transition from a descriptive approach to a predictive and prescriptive one (Provost & Fawcett ; Watson) [14][15]. Kaplan and Norton's Balanced Scorecard [8] embodied this evolution by integrating financial, operational, and strategic dimensions into performance management.

African studies, notably those by Okeke-Uzodike and Chitakunye [16] and Hartley's master's thesis [17], confirm that the adoption of modern decision-making systems improves transparency, strategic planning, and decision quality, despite persistent technological and organizational constraints.

All of this work converges on a unanimous observation: organizational performance fundamentally depends on the ability of decision-making systems to transform data into reliable indicators and strategic actions, thereby strengthening the efficiency, responsiveness and consistency of decisions.

2.6. Analysis of existing work

Previous work clearly demonstrates the growing importance of decision-making systems and data mining techniques for improving organizational performance assessment. However, several limitations can be identified. For example, while the ILO [11] offers a comprehensive methodological framework for monitoring the performance of employment agencies, this framework remains general and does not provide an operational model integrating a data warehouse and automated analytical tools.

The work of Al-Radaideh & Al-Nagi [12] demonstrates the effectiveness of clustering for identifying homogeneous groups, but it focuses primarily on external performance (user or employee satisfaction), not on the internal operational performance of organizations.

Furthermore, Provost and Fawcett [14] and Osunrinde and Tiamiyu [10] highlight the importance of data quality, but they do not address the challenges related to integrating historical data from operational systems into a decision-making architecture. The gap between theory and practice remains significant, particularly in African public administrations, where decision-making tools are rarely deployed systematically.

Although Koch-Rogge [13] demonstrates the value of the DEA method for evaluating the effectiveness of service units, this approach has a major limitation: it is a static approach that does not allow for tracking performance changes over time.

A major limitation lies in the lack of integration between business expertise and algorithmic models. None of the reviewed studies proposes a hybrid approach combining clustering and business rules adapted to the specificities of the domain, which can affect the operational relevance of the solutions.

Finally, the practical challenges of integrating historical data from operational systems into a comprehensive decision-making architecture remain insufficiently addressed, particularly in the African context where the systematic deployment of decision-making tools is lacking.

Thus, the existing literature highlights the necessary theoretical foundations (performance indicators, data mining, segmentation, decision-making systems), but reveals gaps in adapting these concepts to a complete decision-making model dedicated to monitoring and early detection of underperformance within companies/organizations. These gaps justify the development of an integrated model combining theoretical foundations and practical implementation.

3. Materials and methods

3.1. Experimental framework of the model

The proposed model is validated using data from a public Beninese agency responsible for implementing national employment and professional insertion policies. It serves as an intermediary between job seekers and employers, while also contributing to employment promotion through its 13 departmental branches.

To improve the management and monitoring of its activities, this agency uses the Client Information System (SICA), an integrated platform that automates the entire intermediation process between job seekers and employers. It enables the digital management of registrations (for both job seekers and employers), job offers and applications, placements, interviews and assessments, training activities, and all operations carried out within its departmental branches. It covers the full operational cycle, from a job seeker's registration to placement or follow-up within an insertion program.

Although highly efficient and containing a large volume of data, SICA remains primarily an operational management tool. The reports it generates offer no real capability for dynamic, predictive, or comparative analysis across branches. This makes it difficult to quickly identify underperforming areas and anticipate performance gaps.

The design of a decision-support system is therefore necessary to leverage the existing data in SICA. Such a system will provide analytical and interactive monitoring of the performance of the departmental branches, enable early detection of underperformance risks, and thereby improve responsiveness in the management of the agency's branches.

3.2. Methodological approach for the design of the decision-making system

The chosen methodological approach is based on the principles of CRISP-DM. This iterative process comprises six phases: organizational context analysis, data preparation, conceptual modeling, technical development, evaluation, and deployment.

The system's technical architecture is based on a layered structure. The source layer integrates standardized CSV files from operational systems. The ETL layer manages the data extraction, transformation, and loading processes.

The data warehouse uses a star schema (Figure 1), while the analytical layer integrates visualization tools and predictive models. Finally, the business intelligence interface provides an interactive dashboard for end users.

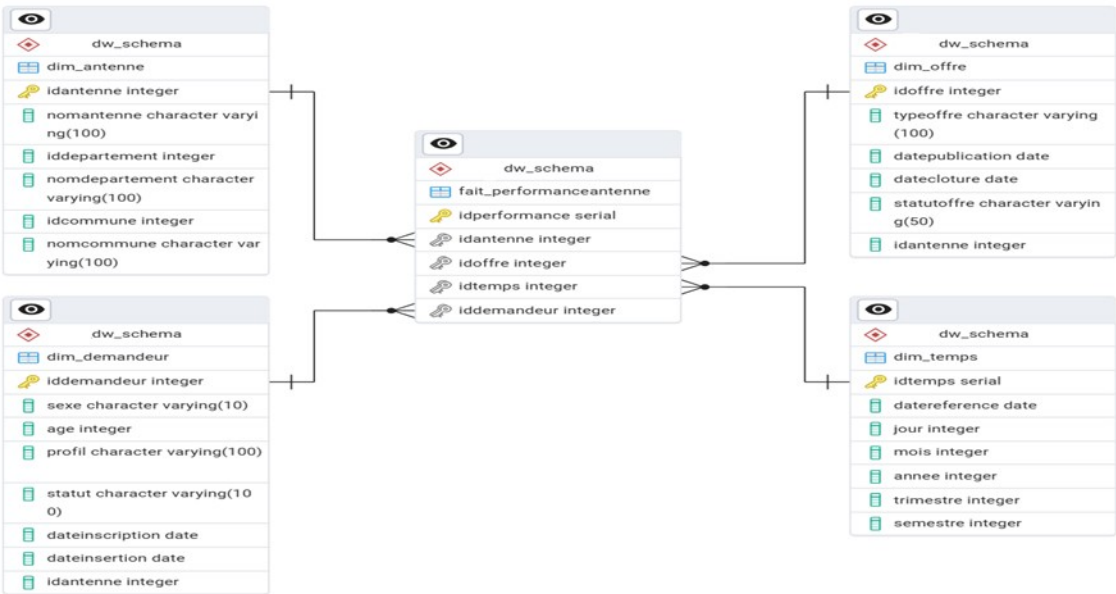

Figure 1: Star diagram model of the warehouse

The conceptual modeling adopts a star schema with a central fact table (Performance) and four main dimensions: Time for temporal analyses, Antenna or focal point for organizational divisions, Offer for the catalog of offers, and Demander for customer or user profiles. Strict protocols ensure data quality, including duplicate removal, format standardization, handling of missing values, and validation of temporal consistency.

3.3. Technical implementation

The PostgreSQL technology tool is used for the business intelligence warehouse, Talend Open Studio for data integration, Power BI for interactive visualization, and Python enhanced with specialized libraries for predictive analytics.

The ETL process follows three distinct steps. The extraction phase collects data from standardized CSV files, verifying source integrity and logging metadata. The transformation phase implements data cleaning via specialized components, standardization of formats (dates, text, numeric), and aggregation according to target dimensions. The target database loading phase manages the incremental population of dimension tables and the updating of the fact table, with an error handling and incident recovery system. Automation is ensured by a master job that sequentially orchestrates the loading of reference dimensions, the population of the temporal dimension, the updating of the fact table, and the verification of referential consistency (Figure 2).

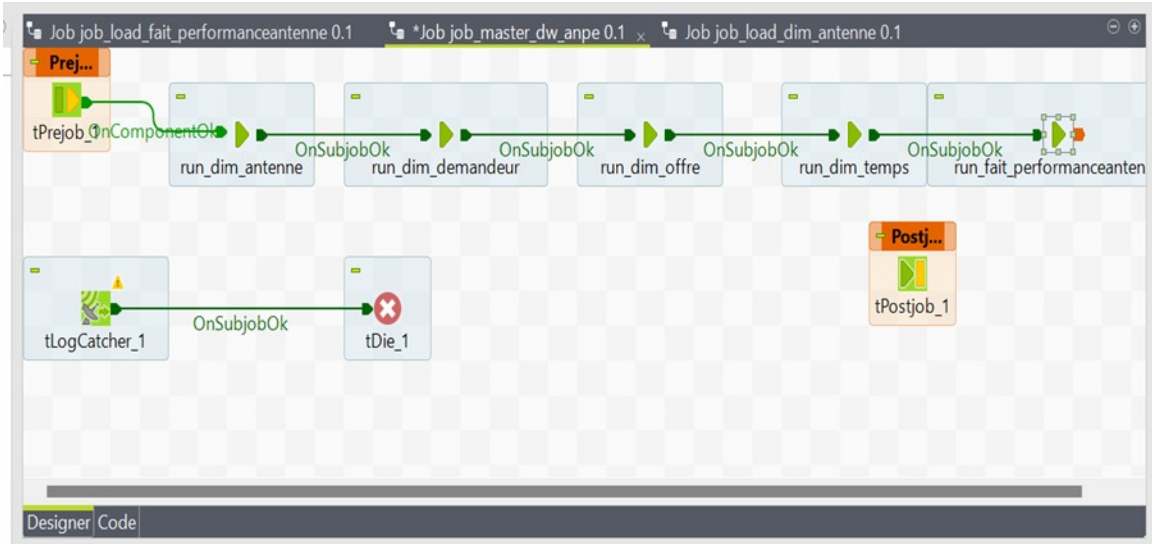

Figure 2: Warehouse Loading Job Master

DAX metrics allow you to calculate key indicators in Power BI: overall and segmented insertion rate, average processing times, satisfaction rate and other detailed performance indicators.

3.4. Risk detection and validation

The proposed hybrid approach to detecting the risk of underperformance integrates unsupervised clustering (K-Means) (Figure 4) for automatic segmentation and business rules for contextual validation. Analytical data preparation includes selecting key variables or indicators (insertion rate, satisfaction, lead times), aggregating indicators by organizational unit, and performing min-max normalization for comparative analysis.

Determining optimal clusters uses the elbow method, testing k values from 2 to 5 and calculating silhouette scores. The selected configuration (k=3) (Figure 3) has a mean silhouette score of 0.35, indicating acceptable separation.

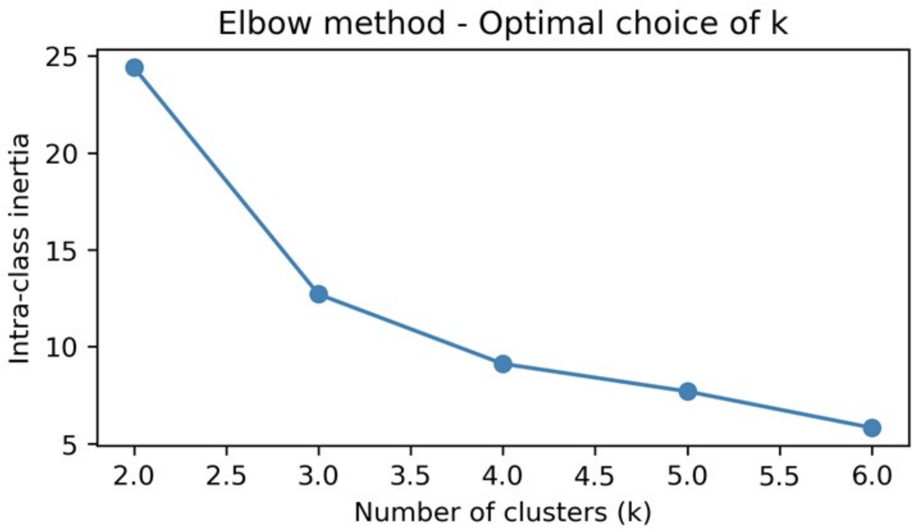

Figure 3: Elbow method (bend at K=3)

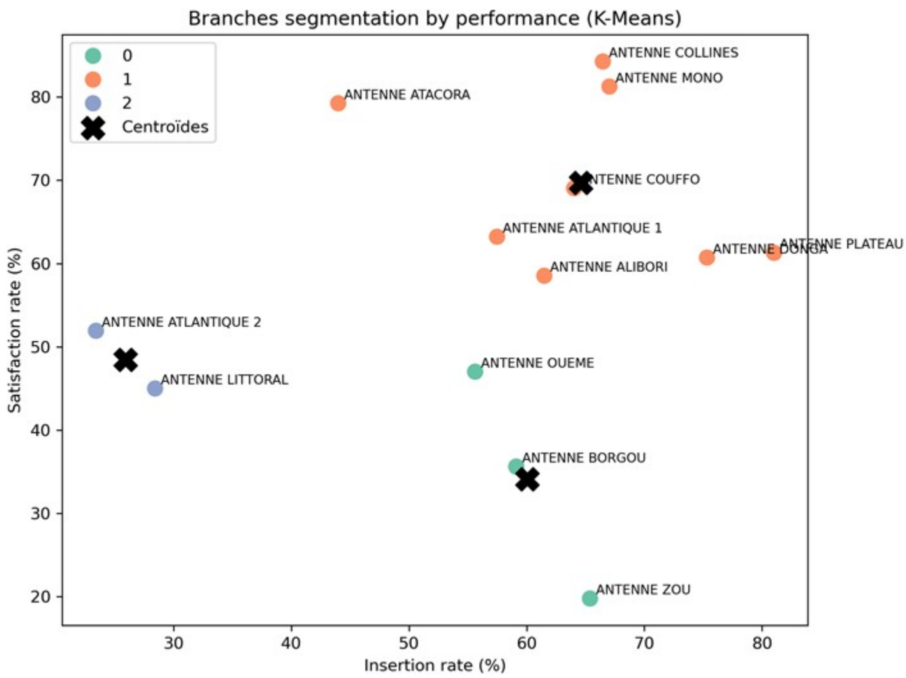

Figure 4: K-Means branches Classification

The integration of business rules defines critical thresholds based on expertise: an insertion rate below 50%, a satisfaction rate below 30%, or a processing time exceeding 15 days triggers a high risk. The combination of scores allows for the calculation of an overall risk index (Figure 5)

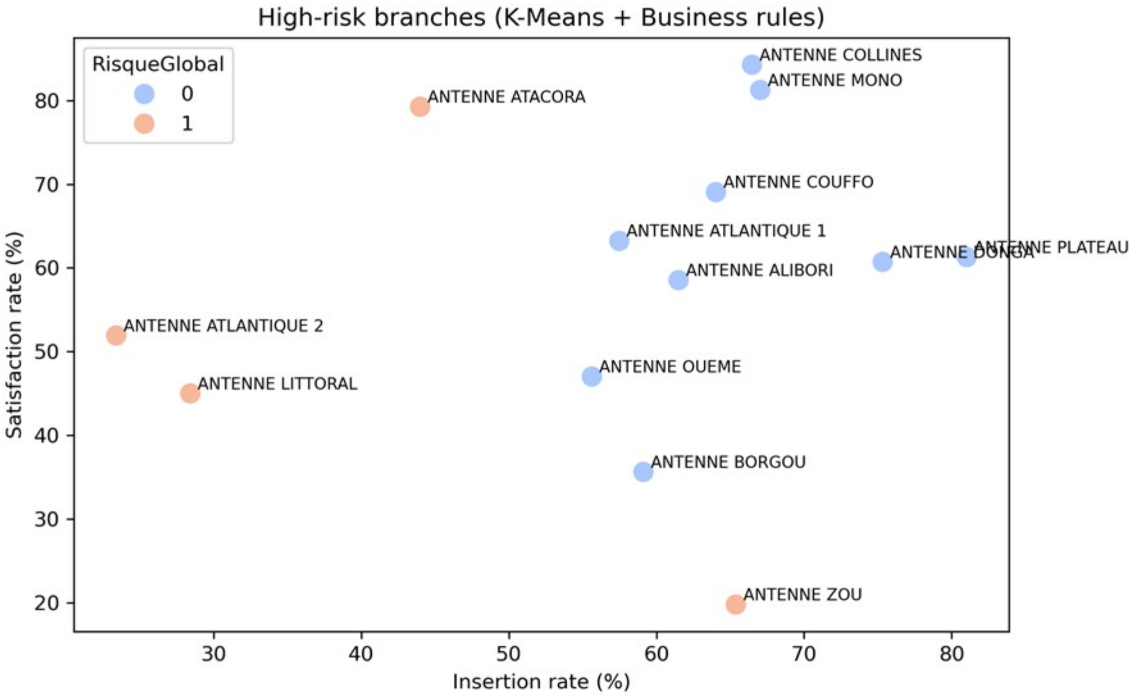

Figure 5: Branches classification using K-Means + Business rules

The cross-validation method used compares the clustering results with business rules and shows a concordance rate exceeding 80% in identifying branches at risk of underperformance. This process identifies priority risk structures or branches, and the result is directly and dynamically visible in the final dashboard for users.

4. Results and discussions

4.1. Results

Our system features an interactive dashboard organized into four complementary modules.

The main dashboard provides a consolidated view of key performance indicators, including registration volumes, placement rates segmented by profile, satisfaction indicators, and average processing times (Figure 6). Dynamic filters allow for comparative analysis by time and location.

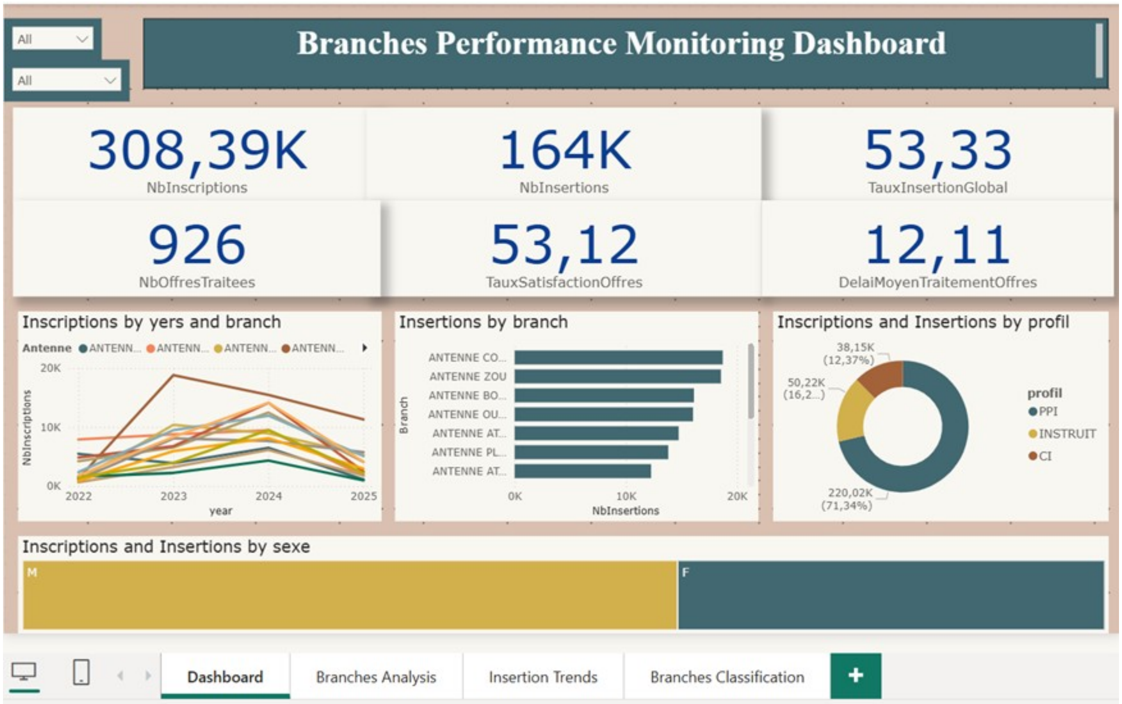

Figure 6: Monitoring dashboard

The model combines K-Means clustering and business rules, achieving an 84.6% agreement rate between the two approaches.

The segmentation identifies three distinct performance profiles:

- The lowest-performing profile is clearly defined : it includes the Littoral branch and the Atlantique 2 branch, both characterized by insertion rates below 30% and customer satisfaction below 55% (cluster 2).
- Conversely, branches such as Plateau, Donga, Alibori, Couffo, Atlantique 1, Mono, and Collines exhibit rates above 55% and constitute the good or very good performance profiles (cluster 1).
- Between these two extremes, some branches show average or contrasting performance (e.g., good insertion but low satisfaction), such as Ouémé, Borgou, and Zou (cluster 0). Business rules, based on critical thresholds (insertion rate <50%, satisfaction <30%, lead times >15 days), allow for operational risk assessment.

The following branches are considered at risk based on the combination of these two approaches :

Table 1: Ranking of detected banches according to risk score

Branches	**Insertion rate (%)**	**Satisfaction rate (%)**	**Processing time (days)**	**Risk score**
Branche Atlantique 2	23,40	51,93	24,20	57,56
Branche Littoral	28,42	45,00	14,28	55,15
Branche Zou	65,38	19,78	8,68	43,11
Branche Atacora	43,99	79,25	3,36	34,90

The results of the hybrid predictive model (K-Means + business rules) are developed in Python and executed directly in Power BI.
The model identifies branches potentially at risk of underperformance (Figure 7).
- In orange : antennas at high risk (according to K-Means + business rules);
- In blue: average, stable, or high-performing branches.

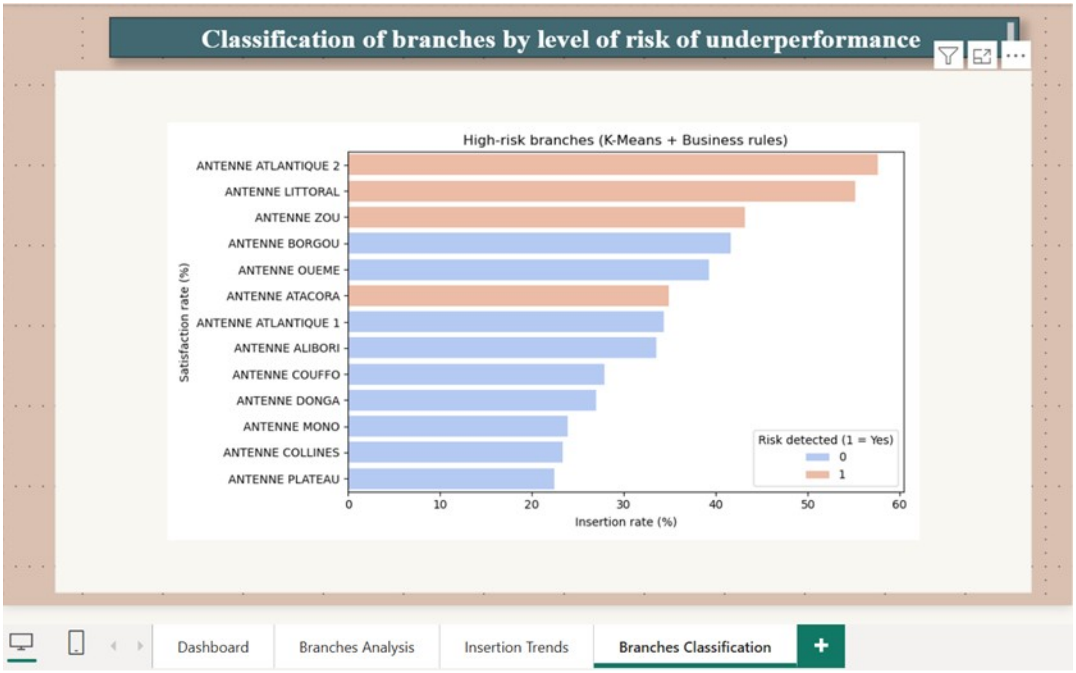

Figure 7: Classification of branches by risk level

The results show that:
- Littoral and Atlantique 2 are clearly at the greatest risk of underperformance, with combined weighted risk scores exceeding 50%.
- Zou and Atacora are in a fragile situation, requiring close monitoring.
- Borgou and Ouémé, although having higher risk scores than Atacora (due to the latter's very good performance in terms of offer satisfaction and processing time), are relatively average or at best out of danger, but require close monitoring.
- The other branches maintain a satisfactory balance between efficiency, satisfaction, and processing time.

The cross-validation matrix (Figure 8) confirms the robustness of the model, with 9 structures classified as high-performing and 2 as high-risk by both methods.
The borderline cases (2 structures) highlight the complementarity of the approaches: clustering detects homogeneous profiles, while business rules capture specific anomalies.

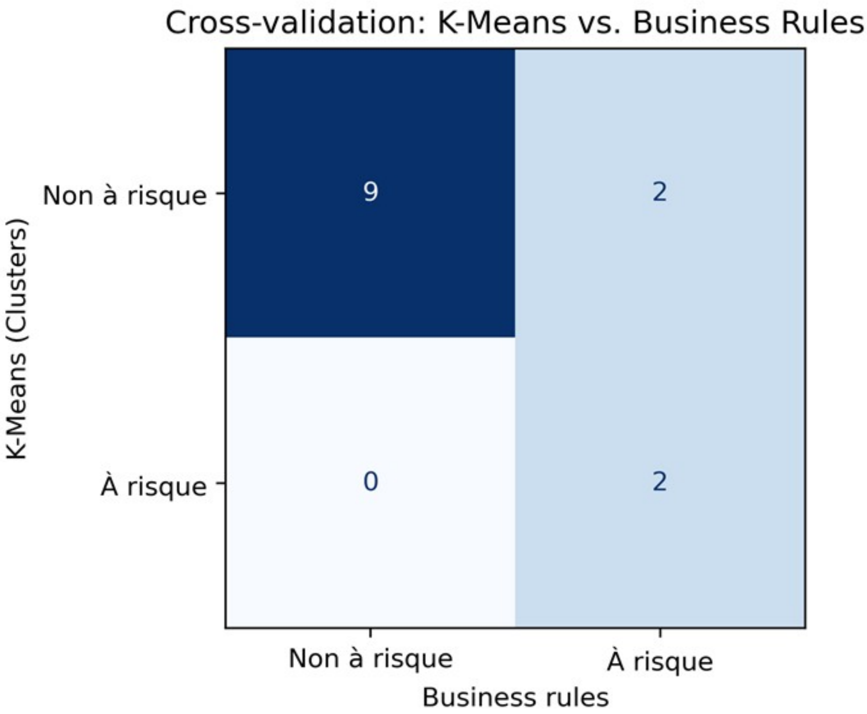

Figure 8: Cross-validation matrix

4.2. Discussions

Despite the satisfactory results obtained using the decision dashboard and the hybrid predictive model (K-Means + business rules), one limitation must be acknowledged:

The absence of contextual variables (geography, economy, demographics): the performance of a branch office does not depend solely on its internal processes.
External factors that could influence the results were not integrated into the model, including :

- The economic dynamism of the department (available job offers, presence of businesses).
- Local demographic characteristics (size of the working population, level of education).

These elements could also explain some of the observed differences between branches and thus limit the statistical interpretation of the model.

5. Conclusion and future work

As part of this research, we designed and implemented a decision-making system for performance monitoring and early detection of underperformance risks within organizations. Our approach integrates a data architecture combining a dimensional warehouse, an automated ETL process, and an interactive dashboard, thus offering a unified and automated solution for strategic management.

The results obtained demonstrate the effectiveness of the deployed system, with a hybrid detection model achieving an 84.6% concordance rate between algorithmic segmentation and business rules. However, some limitations remain, notably the limited integration of geographical contextual variables.

Future work could therefore explore enriching the models with socio-economic or demographic data, developing proactive alerting mechanisms, or implementing time-domain forecasting techniques (ARIMA). Our approach could also be extended to other sectors, thus offering a decision-making solution adaptable to diverse organizational needs.

Declaration on Generative AI

During the preparation of this work, the authors used chatGPT-5 mini for the following activities: language refinement, grammar corrections, and occasional structural suggestions. After using this tool, the authors reviewed and edited all generated content as needed and take full responsibility for the publication's scientific integrity and content.

References

[1] D. J. Cord, The Decline and Fall of Nokia. Schildts & Söderströms, 2014.

[2] C. Mui, « How Kodak Failed », Forbes, 2012, [En ligne]. Disponible sur: https://www.forbes.com/sites/chunkamui/2012/01/18/how-kodak-failed/

[3] M. Diallo, « La faillite d'Air Afrique : causes et conséquences ». 2002.

[4] S. Ogunyemi, « NIPOST Reform and Performance Challenges ». 2017.

[5] W. H. Inmon, Building the data warehouse. John wiley & sons, 2005.

[6] R. Kimball et M. Ross, The data warehouse toolkit: The definitive guide to dimensional modeling. John Wiley & Sons, 2013.

[7] M. J. Lebas, « Performance measurement and performance management », Int. J. Prod. Econ., vol. 41, no 1-3, p. 23-35, 1995.

[8] R. S. Kaplan et D. P. Norton, The Balanced Scorecard: Translating Strategy into Action. Harvard Business School Press, 1996.

[9] M. Golfarelli, S. Rizzi, et C. Pagliarani, Data warehouse design. McGraw-Hill Education, 2009.

[10] O. Osunrinde et M. Tiamiyu, « Information identification, evaluation and utilisation for decision-making by managers in South West Nigeria », South Afr. J. Inf. Manag., vol. 19, no 1, p. 1-8, 2017.

[11] W. A. of P. E. Services, Local Economic and Employment Development (LEED) The World of Public Employment Services Challenges, capacity and outlook for public employment services in the new world of work: Challenges, capacity and outlook for public employment services in the new world of work. OECD publishing, 2016.

[12] Q. A. Al-Radaideh et E. Al Nagi, « Using data mining techniques to build a classification model for predicting employees performance », Int. J. Adv. Comput. Sci. Appl., vol. 3, no 2, 2012.

[13] M. Koch-Rogge, « Assessing service performance: applying data envelopment analysis for evaluating employees' performance in the service industry », PhD Thesis, Anglia Ruskin Research Online (ARRO), 2023.

[14] F. Provost et T. Fawcett, Data Science for Business: What you need to know about data mining and data-analytic thinking. O'Reilly Media, Inc., 2013.

[15] H. J. Watson, « Tutorial: Big data analytics: Concepts, technologies, and applications », Commun. Assoc. Inf. Syst., vol. 34, no 1, p. 65, 2014.

[16] O. Okeke-Uzodike et P. Chitakunye, « Public sector performance management in Africa: Reforms, policies and strategies », Mediterr. J. Soc. Sci., vol. 5, no 26, p. 85-92, 2014.

[17] M. K. Hartley, « An analysis of business intelligence for improved public service delivery », Master's Thesis, University of Cape Town, 2015. [En ligne]. Disponible sur:

https://open.uct.ac.za/handle/11427/15534