Performance Indicators Development for Public Services using an AI-based clustering approach.

Victor de Almeida Xavier[a*], *Felipe* Maia Galvão França[b], *Priscila* Machado Vieira Lima[c,d], *Nelson* Maculan[e],

[a]PESC, COPPE, UFRJ, Brazil, victorx@cos.ufrj.br, 0000-0002-2281-6356.

[b]IT, University of Porto, Portugal, felipe@ieee.org, 0000-0002-8980-6208.

[c]PESC, COPPE, UFRJ, Brazil, priscilamvl@cos.ufrj.br, 0000-0002-8515-9904.

[d]AEP, NCE, UFRJ, Brazil, priscilamvl@cos.ufrj.br, 0000-0002-8515-9904.

[e]PESC, COPPE, UFRJ, Brazil, macualn@cos.ufrj.br, 0000-0002-3897-3356.

Submitted: 31 January 2025, Revised: 26 March 2025, Accepted: 21 April 2025, Published: 30 June 2025

Abstract. Despite the hype for AI-based generative applications, the original pitfalls in implementing AI-based technologies remain: data quality, readiness, and availability. AI results are tied to the quality of the data provided to the underlying algorithms. Inevitably, bad data produces bad results. The same happens for data readiness and availability: without correct preparation and access, even high-quality data cannot be used in AI-based solutions. To help decision-makers in the public sector, performance indicators applied to processes and, therefore, to public services can be used as tools to guide public project selection, allocation of resources, monitoring of results, etc. In this way, performance indicators can be used to standardize the evaluation of public policies, helping citizens and city administrators make more informed decisions regarding the environment they live in. In the present work, we propose a performance indicator development process (PIDP) based on AI clustering techniques, which can be used to gather key performance indicators (KPI) among available data from open data sources and consumed and/or produced by the core processes that implement a public service.

Keywords. Performance Indicators, KPI, AI, Clustering, ClusWiSARD, K-Means, DBSCAN, Hierarquical, DDDM
Poster, DOI: https://doi.org/10.59490/dgo.2025.1055

1. Introduction

Data-driven decision-making (sometimes abbreviated as DDDM) is the process of using data to inform your decision-making process and validate a course of action before committing to it (Stobierski, 2019). Despite its extensive use as a business management tool in the private sector, DDDM can and should be used to help decision makers in the public sector. The main goal of DDDM is to provide mechanisms for stakeholders in the public sector to make more informed decisions. For this reason, the "data" part in DDDM is crucial in the decision-making process. Low-quality data should certainly lead to poor information provided to those stakeholders. Data quality is not easily verifiable during the decision-making process.

Performance indicators (PIs) are one of the most commonly used tools to evaluate processes in terms of effectiveness (Fisher, 2019) and their use is ubiquitous in the public sector (Smith, 1999). Key Performance Indicators (KPIs) are the result of performance indicators development and is tightly correlated to data processing based on artificial intelligence methods and technologies, as pointed by Bilal and Oyedele, 2020. The capacity of a KPI correctly identify in which level a public policy or service stands is closely dependent on the capacity of each PI that composes it. And each identified PI relays on the data that was used to gauge it.

2. Performance Indicators Development Process (PIDP)

The Performance Indicators Development Process (PIDP), as defined in our previous work brought by Xavier et al., 2024, explores and analyzes processes-related data to look for candidates for performance indicators. The PIDP uses clustering techniques to group related data together. First, the data is downloaded from available public data sources. The samples with the standardized data are analyzed to search for performance indicators among the features, present as the name of the columns of the data set. The PIDP also deals with difficulties regarding the quality of the data being processed. The same process is presented by other techniques like data envelopment analysis (DEA), as pointed by Ferreira et al., 2023. The PIDP depicts the basic concepts of PIs implementation. For this aspect, it can be seen as a framework to achieve the PIs implementation's goals, as a stakeholder-centred process developed by Dransfield et al., 1999

Fig. 1 – Performance indicators development process (PIDP) general view

The PIDP workflow overview is synthesized in Figure 1.

2.1. Data sources selection

The first step in the performance indicators development process (PIDP) is to obtain reliable data used in the decision-making process. For example, if the city administrators are interested in green-house gas' (GHG) emissions impacts, the PIDP will download data from databases that contain emissions-related data. Thus, cities represent a minimal viable comparison unit in this work.

In this work, we will use a public database provided by CDP (Carbon Disclosure Project) that contains emissions-related data from 814 cities gathered in 2019.

2.2. Data exploration

The next step in the PIDP is to narrow the emissions-related data into data units. A data unit is an abstraction extracted from the data structure (answers field in the CDP database, for example) that can provide insights on candidates to performance indicators.

The data to be explored should be initially reduced to represent a context (period of time, origin of the data, focus of the study - cities). Thus, the main goal of this scope filtering is to filter the valuable data among the data available from the databases. The process initiates attempts to correct errors such as inconsistencies found in the available data: e.g. wrong data type, empty value in "selection" or "multi-selection" answer type, empty value in "not null" answer. If the error cannot be recovered using other data from the same record, the record is discarded.

2.3. Data pre-processing

The data preprocessing step is responsible for preparing the available data to be correctly used by the clustering algorithms and it requires the data to be cleaned from consistency errors. After an initial inspection of the data in the CDP database forms, inconsistencies and errors were found that could jeopardise the clustering process.

According to Shepperd et al., 2013, the best preprocessing strategy is that in which the problematic data should be treated first. Some cases of either conflicting feature values or implausible values should be discarded before data can be used. Therefore, it was necessary to build a support system to deal with these issues and leverage the quantitative and qualitative analysis steps.

In this work, we use **functional data filtering**, in which filtering parameters are passed to preprocessing execution module to segregate only the information needed in the context of a preprocessing configuration and optimise the drill down during quantitative and qualitative analyses, and **errors mitigation** techniques, which can try to mitigate the effects of typos and other simple errors in the next steps of PIDP.

2.4. Quantitative analysis

The quantitative analysis of the results obtained from the clustering methods can be used to indicate features with better chances to be used as performance indicators. Thus, the clustering results are treated and viewed as an alternative to purely statistical ones. However, the main goal is to search for similarities and answers that indicate different approaches implemented by the cities that are grouped in the same cluster. The nuances of the clustering process, the comparative data generated, and the validation techniques are shown in the following sections.

2.5. Qualitative analysis

The main objective of the qualitative analysis phase is to compare the responses of different cities present in the same cluster, indicating a convergent approach over the data. Another strategy is to compare different responses from cities in different clusters, indicating a divergent approach in this case. The following sections detail the techniques involved in the step of the process.

To achieve this goal, the PIDP uses two techniques: and adaptation of **Grounded Data Theory**, in which the samples are checked for inconsistencies, and **Case Study**, in which a sample can be used to downgrade a PI's candidate.

3. Results

3.1. PIs developed from Emissions Reporting Correlations

Analysing the emissions reporting data correlations are another way to look for performance indicators. After some experiments with broadly used indicators like GDP, HDI and their variants, and others like OECD, C40, GCoM and SCI memberships, some results indicate correlations between then and emissions reported by cities. The cities total emissions clusters distribution is shown in Figure 2.

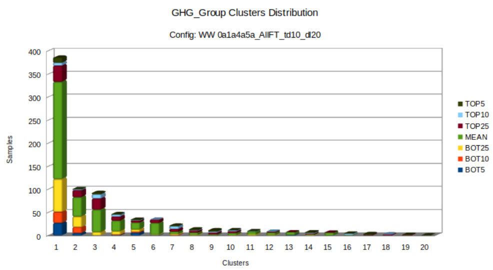

Fig. 2 – Cities total emissions in 2019 clusters distribution using configuration WW_0a1a4a5a_AllFT.

3.1.1. PI: SHDI x Cities Emissions

The correlations between emissions and human development index (HDI) can be seen in studies at the country level. Nevertheless, this work aims to use available sub-national human development index (SHDI) information to investigate the correlation between this index and cities total emissions.

3.1.2. PI: OECD x Cities Emissions

The organisation for Economic Co-operation and Development (OECD)OECDorganisation for Economic Co-operation and Development is a membership of developed countries that cooperate in economic matters. The cities in these countries follow the policies and directives negotiated in forums promoted by OECD. To some extent, part of these policies is applied as is by the cities.

3.1.3. PI: GCoM Membership

The Global Covenant of Mayors for Climate & Energy (GCoM) has more than ten thousand participant cities. It aims to monitor hazardous events related to climate and energy that occur inside cities' boundaries through mitigation plans and other additional information.

During the analysis, the prevalence between GCoM and CDP database was 73%, indicating that 219 cities reached by CDP partnership are still not part of GCoM by 2019. Considering the clusters with a prevalence ratio over 90%, clusters 5, 12, 16, 17, 18, 19, and 20 had 59 samples reprocessed.

However, the data obtained during the implementation of the plans are not uniform. A reason is that even though these cities were selected by representing overall performance, their methodology varies based on many aspects not covered by the questions set. An example of it is the question "4.3:Please give the name of the primary protocol; standard; or methodology ... city-wide GHG emissions".

3.1.4. PI: Smart Cities Index Membership

A consortium of international organisations maintains the Smart Cities Index (SCI) to rank cities in terms of developing connectivity to services provided by the cities. The list of cities in SCI in 2019 is available in the appendix. The cluster distribution of SCI members is shown in Figure 3. The SCI 2019 had 111 cities, and

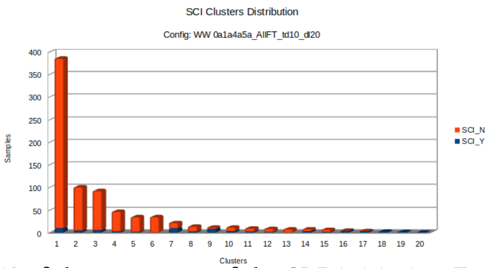

Fig. 3 – SCI clusters distribution using configuration WW_0a1a4a5a_AllFT.

63 of them are part of the CDP initiative. Even though it is not expected that all SCI cities have any direct goal related to emissions reduction, selecting the ones that are part of the CDP initiative introduces a bias that needs to be considered during the analysis. The clusters with SCI minimal prevalence were 1, 7, 9, 16, 18, 19, and 20, and the frequency of positive answers is similar to those of C40, although the execution of the actions is better. The distribution is shown in table 1 for comparison.

The distribution of the answers is a little more stable in this case. Still, 10% of the selected cities does not have a GHG emissions reduction target in place. In clusters 1 and 2, we have situations that represent errors during data processing or cities entered; the value "0" indicates that no information was provided. However, in clusters 4 and 7, there are some indications of "planning and not executing", confirming some cities have problems running the last mile and, for this, not getting the full benefits from emissions reduction initiatives.

The average of SHDI for cities in SCI is very high (879) and it is inclined to some uniformity in answers for at least 90% of them. It can be seen in distributions by clusters and by regions.

Tab. 1 – Selected samples clusters 1, 7, 9, 16, 18, 19 and 20 which are also in SCI database. The answers frequencies to the questions 1.0, 4.0 and 5.0 are shown.

Question	Positive Answers	Negative Answers
1.0: Does your city incorporate sustainability goals and targets (e.g. GHG reductions) into the master planning for the city?	93.65%	6.35%
4.0: Does your city have a city-wide emissions inventory to report	96.83%	3.17%
5.0: Do you have a GHG emissions reduction target in place at the city-wide level?	90.84%	9.52%

4. Conclusion and future works

According to the results of the experiments using PIDP and CDP, GCoM and SHDI databases, we can support the conclusion that it is possible to address the key performance indicators discovery through the development of performance indicators based on the analysis of the data present in those databases and the inter relations expressed in correlations found in them. Despite the correlations are weak in the first view, the qualitative analysis of the data showed problems regarding the quality of the data, mainly in the CDP database, in which data set were found both type of errors: typos and leniency, expressed in the form of the selection of the "first option" of the list box in some answers in which they are not suppose to exist.

To address this finding, one alternative is to develop an evaluation model of the quality of the data capable of predict idiosyncrasies in the fulfilment of the forms that are in the origin of the problematic data. In future works, we are aiming to develop these model, aggregating other AI-based technologies and a time series evaluation mechanisms.

Acknowledgement

- **Contributor Statement** based on Policies (tudelft.nl):
  Victor: Conceptualization, Data Curation, Formal analysis, Investigation, Methodology, Project administration, Software, Visualization, Writing – Original Draft, Writing - Review & Editing
  Priscila: Supervision, Validation
  Felipe: Validation
  Maculan: Supervision
- **Conflict Of Interest (COI)***: "There is no conflict of interest"

References

Bilal, M., & Oyedele, L. O. (2020). Big data with deep learning for benchmarking profitability performance in project tendering. *Journal of Expert Systems with Applications, vol 147, pp. 113194*.

Dransfield, S. B., Fisher, N. I., & Vogel, N. J. (1999). Using statistics and statistical thinking to improve organisational performance (with discussion). *Int. Statist. Rev., 67, 99–150*.

Ferreira, D. C., Figueira, J. R., Greco, S., & Marques, R. C. (2023). Data envelopment analysis models with imperfect knowledge of input and output values: An application to portuguese public hospitals. *Journal of Expert Systems with Applications, vol 231, pp. 120543*.

Fisher, N. (2019). A comprehensive approach to problems of performance measurement. *J. R. Stat. Soc. A, 182: 755-803. https://doi.org/10.1111/rssa.12424*.

Shepperd, M., Song, Q., Sun, Z., & Mair, C. (2013). Data quality: Some comments on the nasa software defect datasets. *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, vol. 39, n. 9, pp. 1208-1215*.

Smith, P. (1999). The use of performance indicators in the public sector. *Journal of the Royal Statistical Society Series A, Royal Statistical Society, vol. 153(1), pages 53-72*.

Stobierski, T. (2019). *The advantages of data-driven decision-making* (tech. rep.). Harvard Business School.

Xavier, V., França, F. M., & Lima, P. M. (2024). Emissions reporting maturity model: Supporting cities to leverage emissions-related processes through performance indicators and artificial intelligence. *RAIRO-Operations Research, v.58 ed.2 pp.1401-1428*.