



POLICY PAPER

# Unlocking the Potential of Zambia's Administrative Data

Rory Hardie, Shahrukh Wani,  
Torsten Figueiredo Walter  
and Benjamin Shawa

This paper outlines the value of administrative data for policymaking in Zambia. Based on interviews with over a dozen government ministries and a review of global best practices, it provides evidence on what constrains the use of administrative data in the Zambian government presently, and a pathway to overcoming these constraints.

FEBRUARY 2024



DIRECTED BY



FUNDED BY



Ministry of Finance and National Planning



---

# CONTENTS

---

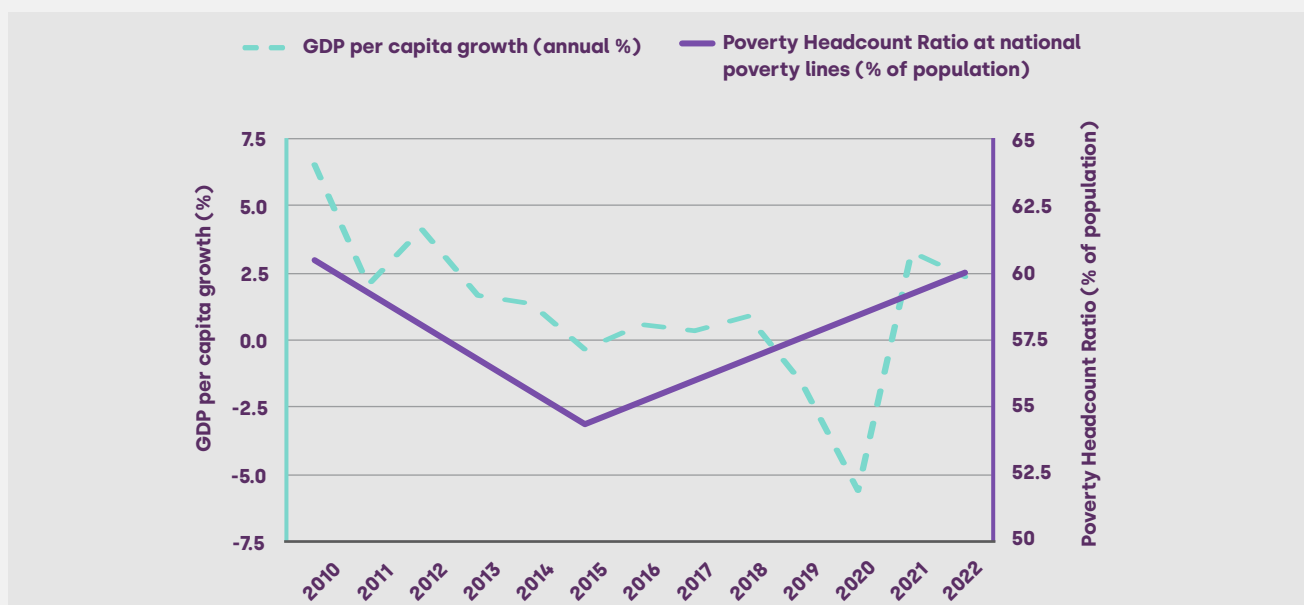
<b>Executive summary</b>	<b>2</b>
<b>1. Policy potential of administrative data</b>	<b>5</b>
1.1 What is administrative data?	6
1.2 How can administrative data be used in the policy lifecycle?	7
<b>2. Zambia's administrative data landscape</b>	<b>13</b>
2.1 Abundant administrative data	14
2.2 Emerging policy use	15
2.3 Untapped policy potential	17
<b>3. Access, quality, and capacity constraints</b>	<b>19</b>
3.1 Access	20
3.2 Quality	22
3.3 Capacity	24
<b>4. Policy agenda</b>	<b>27</b>
4.1 Leveraging research partnerships	28
4.2 Institutionalising data skills training	31
4.3 Formalising data sharing	32
4.4 Centralising administrative data storage	33
4.5 Digitising administrative data collection	36
<b>5. Conclusion</b>	<b>38</b>

---

## Executive summary

Over the past decade, Zambia has experienced sluggish economic growth. From 2010 to 2019, Zambia's average annual GDP per capita growth was only 1.56%.<sup>1</sup> In the same period, Zambia also struggled to progress in reducing the incidence of poverty. The proportion of Zambia's population below the national poverty line only decreased by 0.5 percentage points from 2010 to 2022.<sup>2</sup>

Figure 1: Zambia's Socioeconomic Trends



Sources: World Bank (2024) and Zambia Statistics Agency (2023)

Data could help Zambia's policy stakeholders make effective evidence-based decisions to improve the situation. For example, data can be used to target resources efficiently, monitor policies consistently, and evaluate policies accurately. Zambia's national policy recognises this: a core ambition of the 8th National Development Plan (8NDP) is to strengthen the National Statistical System (NSS) so policymakers can make well-informed decisions.<sup>3</sup> Doing so will involve improving the supply of, and policy demand for, data – from traditional data (such as surveys or censuses) to modern big data (such as satellite imagery or remote sensors).

One critical source of evidence which should not be overlooked in Zambia's NSS is administrative data – that is, the data collected as part of providing public services. **The high frequency and comprehensive coverage of administrative data make it an extremely informative resource in the policy lifecycle.** Administrative data is especially powerful when it complements – or is combined with – other data sources to generate precise policy insights.

1 World Bank. "GDP per capita growth (annual %) – Zambia." Accessed January 22, 2024. <https://data.worldbank.org/indicator/NY.GDP.PCAP.KD.ZG?end=2022&locations=ZM&start=2010>

2 Zambia Statistics Agency. *2022 Poverty Assessment in Zambia: Key Highlights*. 2023

3 Ministry of Finance and National Planning. *Eighth National Development Plan (2022-2026)*. 2022

Given the potential value of administrative data, three questions emerge for Zambia: To what extent is administrative data used for policymaking? What constrains the use of administrative data? How can these constraints be overcome?

The IGC collaborated with Zambia’s Ministry of Finance and National Planning to shed light on these questions. The study involved in-person interviews with fourteen government ministries and the submission of open-ended surveys by thirteen government ministries.<sup>4</sup> The findings are as follows:

### 1. To what extent is administrative data used for policymaking?

- ✓ **Abundant Data** – Zambia’s ministries collect and store an abundance of administrative data which could be informative for policymaking.
- ✗ **Underutilised Data** – Administrative data is not used consistently in the policy lifecycle and is typically limited to policy monitoring.

### 2. What constrains the use of administrative data?

- ✗ **Access** – Siloed data storage and inefficient data sharing limit data access, awareness, and security.
- ✗ **Quality** – Incomplete data, missing metadata, and non-standardised data linkages lower data quality.
- ✗ **Capacity** – Technical capacity gaps in supply and demand weaken data management and reduce data demand.

### 3. How can these constraints be overcome?

Policy Recommendation	Constraint(s) Targeted		
	Access	Quality	Capacity
<b>Leverage research partnerships</b> , that use data for policy insights, to identify how data can be used for policymaking and increase data demand.		✓	✓
<b>Institutionalize data skills training</b> to close skills gaps in the supply of administrative data and improve data management, visualisation, and analysis.		✓	✓
<b>Formalise inter-ministry data sharing</b> , using clear governance frameworks, to increase the efficiency of data exchange, while protecting data security.	✓		
<b>Centralise administrative data storage</b> within and across ministries to reduce the frictions on data access, integration, and analysis.	✓	✓	
<b>Digitise administrative data collection</b> to improve data access and quality, while eliminating the costs associated with manual digitisation.	✓	✓	

<sup>4</sup> See ‘Annex 1’ for a more detailed description of both the context and the method of this study.

The paper is structured as follows.

**Section 1** discusses the policy potential of administrative data, looking at regional examples where administrative data has been used to target, deliver, and evaluate policy.

**Section 2** provides an overview of Zambia's administrative data, describing the type of data that exists and how it is currently used in policymaking.

**Section 3** unpacks key constraints to the policy use of Zambia's administrative data.

**Section 4** offers a policy agenda outlining recommendations for the Zambian government's consideration.



GCShutterstock/E+ via Getty Images

## 1. Policy potential of administrative data

### KEY MESSAGES

- 1. Administrative data is the by-product** of an organisation's day-to-day operational activities, e.g., delivering a service.
- 2. Administrative data has beneficial properties** for policymaking, including high-frequency collection and universal coverage.
- 3. Administrative data can be a complementary** source of information to other data sources, such as surveys or censuses.
- 4. Administrative data can be informative** at each stage of the policy lifecycle: I. Design, II. Delivery, and III. Review

Administrative data can help policymakers make well-informed and targeted decisions. It can be especially valuable when it is combined with other data sources. This section explains what administrative data is and how it can be incorporated into each stage of the policy lifecycle.

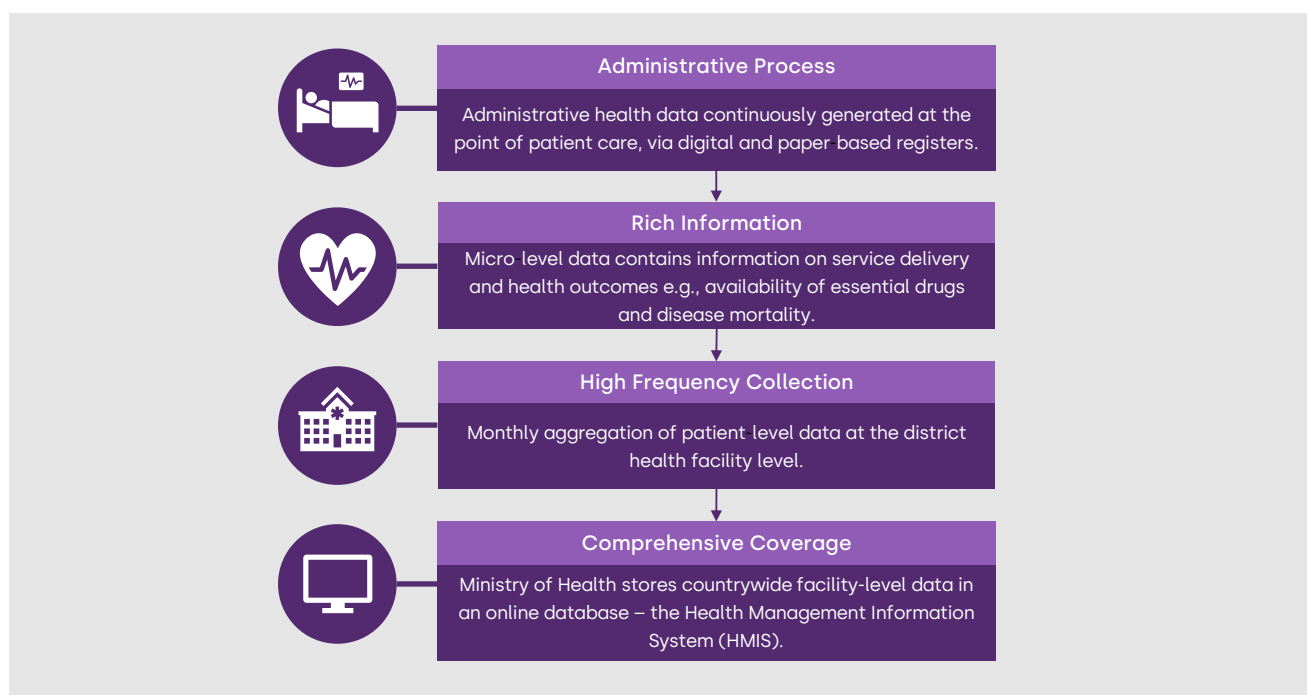
## 1.1 What is administrative data?

Unlike traditional data sources – such as surveys or censuses – administrative data is not originally collected for the purpose of generating statistics.<sup>5</sup> Instead, **administrative data is collected during the delivery of a service for an administrative objective, such as registration, transaction, or record keeping.**<sup>6</sup> Therefore, administrative data is a by-product of an organisation's day-to-day operational activities, and it typically has two defining features:<sup>7,8</sup>

- 1. High-Frequency Collection** – Administrative data is continuously collected, tracking activities or outcomes over time (i.e., the data has a panel structure).
- 2. Universal Coverage** – Administrative data usually comprehensively covers the target population. That means that the data includes all individuals interacting with the service.

**Figure 2** illustrates an example of how administrative data is generated.

**Figure 2: Zambia's Health Administrative Data**



Sources: Ministry of Health interview, Figueiredo Walter (2018)<sup>9</sup>, and Lee et al. (2023)<sup>10</sup>

5 OECD. "Short-Term Economic Statistics (STES) Administrative Data: Two Frameworks of Papers." *OECD Working Paper* (2012)

6 UK Research and Innovation (UKRI). "Key terms glossary," January 17, 2023. <https://www.ukri.org/councils/esrc/guidance-for-applicants/research-ethics-guidance/useful-resources/key-terms-glossary/>

7 Hand, David. "Statistical Challenges of Administrative and Transaction Data." *Journal of the Royal Statistical Society Series A: Statistics in Society* 181, no.3 (2018): 555–605

8 Bachas, Pierre and Anders Jensen. "Administrative data and methodologies for tax policy." *International Growth Centre* (2023)

9 Figueiredo Walter, Torsten. "The spatial distribution of health services in Zambia." *International Growth Centre* (2018)

10 Lee, Na-Mee, Douglas Singini, Craig Janes, Karen Grépin, and Jennifer Liu. "Identifying barriers to the production and use of routine health information in Western Province, Zambia." *Health and Planning* 38, no.9 (2023): 996-1005

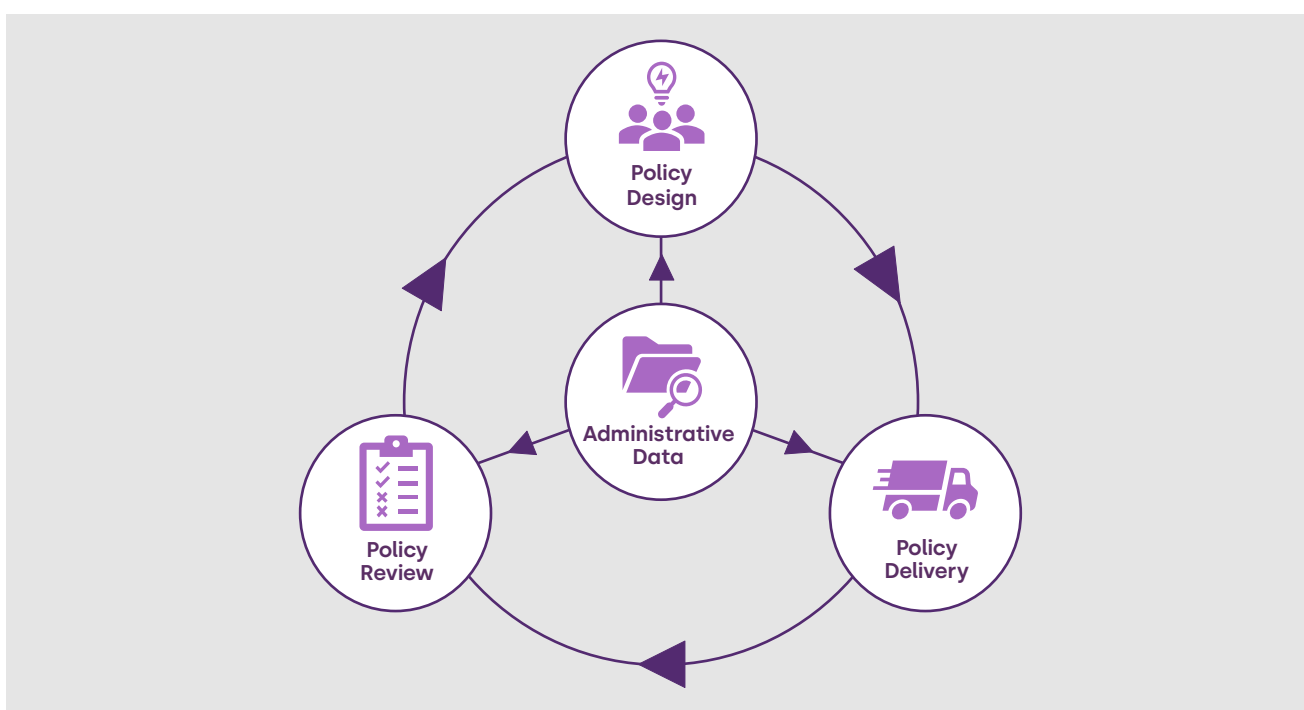


## 1.2 How can administrative data be used in the policy lifecycle?

The frequency and coverage of administrative data make it an attractive resource for researchers and policymakers who seek to produce high-quality and impactful policy insights.<sup>11</sup> For policymakers to maximise the value of existing administrative data – and avoid the costs of unnecessary additional data collection – it is important to understand how it can be used in the policy lifecycle.

There are various ways of framing the policy lifecycle; this report splits it into three phases: **policy design, policy delivery, and policy review.**<sup>12</sup>

**Figure 3: The Policy Lifecycle and Administrative Data**



Administrative data *alone* can only play a limited role in this lifecycle, and its usefulness cannot extend much beyond its original operational purpose. However, **when it is combined or when it complements other data sources, there are substantial gains to be made from employing administrative data** at each stage of this lifecycle.<sup>13</sup> As the case studies will demonstrate, **administrative data can be re-purposed in innovative ways** to close critical information gaps at each stage of the policy lifecycle.

11 Connelly, Roxanne, Christopher Playford, Vernon Gayle, and Chris Dibben. "The role of administrative data in the big data revolution in social science research." *Social Science Research* 59 (2016): 1–12

12 European Network for Rural Development, European Commission. "The role of monitoring and evaluation in the policy cycle." September 23, 2021. [https://ec.europa.eu/enrd/evaluation/back-basics/role-monitoring-and-evaluation-policy-cycle\\_en.html](https://ec.europa.eu/enrd/evaluation/back-basics/role-monitoring-and-evaluation-policy-cycle_en.html)

13 Groves, Robert and George Schoeffel. "Use of Administrative Records in Evidence-Based Policymaking." *The Annals of the American Academy of Political and Social Science* 678, no.1 (2018): 71-80

## What should the policy target?

### Stage 1: Policy design

Designing policies which most effectively address the needs of the target population involves establishing a comprehensive understanding of the socio-economic context.<sup>14</sup> To do this, it is necessary to conduct rigorous ex-ante (pre-policy) evaluations. These evaluations intend to reveal a) priority areas that the policy should target and b) the mechanism that is likely to be most effective.

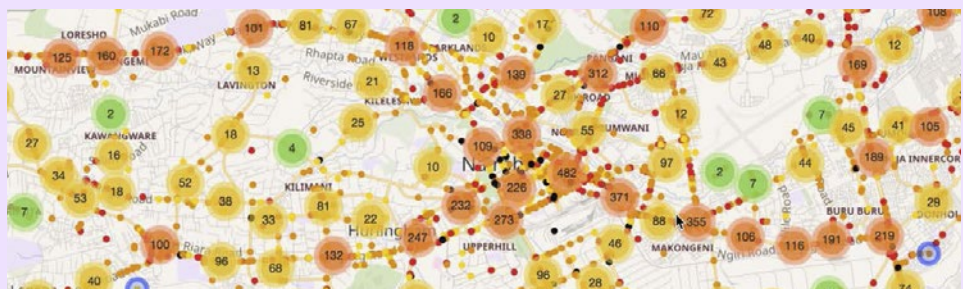
Given its universal coverage and tendency to include targeting data (e.g., disaggregated demographic information on age, gender, location, education, etc.<sup>15</sup>), administrative data can help identify where or to whom policies should be targeted. Once targets have been identified, complementary data sources, like surveys, which can collect specific information for the target group or area, can be used to establish which mechanism is likely to be most effective.

#### CASE STUDY 1: KENYA'S ROAD SAFETY

Improving road safety in Kenya is an urgent policy priority because road traffic accidents are one of the top five causes of death for individuals aged 5-70 and the top cause for boys aged 15-19.<sup>16</sup> To formulate policies that efficiently prioritise road infrastructure investment, Kenya's policymakers need access to detailed high-frequency crash data.

Kenya's National Police Service (NPS) worked with the World Bank's Development Impact Evaluation Group (DIME) to manually digitise 12,546 administrative crash records in Nairobi covering nine years.<sup>17</sup> These records contained the crash location, details, and severity information. This data was innovatively combined with crowdsourced data (e.g., from Twitter, AccuWeather, Google Maps, and Uber) to create the first georeferenced multiyear crash data set and accompanying map for Nairobi.<sup>18,19</sup>

Figure 4: Nairobi's Crash Map



Source: Bedoya, L. et al (2024)

- 14 European Network for Rural Development, European Commission (2021)
- 15 Goldilocks Deep Dive. "Using Administrative Data for Monitoring and Evaluation." *Innovations for Poverty Action* (2016)
- 16 World Health Organisation (WHO). "New initiative to tackle road crash deaths in Kenya." May 25, 2022. <https://shorturl.at/IEQV4>
- 17 Dimewiki. "Administrative and Monitoring Data." World Bank, August 14, 2023. [https://dimewiki.worldbank.org/Administrative\\_and\\_Monitoring\\_Data](https://dimewiki.worldbank.org/Administrative_and_Monitoring_Data)
- 18 Legovini, Arianna and Maria Jones. "Administrative Data in Research at the World Bank: The Case of Development Impact Evaluation (DIME)." In *Handbook on Using Administrative Data for Research and Evidence-Based Policy* edited by Shawn Cole, Iqbal Dhaliwal, Anja Sautmann and Lars Vilhuber. Cambridge (MA): Abdul Latif Jameel Poverty Action Lab, 2021
- 19 Bedoya, Lupe, Arianna Legovini, Sveta Milusheva and Rob Marty. "smarTTrans: Investing in analytics to save lives." World Bank. Accessed January 14, 2024. <https://thedocs.worldbank.org/en/doc/194791576169026323-0050022019/original/RobMartyieConnectWorkshopsmarTtrans.pdf>

This mapping revealed that 200 of the 1,400 crash sites across Nairobi contributed to more than half of all road traffic deaths. As a result of this analysis, Kenya's policymakers can target infrastructure investment and road traffic safety management to these crash hotspots – 150 kilometres of the total 6,200-kilometre road network.<sup>20</sup> Additionally, to understand the risk factors at the crash hotspots in greater detail, primary data was collected via surveys and videos on the road infrastructure at the 200 sites.<sup>21</sup>

*How will the policy be implemented and monitored?*

## Stage 2: Policy delivery

The second stage is policy implementation and monitoring progress towards the desired objectives. The consistent high-frequency collection of administrative data is very important for policy monitoring. Administrative data can help create and monitor statistical indicators of policy performance, which guide policy adjustments. Additionally, if the data is easily accessible, then it is sometimes possible to conduct **real-time policy analysis** – this is especially important in times of crisis, such as droughts or disease outbreaks.

### CASE STUDY 2: CAPE TOWN'S DROUGHT MANAGEMENT

The frequency of droughts, extreme temperatures, and severe storms is expected to dramatically increase as climate change accelerates.<sup>22</sup> In the face of such climate shocks, it is critical that policymakers have tools at their disposal which allow them to adapt and manage the potentially devastating impacts of these events.

In 2017-2018, an unprecedented one-in-400-year drought (which began in 2015) left the City of Cape Town (CCT) on the verge of running out of water completely – the dreaded 'Day Zero' scenario.<sup>23,24</sup> Fortunately, the innovative use of administrative data by CCT's policymakers made it possible to monitor water levels, target pipe maintenance, and reduce water consumption.<sup>25</sup>

Administrative data on dam storage was used to create a Water Dashboard that displayed weekly dam storage percentages, level changes, and average daily water production. This dashboard enabled policymakers to monitor fluctuations in water storage. Also, administrative data on water billing was used to create a Water Usage Map, which displayed household water consumption and neighbourhood

20 Dimewiki (2023)

21 Bedoya, L. et al (2024)

22 Cevik, Serhan and João Jalles. "Eye of the Storm: The Impact of Climate Shocks on Inflation and Growth. *IMF Working Paper* (2023)

23 Cape Town Drought Response Learning Initiative. "The Cape Town Drought Response Learning Initiative." Accessed December 18, 2023. <https://www.drought-response-learning-initiative.org/>

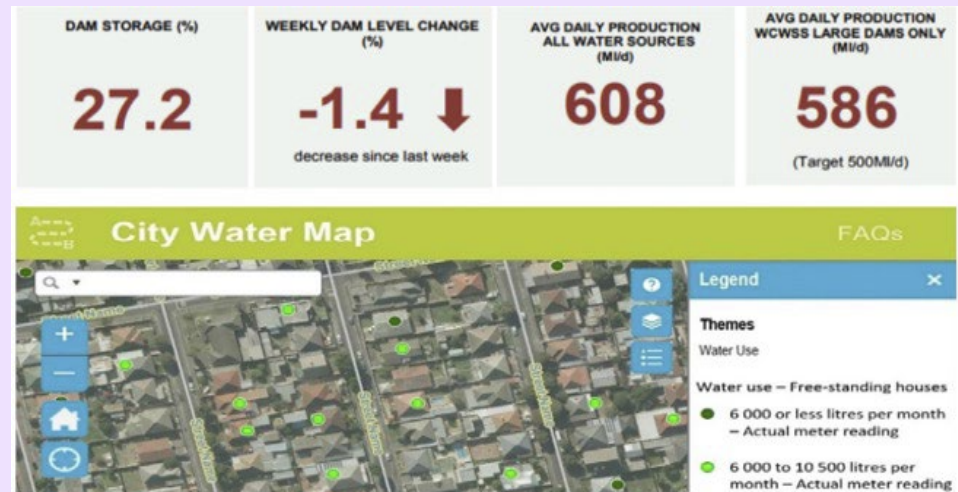
24 Hill-Lewis, Geordin. "Cape Town: Lessons from managing water scarcity." *Brookings*, March 22, 2023. <https://www.brookings.edu/articles/cape-town-lessons-from-managing-water-scarcity/>

25 Wright, Carol, Natasha Primo, Victoria Delbridge, and Kezia Fortuin. "Data and research as key enablers of city outcomes: A case study of the City of Cape Town (2000-2022)." *International Growth Centre*, (2023)

water usage. The map functioned such that households which saved a significant amount of water were given a green dot.<sup>26</sup>

This Water Usage Map leveraged social norms and the desire to conform to community expectations; by creating peer-to-peer monitoring, households were nudged into reducing their water consumption. Ultimately, CCT managed to decrease its daily water consumption by more than 50%, from 1.2 billion litres in February 2015 to 516 million litres in 2018.<sup>27</sup>

**Figure 5: City of Cape Town Water Dashboard**



Source: Wright, C. et al. (2023)

## What was the impact of the policy?

### Stage 3: Policy review

The final stage is policy review. Policymakers conduct ex-post evaluations to establish how well the policy performed against its initial objectives, if it had the intended long-run impact, and if the policy was good value for money. These evaluations are the key to grasping what works and what does not work – they form the building blocks of future policy decisions.

The rich and high-frequency nature of administrative data makes it extremely helpful in detecting subtle changes that result from policies. Having access to lots of data points on an individual over time and a large sample size makes it possible to identify small effect sizes and **accurately estimate the impact of a policy.**

<sup>26</sup> Wright, C. et al (2023)

<sup>27</sup> Wallace, Bryony. "Avoiding Day Zero: How Cape Town cut its water usage by 50% in three years." Oxfam, (2021)

### CASE STUDY 3: RWANDA'S VAT REVENUE

Many developing countries struggle to raise sufficient tax revenue to fund public services and infrastructure.<sup>28</sup> In 2021, the average tax-to-GDP ratio across 33 African countries was 15.6% – significantly lower than the OECD average of 34.1%.<sup>29</sup>

In Rwanda, the government prioritised improving value-added tax (VAT) collection because it is the largest contributor to their tax revenue.<sup>30</sup> In 2013, Rwanda introduced a law which requires VAT-registered firms to provide customers with a certified Electronic Billing Machine (EBM) receipt for each sale. In theory, tax compliance should increase because it enables the Rwanda Revenue Authority (RRA) to monitor the amount of VAT firms should pay; the RRA receives the sales transaction data in real time from the EBMs.<sup>31</sup>

IGC researchers worked with the RRA to evaluate the impact of EBMs on VAT revenue using the RRA's administrative data. Firstly, to assess the change in VAT payments after EBM adoption, they used the RRA's administrative data on all VAT payments from 2012-2014 – containing information on the payment type, date, amount, and the primary firm activity. Secondly, to distinguish whether a firm had adopted EBM in a specific quarter (separating the treated and untreated), they used detailed RRA administrative data on EBM registration (firm name, EBM activation data, location, etc).<sup>32</sup>

The results from the evaluation found that the introduction of EBMs increased VAT payments by 5.4% on average.<sup>33</sup> Additionally, EBMs were found to have the greatest impact on small firms in sectors where – before the policy was introduced – VAT compliance would likely have been lower than in other sectors. This policy review helps to inform future tax policy in Rwanda; it suggests that EBM policies should focus on sectors with lower initial compliance rates, where the tax revenue increase from EBM adoption would be significant.<sup>34</sup>

---

28 Bachas, P. et al (2023)

29 OECD. "Revenue Statistics in Africa 2023." (2023). <https://www.oecd.org/ctp/revenue-statistics-in-africa-2617653x.htm>

30 Ghirmai, Eva, Sarah Logan, and Sally Murray. "The incidence and impact of Electronic Billing Machines for VAT in Rwanda." *International Growth Centre*, (2016)

31 Ghirmai, E. et al (2016)

32 Eissa, Nada, and Andrew Zeitlin. "Incidence and Impact of Electronic Billing Machines for VAT in Rwanda." *International Growth Centre*, (2014)

33 Ghirmai, E et al (2016)

34 Eissa, N. et al (2014)

## Policy lifecycle: summary

Administrative data can play an essential role in all stages of the policy lifecycle – especially when it is combined with other data sources.

Importantly, the **methods of using administrative data to target, monitor, and evaluate policy are all interlinked** and build on one another:

- i. To evaluate policies comprehensively requires adequate monitoring.
- ii. Effectively targeting policies requires precise evaluations.

If the foundations to use administrative data for simple policy monitoring are weak (e.g., poor data quality, access, awareness, management, or linkages), then it becomes much more challenging to use that same data for more complex policy evaluations and targeted interventions in the future.

**Therefore, it is important to identify which administrative data is critical to the policy lifecycle and prioritise improving the usability of these data.**

- For example, some administrative data sources are exceptionally valuable for policymaking. As explored in an IGC paper by Bachas and Jensen (2023), **administrative tax data** has been used to generate insights on policy questions related to international trade, firm-level behaviour, inequality, and public finance.

Fundamentally, **administrative data has significant potential to add value to the policy lifecycle**. Maximising the value of existing administrative data means using it not only to monitor policy but also to evaluate and target decision-making. For this to be possible, it is necessary to match supply and demand – identify which administrative data can be used to provide insights on specific policy questions.



## 2. Zambia's administrative data landscape

### KEY MESSAGES

- Zambia's ministries collect and store an **abundance of administrative data** which could be valuable in policymaking.
- Zambia's ministries' policy use of administrative data is currently **limited relative to what it could be**.
- Zambia's ministries **share a desire for transformational change** in the use of data in policymaking.

Zambia's administrative data has significant policymaking potential. This section provides an overview of the administrative data landscape across Zambia's government ministries, detailing the type of administrative data collected and the way it is currently used in policymaking.

## 2.1 Abundant administrative data

The interviews and surveys reveal that **Zambia's government ministries are collecting and storing abundant administrative data**, which could be valuable in policymaking. As noted by ZamStats, the day-to-day operations of government ministries generate a significant quantity of administrative data.<sup>35</sup> The table below provides three examples from different ministries, which illustrate the type of data being collected:

**Figure 6: Zambia's Administrative Data Examples**

	Ministry of Labour and Social Security	Ministry of Finance and National Planning	Ministry of Transport and Logistics
Administrative Process	<b>Labour Inspection Forms (18c)</b> – filled out by labour officers/inspectors during routine company inspections	<b>Central Securities Depository (CSD) system</b> – automatically collects data on electronically held government securities (e.g., Treasury Bills and Bonds)	<b>Road Traffic Accident (RTA) reports</b> – Zambia police collect RTA data on paper at the accident site or when it is reported
Information	<b>Forms Data</b> – number of employees, types of contracts, wages, employee benefits, social security compliance, industrial relations, and safety/health conditions	<b>CSD Data</b> – centralised government securities registry with data on the ownership, transfer patterns, and overall structure of domestic debt	<b>RTA Data</b> – time of accident, location description, gender, vehicle type, and type of accident (e.g., fatalities, serious injuries, or car damage only)
Coverage & Frequency	Forms aggregated by provinces weekly and consolidated into a national report. Covers all the districts which the ministry has a presence in (64/116)	Continuously collected as and when there are changes in the flows of government securities	National coverage and police departments share the data on a monthly and quarterly basis with the Road Transport and Safety Agency (RTSA)
Storage	Forms filled out online and stored directly on the Labour Market Information System (LMIS), or filled out on paper and then transferred to LMIS	Data from the CSD system is exported to Excel, which enables further analysis by the Ministry of Finance and National Planning and the Bank of Zambia	Paper RTA data digitised by police departments and stored in an Excel offline hard drive by the RTSA

Given that some ministries have much larger administrative operations than others, the volume of administrative data varies substantially from one ministry to the next.

- For example, the newly created Ministry of Small Medium Enterprises Development does not currently store much data because the processes for SMEs (e.g., registration for government support programmes) are disaggregated by sector. While there is demand for an integrated platform for all SME data, right now, SME data is still

<sup>35</sup> Ministry of Finance and National Planning, and the Zambia Statistics Agency. *Second National Strategy for the Development of Statistics*. 2023



held separately by each of the sector representative ministries (e.g., Ministry of Community Development, Ministry of Agriculture, or Ministry of Tourism).

## 2.2 Emerging policy use

The second finding is that there are ministries that **use administrative data for some limited policy applications**. Currently, it seems that the ministries that use administrative data **primarily use** it to monitor policy.

- For example, the Ministry of Fisheries and Livestock uses geolocated administrative data on disease outbreaks to target information dissemination on disease management effectively.

The **use of administrative data at the formation and review stages of the policy lifecycle appears to be less common**. However, there are still some examples, two of which are outlined below.

### CASE STUDY 4: ZAMBIA'S CANCER TREATMENT UNITS

To deliver universal health coverage, a central component of Zambia's National Strategic Health Plan (2022-2026) is the decentralisation of health systems and services.<sup>36</sup> Non-communicable diseases accounted for 35% of deaths in Zambia in 2021.<sup>37</sup> A significant portion of these deaths can be attributed to cancer – which has a very high mortality rate in Zambia.<sup>38</sup> Therefore, the Ministry of Health (MoH) intends to roll out the decentralisation of cancer treatment.



Source: Lusaka Times (2023)<sup>39</sup>

Cancer can only be treated at the Cancer Diseases Hospital (CDH) in Lusaka. The MoH wants to build two more Cancer Treatment Centres and administrative data was critical in deciding their location. The MoH has access to the National Cancer Registry database, an administrative

36 Ministry of Health. 2022-2026 *National Health Strategic Plan*. 2022.

37 World Health Organization (WHO). "Country Disease Outlook: Zambia." August, 2023. <https://www.afro.who.int/sites/default/files/2023-08/Zambia.pdf>

38 Ministry of Health (2022)

39 Lusaka Times. "Cancer Treatment Centre to be constructed in Ndola." *Lusaka Times*, May 19, 2023. <https://www.lusakatimes.com/2023/05/19/cancer-treatment-centre-to-be-constructed-in-ndola/>

data set that records the hospital location where each cancer case is diagnosed (along with the age and type of cancer).

This data showed that, in 2020, the provinces with the second and third highest number of cases (after Lusaka) were the Copperbelt and Southern – accounting for 13% and 10% of all cancer cases, respectively. This trend was the same from 2011 to 2020. Within the two identified provinces, the districts Ndola (Copperbelt) and Livingstone (Southern) have university teaching hospitals with the necessary services to complement a cancer treatment centre. For example, these hospitals have theatres, pathology labs (to identify cancer), resident pathologists, blood products, imaging facilities, etc.

Additionally, these two districts are ideally positioned to receive referral cases from the country's northern (Ndola) and southern or western (Livingstone) regions. Therefore, combining administrative data on the density of cancer cases and information on the existence of complementary facilities led to the decision to construct the cancer treatment units in Ndola and Livingstone.

## **CASE STUDY 5: REALLOCATING ZAMBIA'S LABOUR OFFICERS**

According to the International Labour Organisation (ILO), a key obstacle to global employment is poor working conditions, including inadequate social protection, lack of formal contracts, and low labour standards.<sup>40</sup> Therefore, a core component of the 8th Sustainable Development Goal (SDG) is the attainment of 'decent work for all'.<sup>41</sup> Achieving this goal requires significant improvements in labour market conditions.

Zambia's Employment Code Act of 2019 outlines the laws which protect the welfare of labourers.<sup>42</sup> To ensure companies comply with these regulations, the Ministry of Labour and Social Security (MLSS) oversees labour inspections. Given finite resources, the impact of these inspections is maximised when labour officers and inspectors are deployed to the areas with the highest economic activity.

The MLSS combined administrative data on the number of labour inspections, including the number of officers/inspectors, in each

---

40 International Labour Organization. "Poor working conditions are main employment challenge." February 13, 2019. [https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS\\_670171/lang--en/index.htm](https://www.ilo.org/global/about-the-ilo/newsroom/news/WCMS_670171/lang--en/index.htm)

41 United Nations (UN). "The 17 Goals." Accessed January 20, 2024. <https://sdgs.un.org/goals>

42 Parliament of Zambia. *The Employment Code Act*, 2019. 2019

province with information from both the Labour Force Survey (number of employed persons in each province) and provincial GDP data:

**Figure 7: Redistributing Labour Officers**

Province	Population 2022	Formal Employment 2022	GDP 2021 (wmillion ZMW)	No of Labour Officers
Lusaka	3,079,964	1,086,114	139,453	21
Copperbelt	2,757,539	703,884	96,076	19
N/Western	1,270,028	131,924	67,918	7
Southern	2,381,728	336,559	37,054	14
Central	2,252,483	386,555	33,894	13
Eastern	2,454,788	159,865	21,232	10
Northern	1,618,412	133,037	12,725	5
Luapula	1,514,011	101,844	12,668	11
Western	1,363,520	153,613	11,384	18
Muchinga	918,296	79,729	10,960	6

The MLSS uses this data to plan the restructuring of the ministry in 2024 and the re-distribution of labour officers and inspectors to areas with higher levels of economic activity and employment. This helps with more efficient enforcement of the Employment Code and protects the welfare of more workers.

## 2.3 Untapped policy potential

While the case studies show that there are ministries that use administrative data, the policy use of Zambia's administrative data is still limited relative to its potential.

**Maximising the value of Zambia's administrative data requires ministries to use it more consistently and in ways that extend beyond policy monitoring.**

This does not mean using administrative data for monitoring is unimportant, but rather that there is room to improve how it is currently incorporated into Zambia's public policy lifecycle.

- For example, even on the monitoring side, the National Development Coordinating Committee's Cluster Advisory Groups noted that a significant quantity of administrative data could be used more innovatively to monitor and report on development progress.<sup>43</sup>

<sup>43</sup> Ministry of Finance and National Planning, and the Zambia Statistics Agency (2023)

The qualitative insights from the interviews with ministerial staff confirmed this idea – many noted that **while administrative data is occasionally used in policymaking, it is not frequent**. Considering the costs associated with data collection, it is in the interest of policymakers to ensure that existing administrative data is used to its full policy potential. As highlighted in the first section of this paper, when administrative data can be integrated with other data sources, it has the potential to provide crucial insights for policy evaluation and formation. Currently, Zambia’s ministries do not use administrative data this way.

Promisingly, the interviews and submissions for this paper reveal overwhelming **demand amongst ministries to develop the relationship between administrative data and policymaking**. In line with the 8th National Development Plan, ministries acknowledged the potential of administrative data and are keen to maximise its policy value. One ministry director described policymaking in the absence of such informative data as a task which is ‘no more productive than chasing the wind’.

As the proceeding section will explore, making transformative changes in the use of administrative data in Zambia’s public policymaking requires a detailed understanding of the current constraints on its policy use.



### 3. Access, quality, and capacity constraints

#### KEY MESSAGES

1. **Access** to Zambia's administrative data is challenging because of **siloed data storage** and **inefficient data sharing**.
2. **Quality** of Zambia's administrative data is lowered by **incomplete data**, **missing metadata**, and **non-standardised data linkages**.
3. **Technical capacity gaps** (in both supply and demand) weaken data management and reduce data demand.

This section outlines critical constraints to using administrative data in Zambia's public policy lifecycle. The constraints discussed are not exhaustive, but they represent the prevailing issues raised in the ministry interviews and submissions. Note that this study did not examine the quality of individual data sets.

## 3.1 Access

**X A significant obstacle to the use of Zambia's administrative data in policymaking is the lack of easy, timely, and consistent data access.**

### Siloed data storage

Given that the collection of administrative data by Zambia's government ministries can be paper-based, it is promising to see that **storage has largely been digitised**. For example, the Ministry of Fisheries and Livestock uses paper-based instruments to collect data on their Enhanced Smallholder Livestock Investment Project (E-SLIP), and this data is collated in Excel before being stored on the programme server. Digital storage makes administrative data more accessible for use in the policy lifecycle.

However, as noted in the interviews, the **manual data entry required to transfer paper-based records into a digital format significantly burdens the ministry statisticians** who are responsible for this process. It takes a significant amount of time, limiting the ability of statisticians to work on other essential data tasks, e.g., analysis or visualisation. This process can also undermine the reliability of the data because, under the pressure of time constraints, there is an increased risk of human error during manual digitisation.

Furthermore, some of Zambia's administrative data is **stored in isolated locations – on individual computer hard drives, spread across ministry departments, with limited backup**. Notably, there is significant variation across ministries, with some storing data in silos and others storing all their data in central online servers. Siloed data storage is worrying for multiple reasons:

- 1. Data Awareness** – Data visibility is reduced when it is stored in isolated locations, making it much harder for policymakers to be aware of all the data that is at their disposal.
- 2. Data Access** – Finding data is challenging, and having only one or two individuals in charge of supplying that data to all the relevant internal and external parties means data access is unlikely to be timely or even possible.
- 3. Data Security** – The risk of unauthorised access to sensitive information increases as there are unlikely to be consistent security standards to safeguard the privacy of the subjects in the data.
- 4. Data Longevity** – When individual data holders move to new roles, 'the data dies with them' (as one ministry stated), which means significant quantities of administrative data are lost.
- 5. Data Coordination** – Siloed storage, as one ministry noted, breaks down coordination, where departments and individuals are unsure who is responsible for collecting and storing specific administrative datasets.

## Inefficient data sharing

The survey submissions seem to suggest data sharing occurs between Zambia's government ministries. Using the Ministry of Fisheries and Livestock's submission, it is possible to map out the ministries and quasi-governmental institutions they exchange data with:

**Figure 8: Ministry of Fisheries and Livestock Data Sharing**



- For example, the Ministry of Fisheries and Livestock receives administrative data on the household beneficiaries of livestock stocking and re-stocking from the Ministry of Labour and Social Security.

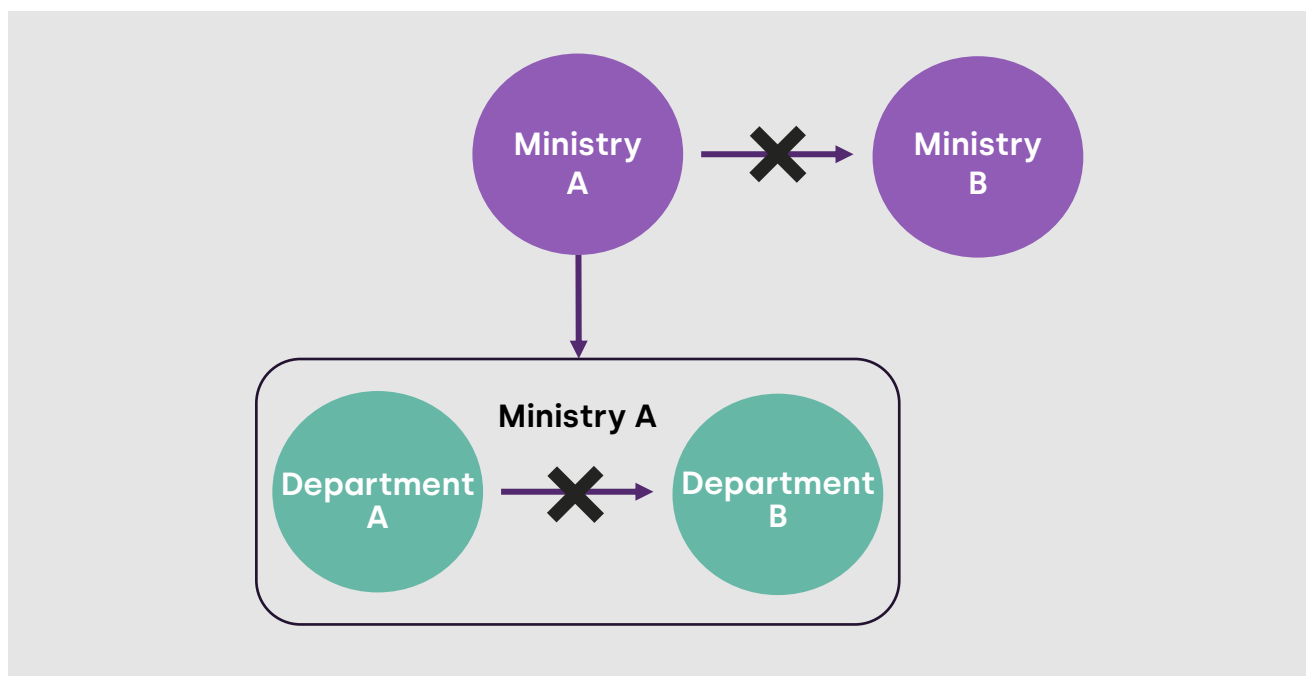
However, this top-level view needs to account for whether this data sharing is *efficient*. Insights from the interviews reveal that **data sharing can be slow, inconsistent, and ad-hoc**:

- 1. Outdated Processes** – Ministries continually send physical letters to request data, and often, these letters get caught up in the chain of command. This outdated and inefficient process means that the data is never shared or there is a costly lag in data exchange.
- 2. Unreliable Relationships** – Ministries noted that whether data is shared is often determined by the strength of inter-ministry relationships. One ministry indicated that, sometimes you go to an institution ten times and they still do not share the data.
- 3. Ad-hoc Sharing** – It appears data requests are sometimes ad hoc. This prevents consistent and periodic data exchanges between ministries.

Data sharing seems to be a time-consuming bargaining process, where ministries constantly go back and forth with data requests. When

combined with the insight that data storage can be siloed within ministries, it is evident that **accessing administrative data is constrained by inefficient intra-ministry and inter-ministry data sharing.**

**Figure 9: Inefficient Intra- and Inter-ministry Administrative Data Sharing**



### 3.2 Quality

**X** While this study did not assess the quality of individual datasets, the interviews and submissions reveal that incomplete data, missing metadata, and weak linkages limit the usability of existing data.

#### Incomplete data

There are two ways administrative data can be incomplete: a) specific fields are missing, and b) sections of the population are missing.<sup>44</sup>

Both issues seem to be present in Zambia's administrative data:

- 1. Incomplete Fields** – There are errors in data inputting where fields are missed, and the variables are not captured. Ministries acknowledged that the individuals in charge of inputting data (often at the district or provincial level) sometimes do not have the necessary training to conduct this process reliably.

<sup>44</sup> Hand, D. (2018)



- 2. Incomplete Population** – Ministries acknowledged that administrative data fails to capture the informal sector. For example, the Ministry of Labour and Social Security noted that it is challenging to gauge employment because informal businesses are not registered accurately.

These issues with the completeness of some of Zambia's administrative data constrain the ability of users to produce accurate data-driven policy insights. If data inputting for specific fields is inconsistent, then it becomes challenging to use administrative data to analyse policies over time. Equally, if significant sections of the population are missing from the data, then any policy analysis that results from the data could be misleading.

## Missing metadata

Metadata describes the specific administrative dataset (or any other data source). Typically, documentation on the metadata in a particular dataset includes, at a *minimum*, information on how the data was/is collected, when the data was/is collected, where the data was/is collected, and definitions of the individual variables.<sup>45</sup>

- For example, for administrative data on labour inspections, a metadata repository or data dictionary would be expected to contain information on how firms are selected for inspection, what questions are asked on the form, and how any of these practices have changed over time.

Unfortunately, it appears Zambia's government ministries do not systematically store metadata or have adequate data dictionaries. Often, the documentation that describes how the data is collected is missing or extremely challenging to access. The storage of metadata appears to be inconsistent and unreliable – sometimes, the only metadata available is an individual's recollection of the data generation process.

Access to this information is crucial for effectively using administrative data in the policy lifecycle. For instance, administrative data definitions are subject to change over time. If the definition of the variables contained in an administrative data change, it can invalidate comparisons of data points over time.<sup>46</sup> Therefore, **missing metadata creates uncertainty, which can constrain the use of administrative data in policymaking.**

## Non-standardised data linkages

As previously highlighted, **to maximise the policy value of existing administrative data, combining it with other data sources must be possible.** Integrating different data sources with one another requires strong linkages across the datasets – this facilitates the accurate matching of individuals, or units, from one dataset to the next.

It has been noted that **concepts and definitions are not standardised across Zambia's administrative data.** Many sectors do not follow international

---

<sup>45</sup> University of North Carolina. "Metadata for Data Management: A Tutorial: Definition." Accessed January 28, 2024. <https://guides.lib.unc.edu/metadata/importance>

<sup>46</sup> Hand, D. (2018)

standards and methodologies when compiling data.<sup>47</sup> This makes it challenging to merge administrative data with other data sources.

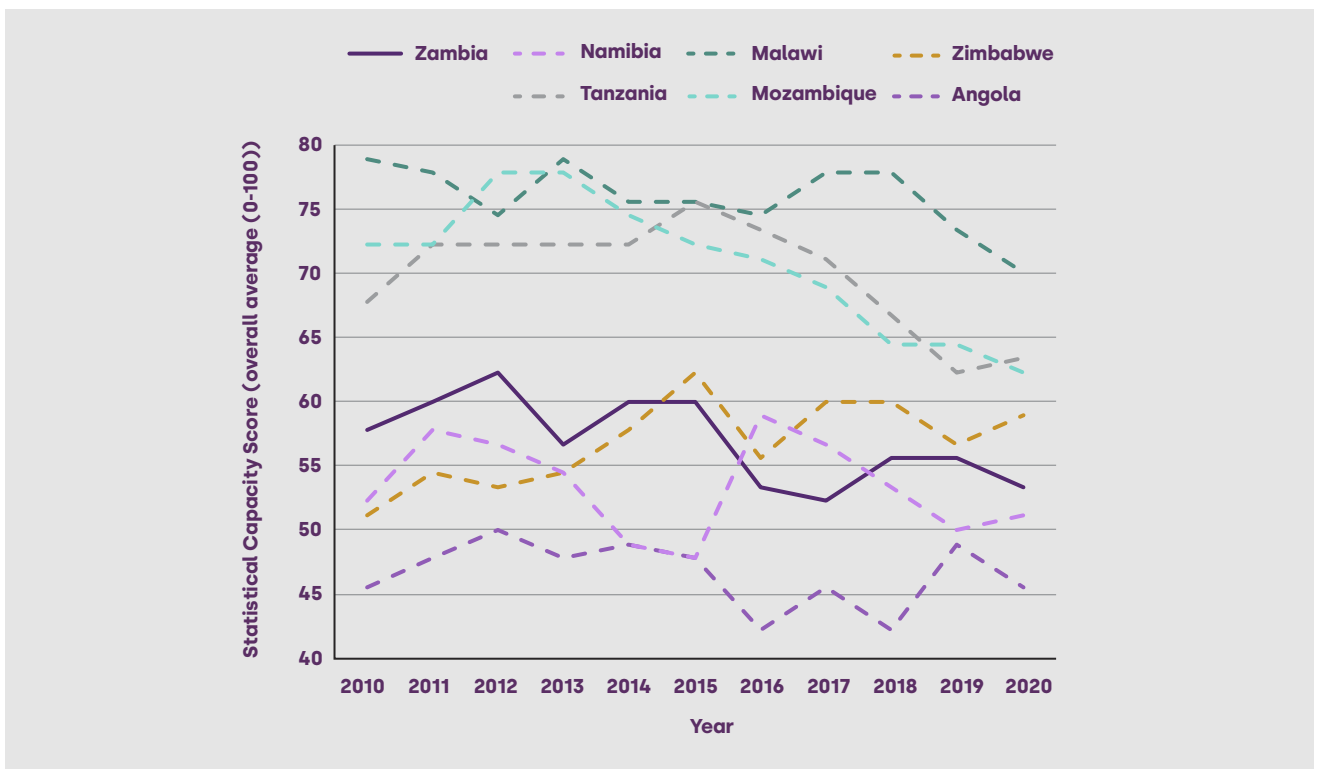
- For example, using data to evaluate policies related to public infrastructure relies on having unique identifiers for each public facility. Combining the data for policy analysis becomes exceptionally challenging if there are inconsistent identifiers across the different data sources (e.g., administrative data and survey data).

### 3.3 Capacity

#### X Technical capacity gaps (in demand and supply) weaken data management and reduce data demand.

The World Bank's annual statistical capacity score should reflect a country's capacity to collect, analyse, and disseminate high-quality information about its population and economy.<sup>48</sup> Figure 11 illustrates the annual changes in the statistical capacity score for Zambia and six of the neighbouring countries from 2010 to 2020:

**Figure 10: Zambia's Statistical Capacity Score (compared with neighbouring countries)**



Source: World Bank (2024)

47 Ministry of Finance and National Planning, and the Zambia Statistics Agency (2023)

48 World Bank. "Statistical Capacity Score (Overall Average) (scale 0-100)." Accessed January 20, 2024. <https://databank.worldbank.org/metadataglossary/world-development-indicators/series/IQ.SCI.OVRL>

These estimates show that **Zambia's statistical capacity declined 7.7% from 2010 to 2020**. This trend of capacity decline and stagnation is relatively consistent across Zambia's neighbours. However, Zambia still seems to lag many of its regional counterparts during this period, such as Malawi, Tanzania, and Mozambique. The capacity deterioration has been attributed to a failure to mainstream statistics into the policymaking process and the resulting lack of funding for statistics.<sup>49</sup>

## Technical capacity gaps (supply side)

**Ministries vary in their capacity to handle data depending on the data burden of the individual ministry.** For example, the Ministry of Small Medium Enterprises currently has only one senior statistician in charge of managing the data they store and receive. In contrast, the Ministry of Health has thirteen national monitoring and evaluation officers that maintain the health information system and coordinate the use of administrative data in evidence-based decision-making.

While staffing may vary, almost all the ministries noted that they have **inadequate capacity to manage the data they collect:**

- 1. Insufficient Statistics Staff** – Only a few statisticians within each ministry work with the administrative data. They have limited time to get through all the data requests while managing, visualising, and analysing the data effectively.
- 2. Inadequate Data Management Skills** – While some ministries have statistical departments well equipped for cleaning, visualising, and analysing data (with training in software such as R, Python, and MatLab), many acknowledge a lack of basic data management skills.

## Technical capacity gaps (demand side)

**Ministry staff who influence and implement policy can use administrative data in policymaking. These individuals form policy questions and can work with statisticians to answer them using available data.**

There seems to be a lack of policy stakeholders within ministries who can identify how administrative data can be used and combined with other data sources to inform decisions at each stage of the policy lifecycle. Ministries noted that more capacity is needed to interpret and apply data to policy settings. Ultimately, **the disconnect between policymakers and data reduces the policy demand for administrative data.**

- For example, one ministry stated that people 'forget' that administrative data can be used to produce data-driven policy outputs. The government has noted that Zambia's ministries have neglected this last stage of the data value chain – the use of data for policy impact.<sup>50</sup>

---

<sup>49</sup> Ministry of Finance and National Planning, and the Zambia Statistics Agency (2023)

<sup>50</sup> Ministry of Finance and National Planning, and the Zambia Statistics Agency (2023)



**“Improving the supply of data will have a limited impact on policymaking in the absence of sufficient policy demand for administrative data.”**



## 4. Policy agenda

### POLICY OPTIONS

- 1. Leverage research partnerships**, that use data for policy insights, to identify how data can be used in policymaking and increase data demand.
- 2. Institutionalise data skills training** to close skills gaps in the supply of administrative data and improve data management, visualisation, and analysis.
- 3. Formalise inter-ministry data sharing**, using clear governance frameworks, to increase the efficiency of data sharing while protecting data security.
- 4. Centralise administrative data storage** within and across ministries to reduce the friction on data access, integration, and analysis.
- 5. Digitise administrative data collection** to improve data access and quality while eliminating the costs associated with manual digitisation.

Overcoming the constraints on using administrative data in policymaking will require coordinated action from Zambia's government ministries, quasi-government agencies, and the Zambia Statistics Agency. This section identifies five policy options, which, when combined, could help to bridge the gap between Zambia's administrative data and the public policy lifecycle

## 4.1 Leveraging research partnerships

- ✓ **Leveraging research partnerships could increase data demand by improving the ability of policy stakeholders to identify how existing administrative data can be used in the policy lifecycle.**

To make long-lasting changes in the relationship between administrative data and policymaking, the government should prioritise increasing the policy demand for Zambia's administrative data. Transforming the supply of administrative data will be much easier and more impactful if a constituency of demand exists for the data.

**Research partnerships provide an ideal opportunity to build the capacity of Zambia's policy stakeholders** to link existing data with specific policy questions. The combination of data needed is different for each policy question. These partnerships can show policy stakeholders how administrative data can complement, or be combined with, other data sources to produce policy-relevant data-driven insights.

Ad-hoc research partnerships often involve minimal collaboration during the analytical process:

1. Policymakers ask researchers specific policy questions.
2. Researchers request and combine various relevant data sources.
3. Researchers produce targeted data-driven insights.
4. Researchers share their final insights with policymakers.

In the short run, such research partnerships can be important for policymaking because policymakers can use the targeted insights to adjust their future decisions. Equally, in the long run, these research partnerships can increase the policy demand for data *if* the research methods and insights are disseminated effectively with key stakeholders.

However, when these partnerships are structured in ways that involve greater collaboration during the analytical process, they are more likely to have a long-term impact on the demand for data. Through peer-to-peer learning, policy stakeholders can develop their understanding of how to maximise the value of administrative data in policymaking.

- There could be value in having a **centralised institution in Zambia where researchers work closely with policymakers to create custom-made data products, which can be used to produce insights on specific policy questions.**

In 2018, the OECD ran a workshop that was centred around developing policymakers' skills in evidence-informed policymaking (EIPM). The skills which were mapped out are the clusters of skills which Zambia's policy stakeholders could develop by leveraging research partnerships<sup>51</sup>:

---

51 OECD. "Skills for Policymakers for Evidence-Informed Policy Making (EIPM)." April 9, 2018. <https://www.oecd.org/gov/evidence-informed-policy-making-agenda-2018.pdf>

1. **Understand** – ability to understand the role of evidence in the policymaking process and differentiate between different types of evidence.
2. **Interrogate/Consider** – ability to critically assess the quality of evidence and understand the fundamental principles of the research method.
3. **Engage** – ability to communicate evidence effectively, engaging with both citizens and the evidence producers.
4. **Obtain** – the ability to identify available evidence, collect information from various sources, and frame questions.
5. **Apply** – ability to identify uses for evidence, how different sources can be combined, and develop plans for how change could occur.
6. **Evaluate** – ability to evaluate the impact of an intervention.

Developing these skills by capitalising on research partnerships could create a sustainable demand for Zambia's administrative data.

## CASE STUDY 6: ZAMBIA'S RURAL ELECTRIFICATION

The IGC is collaborating with government ministries to understand the effect of rural electrification on public infrastructure across Zambia. Studying rural electrification in low-income countries has historically been hampered by the lack of high-quality geo-identified data at the national level. Such data is crucial for evaluating rural electrification because it enables the accurate tracking of the progress and expansion of the electric grid infrastructure.

In the case of Zambia, the Rural Electrification Authority (REA), which sits under the Ministry of Energy, is coordinating the electrification of rural Zambia in the pursuit of universal access to electricity. REA's rollout of electrification has been done using the Rural Electrification Masterplan (REMP) since 2009. A key requirement of the original REMP was that public infrastructure, such as schools and health facilities, get connected in every village. There is a lack of historical records of what has been built and when, so administrative data from government ministries has helped to fill this void.

Figueiredo Walter and Moneke (2024)<sup>52</sup> showcase how administrative data from the Ministry of General Education and the Ministry of Health can be used to track the arrival of the electric grid across schools and health facilities throughout rural Zambia. In particular, the researchers accessed annual data from the Education Management Information System (EMIS) on schools' electrification status. Likewise, the Ministry of Health stores similar information on rural health facilities in the Health Management Information System (HMIS).

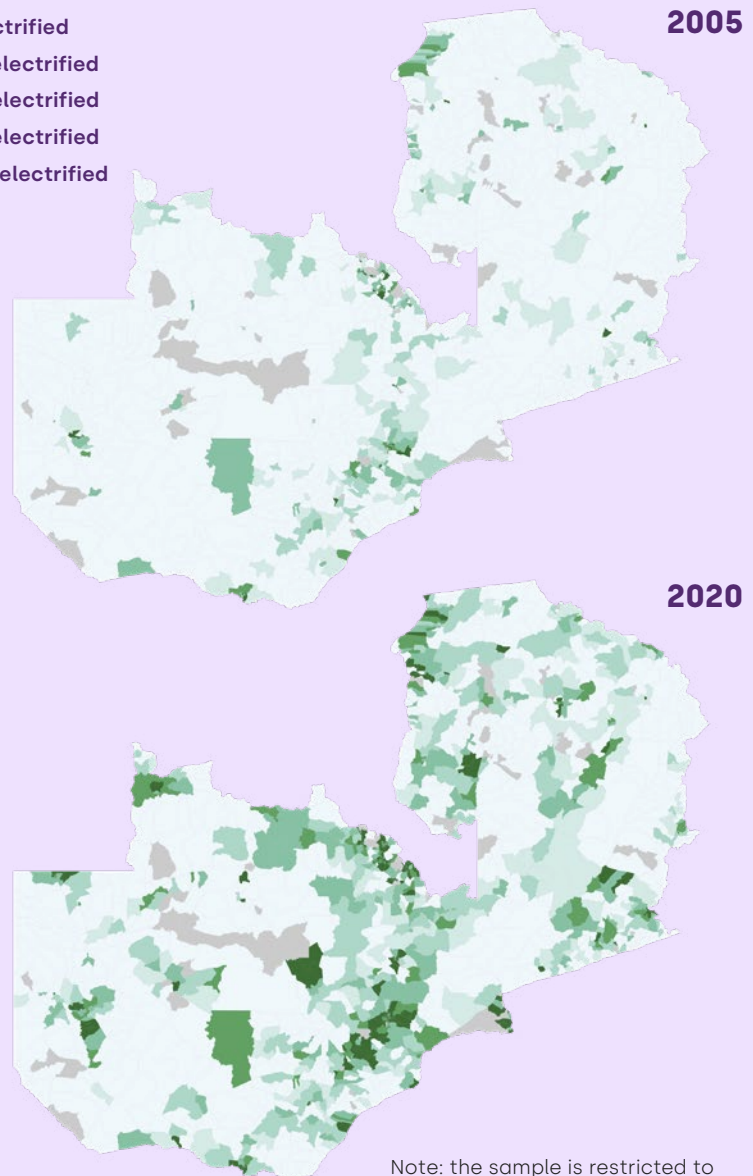
By geo-identifying schools and health facilities and mapping these to rural growth centres (the target of REA's electrification rollout), a novel method of tracking rural electrification was developed.

---

52 Figueiredo Walter, Torsten and Niclas Moneke. *Unpublished manuscript*. (2024)

**Figure 11: Electrification Percentage by Ward – Derived from EMIS School Electrification Status**

- No electrification
- 0-20% electrified
- 20%-40% electrified
- 40%-60% electrified
- 60%-80% electrified
- 80%-100% electrified



Note: the sample is restricted to schools that existed from 2005 to 2020.

This combination of administrative data sources allowed the researchers to work closely with government ministries and analyse the effects of rural electrification in the electrified communities once the grid had arrived – especially when combined with secondary data on social and economic indicators from ZamStats.

Findings from this research highlight novel insights about the specific circumstances under which rural electrification improves rural outcomes in Zambia. REA is working closely with the IGC researchers throughout this process to develop sustainable in-house capacity and fill technical skills gaps that will allow REA to derive similar analyses for future planning and operations.



## 4.2 Institutionalising data skills training

✓ **Consistent data skills training could close skills gaps in the supply of administrative data and improve data management, visualisation, and analysis.**

**Ministerial staff handling data – especially statisticians – could benefit from continually updating or refreshing their skills.** While one-off training days or workshops are essential methods of plugging skills gaps in the short run, they are not sustainable methods of maintaining the necessary data skills across government ministries.<sup>53</sup>

For data skills training to be effective and cost-efficient, it should be systematic:

1. Conduct thorough **skills assessments** across the ministries (e.g., assess data management, visualisation, and analysis skills).
2. Identify individuals who either a) urgently **need training** or b) have the requisite skills to **lead training**.
3. Deploy these second individuals as **'multipliers of knowledge'**<sup>54</sup> to help run the skills training.
4. Ensure data training is implemented consistently and evolves with modern data practices.

Private sector actors can support this training, but it is much more cost-effective and sustainable to have it driven by ministerial staff who are already capable of training – capitalising **on existing government capacity**. Consistent collective training of this nature can also help to foster **better data and policy linkages between ministries**.

Promisingly, a vital component of the Second National Strategy for the Development of Statistics (NSDS2) is empowering people through training.<sup>55</sup> ZamStats is the institution in charge of delivering a three-year rolling Statistics Training Plan, reviewed annually based on a skills assessment. As part of this plan, ZamStats intends to develop and implement a policy for ongoing in-service training which develops statistical capacity and skills. Importantly, this will include computer, data analysis, and dissemination skills training – including creating a data analysis unit that trains statisticians in SPSS, Stata, Excel, and CPro.<sup>56</sup>

---

53 PARIS21, and the Mo Ibrahim Foundation (MIF). "Working Paper: Bridging the Data-Policy Gap in Africa." *PARIS 21 and MIF Working Paper*, (2021)

54 PARIS21, and the Mo Ibrahim Foundation (MIF) (2021)

55 Ministry of Finance and National Planning, and the Zambia Statistics Agency (2023)

56 Zambia Statistics Agency. *Sector Strategy Plan for Statistics (SSP)*, 2022-2026. 2023. 2023.

## 4.3 Formalising data sharing

- ✓ **Formalising inter-ministry data sharing, using clear governance frameworks, could increase the efficiency of data exchange while protecting data security.**

For policy stakeholders to access unique combinations of data, which enable them to answer specific policy questions, data sharing needs to be efficient. Data sharing should aim to strike the optimal balance between free data flows and data protection<sup>57</sup>; current data sharing relationships appear to achieve neither – **administrative data is not flowing freely, and its security is not guaranteed.**

A first step towards improving data sharing is increasing awareness of what data sets exist and where they are held – one ministry noted that such awareness is vital to trust in sharing relationships. While the interviews and submissions reveal some of the existing administrative data, this mapping is incomplete.

To close the knowledge gap, it could be beneficial to generate a comprehensive map of all the administrative data sets collected by each ministry. This map, or register, of administrative data sets could be used by ministries to identify which datasets (collected by other ministries) would add value to their policy lifecycle.

A second step is promoting the use of long-term formal data-sharing agreements (e.g., a Memorandum of Understanding), which set out clear expectations and increase the predictability of data sharing. Using the framework outlined by the UN, these sharing agreements should specify<sup>58</sup>:

- **Parties/Focal Points** – ministries involved and the key personnel within each ministry in charge of the data exchange.
- **Purpose** – rationale for the data sharing and the intended use of the data.
- **Data** – describe the data content (variables, metadata, etc.), data format (e.g., Excel, CSV, etc.), and the means of transmission (e.g., web interface, email, etc.).
- **Duration/Periodicity** – length of data sharing agreement and how often new data should be shared (e.g., monthly, annually, etc.).
- **Obligations** – expectations for each party, specifying the personnel who will fulfil the agreed actions.
- **Legal Basis** – refer to sections of the Zambia Statistics Act (2018), Zambia Data Protection Act (2021), and other legal documents determining data sharing.
- **Confidentiality** – when microdata is involved (often the case with

---

57 OECD. "Mapping Approaches to Data and Data Flows." *OECD*, (2020)

58 UNSD-DFID. *UNSD-DFID Project on SDG Monitoring*. United Nations (UN). September 30, 2019. <https://unstats.un.org/capacity-development/UNSD-FCDO/files/Data-Sharing-Arrangement-Guidelines-v3.pdf>

administrative data), specify the obligations for confidentiality.

- **Amendments** – outline the procedure for amending or updating the agreement.

These agreements improve data sharing by ensuring the following:

1. **Privacy Safeguards** – One ministry noted that in the absence of clear safeguards for data privacy and security, they are less likely to share administrative data. By formalising data sharing and being clear on the safeguards which protect confidentiality, ministries will be more confident in sharing the administrative data they collect. Equally, protecting personally identifiable information in the modern data ecosystem is challenging, which is why governance frameworks and regulatory policies cannot be overlooked.<sup>59</sup>
2. **Consistent Sharing** – Long-term overarching agreements for sharing specific administrative data should reduce uncertainty and the costly inertia generated by continually sending physical letters requesting the same data. When sharing is formalised, efficiency should increase as the ministries involved clearly understand the frequency and timeliness expected in the data exchange. Of course, the more interoperable the administrative data storage system is (i.e., if it is on an online server and not a personal hard drive), the easier this data sharing becomes.<sup>60</sup>

**Digitising Data Sharing:** Moving from paper to digital processes could also improve the process of data requests, signing data sharing agreements, and data exchange. Digitising the process could significantly reduce the time between requested data and shared data. Digitising these processes would have to be part of broader public sector reform, ensuring civil servants possess the necessary digital skills.

## 4.4 Centralising administrative data storage

- ✓ **Centralising administrative data storage within and across ministries could reduce friction in accessing, integrating, and analysing data.**

Given the siloed storage of administrative data within ministries and the importance of being able to combine administrative data with other data sources, there could be value in centralising data storage in two tiers:

- i. **Ministry Level** – Centrally store and back up all the ministry's administrative datasets from different departments and individual hard drives.
- ii. **National Level** – Create a central repository for administrative data sets

---

59 PARIS21, and the Mo Ibrahim Foundation (MIF) (2021)

60 World Bank. "Interoperability: Towards a Data-Driven Public Sector." World Bank, (2022)

from various ministries, especially those useful in many different policy contexts (e.g., tax data).

Promisingly, one core element of the Second National Strategy for the Development of Statistics is establishing **Statistics Units** in ministries where they are absent and enhancing their functionality in ministries where they already exist. ZamStats plans to provide the template architecture for how these Statistics Units should function – their primary purpose is coordinating data collection, storage, and use.<sup>61</sup>

Firstly, by increasing coordination between policy stakeholders and statisticians, **these Statistics Units could help ensure each ministry's data-driven outputs are demand-led** – that is, they offer insights relevant to the ministry's policy decisions.

Secondly, these **Statistics Units could provide the perfect platform for centralising data**. This centralisation could improve:

1. **Access** – Reduces the time spent requesting data from different departments and searching out specific individuals who hold the data.
2. **Awareness** – Increases data visibility and makes it easier for policy stakeholders to be aware of what data is at their disposal.
3. **Security** – If centralisation involves standardising data governance frameworks (e.g., policies on data access and backing up data), it could reduce the risk of unauthorised access to sensitive data and data loss.
4. **Quality** – Having all the data in one location makes it easier to standardise linkages across different datasets so they can be easily merged. Centralisation could also involve the central storage of data dictionaries to ensure that the way each data set is generated is clear.

Data centralisation could simply involve storing Excel sheets centrally or a more complex process of moving data storage to a cloud-based server. Online servers with rigorous access and authorisation controls can enable easy and efficient data access. These servers can be accessed remotely, which removes the need for continuous in-person data extraction, reducing the friction of data access and the burden of internal data requests on ministry staff.

- Importantly, central storage platforms should be designed to secure administrative data. For example, encrypting files (preventing other parties from understanding the data unless they have the encryption key) increases data security when using cloud servers. Equally, ensuring the operating system is regularly updated maintains defence against potential security breaches.<sup>62</sup>

---

61 Ministry of Finance and National Planning, and the Zambia Statistics Agency (2023)

62 Cole, S *et al.* (2021)

## CASE STUDY 7: ZAMBIA'S EDUCATION MANAGEMENT INFORMATION SYSTEM

Administrative data on schools is essential for accurately monitoring and evaluating education policies, resource allocation, and public infrastructure development. In 2001, in collaboration with AKROS and UNICEF, Zambia's Ministry of Education began working on an Education Management Information System (EMIS).<sup>63</sup> EMIS is still in place today, and it works in four stages:<sup>64</sup>

1. **Collection** – School headteachers record monthly data on student attendance, teacher attendance, student test scores, school feeding, and access to water and sanitation.
2. **Collation** – School headteachers meet area supervisors who collate the administrative data on a mobile platform (using a tablet).
3. **Centralisation** – Administrative data is sent and stored on the cloud-based server (EMIS).
4. **Insights** – EMIS is used to produce and share insights with various actors, e.g., Ministry of Education officials and district education officers.

Therefore, EMIS is a centralised online server where relevant ministry staff can access administrative data efficiently.

**Centralising administrative data at the national level could make it easier to combine administrative data with other data sources (e.g., surveys or censuses) and produce targeted data-driven insights.** Centralising administrative data that is especially valuable for policymaking could be beneficial. Indeed, to develop the National Statistical System (NSS), **ZamStats intends to start building a state-of-the-art centralised administrative data repository in 2024.**

- Significantly, such centralisation of administrative data should not restrict access to sufficiently disaggregated information. A prominent concern raised in the interviews with ministries is that centralisation could make it challenging to access information at the right level of disaggregation. To facilitate precise policy analysis, policy stakeholders still need to be able to access such disaggregated data.

---

<sup>63</sup> Warrick, R. Drake, George Caldwell, and Jennifer Anderson. *Education Management Information System Program (EMIS) and Related Activities: Final Program Report*. United States Agency for International Development (USAID). 2024.

<sup>64</sup> AKROS. "Education Management Information Systems." Accessed January 25, 2024. <https://akros.com/informatics/emis/>

## 4.5 Digitising administrative data collection

- ✓ **Digitising critical administrative data at the point of collection could improve data access and quality while eliminating the costs of manually transferring paper-based records onto digital platforms.**

According to UNCDF estimates, **Zambia's 'Digital Economy Score' was 55% in 2022, indicating that Zambia is in the phase of digital expansion.**<sup>65</sup> While much of Zambia's administrative data is stored digitally (e.g., on Excel), the collection is often still reliant on paper-based methods. The government has noted the burden current paper-based practices put on the individuals responsible for administrative data collation and storage.<sup>66</sup>

The survey submissions reveal that to implement digitisation, some ministries would welcome increased investment in ICT infrastructure at provincial centres. Of course, under the pressure of financial constraints, such investments may seem too costly. However, *if* the data being digitised is essential for policy making, **the initial fixed cost of digitising administrative data collection can be recouped via long-run operational efficiency gains.**

Digitisation can potentially be a cost-saving venture for Zambia's ministries – freeing up time for statisticians to complete other essential data tasks. Equally, suppose digitisation means the data is transmitted immediately to a central platform. In that case, it can increase the timeliness of data access and the ease with which policy planners can make evidence-based decisions.<sup>67</sup>

- Policymakers in the City of Cape of Town (CCT) gained from digitisation. After implementing the 2002 Smart City Strategy, digitising data collection made it possible for CCT's policymakers to access real-time information on sectors and specific issues, facilitating faster operational decisions and policy responses.<sup>68</sup>

### **CASE STUDY 8: ZAMBIA INTEGRATED LAND ADMINISTRATION SYSTEM**

Land administration data is extremely valuable for informing policy decisions on issues related to tax, natural resource management, rural development, and much more. In 2023, the Ministry of Land and Natural Resources (MLNR) deployed the Zambia Integrated Land Administration System (ZILAS). ZILAS transitioned Zambia's land administration from a paper-based system to a modern digital platform.

65 United Nations Capital for Development. "Inclusive Digital Economy Scorecard." Accessed January 12, 2024. <https://ides.uncdf.org/>

66 Ministry of Finance and National Planning, and the Zambia Statistics Agency (2023)

67 World Bank. *Accelerating Digital Transformation in Zambia: Digital Diagnostic Report*. Washington DC: World Bank, 2020.

68 Wright, C. et al. (2023)

To use ZILAS, citizens register an account on the digital government services platform ZamPortal; then, they can access any of the 60 ZILAS digitised services with built-in payment and digital signature features. For example, ZILAS can be used to apply for a land lease and title registration.<sup>69</sup>

While there are still some issues with lagged updating, ZILAS has revolutionised how land administration is managed in Zambia. MLNR noted that the efficiency gains are significant, and citizens no longer lose their land registration information. Equally, by reducing physical interactions with ministry staff, ZILAS has eliminated opportunities for corruption.<sup>70</sup> From a policy standpoint, the platform has also brought much greater disaggregation and specificity. For example, MLNR highlighted that they can now analyse the breakdown of land registration by gender.

However, there are essential points to consider when digitising administrative data collection:

- 1. Digitising administrative data collection is reliant on technology and electricity.** The Ministry of Health highlighted that a challenge to their administrative data collection is the need for more stable connections in rural health centres. Using digital health registers becomes ineffective when the data collectors have to re-input the data multiple times during power cuts. Paper-based collection may remain the more viable option in these areas until power becomes more reliable.
- 2. Digitising administrative data collection requires sufficient digital literacy** amongst the population interacting with the system. If ministerial officers conduct data collection, then the issue of digital literacy pertains to a small group. However, suppose data collection is self-service (i.e., in the case of ZILAS). In that case, it is necessary to identify the national 'digital divide' – who can use the platform and who will not.<sup>71</sup> As a guide, UNCDF estimates suggest Zambia's digital literacy was 49% in 2022.<sup>72</sup>

---

69 Ministry of Lands and Natural Resources (MLNR). (n/a). "Zambia Integrated Land Administration System." Accessed 15 January, 2024. [https://www.mlnr.gov.zm/?page\\_id=2709](https://www.mlnr.gov.zm/?page_id=2709)

70 Mwale, M. *Ministerial Statement to Parliament on the Deployment of the Zambia Integrated Land Administration System (ZILAS)*. Ministry of Lands and Natural Resources. 2023.

71 Wright, C. et al. (2023)

72 United Nations Capital for Development. "Zambia 2022: Inclusive Digital Economy Score." Accessed January 10, 2024. <https://ides.uncdf.org/2022/ZMB>

---

## 5. Conclusion

While Zambia's administrative data has the potential to close critical evidence gaps in public policymaking, its use is currently limited relative to what it could be. Promisingly, some ministries are capitalising on the value of administrative data for policy monitoring, but there is still room to expand its use in targeting and evaluating policy. Transforming the relationship between administrative data and decision-making in Zambia's ministries will require reforms which address the underlying issues related to data access, quality, and capacity.

Arguably, Zambia's policymakers should prioritise actions that reduce capacity gaps on the demand side of data. Leveraging research partnerships is one-way policy stakeholders can improve their ability to identify how different data can be combined to answer specific policy questions. As a result, such partnerships could help to create a consistency of demand for Zambia's administrative data. Enacting policies that streamline the supply of administrative data will have a much more significant impact in the long run if high demand exists for that data.

Regarding data access, there are low-hanging efficiency gains to be made from formalising data-sharing relationships between ministries. In addition to digitising data requests, having formalised agreements which safeguard privacy and ensure consistent data sharing will reduce friction in data exchange. Equally, centralising administrative data storage, both within and across ministries, could increase the ease with which administrative data can be accessed and integrated with other data sources – facilitating targeted policy analysis.

On the issue of data quality, there is a blend of options which could mitigate problems with incomplete data, missing metadata, and weak data linkages. One option could be institutionalising data skills training, where ministries can consistently learn from one another and ensure their data management practices are current. Additionally, where appropriate, digitising administrative data collection could avoid reliability issues with manual digitisation and increase timely data access.

Looking forward, this paper aims to open up the discussion on how Zambia's ministries can unlock the potential of existing administrative data. The constraints and policy options cited in this paper are not exhaustive, and there is a need for more thorough assessments of how individual ministries and their respective policymakers can better use the data at their disposal. Ultimately, driving forward the culture of evidence-based policymaking – including the use of administrative data – will require coordinated action from Zambia's government ministries, quasi-government agencies, and the Zambia Statistics Agency.



## Annex 1: Study context & method

The 8th National Development Plan (8NDP) emphasises the importance of using quality and timely data to make informed policy decisions. In this vein, one of the critical priorities of the 8NDP is strengthening Zambia's Monitoring and Evaluation (M&E) framework – this should increase the ease with which policymakers can track national development progress and adjust public policies accordingly.<sup>73</sup>

Strengthening M&E will involve developing the national data infrastructure, including the collection, processing, and policy use of administrative data. Indeed, Zambia's Second National Strategy for the Development of Statistics (NSDS2) highlights the need to improve administrative data's coverage, availability, and objectivity.<sup>74</sup> Before implementing transformational changes in the administrative data space, it is important to establish a foundational understanding of the current administrative data landscape in Zambia.

The Zambia Statistics Agency (ZamStats) is responsible for the publication of official statistics. ZamStats' mandate includes the collection, compilation, and dissemination of national-level data – primarily taken from surveys and censuses.<sup>75</sup> While ZamStats intends to establish a central repository for administrative data in 2024, currently **most public administrative data sits within the 25 government ministries.**

Therefore, the IGC Zambia team, in collaboration with the Ministry of Finance and National Planning, systematically studied the relationship between administrative data and the public policy lifecycle in Zambia's government ministries. Note: this study did not assess the quality of individual data sets and instead focussed on building an overarching picture of Zambia's administrative data landscape.

The study involved two stages:

- 1. In-person Interviews** – with 14 government ministries. To ensure the insights from the interviews were representative, a range of ministerial staff were involved, from statisticians and economists to planners and directors. The primary purpose of these interviews was to understand the relationship between administrative data and decision-making while also gauging the level of inter-ministry data sharing.
- 2. Follow-up Surveys** – submitted by 13 government ministries. The surveys were open-ended and designed to reveal qualitative insights, including specific examples of the administrative data which is being collected and how it is both stored and used in policymaking. Equally, these surveys intended to reveal the technical capacity ministries have to handle the administrative data they collect.

---

73 Ministry of Finance and National Planning (2022)

74 Ministry of Finance and National Planning, and the Zambia Statistics Agency. *Second National Strategy for the Development of Statistics*. 2023. [t.ly/5jpgD](https://t.ly/5jpgD)

75 Zambia Statistics Agency (ZamStats). 'About Us.' Accessed December 20, 2023. [t.ly/4IX9o](https://t.ly/4IX9o)







**IGC**

[www.theigc.org](http://www.theigc.org)

Photo: GCS shutter/E+ via Getty Image