Qhala

Catalysing Africa's
Digital Future

Data Maturity Assessment: Results and Findings

Timeline & Outputs

Planning for the assessment

Researching current methods for data maturity assessment and designing a more suited approach

Aug 2021

Aug - Oct 2021

Data Collection

Disseminating questionnaires and following up responses

Oct 2021 - Feb 2022

Data organisation and analysis

Developing the final report on the findings of the data maturity assessment toolkit exploration

Final Workshop

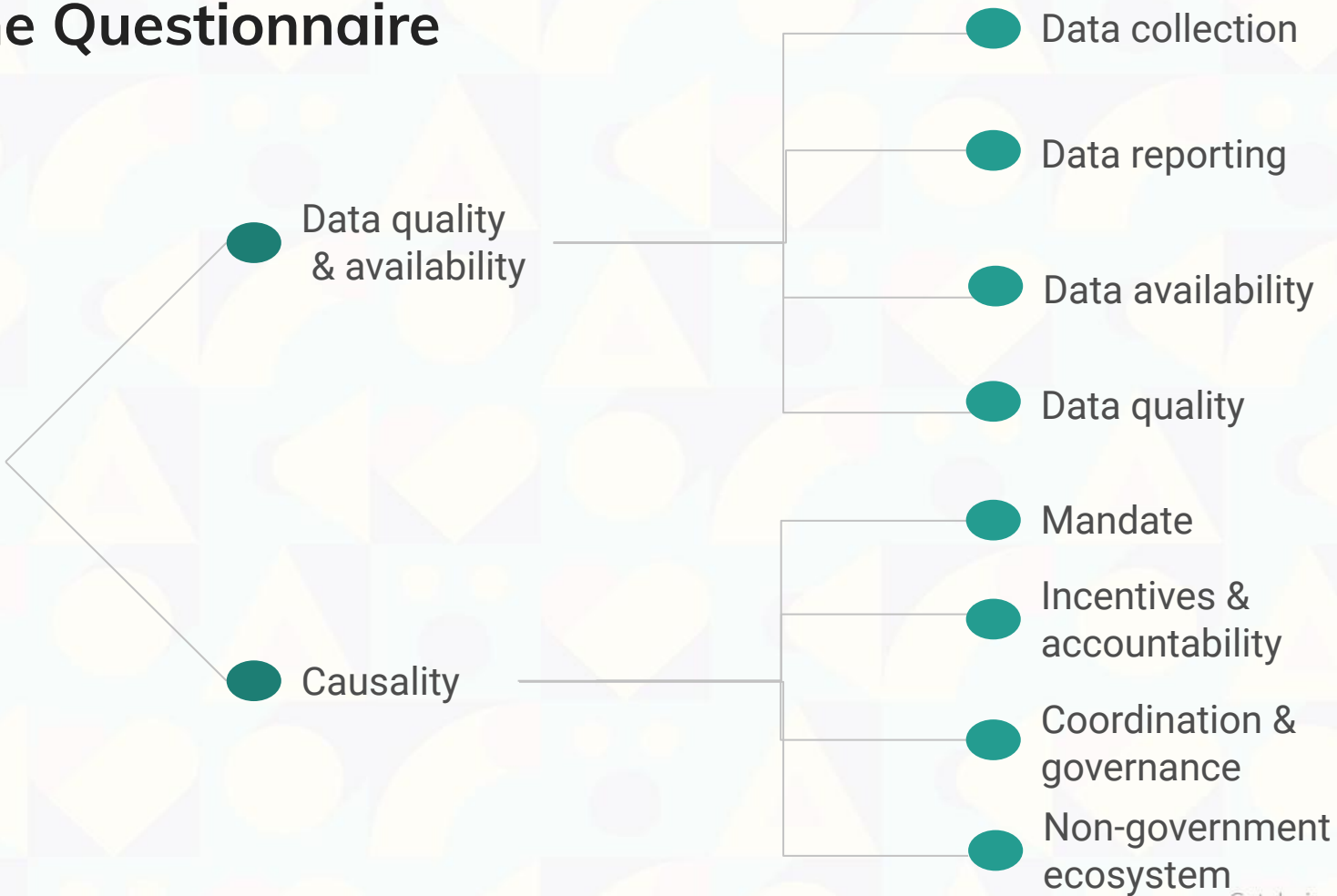
Presenting findings

Presenting final findings/**county-specific plans to follow the final workshop**

Study Objectives

- To understand the challenges and limitations that prevent county governments from collecting, using and sharing data in their operations, ie. **data maturity**
- To measure the level of data maturity in eight Kenyan counties
- To develop a standardised data maturity assessment toolkit that all county governments can use to measure their own level of data maturity
- To provide county governments with the data that demonstrates structured investments needs which which can be leveraged through partnerships for resource mobilization
- To understand the ways in which county governments can be supported to work towards higher levels of data maturity

The Questionnaire



Participants

With a target goal of interviewing participants in 10 counties, we ended up with 8 participating counties within the 4 departments of Governance, Health, Education and Agriculture.

These include;

- Nairobi
- Samburu
- Turkana
- Nandi
- West Pokot
- Nakuru
- Mandera (virtual)
- Makueni (virtual)

Challenges and Limitations

1. The questionnaire was too lengthy for an exercise with no tangible incentives and some indicators were redundant or unclear. We developed a second version of the tool from this feedback which we used for the remaining department in that county, and for subsequent interviews in the other counties.
2. The survey respondents were not incentivised to participate as we felt it may affect the quality and candour of responses. We reviewed the tool and came up with a shorter version that captured as much information as possible and took less time to complete.
3. Each county has different priorities when it comes to sectoral indicators and hence it was difficult to compare responses from different counties.
4. Some of the interviews were conducted virtually or filled in by the officials due to connectivity issues and/or logistical and scheduling challenges.

DATA MATURITY MODEL

Stages of the Data Maturity Model

Seedling

This stage is the initial step that counties take towards building their data infrastructure. Data is collected as part of operations but without a specific use or objective, it is not processed for use or consumption.

Growing

Data infrastructure is improving at this stage to accommodate regular collection and reporting. The use of data is integrated into departmental processes and regulations. The system has expanded to include data from other sources within the local and national government.

Flowering

Data-driven decision making is a priority at this stage of maturity. The use of technology is integrated into data management practices. The scope of data collection is widened to include partnerships outside the government. The data is reliable and available to all interested parties.

Fruitful

At the final stage of maturity, data is used for innovation and strategy. High quality data is freely available and up to date due to an intentional boost in resourcing and supportive legislation. Contextual data such as location and demographics is included for more robust analysis.

1) Seedling

- Physical data collection
- Data is collected infrequently, and not on a regular schedule
- Data is stored in a format that cannot be easily accessed or used by most stakeholders
- Data comes from a single source or only within the county govt

2) Growing

- Physical/automated data collection
- Data is collected infrequently, but on a regular schedule
- Data is reported quarterly/yearly
- Data is stored in a format that can be accessed within government, but with restrictions
- Data comes from multiple internal sources, with no formal mechanism for collaboration
- Data collection is required for administrative reporting

3) Flowering

- Physical/automated data collection
- Data is collected infrequently, but on a regular schedule
- Data is reported quarterly/yearly
- Data is stored in a format that can be accessed within government, but with restrictions
- Data comes from multiple internal sources, with no formal mechanism for collaboration
- Data collection is required for administrative reporting

4) Fruitful

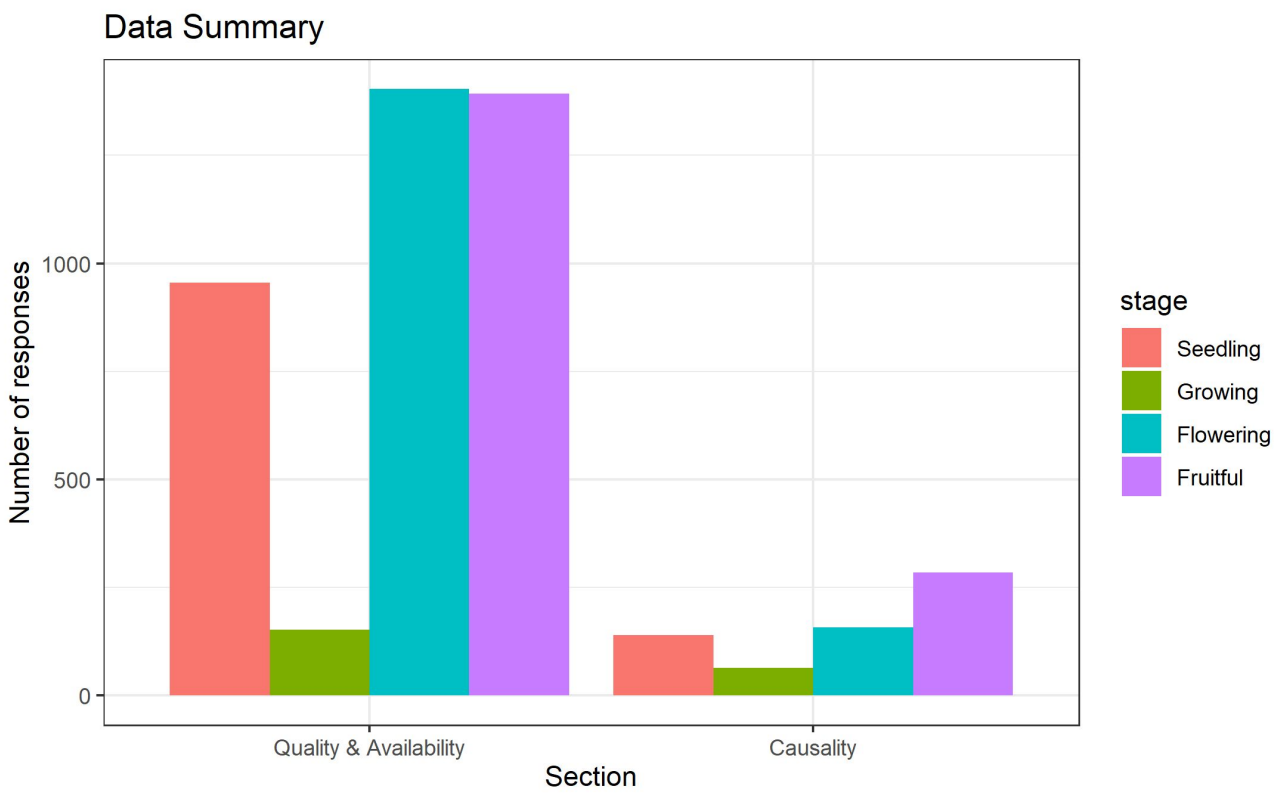
- Data collection is legally mandated, with penalties for not reporting data
- Database is updated in real time
- Data is stored in a format that can be used by all stakeholders without restrictions
- Collaborators can contribute to databases and use available data for new projects
- Audit reports are also available to the public
- Stakeholders have clear channels for feedback based on shared data
- The county has a budget with an extensive funding stream for data functions
- The county elected representatives and management prioritise data quality and availability in their operations
- Marginalised groups are considered in data-driven decision making

Scoring - example

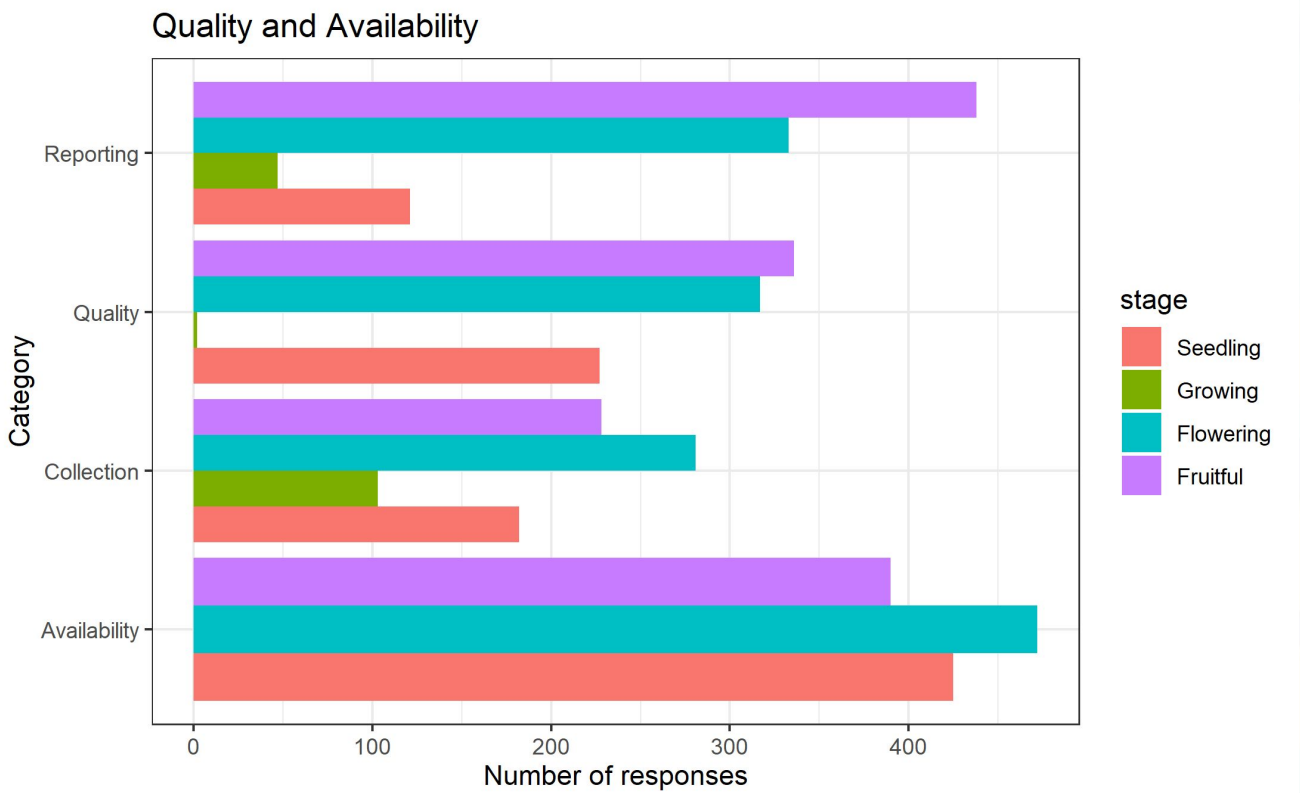
Answer	Score	Stage in Model	Property
Editable soft files, online.	4	Fruitful	Data is stored in a format that can be used by all stakeholders without restrictions
Non editable (e.g. PDF) soft files, online.	3	Flowering	Data is stored in a format that can be used by all stakeholders with some restrictions
Editable soft files, CD/DVD/Flashdisk.	2	Growing	Data is stored in a format that can be accessed within government, but with restrictions
Non-editable soft files, CD/DVD.	2	Growing	
Paper	1	Seedling	Data is stored in a format that cannot be easily accessed or used by most stakeholders

RESULTS AND FINDINGS

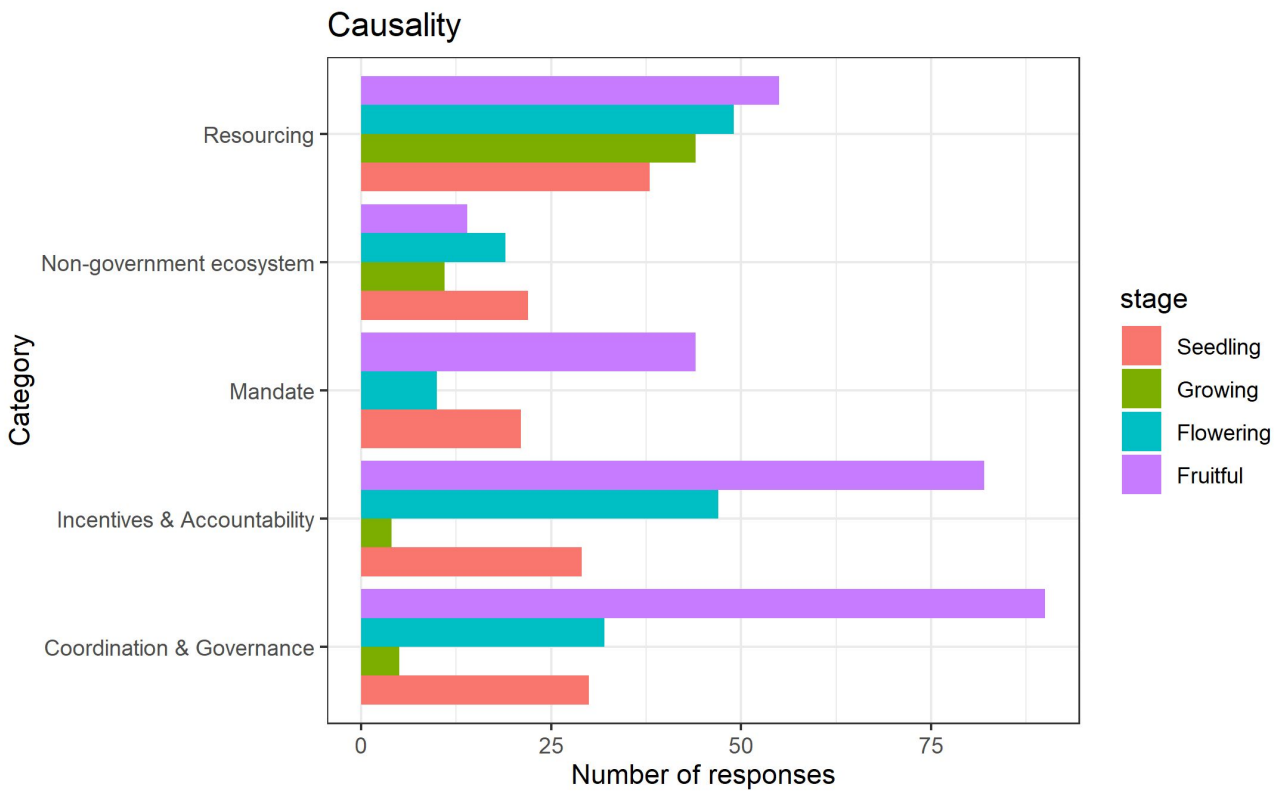
Summarizing all sectors - Insights



All sectors - Data quality and availability



All sectors - Causality

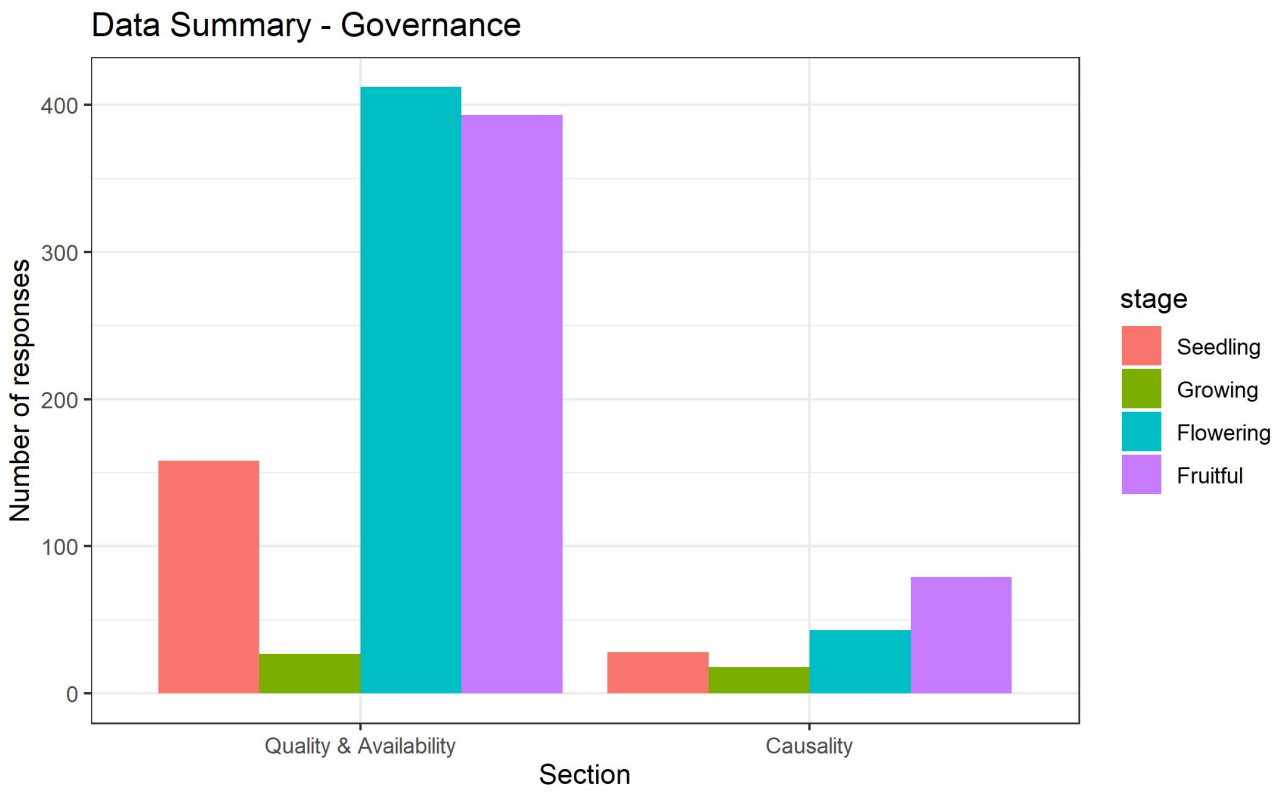


Governance

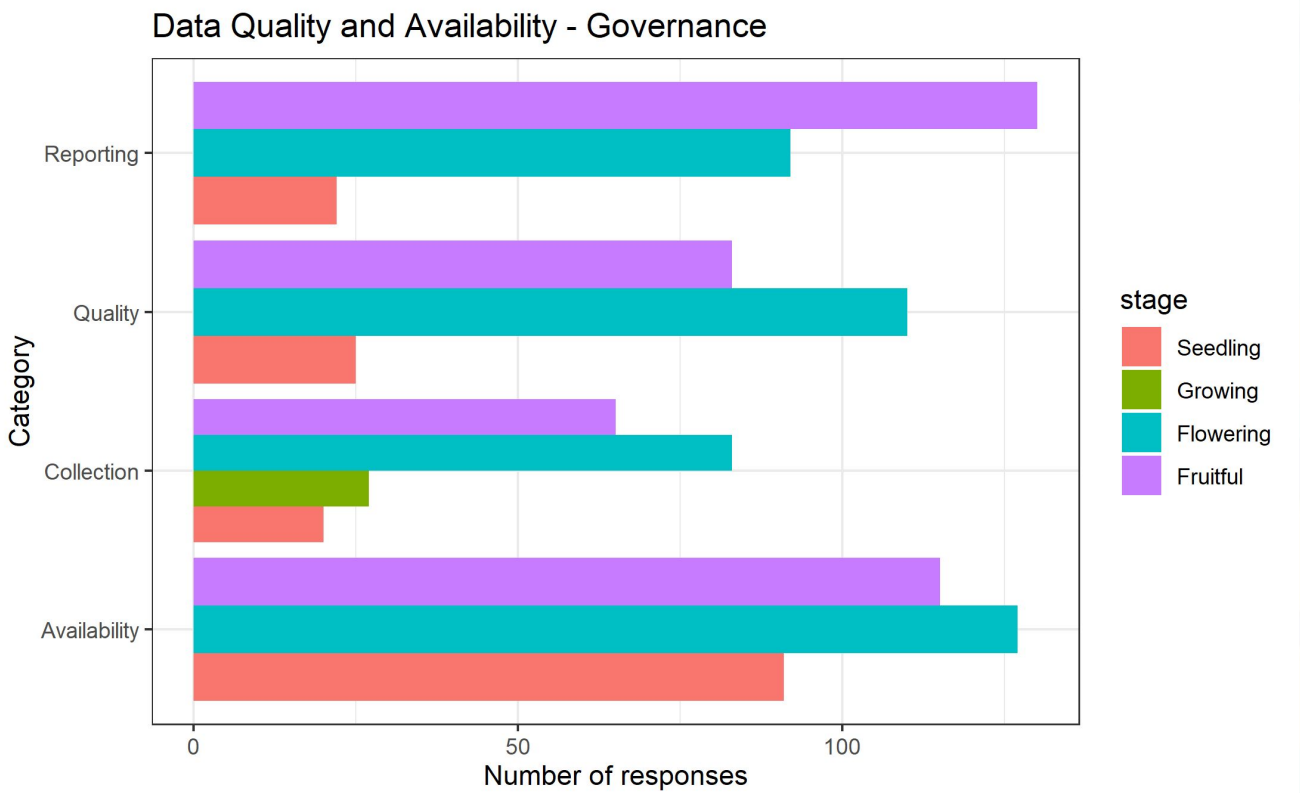
Governance - Indicators

1. County budget - development and recurrent
2. County annual reports
3. Amount of ‘own’ revenue as a percentage of total budget
4. Collection efficiency
5. Total county employees per 1000 population, disaggregated by level of education (primary, secondary, university)
6. No. of Services under E-Governance
7. No. of town hall meetings held
8. No of civic education events held in the County
9. Annual number of public local government meetings and total attendance
10. No. of policies and laws passed per year
11. Does the county have an Ombudsman to address Citizen grievances
12. Number of live broadcasts of House proceedings

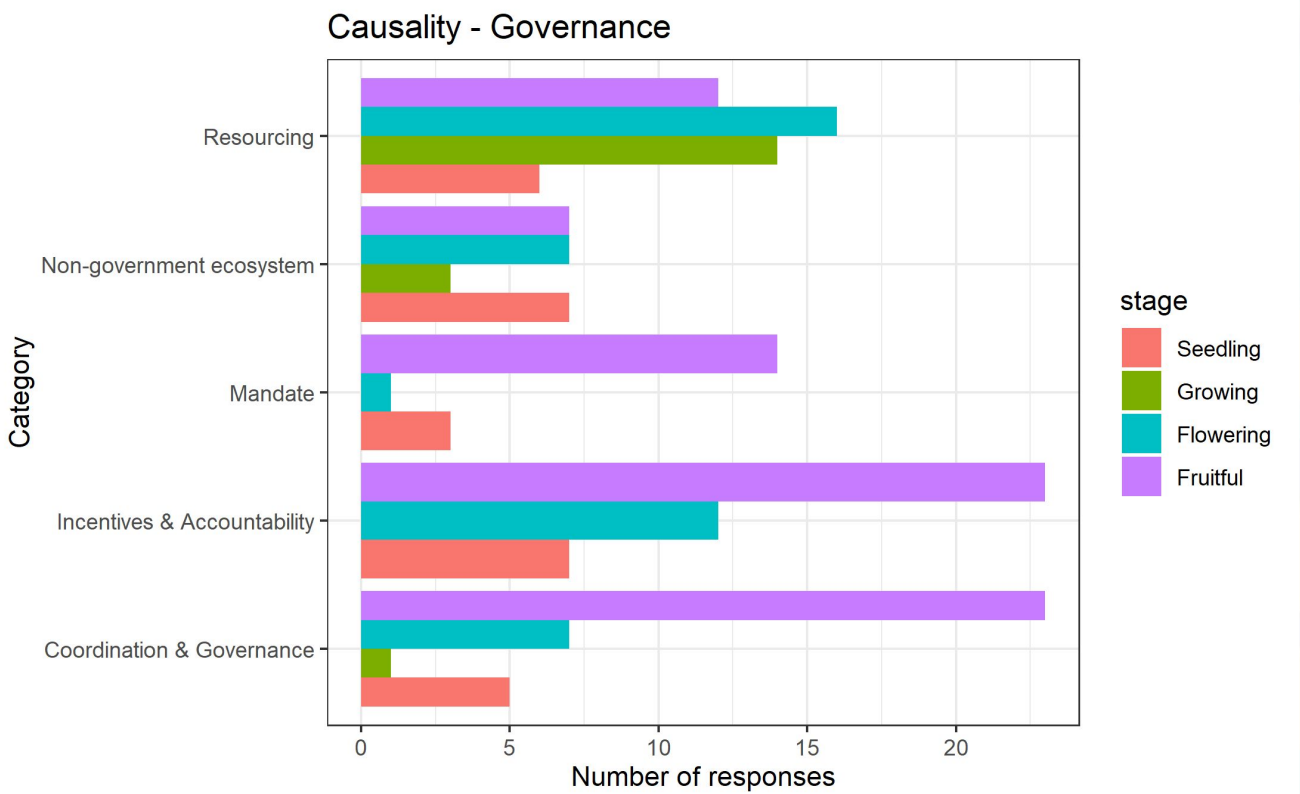
Governance - Insights



Governance - Data quality and availability



Governance - Causality



Governance - Recommendations

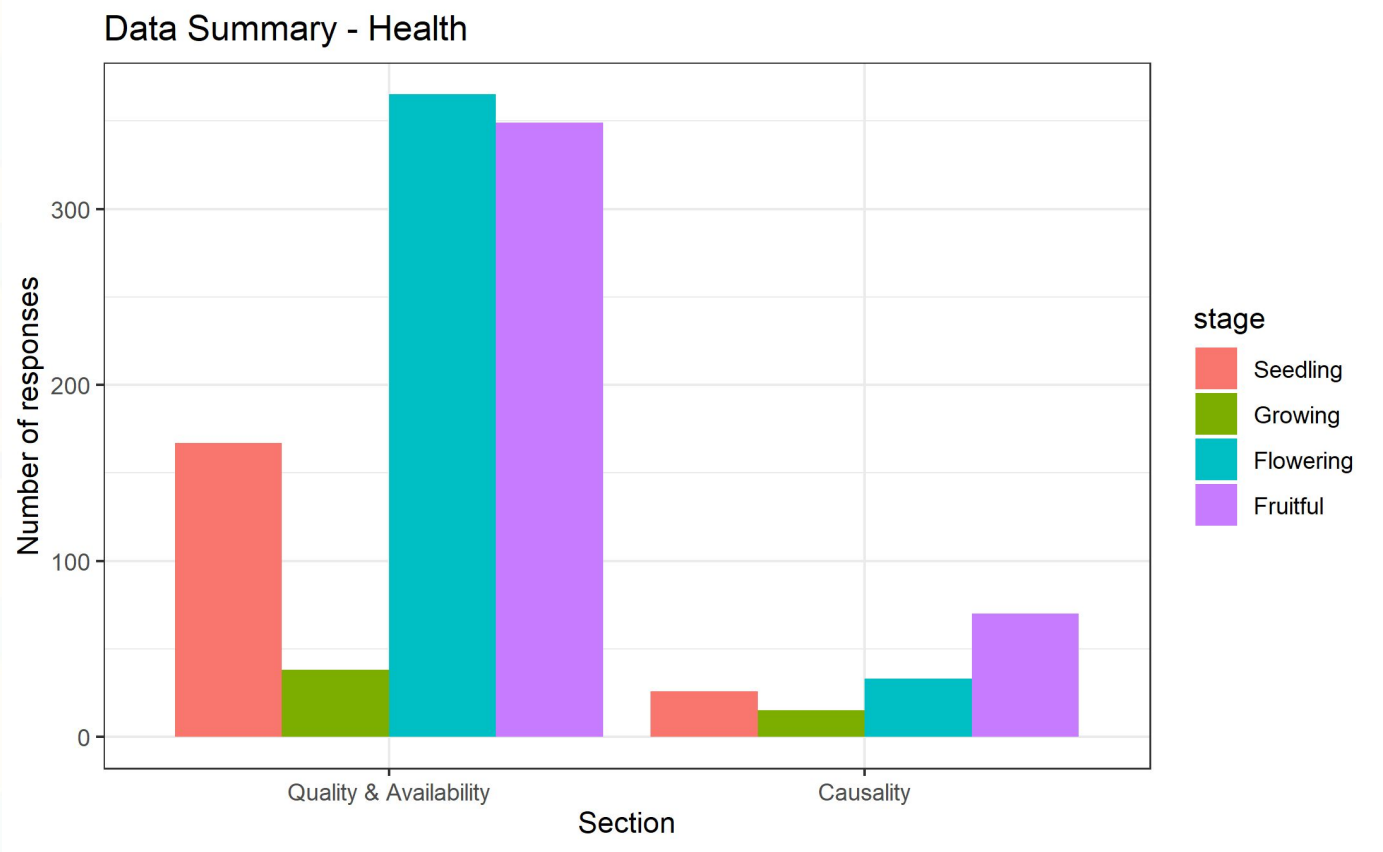
- Responsible departments should find more innovative uses for governance data beyond financial planning, monitoring and evaluation.
- Records on public participation (the data that is published/or collected during these sessions) should be digitised for easier access.
- The Commission on Administrative Justice has developed and issued a model law on access to information for county government, and engaged a number of counties to enact their specific ATI laws (for example Kilifi and Makueni). This is one of the formative steps in ensuring data governance is instituted.
- Even when such data is collected and published, there should be intentionality around reaching the majority especially the marginalised and disadvantaged groups in the communities.

Health

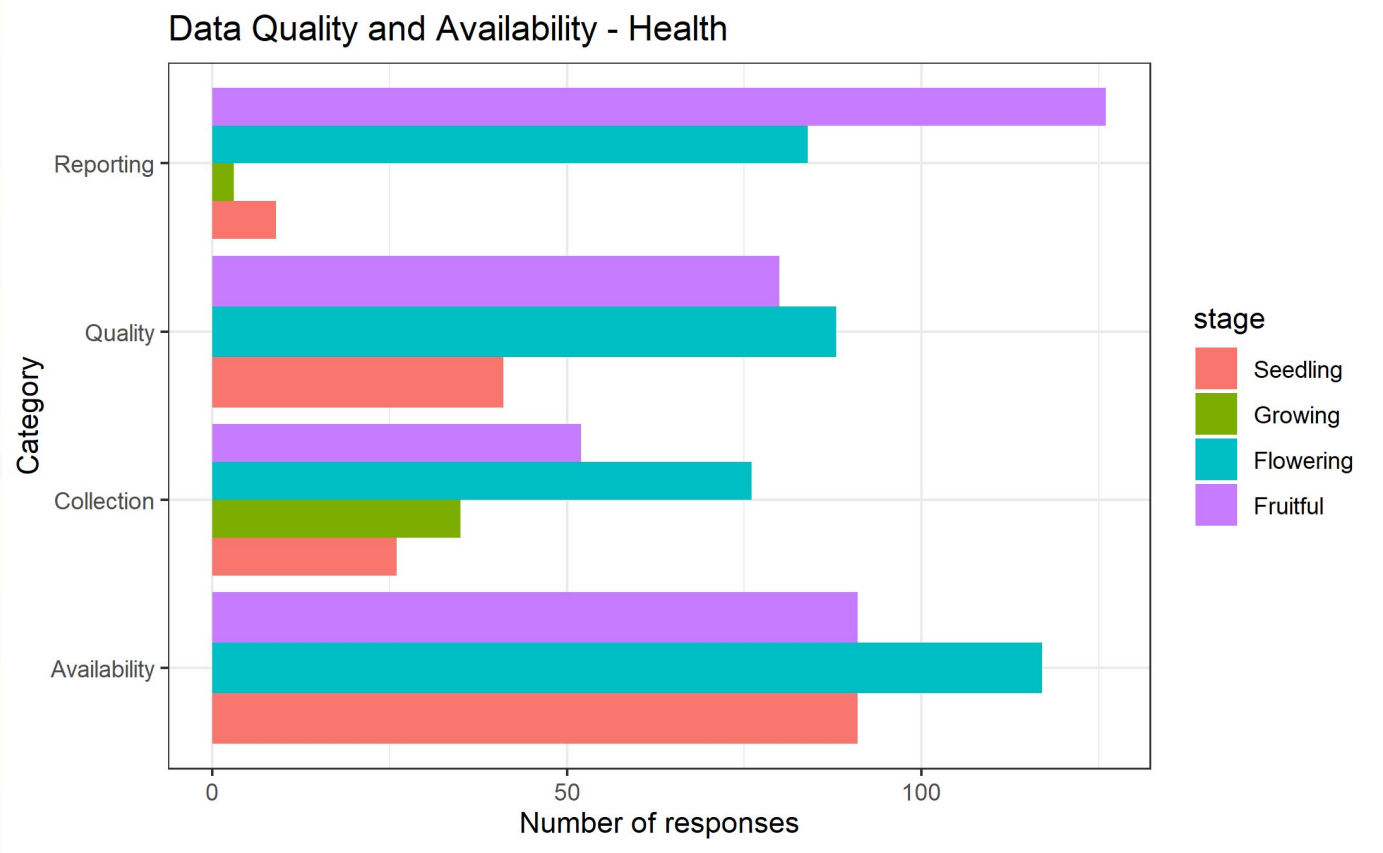
Health - Indicators

1. Doctor:population ratio
2. Health facility numbers in the County
3. Proportion of deliveries conducted by skilled birth attendants
4. Proportion of children under the age of 1 fully immunized
5. Proportion of general population testing positive for malaria
6. % of villages declared Open Defecation Free
7. Number of environment & public health workers
8. No of TB cases identified and put on treatment
9. Proportion of pregnant women attending 4th ANC visit
10. Proportion of women of reproductive age accessing family planning services
11. HIV prevalence rate
12. Percentage of children under 5 that are underweight

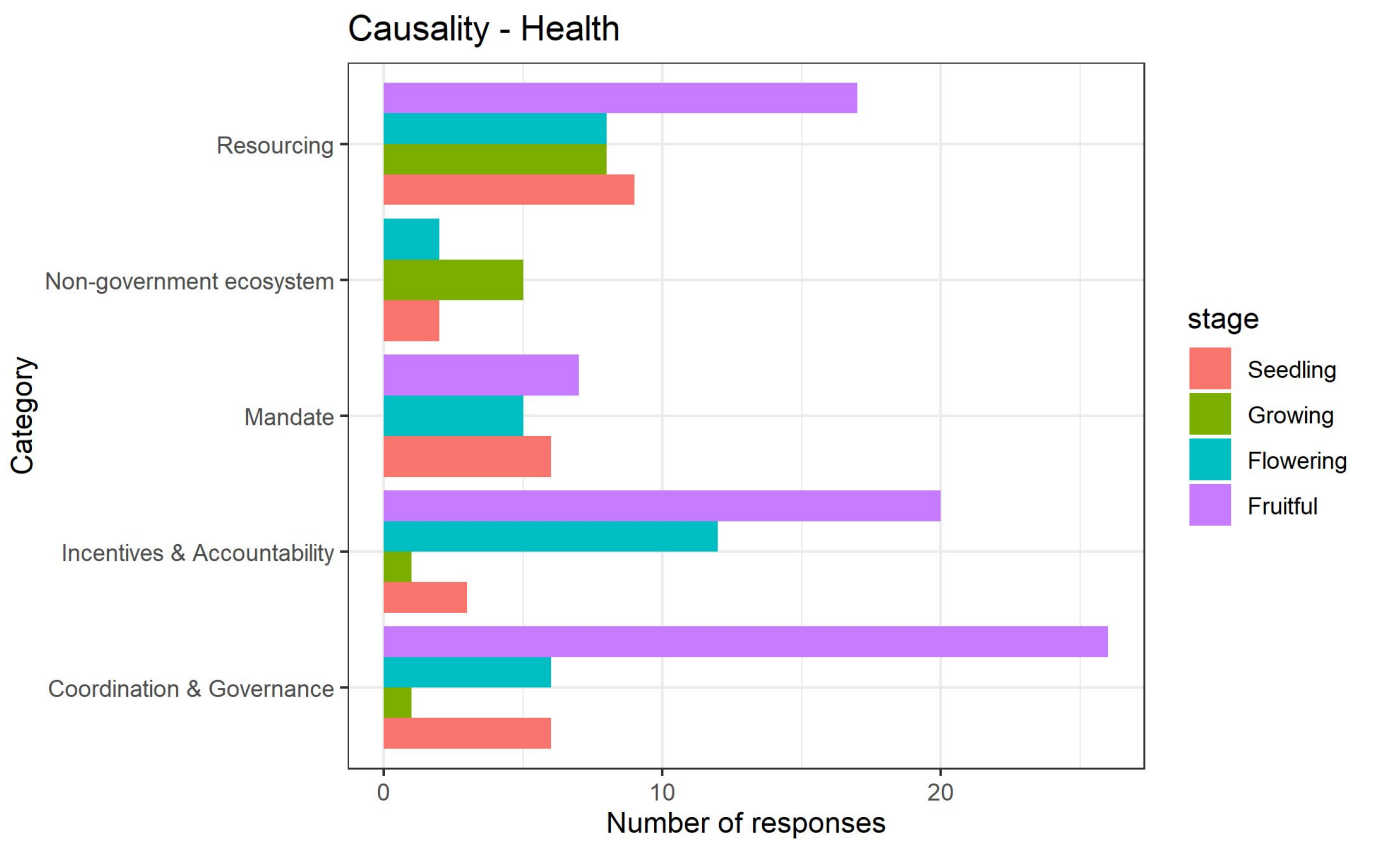
Health - Insights



Health - Data quality and availability



Health - Causality



Health - Recommendations

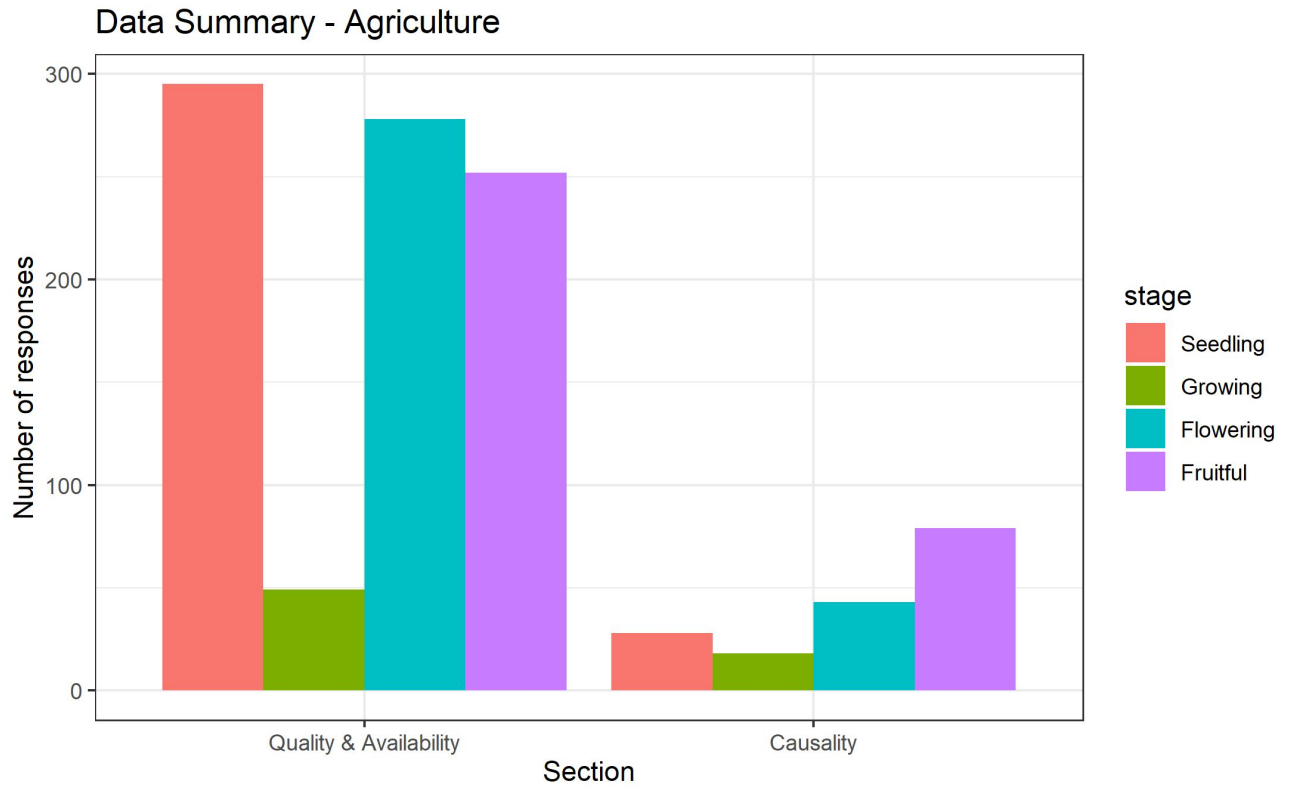
- Health data is readily available for indicators included in the KHIS. Hence it was suggested by respondents that the questionnaire can be amended to include less common indicators such as neglected tropical diseases.
- Health data collection and storage can be made more efficient by automation that links Electronic Management Systems (EMRs) at the facility level with the KHIS database in real time.

Agriculture

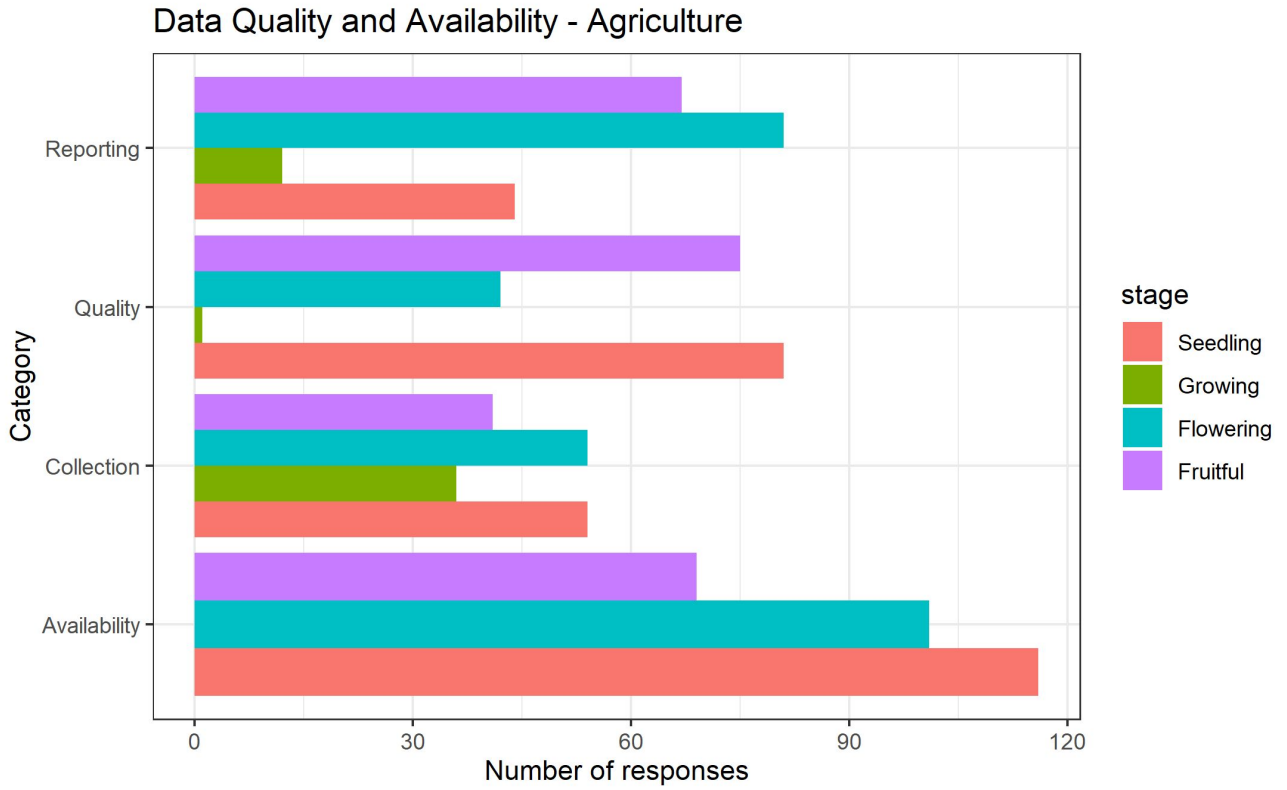
Agriculture - Indicators

1. Tonnes of Maize Produced annually
2. Percentage of population receiving food aid
3. Proportion of animals vaccinated
4. No of annual trainings for farmers
5. Number of active farmer associations
6. Hectares of arable land under crop production
7. Tonnes of fish produced
8. No. of forests conserved, managed and protected
9. Ha of crops under irrigation
10. Number of beneficiaries accessing farm inputs (tools, seeds, fertilizers, pesticides)

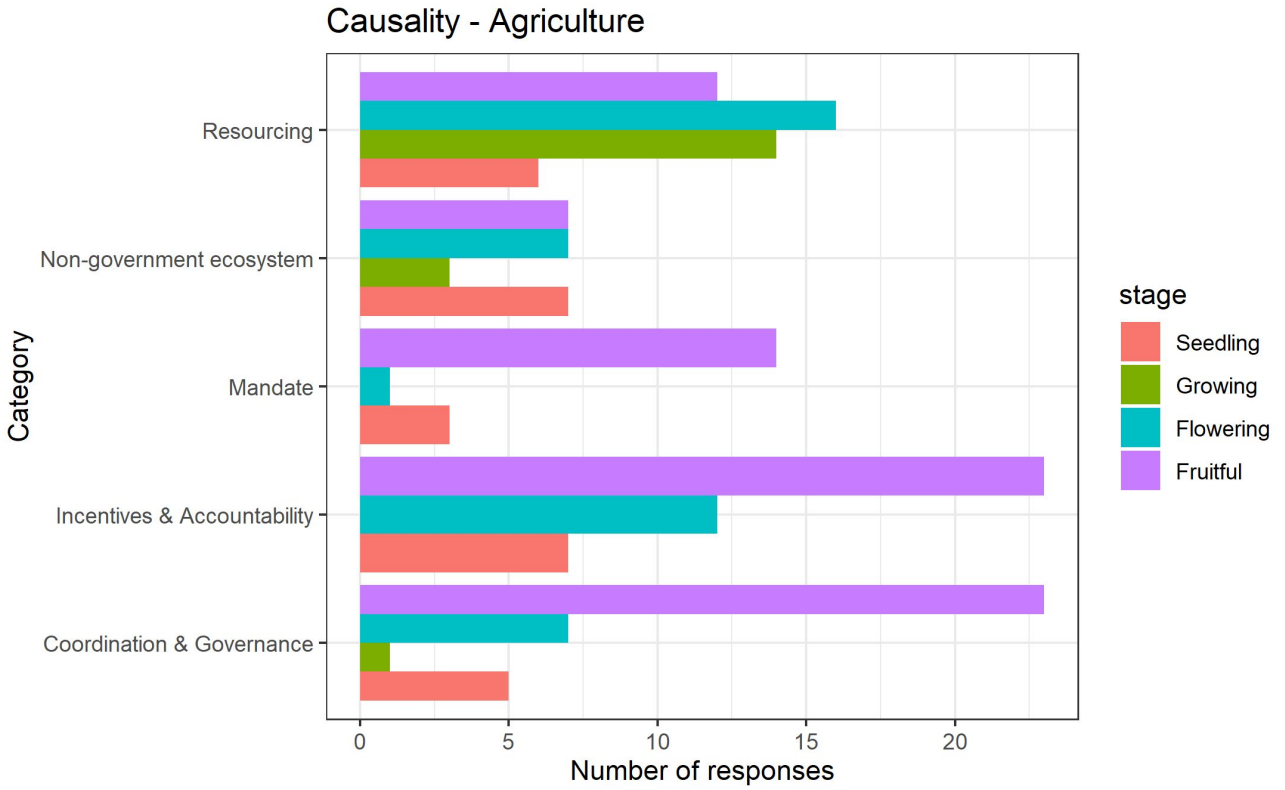
Agriculture - Insights



Agriculture - Data quality and availability



Agriculture - Causality



Agriculture - Recommendations

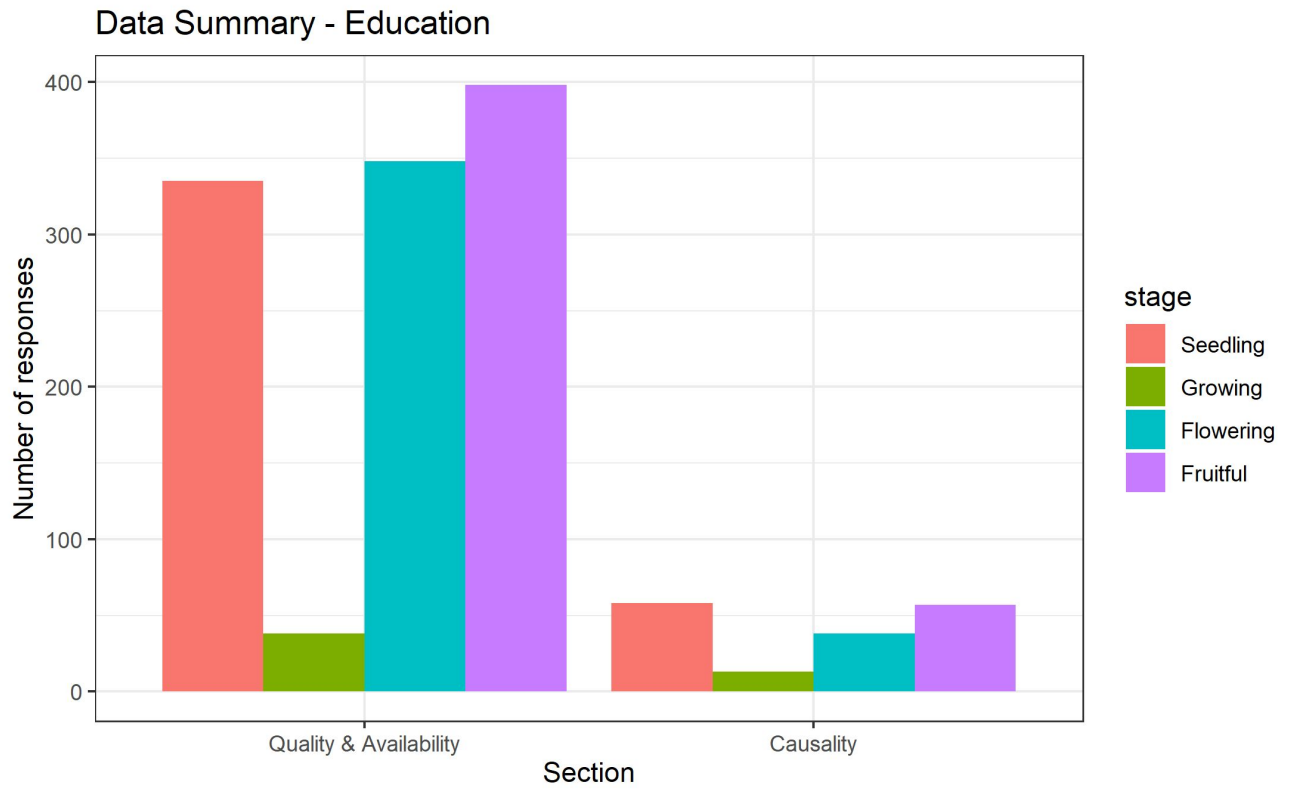
- County departments of agriculture can be assisted to come up with standardised templates and centralised secure web-based databases for Ag. data management - across entire value chains.
- Agriculture departments can collaborate with actors in the private sector to collect and use additionals/alternative data for planning. This, for example, could help with forecasting and forecasting extreme events that affect agricultural outputs.

Education

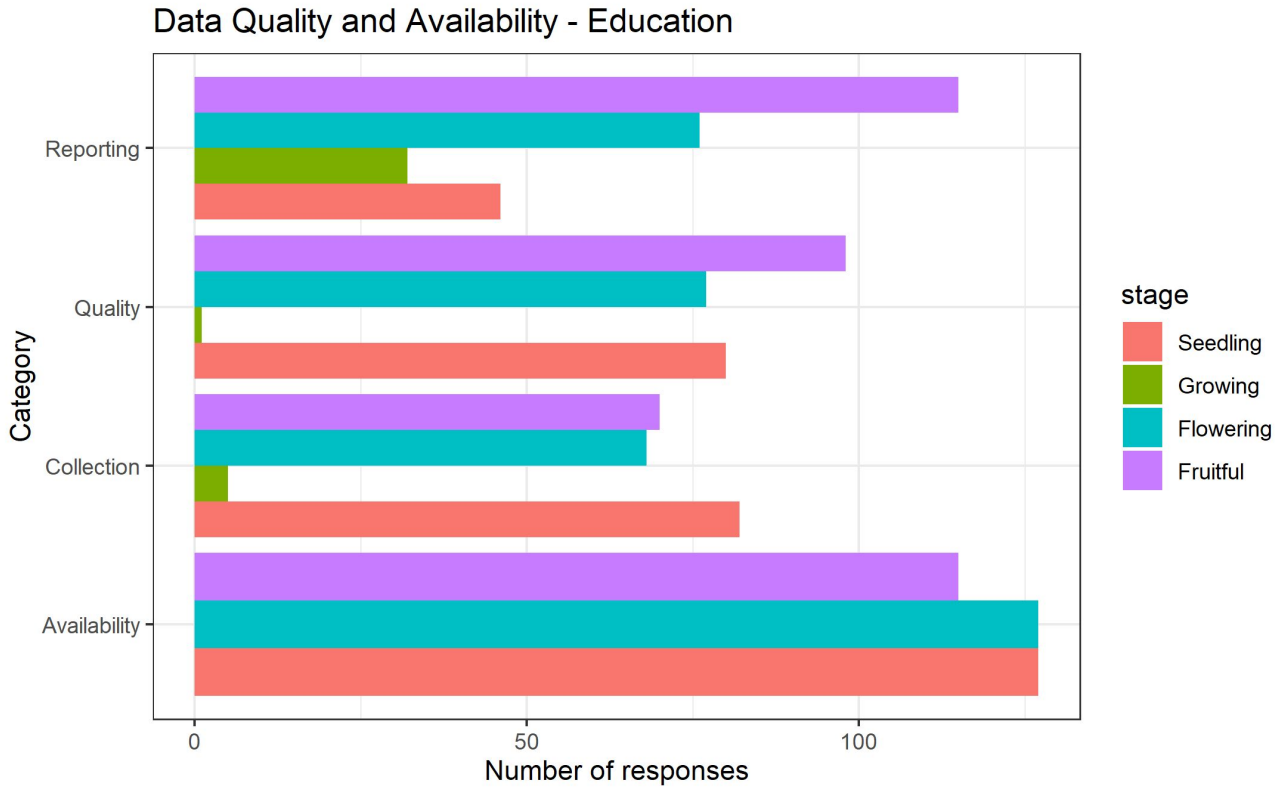
Education - Indicators

1. ECDE enrolment rate
2. ECDE teacher/Pupil ratio
3. Rate of ECDE student transition to primary school
4. Number of ECDE classrooms constructed and equipped in the last financial year
5. Number of ECDE schools with a feeding programme
6. Amount spent per ECDE student
7. ECDE learners with special needs(%)
8. TVET enrolment rate
9. TVET instructor:trainee ratio
10. TVET course completion rate
11. Number of vocational training centres equipped with modern tools and equipment
12. Proportion of TVET centres with access to internet connectivity

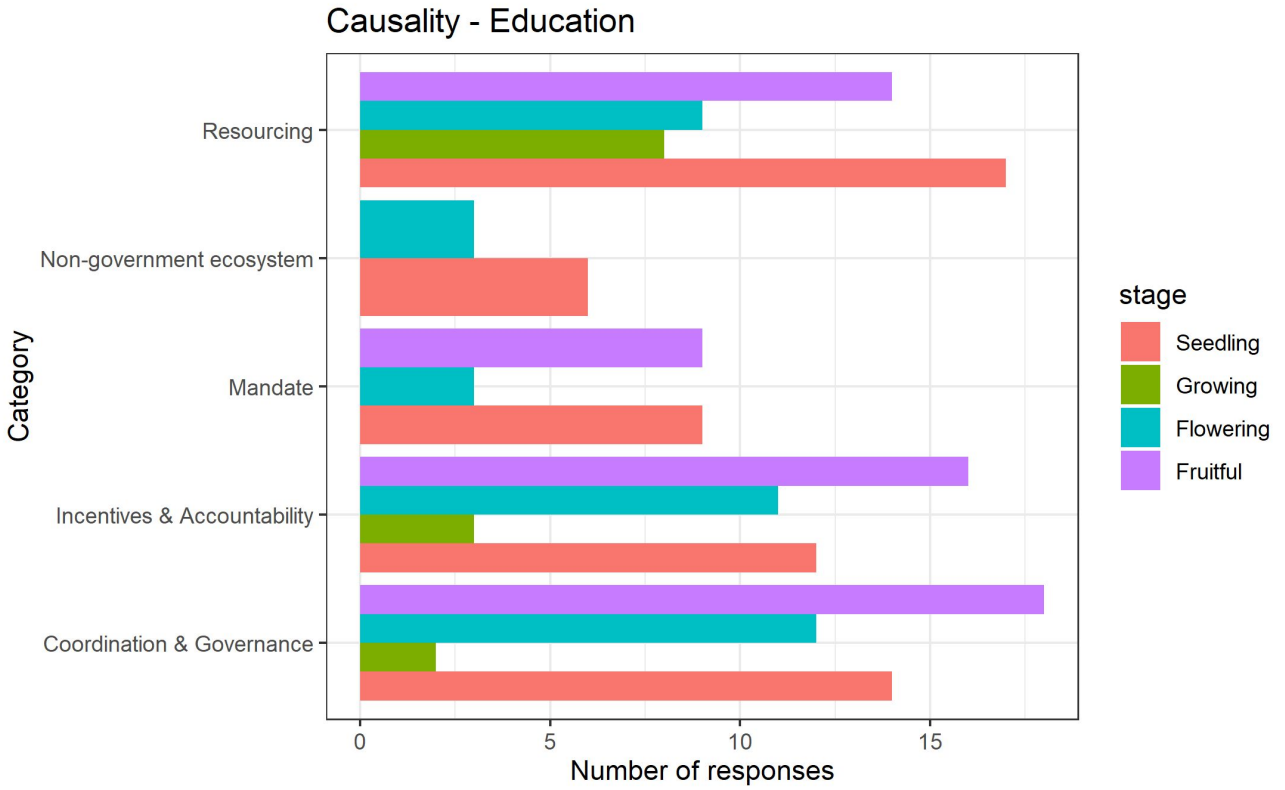
Education - Insights



Education - Data quality and availability



Education - Causality

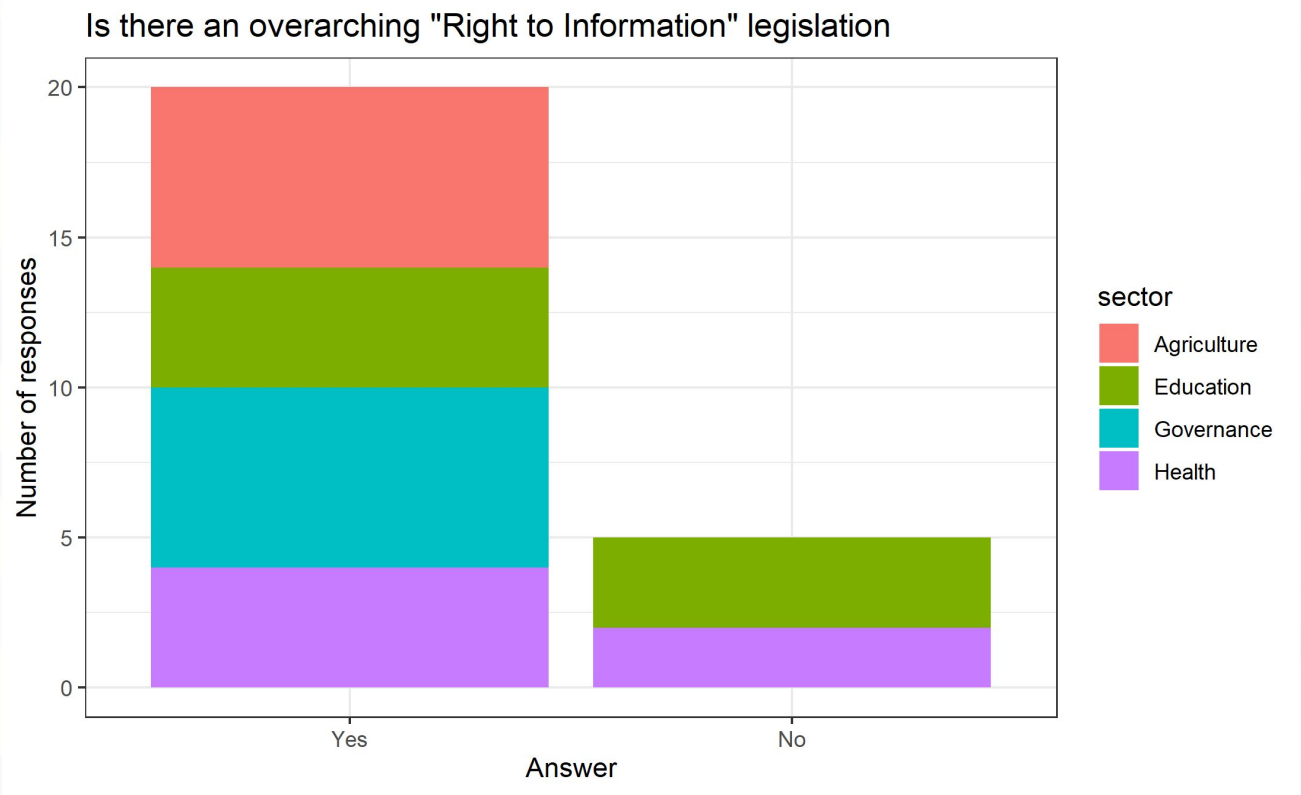


Education - Recommendations

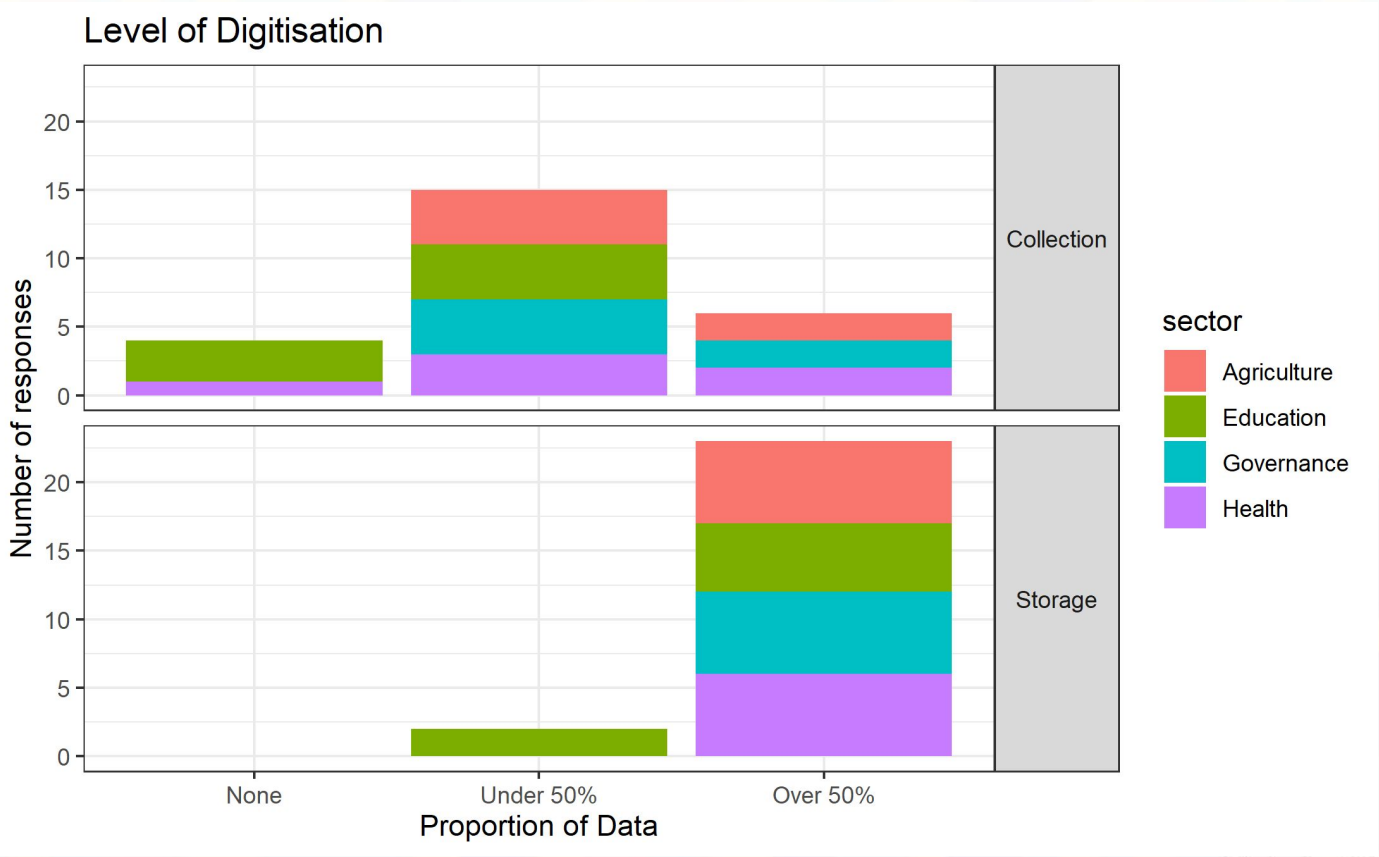
- County governments should collaborate more with private sector and non-governmental actors to collaborate on data collection and use; besides delivering on the education mandate. This should also be intentional when it comes to data collection for children/learners with special needs (the study reveals it is not frequently updated).
- The indicators tracked under this department can expand to include special needs education among others; so that data is refreshed periodically. This would improve outcomes in this area.

Cross-sectoral observations

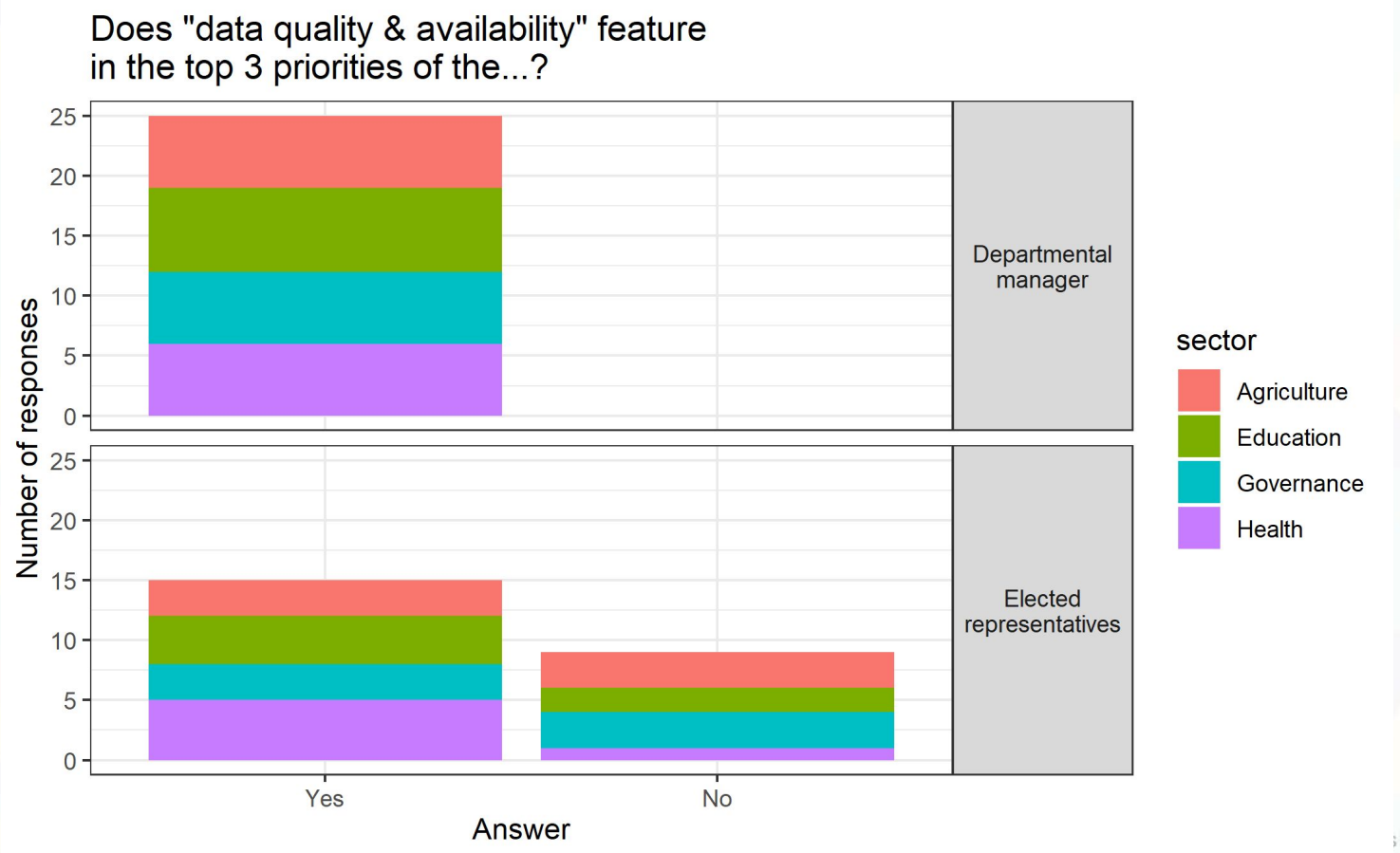
Right to information



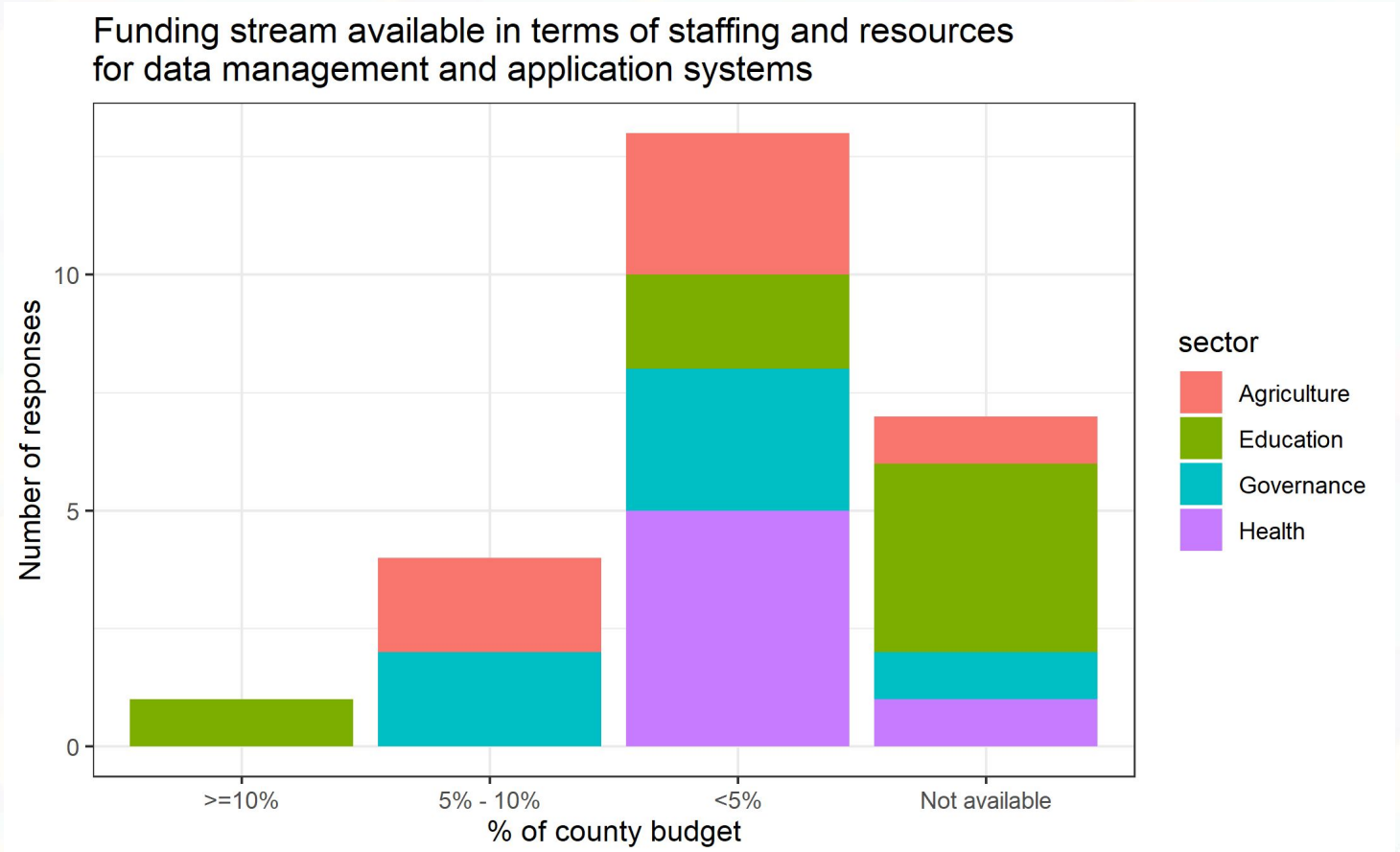
Digitisation



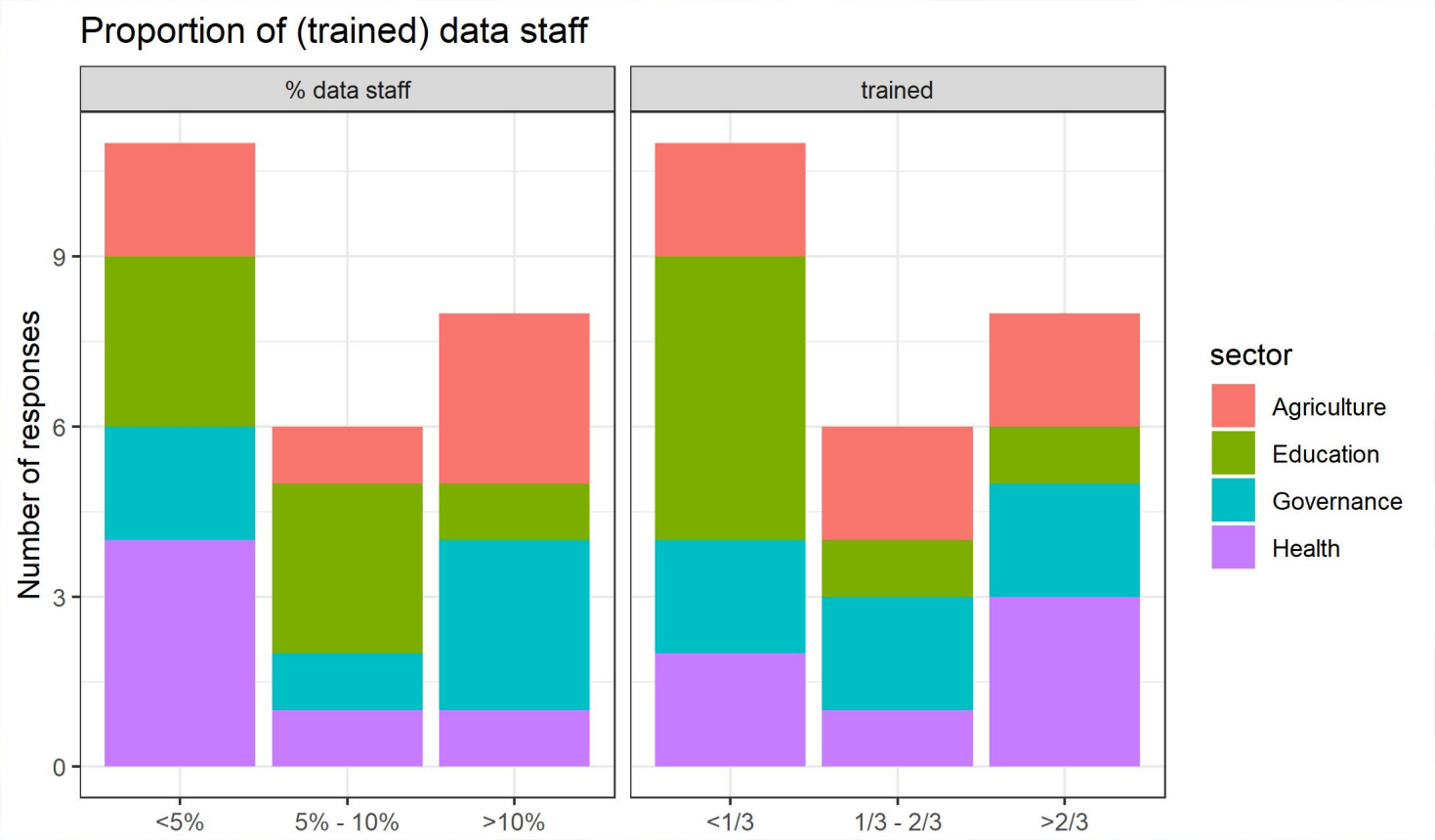
Priority


Does "data quality & availability" feature
in the top 3 priorities of the...?
Number of responses
25
20
15
10
5
0
Departmental manager
Elected representatives
Yes
No
Answer
sector
Agriculture
Education
Governance
Health

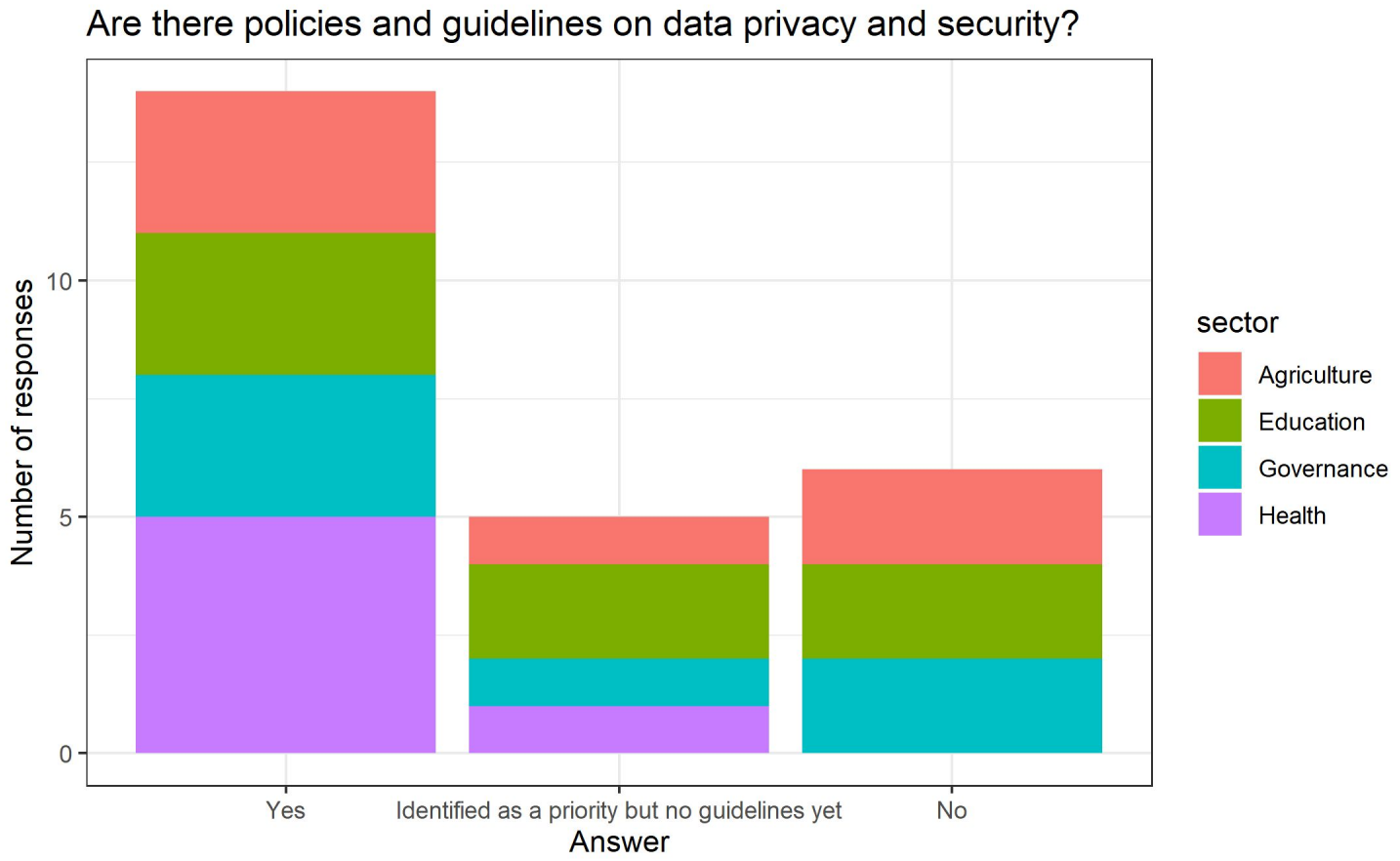
Funding


Funding stream available in terms of staffing and resources for data management and application systems
Number of responses
10
5
0
>=10%
5% - 10%
<5%
Not available
% of county budget
sector
Agriculture
Education
Governance
Health

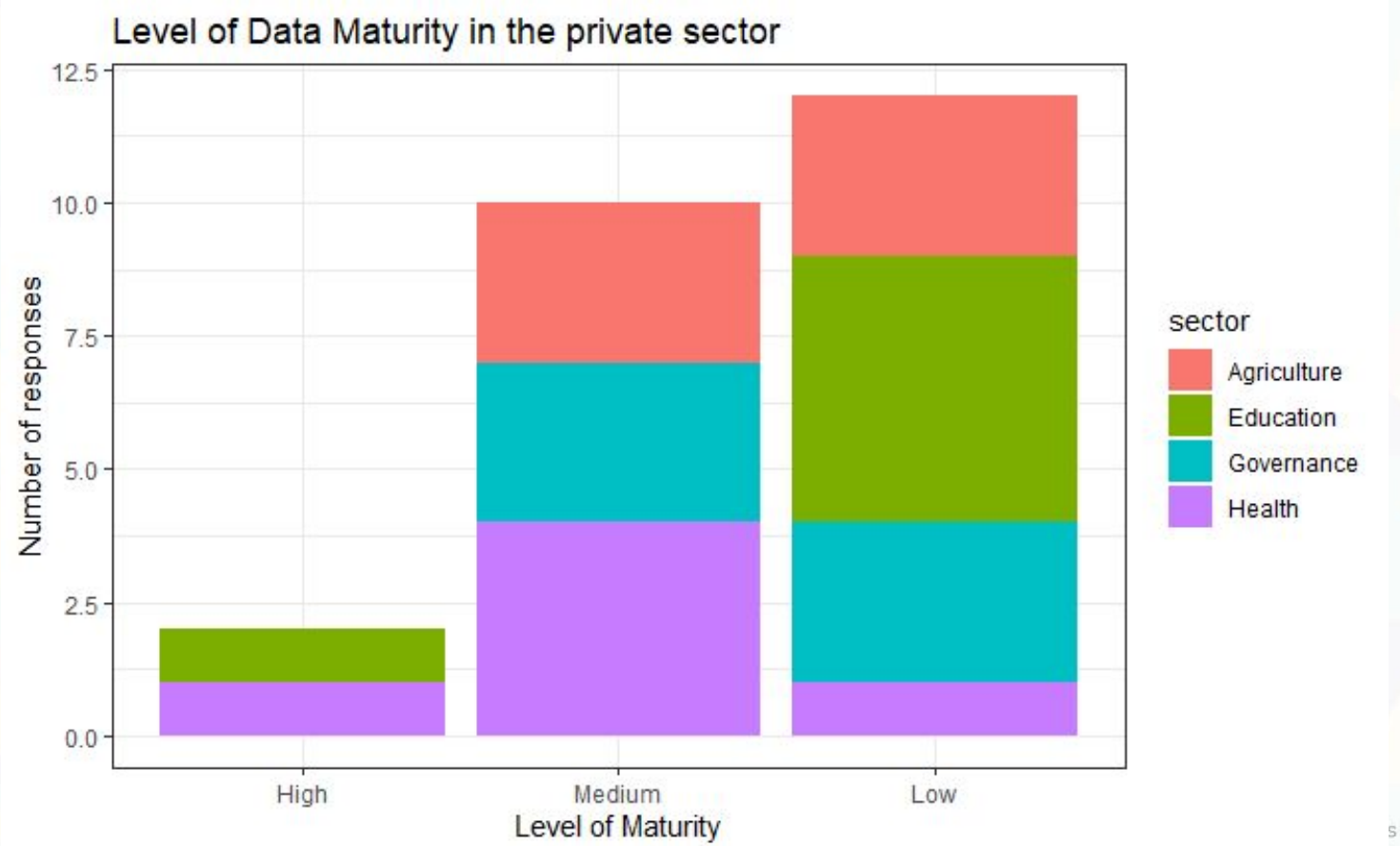
Staffing


Proportion of (trained) data staff
% data staff
trained
Number of responses
0
3
6
9
<5%
5% - 10%
>10%
<1/3
1/3 - 2/3
>2/3
sector
Agriculture
Education
Governance
Health

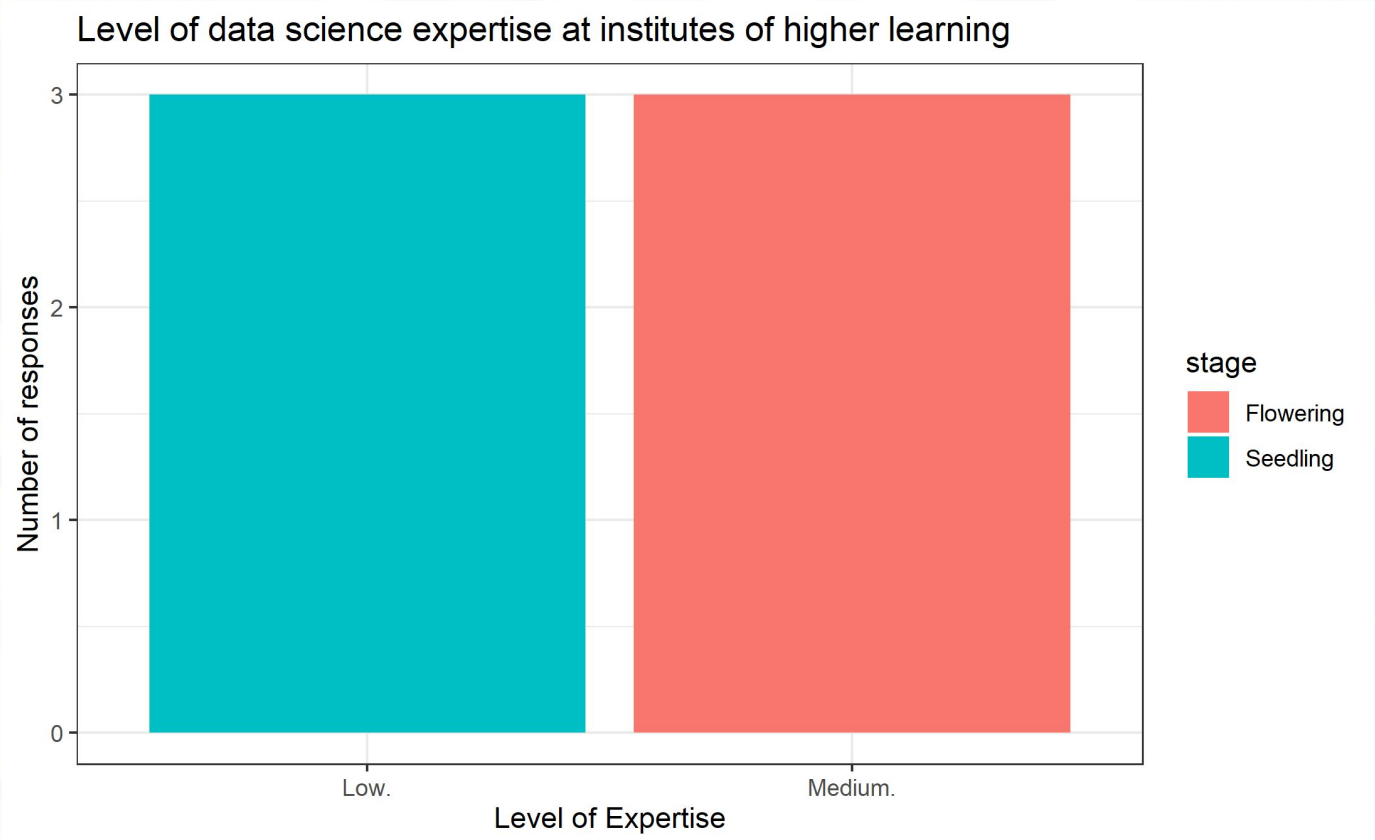
Data privacy and security



Private sector



Data science expertise



Insights - Data quality & availability

Data Collection:

- Most data is collected in physical forms at the facility/event level. Some exceptions to this are using the IFMIS system (that automates the procurement process) for finance data under governance and data collected by some non-governmental entities in different sectors.
- Figures for most indicators in the agriculture sector are calculated based on standard formulas due to the amount of resources required to visit every site.

Data Reporting

- Data is reported quarterly for most of the indicators in this study, and this seems to be the golden standard for all sectors.
- Primary data sources are collated at the sub-county or county level on a monthly basis for most indicators.
- County governments release quarterly and yearly reports on progress towards their targets in PDF format.

Insights - Data quality & availability

Data Availability

- Data was available for most indicators in this study and all departments have an administrative requirement for data collection.
- Data on feedback mechanisms is limited to collection, with hardly any tracing of the subsequent processes addressing the feedback.

Data Quality

- Location and gender data is available for most indicators **where it is relevant**, since the data is collected at the unit/facility level. *But more should be done towards attaining this level of disaggregation.*
- The national government (through Sector Technical Working Groups with KNBS) carries out validation exercises; more can be done to standardize/regularize this at county departmental levels eg. perhaps as part of the audits, etc.

Recommendations - Data quality & availability

Data Collection:
Data collection processes could be improved by digitising more processes to feed into live databases/dashboards.

Data Reporting
Reporting could be improved by adding dashboards and data in more usable formats for the public to conduct their own analyses.

Data Availability
County governments should monitor feedback mechanisms more keenly and actively seek feedback from citizens on their work.

Data Quality
County governments should strengthen data validation as well as data use so as to improve the quality of data that is reported from all sectors.

Insights - Causality

Mandate:

- County departments are required to regularly report on their progress according to their Annual Development Plans (ADPs) and the County Integrated Development Plans (CIDPs).
- However, penalties for not reporting some data are not well outlined for all the sectors.

Incentives and Accountability:

- Financial data is required for disbursement of funds needed to proceed with work as planned.
- Departmental managers prioritise data management but this is not always the case for the County Assembly.
- In most cases, the data is reported by the parties responsible for the outcomes, which could have an effect on data quality without data validation by a third party.

Insights - Causality

Resourcing:

- There is a low level of resourcing - both staff and material resources - specific to data in county governments.
- The governance and health sectors receive systemic data management support from the national government and this is reflected in their scores.
- Data science expertise is low in institutions of higher learning, although students learning related courses have been able to transfer their skills to applications in statistics and planning.

Coordination and Governance:

- County governments work with the national government to collect, store and disseminate data, mostly in the health and governance sectors.
- The data maturity in the private sector (and CSOs) is low, and there is minimal collaboration with county governments on data management.

Recommendations - Causality

Mandate:
Penalties for not reporting data should be clearly outlined and enforced in all sectors.

Incentives and Accountability:
More sensitisation is required to bring elective representatives on board. Helping them to understand the value of a data ecosystem to their ongoing work would incentivise them to champion related initiatives.

Resourcing:
County governments should collaborate with local institutions to build this capacity for their growing data management needs. Governments should also invest in data specific training for county data staff.

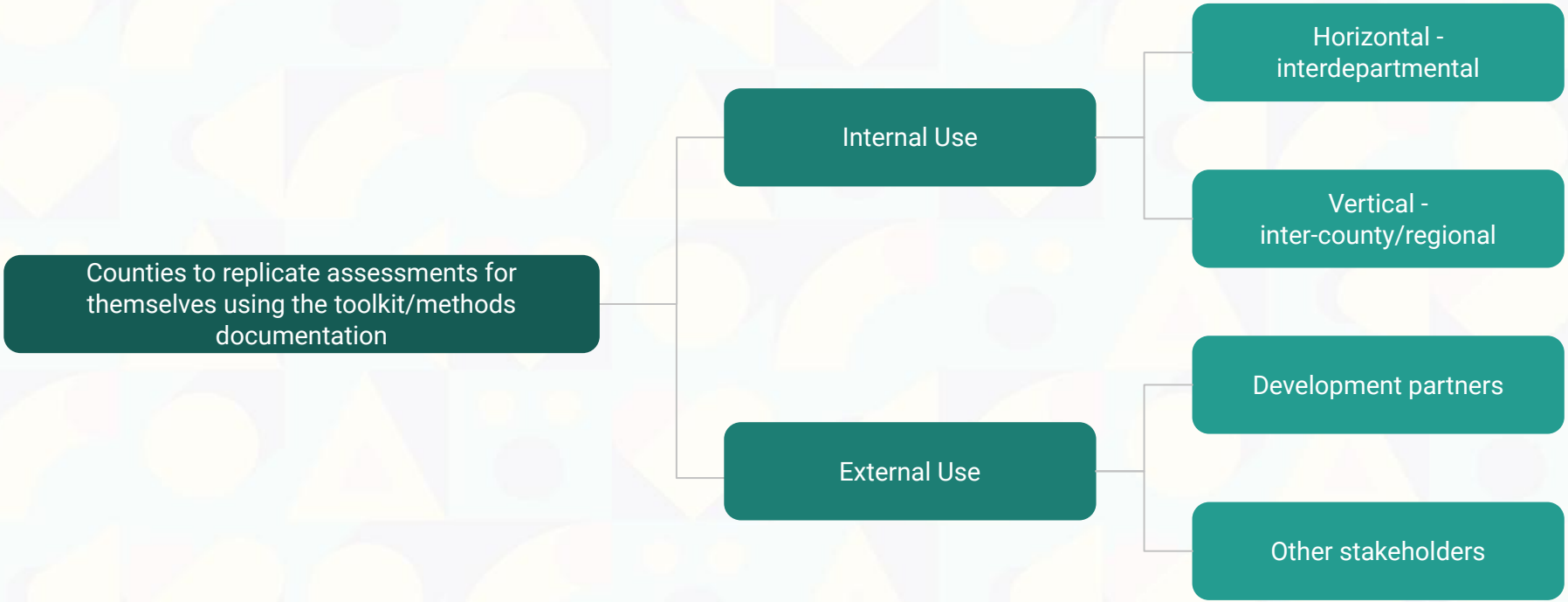
Coordination and Governance:
Departments where the private sector is heavily involved should encourage private entities (CSOs and academia) to collect data as part of their practices and collaborate with them for data collection and validation. They should not just collaborate on activities but also in the data value chain as well.

Recommendations - All counties

- County governments should invest in an open data system where they collect data on the same indicators and disseminate it in user-friendly formats for ease of access and use.
- County governments should invest in tools for automated data collection that forward data for analysis and reporting in real time. This is especially pertinent to departments where data is collected on the ground and on a regular basis such as education, health and agriculture.
- Actors in all sectors should work to ensure that the processes for feedback on the impact of work done by the county government are well documented at every level, to ensure that citizens have access to efficient solutions.

Next Steps...

How can your government use the toolkit?



About Qhala

We help organisations solve their most pressing problems, inspire growth and achieve high results, through digital transformation and embedding innovation into the company culture.

We build and launch digital and data platforms and ventures that solve key societal and economic problems.

A two-in-one company: innovation consulting and venture lab.

Comments/Questions?