Data quality dimensions

August 2021

INTRODUCTION

This overview document is supplementary to the Data Quality Framework Overview document and is intended to provide a more detailed level of support and direction to better understand the three quality dimensions of accuracy, completeness and timeliness.

Any relevant current industry guidance and case studies have been referenced, where they provide more detailed assistance.

WHAT ARE THE DEFINITIONS OF THE DATA QUALITY DIMENSIONS?

Definitions of the three data quality dimensions are provided in the below table.

Dimension	Definition
Accuracy	The data reflects the real-world object or event
Completeness	The data is comprehensive for where it is intended to be used
Timeliness	The data is available when expected and needed

WHY WERE THE DIMENSIONS SELECTED, AND WHY ARE THEY IMPORTANT?

The three quality dimensions were selected for the following reasons:

ACCURACY: The data we hold on our networks and the component assets needs to reflect their physical attributes, condition and demand/use. This is needed to develop evidence-based plans and programmes for their effective management. The metrics testing the accuracy dimension have been selected to interrogate the data we hold reflects the networks and assets we manage.

COMPLETENESS: An appropriate level of detailed understanding is needed for the planning, and reporting on, the effective management of our networks and their component assets. The metrics testing the completeness dimension have been selected to interrogate the level of data we hold is comprehensive for turning into information.

TIMELINESS: The data we use to inform our planning and decision-making processes and reporting needs to be available when these activities occur. The metrics testing the timeliness dimension have been selected to interrogate that data is available at the time it is needed.

KEY POINTS

Data quality dimensions:

✓ The data quality metrics interrogate the three quality dimensions of accuracy, completeness and timeliness

✓ Any improvement plan should be focused on fixing the root cause of the problem and not simply correcting the data record(s)

✓ Supports the development of evidence-based plans and programmes

✓ Supports robust reporting of ONRC Performance Measure results

✓ Consideration needs to be given to the level of quality for each dimension and the appropriate balance between them

Our asset data needs to reflect the networks we manage, to an appropriate level of detail and be available at the time it is needed.

REG THE ROAD EFFICIENCY GROUP

We are. LGNZ.
Te Kāhui Kaunihera ō Aotearoa.

WAKA KOTAHI
NZ TRANSPORT AGENCY

WHAT ACTION SHOULD BE TAKEN FOR RESULTS NOT AT THE EXPECTED STANDARD?

The metrics are intended to flag potential quality or process issues with an RCA's data process or practices. As you read the report, think about why the result may be as it is. What can be done to improve the data for those metrics where the result is not in the "green zone" (the expected standard) and is low when compared to the sector.

Further analysis is expected to determine the root cause, if any, producing a result below the expected standard. A data improvement plan can then be developed. Any actions identified in **THE IMPROVEMENT PLAN SHOULD BE FOCUSED ON FIXING THE ROOT CAUSE OF THE PROBLEM AND NOT SIMPLY CORRECTING THE DATA RECORD(S)**. This approach will achieve more enduring results by tackling the problem at a process and systems-level rather than a 'quick fix'.

EXAMPLES OF COMMON CAUSES FOR POOR DATA

Below are some common examples for a result not being at the expected standard and the quality dimension(s) they are likely to affect:

Cause	Accuracy	Completeness	Timeliness
Poor quality assurance/control associated with additions, updates and deletions in the asset inventory register	✓	✓	✓
Untrained staff collecting data and/or populating the asset inventory register	✓	✓	
Inappropriate data collection methodologies used	✓	✓	✓
Absence of an active data quality standard/manual clearly defining what is to be recorded, by whom and to what level	✓	✓	
Insufficient resources allocated to managing and maintaining the asset inventory register	✓	✓	✓
Poorly defined contract scope requirements for procured services	✓	✓	✓
Updating of asset inventory data not completed to meet key milestones			✓

COMPROMISES BETWEEN QUALITY DIMENSIONS

Consideration needs to be given to the level of quality for each dimension and the appropriate balance between them. For example, accuracy can be improved through very extensive quality assurance and validation processes. However, this will potentially be to the detriment of timeliness with the data not being available when required.

CONCLUSION

The data quality metrics test the quality dimensions of accuracy, completeness and timeliness. These were selected as the data we base our decisions and reporting on needs to reflect the networks we manage, to an appropriate level of detail and be available at the time it is needed.

Any data improvement plans should look to address the root cause of the problem and not simply correct the data record(s). This may include considering a balanced approach in achieving the appropriate quality across the three dimensions.

REFERENCES

Data quality project webpage
Data quality framework overview
Understanding the data quality results overview
Overall Asset Management Score Calculation
Metric library

REG is a collaborative project between Local Government and Waka Kotahi.

For more information, please contact:

Road Efficiency Group
RoadEfficiencyGroup@nzta.govt.nz