INTRODUCTION
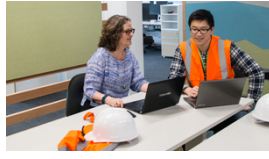
This overview document is intended to provide high level support and direction to better understand the intent and purpose of the data quality framework developed by the Road Efficiency Group (REG).

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

WHAT ARE THE DATA QUALITY FRAMEWORK AND REPORTS?

The data quality of each RCA including Teritorial Local Authorities, DoC and Waka Kotahi by State Highway Region is assessed annually against a suite of data quality metrics interrogating a cross-section of the data used to support our asset management activities.

Each RCA's results are compared against an expected standard and the distribution of all RCAs across New Zealand. For TLAs and DoC the comparison is against other TLAs, and for Waka Kotahi region the comparison is against the other State Highway regions. The results of the annual data quality reports have several uses. The main ones are:

- At a network level to support the identification of data quality improvement opportunities enabling evidence-based decision-making
- Annual monitoring and reporting on data quality at a national level supporting the identification and development of reference and guidance material to support the sector
- For Technical Auditors to reference and for Investment Advisors to consider alongside funding requests.

The intent is for the results to flag potential issues with how the data is being collected, managed and maintained. Further analysis is required on those results not at the expected standard to determine what action (if any) is required. **ANY IMPROVEMENT ACTION SHOULD BE TO ADDRESS THE ROOT CAUSE, RATHER THAN SIMPLY CORRECTING THE RECORD(S).**

WHY HAVE THE METRICS BEEN SELECTED AND WHY ARE THEY IMPORTANT TO ME?

Good quality data is necessary to effectively and efficiently deliver services and manage our assets. The intention is for the results to identify opportunities for improvement in the way both an individual network and the sector collects, manages and uses data to support our investment planning and decision-making processes.

Each metric has been selected to test a cross-section of the data looking at data quality in terms of accuracy, completeness and timeliness. The metrics have been selected as they:

- Have been considered an important input into our asset management decision-making processes
- Interrogate data that underpins the ONRC performance measures.

It is necessary to have confidence in the data when considering the results at an individual level and in the comparative reporting against peers.

KEY POINTS

Data quality framework:

- ✓ Good quality data is necessary to effectively and efficiently deliver services and manage our assets

- ✓ The data quality of each RCA; Local Authority, DoC and Waka Kotahi State Highway region, is assessed annually against a suite of data quality metrics interrogating a cross-section of the data used to support our asset management activities

- ✓ The intent is for the results to flag potential issues with how the data is being collected, managed and maintained

- ✓ Each metric has been selected to test a cross-section of the data looking at data quality in terms of accuracy, completeness and timeliness

- ✓ The result for each metric is considered against grade thresholds providing an indication of the level of quality

- ✓ The framework should not be used to replace quality control/assurance processes required with updates to the asset information system

- ✓ The focus of any improvement action by a RCA should be at the root cause, and not simply correcting the record(s)

- ✓ The annual national and RCA results are used to inform data improvement strategies and develop supporting programmes

The data quality framework and metrics interrogate a targeted sample of the data that supports our asset management decision-making.

REG THE ROAD EFFICIENCY GROUP

We are. LGNZ. Te Kāhui Kaunihera ō Aotearoa.

WAKA KOTAHI NZ TRANSPORT AGENCY

THE DATA QUALITY METRIC FRAMEWORK

The framework interrogates a TARGETED SAMPLE of each RCA's asset data. It is not a complete data quality check or validation of data managed and maintained by each RCA. IT SHOULD NOT BE USED TO REPLACE QUALITY CONTROL/ASSURANCE CHECKS required as part of adding, updating and deleting records in your asset database.

Each metric is associated with a cross-section of the data that supports our asset management processes, including a subset that underpins the ONRC performance measure results. The structure of the report is to group the metrics into data categories, level of importance and the data quality dimensions tested. The reported results at a metric level are aggregated up to these category and dimension levels.

Score

The data quality report has an overall Score. This is a weighted calculated score based on the importance of each metric applicable to an individual RCA. The maximum Score of 100 is achievable by having all metrics at the expected standard level. The current minimum achievable Score is 14. More detail on the SCORE CALCULATION can be found within the resources on the REG website.

Quality Dimensions

The metrics test the quality of a specific set of data against the quality dimensions of accuracy, completeness or timeliness. The table below defines each of these dimensions.

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

There is a subsequent DATA QUALITY DIMENSIONS OVERVIEW document to this, providing a more detailed view of the three data quality dimensions.

Importance

Each metric has been assigned a level of importance based on its likely impact on our asset management decision-making processes. There are three categories; **high**, **moderate** and **low**.

Expected Standard

The results for each metric is considered against grade thresholds proving and an indication of the level of quality. There are three grades; **at the expected standard**, **minor issues** and **major issues**.

Data Categories

There are six categories, each with one or multiple sub-categories interrogating a specific data set via a varying number of metrics:

- Network
- Asset Inventory
- Maintenance Activity
- Condition
- Demand/Use
- Crash

Metrics Not Relevant to a Network

A few of the metrics do not apply to all networks. For example, any metrics relating to the rural network on an urban only network. These metrics have a result of "NA" when not applicable and do not contribute to the aggregated results at a dimension or category level or the Score.

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

The results are intended to flag potential quality or process issues with an RCA's data process or practices. As you read the reports, think about why the result may be as it is. Consider 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 compared to the sector.

CONCLUSION

The data quality framework and metrics interrogate a targeted sample of the data that supports our asset management decision-making. Each metric has been selected to test a cross-section of the data looking at data quality in terms of accuracy, completeness and timeliness. It should not be used to replace quality control/assurance processes required with updates to the asset information system.

REFERENCES

Data quality project webpage
Data quality dimensions 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