
Addressing —

**Part 3:
Address data quality**

Adressage —

Partie 3: Qualité des données d'adresse

STANDARDSISO.COM : Click to view the full PDF of ISO 19160-3:2020



STANDARDSISO.COM : Click to view the full PDF of ISO 19160-3:2020



COPYRIGHT PROTECTED DOCUMENT

© ISO 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

	Page
Foreword	v
Introduction	vi
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Conformance	4
5 Symbols, notation, and abbreviated terms	4
5.1 Unified modelling language.....	4
5.2 Package abbreviated terms.....	4
6 Describing the quality of address data	5
7 Data quality elements for address data	6
8 Measuring the quality of address data	7
8.1 General.....	7
8.2 Completeness.....	7
8.2.1 General.....	7
8.2.2 Commission.....	8
8.2.3 Omission.....	8
8.3 Logical consistency.....	9
8.3.1 General.....	9
8.3.2 Conceptual consistency.....	9
8.3.3 Domain consistency.....	10
8.3.4 Format consistency.....	10
8.3.5 Topological consistency.....	10
8.4 Positional accuracy.....	10
8.4.1 General.....	10
8.4.2 Absolute or external positional accuracy.....	11
8.4.3 Relative or internal positional accuracy.....	11
8.4.4 Gridded data positional accuracy.....	11
8.5 Temporal quality.....	11
8.5.1 General.....	11
8.5.2 Accuracy of time measurement.....	11
8.5.3 Temporal consistency.....	12
8.5.4 Temporal validity.....	12
8.6 Thematic accuracy.....	12
8.6.1 General.....	12
8.6.2 Classification correctness.....	12
8.6.3 Non-quantitative attribute correctness.....	13
8.6.4 Quantitative attribute accuracy.....	13
8.7 Usability.....	14
9 Evaluation of the quality of address data	14
9.1 General.....	14
9.2 Evaluation process.....	14
9.2.1 Metaquality.....	15
9.3 Address data quality evaluation methods.....	16
10 Reporting the quality of address data	16
Annex A (normative) Abstract test suites	17
Annex B (informative) Methods for measuring address data quality	19
Annex C (informative) Examples of evaluating and reporting address data quality	23

Bibliography35

STANDARDSISO.COM : Click to view the full PDF of ISO 19160-3:2020

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular, the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 211, *Geographic information/Geomatics*.

A list of all parts in the ISO 19160 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Addresses provide one of the most common ways to unambiguously determine an object for purposes of identification and location and yet they can vary considerably from country to country. For example, in many Euro-centric countries reference to a road network in the address is common, while addresses in countries such as Japan comprise hierarchies of administrative areas without reference to thoroughfares. Addresses serve many purposes including postal delivery, emergency response, marketing, mapping, utility planning, and land administration to name a few (ISO 19160-1).

A wide variety of address standards and/or specifications are in use around the world. The preparatory work for this document, the review summary of the ISO 19160 stage zero project describes a number of these standards which are well integrated into various operational processes and in some cases legally enforced (ISO 19160-1). In addition, ISO 19160-1 has been developed as an international standard to facilitate interoperability between existing and future address specifications.

There are also many stakeholders involved in addressing: for address assignment (e.g. local governments, postal operators); for address use (e.g. customer service providers and electronic business, local and national governments, utility service providers, election commissions); and for finding addresses (e.g. citizens, delivery and emergency response service providers). Relevant stakeholders were identified during the preparatory work of the 19160 stage zero project as well as during the development of ISO 19160-1. These stakeholders are now either involved or aware of the development of the ISO 19160 addressing standards (ISO 19160-1).

In order to organize, maintain, and provide address records, address authorities must create master address repositories that replace the numerous, often isolated and incomplete address data files from different parts of an organization with one authoritative, integrated address database (e.g. US FGDC Address). Moreover, it is common for address data from local authorities to be aggregated at higher administrative levels (e.g. regional, state, national).

The wide range of uses for addresses, as well as the need to share and aggregate address data, requires a consistent framework for measuring and reporting the quality of address data. The goal of this document is to provide address database managers, address data aggregators, and address data users with this framework. This document also recognizes that the wide range of uses for addresses means that the data quality evaluation process utilized will depend on the use of the addresses being evaluated.

This document is a profile of ISO 19157 and is needed due to the uniqueness and complexity of addressing. The temporal nature of addresses, their components, and the real-world objects they identify can make it difficult to determine if addresses are missing or incorrectly included in a dataset. An address is also a complex collection of information with attributes and components that often reference other objects or are restricted by other components. These are just a few of the challenges faced by address data stakeholders that this document aims to overcome. Its objective is to ensure address data quality is maintained and/or improved by establishing a set of address data quality measures. Furthermore, this document can be used to understand and communicate the quality of an address dataset when address data aggregation or sharing needs to be accomplished.

Addressing —

Part 3: Address data quality

1 Scope

This document

- is a profile of ISO 19157;
- establishes a set of data quality elements and measures for describing the quality of address data;
- describes procedures for reporting data quality;
- provides guidelines for the use of the established set of data quality elements and measures for describing the quality of address data.

This document can be used by those evaluating and reporting the quality of address data such as address data managers, address data aggregators, and address data users.

This document does not attempt to define minimum acceptable levels of quality for address data.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO 19115-1:2014, *Geographic information — Metadata — Part 1: Fundamentals*

ISO 19115-2, *Geographic information — Metadata — Part 2: Extensions for acquisition and processing*

ISO 19157:2013, *Geographic information — Data quality*

ISO 19160-1:2015, *Addressing — Part 1: Conceptual model*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

3.1

accuracy

closeness of agreement between a test result or measurement result and the true value

[SOURCE: ISO 3534-2:2006, 3.3.1, modified — NOTES 1 to 3 have been removed.]

**3.2
address**

structured information that allows the unambiguous determination of an object for purposes of identification and location

[SOURCE: ISO 19160-1:2015, 4.1, modified — Notes 1 to 6 to entry and EXAMPLES 1 to 4 have been removed.]

**3.3
address alias**

one of a set of *addresses* (3.2) unambiguously determining the same addressable object (3.9)

[SOURCE: ISO 19160-1:2015, 4.3]

**3.4
address class**

description of a set of addresses (3.2) that share the same address components (3.5), operations, methods, relationships, and semantics

[SOURCE: ISO 19160-1:2015, 4.4, modified — EXAMPLES 1 and 2 have been removed.]

**3.5
address component**

constituent part of the *address* (3.2)

[SOURCE: ISO 19160-1:2015, 4.5, modified — Notes 1 and 2 to entry have been removed.]

**3.6
address position**

position representing the *address* (3.2)

[SOURCE: ISO 19160-1:2015, 4.7, modified — Note 1 to entry has been removed.]

**3.7
address reference system**

defined set of *address components* (3.5) and the rules for their combination into addresses (3.2)

[SOURCE: ISO 19160-1:2015, 4.8]

**3.8
addressing**

activities involving *addresses* (3.2)

[SOURCE: ISO 19160-1:2015, 4.6]

**3.9
addressable object**

object that may be assigned an *address* (3.2)

[SOURCE: ISO 19160-1:2015, 4.2]

**3.10
conformance**

fulfilment of specified requirements

[SOURCE: ISO 19105:2000, 3.8]

**3.11
correctness**

correspondence with the *universe of discourse* (3.25)

[SOURCE: ISO 19157:2013, 4.5]

3.12**data quality basic measure**

generic data *quality* (3.23) measure used as a basis for the creation of specific data quality measures

[SOURCE: ISO 19157:2013, 4.7, modified — Note 1 to entry has been removed.]

3.13**dataset**

identifiable collection of data

[SOURCE: ISO 19115-1:2014, 4.3, modified — Note 1 to entry has been removed.]

3.14**direct evaluation method**

method of evaluating the *quality* (3.23) of a *dataset* (3.13) based on inspection of the *items* (3.19) within the *dataset* (3.13)

[SOURCE: ISO 19157:2013, 4.10]

3.15**feature**

abstraction of real-world phenomena

[SOURCE: ISO 19101-1:2014, 4.1.11]

3.16**feature type**

class of *features* (3.15) having common characteristics

[SOURCE: ISO 19156:2011, 4.7]

3.17**geographic data**

data with implicit or explicit reference to a location relative to the Earth

[SOURCE: ISO 19109:2015, 4.13]

3.18**indirect evaluation method**

method of evaluating the *quality* (3.23) of a *dataset* (3.13) based on external knowledge

[SOURCE: ISO 19157:2013, 4.17, modified — Note 1 to entry has been removed.]

3.19**item**

anything that can be described and considered separately

[SOURCE: ISO 19157:2013, 4.18, modified — Note 1 to entry has been removed.]

3.20**metadata**

information about a resource

[SOURCE: ISO 19115-1:2014, 4.10]

3.21**metaquality**

information describing the *quality* (3.23) of data quality

[SOURCE: ISO 19157:2013, 4.20]

3.22

profile

set of one or more base standards or subsets of base standards, and, where applicable, the identification of chosen clauses, classes, options and parameters of those base standards, that are necessary for accomplishing a particular function

[SOURCE: ISO 19106:2004, 4.5, modified — The NOTE has been removed.]

3.23

quality

degree to which a set of inherent characteristics of an object fulfils requirements

[SOURCE: ISO 9000:2015, 3.6.2, modified — Notes 1 and 2 to entry have been removed.]

3.24

standalone quality report

free text document providing fully detailed information about data *quality* (3.23) evaluations, results and measures used

[SOURCE: ISO 19157:2013, 4.23]

3.25

universe of discourse

view of the real or hypothetical world that includes everything of interest

[SOURCE: ISO 19101-1:2014, 4.1.38]

4 Conformance

Any address product claiming conformance to this document shall pass all the requirements described in the abstract test suite presented in [Annex A](#) as follows:

- a) Address data quality shall be measured and described as outlined in [A.1](#) and [A.2](#);
- b) An address data quality measure shall pass the tests outlined in [A.3](#);
- c) An address data quality evaluation process shall pass the tests outlined in [A.4](#);
- d) Metadata used to report the quality of address data shall pass the tests outlined in [A.5](#);
- e) A standalone address data quality report shall pass the tests outlined in [A.6](#).

5 Symbols, notation, and abbreviated terms

5.1 Unified modelling language

In this document, conceptual schemas are presented in the Unified Modelling Language (UML). ISO 19103 Conceptual schema language presents the specific profile of UML used in this document.

5.2 Package abbreviated terms

Abbreviated terms are used to denote the package that contains a class. Those abbreviated terms precede class names, connected by a “_”. The international standard in which those classes are located is indicated in parentheses. A list of those abbreviated terms follows.

- DQ Data Quality [ISO 19157]
- MD Metadata [ISO 19115-1]
- CI Citation [ISO 19115-1]

6 Describing the quality of address data

This document is a profile of ISO 19157. The quality of address data is expressed by data quality elements which are described by established data quality measures, the evaluation processes for carrying out those measures, and the results of the evaluations. In addition, it is possible in some cases to describe the quality of the measures or evaluation processes themselves. Finally, address data quality is reported as metadata or in standalone quality reports.

Figure 1 provides an overview of quality for geographic data, which also applies to address data.

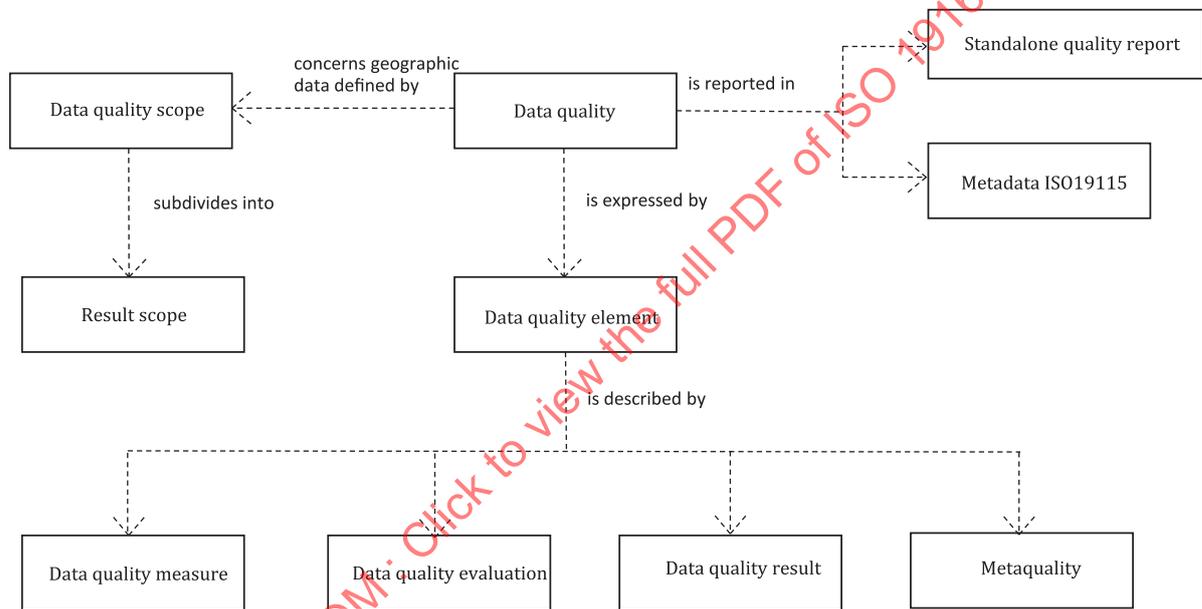


Figure 1 — Conceptual model of quality for geographic data (Source: ISO 19157:2013)

Different quality elements and different subsets of the data may be considered when describing the quality of geographic data. In order to describe these, data quality units are used. A data quality unit is the combination of a scope and data quality elements. Figure 2 describes a data quality unit.



Figure 2 — Data quality unit (Source: ISO 19157:2013)

The concept of a data quality unit is important for measuring the quality of address data. Address datasets often include data from multiple sources as well as different types of addresses (known as address classes in ISO 19160-1). For example, an address data aggregator could be managing addresses used solely for mailing as well as addresses used only for location purposes. Data quality measures appropriate for one use or classification may not be appropriate for the other. In this case, the data quality unit establishes the scope (e.g. postal addresses vs. emergency services, post boxes vs. thoroughfare addresses) and the element being measured.

Requirement 1. The quality of address data shall be described per data quality unit and data quality element, in conformance with ISO 19157.

7 Data quality elements for address data

This clause establishes the set of data quality elements from ISO 19157 for describing the quality of address data. The quality of address data is described according to data quality elements defined in ISO 19157:

- completeness;
- logical consistency;
- positional accuracy;
- temporal quality;
- thematic accuracy;
- usability.

The quality elements of address data are identical to ISO 19157 and are shown in [Figure 3](#).

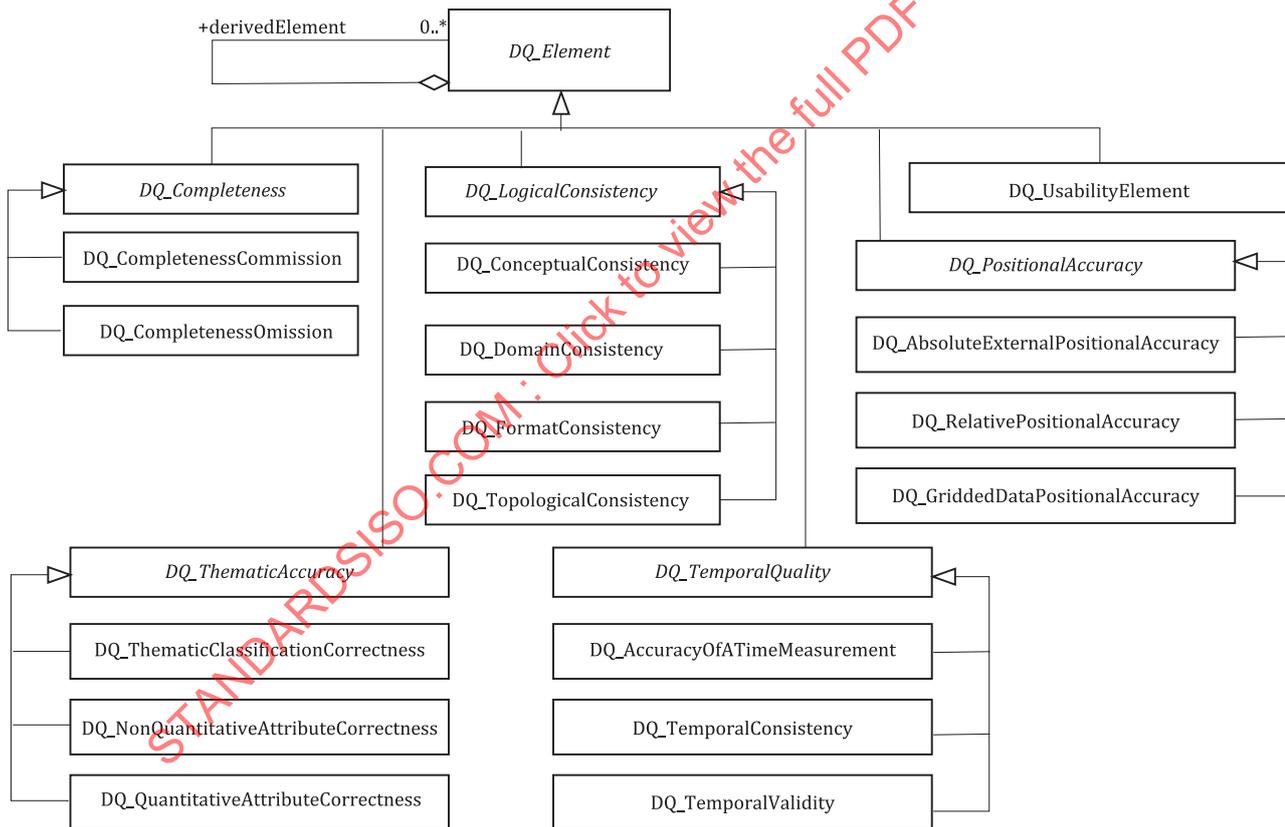


Figure 3 — Quality elements of address data (Source: ISO 19157:2013)

8 Measuring the quality of address data

8.1 General

This clause establishes measures that are relevant to address data. According to ISO 19160-1, address data consists of the following general classes:

- a) Address;
- b) Address component;
- c) Addressable object;
- d) Reference object.

Within these classes, errors may occur at one or more of the following levels:

- item (e.g. an address or component missing or incorrectly present in a dataset);
- attribute (e.g. an incorrect primitive type or value);
- relationship (e.g. an incorrect or missing relationship between addresses, address classes, or address components).

Requirement 2. A data quality measure for address data shall test for errors at the item, attribute, and/or relationship levels for one or more general address classes (Address, Component, Addressable Object and Reference Object) defined in ISO 19160-1:2015, 6.2.

The measures for quality elements of address data are identical to the standardized data quality measures in ISO 19157. However, due to the nature of data quality and geographic data, this document recognizes that a standardized list of data quality measures cannot be complete. It may be necessary for an address data quality evaluator to define additional data quality measures. If so, these measures shall be described according to the structure given in ISO 19157.

Requirement 3. Any measure specific to addressing shall be described according to the components defined in ISO 19157.

8.2 Completeness

8.2.1 General

Completeness is defined (ISO 19157) as the presence and absence of features, their attributes and relationships. It consists of two data quality elements:

- a) commission: excess data present in a data set;
- b) omission: data absent from a data set.

Completeness errors may be reported as:

- individual errors;
- total number of errors;
- rate of error (number of errors in relation to the number of relevant items).

Completeness should mainly be used at the general address class (Address, Component, Addressable Object and Reference Object) level. While it is possible to measure completeness of attributes of addresses or relationships between address classes, it is recommended that careful consideration first be given to logical consistency before using completeness for this purpose (e.g. using the logical consistency measure to identify a child address that has a relationship to more than one parent address).

Methods for measuring completeness of address data are provided in [Annex B](#).

8.2.2 Commission

Measuring for commission in address data involves determining if there are excess items in the general address classes or the attributes and relationships of those items. There are several approaches to measuring for commission in address data such as checking for items that are out of scope based on product specification.

EXAMPLE 1 Address positions are checked against specified administrative boundaries for excess items (i.e. if address positions outside specified boundaries are present in the dataset).

EXAMPLE 2 An address dataset is tested for excess items to ensure that only specified addressable object types are present.

The conceptual schema of the dataset may also be used to check for excess items.

EXAMPLE 3 Address component values are validated with the reference object to which the component references to ensure there are no excess values present.

Checking for duplicate items is another approach to measuring for commission in an address dataset that should be undertaken.

EXAMPLE 4 Inspection for duplicate address records is carried out by exact matching the components and attributes of each address to all other addresses in the dataset.

It is also possible to conduct non-quantitative commission checks by comparing the number of items present in an address dataset to expectations based on other sources. Non-quantitative commission checks may only indicate potential completeness errors and often require further investigation (e.g. interactive review, field survey).

EXAMPLE 5 Record counts in a residential address dataset of a city are compared to population counts and population per housing unit estimates obtained from a recent survey to determine if the number of records seems excessive given the population/population per unit estimates of the city (i.e. potential excess items).

EXAMPLE 6 A count of recently added records to a jurisdiction's address dataset is compared to housing permits issued in that jurisdiction to determine if the number of added records seems excessive given the number of permits issued (i.e. potential excess items).

8.2.3 Omission

Measuring for omission in address data involves determining if there are missing items in the general address classes or the attributes and relationships of those items. Checking for missing items may be done by:

- a) Checking that all items that are in scope based on product specification are included in the dataset.

EXAMPLE 1 An address dataset is tested for missing items by validating that all expected addressable object types are present.

- b) Checking that all mandatory components, attributes, and relationships based on the conceptual schema of the dataset are included.

EXAMPLE 2 Address component values are validated with the reference object to which the component references to ensure that all expected values are present.

- c) Checking that all items present in a source known to be correct are included in the dataset.

EXAMPLE 3 Address component values are validated with the reference object to which the component references to ensure that all expected values are present.

- d) Comparing the number of items to expectations based on other sources (non-quantitative).

EXAMPLE 4 A count of records in a jurisdiction's address dataset is compared to housing permits issued in that jurisdiction to determine if the number of records seems exceedingly low given the number of permits issued (i.e. potential missing items).

8.3 Logical consistency

8.3.1 General

Logical consistency is defined in ISO 19157 as the degree of adherence to logical rules of data structure, attribution, and relationships. It consists of four data quality elements:

- conceptual consistency: adherence to rules of the conceptual schema;
- domain consistency: adherence of values to the value domains;
- format consistency: degree to which data are stored in accordance with the physical structure of the data set;
- topological consistency: correctness of the explicitly encoded topological characteristics of a data set.

Evaluating logical consistency of a dataset requires knowledge of the logical rules (conceptual, logical, and physical) that apply to the dataset. This document uses the conceptual model in ISO 19160-1 to establish how logical consistency is applied to address data.

8.3.2 Conceptual consistency

Measuring conceptual consistency involves determining whether or not an item, attribute, or relationship adheres to the rules of the conceptual schema of the address dataset being evaluated.

EXAMPLE 1 All addresses in a dataset are checked to ensure that they have at least one address component.

EXAMPLE 2 All address aliases in a dataset are checked to ensure that they have an AddressAliasType.

EXAMPLE 3 All child addresses in a dataset are checked to ensure that they have only one parent address identified in the dataset.

Address assignment rules are often used by addressing authorities to assign addresses to locations or objects. These rules may apply at a national, regional, or other administrative level. Due to the variation of address assignment schemes around the world, defining address rules is outside the scope of this document. However, checking for adherence to address assignment rules when they exist is an important part of address data quality.

EXAMPLE 4 Address numbers are checked for conformance to thoroughfare side/parity rules.

EXAMPLE 5 Thoroughfare name types of addresses in a dataset are consistent with the cardinal direction rules (e.g. all boulevards must run east/west) of the reference object.

Other data quality elements (domain consistency, positional accuracy, thematic accuracy) may also be considered when checking for adherence to address assignment rules. These other data quality elements should be considered when the address assignment rules are known but are not included in the conceptual schema (e.g. address assignment rules provided as metadata of the address dataset).

Some address components relate to other components and these relationships can be tested for conformance using conceptual consistency measures.

EXAMPLE 6 Box numbers could be in scope of a post office. If each box number at a post office is unique, the post office component is the scopeComponent (superordinate component) and the box number is the valueComponent (subordinate component).

Address component relationships may be topological in nature. These constraints may be expressed explicitly or implicitly in the conceptual model of an address dataset. In either case, conceptual consistency measures can be used to test for conformance to the conceptual model.

EXAMPLE 7 Checking that the city associated with an address is found within the associated state.

EXAMPLE 8 Checking that the thoroughfare name associated with an address is found within the associated municipality.

8.3.3 Domain consistency

Address data includes several valid value domains that establish the set of accepted values for address components and attributes. Some value domains may be found in the conceptual schema as codelists (see ISO 19160-1 for Core and Lifecycle codelists) or valid datatypes for specific fields (e.g. preferenceLevel: Integer) while others may be established through product specification (e.g. geographic or administrative scope) or with locally defined values. Domain consistency measures are used to check that address data is consistent with the value domains established by the conceptual schema or otherwise specified. However, even without these specified domains, domain consistency can still be used to indicate errors (e.g. finding country names in the data that do not exist in the real world) or for indicating potential errors (e.g. values in a field that are inconsistent with the dataset as a whole).

EXAMPLE 1 AddressStatus is checked for values other than 'unknown', 'official' or 'unofficial'.

EXAMPLE 2 AddressComponentType thoroughfareName is validated against its reference object (e.g. street centerline file).

EXAMPLE 3 AddressComponentType postCode is verified with the postOfficeName associated with the address.

8.3.4 Format consistency

Address data may be stored and shared in many different formats (e.g. rDBMS, text, XML). Product specifications state the required format while address dataset specifications state the supposed actual file format. Format consistency can be used to check that the address dataset is in the format stated by its specification or the format required by a product specification. It is also possible to measure format consistency of attributes of address data (e.g. attribute data types).

EXAMPLE 1 The file type of an address dataset is confirmed to be in the format required by a product specification.

EXAMPLE 2 A 5-digit, numeric post code field is verified that it contains only five numbers.

8.3.5 Topological consistency

Topological constraints in address data are typically defined (explicitly or implicitly) in the conceptual schema of the address dataset and should be reported as conceptual consistency. The topological consistency measures in ISO 19157 generally do not apply to address data. It may be necessary for an address data quality evaluator to define additional data quality measures. If so, these measures shall be described according to the structure given in ISO 19157.

8.4 Positional accuracy

8.4.1 General

Positional accuracy is defined as the accuracy of the position of features within a spatial reference system. It consists of three data quality elements:

- absolute or external positional accuracy: closeness of reported coordinate values to values accepted as or being true;

- relative or internal positional accuracy: closeness of the relative positions of features in a dataset to their respective relative positions accepted as or being true;
- gridded data positional accuracy: closeness of gridded data spatial position values to values accepted as or being true.

8.4.2 Absolute or external positional accuracy

In most cases, addresses are represented as points on maps or in mapping applications. The positional accuracy of an address point or an addressable object point may be checked against an external source that is considered to be accurate, for example source mapping or by field survey. The differences between such pairs of measurements gives the absolute positional accuracy, which may be recorded as a root mean square error, or as the number of errors in excess of a pre-defined level (i.e. gross errors).

EXAMPLE 1 An address position with an AddressPositionType of “parcel centroid” is checked against a reference dataset (parcel layer) that has parcel centroids accepted as being true.

EXAMPLE 2 The position of an addressable object with AddressPositionType of “front door” is verified in the real world with survey tools against the position of the real world object (front door).

8.4.3 Relative or internal positional accuracy

Relative positional accuracy (the relative positioning of points relative to one another), is not usually applicable to address data, but could be used for positional/coordinate values relative to addresses within the same dataset that exhibit the same address components (e.g. similarly numbered properties within the same thoroughfare).

EXAMPLE Address numbers along a thoroughfare that are assigned in sequential order can be checked that the position of each is correct in relation to the others (e.g. horizontal position of address ‘5 Cherbourg Ave’ is between the horizontal positions of addresses ‘3 Cherbourg Ave’ and ‘7 Cherbourg Ave’).

8.4.4 Gridded data positional accuracy

Address data is not usually held as gridded data, therefore gridded data positional accuracy is not relevant for address data quality.

8.5 Temporal quality

8.5.1 General

Temporal quality is defined in ISO 19157 as the quality of the temporal attributes and temporal relationships of features. It consists of three data quality elements:

- accuracy of a time measurement: closeness of reported time measurements to values accepted as or known to be true;
- temporal consistency: correctness of the order of events;
- temporal validity: validity of data with respect to time.

NOTE The date/time format is defined in ISO 8601-1.

8.5.2 Accuracy of time measurement

The point or period of time at which an address, addressable object, or address component is inserted into or superseded in an address database may be captured for data management purposes (i.e. beginLifespan and endLifespan). The time a real world addressable object becomes valid or invalid may also be captured (i.e. validFrom and validTo). Accuracy of time measurements can be used to check for the closeness of these times compared to what is known to be true.

8.5.3 Temporal consistency

According to ISO 19160-1, addresses, address components, and addressable objects all have the attributes lifecycle and lifecycleStage. The attributes of the Lifecycle type can be checked using temporal consistency measures. The temporal relationships between classes may be tested as well.

EXAMPLE 1 The validFrom date of an address is confirmed to be earlier than its validTo date.

EXAMPLE 2 The beginLifecycle date of an addressable object is confirmed to be earlier than the date from which it is unambiguously determined by its associated address (i.e. addressedFrom).

8.5.4 Temporal validity

Any address data within the DateTime domain can be checked with temporal validity measures.

EXAMPLE 2018-03-21 is a valid date.

The rules describing the validity of data with respect to time may be part of the conceptual schema. It may be reported either as temporal validity or as conceptual consistency if the rules are part of the conceptual schema.

8.6 Thematic accuracy

8.6.1 General

Thematic accuracy is defined (ISO 19157) as the accuracy of quantitative attributes and the correctness of non-quantitative attributes and of the classifications of features and their relationships. It consists of three data quality elements:

- a) classification correctness: comparison of the classes assigned to features or their attributes to a universe of discourse (e.g. ground truth or reference data);
- b) non-quantitative attribute correctness: measure of whether a non-quantitative attribute is correct or incorrect;
- c) quantitative attribute accuracy: closeness of the value of a quantitative attribute to a value accepted as or known to be true.

Thematic accuracy errors may be reported as:

- errors (individual);
- number of errors (total);
- rate of error (number of errors divided by number of items, which may be expressed as a fraction or as a percentage).

8.6.2 Classification correctness

Determining classification correctness for address data means checking that all items (occurrences) for each object class are classified correctly.

The object classes in ISO 19160-1 and corresponding classifications are shown in Table 1.

Table 1 — Classifications of address object classes in ISO 19160-1

Object class	Object classification
Address	Address Class
Address Alias	Address Alias Type
Address Component	Address Component Type

Table 1 (continued)

Object class	Object classification
Addressable Object	Addressable Object Type
Reference Object	Reference Object Type

Checking these may be done by:

- a) Checking values for reasonableness, i.e. are values different from those that are to be expected from the scope of the dataset (e.g. are there Addresses with a different Address Class from that specified in the dataset scope?);

EXAMPLE 1 The value of an attribute falls outside of the domain specified. For example, an address is classified as being of Address Class "basic address" when no such class of Address exists.

- b) Checking values against an external source of data that is considered to be true, e.g. from field observation or survey, or other records.

EXAMPLE 2 An addressable object classified as a building is determined through field observation to be a vacant land parcel.

8.6.3 Non-quantitative attribute correctness

For address data, non-quantitative attribute correctness means, for each object class included in the dataset (in ISO 19160-1, Address, Address Component, Addressable Object, Reference Object), checking if all attributes for each occurrence are correct.

Checking this may be done by:

- a) Checking values for reasonableness, i.e. are values different from those that are to be expected from the scope of the dataset.

EXAMPLE 1 An address has a component thoroughfare name that does not exist in the corresponding locality or administrative area.

EXAMPLE 2 There are slight differences in spelt names (e.g. of thoroughfares, localities) which are likely to be typographic errors.

EXAMPLE 3 The value of an address component does not match with the expected value based on the value of other address components. For example, an address has a road name of "Mildred Road", however its number value is consistent with the address range for "William Avenue".

- b) Checking values against an external source of data that is considered to be true, e.g. from field observation or survey or other records.

EXAMPLE 1 The AddressComponentType is of a certain value, but the ReferenceObject dataset type does not match with that value. For example, an address contains a component called "suburb", however, the related ReferenceObject dataset is called "city".

EXAMPLE 2 Imagery used against address to check if the addressable object exists.

EXAMPLE 3 An address position with an AddressPositionType of "front door" is confirmed through imagery to be at the front door of the addressable object.

8.6.4 Quantitative attribute accuracy

Quantitative attributes in address data are rare. However, the conceptual model in ISO 19160-1 includes the quantitative attribute *preferenceLevel* found in the Address class and in the Core type *AddressComponentValue*. These attributes and those found in profiles of ISO 19160-1 with quantitative data types (e.g. integer) can be checked using quantitative attribute accuracy measures.

EXAMPLE An address in an address dataset has a *preferenceLevel* of 2 but it does not have a related (alias) address with *preferenceLevel* of 1.

8.7 Usability

The usability data quality element may be used to indicate an address dataset's suitability for a particular application. It may also be used to declare an address dataset's conformance to a specification. This is accomplished by aggregating several quality results and reporting them as a single result for the usability element.

EXAMPLE 1 An address authority declares their dataset suitable for mailing purposes.

EXAMPLE 2 An address dataset provider indicates that all requirements in the referred specification are fulfilled.

The usability element may also be used for a quality evaluation based on user requirements which cannot be covered by the other data quality elements.

EXAMPLE 3 An evaluation method is developed that assigns a confidence score based on levels of quality.

See [C.5](#) for an example of aggregating several quality results and reporting them as a single result for the usability element.

9 Evaluation of the quality of address data

9.1 General

Evaluating the quality of address data requires identifying the applicable elements and measures for the dataset being evaluated. Address data quality elements provide a means for determining how well an address dataset meets the requirements stated in its data product specification or user requirements. Not all address data quality elements are applicable to all types of address datasets. Some are more applicable at the file level (e.g. format consistency), while others are more applicable at the instance level (e.g. non-quantitative attribute correctness). Furthermore, address data quality elements are often interrelated. For example, an incorrect place name on an address record could generate multiple errors (e.g. completeness, domain consistency, thematic accuracy). The address data quality evaluator is responsible for identifying the applicable elements and measures for the dataset being evaluated while adhering to the following rule:

- for an address dataset with a data product specification or user requirements, all data quality elements and measures necessary to determine conformance to the data product specification or user requirements shall be considered applicable.

Requirement 4. All applicable address data quality measures shall be applied to the identified address data quality units.

Requirement 5. The evaluation of a data quality element of address data shall be described according to a measure, evaluation method, and result, in conformance with ISO 19157.

9.2 Evaluation process

Requirement 6. An evaluation process for address data shall be defined according to ISO 19157.

The ISO 19157 data quality evaluation process is described in [Figure 4](#).

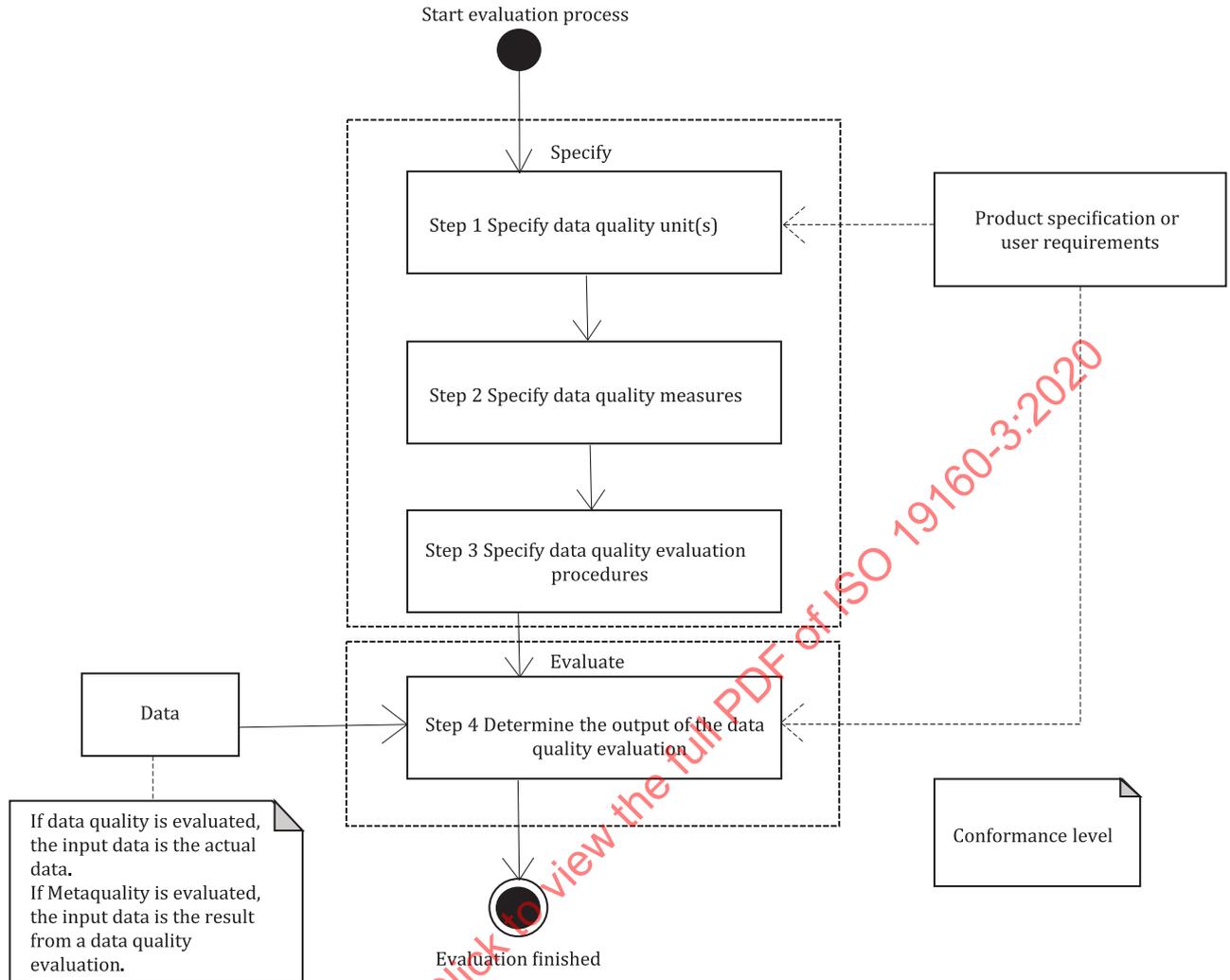


Figure 4 — Data quality evaluation process (Source: ISO 19157:2013)

9.2.1 Metaquality

According to ISO 19157, metaquality is a set of statements about a quality evaluation and its result. Knowledge about the quality and suitability of the method, measure, and given result of an evaluation process may be just as important as the result itself. Metaquality may be described with the following elements:

- confidence: trustworthiness of a data quality result;
- representivity: degree to which the sample used has produced a result which is representative of the data within the data quality scope;
- homogeneity: expected or tested uniformity of the results obtained for a data quality evaluation.

EXAMPLE A dataset consisting of the addresses in all suburban areas of a city is checked for commission using a sample of addresses (a subset of suburban areas). The evaluation method includes several trained field representatives canvassing the areas and collecting address information. The field work includes quality control methods. Metaquality is reported as high confidence (field work) and homogeneity (e.g. comparison of all field work shows low root mean square errors). The sample is also considered to be representative due to geography, zoning laws, time periods of data collection, etc.

A metaquality element is described by the same descriptors as for the quality element (measure, evaluation method and result) with an additional descriptor, related quality element. The related quality element is the element on which the metaquality element applies.

9.3 Address data quality evaluation methods

The data quality evaluation method describes those procedures which are applied to the geographic data to arrive at a data quality result. Different evaluations are often used for the various data quality elements. A data quality evaluation method should be included for each applied data quality measure.

Data quality evaluation methods can be divided into two classes: direct and indirect. Additional results may be produced through aggregation or derivation of existing results (see ISO 19157:2013, 9.3).

A direct evaluation method is a method of evaluating the quality of a dataset based on inspection of the items within the dataset and can be further classified as internal or external. Direct evaluation can be based on a full data survey or on a sample of a dataset.

- Checking for adherence to a conceptual schema, the correct use of data types, and classification correctness are examples of internal direct evaluation.
- Utilizing parcel or building footprint geometry to check the accuracy of an address position or conducting field checks are examples of external direct evaluation.

External direct evaluation checks rely on the quality of reference data being sufficient enough to meet the purpose of the checks. Measuring the quality of external data is outside the scope of this document.

An indirect evaluation method is a method of evaluating the quality of a dataset based on external knowledge or experience of the data product and can be subjective.

Requirement 7. An evaluation method for address data shall be defined according to ISO 19157.

See [Annex C](#) for examples of evaluating and reporting address data quality.

Requirement 8. At least one data quality result shall be provided for each data quality element. This could be a quantitative result, a conformance result, a descriptive result or a coverage result.

Results are described in detail in ISO 19157:2013, 7.4.4.

10 Reporting the quality of address data

Requirement 9. The quality of address data shall be reported as metadata and standalone quality reports as specified in ISO 19115-1 and ISO 19115-2.

Requirement 10. A standalone address data quality report shall include all relevant components as specified in ISO 19157.

See [Annex C](#) for examples of evaluating and reporting address data quality.

Annex A (normative)

Abstract test suites

A.1 Test case identifier: Address data quality units and elements

- a) Test purpose: To verify that ISO 19160-3 (this document), requirement 1 is met.
- b) Test method: Check whether the quality of address data is described per data quality units and data quality elements in conformance with ISO 19157.
- c) Reference: ISO 19157:2013, 7.2, 7.3.
- d) Test type: Basic.

A.2 Test case identifier: Application of address data quality measures

- a) Test purpose: To verify that ISO 19160-3 (this document), requirements 2 and 5 are met.
- b) Test method: Check that all applicable address data quality measures are applied to the identified address data quality units and that these measures test for errors at the item, attribute, and/or relationship levels for all general address classes defined in ISO 19160-1.
- c) Reference: [Clauses 7](#) and [9.1](#).
- d) Test type: Basic.

A.3 Test case identifier: Description of address data quality measures

- a) Test purpose: To verify that that ISO 19160-3 (this document), requirement 3 is met.
- b) Test method: Check abstract test suite provided in ISO 19157:2013, A.5.a, A.5.b, A.5.c, A.5.d.
- c) Reference: ISO 19157:2013, Clause 8 and Annex C.
- d) Test type: Basic.

A.4 Test case identifier: Address data quality evaluation process

- a) Test purpose: To verify that ISO 19160-3 (this document), requirements 4, 6, 7 and 8 are met.
- b) Test method: Check abstract test suite provided in ISO 19157:2013, A.1.a, A.1.b, A.1.c, A.1.d.
- c) Reference: ISO 19157:2013, 9.1.
- d) Test type: Basic.

A.5 Test case identifier: Reporting address data quality as metadata

- a) Test purpose: To verify that that ISO 19160-3 (this document), requirement 9 is met.
- b) Test method: Check abstract test suites provided in ISO 19115-1:2014, A.2.1, A.2.2, A.2.3, A.2.4, A.2.5

ISO 19160-3:2020(E)

- c) Reference: ISO 19115-1:2014, A.2.1, A.2.2, A.2.3, A.2.4, A.2.5
- d) Test type: Basic.

A.6 Test case identifier: Standalone address data quality report

- a) Test purpose: To verify that that ISO 19160-3 (this document), requirement 10 is met.
- b) Test method: Check abstract test suite provided in ISO 19157:2013, A.4.a, A.4.b, A.4.c, A.4.d.
- c) Reference: ISO 19157:2013, A.4.a, A.4.b, A.4.c, A.4.d.
- d) Test type: Basic.

STANDARDSISO.COM : Click to view the full PDF of ISO 19160-3:2020

Annex B (informative)

Methods for measuring address data quality

B.1 General

An address is a complex combination of information. It contains references to a number of separate features (e.g. road name, suburb), recorded as attributes collected together in a single address feature, along with additional attributes specific to that address (e.g. the number). That address shall take into consideration such things as:

- its own unique attributes;
- its own location;
- the location of other addresses around it;
- the attributes of the other addresses around it;
- the values of the attributes it draws from other reference datasets;
- the consistency between the attributes it draws from reference datasets;
- its location compared to the geometry of the feature in the reference dataset that the attribute in the address references.

Given that an address is a complicated collection of information, tests that span multiple data quality elements conducted in a specific order are often needed. For example, if an address number needs to be checked to make sure it is sequentially correct compared to the other addresses on the same street, the street name associated with the address should be checked for thematic accuracy prior to conducting the address number checks.

B.2 Methods for measuring the completeness of address data

B.2.1 Boundary-based methods

Boundary-based completeness methods check the address data against the specified area to which the dataset is associated. Depending on the data available, excess item measures can be conducted using geometry (e.g. address points outside specified city boundaries) or with the data attributes (e.g. place name on an address is not the name specified for the dataset). Boundary-based evaluation methods are external direct and can be used to apply the commission and omission data quality measures described in [8.2](#).

B.2.2 Classification-based methods

Classification-based completeness methods check the address data against the specified classes of addresses or addressable objects included in a dataset. The scope or purpose of an address dataset being evaluated must be known when using these methods. For example, a city tax assessor may maintain an address dataset that only includes the mailing addresses of property owners. This does not mean that other valid addresses associated with the city do not exist. Address datasets comprised only of addresses with official status or valid from/valid to dates are also examples of datasets that can be evaluated with classification-based methods. Classification-based methods are internal direct and can be used to apply the commission and omission data quality measures described in [8.2](#).

B.2.3 Comparison-based methods

Comparison-based methods check the address data unit against a reference file (e.g. universe of discourse, population estimates). Comparison-based methods may be external direct or indirect (depending on the reference file being compared to) and can be used to apply the commission and omission data quality measures described in 8.2.

B.3 Methods for measuring duplicate items in address data

B.3.1 Methods for measuring duplicate addresses

A number of duplicate address measures can be conducted by checking for exact duplication of all data (e.g. attributes, components, addressable object, position) on multiple records. Depending on the dataset specification and the tools available (e.g. address matching software), there are several variations on duplicate address measures. For example, two or more addresses with different component values determining the same addressable object with no indication that they are related might indicate duplicate records. Another example is the use of address matching software to determine fuzzy matches of one or more records.

Methods for measuring duplicate addresses may be internal direct (e.g. all addresses in a dataset are matched to each other) or external direct (e.g. addresses in a dataset are matched to a reference file) and can be used to apply the commission data quality measures described in 8.2.

EXAMPLE 1 An address data quality evaluation checks that an address dataset does not contain two addresses, with exactly the same values in exactly the same set of components, referencing the same addressable object.

EXAMPLE 2 An address data quality evaluation checks that an address dataset does not contain two addresses, with exactly the same values in exactly the same set of components, referencing the same geographic location (x,y).

Careful consideration should be taken when identifying duplicate addresses that do not share exact duplication of all data. This may include obtaining results from multiple data quality elements and measures conducted in a specific order before being able to determine address duplication.

EXAMPLE 3 The Australian Bureau of Statistics uses the method of matching and identifying duplicate addresses that is specified by the national postal agency, Australia Post. Most national postal agencies have a requirement to identify exact and inexact matches against an address list for the purposes of accurate mail delivery. Australia Post specify this requirement as a set of rules published as part of the Address Matching Approval System (AMAS). The AMAS specifies a series of hierarchical rules that determine when synonyms can be used and therefore when to recognize an address that does not exactly match as a valid match and therefore a duplicate.

An example is as follows. Tables B.1, B.2 and B.3 below provide a description of the typical input information that is used to apply AMAS rules to the input addresses listed in Table B.4. This illustrates that 267 MAROONDAH HWY NUNAWADING VIC 3131 is identified as a duplicate of 267 WHITEHORSE RD NUNAWADING VIC 3131 because a known street synonym can be used to generate an exact match under the AMAS rule set.

Table B.1 — Index records

Record number	Index address
Record#1	267 WHITEHORSE RD NUNAWADING VIC 3131
Record#2	433 WHITEHORSE RD MITCHAM VIC 3132

Table B.2 — Street synonyms

PAF street	PAF street type	Alternate street	Alt street type	Index locality	Index Pcode
WHITEHORSE	RD	MAROONDAH	HWY	NUNAWADING	3131
WHITEHORSE	RD	MAROONDAH	HWY	MITCHAM	3132

Table B.3 — Locality synonyms

Locality	Synonym	Postcode
MITCHAM	MITCHAM EAST	3132

Table B.4 — Input addresses, match result and explanation of the application of AMAS rule

Input address	Match to	Explanation the application of AMAS rules
267 MAROONDAH HWY NUNAWADING VIC 3131	267 WHITEHORSE RD NUNAWADING VIC 3131	An exact match is made using the Street Alternate.
L 1 433 MAROONDAH HWY MITCHAM VIC 3132	No Match	An exact match must be made on all address components of the customer/test record. PPM Rule 1 is not permitted.
267 MAROONDAH RD NUNAWADING VIC 3131	No Match	An exact match must be made on all address components in order to use the Street Alternate.
267-271 MAROONDAH RD NUNAWADING VIC 3131	No Match	An exact match must be made on all address components in order to use the Street Alternate.
4/267 MAROONDAH RD NUNAWADING VIC 3131	No Match	An exact match must be made on all address components in order to use the Street Alternate.
433 MAROONDAH HWY MITCHAM VIC 3131	No Match	Must have a valid Locality, Postcode combination to use the Street Alternate.
433 MAROONDAH HWY MITCHAM EAST VIC 3132	433 WHITEHORSE RD MITCHAM VIC 3132	Synonyms may be used to find the correct Locality and match using the Street Alternate record

B.3.2 Methods for measuring duplicate addressable objects

A number of duplicate addressable object measures can be conducted by checking for exact duplication of all data (e.g. attributes, position) on multiple records. Depending on the data product specification, it is also possible to determine duplicate addressable objects without exact duplication of all data.

Methods for measuring duplicate addressable objects may be internal direct or external direct (e.g. address positions are displayed over imagery) and can be used to apply the commission data quality measures described in 8.2.

EXAMPLE 1 An address data product specification states that all real world objects (houses) are to be represented only once in the address dataset. Inspection of the dataset reveals that multiple addressable objects are referencing the same house but with different address position types (front door, building centroid) and coordinates.

Careful consideration should be taken when identifying duplicate addressable objects that are not referenced by the same address. This may include obtaining results from multiple data quality elements and measures conducted in a specific order before being able to determine addressable object duplication.

EXAMPLE 2 Two or more addresses with different component values determining the same real world object with no indication that they are related in the dataset (aliases, versions) could result in apparent duplicate addressable objects in the dataset. These could be considered as other data quality errors (conceptual consistency) prior to determining duplication of addressable objects.

B.3.3 Reporting duplicate items in address data

Consideration should be given to reporting duplicate addresses or addressable objects in order to avoid counting errors multiple times. This is particularly important if an aggregated data quality result is being reported. Customer requirements may specify how duplicate items are to be handled and reported.

EXAMPLE For any group of addresses with the same attributes, process one as a valid address and report the rest as duplicates.

Data quality units in the address data quality evaluation may also impact how duplicates are handled. For example, all addresses that fall outside the geographic extent of the product specification (e.g. city boundaries) could be reported as excess items even if some of the addresses are identical, while duplicate item checks are only conducted on addresses within the geographic extent of the product specification. In such cases, the address data quality evaluator should describe the methods for determining duplicate items in the data quality metadata or standalone quality report.

STANDARDSISO.COM : Click to view the full PDF of ISO 19160-3:2020

Annex C (informative)

Examples of evaluating and reporting address data quality

C.1 General

This Annex provides examples of evaluating and reporting the quality of address data.

C.2 Example data product specification for the address dataset

C.2.1 General

A dataset of street addresses in the City of Tshwane Metropolitan Municipality in South Africa.

C.2.2 Feature types

The street addresses are represented as point features, and each street address has at least an address number, a street name and a suburb name (i.e. the mandatory address components for a street address, as specified in SANS 1883-1:2009). The data conforms to SANS 1883-1:2009 and SANS 1883-2, a profile of the conceptual model for addresses defined in ISO 19160-1. According to SANS 1883-1:2009, the preferenceLevel attribute of an address indicates the ranking of the address in a set of address aliases; 1 indicates highest ranking.

C.2.3 Rules

The addresses in the dataset shall adhere to the following rules:

- The suburb name in the address shall be the same as the name of the enclosing suburb boundary;
- Each cadastral land parcel shall be associated with at least one address;
- Each cadastral land parcel may be associated with more than one address, but only one of them has its preferenceLevel set to 1.

C.2.4 Quality requirements

The overall data quality requirement shall be conformant with the following data quality requirements:

- a) only feature types and attributes defined in this data product specification can be present in the data set;
- b) max 200 items can be missing;
- c) max 200 items can be in excess;
- d) all items are classified as street address;
- e) max 200 items can have incorrect suburb names;
- f) each item has the mandatory address components for a street address;
- g) max 200 items are incorrectly associated with cadastral land parcels;
- h) max 200 items have suburb names that do not match the enclosing suburb boundary.

C.3 Process to evaluate the quality of address data

C.3.1 Specify data quality units

A single data quality unit is identified, comprising quality elements completeness (commission and omission), thematic accuracy (thematic classification correctness), logical consistency (domain consistency, conceptual consistency and topological consistency); and evaluated on the whole dataset (scope).

C.3.2 Specify data quality measures

The data quality measures used for the different quality elements in this example are listed in [Table C.1](#).

Table C.1 — Data quality measures

Data quality element	Data quality measure
Commission	Excess items
Commission	Number of excess items
Omission	Missing items
Omission	Number of missing items
Thematic classification correctness	Items that are incorrectly classified
Thematic classification correctness	Number of incorrectly classified items
Domain consistency	Items with incorrect suburb names
Domain consistency	Number of items with incorrect suburb names
Conceptual consistency	Items with one or more mandatory address components omitted
Conceptual consistency	Number of items with one or more mandatory address components omitted
Topological consistency	Items that are topologically inconsistent with the cadastral land parcels
Topological consistency	Number of items that are topologically inconsistent with the cadastral land parcels
Topological consistency	Items with incorrect suburb names
Topological consistency	Number of items with incorrect suburb names

C.3.3 Specify data quality evaluation methods

Commission: Excess items

External direct evaluation is applied by means of a boundary-based method and full inspection. Any address that is not within the municipal boundary of the City of Tshwane is identified as an excess item.

Commission: Number of excess items

External direct evaluation is applied by means of a boundary-based method and full inspection. Any address that is not within the municipal boundary of the City of Tshwane is counted as an excess item.

Omission: Missing items

External direct evaluation is applied by means of a comparison-based method and full inspection. The cadastral land parcels serve as reference. The centroid of any cadastral land parcel that does enclose an address is identified as a missing item.

Omission: Number of missing items

External direct evaluation is applied by means of a comparison-based method and full inspection. The cadastral land parcels serve as reference. Any cadastral land parcel that does not enclose an address is counted as a missing item.

Thematic classification correctness: Items that are incorrectly classified

Internal direct evaluation is applied by means of a classification-based method and full inspection. Any address that is not classified as 'street address', is identified as an item that is incorrectly classified.

Thematic classification correctness: Number of items that are incorrectly classified

Internal direct evaluation is applied by means of a classification-based method and full inspection. Any address that is not classified as 'street address', is counted as an item that is incorrectly classified.

Domain consistency: Items with incorrect suburb names

External direct evaluation is applied by means of comparison-based method and full inspection. A list of suburb names provided by the municipality serves as reference. Any address with a suburb name not in this list is identified as an item with incorrect suburb name.

Domain consistency: Number of items with incorrect suburb names

External direct evaluation is applied by means of comparison-based method and full inspection. A list of suburb names provided by the municipality serves as reference. Any address with a suburb name not in this list is counted as an item with incorrect suburb name.

Conceptual consistency: Items with one or more mandatory address component omitted

An internal direct evaluation method is applied by means of full inspection. Each item that does not comprise an address number, street name and suburb name is identified.

Conceptual consistency: Number of items with one or more mandatory address component omitted

An internal direct evaluation method is applied by means of full inspection. Each item that does not comprise an address number, street name and suburb name is counted.

Topological consistency: Items that are topologically inconsistent with the cadastral land parcels

External direct evaluation is applied by means of a boundary-based method and full inspection. For each cadastral land parcel, the addresses within its boundaries for which the preferenceLevel is set to 1 are identified. If there is more than one such address, they are identified as items that are topologically inconsistent with the cadastral land parcels.

Topological consistency: Number of items that are topologically inconsistent with the cadastral land parcels

External direct evaluation is applied by means of a boundary-based method and full inspection. For each cadastral land parcel, the addresses within its boundaries for which the preferenceLevel is set to 1 are identified. If there is more than one such address, they are counted as items that are topologically inconsistent with the cadastral land parcels.

Topological consistency: Items with incorrect suburb names

External direct evaluation is applied by means of a boundary-based method and full inspection. The suburb name of the address is compared to the name of the enclosing suburb. If the names differ, the address is identified as an item with incorrect suburb name.

Topological consistency: Number of items with incorrect suburb names

External direct evaluation is applied by means of a boundary-based method and full inspection. The suburb name of the address is compared to the name of the enclosing suburb. If the names differ, the address is counted as an item with incorrect suburb name.

C.3.4 Determine the output of the data quality evaluation result

The methods are applied to identify and count items.

C.4 Report of the quality of the address data

The quality results are reported in [Table C.2](#).

Table C.2 — Report of the data quality results

Data quality element	Data quality measure	Data quality requirement	Data quality measured	Pass?
Commission	Number of excess items	200	50	Yes
Omission	Number of missing items	200	300	No
Thematic classification correctness	Number of incorrectly classified items	0	0	Yes
Domain consistency	Number of items with incorrect suburb names	200	150	Yes
Conceptual consistency	Number of items with one or more address component omitted	0	10	No
Topological consistency	Number of items that are topologically inconsistent with the cadastral land parcels	200	50	Yes
Topological consistency	Number of items with incorrect suburb names	200	150	Yes

C.5 Example data product specification for the address dataset

C.5.1 General

A dataset of single-family houses and their associated addresses used for Census purposes in the US.

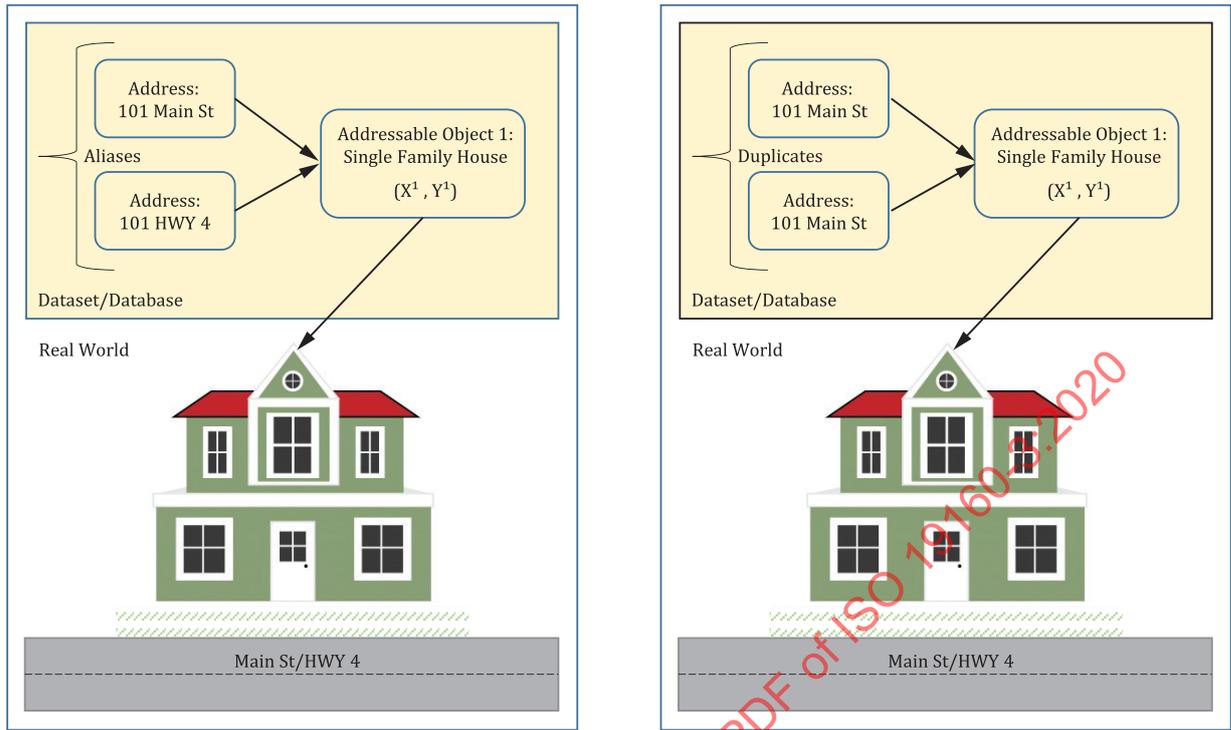
C.5.2 Feature types

Each address has at least an address number and a street name. Each address is associated with an addressable object, i.e. a single-family house. The addressable object is represented as a point feature.

C.5.3 Rules

The addresses in the dataset shall adhere to the following rules:

- a) Each single-family house shall have at least one address associated with it;
- b) An addressable object may have more than one address associated with it, as long as they are different and an *aliasType* is provided. These are address aliases. See [Figure C.1 a\)](#). If they are identical, they are referred to as duplicates. See [Figure C.1 b\)](#);
- c) Street names shall not have spelling mistakes in them. See [Figure C.2](#). Note that once the spelling mistake in this figure is corrected, the identical addresses contravene rule 2);
- d) An addressable object shall not be duplicated, i.e. a single-family house is represented by a single addressable object in the dataset. An addressable object may have more than one *position*. See [Figure C.3](#).



a) Alias addresses

b) Duplicate addresses

Figure C.1 — A single family house associated with more than one address

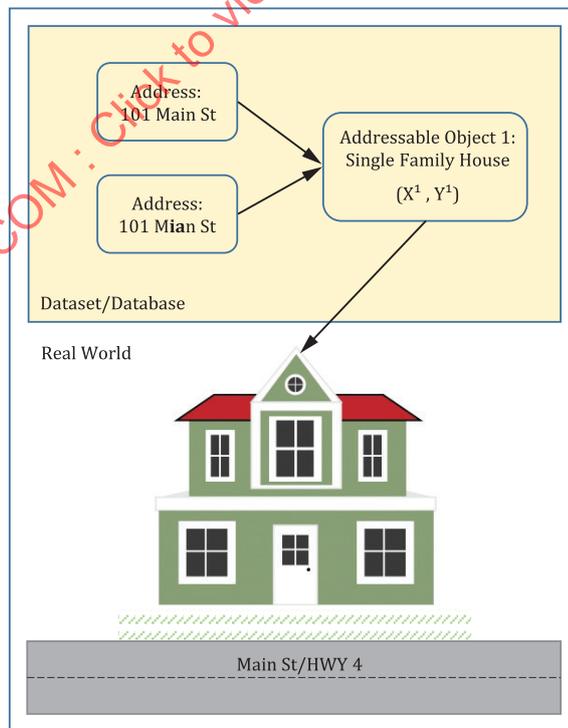


Figure C.2 — A spelling mistake in an address