STANDARD ON

Data Quality

(Approved April 2021)

IAAO
INTERNATIONAL ASSOCIATION
of ASSESSING OFFICERS
Valuing the World

STANDARD ON DATA QUALITY

STANDARD ON DATA QUALITY

Approved April 2021

International Association of Assessing Officers

IAAO assessment standards represent a consensus in the assessing profession and have been adopted by the Board of Directors of IAAO. The objective of IAAO standards is to provide a systematic means for assessing officers to improve and standardize the operation of their offices. IAAO standards are advisory in nature and the use of, or compliance with, such standards is voluntary. If any portion of these standards is found to be in conflict with national, state, or provincial laws, such laws shall govern. Ethical and/or professional requirements within the jurisdiction may also take precedence over technical standards.

About IAAO

IAAO, formerly the National Association of Assessing Officers, was founded for the purpose of establishing standards for assessment personnel. IAAO is a professional membership organization of government assessment officials and others interested in the administration of the property tax. Over the years IAAO members have developed assessment practice and administration standards, and many of these standards have been adopted by state and international oversight agencies, and some have been incorporated into legislation.
IAAO continues at the forefront of assessment in North America and has been expanding its reach to the global community for the last five decades. Because standards form the rules by which North American assessors perform their duties, they may not be directly applicable to an overseas audience. The standards have been updated to also present the broad principles upon which the rules are based. IAAO believes those principles may be adapted to many differing statutory and regulatory scenarios worldwide.

Acknowledgments

The Standard (2019–2020) was developed and written through the dedication of the Data Quality Standard Task Force Phase 1 chaired by Travis Horne, MAS, LAAS, CRA, P.App, AIMA, AMAA; and members Doug Pack; Hart Mauritz, MBA, Ph.D; Jeff Caddell; Rob Ross; Jennifer Rearich, MAS, RES; Cory Yemen; Bill MacGougan, AACI, P.App, MBA; and Paul Bidanset; staff liaison. Phase 2 was chaired by Travis Horne, MAS, LAAS, CRA, P.App, AIMA, AMAA; with members Bill MacGougan, AACI, P.App, MBA; Hart Mauritz, MBA, Ph.D; Mel Obbink, ICA; Jeff Caddell; Rob Ross; and Paul Bidanset; staff liaison. The chair of the Research and Standards Committee was Douglas Warr, AAS.

Revision notes

This is the first standard.

Published by
International Association of Assessing Officers
314 W 10th St
Kansas City, Missouri 64105-1616
816-701-8100 • Fax: 816-701-8149
www.iaao.org
Library of Congress Catalog Card Number: ISBN 978-0-88329-255-6

Produced in the United States of America.

CONTENTS

STANDARD ON DATA QUALITY

1. SCOPE

The IAAO *Standard on Data Quality* defines quality requirements for critical elements of assessment data and provides guidance for meeting those requirements for use in mass appraisal. Assessment data for the purposes of this standard can be separated into property characteristic, economic, and geographic.[1]

This standard addresses quality for both the collection of new inventory and the management of existing inventory. For all types of assessment data, this standard outlines the general requirements for accuracy, currency, consistency, and completeness. Specific local standards should be developed in accordance with this general standard.

This standard is concerned with data as an input into the assessment process. It is not intended to provide guidance for data related to the state of an assessment process (e.g., inspections and other operational processes) or other secondary uses. Assessors should decide whether to include these factors in the development of their own local standards. The quality of the outcomes of assessment processes (e.g., as reported in ratio study results) are well covered in other IAAO Technical Standards and adopted by the assessment industry,[2] and therefore outside the scope of this standard.

1 Eckert, J., R. Gloudemans, and R. Almy, ed. 1990. *Property Appraisal and Assessment Administration*, Chapter 5. Chicago, IL: IAAO.

2 IAAO. 2020. "IAAO Technical Standards." https://www.iaao.org/wcm/Resources/Publications_access/Technical_Standards/wcm/Resources_Content/Pubs/Technical_Standards.aspx?hkey=9c330567-135b-4adc-a772-00008232ab90.

2. OVERVIEW

The importance of data quality has been magnified by the recent and rapid advancements in data analytics as a discipline, transcending all industries in economies around the world. Consequently, this standard provides a common framework for the use of data in the mass appraisal of real property.

This standard addresses a common gap in the current assessment literature: data quality is not succinctly defined, and concrete measures or guidance for the creation of measures are absent. The gap is particularly notable for data as input for assessment or mass appraisal professionals. Data quality is often only inferred from measured outcomes of the assessment process. For example, ratio studies provide indicators of the quality of the assessments themselves, but do not measure the quality of underlying data as an input.

Conclusions on the data quality of the assessor's *input* based solely on the *outcome* (assessments) can be treacherous. There may be cases in which an assessment roll achieves quality outcome measures but the underlying data are subpar. In this fairly common state of affairs some assessments may be of low quality but simply not reflected in the measures. The good measures but poor data combination can be the result of several factors affecting the outcome measures, including sales-chasing, miss-applied mass appraisal methods or nonrepresentative data. Conversely, low-quality assessments do not necessarily equate to low-quality data. This may be caused by the valuation model misinterpreting the data.

While the use of good-quality data by assessors does not guarantee high-quality assessments, the use of poor-quality data renders the task considerably more challenging. Therefore, this standard provides tools for reducing roll quality risk caused by unseen data problems. Beyond assessment, good-quality data matter for making any fact-based decisions and affects numerous users of the information, including taxing jurisdictions, taxpayers, tax agents, public policy makers, academics, researchers, and the broader real estate community.

2.1 RELATIONSHIP TO OTHER STANDARDS

Other IAAO Standards, publications, and authorities ought to be considered and interpreted in conjunction with this standard unless expressly stated otherwise. Local practices and jurisdictional exceptions should also be understood to align when possible.

2.2 KEY ASSUMPTIONS

The following assumptions underlie the requirement for this standard:

- High-quality data used as an input foster high-quality assessments.[3]
- The quality of recorded data erodes over time as physical, economic, governmental, and societal factors change.

3 IAAO. 2013. Section 3.1, "Uniform and accurate valuation of property requires correct, complete, and up-to-date property data." Kansas City: MO: IAAO.

2.3 PRINCIPLES OF DATA QUALITY MANAGEMENT

- All data elements within an assessment system must have a commonly understood definition.
- The different dimensions of data quality must be specified and clearly understood.
- All data elements within an assessment system should have defined criteria for measuring each dimension of data quality.
- Identified dimensions of data quality and their measures need to be simple in order to be practically applied and understood by stakeholders.
- Assessment data management systems should include a means for validating the quality of all assessment data and clear standards against which existing data are to be measured.
- All data within an assessment system should be methodically tested to confirm that each dimension of quality is achieved for each data element.
- All assessment functions that interact with assessment data should strive for compliance with data quality standards.
- Data that do not meet the data quality standard should be corrected in a timely manner.
- Quality data *and* consistent practices are required and ensure that assessments are accurate and fair, even in those areas in which other measures are lacking,

2.4 DIMENSIONS OF DATA QUALITY

For the purposes of this standard, data quality is viewed as data that are accurate, current, consistent, and complete.

Dimension 1: Accuracy

The degree to which records reflect reality. To maintain credibility and precision, all assessment data should be accurate according to the specifications set by the assessor.[4]

Dimension 2: Currency

Currency refers to how the specific data elements reflect the current market preferences (fit for purpose) and how the data itself reflect current property characteristics (up-to-date).

- **Fit for (Current) Purpose.** The degree to which data capture a set of features of a property, parcel, improvement, or other taxable real estate asset that have an impact on its value. All data should be useful and fit for their purpose: to estimate property value through one of the three common valuation approaches. It recognizes that markets are constantly changing. As a result, what is or is not appropriate data should be periodically reconsidered. The assessor should maintain property data quality for property characteristics that affect value and consider ceasing to maintain data that do not materially affect value.
- **Up-to-Date:** The characteristics of any particular parcel can change over time. It is important these changes are reflected in the assessor's data.[5]

4 For example, the size of a building should be accurate within the specified tolerance.

5 For example, a building may be demolished resulting in the need to update data to reflect vacant land.

Dimension 3: Consistency

The degree to which a characteristic is consistently applied.[6]

Dimension 4: Completeness.

The degree to which data for a particular data element have been recorded in the data set for all properties, for example, missing basic information on land size, finished area, or property age for a set of properties in the database.

> The level of detail in providing data elements to capture the relevant aspects for assessing a property is also a consideration.[7]

6 For example, a residential building recorded as Excellent Condition should be in the same physical condition, or state of repair, as all other properties in the same model that are recorded as Excellent Condition.

7 For example, land data may include size, location, topography, zoning, access, and so on.

3. DATA QUALITY MANAGEMENT

Monitoring and reviewing data quality is a fundamental part of a successfully functioning mass appraisal and assessment process. The rate at which the quality of assessment data erodes is highly variable, but it is certain that the gap between what actually exists in the world versus what is in the assessor's records grows over time. In addition to maintaining data to a specified standard and determining areas of strength and/or weakness of data, the results may be used to determine how raw data, stratification of data, data sources, or data collection efforts can be enhanced to produce better future performance.

Regardless of the source, the assessor needs to have in place quality assurance and control mechanisms to ensure the quality of assessment data and to detect and correct deficiencies in the quality of data prior to valuation. These components, when built into the regular reassessment cycle, can assist in providing the necessary confidence in the assessment product.

Most measures of data quality are objective, although some measures can be subjective. Prioritization of the data quality review cycle may be based upon their importance or impact on assessments and/or relevant stakeholders.

3.1 METHODS FOR DATA QUALITY CONTROL AND ASSURANCE

A combination of factors such as reassessment frequency, size of the jurisdiction, and legislated or policy requirements directing the assessor shapes data quality assurance and quality control programs.

All assessment data quality management functions should be documented as part of a broader enterprise-level quality management framework (see figure 1) that contains quality assurance and quality control elements related to (1) quality of the data itself, (2) the quality of data collection, and (3) the quality of data analysis:

1. Clear, up-to-date policy and procedures documentation that includes:
   - Specifications for the data elements to be collected and stored;
   - Standard definitions for all data elements and related terms;
   - Acceptable methods for the uniform collection and recording of all assessment data;
   - Controls on the output for each data-related process or subprocess;
   - Standards for the ongoing testing and maintenance of existing data as they age;
   - Policy compliance testing and reporting function; and
   - Regular procedural reviews.

2. Organizational structure and standard operational workflow to support all related functions.

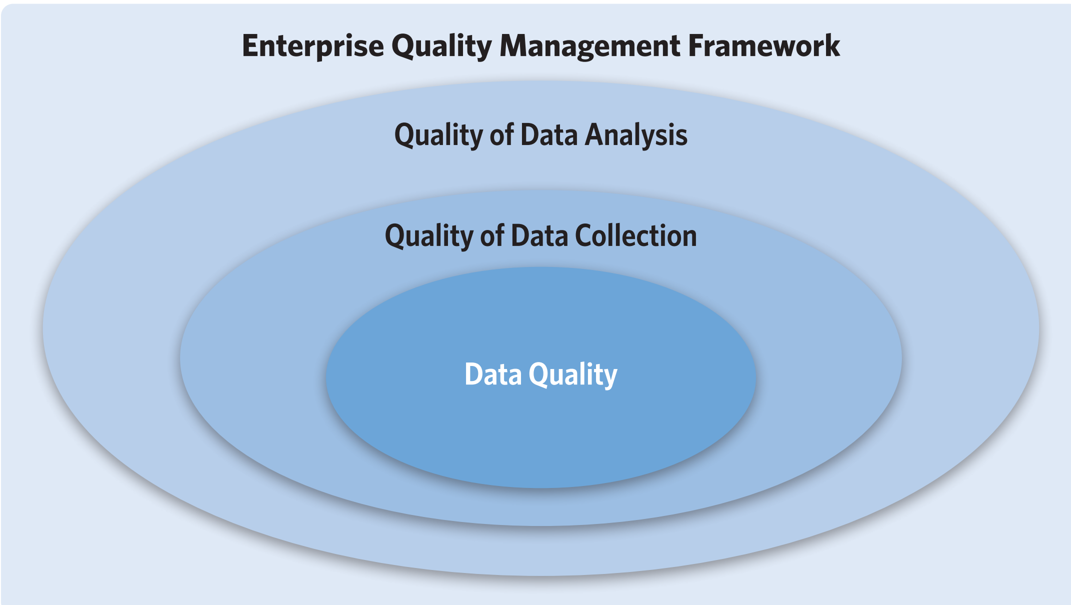
3. Adequate staffing to carry out all the planned functions within the assessor's data quality management framework.

4. Training and performance plans that reinforce and support the data quality management function at an individual (and team) level—whether contracted or in-house services.

5. Re-inspection specifications such as a requirement for all data to be re-inspected or otherwise tested for accuracy over a given cycle.

6. Tools and technology that support all the functions to be carried out.

FIGURE 1. Enterprise Quality Management Framework



3.2 DATA QUALITY MEASUREMENT AND ANALYSIS

The purpose of data quality measurement and analysis is to understand the condition of data in relation to expectations. Assessment data quality measures support various analyses. Descriptive and inferential statistics are central to this process.

Descriptive statistics provide basic information about the raw data and, through additional analysis, insights into current data quality.[8] These statistics may be extracted from audits as part of continuous data quality monitoring (see Appendix G) and describe entire populations. Or they may describe sample data sets and provide insight into representativeness or other descriptive characteristics of the data.

Inferential statistics allow the assessor to use data from a sample to make predictions (inferences) about the population. These can be inferences about the level or uniformity of assessments based

8 For example, the number of sales used in a ratio study or the number of properties in an entire population that have basement finish but no basement area.

on a ratio study or a more complete and accurate description of the location or the nature of data quality based on a property data sample. These insights can then be used to ascertain whether a standard has been met, or, if not, most effectively to target data collection resources and inform decisions such as which data refresh technique to use to most efficiently bring the data quality up to a defined standard.

Employing a methodical data quality measurement and analysis process, an assessment jurisdiction can support a proactive data maintenance program and potentially set the stage for a preventive data quality program that helps diagnose the root cause of data quality issues, accurately predicts where they are most likely to occur in the future, and prescribes the most effective methods to ensure data quality is continuously maintained.

A number of methods for measuring the quality of assessment data are outlined in Table 1 and explained in more detail in the Appendixes.

TABLE 1. Methods for measuring the quality of assessment data

Quality Dimension	Description	Methods
Accuracy	Data on record reflect reality	• Data sampling • Anomaly detection • Continuous monitoring • Data exploration
Currency	1. Fit for purpose 2. Up-to-date	• Data sampling • General quality score • Data exploration
Consistency	Observed data are captured and coded in a consistent well-defined manner.	• Continuous monitoring • Anomaly detection • Sold and unsold
Completeness	All data fields contain appropriate units of data, and data elements collectively describe the subject	• Continuous monitoring • Anomaly detection

SUGGESTED READING

Eckert, J. K., R. J. Gloudemans, R. R. Almy, eds. 1990. *Property Appraisal and Assessment Administration.* Chicago, IL: IAAO.

BIBLIOGRAPHY

Eckert, J. K., R. J. Gloudemans, R.R. Almy, eds. 1990. *Property Appraisal and Assessment Administration.* Chicago, IL: IAAO

IAAO. 2013. *Standard on Mass Appraisal of Real Property.* Kansas City: MOIAAO.

IAAO. 2020. *02: Collecting and Maintaining Property Data, Apendium Overview.* https://www.iaao.org/apendium/02.html.

"IAAO Technical Standards." 2020. International Association of Assessing Officers. https://www.iaao.org/wcm/Resources/Publications_access/Technical_Standards/wcm/Resources_Content/Pubs/Technical_Standards.aspx?hkey=9c330567-135b-4adc-a772-00008232ab90.

Thimgan, G., ed. 2010. *Property Assessment Valuation.* 3rd ed. Kansas City, MO: IAAO.

University of British Columbia, Real Estate Division, Sauder School of Business, ed. 2018. *Foundations of Real Property Assessment and Mass Appraisal.* Vancouver: Sauder School of Business.

GLOSSARY

GLOSSARY

Blind Sampling. To reduce bias in data collection for the purposes of testing data quality, the data is collected without the aid of existing property record information. The data collector must be "blind" to the current data on file.

Change Detection. A comparison of two aerial images of the same area at different times to produce a third image highlighting areas of possible change.

False Negative. An error in which a test result improperly indicates no presence of a condition when in reality it is present. Also known as a type II error.

False Positive. An error in data reporting in which a test result improperly indicates the presence of a condition when in reality it is not present. Also known as a type I error or "false alarm".

Field Inspection. The process of physically visiting a property to collect data needed for assessment.

Gentrification. The process of changing the character of a neighborhood through the influx of more affluent residents and businesses.

Hard Edit. A data error sufficiently serious to trigger rejection of the entry.

Jaccard Similarity Measure. A statistical measure of similarity between two sets of data to see which members are shared and distinct, evaluated on a 0-100% scale. The higher the percentage, the more similar the two populations. Also known as a Jaccard Coefficient or the Jaccard Index.

Land Registry. A system in which matters pertaining to land ownership, possession, and rights can be recorded with a government agency as a means of facilitating transactions and providing evidence of title.

Minkowski Metric. Any of a family of possible ways of measuring distance. Euclidean distance, a member of this family, computes straight-line distances (as the crow flies) by squaring differences in like coordinates, summing them, and taking the square root of the sum. In mass appraisal model building, Minkowski metric usually refers to the sum of absolute differences (not squared) in each dimension and resembles a "taxicab" or city block pattern. Other alternatives are possible, including the distance as calculated only for the dimension of greatest difference, but the city block distance is most common.

Quality Assurance. The proactive maintenance of a desired level of quality in a service or product, especially by means of attention to every stage of the process of delivery or production.

Quality Control. A reactive process through which a business seeks to ensure that product quality is maintained or improved with either reduced or zero errors.

Relational Database Management System (RDBMS). A type of database which uses a structure that allows users to identify and access data in relation to another piece of data in the database. Often, data in a relational database is organized into tables.

Snap Shot. A verbatim record reflecting a specific point in time.

Soft Edit. A data edit signaling an atypical entry and triggering a warning message.

APPENDIX

APPENDIX A. Users

Assessment data are frequently shared with other entities, including other governmental agencies, and serve as an important community resource. Maintaining high levels of data quality ensures not only that the assessment process functions properly, but also that other users of the data can be confident in using it for their purposes.

<table>
<tr><th>USERS</th><th>USES</th></tr>
<tr><td>

Tier One Users

Assessor’s Office
Oversight Body
Appeals Body

Tier Two Users

Jurisdiction’s Planning Department
Jurisdiction’s Transportation Department
Jurisdiction’s Economic Development/ Commerce Department
Jurisdiction’s Revenue Department - Collections focus
Jurisdiction’s GIS Department
Jurisdiction’s EMS and Emergency Management
Jurisdiction’s Buildings/Licensing and Inspection Department
Jurisdiction’s Office of Management and Budget

Tier Three Users

Public/Open Data/Freedom of Information
Neighboring and related assessing jurisdictions
Data aggregators and purchasers

</td><td>

Valuation
Taxation
Assessment certification and quality control
Appeals process

At oversight level
At internal petition process level
At public petition level

Independent petitioner
Represented petitioner

Strategic planning, zoning, comprehensive planning
Transportation improvement planning
Tax collection and arrears management
Economic development, takings, and government driven parcel assemblage
Building code enforcement
Licensing enforcement
Budgeting
Emergency response
Mapping
Mortgage approval and securitization
Public curiosity and interest
Real estate agents, brokers, and related web AVMs[a]

</td></tr>
</table>

a For example, Corelogic, HouseCanary.

APPENDIX B. Risks and Benefits

Data quality has wide-reaching implications across not only all aspects of assessment but also tertiary uses of the data. This section discusses the risks and benefits for various stakeholders in the assessment process.

THE PUBLIC

The most public aspects of data quality can make or break the trust of property owners, tenants, and taxpayers. Large data discrepancies are justifiably problematic, but even small errors can start an erosion of trust that is difficult to rebuild. Although a property owner's main concern may typically be about the appropriate representation of their property by the assessor, it can be just as important for their neighbor's property to be reflected with the same quality information so that the tax burden can be equitably distributed. The state of data quality will likely be directly correlated with the necessity of public relations efforts: high-quality data builds and maintains public trust, while poor-quality data can undermine efforts to establish public trust.

APPEALS

Appeals and protests can indicate data may not be providing an accurate indication of the value of a property. Conversations with property owners throughout the appeal process can help determine the level of data quality. When factual errors are present, they may cause complications for explaining and supporting valuations to both the property owner and the trier of fact. The quality of data is directly correlated to assessor confidence in explaining and supporting property value in an appeal.

ASSESSOR'S OFFICE/JURISDICTION

Data collection is the foundation upon which all assessment processes are reliant. Establishing clear data collection procedures and maintaining a quality assurance program are paramount to maintaining high-quality data. Clear guidance minimizes subjectivity or differences of opinions among assessors/data collectors, while quality assurance processes review collected data to ensure procedures are followed and consistency is maintained. In addition, maintaining high-quality data assists in other evaluations of the office, such as performance audits.

MODELING

In mass appraisal, valuations are typically calculated through a model, whether it relies on a cost system, multiple regression, machine learning, artificial intelligence, or some combination thereof. If consistent, reliable data are not available, constructing and utilizing a model for purposes of estimating values may prove extremely challenging. When data are of poor quality, not only does this affect the final estimate, but also it has implications for the model-*building* process, particularly if data errors are above acceptable norms. In conducting data analysis as part of the model-building process, assessors take care to identify outliers and/or unreliable data. While assessors ought to take every precaution to minimize impact, poor data quality is not a problem that can be easily resolved by the assessor in the modeling process.

BUDGETING

Budgets allow assessors to undertake special projects as well as maintain an assessment system. Data collection is a time-intensive, costly aspect of assessment, but the costs of poor data quality are far more expensive than the costs of maintaining high-quality data. Effective, efficient, and economical use of financial resources favor the adoption of collection processes that facilitate high-quality data.

A desktop review process may include capturing and updating property data by integrating the local computer-assisted mass appraisal (CAMA) system with oblique imagery, street-front photos, geographic information system (GIS) layers, and orthographic imagery with geo-referenced sketches. Baseline comparisons aid in tracking data quality review from a data comparison standpoint. Reviews referencing baseline results via the desktop review processes may result in positive or negative impacts on the assessment role. This type of impact could result in more or less annual tax revenue. From a financial perspective, the review—if the results are negative—still yield a positive conclusion based on an identification or correction within the data. Up or down, the results of the review are primarily focused on correcting or identifying a misallocation of the tax burden, thereby promoting uniformity across the tax base.

Actual review costs are typically not known during initial reviews or establishment of baselines. Iterative reviews that did follow a set of review standards and methods may typically decrease the costs as internal efficiencies and reporting required are refined and updated within the jurisdiction. Recoding the costs to review data quality enables metrics to quantify efficiencies and measure any new processes implemented in the data collection or review. Sampling may be the lowest cost review with the tangible results in which the measurements are acceptable; however, it is not the only option.

AUDIT

Assessors are typically subject to oversight compliance through a variety of angles, many of which are highly dependent on data quality or directly evaluate data collection procedures. While ratio studies and existing methods relying on those studies are some of the primary tools for ensuring assessments are fair and equitable, audits can also evaluate whether the underlying data accurately reflect property characteristics. Most avenues of auditing are in some way connected to the quality of data underlying any particular assessment procedure or process.

PROFESSIONALISM

The maintenance of high-quality data indicates the utmost levels of professionalism by assessors. Ethical and professional standards require assessors to take care in their mass appraisal process to avoid errors and omissions.

APPENDIX C. Summary of Methods

PROPERTY DATA SAMPLING

By fully inspecting a randomly selected sample of properties, the assessor can make reliable inferences about the accuracy of property data for the population. This can provide for a comprehensive standards-based program in which a complete re-inspection of all properties is not feasible for each revaluation.

ASSESSMENT TO REVIEW RATIO

This method is applicable to considering data quality when objective information is insufficient or has historically been based upon subjective opinions. The results of this method allow prioritization of data review or collection. This method may apply within groups or between them [City wide or comparing market 1 to market 2]. It may also apply across groups based on specific characteristics [bedrooms]. The results are objective, enabling greater consideration. Results may be stored as a data set or plotted on maps. Results and subsequent efforts by the assessor are measurable on a continuous basis.

SOLD AND UNSOLD

The Sold and Unsold Analysis uses simple tests to identify whether a statistically significant difference in value change exists between sold and unsold properties within a particular market area. In the sales review process, data elements of properties that have sold may have been updated to reflect the property at the date of sale. If sales required substantial updating of subject properties' data, then unsold properties' data may also require updating. The results of the Sold and Unsold Analysis should inform the assessment organization's decisions on future data improvement efforts.

CONTINUOUS MONITORING

Continuous data quality monitoring utilizes a rules-based approach with queries developed to extract erroneous or suspect data from the CAMA system on a schedule appropriate for the jurisdiction (e.g., daily, biweekly, quarterly or annually). For example, queries can be developed based on rules such as a 2,000-sq-ft dwelling would typically be expected to have more than one bathroom. Continuous data quality monitoring allows proactivity rather than reactivity in identifying and resolving individual data errors and broader data quality issues before valuations are finalized.

ANOMALY DETECTION

Anomaly detection enables an assessor to analyze large amounts of data to identify properties/parcels/folios with unusual characteristics, and steers the assessor's efforts in addressing potential data errors in an efficient and economical manner. The method may be used on continuous, categorical, or binary elements. It does not require assessments or sales to occur in order to apply. The return on investment of employing this method is dependent on the assessor's selection of elements that have the greatest impact on roll value or the public's confidence in the underlying data. A good understanding of the market under review through experience, exploratory analysis, or sampled data is a prerequisite for its application.

EDGE CASE IDENTIFICATION

The Edge Case Identification method is used to identify extreme sales ratios that may stem from many sources: nonmarket influences in the market data, gaps or weakness in the valuation method, and incomplete or incorrect data. While some extreme ratios and variation are typically expected, the presence of systemic error on a data element may heavily affect the quality of the assessment roll and introduce bias when applied to the population. A small number of data elements may have a large influence in all valuation methods, for example, location, size, quality, style, and effective age. Identifying over-representation in extreme sales ratios identifies areas of weakness in the data, which may be used to prioritize review and correction in the population data.

GENERAL QUALITY SCORE (GDQS)

The General Data Quality Score method allows for comparison of data quality across different sets of data using a single measure ranging from 0 to 100. This method may be used to grade the quality of data after conducting a careful audit or inspection of properties in the field.

IMPACT IDENTIFICATION

The Data Quality Impact Identification method simulates the impact of low data quality on roll quality measures. Methodically injecting increasing amounts of errors into a known sample reveals the impact of data quality on statistics like the coefficient of dispersion (COD) and price-related differential (PRD). This allows the prioritization of certain data elements for further investigation.

APPENDIX D. Property Data Sampling

The premise of property data sampling is straightforward: the best way to know the accuracy of property data is to go out in the field and test it. By fully inspecting a randomly selected sample of properties, the assessor can make reliable inferences about the accuracy of property data for the population. Many jurisdictions carry out some form of sampling from time to time for various purposes. However, to help drive a standards-based data maintenance program, the sampling effort ought to be a well-designed process that is organized and ongoing.

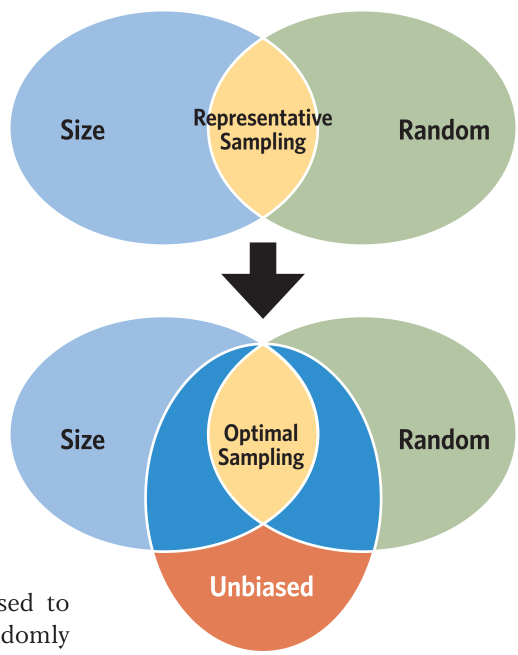
In order to use sampling as part of a comprehensive program and to provide uniformity of practice and results across the assessment jurisdiction, the data accuracy standard within this standard, and/or another local standard, are reliable guidelines. A locally developed standard must identify the data elements, and the degree of accuracy required for each, to serve as a consistent determination of acceptable data accuracy and to support in subsequent analysis.

SAMPLE REPRESENTATIVENESS

The first, and often most important, issue is the generation of a random sample of sufficient size. The next piece is a verifiably unbiased data collection process. Having such a repeatable function allows the application of sampling as a method to validate the accuracy of property data for most stratifications as well as the entire jurisdiction. It is recommended that a 90 percent confidence level be used with an allowable error of ±10 percent combined with "blind" unbiased (and quality controlled) data collection.

Notes on Valid Sample Sizes

- No sample tells the assessor the exact characteristics of the population.
- The confidence interval expresses the range of values on either side of the sample statistic within which a certain degree of statistical evidence (or confidence) suggests that the corresponding statistic for the true population is contained.
- A 90 percent confidence interval indicates 90 percent confidence that the true population statistic falls within the calculated confidence interval.
- An allowable margin of error of 10 percent indicates the assessor is 90 percent confident the true population statistic is no more than ±10 percent different from the calculated sample statistic. That is, if a sample indicates building area property data is accurate 89 percent of the time, the assessor is 90 percent confident that the true population is no more accurate than 99 percent and no less accurate than 79 percent.



A random sample generator report should be used to calculate a statistically valid sample size and randomly select properties from the population to be sampled. The required sample size depends on the location of the population, that is, urban and more homogeneous or rural and more heterogeneous.

Depending on the CAMA system used, the property data collection process likely requires the creation of a snapshot of the data present on a property record prior to inspection and the provision of an inspection template to the data collector with specific information to identify and locate the property. In order to avoid potential bias, the template must not include any of the primary valuation drivers such as building class, square footage, or effective age.

Data collectors are instructed to follow all applicable inspection procedures during sample data collection except when that might involve the data collector viewing or otherwise learning about the existing property data on record. Every attempt should be made to ensure that the complete property data are being collected, just as it would be recorded on a first-time inspection; therefore, a full interior inspection is strongly preferred.

From the field inspection, the data are entered into the previously created blank property record. The new information can then be compared to the property record and a Property Data Accuracy Report generated. This report should be set up to list the discrepancies between the two data sets for all key data elements and highlight where a given data element is outside of the acceptable range to meet the standard.

PROPERTY DATA ACCURACY REPORTING

These reports should indicate at a high level whether a sampled property group clearly meets the data accuracy standard. They should also array the data comparisons (current collected versus property record) for the key data elements, as well as ratio calculations, and measures of central tendency and dispersion in a way that helps the assessor review the results in detail when a population fails to meet the standard or the conclusion is not clear. Different reports may be required for different property types.

While there is a general tendency for out-of-date property data to result in ratios below 1.00 (i.e., it is more likely that updated data result in a higher value) absolute variance should be referenced where practical for the clearest picture of data accuracy and to best support subsequent analysis of results.

It is worth emphasizing the value of collecting all data that make up a complete property record from sampled properties. While collecting partial data can, at times, provide a reliable indicator of whether an identified property group meets the referenced data standard, having complete data allows full value calculations and more in-depth analysis of the results to help confirm data issues. Such analysis helps to support quality control on the process, ensure that the results are valid., and isolate specific data issues, allowing for a more targeted data-updating solution. Ultimately, complete data help ensure the most effective and efficient use of resources, the best potential outcomes for accurate data and fair and accurate assessments, and continuous process improvement.

DATA UPDATE RESULTS REPORT

In addition to property data accuracy reports, data update reports should be developed to provide a valid and meaningful measure of updating results. By taking a snapshot before and after a project (similar to the development of an accuracy report from sampling), the assessor can identify the specific property data that changed as a result of the update and report on the results.

There is more value in this reporting than only the results of the data update, however. The added value of this information is in the analysis it supports for the data maintenance function (e.g., to learn about which data update methodology is most effective in different circumstances), to validate sampling results, and to provide feedback to the sampling process itself supporting continuous process improvement.

APPENDIX E. Assessment-to-Review Ratio

OVERVIEW

This method is applicable in considering data quality in which objective information is insufficient or has been historically based upon subjective opinions. The results of this method allow prioritization of data review or collection. This method may apply within groups or between them (city-wide or comparing market 1 to market 2). It may also apply across groups based on specific characteristics (bedrooms). The results are objective, enabling greater consideration. Results may be stored as a data set or plotted on maps. Results and subsequent efforts by the assessor are measurable on a continuous basis.

DATA TYPES

Data suitable for this method are property data (land, improvements, and characteristics).

QUALITY DIMENSIONS

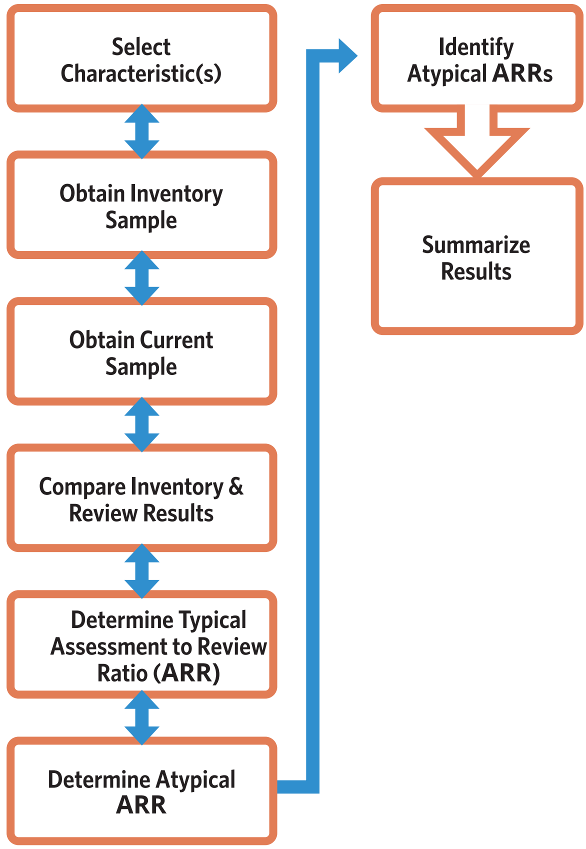
This method focuses on an overlap between accuracy, currency, and consistency. Reviewing completeness is not well suited to this method because the assessment-to-review ratio (ARR) is dependent upon the data selected by the assessor.

PREREQUISITES

Sampling at a minimum of 3 percent of the inventory is necessary.

DESCRIPTION

1. Select a characteristic(s) to review
2. Obtain a sample from the assessor's records
3. Obtain sample from [**2**] based on the characteristic in [**1**]
4. Compare the results of [**2**] and [**3**]
5. Determine typical ARR
6. Determine atypical ARR
7. Identify atypical ARRs
8. Summarize and report results.



OUTCOME MEASURES

Atypical ARRs indicate the frequency and degree of deviation of the characteristic selected. Establishing a benchmark typical ARR for inventory data enables consideration time forward. Using a benchmarked ARR enables application of this method without dedicated data collection projects.

INTERPRETATION OF RESULTS

The results of the ARRs allow objective consideration of inventory review by property group or characteristic across property groups.

DETAILED DESCRIPTION

1. Select a characteristic to review:

Basement finish

2. Obtain a sample from the assessor's records

Account	Type	Above Grade Area sq ft	Basement	Inventory
				Basement Finished Area
6685169	Bng	3182.5	Yes	1495.775
6690804	2S	3182.5	Yes	1495.775
6696439	1S	3182.5	Yes	1495.775
6702074	Bng	3182.5	No	0
6707709	2S	3182.5	Yes	1495.775
6713344	1S	3182.5	No	0
6718979	Bng	3182.5	Yes	1495.775
6724614	2S	3182.5	Yes	1495.775
6730249	1S	3780	Yes	888
6735884	Bng	3640	Yes	1710.8
6741519	Bng	3640	Yes	1710.8
6747154	Bng	3640	Yes	1710.8
6752789	Bng	1639	Yes	770.33
6758424	Bng	3048.5	Yes	1432.795

3. Obtain sample from [2] based on the characteristic in [1]

Account	Type	Above Grade Area sq ft	Basement	Inventory	Basement	Sample [Inspection]
				Basement Finished Area		Basement Finished Area
6685169	Bng	3182.5	Yes	1495.775	*Yes*	*1346.1975*
6690804	2S	3182.5	Yes	1495.775	*Yes*	*1495.775*
6696439	1S	3182.5	Yes	1495.775	*Yes*	*1495.775*
6702074	Bng	3182.5	No	0	*No*	*0*
6707709	2S	3182.5	Yes	1495.775	*Yes*	*1495.775*
6713344	1S	3182.5	No	0	*No*	*0*
6718979	Bng	3182.5	Yes	1495.775	*Yes*	*1346.1975*
6724614	2S	3182.5	Yes	1495.775	*Yes*	*1495.775*
6730249	1S	3780	Yes	888	*Yes*	*888*
6735884	Bng	3640	Yes	1710.8	*Yes*	*1710.8*
6741519	Bng	3640	Yes	1710.8	*Yes*	*1539.72*
6747154	Bng	3640	Yes	1710.8	*Yes*	*1710.8*
6752789	Bng	1639	Yes	770.33	*Yes*	*770.33*
6758424	Bng	3048.5	Yes	1432.795	*Yes*	*1432.795*

4. Compare the results of [2] and [3]

[2] ÷ [3] = ARR

Account	Type	Above Grade Area sq ft	Basement	Inventory Basement Finished Area	Basement	Sample [Inspection] Basement Finished Area	Ratio
6685169	Bng	3182.5	Yes	1495.775	Yes	1346.1975	*1.111111*
6690804	2S	3182.5	Yes	1495.775	Yes	1495.775	*1*
6696439	1S	3182.5	Yes	1495.775	Yes	1495.775	*1*
6702074	Bng	3182.5	No	0	No	0	*1*
6707709	2S	3182.5	Yes	1495.775	Yes	1495.775	*1*
6713344	1S	3182.5	No	0	No	0	*1*
6718979	Bng	3182.5	Yes	1495.775	Yes	1346.1975	*1.111111*
6724614	2S	3182.5	Yes	1495.775	Yes	1495.775	*1*
6730249	1S	3780	Yes	888	Yes	888	*1*
6735884	Bng	3640	Yes	1710.8	Yes	1710.8	*1*
6741519	Bng	3640	Yes	1710.8	Yes	1539.72	*1.111111*
6747154	Bng	3640	Yes	1710.8	Yes	1710.8	*1*
6752789	Bng	1639	Yes	770.33	Yes	770.33	*1*
6758424	Bng	3048.5	Yes	1432.795	Yes	1432.795	*1*

5. Determine typical ARR

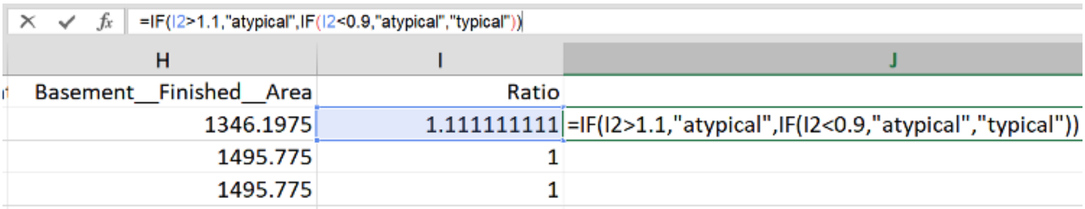
Count	137
Mean	0.95
Median	1
SE	0.02
COD	8.316
Min	0
Max	1.11
25th Percentile	1
75th Percentile	1

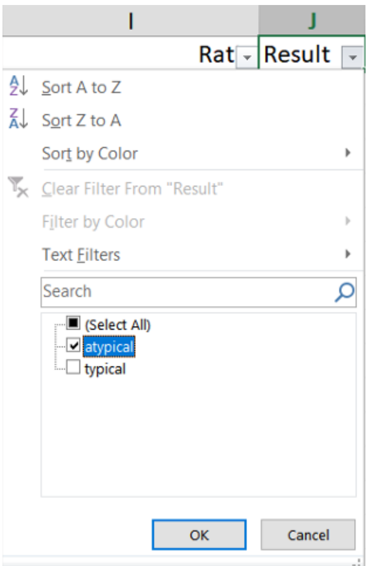
Typical between 0.90 and 1.10.

6. Determine atypical ARR

Atypical >1.10 and <0.90

7. Identify atypical ARRs





Account	Type	Above Grade Area sq ft	Basement	Basement Finished Area	Basement	Basement Finished Area	Ratio	Result
6685169	Bng	3182.5	Yes	1495.775	Yes	1346.1975	1.111111111	atypical
6686693	Bng	3182.5	Yes	1495.775	Yes	1346.1975	1.111111111	atypical
6687709	Bng	3640	Yes	1710.8	Yes	1539.72	1.111111111	atypical
6689487	1S	2730	Yes	1	Yes	2593.5	0.000385579	atypical
6691773	1S	2715	Yes	1276.05	Yes	1148.445	1.111111111	atypical
6692535	1S	3620	Yes	1701.4	Yes	1531.26	1.111111111	atypical
6692789	Bng	1845	Yes	1734.3	Yes	1560.87	1.111111111	atypical
6694059	1S	1845	Yes	1734.3	Yes	1560.87	1.111111111	atypical
6694821	1S	1845	Yes	1	Yes	1752.75	0.000570532	atypical
6696091	2S	2114	Yes	1987.16	Yes	1788.444	1.111111111	atypical
6696345	1S	1845	Yes	1	Yes	1752.75	0.000570532	atypical
6696853	2S	1845	Yes	1734.3	Yes	1560.87	1.111111111	atypical

8. Summarize and report results

Account	Type	Above Grade Area sq ft	Basement	Basement Finished Area	Basement	Basement Finished Area	Ratio	Result
6689487	1S	2730	Yes	1	Yes	2593.5	0.000385579	atypical
6691773	1S	2715	Yes	1276.05	Yes	1148.445	1.111111111	atypical
6692535	1S	3620	Yes	1701.4	Yes	1531.26	1.111111111	atypical
6694059	1S	1845	Yes	1734.3	Yes	1560.87	1.111111111	atypical
6694821	1S	1845	Yes	1	Yes	1752.75	0.000570532	atypical
6696345	1S	1845	Yes	1	Yes	1752.75	0.000570532	atypical
6698123	1S	2114	Yes	1	Yes	2008.3	0.000497934	atypical
6703711	1S	1845	Yes	1	Yes	1752.75	0.000570532	atypical
6705489	1S	2114	Yes	1	Yes	2008.3	0.000497934	atypical
6711839	1S	1845	Yes	1	Yes	1752.75	0.000570532	atypical
6713617	1S	2114	Yes	1	Yes	2008.3	0.000497934	atypical
6696091	2S	2114	Yes	1987.16	Yes	1788.444	1.111111111	atypical
6696853	2S	1845	Yes	1734.3	Yes	1560.87	1.111111111	atypical
6699139	2S	2652	Yes	2492.88	Yes	2243.592	1.111111111	atypical
6701171	2S	2652	Yes	2492.88	Yes	2243.592	1.111111111	atypical
6703203	2S	2652	Yes	2492.88	Yes	2243.592	1.111111111	atypical

APPENDIX F. SOLD AND UNSOLD

OVERVIEW

The Sold and Unsold Analysis uses simple tests to identify whether a statistically significant difference in value change exists between sold and unsold properties within a particular market area. In the sales review process, data elements of properties that have sold may have been updated to reflect the property at the date of sale. If sales required substantial updating of subject properties' data, then unsold properties' data may also require updating. The results of the Sold and Unsold Analysis should inform the assessment organization's decisions on future data improvement efforts.

Property Group of Interest	Sale/ Nonsale	Count	Mann-Whitney Test Result	Weighted Mean Value Change	Overall Result
Group 1	Sale	511	0.00	4.1%	Pass (5%)
	Nonsale	26,990		0.2%	
Group 2	Sale	157	0.66	1.9%	Pass (both)
	Nonsale	3,693		0.7%	
Group 3	Sale	16	0.10	-14.2%	Pass (M-W)
	Nonsale	741		-2.0%	
Group 4	Sale	535	0.00	5.8%	Fail
	Non-sale	27,095		0.3%	
Group 5	Sale	9			N/A
	Nonsale	402			Insufficient data

DATA TYPES

- Market data, specifically the previous and current assessment roll value, as well as a catalogue of accounts that sold between the respective base dates.
- Optionally, property data, specifically data to identify nonmarket changes to the above accounts in order to exclude them from the analysis (e.g., building permit data).

QUALITY DIMENSIONS

- Consistency
- Accuracy
- Currency
- Completeness

PREREQUISITES

Performing Sold and Unsold Analysis requires

- A complete or near-complete set of reassessment values for the previous and current rolls
- A complete or near-complete list of properties that sold between the respective base dates, and, optionally,
- Reliable data to identify nonmarket changes.

DESCRIPTION

- Assemble a data set with previous and current roll values, identifying properties that sold between the respective base dates.
- Exclude properties with known nonmarket changes from the final analysis.
- Groups of interest with few accounts or little sales activity may have insufficient sample size to calculate reliable statistics. A group with a small sample size (e.g., fewer than 11 Sales or Unsold properties) is deemed to have insufficient data for statistical analysis. In this case, attempt to pool together similar groups of interest to achieve a minimal viable sample.
- Calculate the percentage value change for each property as:

$$\frac{\textit{(Current Roll Value – Previous Roll Value)}}{\textit{(Previous Roll Value)}}$$

- Identify all property groups of interest on which to perform Sold and Unsold analysis (e.g. market or geographic areas, property types).
- Calculate the Mann-Whitney U test statistic comparing percentage value changes for Sold and Unsold properties for each of the groups of interest.
- Calculate the Weighted Mean Value Change for Sold and Unsold properties for each of the groups of interest.
- Determine the result of the Sold and Unsold analysis for each group of interest, and assign a Pass or Fail rating
  - A Mann-Whitney U test statistic of 0.05 or greater indicates no statistically significant difference in the distribution of the Sales and Unsold properties and is deemed a pass.
  - A difference in weighted means of 5 percent or lower indicates that the value change of Sold and Unsold properties is similar and may be deemed a pass.
  - A group of interest with a Mann-Whitney U test statistic of 0.05 or greater and a difference in weighted means of 5 percent or lower passes both criteria.
  - A group of interest with a Mann-Whitney U test statistic less than 0.05 and a difference in weighted means greater than 5 percent is deemed a fail.

OUTCOME MEASURE

A Mann-Whitney test statistic greater than 0.05 indicates that there is no statistically significant difference in the distribution of value changes when sold and unsold properties are compared. Large sample sizes cause the acceptable difference between samples to shrink, which may identify inequity even when there is a small difference (less than 5 percent). Because of this, the jurisdiction may also wish to compare the overall weighted mean value change of Sold and Unsold properties and deem equitable treatment if the difference in these weighted means is less than 5 percent.

INTERPRETATION OF RESULTS

If the market area fails the Mann-Whitney test and/or there is more than a 5 percent difference between the weighted mean change of sold and unsold properties, (the jurisdiction should investigate further those market areas to focus and prioritize data improvement efforts.

DETAILED DESCRIPTION

In the course of the sales review process, properties that have sold receive data updates that unsold properties may not receive in preparation for the reassessment. This difference in data updates may contribute to inconsistent data and inequity in values between properties that sold and those that did not sell in subsequent rolls. Sold and Unsold analysis aims to identify whether a statistically significant difference in value change exists between these types of properties. The assumption is that if the sale filed required substantial data updating, the unsold file potentially also requires data updating. The process described identifies areas to be identified where data inconsistencies exist and updating is required, allowing the jurisdiction to prioritize data improvement efforts.

The Mann-Whitney test of statistical significance is utilized to compare the appraisal performance of properties that sell in a roll year to those that did not by comparing the relative change in assessment value between the two groups. The goal of this test is to demonstrate that sold and unsold properties have been treated equitably through the year’s reassessment activities. The assumption is that sold and unsold properties come from the same population as far as reassessment is concerned. A threshold for reasonable difference in the weighted mean value change (e.g., 5 percent) may be used in conjunction with or in place of the Mann-Whitney test to provide additional evidence for potential data inconsistencies.

APPENDIX G. CONTINUOUS MONITORING

OVERVIEW

Continuous data quality monitoring utilizes a rules-based approach with queries developed to extract erroneous or suspect data from the CAMA system on a schedule appropriate for the jurisdiction (e.g., daily, biweekly, quarterly or annually). For example, queries can be developed based on rules such as a 2,000-sq-ft dwelling would typically be expected to have more than one bathroom. Continuous data quality monitoring allows proactivity rather than reactivity in identifying and resolving individual data errors and broader data quality issues before valuations are finalized.

DATA TYPES

For property data and market data. Any data field that may have a rule developed to identify errors or potential errors is a candidate for Continuous Monitoring.

QUALITY DIMENSION

Continuous Monitoring is suitable for addressing quality issues related to the following dimensions of data quality:

- Consistency
- Accuracy
- Completeness.

PREREQUISITES

Continuous Monitoring may apply to any assessment jurisdiction's database. Periodic reports require only the capability to query the database for specific scenarios, which can then be distributed for review and correction.

DESCRIPTION

1. Define the property type or characteristic for consideration.

2. Define the expected result or range of tolerance within the Step 1 query. Based on rules or scenarios relevant for the jurisdiction.

3. Extract the data based on Step 1.

4. Create and apply rules determined in Step 2 to the data.
   - Review the results for the following types of errors:
   - Objective error: errors in accuracy or completeness that cannot be correct (e.g., missing mandatory data fields, dwelling area of 0 sq ft)
   - Probable error: scenarios that are likely to be an error but are possibly correct (e.g., basement living area greater than the total of ground floor living area plus additions)
   - Possible error: scenarios that are not necessarily indicative of an error, but are of significant value or risk to the roll (e.g., extreme value changes, change in override status).

5. Save the results in a format that can be easily reviewed and shared.

6. Determine:
   - A reporting schedule of the results (e.g., daily, weekly, quarterly) for the purpose of continuous data quality monitoring
   - Staff to receive the reports based on Step 7a.

7. Revisit the usefulness of reports routinely to optimize the return on investing time and effort in continuous monitoring

OUTCOME MEASURE

Continuous Monitoring results vary upon the type. Periodic reporting allows for batch review and trend identification prior to finalizing the roll.

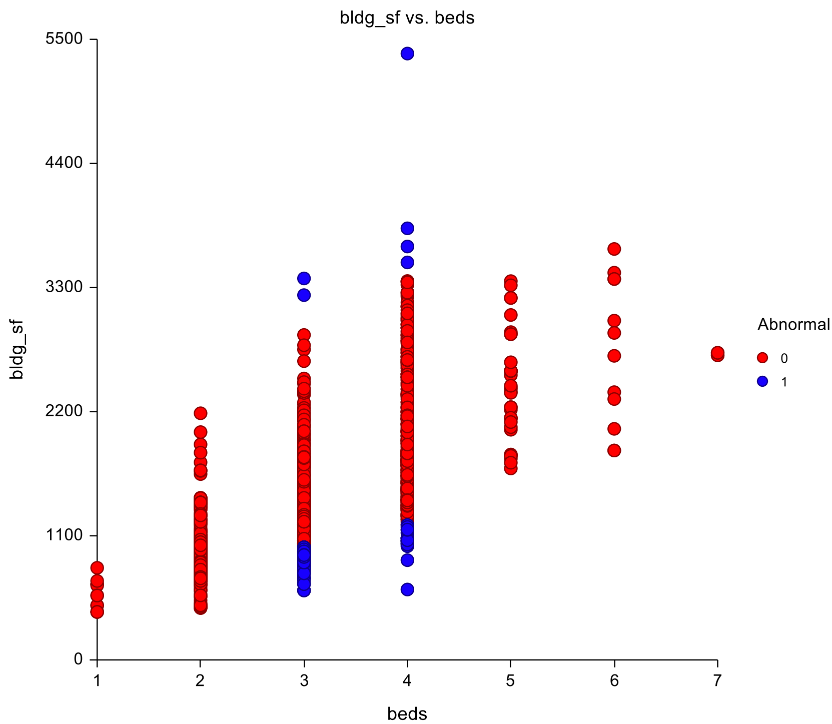
Proactive Continuous Monitoring of data quality not only assists with improving the quality of data in the CAMA system upon which valuations are based, but also serves the important function of identifying training opportunities for data collection and data entry personnel, as part of the continuous improvement function within the jurisdiction.

INTERPRETATION OF RESULTS

Continuous Monitoring results may be interpreted depending on their type. Objective errors must be corrected. Probable errors and reviews require follow-up by the reviewer with the assessor or data entry personnel that made the change. Possible errors should be reviewed by subject matter experts and assessors to determine whether the findings of the query point to data quality issues or are factual. Identifying repeated errors may reveal gaps in training, technology, and/or procedures.

APPENDIX H. ANOMALY DETECTION

OVERVIEW



Anomaly detection enables an assessor to analyze large amounts of data to identify properties/parcels/folios with unusual characteristics, and steers the assessor's efforts in addressing potential data errors in an efficient and economical manner. The method may be used on continuous, categorical, or binary elements. It does not require assessments or sales to occur in order to apply. The return on investment of employing the method is dependent on the assessor's selection of elements that have the greatest impact on roll value or the public's confidence in the underlying data. A good understanding of the market under review through experience, exploratory analysis, or sampled data is a prerequisite for its application.

DATA TYPES

The Anomaly Detection method is suitable for a wide range of data types, particularly property and market data.[9] In its current form, the method is not suitable for the analysis of unstructured data.

9 IAAO. 2020. *02*: *Collecting and Maintaining Property Data,* Apendium Overview. https://www.iaao.org/apendium/02.html.

QUALITY DIMENSION

The Anomaly Detection method is a primary basic tool for data quality analysis and enhancement. It provides direct insights into the accuracy of data by flagging unusual observations on the data set. Further analysis of anomalies may provide insights into the currency and consistency of data. Information about completeness is only a side product of applying this method:

1. Accuracy
2. Currency
3. Consistency
4. Completeness

At a minimum the element(s) reviewed may indicate a non-issue that those individual elements providing guidance that data quality issues are not underlying these specific elements.

PREREQUISITES

A spreadsheet containing the data for the elements to be reviewed, *and* a benchmark or typical value for those elements gained through experience, exploratory analysis, or sampled data. The method may be applied routinely, as a result of surveys (e.g., requests for information, self-reporting), trends in taxpayers' inquiries, or appeals.

DESCRIPTION

1. Select data characteristic(s) to review.[10]
2. Determine the appropriate unit of comparison based on Step 1.
3. Select stratification(s) for review.[11]
4. Save a snapshot of the data set.
5. Plot the data.
6. Identify abnormal results.[12]
7. Summarize abnormal results.

10 For example, property and/or market characteristics.

11 For example, market area 1 and market area 2, or bedrooms, city-wide.

12 Benchmarks may be established once the first review is completed in order to apply and/or refine time forward.

OUTCOME MEASURE

Outcome measures would be for this method the unusual/abnormal values for data elements observed in the data set. What constitutes *unusual* or **abnormal** needs to be defined by way of a comparator/benchmark or statistical analysis (e.g., flags statistically valid outliers).

INTERPRETATION OF RESULTS

Abnormal results should be considered within the scope of impact they may have on assessments. Typically, the elements selected for review are based on a degree of impact already found to be material by the assessor. The results may be inspected specifically for these abnormalities or added to regular inspections schedules where they were previously absent. Best practices suggest that an inspection of a single element may not be as fruitful as a complete inspection of the entire property, correcting or confirming property elements data not specifically reviewed.

Once the data have been corrected or verified, the resulting data can provide an updated baseline sample that may be a suitable baseline for the methods application in other similar classes of property.

Further Resources

2019 IAAO U40 Leadership Lab "Improving Data Quality through Analytics." IAAO Annual Conference

DETAILS

Assessor Data

1. Select data characteristic(s) to review

   [Bedrooms]

   Number of bedrooms

2. Determine the appropriate unit of comparison based on Step 1.

   [BLDG_SF]

   Building square feet, as measured from the exterior of the property

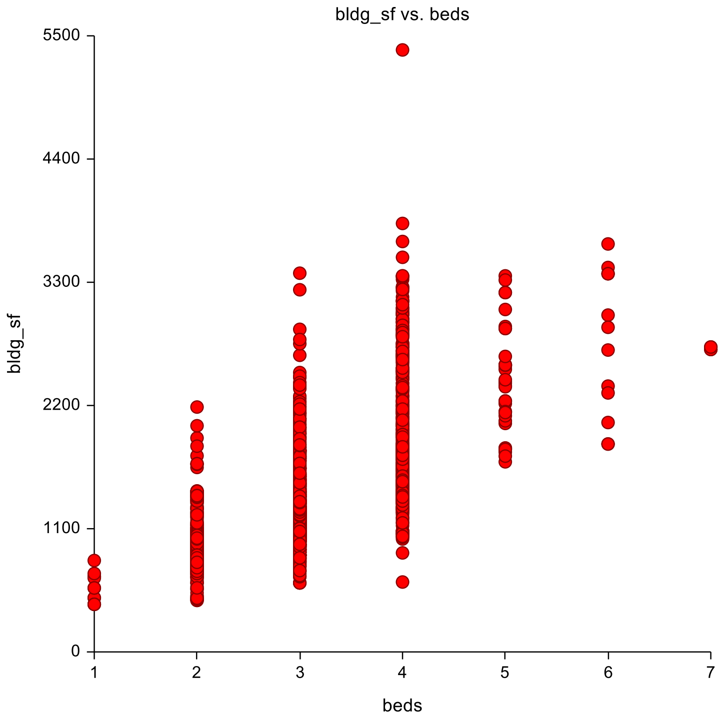
3. Select stratification(s) to review

   Jurisdiction: Assessor Data

   Market Area: Township A

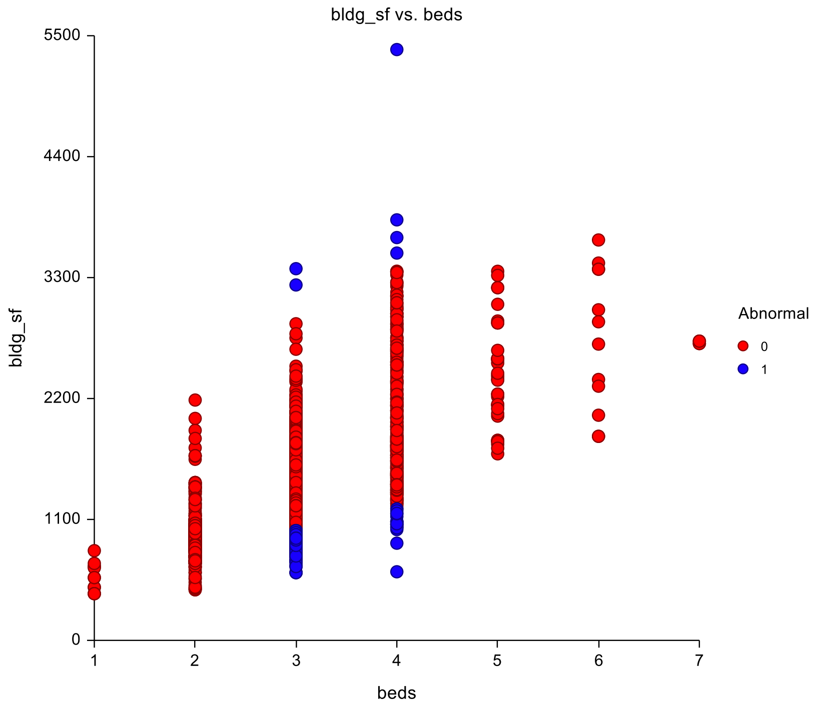
4. Save a snapshot of the data set

5. Plot the data in Step4



6. Identify abnormal data

Beds	Minimum sq ft	Maximum sq ft
1	400	1,100
2	600	2,400
3	1,000	3,000
4	1,200	3,500
5	1,400	3,500
6	1,600	3,700
7	2,300	6,000



7. Summarize abnormal results

Beds	Bldg_sq ft Count	Bldg_sq ft Mean	Bldg_sq ft Median	Bldg_sq ft Min	Bldg_sq ft Max
TYPICAL					
1	12	575	571	425	816
2	615	947	942	460	2,186
3	1,877	1,267	1,200	1,000	2,880
4	410	2,007	1,838	1,200	3,358
5	37	2,410	2,400	1,697	3,357
6	12	2,739	2,797	1,856	3,642
7	4	2,711	2,711	2,700	2,722
ABNORMAL					
3	283	922	931	616	3,381
4	24	1,558	1,078	624	5,374

Pin	Beds
19061030070000	1
19061050070000	3
19061080130000	1
19061100140000	1
19061120070000	1
19061180310000	1
19061190130000	1
19062030480000	1
19062060260000	2
19062080390000	1
19062080430000	1
19062120550000	1
19062130210000	1
19062150450000	3
19062150470000	2
19062150510000	1

APPENDIX I. Edge Case Identification

METHOD OVERVIEW

The Edge Case Identification method is used to identify extreme sales ratios that may stem from many sources: nonmarket influences in the market data, gaps or weakness in the valuation method, and incomplete or incorrect data. While some extreme ratios and variation are typically expected, the presence of systemic error on a data element may heavily affect the quality of the roll and introduce bias when applied to the population. A small number of data elements may have a large influence in all valuation methods, for example, location, size, quality, style, and effective age. Identifying over-representation in extreme sales ratios identifies areas of weakness in the data, which may be used to prioritize review and correction in the population data.

EDGE CASE ANALYSIS, QUALITY GRADE, MARKET AREA 1

Assessment-to-Sales Ratio	Number of Observations	A	B	C	D	E
< 0.49	45	0%	−7%	−26%	16%	17%
0.50–0.59	72	1%	−3%	−30%	19%	14%
0.60–0.69	110	0%	−3%	−17%	18%	2%
0.70–0.79	748	0%	0%	−15%	10%	6%
0.80-0.89	4,156	0%	−1%	−3%	3%	1%
0.90–0.99	13,733	0%	0%	3%	−2%	−1%
1.00–1.09	11,966	0%	1%	2%	−2%	−1%
1.10–1.19	3,913	0%	−1%	−3%	2%	1%
1.20–1.29	1,183	0%	−4%	−8%	7%	4%
1.30–1.39	445	0%	−5%	−10%	11%	4%
1.40–1.49	98	0%	−4%	−14%	9%	10%
1.50–1.59	64	0%	−3%	−18%	13%	8%
> 1.60	223	0%	−3%	−18%	12%	10%

DATA TYPES

Market data, specifically assessment-to-sales ratios and key data elements for each sale at the time of sale.

QUALITY DIMENSION

- Consistency
- Accuracy

PREREQUISITES

Edge Case Identification should be considered when there are concerns about value consistency in assessment across key variables. Edge Case Identification may also be used proactively, to identify data elements with potential for improvement early in the sales review and reassessment processes.

This analysis requires a data set of sales with assessment-to-sales ratios and key data elements for each sale at the time of sale.

DESCRIPTION

- Gather a data set of sales, with assessment-to-sales ratios and key data elements for each sale at the time of sale.
  - Key data elements may include location, size, quality, style, effective age, and any other variable of interest to the assessment jurisdiction.
- For each key variable, create a table counting the number of sales by assessment-to-sales ratios range by categories within the given key variable.
- Using these tables of counts, create a second table for each that calculates the percentage of over- or under-representation. Each cell of these tables is calculated as:

$$\frac{\textit{(Number Observed)}}{\textit{(Ratio Group Total)}} - \frac{\textit{(Category Group Total)}}{\textit{Overall Total}}$$

- Overlay a color heatmap to assist in identifying the areas of over- and under-representation.
- Catalogue which key variables and categories within key variables are over-represented in the extreme sales ratios, keeping mind of the relative sales sample sizes.
  - Optionally, determine a threshold of overrepresentation of, say, 10 percent to qualify as requiring review.

OUTCOME MEASURES

Edge Case Identification provides evidence on the consistency of assessment-to-sales ratios across key variables, identifying areas or categories of sales that appear disproportionately in the extreme ratios.

Using a threshold of, say, 10 percent is a consistent method to determine if a key variable is over- or underrepresented in the extreme sales ratios.

INTERPRETATION OF RESULTS

Some over- and under-representation of extreme ratios for certain key variables may be expected, and the overall sales sample size and number of sales for each category within each key variable must be considered. A threshold for over- or under-representation is recommended, of, say, 10 percent.

Data review focusing on these types of elements may identify missing, inconsistent, or outdated information, and this improves the fit of the estimates. These data elements are also strong candidates for review and update in the population file.

EDGE CASE ANALYSIS, QUALITY GRADE, MARKET AREA 1

Assessment-to-Sales Ratio	Number of Observations	A	B	C	D	E
< 0.49	45	0%	−7%	−26%	16%	17%
0.50–0.59	72	1%	−3%	−30%	19%	14%
0.60–0.69	110	0%	−3%	−17%	18%	2%
0.70–0.79	748	0%	0%	−15%	10%	6%
0.80–0.89	4,156	0%	−1%	−3%	3%	1%
0.90–0.99	13,733	0%	0%	3%	−2%	−1%
1.00–1.09	11,966	0%	1%	2%	−2%	−1%
1.10–1.19	3,913	0%	−1%	−3%	2%	1%
1.20–1.29	1,183	0%	−4%	−8%	7%	4%
1.30–1.39	445	0%	−5%	−10%	11%	4%
1.40–1.49	98	0%	−4%	−14%	9%	10%
1.50–1.59	64	0%	−3%	−18%	13%	8%
> 1.60	223	0%	−3%	−18%	12%	10%

The heatmap visually summarizes the results. Green negative numbers indicate that quality grade C sales are under-represented in the extreme ratios—fewer quality grade C sales than expected if all ratios were evenly distributed. The red positive numbers indicate that quality grades D and E are over-represented in the extreme ratios, implying that estimates both over- and underestimate these elements at a rate higher than that for quality grades A, B, and C. The assessed values for quality grade D and E sales do not capture the sale rice as consistently as the other quality grades, implying a potential need for data improvement on these sales and the population data for quality grade D and E properties.

DETAILED DESCRIPTION

To calculate the heatmap values, create a table with the number of observations in each category with row and column totals.

EDGE CASE ANALYSIS, QUALITY GRADE, MARKET AREA 1

Assessment-to-Sales Ratio	Number of Observations	A	B	C	D	E
<0.49	45	0	1	23	12	9
0.50–0.59	72	1	4	34	21	12
0.60–0.69	110	0	7	67	31	5
0.70–0.79	748	4	64	468	150	62
0.80–0.89	4,156	29	335	3,101	544	147
0.90–0.99	13,733	54	1,271	11,047	1,121	240
1.00–1.09	11,966	48	1,205	9,481	1,023	209
1.10–1.19	3,913	17	327	2,937	494	138
1.20–1.29	1,183	2	65	827	210	79
1.30–1.39	445	0	18	301	96	30
1.40–1.49	98	0	5	62	19	12
1.50–1.59	64	0	4	38	15	7
>1.60	223	0	14	132	49	28
Total	**36,756**	**155**	**3,320**	**28,518**	**3,785**	**978**

The percentage of over- or under-representation in a group is calculated as

$$\frac{\text{(Number Observed)}}{\text{(Ratio Group Total)}} - \frac{\text{(Category Group Total)}}{\text{Overall Total}}$$

For example, the calculation for grade Ds in the ratio group of 0.70–0.79 is

$$\frac{150}{748} - \frac{3,785}{36,756} = 10\%$$

There are 10 percent more grade D sales with ratios between 0.70 and 0.79 than the assessor would expect if grades and ratios were evenly distributed throughout all groups.

APPENDIX J. General Data Quality Score

METHOD OVERVIEW

The General Data Quality Score (GDQS) method allows for comparison of data quality across different sets of data using a single measure ranging from 0 to 100. This method may be used to grade the quality of data after a careful audit or inspection of properties has been conducted in the field.

DATA TYPES

The GDQS method is useful for property data, especially the property's characteristics of improvements and land.

QUALITY DIMENSIONS

The GDQS method is suitable to characterize the overall quality of data, that is, a global measure of quality along all four dimensions of quality as defined by this standard: Accuracy, Currency, Consistency, and Completeness.

PREREQUISITES

Clear guidelines defining conditions under which a particular data element meets the standard should exist.

DESCRIPTION

1. Randomly select a set of properties to audit. The sample should be a random selection in order to obtain a general characterization of the quality of the data. It is important to select the properties in a way that it does not bias the GDQS. For example, selecting properties that have recently been audited may bias the GDQS upward, since errors are less likely in this set. On the other hand, selecting properties in an area with known data quality issues may bias the GDQS downward. See the details below for a procedure for sampling randomly using excel.

2. Conduct an audit of the properties' characteristics.

3. For each property, and each characteristic, compare the results of the audit to the previous information on the property.

4. Determine whether each characteristic of each property met each element of the standard prior to the audit: Accuracy, Currency, Consistency, and Completeness.

5. Calculate the GDQS of this sample by totaling the number of Yes responses to Step 4 above, and dividing by four times the number of properties in the sample.

6. Optional: If there is variation in the number of data elements per property, the DQS may be weighted to better characterize the data quality of properties.

INTERPRETATION OF RESULTS

The GDQS is a general indicator for data quality of a data set, with 0 indicating that no data in the sample meets the standard in any of the four measures, and 100 indicating that all data in the sample meet the standard in each of the four measures. The table below shows how GDQS characterizes general data quality; it provides general guidance and may be adjusted up or down by a jurisdiction to reflect particular circumstances of the jurisdiction.

GDQS Score	Characterization
≥ 90%	Good
≥ 80% and < 90%	Fair
< 80%	Poor

DETAILS

Random Sampling

In order to take a random sample of any number, n, properties from some overall population, *N*, using Excel, take the following the steps:

1. Load the entire set into Excel.

2. Generate a new column containing a random number =RAND()

3. Sort data on this column. Take the first *n* rows as the random sample.

Jaccard Similarity

The GDQS makes use of the Jaccard similarity measure, which constructs an index of shared set membership, standardized by the potential number of set elements that may be shared. For data quality measurement purposes, the comparison is from existing data to ideal data. As such, the existing data are compared to the ideal state, that they are accurate, current, consistent, and complete.

Mathematical Definition

Ideal data are accurate, current, consistent, and complete, containing all four. As such, the ideal set for datum is

- **Ideal Datum:** An assessment jurisdiction's existing data may be accurate, current, consistent, and complete. As such, the set for existing datum is

$$I = \{\, accurate,\ current,\ consistent,\ complete \,\}$$

- **Existing datum:**

$$E = \{e | e \in \{\, accurate,\ current,\ consistent,\ complete \,\}\}$$

Therefore, the Jaccard similarity measure for a data element is

$$J = \frac{(E \cap I)}{(E \cap I)}$$

Relatedly, if the Jaccard similarity measure captures how close data are to ideal, $1 - J$ is how far they are from ideal.

- The **GDQS** is

$$Score = \frac{1}{n} \sum_{i=1}^{n} J_i$$

Similarly, the **General Data Quality Gap (GDQG)** is:

$$Gap = \frac{1}{n} \sum_{i=1}^{n} (1 - J_i)$$

Practical Example

Parcel	Datum	Actual Value	Existing Value	Audit Notes	Accurate	Current	Consistent	Complete	J_i
1	Living Area	2300	2300	Measured using accepted/local standard	1	1	1	1	$\frac{4}{4} = 1$
1	Land Area	4800	4800	Measured using County standard	1	1	1	1	$\frac{4}{4} = 1$
1	Percent Good	0.84	0.9	Depreciation miscalculated and out of date	0	0	0	1	$\frac{1}{4} = .2$
1	Construction Grade	A	B	Grade not consistent with County standard	0	1	0	1	$\frac{2}{4} = .5$
2	Living Area	1850	1825	Measuring using County standard	1	1	1	1	$\frac{4}{4} = 1$
2	Land Area	43550	43000	Not measured using County standard	1	1	0	1	$\frac{3}{4} = .7$
2	Percent Good	0.79	0.79	Depreciation miscalculated and out of date	0	0	0	1	$\frac{1}{4} = .2$
2	Construction Grade	B	—	Grade missing	1	1	1	0	$\frac{3}{4} = .7$
2	Frontage	660	660	Measured using County standard	1	1	1	1	$\frac{4}{4} = 1$

It is important to establish clear guidelines that help auditors determine whether a data element matches each of the four criteria. In the example data, missing construction grade on parcel 2 is not counted as an accuracy defect, just a completeness defect. On parcel 1, where construction grade is not consistent with the way the jurisdiction grades parcels, it is treated as both an accuracy and consistency defect. Construction grade is counted as inaccurate, but land area is close enough to be considered accurate. These decisions are within the standard so long as they are consistently applied throughout the audited sample.

In the data above, the GDQS can be calculated in two ways. First, treating all data elements equally:

$$GDQS = \frac{1 + 1 + 0.25 + .5 + 1 + 0.75 + 0.25 + 0.75 + 1}{9} = \frac{6.5}{9} \approx 0.722 \equiv \text{Poor quality data}$$

Similarly, the GDQG is

$$\sim 1 - 0.722 = 0.278$$

Because parcels can have differing numbers of data elements, it may be more useful to equally weight parcels, rather than datum. In this case the GDQS is

$$\frac{\frac{1+1+0.25+0.5}{4}+\frac{1+0.75+0.25+0.75+1}{5}}{2}=\frac{1.4375}{2}\approx 0.719$$

And Gap is

$$\sim 1-0.719=0.281$$

Making the Measures Comparable Across Jurisdictions

Oversight bodies may use the GDQS to compare jurisdictions. For an accurate comparison, oversight bodies must provide clear guidance on the auditing procedures for each data element, and weighting schemes in calculating the GDQS.

APPENDIX K. Impact Identification

METHOD OVERVIEW

The Data Quality Impact Identification method simulates the impact of low-data quality on roll quality measures. Methodically injecting increasing amounts of errors into a known sample reveals the impact of data quality on statistics like COD and PRD. This allows the prioritization of certain data elements for further investigation.

DATA TYPES

This method is useful for any characteristics of properties.

QUALITY DIMENSIONS

This method characterizes the overall quality of data without specific reference to any of the dimensions of data quality outlined in this standard.

PREREQUISITES

Property characteristics must be stored in a machine-readable format. Staff must be able to perform regressions and Monte Carlo simulations.

DESCRIPTION

1. Select a representative sample of properties for regression modeling. The sample should be large and diverse enough to characterize the jurisdiction's modeling set within a specific group (e.g., residential, residential condominiums, retail), but not so large as to unnecessarily prolong computation times.

2. Select a reasonably high-quality modeling approach that yields good ratio statistics.

3. Identify the characteristics of interest to include in the study. These can be the largest determinants of value, or the most likely to be low quality.

4. For each characteristic, identify the most likely types of errors. For example, it may be more likely that finished basements in reality are coded as unfinished in the modeling sample.

5. Create a script that iteratively simulates each error identified in Step 4 for each corresponding characteristic identified in Step 3. The script should simulate a small error at first, and increase the frequency and magnitude of the error with each iteration.

6. On each iteration
   - Replace the original modeling data with the intentionally corrupted data.
   - Capture a measure of the amount of data intentionally corrupted.
   - Calculate sales ratio statistics, and save them.

7. The data resulting from following Steps 1–6 contains sales ratio statistics corresponding to different levels of data corruption. Regressing a sales ratio statistic against the corruption indicator results in a coefficient that shows the marginal impact of an additional 1 percent of data corruption on the sales ratio statistic of interest.

8. Ranking the characteristics from Step 4 by the marginal effects estimated in Step 7 gives a ranking of the relative importance of data corruption in each of the characteristics chosen in Step 4. This can be used to direct resources to data quality issues, and to discuss such issues with internal and external stakeholders.

OUTCOME MEASURES

Typical values for this procedure range between 0 and 0.5. Values outside of this range should be disregarded.

INTERPRETATION OF RESULTS

The coefficients estimated in Step 7 show the marginal impact of a 10 percent decrease in data quality in any given data vector on sales ratio statistics. Suppose a regression of COD on the percentage decrease in data quality for building square footage is 0.5. This would indicate that a 10 percent decrease in data quality in that characteristic lowers COD by a half a point.

The ranking produced by Step 8 may be used to select attributes for data-cleansing efforts.

DETAILS

Finding incorrect data can be costly. In order to maximize the returns of resources spent finding and correcting bad data, an assessor may want to prioritize types of data that, when incorrect, have a large impact on IAAO sales ratio statistics. This appendix outlines one method an assessor can use to determine which data fields they may want to prioritize for review.

To fix ideas, consider a binary variable that takes one of two values: yes, no. Each observation of this variable may be accurate or inaccurate in reality. The assessor can construct a reference table to show the overlap between data and fact:

	Real = 1	**Real = 0**
Data = 1	Accurate	False positive
Data = 0	False negative	Accurate

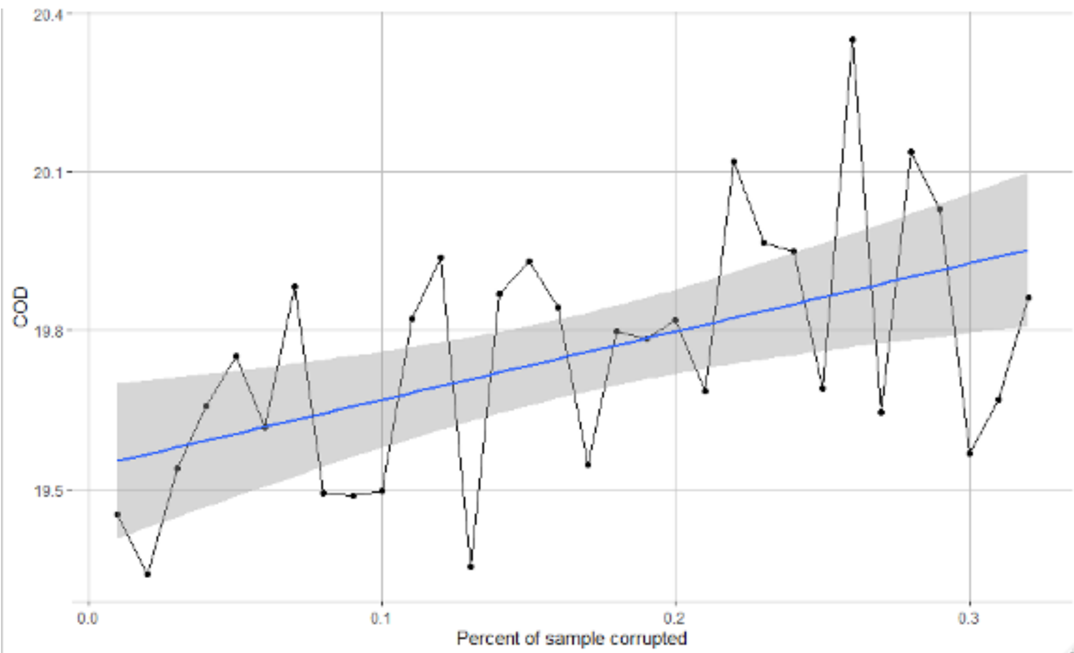
All the data on this particular variable fall into one of these four categories. However, it may be difficult, or even impossible, to determine what share of the data fall into each category. In order to get a better idea of the impact of corruption in this particular variable, the assessor can intentionally introduce increasing amounts of corruption and measure the impact on ratio statistics.

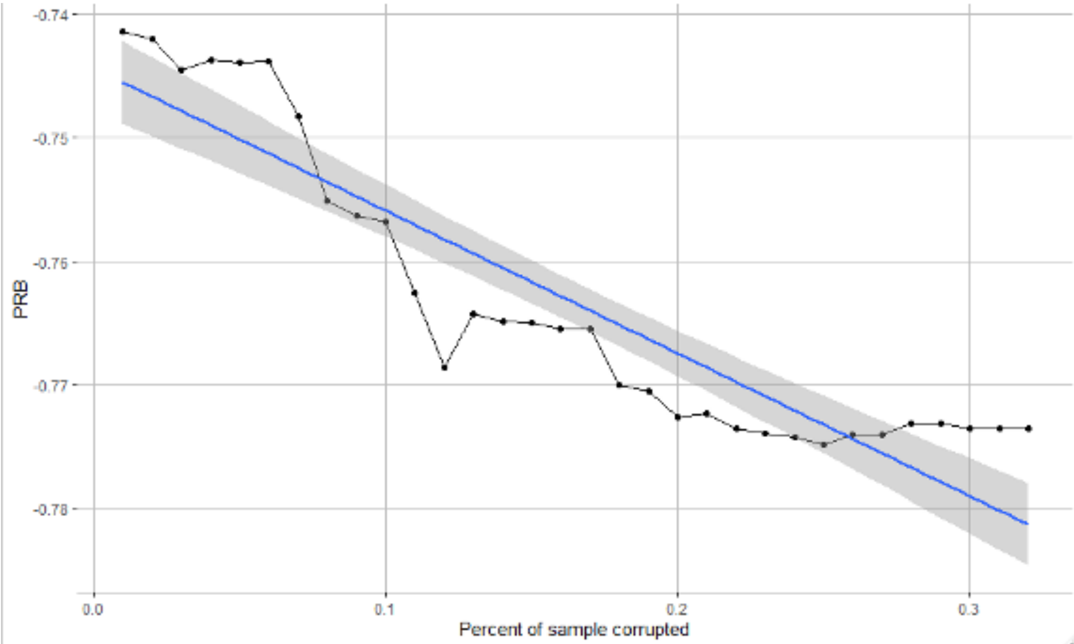
Data may be more likely to contain false negatives than false positives, since taxpayers are more likely to alert the office in the event that it has a false positive rather than a false negative.

At baseline, about 32 percent of homes in the sample have a finished basement. To assess the impact of data corruption in this field, the assessor may iteratively and intentionally introduce errors into the data, predict property values, and measure the IAAO statistics. Starting at 1 percent of the sample, and increasing in intervals of 1 percent, the assessor may randomly switch off the indicator for finished basement for an increasingly large amount of the sample. The assessor may then re-run a straightforward Ordinary Least Squares [OLS] model, predict property values, and measure the sales ratio statistics.

The following graphs show the COD and PRD calculated at increasing levels of data corruption. The slope of the fit line in each graph indicates the impact of data corruption on each statistic. Roughly, a 10 percent increase in data corruption in this variable, in this sample, corresponds to a 0.5 percent increase in COD, and a –1.3 percent impact on PRB.

SIMULATED COD WITH INCREASING DATA CORRUPTION: FINISHED BASEMENT





How does this effect compare to a more *important* predictor, like building square feet? To see this, the assessor performed a similar procedure on the same sample data. First, the assessor randomly selected an increasing percentage of the sample to introduce an error to. Then, to each property selected for an error, the assessor added a random error ~N(–150, 300) to building square feet. To put this into context, the mean building size in the sample is 1,259 ft. The error is roughly a 10 percent error.

Again, with each round of corruption, the assessor predicted home values and calculated IAAO ratio statistics. The results are presented below. A 10 percent increase in the rate of error is associated with a 0.6 percent increase in the COD, and a 1.5 percent decrease in the PRB. Comparing these results with the one above might lead an assessor to conclude that they should focus their resources on whichever variable is less costly to correct, since they have roughly the same impact on ratio outcomes.

Now consider an unobservable vector (real), which holds the truth about the world. Suppose the assessor constructs modeling data such that $y = 250{,}000 + 10{,}000 \times (\text{data}) + e$, where $e \sim N(0, 500)$. Estimating this model recovers the constructed parameters.

MODEL INFO

Observations: 10,000

Dependent Variable: df$y

Type: OLS linear regression

MODEL FIT

$F(1,9998) = 649139.41$, $p = 0.00$

$R^2 = 0.98$

Adj. $R^2 = 0.98$

Standard errors: OLS

	Est.	S.E.	*t* val.	*p*
(Intercept)	250016.42	5.66	44163.33	0.00
df$data	9991.41	12.40	805.69	0.00

Example 1. *Classical* measurement error is a mean-zero (unbiased) error. Classical error will bias estimated coefficients toward 0. Introducing an error rate of 10 percent into the constructed data reduces the magnitude of the estimated effect by about 30 percent in the constructed data set.

```
MODEL INFO
Observations: 10,000
Dependent Variable: df$y
Type: OLS linear regression

MODEL FIT
F(1,9998) = 11710.00, p = 0.00
R² = 0.54
Adj. R² = 0.54

Standard errors: OLS
----------------------------------------------------
                  Est.    S.E.   t val.     p
---------------- ----------- ------- --------- ------
(Intercept)      250311.41  32.30  7750.30  0.00
df$data          6834.47   63.16   108.21  0.00
----------------------------------------------------
```

Example 2. In practice, errors are likely not classical. This is so because taxpayers have a greater incentive to point out errors when they result in higher estimates of their property values. It is likely that the typical error for most offices is negatively biased. Using the constructed data from above, the assessor introduced a biased error. A 1 percent error rate applied to properties selected for intentional data corruption had a strong impact on the estimated slope coefficient

MODEL INFO

Observations: 10,000

Dependent Variable: df$y

Type: OLS linear regression

MODEL FIT

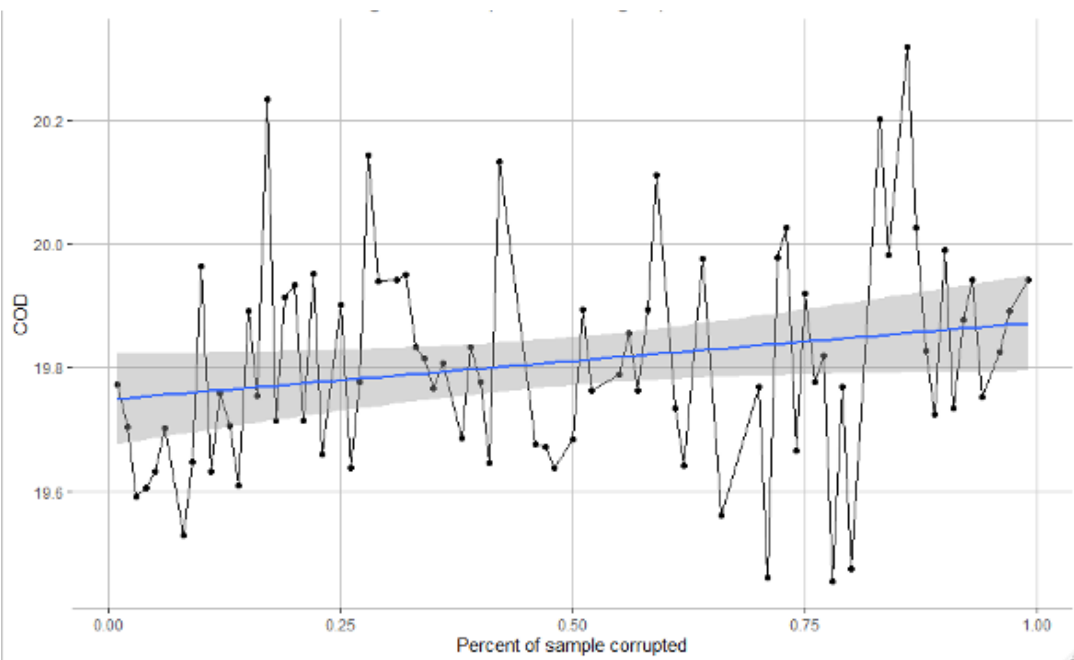
F(1,9998) = 63854.14, p = 0.00

R^2 = 0.86

Adj. R^2 = 0.86

Standard errors: OLS

```
------------------------------------------------------
                        Est.     S.E.    t val.      p
---------------- ----------- ------- ---------- ------
(Intercept)        250235.59   15.94   15700.75   0.00
```



Example 3. By repeating example 2 for ten selected characteristics, the assessor can see the impact of data corruption on COD. The following table shows the slope coefficient for ten different characteristics intentionally corrupted in a sensible way. Based on this table, a data-cleansing exercise should focus on ensuring that transaction dates are accurate, and that land square footage and building age are accurate. Improving data quality in these areas leads to larger improvements in CODs.

Variable	Coefficient
Land square feet	0.33
Sale quarter	0.28
Building Age	0.2
Basement finish status	0.18
Full Bathrooms	0.08
Building square feet	0.08
Neighborhood code	0.02
Half bathrooms	0.02
Room count	0
Garage size	0.03
Attic finish status	0.06
Sale quarter of year	0.07

ASSESSMENT STANDARDS OF THE INTERNATIONAL ASSOCIATION OF ASSESSING OFFICERS

Guide to Assessment Standards

Standard on Assessment Appeal

Standard on Automated Valuation Models

Standard on Contracting for Assessment Services

Standard on Data Quality

Standard on Digital Cadastral Maps and Parcel Identifiers

Standard on Manual Cadastral Maps and Parcel Identifiers

Standard on Mass Appraisal of Real Property

Standard on Oversight Agency Responsibilities

Standard on Professional Development

Standard on Property Tax Policy

Standard on Public Relations

Standard on Ratio Studies

Standard on Valuation of Personal Property

Standard on Valuation of Property Affected by Environmental Contamination

Standard on Verification and Adjustment of Sales

TO DOWNLOAD THE CURRENT APPROVED VERSION OF ANY OF THE STANDARDS LISTED ABOVE, VISIT IAAO.ORG