Fannie Mae

Beyond the Guide

Section 8

QC in a Data-Driven Environment

Section 8

QC in a Data-Driven Environment

Selling Guide D1-3-03

Technology is transforming the financial world around us, and quality control (QC) must find ways to adapt to the transformed landscape. New terminology, methods, and even currency have become accepted, and being "data-driven" is a necessity.

Fannie Mae regularly surveys the market to better understand its changes. A 2018 survey of mortgage lenders' leadership revealed the primary objectives of investing in artificial intelligence/machine learning (AI/ML) were to improve operational efficiency and enhance consumer/borrower experience. That focus is wise since a 2021 homebuyer survey[1] revealed that 62% of homebuyers were *'Somewhat Interested'* or *'Very Interested'* in a more digital or fully digital mortgage process. To compete and stay profitable, mortgage lenders must adapt and transform their way of working. In the past decade, many elements of the mortgage manufacturing process have become partially or fully automated.

Task		Technology solution
Document management	→	Scanners, optical character recognition (OCR) including AI/ML
Loan application data entry	→	Web-based e-forms that consumers complete
Executing disclosures, notes, and deeds of trust	→	Secure e-sign and e-note technology
Employment and income verification	→	Automated underwriting system (AUS) validation of income/employment data direct from source
Income processing and analysis	→	Income analysis software with upload and integration capability
Asset verification and payment calculations	→	Vendor and AUS software with AI that obtains and calculates sufficiency of assets to close and recurring payments for rent and other debts
Collateral risk management	→	Leverage standardized appraisal data (based on the Uniform Appraisal Dataset) to enable risk-based, data-centric approaches to real estate appraisals as well as reporting available in Collateral Underwriter® (CU®), like CU Risk Score, CU Risk Flags, and CU Messages, to understand your current pipeline risk.

These advances continue to improve methods of doing business and introduce new ways to think about loan manufacturing quality.

---

[1] National Housing Survey: COVID-19, Mortgage Digitization, and Borrower Satisfaction Q12021

Current state

QC testing is still performed largely to test physical or scanned copies of paper files. Through manual reverifications and reviews, QC is designed to ensure the AUS data is accurate and valid, and that the information relied upon for the loan decision is factual.

Future state

QC needs to evolve as industry practices do. QC testing should add new methods of testing data to traditional reverification practices as more loan data is received digitally from new and increasing sources. Caution should be used to ensure we are not lulled into assuming that because data is obtained electronically it meets the needs or expectations for data accuracy. QC should constantly evolve to test new risks, which includes understanding the accuracy and reliability of information that is digitally sourced. Once accuracy and reliability of new data sources are confirmed, data aggregation and data scanning techniques will allow for richer exception-based sampling. This enables a stronger ability to detect defects.

Data integrity

As more of the mortgage process migrates toward digital solutions, managing and testing data integrity becomes more critical to loan quality and QC processes. Before testing can occur, QC sampling processes rely on the lender origination loan data (e.g., loan product, purpose, channel, etc.) to ensure accurate sampling for post-closing reviews.

> **QC should constantly evolve to test new risks, which includes understanding the accuracy and reliability of information that is digitally sourced.**

Data gaps or inaccuracies can cause insufficient sampling, resulting in adjusted testing, reporting, and investor noncompliance. An important part of the QC review is checking data integrity, which focuses primarily on confirming the correct information was entered into Desktop Underwriter® (DU®). As more digital solutions – like Fannie Mae's DU validation service – are introduced to the loan origination process, QC must reexamine what checking data integrity encompasses. The more data that is delivered via intermediary sources, the more opportunity exists for variation within the data.

Lenders with a high level of confidence in the accuracy of their data can feel comfortable that the testing and reporting based on that data is a true and accurate reflection of their risks and opportunities. For QC managers, understanding your organizations' loan data integrity confidence is critical to understanding data-quality risks and how to appropriately test for them.

Finding data

Lenders collect and house more loan data than ever before. Since much of it is standardized under the Uniform Mortgage Data Program® (UMDP), it can be used to help QC explore new methods of leveraging currently stored data to enhance overall loan quality. Having access to more data can provide a view of:

- patterns of activity between loan parties not visible from the loan-level view
- data discrepancies between segmented departments like origination, delivery, and servicing
- data that is inconsistent with expected results
- how data quality impacts costs

Since current and future technology systems are highly dependent on accurate data, the value of leveraging this stored data is contingent on its accuracy and standardization. QC practitioners must continue to increase and broaden their understanding of how data is obtained and stored so that testing keeps pace with the true sources of all data and how it impacts QC.

Leveraging stored data from third-party sources such as digital income/employment vendor reports is an excellent example of data that provides insight into loan manufacturing data accuracy and integrity. Study the linked materials in the Resources section to understand what standardized loan data is currently required of lenders delivering loans to Fannie Mae.

Internal data standardization

Data standardization is about making sure that data is internally consistent. This means each data type has the same label, format, and meaning. The implementation of UMDP streamlined data standardization for the mortgage industry, but lenders should be highly motivated to find more ways to standardize as much of their data as possible. Standardized values help track data that is not easy to compare. When the same address is formatted differently – such as 123 North Maple Street versus 123 N. Maple St. – this one data point becomes two. If property addresses are standardized by using a single source of truth when populating data into the origination system (e.g., U.S. Postal Service or title data), the advantages of standardized values become evident. Investing resources into the data standardization process provides benefits such as:

- Systematically testing origination files for compliance against established guidelines.

- Identifying correlations between the manufacturing loan quality of various new credit policies and the pricing implications.

- In conjunction with fraud detection software, identifying bad actors in transactions and stopping impacted loans from closing before human review.

These are only a fraction of the benefits of expanding data standardization beyond the industry minimum requirements established by UMDP.

QC's role in data integrity

QC's role in managing data integrity should be to identify instances and root causes of data discrepancies within the scope of QC's activities and to help the organization develop plans to mitigate or eliminate those discrepancies. A good example of new data sources that may need testing is digital third-party income/employment/asset reports. As lenders search for ways to gain efficiency while maintaining accuracy, management is more likely to expand their use of digital solutions. QC's testing methodology must evolve to identify potential types of new deficiencies that may not be visible on flat documents and to help create ways to remediate those deficiencies and underlying issues. This requires understanding the various considerations of digital source data used today.

QC and DU validation services

An excellent example of the balance between digital solutions and current state QC testing is DU validation services. Lenders that opt in to use the services may be relieved of certain QC requirements, which can result in reduced costs and improved efficiencies. It is important to note that lenders always remain fully responsible for reviewing the whole file and reconciling any contradictory or conflicting information. What does contradictory or conflicting information look like? Consider these examples that QC can identify:

- Output from data analysis tools conflicts with vendor employment or income report.

- Automated Clearing House payroll deposits on bank statements or vendor asset reports do not align with a report (such as from The Work Number) used by DU for income validation or paystubs in the loan file.

These examples help highlight when discrepancies are driven by information the system cannot read or assess versus actual data discrepancies. Day 1 Certainty® resources include eLearning examples, useful insights, and training.

Using QC data

Digital solutions and their potential efficiencies aren't exclusive to loan production or the consumer experience. QC must consider ways to leverage the digital environment and stored data to increase QC process efficiencies, identify quality issues more effectively, and reduce costs by managing and analyzing loan data in new ways. In recent years, increased use of Fannie Mae's standardized data sources has surfaced new benefits to lenders, such as:

- lenders using the DU validation service could see a reduction of some reverification costs, and

- decreased need for field reviews as a result of implementing strong collateral risk assessment practices.

These benefits are specific to Fannie Mae, but how can lenders find other ways to use QC data with their own resources and data?

☑ Aggregate the data

Each lender houses volumes of loan data that may be stored and managed in various areas of an organization. Because most closed loan data must be standardized, aggregating the origination and QC data into a single location (database or spreadsheet) creates a powerful resource for identifying and assessing known and unknown risks. If you can supplement this data with investor reviews and their results, that resource becomes even richer.

☑ Build rules

Leverage known risks such as common defects or layered credit risks to build rules to search the data for loans with similar attributes. Once identified, these loans can be tested sooner and more strategically to better identify loans with higher risks. This focuses limited QC resources on loans with the highest likelihood of having defects.

☑ Leverage current tools to inform sampling

If OCR software with data scraping (text extraction) is currently used in operations or document management, consider repurposing the software for QC as a method of increasing confidence in loan data and targeted discretionary sampling, like this example:

- Run 100% of each month's loan production through OCR software to find loans containing both a verification of income and a borrower paystub. Use data scraping to capture the income amounts from both documents.

- Transfer the captured data to a calculator that reconciles the captured income amounts. Loans with discrepancies can be reviewed as part of a QC targeted component review.

- Loans run through the calculator without a discrepancy might be reported as having a higher income confidence level. As an added benefit, this process can measure the effectiveness of the OCR/data scraping software.

☑ Design processes to address outliers

As rules are designed to sort loans more strategically for targeted sampling, specific methods of reviewing and testing can be designed to improve the effectiveness and efficiency of the targeted QC reviews.

> QC must consider ways to leverage the digital environment and stored data to increase QC process efficiencies, identify quality issues more effectively, and reduce costs by managing and analyzing loan data in new ways.

Increase QC digital awareness

Numerous free resources exist that can help expand QC's understanding of the evolving digital market and how data analysis methods can be streamlined through more efficient ways of using readily available tools such as Excel and Power BI. Learning to use these tools can increase efficiency of analysis and reporting. Consider these methods:

- YouTube – search for Excel or Power BI training. Nearly all are free, and many can be completed in 30 minutes or less.

- Microsoft – offers free training for both Excel and Power BI.

- Industry publications – online mortgage publications and blogs can provide insight on new and emerging digital trends.

- Encourage staff to be curious – QC managers can create a constant learning environment by having each employee demonstrate a new technology skill or topic monthly or quarterly that is shared with the team as a part of employee goals.

- Fannie Mae's Loan Quality page – contains links to training, FAQs, and other resources that can expand knowledge and insight on navigating a digitized QC landscape.

Data-driven QC

A data-driven QC department maintains a focus on precision to identify and avoid current and future quality issues. QC should capture all available data, not just the data we use today. It has the power to inform business decisions by using data for continuous improvement. Data-driven insight helps inform action plans to reduce current defect levels and can also help identify QC opportunities when changes are on the horizon.

> Data-driven insight helps inform action plans to reduce current defect levels and can also help identify QC opportunities when changes are on the horizon.

Resources

2018 survey "How Will Artificial Intelligence Shape Mortgage Lending?"

Selling Guide D1-3-03

Uniform Mortgage Data Program (UMDP)

Uniform Loan Delivery Dataset (ULDD)

Uniform Residential Loan Application (Form 1003)

Uniform Closing Dataset

Uniform Appraisal Dataset (UAD)

2022 Appraisal Modernization Article

Desktop Underwriter Validation Service

Day 1 Certainty resources

Loan Quality page