



Assessing data quality is not only about identifying missing data. It is knowing whether deficiencies will materially impact insured losses. Do assessments adequately address the details in the data? Are metrics consistent and objective? Can they systematically correlate to loss?

The RMS® Data Quality Toolkit identifies key drivers of data quality and assesses their impact on modeled results, allowing users to enhance data where it has the greatest impact on risk pricing and portfolio management.

Data Quality Toolkit

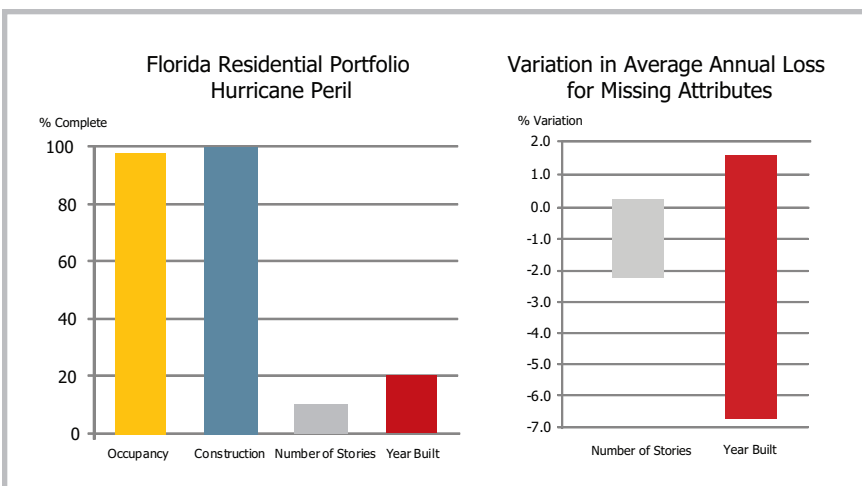
The RMS Data Quality Toolkit evaluates and enhances data quality, helping to reduce uncertainty around loss estimates and improve risk differentiation for more risk-appropriate decisions.

IDENTIFY DRIVERS OF DATA QUALITY

When assessing data across thousands of locations, it can be simple to identify where building attributes are missing or incomplete, but harder to determine if these discrepancies impact expected loss. Data Quality Toolkit metrics evaluate data based on geocoding and vulnerability characteristics most relevant for perils and regions. Unlike in-house or third-party metrics, RMS evaluates data quality based on the science used to build the models, offering previously unattainable insight into underwriting and pricing decisions:

KEY FEATURES AND APPLICATIONS

- Improve the accuracy of loss results by identifying drivers of data quality using RMS metrics
- Examine data credibility and build intuition around data assumptions and coding practices
- Evaluate and enhance data with the RMS® ExposureSource database
- Benchmark and compare data from the aggregate to the location level using reports and sensitivity tests



- Geocoding metrics assess the relative importance of geocoding resolution based on the severity and gradient of hazard (e.g., would capturing data at street address level improve risk assessment)
- Vulnerability assessments evaluate the impact of missing building characteristics weighted by hazard and location

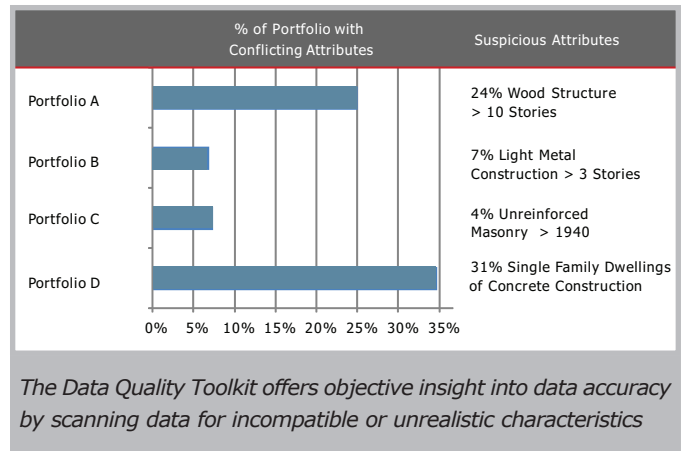
Results can be viewed at the aggregate level, and users can drill down to identify the data quality issues that are driving modeled loss estimates.

Because most Florida residences are single-story, missing data on number of stories has a minimal impact on model results; however, missing data on year built significantly increases uncertainty around expected loss

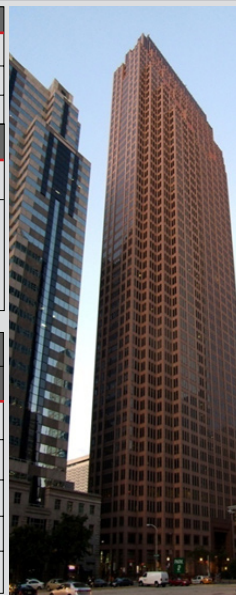
EXAMINE DATA CREDIBILITY

The Data Quality Toolkit allows users to identify unintentional errors in data as well as validate data assumptions based on organizational practices:

- Identify unusual or illogical combinations of building, financial, and policy attributes using RMS or user-defined rules
- Compare data to industry standards to flag potentially aggressive or conservative coding of building characteristics



ExposureSource Database Attributes for Bell Atlantic Tower					
Address		1717 Arch St, Philadelphia, PA 19103			
Building Name		Bell Atlantic Tower			
Building Area		1,200,000 sq ft			
RMS Data	Construction	Occupancy	Number Stories	Year Built	AAL Ratio
Exact Match	Steel Frame (4A)	Office	55	1989	100%
Additional Acceptable Values	Moment Steel Frame	General Commercial Conference Center Professional, Technical, Business Services	41-120	1989-1994	100%



Client Attributes for Bell Atlantic Tower					
Client	Construction	Occupancy	Number Stories	Year Built	AAL Ratio
1	Unknown	Telephone & Telegraph	Unknown	Unknown	297%
2	Unknown	Communication (Radio and TV)	Unknown	Unknown	297%
3	RM Shear Wall	Telephone & Telegraph	Unknown	Unknown	254%
4	Fire Resistive	Telephone & Telegraph	1	1995	164%
5	Moment Steel Frame	General Commercial	52	1989	100%

Red: Incorrect/Unacceptable Value Black: Acceptable Value

Comparison of client data against the ExposureSource database illustrates that, even for high-profile properties, data and resulting loss estimates can vary significantly. The ExposureSource database applies an intelligent approach to client data analysis—only variations in data that impact loss affect the resulting AAL ratio.

ENHANCE EXISTING DATA

Assess accuracy by comparing data against the ExposureSource database, an extensive database of U.S. commercial and residential properties. Data is validated and assigned confidence thresholds based on field surveys, satellite imagery, and other research, allowing users to make informed decisions when comparing and enhancing existing data. Improve your data by replacing missing or incorrect building attributes with values from the database.

BENCHMARK BASED ON YOUR STANDARDS

The Data Quality Toolkit provides the ability to compare and benchmark data based on your criteria. Portfolios can be grouped for analysis across multiple Exposure Data Modules (EDMs) to measure and report on data quality over time or by groupings such as portfolio, cedant, underwriter, state, county, account, or location.

To examine the potential impacts of data on losses, users can run sensitivity tests based on best and worst case scenarios, RMS recommended rules, or user-selected criteria. Results are generated in much less time than required to run the full stochastic model.

