



Quantitative Methods for QA & QC


MBA Quality Assurance Meeting
October 1-2, 2007

Ali Moazami

amozami@yahoo.com

(917) 670-0963

Discussion Plan

- 
- Quality Assurance
 - Quality Control
 - Data Capture & Access
 - Database Validation
 - Sampling Methodology
 - Model-Based Risk Assessment
 - Main Driver Analysis
 - Correlation
 - Multi-Factor Analysis
 - Descriptive and Analytical Reporting

- Re-underwriting of a subset of approved mortgages by a dedicated underwriting team to:
 - Provide timely and targeted qualitative feedback on the soundness and reasonableness of underwriting judgments
 - Review targeted higher risk transactions identified by screening key data elements
 - specific underwriters
 - branch offices
 - watch list brokers
- Sample drawn to be representative of the broad originated population with over-sampling of layered risk and other targeted groups
- Results tabulated by branch and reviewed by regional production and credit managers
- Equivalent to pre-sale due diligence performed by loan buyers:
 - Typically covers 10-50% of all originations depending on loan type and experience
 - Includes a targeted review of highest risk files
 - Frequently validates key loan characteristics against available databases (AVMs, fraud checks, etc.)

- Detailed reprocessing and revalidation of:
 - Information elements supporting loan decision including collateral, credit, capacity, etc.
 - Soundness of the underwriting decision
 - Adherence to investor and internal guidelines and procedures
 - Completeness and enforceability of closing documents
 - Compliance with all applicable laws and regulations
- Monthly audit of a sample of all originated loans using a statistically representative sample (approximately 4% of all loans originated)
- QC manager typically responsible for producing monthly report describing findings, action items, their resolution and any action plan
- Actions taken by management formally documented by citing each deficiency, identifying the cause of the deficiency, and providing management's response or actions taken

- Successful completion of QA/QC objectives require application of a variety of quantitative methods to:
 - Describe population and its performance
 - Validate information supporting underwriting decision
 - Identify high risk and “abnormal” behavior
 - Relate risk/behavior to potential causes so as to identify action items
- Quantitative Methods broadly grouped into:
 - Data Manipulation
 - Capture
 - Quality & Validation
 - Access
 - Statistical Inference
 - Sampling
 - Risk Assessment
 - “Main Driver” analysis
 - Description and Reporting
 - Conventional analysis and trending

- Electronic capture of key (risk) characteristics
 - “Flat” files
 - Relational Databases
 - Include analytical dimensions
 - Systematic imaging will streamline handoffs
- Data Quality
 - Critical step, often neglected in upstream systems
 - Internal consistency, typically rule-based
 - External validation
- Access
 - Bulk rather than one file at a time
 - Spreadsheet
 - Specialized tools
 - Advanced graphing and visualization
 - Online Analytical Processing (OLAP), “slicing & dicing”
 - Statistical modeling

- Natural extension of internal data quality processes
- May be sample-driven or universal
- Special fraud application in QA/QC
- Keep an eye on potential matching issues
 - Is there any bias towards high-risk files?
- Applicable to many key characteristics:
 - Identity
 - Social Security numbers verified against Social Security Administration, OFAC and Patriot Act files
 - Income
 - IRS's 4506-T
 - Bureau Score
 - Property Value

- Random Sample is required to make an inference about a population without looking at every single file
 - Refers to taking a number of independent observations from the population
 - Allows measures of uncertainty for the estimates
- Types of Random Sample
 - Simple Random Sample: every file has an equal chance of being selected
 - Stratified Sample: Select independent samples from a number of subpopulations, groups or strata within the population to
 - ensure that particular groups are adequately represented
 - improve efficiency by gaining greater control on the composition of the sample
- Methods of producing Random Samples
 - Random number table
 - Random number generator (mathematical algorithms)
- Required Sample Size depends on:
 - Confidence Level: typically 95%
 - Defective Rate: estimated % of loans with adverse QC findings
 - Accuracy Range: acceptable Sampling Error around true Defective Rate is plus or minus x%
 - Size of the audit population

- Distills large number of information elements into simple measures of risk:
 - Rule-based
 - Statistical
 - Pattern Recognition
- Major categories within QA/QC include:
 - Property Assessment
 - Automated property valuation (AVM)
 - Overall transaction risk, based on property and neighborhood characteristics
 - Credit re-verification
 - Comprehensive Transaction Assessment
 - May combine data validation with risk modeling (e.g. WRAP)
- May simplify targeting for QA/QC review, pre or post-funding.
 - Ask for validation studies, general vs. your production
 - Understand the information elements used in the modeling
 - Refine how the tool is applied in your installation
 - Processes & Procedures
 - Custom thresholds
 - Constant tracking and validation

- Conventional analysis typically focuses on a series of two-way correlations
 - Correlation is a number between -1 and +1 that indicates the strength and direction of the relationship between two (random) variables
 - Easy to perform and intuitively appealing but,
 - Correlation is not causation
 - Neglects interaction between multiple drivers
- Instead, multiple regression seeks to learn about the relationship between several independent or explanatory variables and a dependent or response variable
 - Recognizes the complexity of interactions
 - Formal statistical analysis can simplify and enhance the creation of insightful reports and
 - Focus attention on the most powerful action items
- Data Mining is broader analytic process designed to explore large amounts of data in search of consistent patterns and/or systematic relationships between variables

- Management needs to understand:
 - What are the general characteristics of the population under review?
 - Where is it representative of the broader population?
 - What are the main findings?
 - How do they break by the dimensions of interest?
 - How has it evolved over time?
 - Which of the findings may safely be applied to the population at large?
- Most important of all, the conclusions and action items have to flow naturally from the presented analysis