MOODY'S ANALYTICS

KENYA BANKERS ASSOCIATION

Better Faster Decisions

Credit Risk Scoring Perspective on East African Banks

09 February 2020

We play an important role in this world

Mission

Empowering our customers to make better, faster decisions

Vision

To be the global leader in solving critical business problems

Values

Customer focus, excellence, open mindset, and teamwork

Credit Risk Scoring Perspective on East African Banks

Muhammad Jafree
Solutions Specialist
Moody's Analytics

Bogdan Pleshkevich
Economicst/Risk Modeller
Moody's Analytics

Andrew Njeru
Head of Credit
Kenya Commercial Bank

Moody's ANALYTICS

Better Faster Decisions

Retail Credit Scoring

Bogdan Pleshkevich, Assistant Director, Economics & Business Analytics, **Moody's Analytics**

09 February 2020

Agenda

1. Introduction
2. Importance of Credit Scoring for Retail Exposures
3. Credit Scoring: Setting and Challenges
4. Implications and Applications
5. Q&A

1 Introduction

2

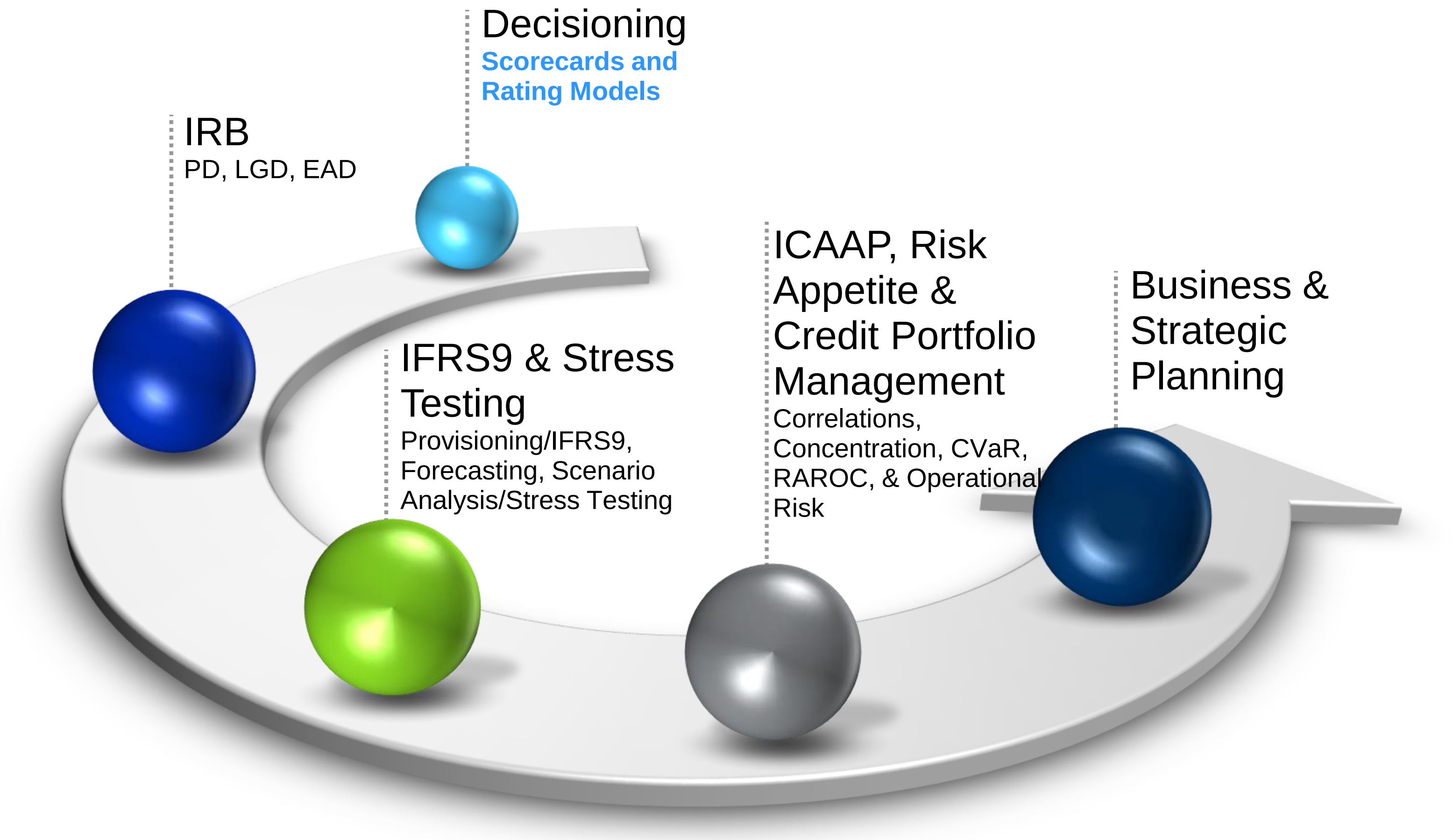
Importance of Credit Scoring for Retail Exposures

Why do we care about measuring Credit Risk?

- **Credit Risk** is associated with a potential event when a counterparty will be unable to meet its financial obligations in full → we need to control arising credit risks
- Proper Credit Risk Management covers
  - Identification, evaluation and **measurement** of Credit Risk
  - Monitoring, reporting and effective communication
  - Actions and efforts to meet regulatory and underwriting standards
- Existence of an efficient set of Credit Risk models is a **must** in the current global situation
- Credit Scoring model is a essential piece of such modelling set

A Holistic Modelling Approach

Risk modelling to drive the business model of a bank



3

Credit Scoring: Setting and Challenges

Scorecard Introduction

- Scorecard models are built to rank order borrowers in terms of their default risk.
- It does not identify “good” or “bad” applications on an individual basis. It provides statistical odds, or probability, that an applicant with any given score will be “good” or “bad”.
- Scorecards are developed using the assumption that “future performance will reflect past performance.” Based on this, the performance of previously opened accounts is analyzed in order to predict the performance of future accounts.

Application Scorecard

Behavior Scorecard

Collection Scorecard

Characteristic	Class	Coeff	Points
LTV	0-50%	-1.258	45
	50-70%	-0.895	33
	70-90%	0.589	-10
	90%+	1.089	-28
Region	North	0.689	-14
	Center	0.213	-2
	South	-0.325	9
Marital Status	Single	0.8	-30
	Married	-0.8	30
Months on Book	<2 years	-0.986	24
	≥ 2 years	0.589	-16
Constant		-3.258	420

Score = 461

Scorecard Types

Data is a key but alternatives are present

Data-based

Approach based on historical (3-5 years) data for Retail exposures.

Requires: High-quality data on defaulted and performing facilities including customer and loan-level characteristics

Hybrid

Approach based both on historical (3-5 years) data and expert surveys

Requires: Part of data on defaulted and performing facilities; Expert surveys on expected default patterns

Judgmental

Approach based exclusively on expert surveys

Requires: Extensive surveys on expected default patterns

Scorecard Model Build Methodology

Six Steps Approach

One — **Sample Building.** To ensure that the data used mirrors the Bank's experience and business.

Two — **Variable Selection.** To consider the most significant variables as potential predictors.

Three — **Binning.** To discretize each variable's information. This ensures simplicity to the tool.

Four — **Model Build**. To select the most important predictors and assign them the correct weight in predicting the default event.

Five — **Model Performance.** To assess the capacity of the model in discriminating good and bad accounts.

Six — **Model Validation**. To validate the efficacy of the model on independent samples.

Scorecard Key Components

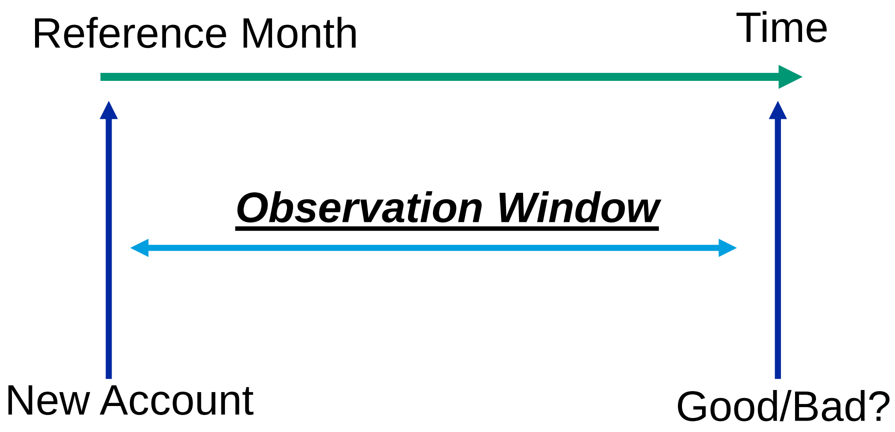
Target Variable

The target variable is the **default flag.** The flag is defined internally in each institution.

- » Days past due
- » Unlikeliness to pay
  - – Forbearance
  - – Litigation
  - – Deceased
  - – Bankruptcy

Observation Window

Gather data for accounts opened during a specific time frame, and monitor their performance for **time window** to determine if they were good or bad.



Scorecard Characteristics

Relevant **credit and loan characteristics** should be included to be selected as predictors.

Internal data

- » Performance
- » Product
- » Loan / Customer characteristics
- » Internal scores

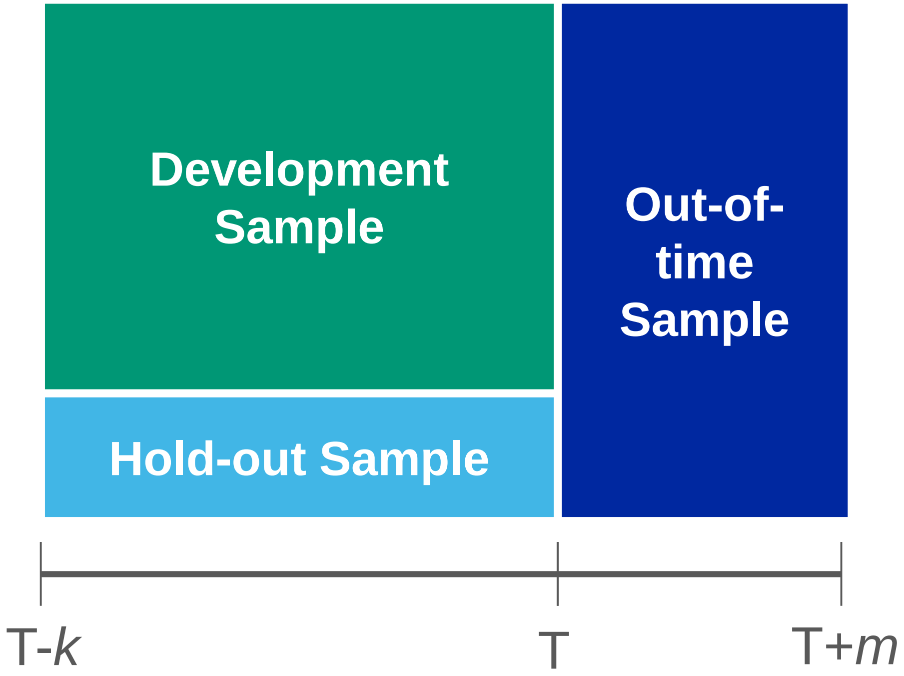
External data

- » Bureau or partner data
- » Other sources

Scorecard Sample Building

Step 1

- Development and holdout samples should cover
  - The longest possible time frame in terms of observation dates
  - The same time-frame
- Development sample should contain more records than the holdout
- Typically: 60/40, 70/30, 80/20 split depending on the sample size. It is important that the holdout contain enough goods/ bads
- Where sample sizes are small, the scorecard can be developed using 100% of the sample and validated using several randomly selected samples of 50% to 80% each
- Out-of-time sample contains observations completely separated in time from the development/holdout sample; typically 6-12 months
- Typically about 2,000 each of goods, bads, and rejects are sufficient for scorecard development.
- Distribution of goods and bads should not be statistically different in development and validation samples (Kolmogorov-Smirnov test can be used for this purpose).



Variable Selection

Step 2

Variables Selection

Methodologies

1 Univariate Analysis

Regress each driver against the target variable. IV and GINI to track which are the most powerful drivers.

2 Correlation Matrices

For categorical variables, it helps to reduce the number of redundant variables.

3 Clustering Algorithms

Based on correlation techniques, they detect groups of variables that share similar behavior. The variable with higher predictive power is eligible to be representative of the whole group.

4 Expert Judgment

Fundamental to prioritize variables that have greater relevance from a business perspective. This can overtake the predictive power based on a standalone statistical basis.

Binning Approaches

Step 3

Experience-Based

Based on experience and industry knowledge

May be subjective and difficult to defend

Statistically-Based

Using statistical techniques to define unique bins

More objective and rigorous

Model Building

Step 4

- The most popular model estimation techniques are:
  - **Logistic regression**
    - Straightforward and sufficiently robust.
    - Algorithms to control for tolerance, stepwise entrance of the drivers
    - Results between 0 and 1.
  - **Decision Tree algorithms**
    - Finds the best split between variables to maximize discrimination
    - Not easy to control in the split/weight they provide
  - **Machine learning algorithms**

Model Performance

Step 5

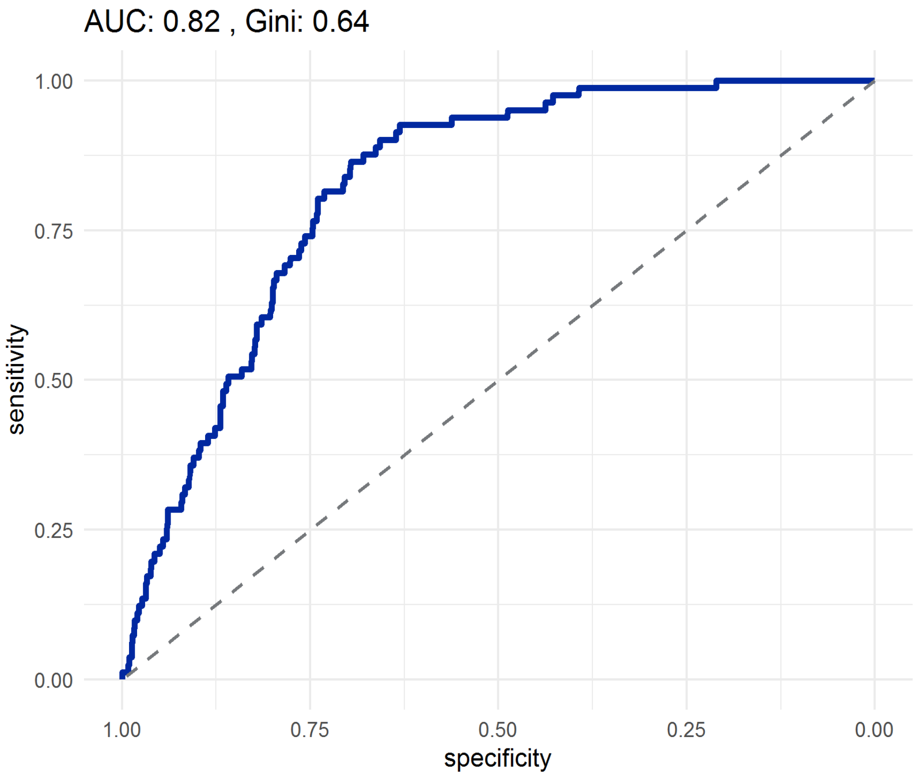
The effectiveness of the scorecard can be verified under the following risk-metrics:

» **Gini Coefficient:** ability to measure the rank-order capability (i.e. distinguishing good/bad accounts).

  Benchmark values: 40-60% (Application scorecard), 70-80% (Behaviour Scorecard)

» **AUROC:** ability to measure accuracy of classification.

  Benchmark values: 50% (Random guess), 100% (perfect correct predictions)



Both these metrics are recommended by the Basel Committee (Working paper n.14- Studies on the Validation of Internal Rating System)

Model Validation

Step 6

The model is built on the development sample. It is therefore important to test on independent samples its effectiveness, using the same risk-metrics describes previously.

Out-of-Sample validation

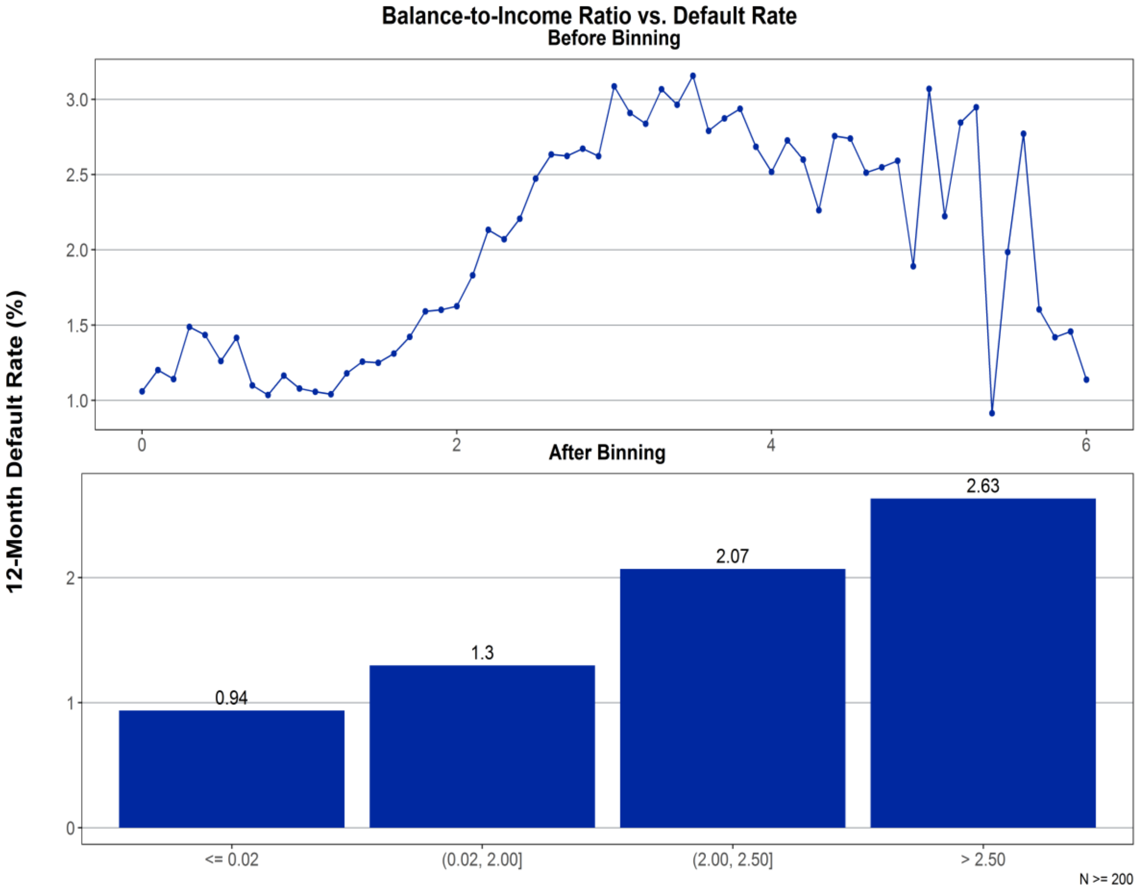
- » Testing the model on the holdout sample. Same time-frame but independent observations.

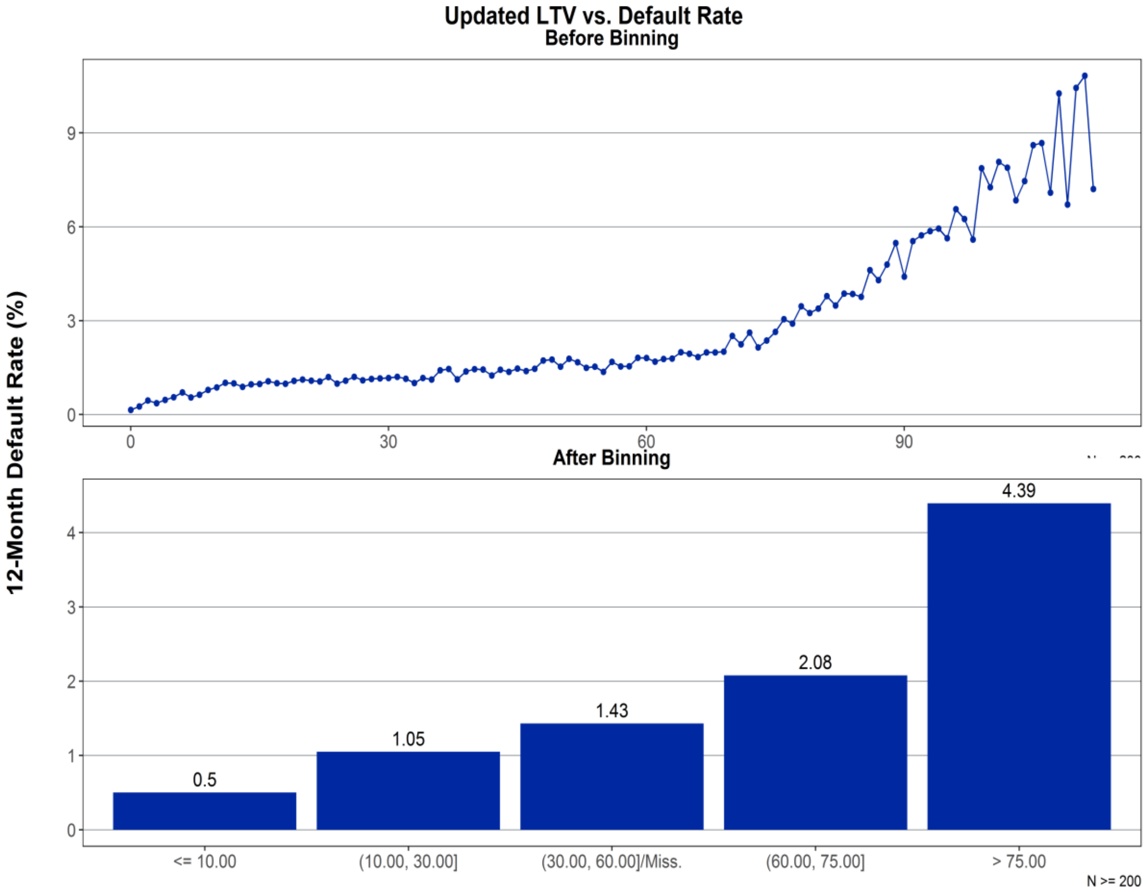
Out-of-Time validation

- » Testing the model on the out-of-time sample. This sample is taken from a different time-frame.

Case Study

Sample Binning: UK Behavioral Scorecard Model

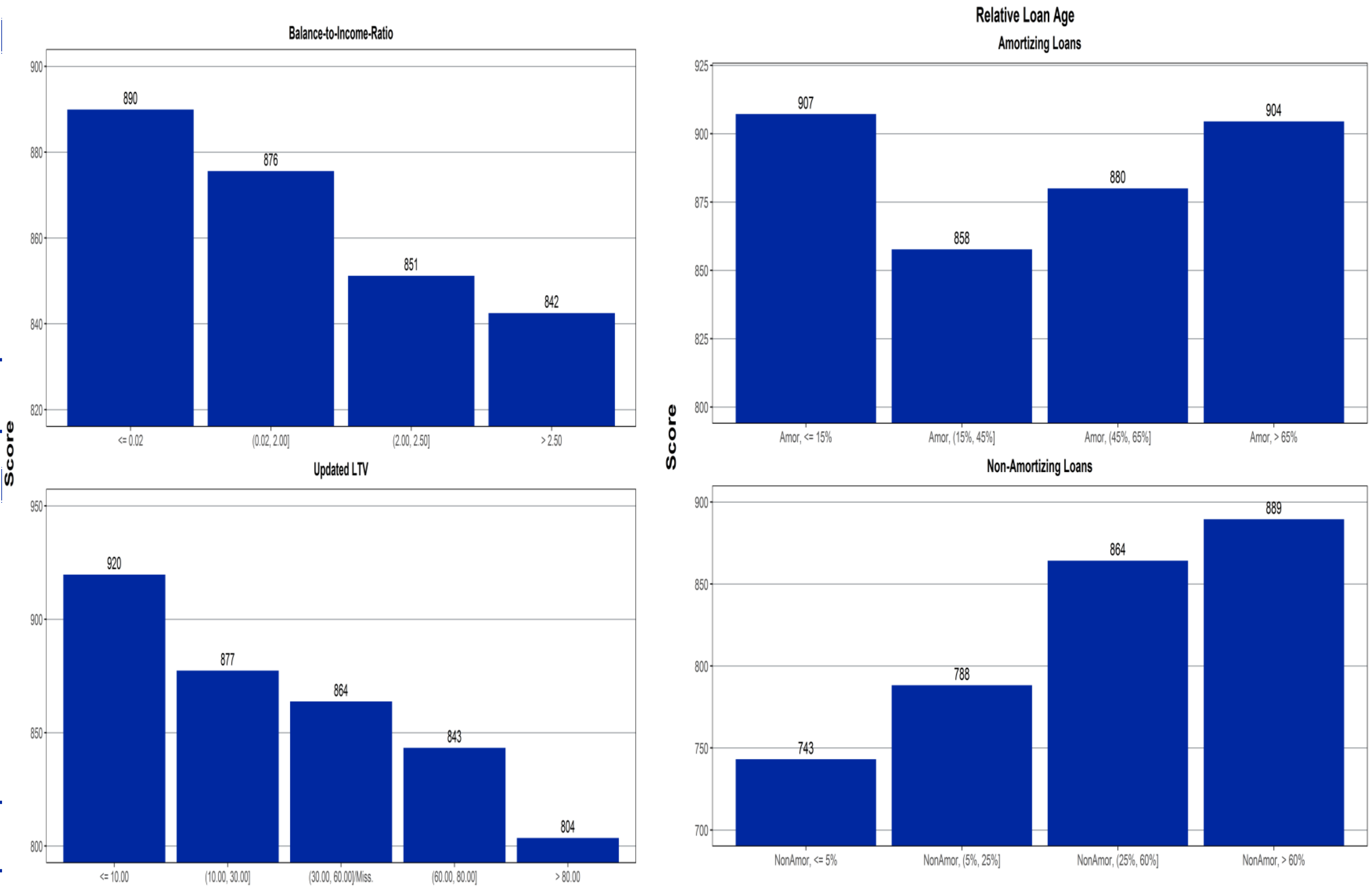




Case Study: UK Behavioral Scorecard Model

ACCOUNT-1 - GOOD ACCOUNT	
Payment Type	Linear Installments
Balance-to-Income	0.2
CCJ	No
Bureau Credit Score	Excellent
Employment Status	Full Time Employed
Relative Loan Age	70%
Months in Arrears	Current
Occupancy Type	Second Home
Updated LTV	8%
PD	**0.15%**
Score	**970.25**

ACCOUNT-2 - BAD ACCOUNT	
Payment Type	Increasing Installments
Balance-to-Income	> 2.5
CCJ	Yes
Bureau Credit Score	Fair
Employment Status	Employed w/ Partial Support
Relative Loan Age	4%
Months in Arrears	Current
Occupancy Type	Owner Occupied
Updated LTV	75%
PD	**51.05%**
Score	**496.96**



Case study: SME Scorecard build based on size

A business with small/medium size and the business with financial data

Very small business with almost no reliable financial or enterprise information

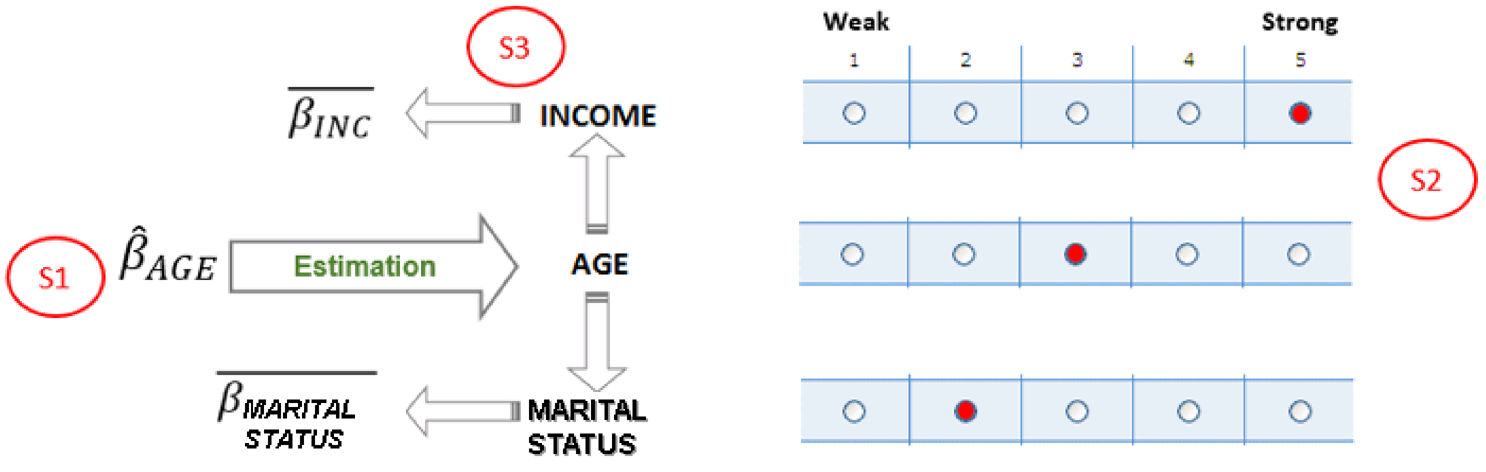
- **Examples**: mid-size family business, multi-partner small business
- **Typical Scorecard Characteristics:**
  - firm-level data, including SME's financial reports, enterprise information such as taxation, shareholder change, foreign investment, legal entity data
  - person-level data, including primary owner and his related parties' personal credit history
- **Modelling Approach**: traditional scorecards building

- **Examples**: one-man business such as a coffee shop, small family workshop
- **Typical Scorecard Characteristics:**
  - quantitative measures of owner's personal characteristics and any available firm-level data
  - qualitative measures such as business prospects, owner's management skills, market recognition
- **Modelling Approach**: hybrid model with SME owner's personal scorecards + qualitative expert overlay or traditional scorecards building if enough data available

Case study: Scorecard Building with limited information

Hybrid Approach

- Quite often, full data are not available and the standard modelling approach might provide rather limited information for a proper scorecard.
- To deal with such cases, the missing information can be supplemented by expert knowledge.
- We overlap information to integrate qualitative data into the empirical model.



$$PD_{hybrid\ model} = \frac{1}{1 + \exp[-(b_0 + \underbrace{b_{11} * x_{11} + \cdots + b_{k1} * x_{k1}}_{\text{WOE of variables from quantitative model}} + \underbrace{b_{12} * x_{12} + \cdots + b_{k2} * x_{k2}}_{\text{WOE of variables with no data available}})]}$$

Step 1. We estimate an empirical model on the limited data available and ask the Bank to provide the ranking and importance of other variables for the modelled risk metrics.

Step 2. Based on the rank ordering of the variables, build-back the coefficient estimates for the variables from the expert questionnaire.

Step 3. Calibrate the model to coefficients to match the default rate in the data.

4 Implications and Applications

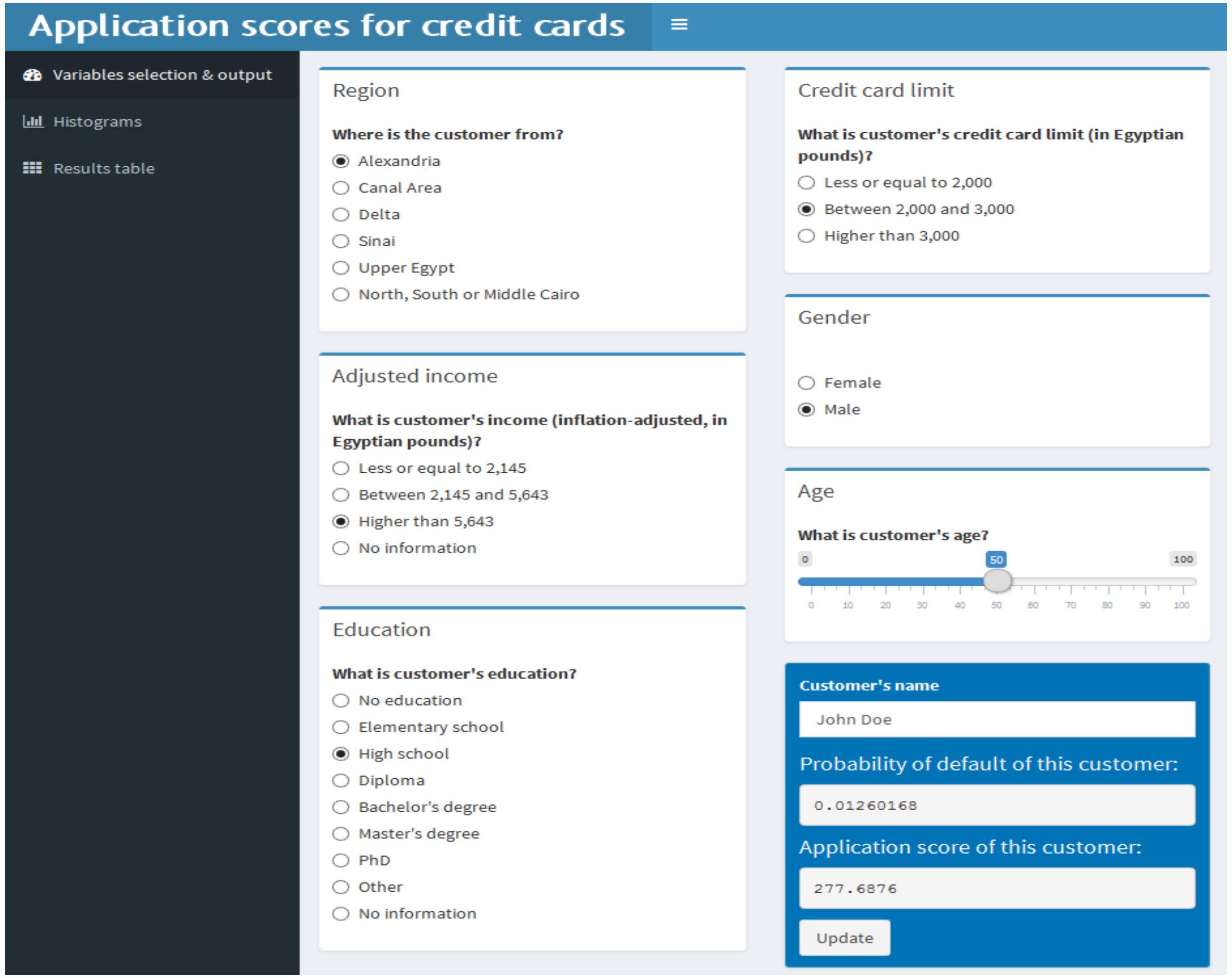
Challenges and nuances

- *Implementation* ← should be driven by model, data and environment
  - **Data:** what is collected currently? What will be collected in future?
  - **Environment:** who will be providing data and running model? How frequent? Does it involve Cloud solution?
  - **Purpose:** what is the model use? How results of the run will be used? Will the model be linked to other models in placed? What is post analysis?
  - **Model:** how complicated the model is?
- *Monitoring* ← should act as regular review of model output and performance
  - Is it accurate? Is it performing well? How stable are the results?

Implementation

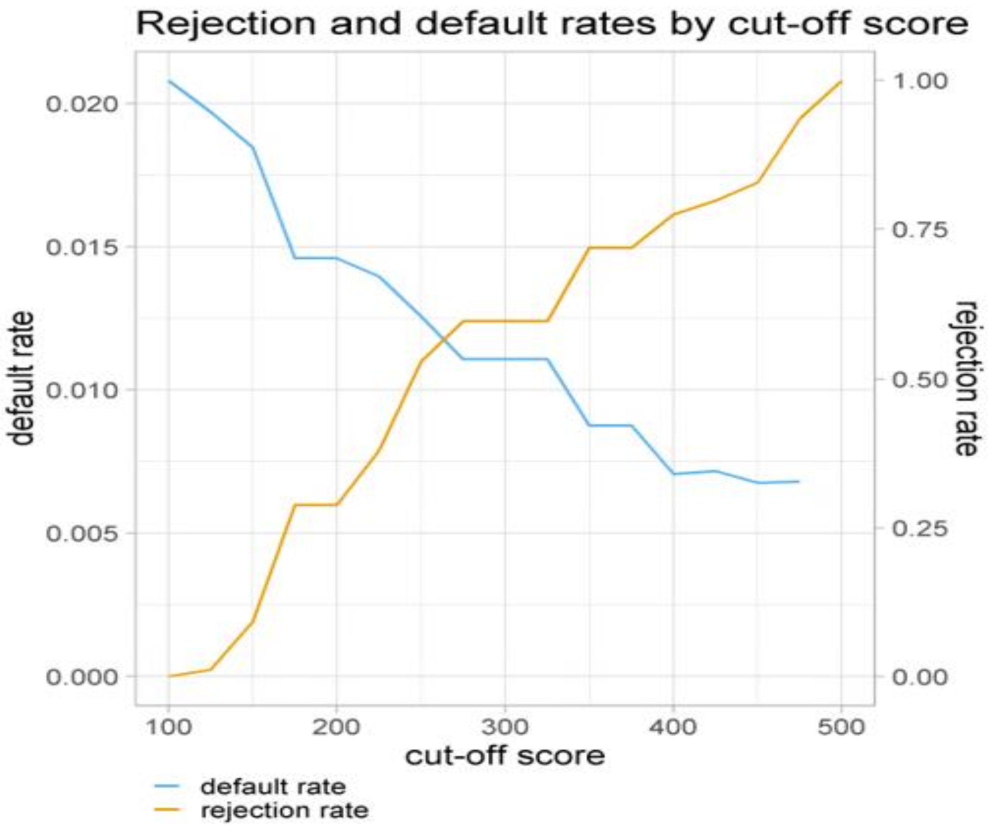
Graphical User Interface

- Example of GUI which designed to be located on local Bank's server
- PD and Score are calculated based on provided inputs
- Other tabs can present joint portfolio analysis or batch run (scoring whole portfolio at once)



Implementation and Use

Model use strategy to support faster decision making and limit setting



Scorecard prediction	Actual performance: Good	Actual performance: Bad/Fraud
Good	1 Correctly passed	2 False negative cases
Bad/ Fraud	3 False positive cases	4 Correctly suspended

Monitoring Model Performance

An established approach to monitoring model performance

Discrimination	Accuracy	Stability
Model ability/power to discriminate between events and non-events, e.g., defaults and performers, and the power to rank-order risk. Applicable to choice models with binary outcome (e.g., PD or scorecard models).	Model ability to deliver accurate best estimate/prediction of output. Applicable to virtually all models with quantifiable output and an observable real-world counterpart.	Comparison of distributional aspects of development sample, on the one hand, with those of any other sample, usually production.
» Gini/ROC – K-S Statistic » Brier Score	» Deviation of Actual from Predicted » HL/Chi-square test	» Population Stability » Characteristic Stability

Moody's ANALYTICS

Better Faster Decisions

Contact Us: Economics & Business Analytics Offices

West Chester, EBA-HQ
+1.610.235.5299
121 North Walnut Street, Suite 500
West Chester PA 19380
USA

New York, Corporate-HQ
+1.212.553.1653
7 World Trade Center, 14th Floor
250 Greenwich Street
New York, NY 10007
USA

London
+44.20.7772.5454
One Canada Square
Canary Wharf
London E14 5FA
United Kingdom

Toronto
+1.416.681.2133
200 Wellington Street West, 15th Floor
Toronto ON M5V 3C7
Canada

Prague
+420.22.422.2929
AFI Butterfly
B1 3rd Floor, Pernerova 691/42
186 00 Prague 8 - Karlin
Czech Republic

Sydney
+61.2.9270.8111
Level 10, 1 O'Connell Street
Sydney, NSW, 2000
Australia

Singapore
+65.6511.4400
6 Shenton Way
#14-08 OUE Downtown 2
Singapore 068809

Shanghai
+86.21.6101.0172
Unit 2306, Citigroup Tower
33 Huayuanshiqiao Road
Pudong New Area, 200120
China

help@economy.com

www.economy.com