Building Credit Scorecards for Small Business Lending in Developing Markets

Dean Caire, CFA
Bannock Consulting

November 2004

This article presents seven steps to building scorecards for small business lending in developing credit markets such as Central and Eastern Europe and Russia. Such markets lack the credit bureaus and rating agencies that advanced market scorecards rely on. Until the third-party information infrastructure develops, a bank must mine its own institutional knowledge and historical portfolio data to develop scorecards that suit its strategies for the small business segment.

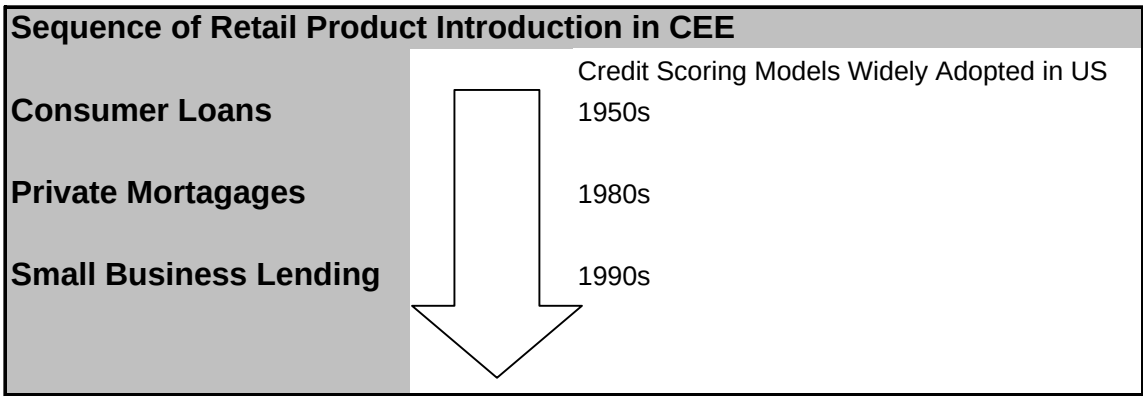
For more information, please contact Greta Bull, Greta_Bull@Bannock.co.uk, phone: 20 7535 0200, Bannock Consulting, 47 Marylebone Lane, London, W1M 6LD

Acknowledgments: thanks to Mark Schreiner of Microfinance Risk Management and Stewart Pirnie of Bannock Consulting for their editing and useful suggestions.

A Short History of Credit Scoring in Central and Eastern Europe

Retail product introduction and credit scoring adoption in the rapidly developing credit markets of Central and Eastern Europe (CEE) has followed the same sequence as in the markets of North America and Europe, as shown in Illustration one below. The key difference is that credit bureaus are either weaker or do not yet exist in rapidly developing markets. Without reliable third-party information sources, developing market banks[1] cannot use advanced market scorecards, which tend to rely heavily on credit bureau information. Instead they must develop scorecards using their market knowledge, experience and internal data.

Illustration 1: Retail Product Introduction and Credit Scoring Adoption in CEE and USA



Banks in developing markets, many of which are owned by Western European banks, have managed to adapt their parent bank's scorecards or create similar scorecards for consumer loans and private mortgages. Consumer loans and private mortgages are both small ticket, homogenous products ideal for simple risk-factor models. Consumer loan limits depend on a borrower's documented salary history, while private mortgage decisions are based on documented salary history, downpayment, and property value.

Small business lending decisions are tougher than consumer and mortgage loan decisions. Loan officers consider a wider range of factors such as financial capacity to repay the loan, willingness to repay the loan, collateral pledged, and the specific terms and conditions of the loan contract. CEE financial statements differ both among countries and with Western Europe and North America, and other market specific risk factors make it difficult to adapt scorecards that weren't developed locally. At the same time, bankers do not want analysts to spend hours spreading a small company's financial statements to underwrite a $20,000 loan. The most appropriate way to underwrite a large book of small-business loans is with a simple scorecard that evaluates a mix of financial and non-financial factors and is customised to specific local conditions of the country and lender.

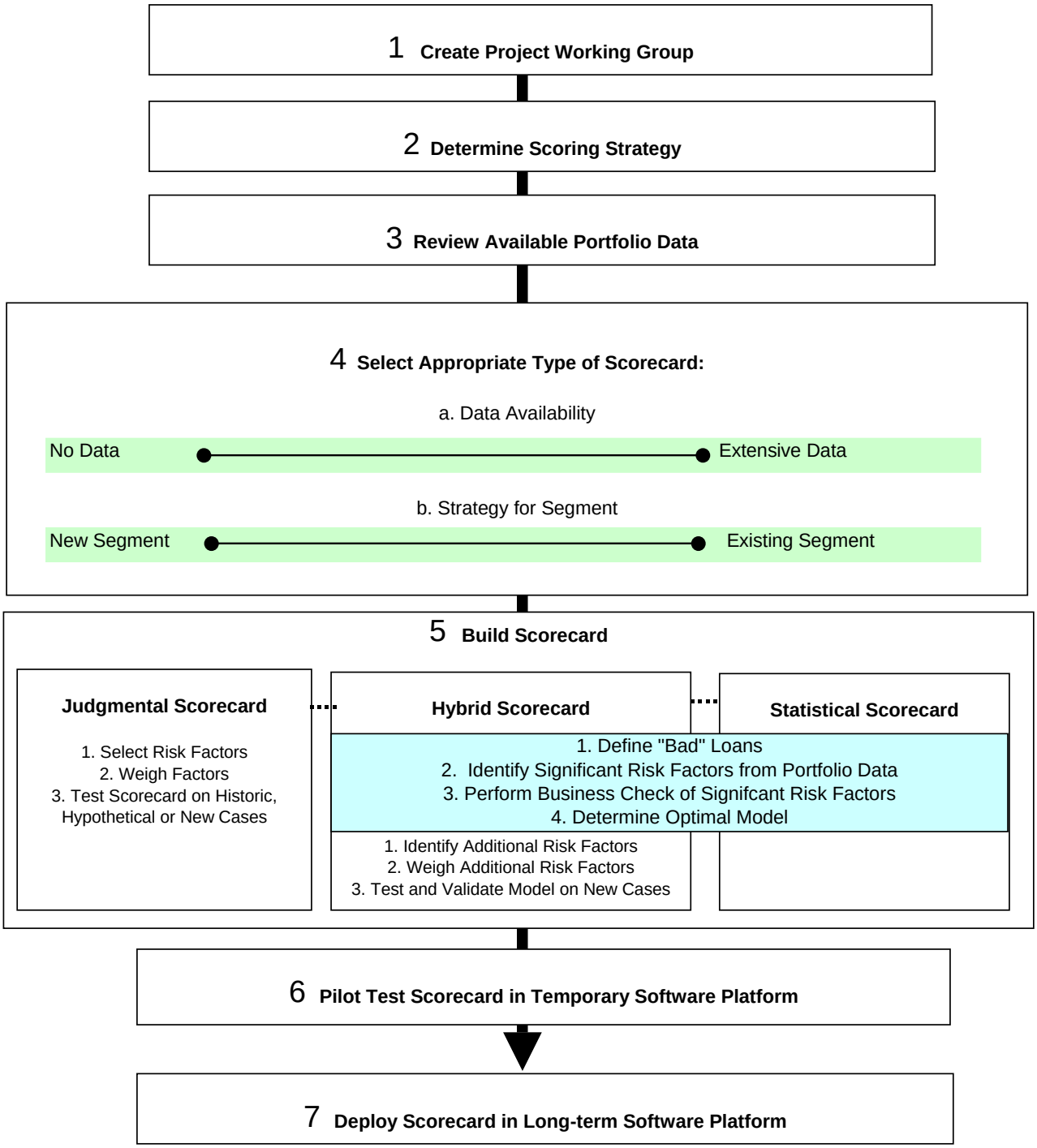
How can a bank in a developing market use its own data, experience, and small business strategy to design custom credit-risk scorecards? The answer is provided in the sections below.

[1] The terms "banks" is used throughout the article for simplicity, but most of the concepts discussed also apply to other finance companies, such as leasing companies.

The Big Picture on Building Scorecards

The diagram below illustrates a seven-step process to building scorecards for the small business segment in developing markets. The balance of this article goes through each of the steps, with a particular focus on step four, how to build scorecards.

SCORECARD BUILDING PROCESS DIAGRAM



1.Create a Project Working Group

First, form a working group. It should include representatives from credit, risk management, marketing, and information technology (IT). This group will plan the scoring project, guide and monitor its progress, and procure the necessary resources to keep things on schedule. It is essential that a sufficiently senior banker "champion" the scoring project to overcome any institutional obstacles or simple resistance to change.

2. Determine Scoring Strategy

Second, the working group should agree on a scoring strategy.

What is a scoring strategy? It is a statement defining how and for what purpose the scorecard will be used. As a hypothetical example, Ex-Am Bank has occasionally lent to small companies in the past using its corporate credit procedures, but now it wants to simplify and modify procedures to issue a large number of two standard credit products: credit lines and term loans of up to $50,000. Ex-Am wants to target businesses with a track record of at least one year, as opposed to the two years it requires for its corporate customers, and it wants to pay attention not only to the business's banking history, but also to the business owner's personal banking history and the owner's years of experience in business. The bank plans to launch a promotional campaign, "fast loans for growing businesses", promising a loan decision in no more than one day. A scoring strategy for this bank could be articulated as: "to develop a scorecard to evaluate applications for small business credit lines and term loans of up to $50,000 and to provide applicants with a loan decision in no more than 24 hours."

3. Review Available Portfolio Data

Third, the working group needs to understand the quality and quantity of information available about past borrowers.

Following our example, Ex-Am has made some loans to small companies in the past and should have some application, financial, and payment data on its small loan portfolio. The working group needs to determine how much data is available and in what format. The bank should generate a portfolio report of all loans made for less than $50,000 and take a "data inventory" of what information is available in the banking computer system, the credit database, and, last of all, in the hard copy, or paper, credit files. Ex-Am bank has 1,450 loans for less than $50,000 original disbursement value issued to 1,316 legal persons. For each client it has electronic data on contact, account and payment information from the banking system. Financial information is stored in hard copy in the credit files, and for clients with outstanding loans, the responsible credit officer updates spreadsheets with periodic financial information. Credit memoranda, legal documents, and any other information are kept only in hard copy in the credit files.

4. Select Appropriate Type of Scorecard

Fourth, the bank must decide what kind of scorecard it will build. There are three main types of credit scorecards that can be developed using only a financial institution's internal data:

1. **Judgmental:** sometimes referred to as expert systems, judgmental scorecards structure credit policies and management risk preferences into a mathematical model that ranks applicants according to risk. A judgmental model can be created without any historical data, so it can be applied to new segments.
2. **Statistical:** statistical scorecards are derived from data on thousands of past applicants in the target sector. Statistical techniques vary, but some of the most popular techniques are decision trees, artificial neural networks, and logistic regression.[2] A statistical model can be developed only for products for which the financial institution has collected a substantial amount of historical data on both its good and problematic clients.
3. **Hybrid:** hybrid scorecards are statistically derived models augmented with judgmentally weighted variables. A hybrid scorecard requires extensive historical data, but provides flexibility to incorporate new risk factors related to a new product or segment.

As shown in illustration two above, the choice of appropriate scorecard is driven by the quality and quantity of data available and the strategy for the segment. If there is little or no historical data, such as when the bank is entering a completely new segment, the only option is a judgmental scorecard. When there is ample historical data, a statistical scorecard is preferable because it can quantify the probability of a "negative credit event". If a bank finds itself with not quite enough data for a statistical scorecard or if it has developed a reasonably predictive scorecard but would like to incorporate additional factors related to a new target client or sector, then a hybrid scorecard may be appropriate.

To return to the Ex-Am Bank example, it wanted to target growing small businesses and focus on some factors not emphasised in its standard credit analysis for corporate clients. The bank had relatively little data in electronic format for its loans of less than $50,000. A considerable amount of other information was scattered around the organisation in various spreadsheet, word processor, and hard-copy credit files. This information could be collected and keyed into a database, but there would likely be trouble with data consistency and missing values, as well as the cost of gathering and keying in the data. Of the 1,450 total loans it had issued, only 560 had been repaid and of those only 25 had experienced arrears of more than 60 days. Another 45 of the currently outstanding loans had had repayment problems. What is the correct choice of scorecard for Ex-Am Bank?

Ex-Am bank is typical of many banks in developing markets – it has experience lending to small businesses, knows its market, and recognizes the need to automate small business lending for it to be profitable. At the same time, it lacks the data required for statistical credit scoring. The appropriate choice appears to be a judgmental scorecard.

[2] A detailed explanation of each technique can be found in Brendan, J. "Applying Data Mining Techniques to Credit Scoring."

5. Build the Scorecard

Fifth, someone needs to sit down, take a deep breath, and get ready to build the scorecard. The steps in scorecard building and the time required for each step differ according to the type of model selected. The sections below describe the main steps in building judgmental, statistical and hybrid scorecards.

Judgmental Scorecards

Judgmental scorecards quantify an organisation's credit policy, market knowledge, risk preferences and segment strategy. A very simple judgmental scorecard might be a checklist of minimum criteria a potential borrower must meet, while a more sophisticated card would combine and weigh all key underwriting criteria such that scorecard decisions generally agree with those made by credit officers.

Because judgmental scorecards rely more on human expertise and organisational knowledge than on statistical relationships, building them may require more time and input from senior management. While some factors should be common to almost all judgmental scorecards, every bank should have a different card – a custom judgmental scorecard is not a "generic" scorecard of the type built from pooled bank data in advanced markets.

There are three main steps to building a judgmental scorecard:

1. select risk factors
2. weigh risk factors
3. test model on historic, hypothetical or new cases

1. Select Risk Factors

The working group should select the risk factors to include in the scorecard. To get started, a bank may want to detail its minimum lending criteria. The minimum lending criteria for Ex-Am Bank are shown in the table below.

Criteria	Acceptable Value
Loan To Collateral Value	< 70%
Annual Turnover	3 times Loan Value
Borrower Years in Business	at least 1 Year
Current Ratio	> 0.5
Total Assets	> €100,000

Surely Ex-Am considers many more risk factors and it would want to list them as well. Risk factors should include items from the financial statements as well as non-financial statements information related to borrowers' banking history, demographics, etc. To keep this example manageable, we will assume that the above table contains all risk factors Ex-Am considers in its small loan underwriting decisions.

2. Weigh Risk Factors

The simplest form of scorecard weighting is equal weighting. If we were to weigh Ex-Am's 5 risk factors so that values within the "Acceptable Value" range receive one point and values outside of it receive zero points, we would have the very rudimentary scorecard pictured below in illustration 3.

This equal weighted scorecard is sometimes referred to as a "gateway" underwriting system. An applicant must pass through a series of "gateways" or checkpoints in order to qualify for financing, with the assumption being that only applicants clearing all 5 of the gateways and scoring 5 points pass the model. The advantage of such a simple model is its simplicity, but it obviously fails to capture the risk trade-offs that underwriters are usually willing to make. Some factors or some combinations of factors are more important than others.

Illustration 3: Equal Weighted Scorecard

JUDGMENTAL SCORECARD 1			
Variable			
1	Loan to Collateral Value		
	>70%	<70%	
	0	1	
2	Annual Turnover to Loan Value		
	<3x	>3x	
	0	1	
3	Years in Business		
	<1	>1	
	0	1	
4	Current Ratio		
	< 0.5	> 0.5	
	0	1	
5	Total Assets (EUR)		
	<100K	> 100K	
	0	1	

To capture the trade offs between and among factors, it is common to give different weights to each factor, either by giving more points to some factors than to others or by expanding the number of categories for each factor and giving positive or negative weights to values that are considerably above or below the minimum acceptable value. Judgmental Scorecard Two, pictured below, takes the first approach and Judgmental Scorecard 3 takes the second approach.

Illustration 4: Two Variable Weighted Judgmental Scorecards

JUDGMENTAL SCORECARD 2			
Variable			
1	Loan to Collateral Value		
	>70%	<70%	
	0	3	
2	Annual Turnover to Loan Value		
	<3x	>3x	
	0	2	
3	Years in Business		
	<1	>1	
	0	1	
4	Current Ratio		
	< 0.5	> 0.5	
	0	1	
5	Total Assets (EUR)		
	<100K	> 100K	
	0	1	

JUDGMENTAL SCORECARD 3			
Variable			
1	Loan to Collateral Value		
	>70%	50-70%	<50%
	0	1	2
2	Annual Turnover to Loan Value		
	<3x	3-5x	>5x
	0	1	2
3	Years in Business		
	<1	2-4	>4
	-2	1	2
4	Current Ratio		
	< 0.5	0.5 - 1	> 1
	0	1	2
5	Total Assets (EUR)		
	<100K	100-500K	> 500K
	0	1	2

For Judgmental Scorecards 2 and 3, the passing score might still be 5 points, but now an applicant who would have failed Judgmental Scorecard 1 for having a current ratio of less than 0.5 would still pass cards 2 and 3 by virtue of having Loan to Collateral Value of less than 50%.

3.Test the Scorecard on Historical, Hypothetical or New Cases
After settling on a set of variables and weights that best represent its underwriting preferences, Ex-Am Bank would need to test the scorecard to see whether, on average, it reaches the same decisions credit officers would make.

Testing can be performed using data from historical cases (sometimes called “back-testing”), hypothetical cases, or on new applicants in a pilot testing phase. In many cases, some combination of the three approaches is used. For example, Ex-Am bank has some historical data. Even if all of the necessary data is not available electronically, it should be possible to manually collect a reasonable sample of 50-100 test cases from hard copy files. In a historical test, a well-weighted card should produce a full range of possible passing scores since all the test cases were, in fact, approved by the bank. If Ex-Am’s existing borrowers cannot pass the new applicant scorecard, the weighting must be adjusted. At

the same time, it is possible to test a card's weighting by making up common client profiles and seeing if imagined "good" clients pass and if "undesirable" clients fail. Finally, a card can be tested on data captured from new applicants in a pilot-testing phase, although this last approach requires additional time.

Statistical Scorecards

Creating statistical scorecards requires extensive data on past borrowers in the same segment that will be scored. In the case of Ex-Am and in the case of many banks in developing markets, there is insufficient data, particularly on problematic, or "bad", clients. But in many finance companies in CEE, such as leasing companies, there is sufficient data on thousands of deals including at least 300-1,000 "bads", a number often cited as the critical mass necessary for statistical modeling.

1. Define "Bad" Loans

A statistical model outputs the probability that a loan application will go "bad", but there is no universal definition for a "bad" loan. "Bad" should be a loan that incurs costs above and beyond any expected profit. While it may be convenient for Basel 2 purposes to use the common default definition of greater than 90 days in arrears, for application scoring each finance company should come up with its own definition based on an understanding of its costs and its experience dealing with problem clients. The definitions should encompass all cases that lose money for the finance company, and as such may be complex: for example, "bad" loans are loans with more than 90 days in arrears or more than 2 spells of arrears over 60 days or an average arrears of 50 days or more.

2. Identify Significant Risk Factors from Portfolio Data

Once a bank has defined what a "bad" loan is, it must begin the process of gathering, cleaning, and, finally, analysing the data. The software and methods for statistical data analysis are many and are beyond the scope of this article, but the result of exploratory statistical analysis should be an equation built from a group of factors persistently related to "bad" loans in the data set. A list of the factors, their relative strength and the direction of their relationships to "bad" loans should be presented to the working group for what is called a "business check".

3. Perform Business Check of Significant Risk Factors

"Business Check" is another way of saying "make sure the risk factors uncovered during the analysis in step two make sense in the real world." The data modeller should definitely know the statistics needed to uncover the most meaningful variables, but practitioners are the best people to spot relationships that don't meet with their everyday experience or their business sense. Disagreements between what the data says and what practitioners say might result from misleading data labels, an error in cleaning or modelling the data, or might not be errors at all, but practical issues that need to be taken into consideration. For example, a highly predictive variable may have to be dropped from a model because it is not possible to collect at the application stage. The result of the business check should be agreement in the working group that the factors in the models make reasonable sense and are feasible to collect from applicants.

4. Determine Optimal Model
As noted above, there are a number of statistical modelling techniques used in the credit industry. If a bank has time and resources, it is a good idea to model the data using several different techniques and to choose the optimal model based on predictive power and ease of model deployment. Predictive power is a measure of the model's ability to separate good cases from bad cases, and ease of deployment is based on the complexity of programming a user-friendly scorecard tool into the end user software program and maintaining it thereafter.

Hybrid Scorecard
A hybrid scorecard is a statistical scorecard modified to include variables that were either not available or not collected historically. New variables must be added and weighed in much the same way as with judgmental scorecards, making the resulting scorecard a combination, or hybrid, of the other two techniques.

Hybrid scorecards take the optimal statistical scorecard and lead it further through the three steps of judgmental scorecard development:

1. Identify additional risk factors
2. Weigh the additional risk factors
3. Test and validate the model on new cases

The main dilemma in creating hybrid scorecards is how to manually assign weights to one or two factors without distorting the significant relationships derived from the statistical analysis. While the author has seen no literature covering hybrid scorecards, one practical approach to weighing the additional judgmental variables is to consider the relative weights of the statistically derived model factors and, using judgment and common sense, decide how much more or less important the new variables are to the credit decision and weigh them accordingly. The relative weights for a statistical card can be determined by looking at the total number of points possible for the card and at the maximum possible number of points for each factor. For example, if a card's total number of possible points is 500 and the factor collateral coverage can score a maximum of 100 points, then collateral coverage accounts for as much as 20% (100/500) of the total scorecard decision.

There is no way to know whether judgmental weightings are "correct" except to test them, and since there was no historical data on the additional variables, this means tests must be carried out on new applicants on a trial basis. Rather than waiting years to validate hybrid models with data from a mature portfolio, it is possible to take of sample of loans scored with a hybrid card and to compare the card's decisions with loan officers' subjective judgment. The judgmental weights can be adjusted and this exercise repeated until such time as the bank reaches a sufficient comfort level with the scorecard's decisions.

The hybrid scorecard's mixing of methods may seem pure madness to scoring purists, yet it is no more than the application of scoring best practices to situations where historical data is meaningful, but incomplete for the purposes of reaching a new segment or tailoring the underwriting criteria for a new segment. Such situations are common in developing markets.

6. Pilot Test the Scorecard in a Temporary Software Platform

Sixth, every scorecard needs to be tested before it is widely used. Some common strategies are to test the scorecard in a limited number of branches or to first run the scorecard in parallel with standard procedures without relying on it as a decision tool. A simple, temporary software solution, such as a dedicated program with data collection functionality in Excel, works for pilot testing. One or several branches can begin implementing the scorecard and provide feedback on:

1. if the scorecard recommendations match loan officers' instinct or gut feeling
2. if procedures for scorecard use require any modification

The bank should collect all pilot test data and periodically review it to ensure that users are correctly entering the data and the scorecard program accurately calculates scores and otherwise works properly.

Once a scorecard has been "live" tested and the bank has made any necessary adjustments to the scorecard or its procedures for use, the bank should develop a long-term software solution that is integrated with the bank's existing systems.

7. Deploy the Scorecard in Long-term Software Platform

Seventh, the long-term success of a scoring project depends on a well designed, easy-to-use scoring tool that streamlines credit processes. A scoring model should fit seamlessly into the bank's application processing system. The preferred software platform for application processing is an SQL or comparable database program with user-friendly interfaces. A scorecard or multiple scorecards can be integrated into the application process flow, with loan documentation, reporting, data collection and data warehousing also handled by the database program. Many vendors sell off-the-shelf application processing software that includes a module for deploying scorecards, or a bank can develop its own database software solution – in either case, banks should plan software development well in advance, as software development and/or implementation projects tend to run over schedule.

Conclusion

For banks in developing markets just starting down the scoring road, a judgmental scorecard running in the appropriate database software is only the beginning of a journey that will lead to ever more powerful risk models that open an array of possibilities for optimising risk management and maximising retail segment profits.