Paper 3554-2019

Developing a Credit Risk Model Using SAS®

Amos Taiwo Odeleye, TD Bank

ABSTRACT

A credit risk score is an analytical method of modeling the credit riskiness of individual borrowers (prospects and customers). While there are several generic, one-size-might-fit-all risk scores developed by vendors, there are numerous factors increasingly driving the development of in-house risk scores. This presentation introduces the audience to how to develop an in-house risk score using SAS®, reject inference methodology, and machine learning and data science methods.

INTRODUCTION

According to Frank H. Knight (1921), University of Chicago professor, risk can be thought of as any variability that can be quantified. Generally speaking, there are 4 steps in risk management, namely:

- **A**ssess: Identify risk
- **Q**uantify: Measure and estimate risk
- **M**anage: Avoid, transfer, mitigate, or eliminate
- **M**onitor: Evaluate the process and make necessary adjustment

In the financial environment, risk can be classified in several ways. In 2017, I coined what I termed the **C**-**L**-**O**-**M**-**O** risk acronym:

- **C**redit Risk, e.g., default, recovery, collections, fraud
- **L**iquidity Risk, e.g., funding, refinancing, trading
- **O**perational Risk, e.g., bad actor, system or process interruption, natural disaster
- **M**arket Risk, e.g., interest rate, foreign exchange, equity, commodity price
- **O**ther Risk not already classified above e.g., Model risk, Reputational risk, Legal risk

In the pursuit of international financial stability and cooperation, there are financial regulators at the global and national levels.

Examples of international regulators:

- The Bank of International Settlements (BIS or Basel Accords)
- The Financial Stability Board (FSB)

Examples of national regulators (such as in the United States):

- The Federal Reserve Bank (the United States Central Bank)
- The United States Treasury's Office of the Comptroller of the Currency (OCC)

Credit Risk Management, in a consumer lending environment, encapsulates the **Prospect-to-Customer Lifecycle** stages with specific risk score used at each stage:

- Marketing of financial products to prospects (Response risk score)
- Origination or underwriting of new accounts as customers (Acquisition risk score)

- Existing Customer Management (Behavior risk score)
- Collections and Recovery (Collections or Recovery risk score)
- Fraud Management (Fraud risk score)

In 1956, Fair & Isaac Company (FICO) pioneered the development of a risk score, **FICO Score**, to manage credit risk. FICO Score, a generic score, was designed to rank-order prospect's or customer's risk based on the information in their credit file at the credit bureau (CB) or credit reporting agencies (CRA) such as Experian, Equifax, or TransUnion (in the U.S.). While there are several generic, one-size-might-fit-all risk scores developed by vendors, there are numerous factors increasingly driving the development of in-house risk scores.

Predictive modeling, machine learning, and data science methods are at the core of credit risk management and are used throughout the credit risk model development process. These include but not limited to logistic regression, decision tree, neural network, discriminant analysis, support vector machine, factor analysis, principal component analysis, clustering analysis and bootstrapping.

There are many analytical software that can be used for credit risk modeling, risk analytics and reporting so why SAS®? SAS® provides **S**tandardization, **A**ssurance of quality and **S**calability.

CREDIT RISK SCORE

Credit Risk Score ("Scorecard" or simply "Risk Score") is a predictive modeling approach used to evaluate the level of credit riskiness associated with prospects or customers. It does not specifically identify "good" (positive behavior) or "bad" (negative behavior) individuals.

Credit Risk Score is a risk rank-ordering estimator that provides a statistical measure (odds or probability) that an individual with given attributes will be "good" or "bad." This statistical measure, usually transformed or "scaled" into a score along with other business and strategy considerations are used as basis for in making credit and financial decisions. In developing a risk score, there are two major considerations to discuss and establish:

- **Model developer:** Who will develop your risk score model (internal modelers or outsourced to external agency)?
- **Model development data:** Which data will be used for model development (internal, external, or a combination of both)?

Credit Risk Score: How is it developed?

- **Generic Score:** Developed by external modelers using only credit bureau data. Examples are FICO score and Vantage score.
- **Vendor Score:** Developed by external modelers (FICO, CRA, FinTech, etc.) specifically for a bank or financial institution using credit bureau and/or in-house data.
- **In-house Score:** Developed by the bank's or a financial institution's modelers using credit bureau and/or in-house data.

Credit risk Score: Why develop it in-house?

While there are several generic, one-size-might-fit-all risk scores developed by vendors, there are numerous factors increasingly driving the development of in-house risk scores such as reduced costs, increased regulations, accessibility to sizeable and reliable data, availability of educational material or training, better analytical software, to mention few.

Common Machine Learning or Data Science approaches used for Credit Score development:

- Logistic Regression
- Decision Tree
- Neural Network
- Discriminant Analysis
- Support Vector Machine

LOGISTIC REGRESSION MODEL

Linear regression model (y= a + b*x) predicts a continuous dependent or target variable (*y*) using the information contained in the independent variable or predictor (*x's*). There are few statistical assumptions that must be met, including normal distribution assumption. McCullagh & Nelder (1989) coined the term **Generalized Linear Model (GLM)** to incorporate outcome variable are not normally distributed using what they called the "LINK" function. An example of such outcomes is the *Statistical Logistic Regression Model*.

Logistic regression model (or Logit) is a commonly used technique in developing scorecards, where the target variable is categorical. It's known as the gold standard or preferred method, due to the good interpretability of attributes coupled with business implications; mostly applicable to Acquisition or Behavior risk score.

Logistic regression model, like most other machine learning or data science methods, uses a set of independent variables to predict the likelihood of the target variable. **Logit transformation** (that's, the log of the odds) is used to linearize probability and limiting the outcome of estimated probabilities in the model to between 0 and 1. **Maximum Likelihood Estimation (MLE)** algorithm is used to estimate all the regression parameters.

Logistic regression modeling process can be exhaustively executed to find the "best" model using all combinations of available independent variables. However, this can be highly computationally intensive, especially if there are several independent variables available. On the other hand, there are algorithms that can be used in order to find a parsimonious "best" model; this includes techniques such as:

- Forward Selection
- Backward Elimination
- Stepwise algorithm
- Chuckwise algorithm

REJECT INFERENCE

During the origination stage of the prospect-customer lifecycle, an acquisition risk score is used in approve-decline decision for all prospects (also known as *Through-The-Door*, **TTD** population).

For any prospect that were previously approved as customers (that's, the **ACCEPTS** population), the credit history and behaviors will be available and known to the credit issuing financial institution; this population is also known as **K**nown **G**ood or **B**ad (**KGB** population). On the other hand, every prospect that was declined (that's, the **REJECTS** population), their credit history and behaviors will be unknown to the credit issuing financial institution.

Any modeling process using just the KGB population will result in a **selection bias** or **missing data** phenomenon. In the light of this, to incorporate the REJECTS population in the model development data, the credit history and behaviors of the REJECTS must be inferred (see Figure 1).

To mitigate this selection bias problem, below are some of the Reject Inference methodologies that have been proposed:

- Parceling
- Nearest Neighbor
- Bureau Performance
- Fuzzy Augmentation
- Simple Augmentation
- Bureau Score Migration
- Approve all applications
- Iterative Reclassification
- Memory-Based Reasoning
- Assign All Rejects to "Bads"
- Manual Adjustment of Weight Of Evidence
- Assign Rejects in the same % to reflect the Accepts

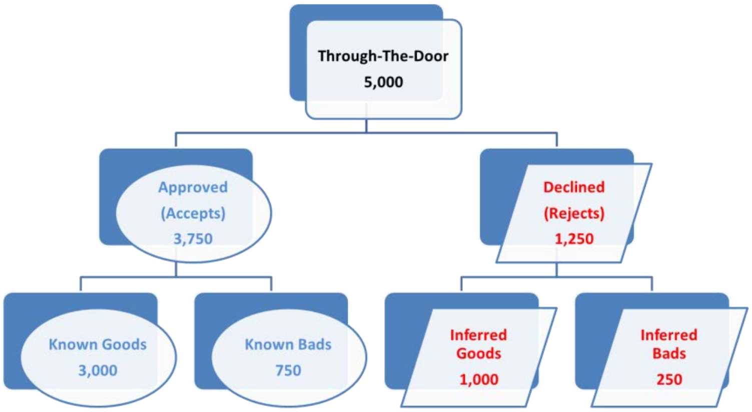


Figure 1. Reject Inference

FUZZY AUGMENTATION

Fuzzy Augmentation algorithm is a two-step Reject Inference approach, which incorporates not only the probability of a REJECTS being "bad," but also the possibility of being "good" and approved in the first place.

Fuzzy Augmentation begins by developing a model using ACCEPTS population data; this is the KGB model. Then, using this model to score the REJECTS population such that every REJECTS will have two probabilities (probability of bad and probability of good) as weights leading to two observations for each REJECTS.

Fuzzy Augmentation Step 1 – Classification:

- Build a model using ACCEPTS or KGB data. This is the KGB model that will predicts p(bad) of the ACCEPTS population
- Use the KGB model to score and infer outcome of the REJECTS to predict the p(bad) of REJECTS
- Compute p(good) of REJECTS = 1 – p(bad) of REJECTS, such that we will have two probabilities for each REJECT, p(bad) and p(good)
- Weigh REJECTS "good" with p(good) and REJECTS "bad" with p(bad)

Fuzzy Augmentation Step 2 – Augmentation:

- Combine inferred REJECTS with ACCEPTS to make up the **K**nown **I**nferred **G**oods and **B**ads (**KIGB**) data

PROBABILITY OF DEFAULT MODEL: ACQUISITION CREDIT RISK SCORE

There are 11 steps in credit risk model development process:

- Event definition
- Data collection
- KGB Data Partition
- KGB Variable Analysis
- KGB Scorecard Modeling
- Reject inference, KIGB
- KIGB Data Partition
- KIGB Variable Analysis
- KIGB Scorecard Modeling
- KIGB Scorecard Scaling
- KIGB Scorecard Evaluation

1. EVENT DEFINITION

When developing a credit risk score, the definition of the event, i.e., default ("bad"), must be clearly established. There are few things to take into considerations when we are thinking of event definition related to default: approach, component, and assumption.

Approaches to Default Definition:

I. Regulatory (e.g., Basel Accords)
II. Risk analytics (e.g., Portfolio Maturity or Strategy Analysis)

Components of Default definition:

I. Default Event: 90 Days Past Due (DPD), Bankruptcy, or Charge off

II. Default Horizon: 12-month, 18-month, or 24-month

Other analyses that can be conducted include Roll Rate Analysis, Current vs. Worst Delinquent Analysis, Ever- vs. Current- "bads" analysis, Vintage Analysis, or Delinquent Maturity Curve.

Assumption of Default Definition:

We assume that "past performance predicts future outcomes." Based on this assumption, first, sample data of approved applicants or accounts will be selected for a specific timeframe. Then, monitor their performances for another specific length of time to determine if they were good (account is current) or "bad" (90 DPD, bankruptcy, or charge off)

Example of a Default definition:

Using a sample of approved accounts from 2016'Q1 that were 90 DPD in 24 month on book (MOB) or performance window (2018'Q1) then we can define default event and horizon as follows:

- Default Event: 90 DPD
- Default Horizon: 24-month

2. DATA COLECTION AND PREPARATION

Data collected for credit risk score development should be reliable and representative of the future prospects or customers on which the risk model will be executed. While the quantity of modeling data varies, at minimum, it should fulfill the requirement of statistical randomness and significance in order for appropriate inference to be made. The independent variables along with the dependent variable make up the **Development Data or Sample**.

The S-P-M Modeling Windows:

Figure 2 shows the SPM modeling windows: **S**ample, **P**erformance and **M**easurement window. It illustrates accounts that were approved and rejected at a particular time in the past (e.g., 2016'Q1).

For the approved or "booked' accounts, at some point in the future (e.g., 2018'Q1) using a 24-month window, we can determine if these accounts had been "good" (current) or "bad" (90 DPD). For the rejected prospects, their performance will be inferred.

Sample Window: *This is the timeframe from which model development data were selected*

Performance Window: *This is the timeframe of which performance of the approved accounts selected in the sample window is monitored. This is also known as the Default Horizon*

Measurement Window: *This is the timeframe at which the performance of approved accounts selected in the sample window and monitored in the performance window will be assigned or classified as good ("0") or bad ("1") as the target variable*

To determine each SPM modeling window, there are mechanisms to be taken into consideration such as vintage analysis, delinquent maturity curve, sample size, seasonality, promo, merge and acquisition, or macroeconomic situation.

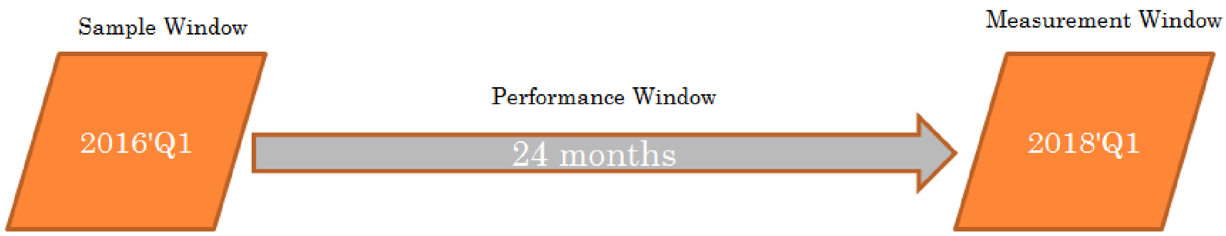


Figure 2. SPM Modeling Windows

3. KGB DATA PARTITION

Model reliability involves the ability of the model to be applicable to future samples and that the conclusions inferred can be generalized. That's, if the final selected model predicts well for subsequent samples from the same or similar population, we can say that the model is reliable. Model reliability is conducted using the **Split-Sample Analysis** method (In-sample validation, **ISV** and the Out-of-Time validation, **OOTV**) in two hierarchical steps:

First, the model development data or sample will be divided into two sub-samples:

- **In-Time sample:** Used for model development and ISV
- **Out-of-Time sample:** Used for OOTV

Second, the In-Time sample will be further divided by randomly assigning each account to one of two groups, the **Training** sample (will be used for model development), or the **Holdout** sample (will be for in-sample validation).

An example of Split-Sample Analysis for model reliability using a Sample window, 2016'Q1:

In-time sample (e.g., Jan. - Feb. 2016)

- Training dataset, 70% of the In-time sample
- Holdout dataset, 30% of the In-time sample

Out-of-Time sample:

- Out of Time validation, e.g., March 2016 data

4. KGB VARIABLE ANALYSIS

Variable analysis includes both variable transformation and reduction. To perform, variable analysis, there are few options available using SAS®:

- SAS® Enterprise Miner
- SAS® PROC HPBIN (High-Performance SAS procedure)
- SAS Macro (there are few SAS macros available online)

Variable Transformation:

Non-Parametric modeling:

- Binning or Weight of Evidence (WOE)
- Widely accepted as the "gold standard"

- Has good interpretability with business implications

Parametric modeling:

- Generalized Linear Model (GLM)
- Generalized Addictive Model (GAM)

Variable Reduction:

- Bootstrapping
- Factor Analysis
- Information Value
- Clustering Analysis
- Kolmogorov–Smirnov
- Principal Component Analysis

5. KGB SCORECARD MODELING

Technically speaking, creating a credit risk score may not be different from other predictive modeling exercises. However, there's a differentiation in the approach of arriving to the final set of predictors in the "best" model.

To determine the final model variables, the selection process requires a blend of business relevance, logical trend, statistical robustness, model implementability and regulatory requirements. This will help to ensure that the model is robust in a way to maximize risk segmentation across different population while also meeting established model risk monitoring and evaluation thresholds.

To develop an Acquisitions or Behavior credit risk score, there are five pillars that are usually considered in selecting the final model attributes by the weight of their contributions:

- Payment History, 35%
- Amount Owed, 30%
- Length of Credit History, 15%
- Credit Mix, 10%
- New Credit, 10%

Logistic Regression modeling using SAS®:

```
/* BUILD A KGB MODEL USING ACCEPTS POPULATION */

proc logistic data=ACCEPTS_KGB_data desc;

model bad_90DPD_24mth = WOE_X1, WOE_X2, WOE_X3, etc. / <options>;

weight <sampling weights>;

run;
```

6. REJECT INFERENCE, KIGB

Fuzzy Augmentation Step 1 (Classification) using SAS®:

```
/* SCORE THE REJECTS USING THE KGB MODEL */
proc logistic data=ACCEPTS_KGB_data desc;
model bad_90DPD_24mth = WOE_X1, WOE_X2, WOE_X3, etc / <options> ;
weight <sampling weights>;
score data=REJECT_RAW_data out = rejects_scored;
run;
/* CREATE DUPLICATE RECORDS FOR EACH SCORED REJECTS*/
data rejects_bad
     rejects_good;
   set rejects_scored;
run;
/* CREATE INFERRED REJECTS BAD DATA AND WEIGHTS FROM SCORED REJECTS (BAD) */
data INFERRED_rejects_bad;
      set rejects_bad;
      bad_KIGB=1;
      wgt_KIGB=sample_wt*PROB_BAD;
      GROUP="REJECT_BAD ";
run;
/* CREATE INFERRED REJECTS GOOD DATA AND WEIGHTS FOR SCORED REJECTS (GOOD)*/
data INFERRED_rejects_good;
     set rejects_good;
     bad_KIGB=0;
     wgt_KIGB=sample_wt*(1-PROB_BAD);
     GROUP="REJECT_GOOD";
run;
```

Fuzzy Augmentation Step 2 (Augmentation) using SAS®:

```
/* KGB population */
data KGB_data;
     set ACCEPTS_KGB_data;
     bad_KIGB= bad_90DPD_24mth;
     wgt_KIGB=sample_wt;
     GROUP="ACCEPT_KGB ";
run;

/* CREATE KIGB data BY MERGING KBG WITH INFERRED REJECTS */
data KIGB_data;
  set KGB_data
      INFERRED_rejects_bad
      INFERRED_rejects_good;
      by id;
run;
```

7. KIGB DATA PARTITION

Same approach used in Step 3 above but now using KIGB data instead of KGB data

8. KIGB VARIABLE ANALYSIS

Same approach used in Step 4 above but now using KIGB data instead of KGB data

9. KIGB SCORECARD MODELING

Same approach discussed in Step 5 above but now using KIGB data and with few minor changes:

```
proc logistic data=KIGB_data desc outest=kigb_test;

model bad_KIGB = KIGB_WOE_X1, KIGB_WOE_X2, KIGB_WOE_X3, etc / <options> ;

weight <sampling weights>;

output out=<such as the modeling data output, say, KIGB_scored_data; the
output of logit, say, kigb_logit; the output of probabilities, say,  prob,
etc.> ;

run;
```

10. KIGB SCORECARD SCALING AND ADVERSE REASON CODES

Scorecard Scaling:

Scorecard scaling is the transformation of log of odds into score using specified **SCALING PARAMETERS**:

- **Base score** (e.g., 200) with a minimum and maximum scale (e.g. 0–1000, 150–350, 300-850)
- **Base Odds** (e.g., 50)
- **Points to Double Odds** (e.g., 20 points)

Simulations and Sensitivity analysis can be performed to determine appropriate score ranges, base score, base odds, and points-to-double-odds (PDO).

It should be noted that the selection of scaling does not affect the predictive ability of the credit risk score. This is an operational decision based on implementability, ease of understanding, continuity with existing scorecards, previously or currently being used in production. Once the final selected score is scaled, the point allocation for each model predictor (or score attribute), and the overall strength of the scorecard should be checked and validated.

Scorecard Scaling using SAS®:

```
/*Sample Scaling Parameters*/
%let PDO = 20;
%let Base_Score = 200;
%let Base_Odds = 50;

* Compute FACTOR and OFFSET */
data KIGB_data;
  set KIGB_scored_data;
```

```
factor= &PDO / log(2);
offset = &Base_Score - (Factor*Log(&Base_Odds));

/* Compute Risk Score */
Credit_Risk_Score = OFFSET  + FACTOR*logit_kigb;

run;
```

Adverse Reason Codes:

In the United States as mandated by regulators, lenders (banks or other financial institutions) are required to provide borrowers with the reasons for declining their credit applications using *Adverse Reason Codes*. These codes can be generated by first obtaining a **Neutral Score**. Any characteristic or attribute for which the applicant scores below the neutral score is then a potential reason for decline.

11. KIGB SCORECARD EVALUATION AND MONITORING

Upon the development of the credit risk model, several statistics and quality metrics are needed for model evaluation and monitoring. These may include:

- Divergence (D)
- Gini's Index (GINI)
- Kolmogorov–Smirnov (K-S)
- EvA: Expected vs. Actual (EvA)
- Population Stability Index (PSI)
- Characteristic Stability Index (CSI)
- Total Population Stability Index (TPSI)
- Log Odds/Point-to-Double (PDO) Analysis
- Area Under Receiver Operating Curve (AUROC or simply ROC)

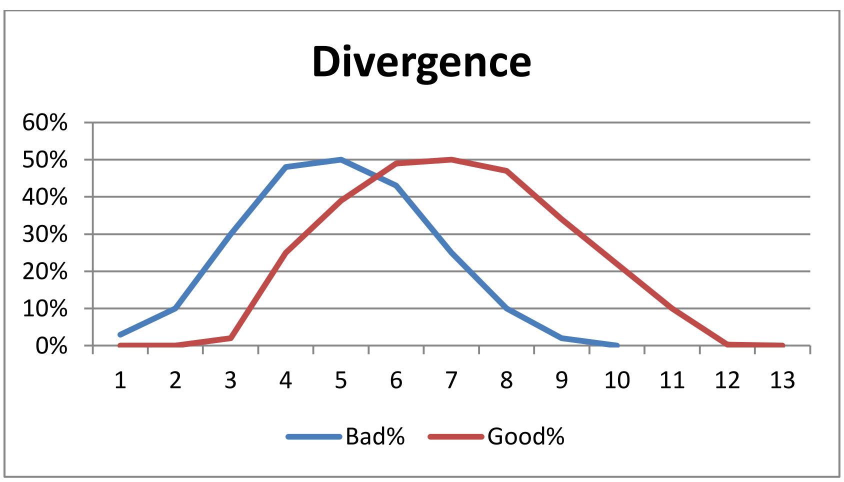


Figure 3. Sample Divergence

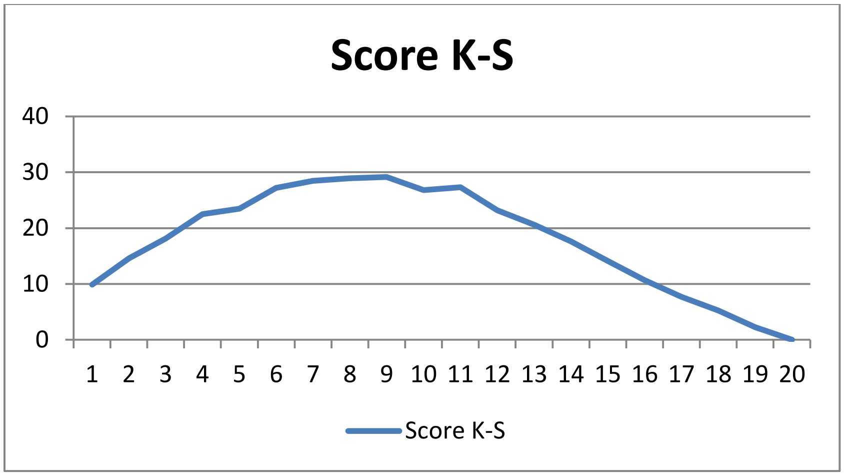


Figure 4. Sample K-S by Cumulative %

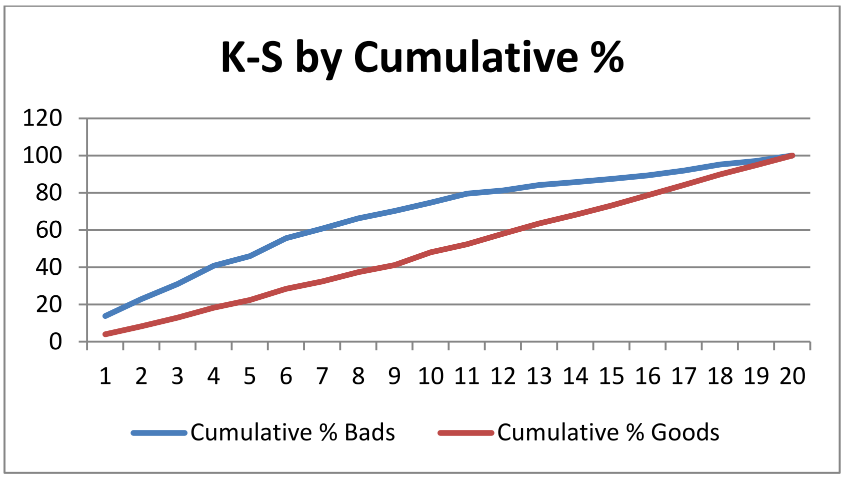


Figure 5. Sample K-S by Cumulative %

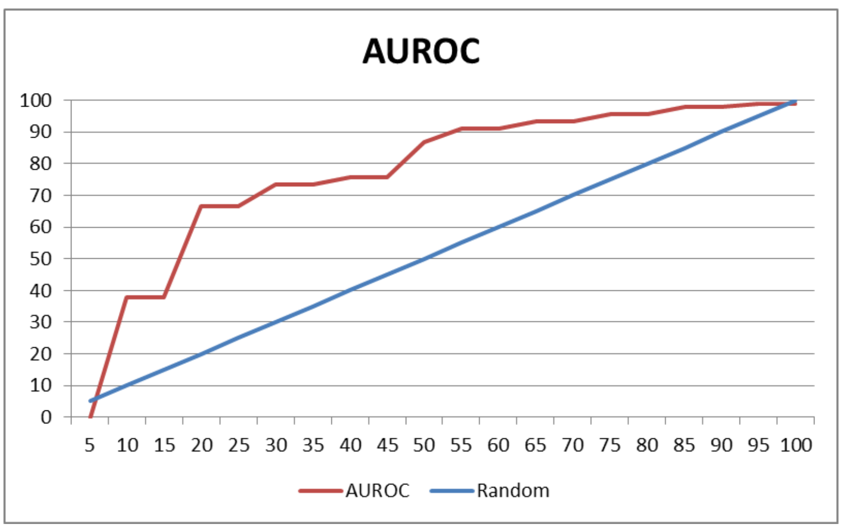


Figure 6. Sample AUROC plot

Population Stability Index								
Risk Score Range	TRAINING freq	HOLDOUT freq	TRAINING %	HOLDOUT %	% Difference	% Ratio	Natural Log of Ratio	Contribution to Index
< 350	2,500	2,345	10.51%	9.77%	0.74%	0.929633	-0.072965	0.000540
351 -< 400	2,503	2,401	10.53%	10.01%	0.52%	0.950694	-0.050563	0.000262
401 -< 450	2,805	2,402	11.80%	10.01%	1.79%	0.848690	-0.164061	0.002929
451 -< 500	2,177	2,403	9.16%	10.02%	-0.86%	1.093965	0.089809	0.000773
501 -< 550	2,444	2,404	10.28%	10.02%	0.26%	0.974856	-0.025465	0.000066
551 -< 600	2,509	2,405	10.55%	10.03%	0.53%	0.949995	-0.051299	0.000271
601 -< 650	2,001	2,406	8.42%	10.03%	-1.61%	1.191680	0.175364	0.002829
651 -< 700	2,512	2,407	10.57%	10.03%	0.53%	0.949651	-0.051661	0.000275
701 -< 750	2,098	2,408	8.82%	10.04%	-1.21%	1.137524	0.128854	0.001564
> 751	2,227	2,409	9.37%	10.04%	-0.68%	1.072074	0.069595	0.000470
							PSI =	1.00%

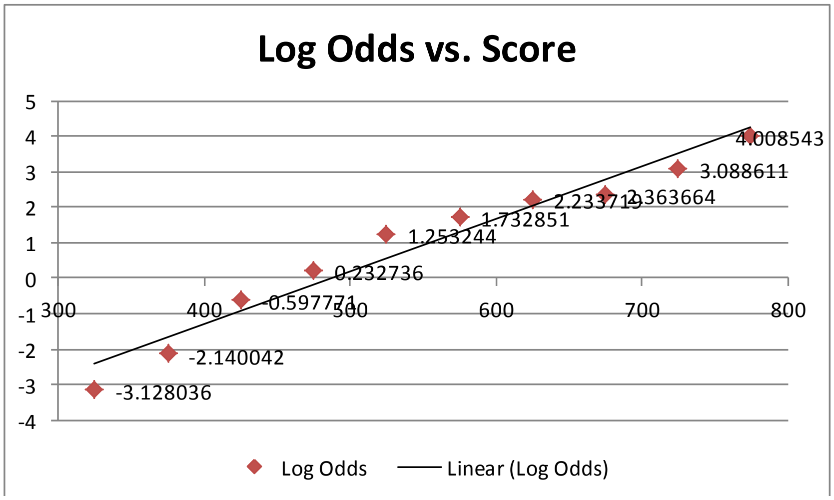
Figure 7. Sample PSI Table

EvA: Expected versus Actual Bad with Rank-ordering								
Credit Risk Score Range	Expected Bads freq	Actual Bads freq	Expected %	Actual %	% Difference	% Ratio	Natural Log of Ratio	Contribution to Index
Low -< 350	8,769	8,339	36.88%	33.09%	3.80%	0.897099	-0.108589	0.004121
351 -< 400	7,261	6,796	30.54%	26.97%	3.58%	0.882936	-0.124503	0.004451
401 -< 450	3,619	4,292	15.22%	17.03%	-1.81%	1.118783	0.112241	0.002029
451 -< 500	1,982	2,239	8.34%	8.89%	-0.55%	1.066012	0.063925	0.000352
501 -< 550	818	1,468	3.44%	5.83%	-2.39%	1.693319	0.526691	0.012562
551 -< 600	523	1,012	2.20%	4.02%	-1.82%	1.827354	0.602869	0.010963
601 -< 650	323	326	1.36%	1.29%	0.07%	0.951828	-0.049371	0.000032
651 -< 700	285	528	1.20%	2.09%	-0.89%	1.745680	0.557144	0.004981
701 -< 750	140	141	0.59%	0.56%	0.03%	0.951828	-0.049371	0.000014
751 >- High	56	62	0.24%	0.25%	-0.01%	1.035721	0.035098	0.000003
							PSI =	3.95%

Figure 8. Sample EvA Table

Log Odd and Point-to-Double-Odds (PDO)								
Risk Score Range	Score	Bads freq	Goods freq	Total freq	Bads %	Goods %	Odds	Log Odds
Low -< 350	325	8,769	1,231	10,000	36.88%	1.62%	0.043804	-3.128036
351 -< 400	375	7,261	2,739	10,000	30.54%	3.59%	0.117650	-2.140042
401 -< 450	425	3,619	6,381	10,000	15.22%	8.37%	0.550036	-0.597771
451 -< 500	475	1,982	8,018	10,000	8.34%	10.52%	1.262048	0.232736
501 -< 550	525	818	9,182	10,000	3.44%	12.05%	3.501686	1.253244
551 -< 600	575	523	9,477	10,000	2.20%	12.43%	5.656758	1.732851
601 -< 650	625	323	9,677	10,000	1.36%	12.69%	9.334518	2.233719
651 -< 700	675	285	9,715	10,000	1.20%	12.75%	10.629828	2.363664
701 -< 750	725	140	9,860	10,000	0.59%	12.94%	21.946574	3.088611
751 >- High	775	56	9,944	10,000	0.24%	13.05%	55.066565	4.008543

Figure 9. Sample PDO



Scaling Parameters:

PDO	20
Base Odds	50
Base Score	200

PDO = Ln(2) / Slope

Figure 10. Sample PDO plot

APPENDIX: MODEL LIFE CYCLE

Model Life Cycle management requires governance, risk and control involving three key stakeholders:

- Model Development (MD) Team
- Model Validation (MV) Team
- Model Risk Management (MRM) Team

***Model Life Cycle process*:**

- Model Identification
- Model Development
- Model Validation and Risk Assessment
- Model Approval for Use
- Model Implementation
- Model Performance Monitoring
- Model Annual Review
- Model Validation Ongoing
- Model Quarterly Attestation
- Model Replacement or Decommissioning

CONCLUSION

This paper started by providing a high-level overview of risk management with application to the financial industry. Then I shed a little light on, as a gentle introduction to, credit risk management and the prospect-customer lifecycle. Further, I defined credit risk score as an analytical method of modeling the credit riskiness of individual borrowers (prospects and customers). Later, I dived into the algorithms and methodologies usually employed to develop a credit risk score. Lastly, I showcased using SAS® the application of Statistical Logistic Regression for probability of default (PD) modeling and Fuzzy Augmentation for reject inference while highlighting weight of evidence (WOE) approach for variable transformation, and bootstrapping, information value, Kolmogorov–Simonov, for variable reduction. While this paper focuses on financial application, the methods, algorithms, and approaches presented can be extended and expanded into other industries such as health care, telecommunication, energy, to mention few.

REFERENCES

Agresti, A. (2013). Categorical Data Analysis, 3rd edition. John Wiley & Sons, Inc., Hoboken, NJ

Allison, P. (2012). Logistic Regression using SAS (Theory & Applications), 2nd edition. SAS Institute, Cary, NC

Baesens B., Rosch D., & Schenle H. (2017). Credit Risk Analytics. Wiley & SAS Institute

Hastie, T.J. & Tibshirani, R.J. (1990). Generalized Additive Models. Chapman and Hall/CRC, Boca Raton, FL

Hayashi, F. (2000). Econometrics. Princeton University Press, NJ

Crouhy M., Galai, D. & Mark R. (2000) Essential of Risk Management. McGraw-Hill/Irwin, New York NY

FICO World 2018 Bootcamp Course. Miami Beach, FL

James, Witten, Hastie, & Tibshirani. (2009). An Introduction to Statistical Learning. Springer, New York NY

Johnson, R.A. & Wichern, D.W. (2007). Applied Multivariate Statistical Analysis, 6th edition. Pearson, Upper Saddle River, NJ

Kutner, Nachtsheim , Neter, & Li (2004). Applied Linear Statistical Models, 5th edition. McGraw-Hill/Irwin, New York NY

Lange, K. (2010). Numerical Analysis for Statisticians. Springer, New York NY

Lawrence & Solomon: Managing Customer Lending Business, 2nd edition. Solomon Lawrence Partners

McCullagh P. & Nelder, J.A. (1989). Generalized Linear Models, Chapman and Hall/CRC, Boca Raton, FL

Siddiqi N. (2018). Intelligent Credit Scoring (Building and Implementing Better Credit Risk Scorecards), 2nd edition, Wiley & SAS Institute

Stokes, M. E., Davis, C. S., & Koch, G. G. (2000). Categorical Data analysis using the SAS System, 2nd edition, SAS Institute, Cary, NC

Thomas, L. C. (2009). Consumer Credit Models (Pricing, Profits and Portfolios). Oxford University Press, New York NY

ACKNOWLEDGMENTS

The author is grateful to Kurt Lessig, Head of U.S. Credit Cards and Unsecured Lending, TD Bank Group, for his leadership and support. Thanks to everyone that have helped and supported me along the way of my professional and educational careers especially my family (home and abroad).

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Amos Taiwo Odeleye
267.461.8599
oduatechnology@gmail.com