



From *Developing Credit Risk Models Using SAS® Enterprise Miner™ and SAS/STAT®*. Full book available for purchase [here](#).

## Contents

<b>About this Book</b> .....	<b>ix</b>
<b>About the Author</b> .....	<b>xiii</b>
<b>Acknowledgments</b> .....	<b>xv</b>
<b>Chapter 1 Introduction</b> .....	<b>1</b>
1.1 Book Overview.....	1
1.2 Overview of Credit Risk Modeling .....	2
1.3 Regulatory Environment.....	3
1.3.1 Minimum Capital Requirements.....	4
1.3.2 Expected Loss.....	5
1.3.3 Unexpected Loss .....	6
1.3.4 Risk Weighted Assets .....	6
1.4 SAS Software Utilized .....	7
1.5 Chapter Summary .....	11
1.6 References and Further Reading .....	11
<b>Chapter 2 Sampling and Data Pre-Processing</b> .....	<b>13</b>
2.1 Introduction .....	13
2.2 Sampling and Variable Selection.....	16
2.2.1 Sampling .....	17
2.2.2 Variable Selection .....	18
2.3 Missing Values and Outlier Treatment.....	19
2.3.1 Missing Values .....	19
2.3.2 Outlier Detection.....	21
2.4 Data Segmentation .....	22
2.4.1 Decision Trees for Segmentation .....	23
2.4.2 K-Means Clustering.....	24

2.5 Chapter Summary .....	25
2.6 References and Further Reading .....	25
<b>Chapter 3 Development of a Probability of Default (PD) Model .....</b>	<b>27</b>
3.1 Overview of Probability of Default.....	27
3.1.1 PD Models for Retail Credit.....	28
3.1.2 PD Models for Corporate Credit .....	28
3.1.3 PD Calibration .....	29
3.2 Classification Techniques for PD .....	29
3.2.1 Logistic Regression.....	29
3.2.2 Linear and Quadratic Discriminant Analysis .....	31
3.2.3 Neural Networks .....	32
3.2.4 Decision Trees .....	33
3.2.5 Memory Based Reasoning.....	34
3.2.6 Random Forests .....	34
3.2.7 Gradient Boosting.....	35
3.3 Model Development (Application Scorecards) .....	35
3.3.1 Motivation for Application Scorecards.....	36
3.3.2 Developing a PD Model for Application Scoring .....	36
3.4 Model Development (Behavioral Scoring) .....	47
3.4.1 Motivation for Behavioral Scorecards.....	48
3.4.2 Developing a PD Model for Behavioral Scoring .....	49
3.5 PD Model Reporting.....	52
3.5.1 Overview .....	52
3.5.2 Variable Worth Statistics .....	52
3.5.3 Scorecard Strength .....	54
3.5.4 Model Performance Measures .....	54
3.5.5 Tuning the Model.....	54
3.6 Model Deployment .....	55
3.6.1 Creating a Model Package .....	55
3.6.2 Registering a Model Package .....	56
3.7 Chapter Summary .....	57
3.8 References and Further Reading .....	58

<b>Chapter 4 Development of a Loss Given Default (LGD) Model.....</b>	<b>59</b>
4.1 Overview of Loss Given Default.....	59
4.1.1 LGD Models for Retail Credit .....	60
4.1.2 LGD Models for Corporate Credit.....	60
4.1.3 Economic Variables for LGD Estimation .....	61
4.1.4 Estimating Downturn LGD .....	61
4.2 Regression Techniques for LGD.....	62
4.2.1 Ordinary Least Squares – Linear Regression .....	64
4.2.2 Ordinary Least Squares with Beta Transformation .....	64
4.2.3 Beta Regression .....	65
4.2.4 Ordinary Least Squares with Box-Cox Transformation .....	66
4.2.5 Regression Trees.....	67
4.2.6 Artificial Neural Networks.....	67
4.2.7 Linear Regression and Non-linear Regression .....	68
4.2.8 Logistic Regression and Non-linear Regression.....	68
4.3 Performance Metrics for LGD.....	69
4.3.1 Root Mean Squared Error .....	69
4.3.2 Mean Absolute Error .....	70
4.3.3 Area Under the Receiver Operating Curve .....	70
4.3.4 Area Over the Regression Error Characteristic Curves .....	71
4.3.5 R-square .....	72
4.3.6 Pearson’s Correlation Coefficient.....	72
4.3.7 Spearman’s Correlation Coefficient .....	72
4.3.8 Kendall’s Correlation Coefficient.....	73
4.4 Model Development.....	73
4.4.1 Motivation for LGD models.....	73
4.4.2 Developing an LGD Model.....	73
4.5 Case Study: Benchmarking Regression Algorithms for LGD .....	77
4.5.1 Data Set Characteristics .....	77
4.5.2 Experimental Set-Up .....	78
4.5.3 Results and Discussion.....	79
4.6 Chapter Summary .....	83
4.7 References and Further Reading .....	84

<b>Chapter 5 Development of an Exposure at Default (EAD) Model .....</b>	<b>87</b>
5.1 Overview of Exposure at Default .....	87
5.2 Time Horizons for CCF .....	88
5.3 Data Preparation .....	90
5.4 CCF Distribution – Transformations.....	95
5.5 Model Development .....	97
5.5.1 Input Selection .....	97
5.5.2 Model Methodology.....	97
5.5.3 Performance Metrics.....	99
5.6 Model Validation and Reporting .....	103
5.6.1 Model Validation .....	103
5.6.2 Reports .....	104
5.7 Chapter Summary .....	106
5.8 References and Further Reading .....	107
<b>Chapter 6 Stress Testing .....</b>	<b>109</b>
6.1 Overview of Stress Testing .....	109
6.2 Purpose of Stress Testing .....	110
6.3 Stress Testing Methods.....	111
6.3.1 Sensitivity Testing.....	111
6.3.2 Scenario Testing .....	112
6.4 Regulatory Stress Testing .....	113
6.5 Chapter Summary .....	114
6.6 References and Further Reading .....	114
<b>Chapter 7 Producing Model Reports .....</b>	<b>115</b>
7.1 Surfacing Regulatory Reports .....	115
7.2 Model Validation.....	115
7.2.1 Model Performance .....	116
7.2.2 Model Stability .....	122
7.2.3 Model Calibration .....	125
7.3 SAS Model Manager Examples.....	127
7.3.1 Create a PD Report .....	127
7.3.2 Create a LGD Report.....	129
7.4 Chapter Summary .....	130

<b>Tutorial A – Getting Started with SAS Enterprise Miner.....</b>	<b>131</b>
A.1 Starting SAS Enterprise Miner .....	131
A.2 Assigning a Library Location .....	134
A.3 Defining a New Data Set.....	136
<b>Tutorial B – Developing an Application Scorecard Model in SAS Enterprise Miner.....</b>	<b>139</b>
B.1 Overview .....	139
B.1.1 Step 1 – Import the XML Diagram .....	140
B.1.2 Step 2 – Define the Data Source.....	140
B.1.3 Step 3 – Visualize the Data.....	141
B.1.4 Step 4 – Partition the Data .....	143
B.1.5 Step 5 –Perform Screening and Grouping with Interactive Grouping .....	143
B.1.6 Step 6 – Create a Scorecard and Fit a Logistic Regression Model .....	144
B.1.7 Step 7 – Create a Rejected Data Source .....	144
B.1.8 Step 8 – Perform Reject Inference and Create an Augmented Data Set .....	144
B.1.9 Step 9 – Partition the Augmented Data Set into Training, Test and Validation Samples .....	145
B.1.10 Step 10 – Perform Univariate Characteristic Screening and Grouping on the Augmented Data Set .....	145
B.1.11 Step 11 – Fit a Logistic Regression Model and Score the Augmented Data Set .....	145
B.2 Tutorial Summary .....	146
<b>Appendix A Data Used in This Book .....</b>	<b>147</b>
A.1 Data Used in This Book.....	147
Chapter 3: Known Good Bad Data.....	147
Chapter 3: Rejected Candidates Data .....	148
Chapter 4: LGD Data .....	148
Chapter 5: Exposure at Default Data .....	149
<b>Index .....</b>	<b>151</b>

From *Developing Credit Risk Models Using SAS® Enterprise Miner™ and SAS/STAT®: Theory and Application*, by Iain Brown. Copyright © 2014, SAS Institute Inc., Cary, North Carolina, USA. ALL RIGHTS RESERVED.