

Business Analytics Using SAS® Enterprise Guide® and SAS® Enterprise Miner™ A Beginner's Guide



Olivia Parr-Rud

Business Analytics Using SAS[®] Enterprise Guide[®] and SAS[®] Enterprise Miner[™] A Beginner's Guide

Olivia Parr-Rud



support.sas.com/bookstore



The correct bibliographic citation for this manual is as follows: Parr-Rud, Olivia. 2014. *Business Analytics Using SAS® Enterprise Guide® and SAS® Enterprise Miner®: A Beginner's Guide*. Cary, NC: SAS Institute Inc.

Business Analytics Using SAS® Enterprise Guide® and SAS® Enterprise Miner®: A Beginner's Guide

Copyright © 2014, SAS Institute Inc., Cary, NC, USA

ISBN 978-1-62959-327-2

All rights reserved. Produced in the United States of America.

For a hard-copy book: No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission of the publisher, SAS Institute Inc.

For a web download or e-book: Your use of this publication shall be governed by the terms established by the vendor at the time you acquire this publication.

The scanning, uploading, and distribution of this book via the Internet or any other means without the permission of the publisher is illegal and punishable by law. Please purchase only authorized electronic editions and do not participate in or encourage electronic piracy of copyrighted materials. Your support of others' rights is appreciated.

U.S. Government License Rights; Restricted Rights: The Software and its documentation is commercial computer software developed at private expense and is provided with RESTRICTED RIGHTS to the United States Government. Use, duplication or disclosure of the Software by the United States Government is subject to the license terms of this Agreement pursuant to, as applicable, FAR 12.212, DFAR 227.7202-1(a), DFAR 227.7202-3(a) and DFAR 227.7202-4 and, to the extent required under U.S. federal law, the minimum restricted rights as set out in FAR 52.227-19 (DEC 2007). If FAR 52.227-19 is applicable, this provision serves as notice under clause (c) thereof and no other notice is required to be affixed to the Software or documentation. The Government's rights in Software and documentation shall be only those set forth in this Agreement.

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513-2414.

October 2014

SAS provides a complete selection of books and electronic products to help customers use SAS® software to its fullest potential. For more information about our offerings, visit support.sas.com/bookstore or call 1-800-727-3228.

SAS® and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

Contents

| | |
|---|------------|
| About This Book | vii |
| About the Author | xi |
| Chapter 1: Defining the Business Objective | 1 |
| Introduction | 1 |
| Setting Goals | 1 |
| Descriptive Analyses..... | 3 |
| Customer Profile..... | 3 |
| Customer Loyalty..... | 4 |
| Market Penetration or Wallet Share | 4 |
| Predictive Analyses..... | 4 |
| Marketing Models..... | 5 |
| Risk and Approval Models..... | 6 |
| Predictive Modeling Opportunities by Industry | 9 |
| Notes from the Field | 13 |
| Chapter 2: Data Types, Categories, and Sources | 15 |
| Introduction | 15 |
| The Evolution of Data..... | 16 |
| Types of Data..... | 17 |
| Nominal Data..... | 17 |
| Ordinal Data | 17 |
| Continuous Data | 18 |
| Categories of Data | 18 |
| Demographic or Firmographic Data..... | 18 |
| Behavioral Data..... | 19 |
| Psychographic Data | 20 |

| | |
|---|-----------|
| Data Category Comparison | 20 |
| Sources of Data | 21 |
| Internal Sources | 21 |
| Storage of Data | 27 |
| External Sources | 28 |
| Notes from the Field | 29 |
| Chapter 3: Overview of Descriptive and Predictive Analyses..... | 29 |
| Introduction | 29 |
| Descriptive Analyses..... | 30 |
| Frequency Distributions | 30 |
| Cluster..... | 33 |
| Decision Tree | 33 |
| Predictive Analyses..... | 35 |
| Linear Regression..... | 36 |
| Logistic Regression..... | 39 |
| Neural Networks | 41 |
| Modeling Process | 43 |
| Define the Objective | 43 |
| Develop the Model..... | 43 |
| Implement the Model | 48 |
| Maintain the Model..... | 49 |
| Notes from the Field | 52 |
| Chapter 4: Data Construction for Analysis | 53 |
| Introduction | 53 |
| Data for Descriptive Analysis..... | 53 |
| Data for Predictive Analysis | 54 |
| Prospect Models..... | 55 |
| Customer Models | 57 |
| Risk Models..... | 59 |
| External Sources of Data..... | 61 |
| Notes from the Field | 61 |
| Chapter 5: Descriptive Analysis Using SAS Enterprise Guide | 63 |
| Introduction | 63 |
| Project Overview | 63 |

| | |
|--|------------|
| Project Initiation | 64 |
| Exploratory Analysis | 65 |
| Importing the Data..... | 65 |
| Viewing the Data | 66 |
| Exploring the Data | 66 |
| Segmentation and Profile Analysis..... | 69 |
| Correlation Analysis..... | 76 |
| Notes from the Field | 77 |
| Chapter 6: Market Analysis Using SAS Enterprise Guide | 79 |
| Introduction | 79 |
| Project Overview | 79 |
| Market Analysis | 80 |
| Project Initiation..... | 80 |
| Data Preparation..... | 80 |
| Penetration and Share of Wallet | 89 |
| Results | 90 |
| Notes from the Field | 91 |
| Chapter 7: Cluster Analysis Using SAS Enterprise Miner..... | 93 |
| Introduction | 93 |
| Project Overview | 93 |
| Cluster Analysis..... | 94 |
| Initiate the Project | 94 |
| Input the Data Source and Assign Variable Roles | 97 |
| Transform Variables | 99 |
| Filter Data | 102 |
| Build Clusters..... | 104 |
| Build Segment Profiles..... | 107 |
| Analyze Clusters and Recommend Marketing or Product Development Actions..... | 109 |
| Notes from the Field | 109 |
| Chapter 8: Tree Analysis Using SAS Enterprise Miner..... | 111 |
| Introduction | 111 |
| Project Overview | 111 |
| Decision Tree Analysis | 112 |
| Initiate the Project | 112 |

| | |
|--|------------|
| Input the Data Source | 114 |
| Create Target Variable | 115 |
| Partition the Data..... | 117 |
| Build the Decision Tree | 118 |
| View the Decision Tree Output..... | 120 |
| Interpret the Findings | 126 |
| Alternate Uses for Tree Analysis..... | 128 |
| Notes from the Field | 128 |
| Chapter 9: Predictive Analysis Using SAS Enterprise Miner | 129 |
| Introduction | 129 |
| Select..... | 130 |
| Initiate the Project | 130 |
| Select the Data..... | 131 |
| Explore | 133 |
| StatExplore..... | 133 |
| MultiPlot..... | 136 |
| Modify..... | 138 |
| Replace Missing Values via Imputation..... | 138 |
| Partition Data into Subsamples | 139 |
| Manage Outliers..... | 140 |
| Transform the Variables..... | 142 |
| Model..... | 145 |
| Decision Tree | 145 |
| Neural Network | 147 |
| Regression | 148 |
| Assess | 151 |
| Notes from the Field | 155 |
| References..... | 157 |

About This Book

Purpose

This book serves as a tutorial for data analysts who are new to SAS Enterprise Guide and SAS Enterprise Miner. It provides valuable hands-on experience using powerful statistical software to complete the kinds of business analytics common to most industries. With clear, illustrated, step-by-step instructions, it will lead you through examples based on business case studies. You will formulate the business objective, manage the data, and perform analyses that you can use to optimize marketing, risk, and customer relationship management, as well as business processes and human resources.

Prerequisites

If you are a savvy business person with a desire to understand what drives your business, then this book can help you get started. You need access to SAS Enterprise Guide or SAS Enterprise Miner software; we provide you with example data to get started, but you will need data to analyze. An understanding of basic statistics is helpful, but not required.

Organization

The book begins by helping you determine and structure the objective of your analysis in accordance with the goals and objectives of your organization or department.

Chapter 2 describes types and sources of data for analysis. Chapter 3 offers an overview of common business analyses, covering both descriptive and predictive analysis. Chapter 4 shows you how to construct a data set for analysis. Chapter 5 details step-by-step instructions for a simple descriptive analysis. Chapter 6 offers the same level of detail for a typical market analysis. Chapters 7 and 8 offer a step-by-step guide to cluster and tree analyses, respectively. Each chapter concludes with a section headed “Notes from the Field,” which offers related business advice and leadership tips.

To conclude, Chapter 9 brings several concepts together in a full step-by-step case study for building and comparing predictive models, culminating in final “Notes from the Field.”