



Modern Analysis of Customer Surveys: with applications using R



Editors
RON S. KENETT
SILVIA SALINI

 **WILEY**

STATISTICS IN PRACTICE

Modern Analysis of Customer Surveys

Statistics in Practice

Series Advisors

Human and Biological Sciences

Stephen Senn

University of Glasgow, UK

Earth and Environmental Sciences

Marian Scott

University of Glasgow, UK

Industry, Commerce and Finance

Wolfgang Jank

University of Maryland, USA

Statistics in Practice is an important international series of texts which provide detailed coverage of statistical concepts, methods and worked case studies in specific fields of investigation and study.

With sound motivation and many worked practical examples, the books show in down-to-earth terms how to select and use an appropriate range of statistical techniques in a particular practical field within each title's special topic area.

The books provide statistical support for professionals and research workers across a range of employment fields and research environments. Subject areas covered include medicine and pharmaceuticals; industry, finance and commerce; public services; the earth and environmental sciences, and so on.

The books also provide support to students studying statistical courses applied to the above areas. The demand for graduates to be equipped for the work environment has led to such courses becoming increasingly prevalent at universities and colleges.

It is our aim to present judiciously chosen and well-written workbooks to meet everyday practical needs. Feedback of views from readers will be most valuable to monitor the success of this aim.

A complete list of titles in this series can be found at <http://eu.wiley.com/wileyCDA/Section/id-300612.html>

Modern Analysis of Customer Surveys

with applications using R

Edited by

Ron S. Kenett

*KPA Ltd., Raanana, Israel, University of Turin, Italy, and
NYU-Poly, Center for Risk Engineering, New York, USA*

Silvia Salini

*Department of Economics, Business and Statistics,
University of Milan, Italy*



A John Wiley & Sons, Ltd., Publication

This edition first published 2012
© 2012 John Wiley & Sons, Ltd

Registered office

John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com.

The right of the author to be identified as the author of this work has been asserted in accordance with the Copyright, Designs and Patents Act 1988.

All rights reserved. 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, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data

Modern analysis of customer surveys : with applications using R / edited by Ron S. Kenett, Silvia Salini.
p. cm.

Includes bibliographical references and index.

ISBN 978-0-470-97128-4 (cloth)

1. Consumer satisfaction – Research – Statistical methods. 2. Consumer satisfaction – Evaluation.
3. Consumers – Research – Statistical methods. 4. Consumers – Research – Data processing. 5. Sampling (Statistics) – Evaluation. 6. Surveys – Statistical methods. 7. Surveys – Data processing. I. Kenett, Ron.
II. Salini, Silvia.

HF5415.335.M63 2011

658.8/3402855282–dc23

2011030182

A catalogue record for this book is available from the British Library.

Print ISBN: 978-0-470-97128-4

ePDF ISBN: 978-1-119-96116-1

oBook ISBN: 978-1-119-96115-4

ePub ISBN: 978-1-119-96138-3

Mobi ISBN: 978-1-119-96139-0

Set in 10/12pt Times by Aptara Inc., New Delhi, India

*To Sima, our children, their families and children;
Jonathan, Alma, Tomer and Yadin, they are my sources
of inspiration and motivation.*

Ron

To Riccardo, Marco and Angela: my cornerstones.

Silvia

In memory of Ennio Isaia

Contents

Foreword	xvii
Preface	xix
Contributors	xxiii

PART I BASIC ASPECTS OF CUSTOMER SATISFACTION SURVEY DATA ANALYSIS

1 Standards and classical techniques in data analysis of customer satisfaction surveys	3
<i>Silvia Salini and Ron S. Kenett</i>	
1.1 Literature on customer satisfaction surveys	4
1.2 Customer satisfaction surveys and the business cycle	4
1.3 Standards used in the analysis of survey data	7
1.4 Measures and models of customer satisfaction	12
1.4.1 The conceptual construct	12
1.4.2 The measurement process	13
1.5 Organization of the book	15
1.6 Summary	17
References	17
2 The ABC annual customer satisfaction survey	19
<i>Ron S. Kenett and Silvia Salini</i>	
2.1 The ABC company	19
2.2 ABC 2010 ACSS: Demographics of respondents	20
2.3 ABC 2010 ACSS: Overall satisfaction	22
2.4 ABC 2010 ACSS: Analysis of topics	24
2.5 ABC 2010 ACSS: Strengths and weaknesses and decision drivers	27
2.6 Summary	28
References	28
Appendix	29
3 Census and sample surveys	37
<i>Giovanna Nicolini and Luciana Dalla Valle</i>	
3.1 Introduction	37

3.2	Types of surveys	39
3.2.1	Census and sample surveys	39
3.2.2	Sampling design	40
3.2.3	Managing a survey	40
3.2.4	Frequency of surveys	41
3.3	Non-sampling errors	41
3.3.1	Measurement error	42
3.3.2	Coverage error	42
3.3.3	Unit non-response and non-self-selection errors	43
3.3.4	Item non-response and non-self-selection error	44
3.4	Data collection methods	44
3.5	Methods to correct non-sampling errors	46
3.5.1	Methods to correct unit non-response errors	46
3.5.2	Methods to correct item non-response	49
3.6	Summary	51
	References	52
4	Measurement scales	55
	<i>Andrea Bonanomi and Gabriele Cantaluppi</i>	
4.1	Scale construction	55
4.1.1	Nominal scale	56
4.1.2	Ordinal scale	57
4.1.3	Interval scale	58
4.1.4	Ratio scale	59
4.2	Scale transformations	60
4.2.1	Scale transformations referred to single items	61
4.2.2	Scale transformations to obtain scores on a unique interval scale	66
	Acknowledgements	69
	References	69
5	Integrated analysis	71
	<i>Silvia Biffignandi</i>	
5.1	Introduction	71
5.2	Information sources and related problems	73
5.2.1	Types of data sources	73
5.2.2	Advantages of using secondary source data	73
5.2.3	Problems with secondary source data	74
5.2.4	Internal sources of secondary information	75
5.3	Root cause analysis	78
5.3.1	General concepts	78
5.3.2	Methods and tools in RCA	81
5.3.3	Root cause analysis and customer satisfaction	85
5.4	Summary	87
	Acknowledgement	87
	References	87

6	Web surveys	89
	<i>Roberto Furlan and Diego Martone</i>	
6.1	Introduction	89
6.2	Main types of web surveys	90
6.3	Economic benefits of web survey research	91
6.3.1	Fixed and variable costs	92
6.4	Non-economic benefits of web survey research	94
6.5	Main drawbacks of web survey research	96
6.6	Web surveys for customer and employee satisfaction projects	100
6.7	Summary	102
	References	102
7	The concept and assessment of customer satisfaction	107
	<i>Irena Ograjenšek and Iddo Gal</i>	
7.1	Introduction	107
7.2	The quality–satisfaction–loyalty chain	108
7.2.1	Rationale	108
7.2.2	Definitions of customer satisfaction	108
7.2.3	From general conceptions to a measurement model of customer satisfaction	110
7.2.4	Going beyond SERVQUAL: Other dimensions of relevance to the B2B context	112
7.2.5	From customer satisfaction to customer loyalty	113
7.3	Customer satisfaction assessment: Some methodological considerations	115
7.3.1	Rationale	115
7.3.2	Think big: An assessment programme	115
7.3.3	Back to basics: Questionnaire design	116
7.3.4	Impact of questionnaire design on interpretation	118
7.3.5	Additional concerns in the B2B setting	119
7.4	The ABC ACSS questionnaire: An evaluation	119
7.4.1	Rationale	119
7.4.2	Conceptual issues	119
7.4.3	Methodological issues	120
7.4.4	Overall ABC ACSS questionnaire assessment	121
7.5	Summary	121
	References	122
	Appendix	126
8	Missing data and imputation methods	129
	<i>Alessandra Mattei, Fabrizia Mealli and Donald B. Rubin</i>	
8.1	Introduction	129
8.2	Missing-data patterns and missing-data mechanisms	131
8.2.1	Missing-data patterns	131
8.2.2	Missing-data mechanisms and ignorability	132
8.3	Simple approaches to the missing-data problem	134
8.3.1	Complete-case analysis	134
8.3.2	Available-case analysis	135
8.3.3	Weighting adjustment for unit nonresponse	135

8.4	Single imputation	136
8.5	Multiple imputation	138
8.5.1	Multiple-imputation inference for a scalar estimand	138
8.5.2	Proper multiple imputation	139
8.5.3	Appropriately drawing imputations with monotone missing-data patterns	140
8.5.4	Appropriately drawing imputations with nonmonotone missing-data patterns	141
8.5.5	Multiple imputation in practice	142
8.5.6	Software for multiple imputation	143
8.6	Model-based approaches to the analysis of missing data	144
8.7	Addressing missing data in the ABC annual customer satisfaction survey: An example	145
8.8	Summary	149
	Acknowledgements	150
	References	150
9	Outliers and robustness for ordinal data	155
	<i>Marco Riani, Francesca Torti and Sergio Zani</i>	
9.1	An overview of outlier detection methods	155
9.2	An example of masking	157
9.3	Detection of outliers in ordinal variables	159
9.4	Detection of bivariate ordinal outliers	160
9.5	Detection of multivariate outliers in ordinal regression	161
9.5.1	Theory	161
9.5.2	Results from the application	163
9.6	Summary	168
	References	168

PART II MODERN TECHNIQUES IN CUSTOMER SATISFACTION SURVEY DATA ANALYSIS

10	Statistical inference for causal effects	173
	<i>Fabrizia Mealli, Barbara Pacini and Donald B. Rubin</i>	
10.1	Introduction to the potential outcome approach to causal inference	173
10.1.1	Causal inference primitives: Units, treatments, and potential outcomes	175
10.1.2	Learning about causal effects: Multiple units and the stable unit treatment value assumption	176
10.1.3	Defining causal estimands	177
10.2	Assignment mechanisms	179
10.2.1	The criticality of the assignment mechanism	179
10.2.2	Unconfounded and strongly ignorable assignment mechanisms	180
10.2.3	Confounded and ignorable assignment mechanisms	181
10.2.4	Randomized and observational studies	181

10.3	Inference in classical randomized experiments	182
10.3.1	Fisher's approach and extensions	183
10.3.2	Neyman's approach to randomization-based inference	183
10.3.3	Covariates, regression models, and Bayesian model-based inference	184
10.4	Inference in observational studies	185
10.4.1	Inference in regular designs	186
10.4.2	Designing observational studies: The role of the propensity score	186
10.4.3	Estimation methods	188
10.4.4	Inference in irregular designs	188
10.4.5	Sensitivity and bounds	189
10.4.6	Broken randomized experiments as templates for the analysis of some irregular designs	189
	References	190
11	Bayesian networks applied to customer surveys	193
	<i>Ron S. Kenett, Giovanni Perruca and Silvia Salini</i>	
11.1	Introduction to Bayesian networks	193
11.2	The Bayesian network model in practice	197
11.2.1	Bayesian network analysis of the ABC 2010 ACSS	197
11.2.2	Transport data analysis	201
11.2.3	R packages and other software programs used for studying BNs	210
11.3	Prediction and explanation	211
11.4	Summary	213
	References	213
12	Log-linear model methods	217
	<i>Stephen E. Fienberg and Daniel Manrique-Vallier</i>	
12.1	Introduction	217
12.2	Overview of log-linear models and methods	218
12.2.1	Two-way tables	218
12.2.2	Hierarchical log-linear models	220
12.2.3	Model search and selection	222
12.2.4	Sparseness in contingency tables and its implications	223
12.2.5	Computer programs for log-linear model analysis	223
12.3	Application to ABC survey data	224
12.4	Summary	227
	References	228
13	CUB models: Statistical methods and empirical evidence	231
	<i>Maria Iannario and Domenico Piccolo</i>	
13.1	Introduction	231
13.2	Logical foundations and psychological motivations	233
13.3	A class of models for ordinal data	233
13.4	Main inferential issues	236

13.5	Specification of CUB models with subjects' covariates	238
13.6	Interpreting the role of covariates	240
13.7	A more general sampling framework	241
13.7.1	Objects' covariates	241
13.7.2	Contextual covariates	243
13.8	Applications of CUB models	244
13.8.1	Models for the ABC annual customer satisfaction survey	245
13.8.2	Students' satisfaction with a university orientation service	246
13.9	Further generalizations	248
13.10	Concluding remarks	251
	Acknowledgements	251
	References	251
	Appendix	255
	A program in R for CUB models	255
A.1	Main structure of the program	255
A.2	Inference on CUB models	255
A.3	Output of CUB models estimation program	256
A.4	Visualization of several CUB models in the parameter space	257
A.5	Inference on CUB models in a multi-object framework	257
A.6	Advanced software support for CUB models	258
14	The Rasch model	259
	<i>Francesca De Battisti, Giovanna Nicolini and Silvia Salini</i>	
14.1	An overview of the Rasch model	259
14.1.1	The origins and the properties of the model	259
14.1.2	Rasch model for hierarchical and longitudinal data	263
14.1.3	Rasch model applications in customer satisfaction surveys	265
14.2	The Rasch model in practice	267
14.2.1	Single model	267
14.2.2	Overall model	268
14.2.3	Dimension model	272
14.3	Rasch model software	277
14.4	Summary	278
	References	279
15	Tree-based methods and decision trees	283
	<i>Giuliano Galimberti and Gabriele Soffritti</i>	
15.1	An overview of tree-based methods and decision trees	283
15.1.1	The origins of tree-based methods	283
15.1.2	Tree graphs, tree-based methods and decision trees	284
15.1.3	CART	287
15.1.4	CHAID	293
15.1.5	PARTY	295
15.1.6	A comparison of CART, CHAID and PARTY	297
15.1.7	Missing values	297
15.1.8	Tree-based methods for applications in customer satisfaction surveys	298

15.2	Tree-based methods and decision trees in practice	300
15.2.1	ABC ACSS data analysis with tree-based methods	300
15.2.2	Packages and software implementing tree-based methods	303
15.3	Further developments	304
	References	304
16	PLS models	309
	<i>Giuseppe Boari and Gabriele Cantaluppi</i>	
16.1	Introduction	309
16.2	The general formulation of a structural equation model	310
16.2.1	The inner model	310
16.2.2	The outer model	312
16.3	The PLS algorithm	313
16.4	Statistical interpretation of PLS	319
16.5	Geometrical interpretation of PLS	320
16.6	Comparison of the properties of PLS and LISREL procedures	321
16.7	Available software for PLS estimation	323
16.8	Application to real data: Customer satisfaction analysis	323
	References	329
17	Nonlinear principal component analysis	333
	<i>Pier Alda Ferrari and Alessandro Barbiero</i>	
17.1	Introduction	333
17.2	Homogeneity analysis and nonlinear principal component analysis	334
17.2.1	Homogeneity analysis	334
17.2.2	Nonlinear principal component analysis	336
17.3	Analysis of customer satisfaction	338
17.3.1	The setting up of indicator	338
17.3.2	Additional analysis	340
17.4	Dealing with missing data	340
17.5	Nonlinear principal component analysis versus two competitors	343
17.6	Application to the ABC ACSS data	344
17.6.1	Data preparation	344
17.6.2	The <i>homals</i> package	345
17.6.3	Analysis on the ‘complete subset’	346
17.6.4	Comparison of NLPCA with PCA and Rasch analysis	350
17.6.5	Analysis of ‘entire data set’ for the comparison of missing data treatments	352
17.7	Summary	355
	References	355
18	Multidimensional scaling	357
	<i>Nadia Solaro</i>	
18.1	An overview of multidimensional scaling techniques	357
18.1.1	The origins of MDS models	358
18.1.2	MDS input data	359

18.1.3	MDS models	362
18.1.4	Assessing the goodness of MDS solutions	369
18.1.5	Comparing two MDS solutions: Procrustes analysis	371
18.1.6	Robustness issues in the MDS framework	371
18.1.7	Handling missing values in MDS framework	373
18.1.8	MDS applications in customer satisfaction surveys	373
18.2	Multidimensional scaling in practice	374
18.2.1	Data sets analysed	375
18.2.2	MDS analyses of overall satisfaction with a set of ABC features: The complete data set	375
18.2.3	Weighting objects or items	381
18.2.4	Robustness analysis with the forward search	382
18.2.5	MDS analyses of overall satisfaction with a set of ABC features: The incomplete data set	383
18.2.6	Package and software for MDS methods	384
18.3	Multidimensional scaling in a future perspective	386
18.4	Summary	386
	References	387
19	Multilevel models for ordinal data	391
	<i>Leonardo Grilli and Carla Rampichini</i>	
19.1	Ordinal variables	391
19.2	Standard models for ordinal data	393
19.2.1	Cumulative models	394
19.2.2	Other models	395
19.3	Multilevel models for ordinal data	395
19.3.1	Representation as an underlying linear model with thresholds	396
19.3.2	Marginal versus conditional effects	397
19.3.3	Summarizing the cluster-level unobserved heterogeneity	397
19.3.4	Consequences of adding a covariate	398
19.3.5	Predicted probabilities	399
19.3.6	Cluster-level covariates and contextual effects	399
19.3.7	Estimation of model parameters	400
19.3.8	Inference on model parameters	401
19.3.9	Prediction of random effects	402
19.3.10	Software	403
19.4	Multilevel models for ordinal data in practice: An application to student ratings	404
	References	408
20	Quality standards and control charts applied to customer surveys	413
	<i>Ron S. Kenett, Laura Deldossi and Diego Zappa</i>	
20.1	Quality standards and customer satisfaction	413
20.2	ISO 10004 guidelines for monitoring and measuring customer satisfaction	414
20.3	Control Charts and ISO 7870	417

20.4	Control charts and customer surveys: Standard assumptions	420
20.4.1	Introduction	420
20.4.2	Standard control charts	420
20.5	Control charts and customer surveys: Non-standard methods	426
20.5.1	Weights on counts: Another application of the c chart	426
20.5.2	The χ^2 chart	427
20.5.3	Sequential probability ratio tests	428
20.5.4	Control chart over items: A non-standard application of SPC methods	429
20.5.5	Bayesian control chart for attributes: A modern application of SPC methods	432
20.5.6	Control chart for correlated Poisson counts: When things become fairly complicated	433
20.6	The M -test for assessing sample representation	433
20.7	Summary	435
	References	436
21	Fuzzy Methods and Satisfaction Indices	439
	<i>Sergio Zani, Maria Adele Milioli and Isabella Morlini</i>	
21.1	Introduction	439
21.2	Basic definitions and operations	440
21.3	Fuzzy numbers	441
21.4	A criterion for fuzzy transformation of variables	443
21.5	Aggregation and weighting of variables	445
21.6	Application to the ABC customer satisfaction survey data	446
21.6.1	The input matrices	446
21.6.2	Main results	448
21.7	Summary	453
	References	455
Appendix	An introduction to R	457
	<i>Stefano Maria Iacus</i>	
A.1	Introduction	457
A.2	How to obtain R	457
A.3	Type rather than ‘point and click’	458
A.3.1	The workspace	458
A.3.2	Graphics	458
A.3.3	Getting help	459
A.3.4	Installing packages	459
A.4	Objects	460
A.4.1	Assignments	460
A.4.2	Basic object types	462
A.4.3	Accessing objects and subsetting	466
A.4.4	Coercion between data types	469
A.5	S4 objects	470
A.6	Functions	472

A.7	Vectorization	473
A.8	Importing data from different sources	475
A.9	Interacting with databases	476
A.10	Simple graphics manipulation	477
A.11	Basic analysis of the ABC data	481
A.12	About this document	496
A.13	Bibliographical notes	496
	References	496

Index		499
--------------	--	------------

Foreword

The key to a successful business is understanding one's customers – knowing what they want and, more generally, how satisfied they are with one's product or service. This book describes how to construct such a key, and how to use it to unlock the door to understanding one's customers. It describes insights, introduces methodology, and outlines the cutting-edge statistical tools necessary for effective customer management.

Customer satisfaction is complementary to actual purchase patterns. The latter will show what the customers *buy*, but the former can show what they *would like to buy*. And a company which does not think ahead, to discover and provide what the customers would like, is a company destined for oblivion. Moreover, while traditionally considerable emphasis has been placed on customer recruitment, recent years have seen a shift towards putting more emphasis on the importance of customer retention. Retaining customers, reinforcing loyalty by providing them with what they want, is cheaper than recruiting new customers. One also has more information on existing customers than on those who 'might' become customers, which means that one can construct much more accurate models of how the customers behave and of what they want.

'Satisfaction' is a perfect example, like 'happiness' and 'well-being', of the sort of concept which triggers debates about the feasibility of measurement in the social sciences. Such attributes defy the traditional hard science notions of 'representational' measurement, in which a formal mapping is established from the system being studied to a mathematical model. Instead, one adopts so-called 'pragmatic' measurement, in which the precise property being measured is defined simultaneously with the procedure for measuring it. Since such concepts can be defined in multiple ways, one has to be very explicit about what exactly one means by 'satisfaction'. The measurement of satisfaction also illustrates some of the other complications with which measurement in the social sciences has to contend, including such things as the dependence on context, how questions are worded, the reactive nature of human beings, and even the time of day. Worse, intrinsic in its very definition is that fact that satisfaction is subjective. There is no concrete property, existing out there in the 'real world', which we are trying to measure. All of these complications add up to make measuring customer satisfaction a difficult and challenging problem – which this book ably and comprehensively tackles.

Moreover, we are now faced with multiple ways in which data can be – and indeed is – collected: the traditional face-to-face interview and postal surveys have been joined by telephone surveys and web surveys of various kinds, as well as comprehensive monitoring of customer behaviour via such things as loyalty cards. These new technologies present exciting new opportunities, but also new challenges: selection bias can be a particular challenge. All of this makes this book very timely. The breadth of statistical tools and models described in this book illustrates perfectly what a rich area this has become.

One of the particularly attractive features of this book is that it is about *practical* statistics. It does not simply describe statistical tools which can be applied in an idealised setting, but confronts the issues which all practicing statisticians have to face – of missing data, outliers, measurement error, selection bias, differing definitions, outdated data, and so on. Its treatment of such issues is cutting-edge. A second particularly attractive feature is its comprehensiveness: it covers all aspects, from basic data sources to high-level statistical modelling. Put together, these two attractive features make it a genuinely valuable book for all researchers concerned with measuring customer satisfaction, be they academic or commercial. Indeed, one might even say, more broadly, that it could serve as a valuable resource for statisticians who work in customer-driven applications in general, and who intend their work to make a difference.

David J. Hand
Professor of Statistics, Imperial College, London
Past President, Royal Statistical Society

Preface

All over the world, organizations are focusing on retaining existing customers while targeting potential customers. Measuring customer satisfaction provides critical information on how an organization is providing products or services to the marketplace. In order to measure customer satisfaction, survey questionnaires are used, in which respondents are asked to express their degree of satisfaction with regard to multiple aspects of the product or service. Statistical analysis of data from these surveys is carried out and measures of various aspects and overall satisfaction are computed. This data is, however, non-trivial to handle because of the subjective nature of the observed variables.

First of all, as described by Ferrari and Manzi (*Quality Technology and Quantitative Management*, Vol. 7, No. 2, pp. 117–133, 2010), the relevance correct weighing of the variables that determine the level of satisfaction are unknown. In addition, these variables often have an ordinal measurement scale which needs to be suitably dealt with. Moreover, the level of satisfaction is generally dependent on both expectations and individual characteristics of respondents as well as on contextual variables. Surveys also contain measurement errors caused by the subjective nature of the variables and by cognitive dissonance that can affect data, with undesired consequences on the reliability of the results. With the objective of handling, or at least to controlling, some of these problems, many different methods to assess customer satisfaction have been proposed in the literature. *Modern Analysis of Customer Satisfaction Surveys* presents, in two parts and an appendix, basic and advanced methods for the analysis of customer satisfaction survey data. We have gathered together contributions from world-class authors who are leading the field with new techniques and innovative ideas.

Most of the chapters include an application of various techniques to a standard set of data collected from 266 companies (customers) participating in the ABC Annual Customer Satisfaction Survey (ACSS). The data refers to a questionnaire consisting of 81 questions. The ABC Company is a typical global supplier of integrated software, hardware, and service solutions to media and telecommunications service providers. The ACSS questionnaire was designed to provide feedback on all company touch points and interactions with customers. It covers topics such as equipment, sales support, technical support, training, supplies, software solutions, customer website, purchasing support, contracts and pricing and system installation. Information on customers includes: country, industry segment, age of ABC's equipment, profitability, customer seniority and position of respondent.

The first part of the book consists of nine chapters. The first two chapters introduce the basics of customer satisfaction survey data analysis. The remaining chapters cover topics such as: sampling, surveys and census; measurement scales; integrated analysis; web surveys;

customer satisfaction and customer loyalty; missing data and imputation methods; and outliers and robustness for ordinal data.

The second part of the book consists of 12 chapters with new and innovative techniques and models. They include: causality models; Bayesian networks; log-linear models; CUB models; Rasch models; decision trees; partial least squares models; nonlinear principal component analysis; multidimensional scaling; multilevel models for ordinal data; control charts applied to customer surveys; and fuzzy methods.

Our interest in statistical analysis of customer satisfaction data has followed different paths. Ron has worked on real problems of customer satisfaction for several decades, doing consulting and academic research on methods designed to generate high-impact conclusions and recommendations. Silvia first dealt with the topic during her studies at the Department of Statistics of the Catholic University of Milan. She then dealt with the evaluation of the quality of education services on multiple projects in her current department at the Faculty of Political Science of the University of Milan. Many of her teachers and supervisors, from the Catholic University of Milan and the University of Milan, have contributed to this book as authors or as reviewers.

The seed for this book was planted on 22 July 2005 when a project to investigate the application of non-standard techniques to survey data was launched at the Department of Statistics and Applied Mathematics of the University of Torino under the auspices of the Diego de Castro Foundation and the strong support of Professor Roberto Corradetti, the head of the department and the foundation general secretary. The late Professor Diego de Castro combined an interest in the development of statistical methodology and its application to business and industrial problems. We believe that this work is well in line with his many contributions to statistics, demography, economics and jurisprudence. Professor Corradetti showed vision and constancy of purpose in making sure that the project we started in 2005 matured and produced important publications. Some of these publications appeared in two special issues of the international journal *Quality Technology and Quantitative Management*.

One of the key participants in the first meetings of this project was Professor Ennio Isaia from the Department of Statistics and Applied Mathematics of the University of Torino, who unfortunately passed away in February 2010. Ennio was a gifted statistician with a unique capability of combining theoretical knowledge, interest in applied problems and a great ability to program in R and other software languages. In a sense, this book is designed to have similar characteristics.

In preparing this book we have enlisted the participation of external reviewers who gave us invaluable feedback. They all deserve sincere thanks. The reviewers were: Anthony Atkinson, Mojca Bavdaž, Maurizio Carpita, Enrico Ciavolino, Benito Vittorio Frosini, Salvatore Ingrassia, Patrik Mair, Monica Pratesi, Matteo Pessione, Susanne Raessler, Christensen Rune Humbo, Marco Scrutari, Roberta Siciliano, Pierpaolo d'Urso, Richard Paul Waterman and Silvia Figini. We are also extremely grateful to Professor David Hand for taking the time to write a foreword that ably summarizes the aims of this book and the challenges of the topics covered. Readers should begin with this foreword which is where the book really starts.

The book provides background for both basic and advanced methods for analysing customer satisfaction survey data. The book chapters are supported by R applications so that practitioners can actually implement the methods covered in the book. The book's website

(www.wiley.com/go/modern_analysis) provides the data sets and R scripts used in the book.

The authors who contributed have helped us present the state of the art in the analysis of customer satisfaction surveys. We thank them for their work and professionalism. While we retain responsibility for any typos and mistakes that may have sneaked in, the chapter authors deserve full credit.

Ron S. Kenett
KPA Ltd., Israel and University of Turin, Italy

Silvia Salini
University of Milan, Italy

Contributors

Alessandro Barbiero Post-doctoral fellow in Statistics, Department of Economics, Business and Statistics, University of Milan, Italy.

Silvia Biffignandi Professor of Business and Economic Statistics, University of Bergamo, Italy.

Giuseppe Boari Professor of Statistics, Faculty of Economics, Università Cattolica del Sacro Cuore, Milan, Italy.

Andrea Bonanomi Assistant Professor of Statistics, Department of Statistics Sciences, Università Cattolica del Sacro Cuore, Milan, Italy.

Gabriele Cantaluppi Associate Professor of Statistics, Faculty of Economics, Università Cattolica del Sacro Cuore, Milan, Italy.

Luciana Dalla Valle Lecturer in Statistics, School of Computing and Mathematics, University of Plymouth, UK.

Francesca De Battisti Assistant Professor of Statistics, Department of Economics, Business and Statistics, University of Milan, Italy.

Laura Deldossi Associate Professor of Statistics, Department of Statistical Sciences, Catholic University of Milan, Italy.

Pier Alda Ferrari Professor of Statistics, University of Milan, Italy.

Stephen E. Fienberg Maurice Falk University Professor of Statistics and Social Science, Carnegie Mellon University, Pittsburgh, USA.

Roberto Furlan Senior methodologist, Kantar Health, Epsom, UK.

Iddo Gal Senior Lecturer, Department of Human Services, University of Haifa, Israel.

Giuliano Galimberti Assistant Professor of Statistics, Department of Statistical Sciences, University of Bologna, Italy.

Leonardo Grilli Associate Professor of Statistics, School of Economics, University of Florence, Italy.

Stefano Maria Iacus Associate Professor, Department of Economics, Business and Statistics, University of Milan, Italy.

Maria Iannario Assistant Professor of Statistics, Department of Theory and Methods for Human and Social Sciences – Statistical Sciences Unit, University of Naples Federico II, Italy.

Ron S. Kenett Chairman and Chief Executive Officer, KPA Ltd., Raanana, Israel; Research Professor, University of Turin, Italy; and international Research Professor at the Center for Risk Engineering of NYU Poly, New York.

Daniel Manrique-Vallier Post-doctoral fellow, Social Science Research Institute and Department of Statistical Science, Duke University, Durham, North Carolina, USA.

Diego Martone President and chief executive, Demia Studio Associato, Trieste, Italy; and contract professor, Department of Psychology ‘Gaetano Kanizsa’, University of Trieste, Italy.

Alessandra Mattei Assistant Professor of Statistics, Department of Statistics ‘G. Parenti’, University of Florence, Italy.

Fabrizia Mealli Professor of Statistics, Department of Statistics ‘G. Parenti’, University of Florence, Italy.

Maria Adele Milioli Associate Professor of Statistics, Department of Economics, University of Parma, Italy.

Isabella Morlini Associate Professor of Statistics, Department of Economics, University of Modena and Reggio Emilia, Italy.

Giovanna Nicolini Professor of Statistics, University of Milan, Italy.

Irena Ograjensek Associate Professor of Statistics, Faculty of Economics, University of Ljubljana, Slovenia.

Barbara Pacini Associate Professor of Statistics, Department of Statistics and Applied Mathematics, University of Pisa, Italy.

Giovanni Perucca Research assistant, Polytechnic University of Milan, Italy.

Domenico Piccolo Professor of Statistics, Department of Theory and Methods of Human and Social Sciences, Statistical Sciences Unit, University of Naples Federico II, Italy.

Carla Rampichini Professor of Statistics, Department of Statistics, University of Florence, Italy.

Marco Riani Professor of Statistics, University of Parma, Italy.

Donald B. Rubin John L. Loeb Professor of Statistics, Department of Statistics, Harvard University, USA.

Silvia Salini Assistant Professor of Statistics, Department of Economics, Business and Statistics, University of Milan, Italy.

Gabriele Soffritti Associate Professor of Statistics, Department of Statistical Sciences, University of Bologna, Italy.

Nadia Solaro Assistant Professor of Statistics, Department of Statistics, University of Milano-Bicocca, Milan, Italy.

Francesca Torti Division of Statistics, Faculty of Economics, University of Parma, Italy.

Sergio Zani Professor of Statistics, University of Parma, Parma, Italy.

Diego Zappa Associate Professor of Statistics, Department of Statistical Sciences, Catholic University, Milan, Italy.

Part I

BASIC ASPECTS OF CUSTOMER SATISFACTION SURVEY DATA ANALYSIS

1

Standards and classical techniques in data analysis of customer satisfaction surveys

Silvia Salini and Ron S. Kenett

Customer satisfaction studies are concerned with the level of satisfaction of customers, consumers and users with a product or service. Customer satisfaction is defined as ‘The degree of satisfaction provided by the goods or services of a company as measured by the number of repeat customers’ (www.businessdictionary.com). Customer satisfaction therefore seems to be an objective and easily measured quantity. However, unlike variables such as revenues, type of product purchased or customer geographical location, customer satisfaction is not necessarily observed directly. Typically, in a social science context, analysis of such measures is done indirectly by employing proxy variables. Unobserved variables are referred to as *latent variables*, whilst proxy variables are known as *observed variables*. In many cases, the latent variables are very complex and the choice of suitable proxy variables is not immediately obvious. For example, in order to assess customer satisfaction with an airline service, it is necessary to identify attributes that characterize this type of service. A general framework for assessing airlines includes attributes such as on-board service, timeliness, responsiveness of personnel, seating and other tangible service characteristics. In general, some attributes are objective, related to the service’s technical characteristics, and others are subjective, dealing with behaviours, feelings and psychological benefits. In order to design a survey questionnaire, a set of observed variables must be identified.

In practice, many of the customer satisfaction surveys conducted by business and industry are analysed in a very simple way, without using models or statistical methods. Typical reports include descriptive statistics and basic graphical displays. As shown in this book, integrating a basic analysis with more advanced tools, provides insights into non-obvious patterns and

important relationships between the survey variables. This knowledge can significantly affect findings and recommendations derived from a survey.

After presenting classical customer satisfaction methodologies, this chapter provides a general introduction to customer satisfaction surveys, within an organization's business cycle. It then presents standards used in the analysis of survey data. Next it gives an overview on the techniques commonly used to measure customer satisfaction, along with their problems and limitations. Finally, it gives a preview and general introduction to the rest of the chapters in this book.

1.1 Literature on customer satisfaction surveys

Survey questionnaire design, data collection approaches, validation of questionnaires, sampling problems, descriptive statistics and classical statistical inference techniques are covered in many books and papers. This book presents such topics, but also provides a large range of modern and non-standard techniques for customer satisfaction data analysis. Moreover, these various techniques are compared by applications to a common benchmark data set, the ABC 2010 annual customer satisfaction survey (ACSS). For details on the benchmark data set and the ABC company, see Chapter 2.

A non-exhaustive list of relatively advanced books dealing with customer satisfaction data analysis includes Grigoroudis and Siskos (2010), Jacka and Keller (2009), Hayes (2008), Allen and Rao (2000), Johnson and Gustafsson (2000), Vavra (1997) and Biemer and Lyberg (2003). Grigoroudis and Siskos (2010) describe service quality models and the Multicriteria Satisfaction Analysis (MUSA), with examples of satisfaction barometers. Hayes (2008) gives special attention to reliability and validity of questionnaires with a link to customer loyalty. The book by Allen and Rao (2000) is most comprehensive in terms of statistical methods. Although not written by statisticians, it provides a useful and well-written description of techniques of descriptive analysis of univariate, bivariate and multivariate data; it also describes dependent models (linear and logistic regression), explanatory techniques (factor analysis, principal component analysis), causal models (path analysis), and structural equation models. Appendix C of Johnson and Gustafsson (2000) presents an interesting comparison of alternative data analysis methods, in particular considering (1) gap analysis, (2) multiple regression, (3) correlation, (4) principal component regression and (5) partial least squares (PLS). Vavra (1997) covers theories of customer satisfaction and loyalty with several examples of scales, analytic procedures and best practices. Biemer and Lyberg (2003) provide a comprehensive treatment of classical design and analysis of sample surveys.

This book is focused on statistical models for modern customer satisfaction survey data analysis. It addresses modern topics such as web surveys and state-of-the-art statistical models such as the CUB model and Bayesian networks (BN). The book chapters, written by leading researchers in the field, use practical examples in order to make their content also accessible to non-statisticians. Our ultimate goal is to advance the application of best practices in the analysis of customer satisfaction survey data analysis and stimulate new research in this area. As stated in the book foreword by Professor David Hand, we aim to make a difference.

1.2 Customer satisfaction surveys and the business cycle

Statistical analysis is a science that relies on a transformation of reality into dimensions that lend themselves to quantitative analysis. Self-administered surveys use structured questioning