Many texts are excellent sources of knowledge about individual statistical tools (general linear models, survival analysis, missing data, etc.), but the art of data analysis is about choosing and using multiple tools. Instead of presenting isolated techniques, this text emphasizes problem solving strategies that address the many issues arising when developing multivariable models using real data and not standard textbook examples. It includes imputation methods for dealing with missing data effectively, methods for dealing with nonlinear relationships and for making the estimation of transformations a formal part of the modeling process, methods for dealing with "too many variables to analyze and not enough observations," and powerful model validation techniques based on the bootstrap. This text realistically deals with model uncertainty, and its effects on inference, to achieve "safe data mining." It also presents many graphical methods for communicating complex regression models to non-statisticians.

REGRESSION MODELING STRATEGIES presents full scale case studies of non-trivial datasets instead of over-simplified illustrations of each method. These case studies use freely available high-level S-PLUS and R functions that make the multiple imputation, model building, validation, and interpretation tasks described in the book relatively easy to do. Most of the methods in this text apply to all regression models, but special emphasis is given to some of the most popular ones: the linear multiple regression model based on ordinary least squares, the binary logistic model, two logistic models for ordinal responses, parametric survival regression models, and the Cox semi-parametric survival model.

This text is intended for Masters' or PhD level graduate students who have had a general introductory probability and statistics course and who are well versed in ordinary multiple regression and algebra. The book is also intended to serve as a reference for data analysts and statistical methodologists, as it contains an up-to-date survey and bibliography of modern statistical modeling techniques.

The author, Frank E Harrell Jr, is Professor of Biostatistics and Chair, Department of Biostatistics, Vanderbilt University School of Medicine, Nashville. He received his Ph.D. in biostatistics from the University of North Carolina, Chapel Hill. He has published numerous predictive models and articles on applied statistics, medical research, and clinical trials. He is an Associate Editor of *Statistics in Medicine*, an editorial consultant for the *Journal of Clinical Epidemiology*, a member of the editorial board of the *American Heart Journal*, a Fellow of the American Statistical Association, and a consultant to the US Food and Drug Administration and to the pharmaceutical industry. He teaches graduate level courses in biostatistical modeling, statistical computing and graphics, and advanced data analysis.