

Programs and Data to Accompany "Modeling Techniques in Predictive Analytics" Miller (2014)

Note that many chapter programs contain library commands for bringing in R functions included in packages. To run these programs, the user needs to first install the packages in his/her R environment. Programs were tested under R 3.0.0 on Mac OS 10.6.8.

Book Location	Description of File and Notes	File Name
Chapter 1	R Program for the Anscombe Quartet (program)	chapter_1_program.R
Chapter 2	Shaking Our Bobbleheads Yes and No (data) Shaking Our Bobbleheads Yes and No (program)	dodgers.csv chapter_2_program.R
Chapter 3	Measuring and Modeling Individual Preferences (data) Measuring and Modeling Individual Preferences (program) [Preliminary. Run utility program appendix_c1_program.R to obtain the spine chart plotting function in the binary R file: mtpa_spine_chart.Rdata]	mobile_services_ranking.csv chapter_3_program.R
Chapter 4	Market Basket Analysis of Grocery Store Data (program)	chapter_4_program.R
Chapter 5	Working with Economic Data (program) [Data are gathered from the Web for this program, so the user should ensure that there is an active Internet/Web connection before running the program. Also, modification of date ranges may be needed in order to bring the program up to date for the data being gathered.]	chapter_5_program.R
Chapter 6	Call Center Scheduling Problem and Solution (shift data) Call Center Scheduling Problem and Solution (call center data) Call Center Scheduling Problem and Solution (program) [Preliminary. Run utility program appendix_c3_program.R to obtain the split-plotting utility functions in the binary R file: mtpa_split_plotting_utilities.Rdata Also run the utility program appendix_c4_program.R to obtain the wait-time ribbon utility functions for ggplot2 work: mtpa_wait_time_ribbon_utility.Rdata]	data_anonymous_bank_shifts.csv data_anonymous_bank_february.txt chapter_6_program.R

Chapter 7	<p>Text Analytics of Movie Taglines (data)</p> <p>Text Analytics Book R Code (data for world cloud)</p> <p>Text Analytics of Movie Taglines (program) [Preliminary. Run utility program <code>appendix_c3_program.R</code> to obtain the split-plotting utility functions in the binary R file: <code>mtpa_split_plotting_utilities.Rdata</code>]</p>	<p><code>taglines_copy_data.txt</code></p> <p><code>MTPA_R_code.txt</code></p> <p> </p> <p><code>chapter_7_program.R</code></p>
Chapter 8	<p>Sentiment Analysis and Classification of Movie Ratings (Hu and Liu negative word list)</p> <p>Sentiment Analysis and Classification of Movie Ratings (Hu and Liu positive word list)</p> <p>Sentiment Analysis and Classification of Movie Ratings (directories and subdirectories of text files under the directory reviews)</p> <p> </p> <p>Sentiment Analysis and Classification of Movie Ratings (program) [Preliminary. Take the code from Appendix C under the file name <code>appendix_C5_program.R</code> is saved in this directory as well. This code will be sourced into <code>chapter_9_program.R</code> a number of times during execution.</p> <p>Word Scoring Code for Sentiment Analysis (program, same as in Appendix C)</p>	<p><code>Hu_Liu_negative_word_list.txt</code></p> <p> </p> <p><code>Hu_Liu_positive_word_list.txt</code></p> <p> </p> <p><code>train/unsup</code></p> <p><code>train/pos</code></p> <p><code>train/neg</code></p> <p><code>test/pos</code></p> <p><code>test/neg</code></p> <p><code>test/tom</code></p> <p> </p> <p><code>chapter_8_program.R</code></p> <p> </p> <p><code>appendix_c5_program.R</code></p>
Chapter 9	<p>Winning Probabilities by Simulation (Negative Binomial Model) (program)</p>	<p><code>chapter_9_program.R</code></p>

Chapter 10	Computer Choice Study: Training and Testing with Hierarchical Bayes (data)	computer_choice_study.csv
	Computer Choice Study: Training and Testing with Hierarchical Bayes (program) [Preliminary. Run utility program appendix_c2_program.R to obtain the simulation utility functions in the binary R file: mtpa_market_simulation_utilities.Rdata]	chapter_10a_program.R
	Preference, Choice, and Market Simulation (program) [Preliminary. Run utility program appendix_c2_program.R to obtain the simulation utility functions in the binary R file: mtpa_market_simulation_utilities.Rdata Also run the utility program appendix_c3_program.R to obtain the split-plotting utility functions for ggplot2 work: mtpa_split_plotting_utilities.Rdata]	chapter_10b_program.R
Chapter 11	California Housing Values: Regression and Spatial Regression Models (data)	houses_data.txt
	California Housing Values: Regression and Spatial Regression Models (program)	chapter_11_program.R
Appendix C	Conjoint Analysis Spine Chart (program)	appendix_c1_program.R
	Market Simulation Utilities (program)	appendix_c2_program.R
	Split-plotting Utilities (program)	appendix_c3_program.R
	Wait-time Ribbon Plot (program)	appendix_c4_program.R
	Word Scoring Code for Sentiment Analysis (program)	appendix_c5_program.R
	Utilities for Spatial Data Analysis (program)	appendix_c6_program.R