Developments in Business Simulation and Experiential Learning, Volume 31, 2004

# ONLINE SALES FORECASTING WITH THE MULTIPLE REGRESSION ANALYSIS DATA MATRICES PACKAGE

**Aspy Palia**
**University of Hawaii at Manoa**
**aspy@cba.hawaii.edu**

## ABSTRACT

*The Web-based Multiple Regression Analysis Data Matrices Package (developed jointly with Justin Yost) enables competing participant teams in the marketing simulation COMPETE to apply their knowledge of multiple regression analysis in sales forecasting. Participants with Web-access use this package to create nine data matrices (one data matrix for each strategic business unit) consisting of relevant predictor and response variables for each of the prior decision periods. Next, the data are screened for potential multicollinearity using correlation analysis. Then, the top two predictor variables that satisfy multiple regression analysis assumptions are used to build a linear unrestricted single-equation multiple regression model. The results are checked for potential heteroskedasticity.*

## INTRODUCTION

The Multiple Regression Analysis Data Matrices Package is an online package that enables competing participant teams in the marketing simulation COMPETE to apply their knowledge of multiple regression analysis (Freund & Williams, 1977; Lilien & Rangaswamy, 2003; Neter & Wasserman, 1974; Pfaffenberger & Patterson, 1977; Tatsuoka, 1971) in sales forecasting (Enrick, 1969; Makridakis, Wheelwright, & McGee, 1983; Willis, 1987). Participants with Web-access can download their prior decision period results, and use this package to automatically create a data matrix for each of their nine strategic business units (SBUs). These data matrices are screened for potential multicollinearity among the predictor variables using correlation analysis. Then, the top two predictor variables that satisfy multiple regression analysis assumptions are used to build a sales forecasting model for each of the 9 SBUs, to predict sales.

The primary purpose of this paper is to present this new user-centered learning tool that helps to prepare students for sales forecasting and marketing decision-making responsibilities in their future careers. The objective is to provide participant teams the opportunity (1) to plan, implement, and control a marketing program for their products and (2) to apply sales forecasting and multiple regression analysis to forecast sales in a dynamic, complex and uncertain simulated competitive environment.

## SALES FORECASTING

A major responsibility of marketing is the preparation of sales forecasts. First, market opportunities are identified through marketing research. Then, the size, growth and profitability of each market opportunity are measured and/or forecasted. Sales forecasts are used (a) by finance to raise the needed cash for investment and operations, (b) by manufacturing to establish capacity and output levels, (c) by purchasing to acquire the necessary supplies, and (d) by human resources to hire the needed workers (Kotler 2003).

Accurate sales forecasts facilitate effective and efficient allocation of scarce resources. Over-estimates of demand lead to several problems. First, excess inventory uses up valuable shelf space and leads to obsolescence. Next, scarce working capital blocked up in inventory carrying charges [funds used or borrowed (a) by manufacturers to produce goods, or (b) by retailers to purchase goods] cannot be used for other purposes such as R&D or promotional expenses. Third, storage charges are incurred to store excess inventory in public or private warehouses. Finally, margins are reduced when excess inventory is removed through end-of-year clearance sales.

Under-estimates of demand lead to a different set of problems. First, stock-outs lead to wasted shelf space. Next, insufficient inventory leads to lost sales and consequent lost margins. Third, failure to keep up with customer demand may necessitate the use of limited and expensive overtime production leading to lower profitability. Finally, and most importantly, the firm may lose customers, when prospects facing an empty store shelf, try an alternative brand or go to an alternative store, and are satisfied by the competitive offering. Given the detrimental impact of inaccurate forecasts, marketers use a variety of sales forecasting techniques in order to forecast sales accurately.

## SALES FORECASTING TECHNIQUES

Marketers forecast sales by (a) extending past behavior, and/or by (b) predicting future behavior. Extending past behavior techniques are applied when past sales data are available. These techniques tend to be more quantitative. They extend past data into the future, and assume that the future will be like the past. These techniques include trend extension, the factor method (using one or more factors such as the Buying Power Index, and the SIC code to forecast the sales of industrial products), time series analysis, the use of leading series, and indices such as the consumer price index, producer price index, and the index of leading economic indicators (McCarthy and

Perreault 1987).

Predicting future behavior techniques do not rely on the availability of past data. Instead, they rely on judgment, and are used when there are changing conditions in the marketplace or changes in the marketing mix used. They are also used to forecast the sale of unstable (fashion) goods and new products. These techniques include the Jury of Executive Opinion, salespeople's estimates, surveys of final buyers, retailers and/or wholesalers, panels of stores and/or final consumers, market tests of existing products, test markets of new products, the substitute method, and needs analysis (Perreault and McCarthy 1996). There is no best method of forecasting in all circumstances. Confidence in the accuracy of sales forecasts is derived by corroborating the results using two or more methods (McCarthy and Perreault 1984).

## SALES FORECAST MODEL BUILDING

A model is an explanation of causal relationships among a set of factual phenomena during a specified time period. Competing participant teams in the COMPETE marketing simulation build linear unrestricted single-equation multiple regression models in order to forecast the sales of each of their nine SBUs. First, the model is linear since the dependent variable sales as well as the predictor variables price, advertising, sales force, quality, average competitor price, average competitor advertising, average competitor sales force, average competitor quality and seasonal variation are all of the first order. Second, the model is unrestricted as there is no a priori theoretical reason to constrain the hyper-regression surface through the origin. Third, a single-equation model is used instead of a simultaneous equation model. Finally, the model is a multiple regression model since more than one predictor variable is used to predict sales.

The Multiple Regression Analysis Data Matrices Package is used to prepare nine data matrices (one for each of the nine strategic business units) for each participant team based on the performance output of the prior periods in the COMPETE simulation. Later, Microsoft EXCEL is used to analyze each strategic business unit (SBU) data matrix, build a linear unrestricted single-equation multiple regression model, and forecast sales.

The data (in each data matrix) are first screened for outliers. Next, correlation analysis is used to screen the predictor variables for potential multicollinearity, and to identify the top predictor variables that satisfy the assumptions underlying multiple regression analysis. Then, multiple regression analysis is used to build a sales forecasting model. The goodness-of-fit of the sales forecasting model is checked with such statistics as $R^2$, Adj. $R^2$, F-statistic with significance, and the Standard Error of Regression relative to the mean and standard deviation of the response variable sales. Later, the partial explanatory power of each predictor variable is checked for expected sign and significance. The error terms are scanned for potential heteroskedasticity (serial auto correlation of the error term) in order to satisfy the assumptions underlying the use of multiple regression analysis. Finally, the multiple regression sales forecasting model is used to forecast sales.

## THE MARKETING SIMULATION COMPETE

COMPETE (Faria, Nulsen, & Roussos, 1994) is a widely used marketing simulation designed to provide students with marketing strategy development and decision-making experience. Competing student teams are placed in a complex, dynamic, and uncertain environment. The participants experience the excitement and uncertainty of competitive events and are motivated to be active seekers of knowledge. They learn the need for and usefulness of mastering an underlying set of decision-making principles.

Competing student teams plan, implement, and control a marketing program for three high-tech products in three regions within the United States. The features and benefits of each product and the characteristics of consumers in each region are described in the student manual. Based on a marketing opportunity analysis, a mission statement is generated, specific and measurable company goals are set, and marketing strategies are formulated to achieve these goals. Constant monitoring and analysis of their own and competitive performance helps the teams better understand their markets and improve their decisions.

Each decision period (quarter), the competing teams make a total of 74 marketing decisions with regard to marketing their three brands in the three regional markets. These decisions include nine pricing decisions, nine shipment decisions, three sales force size decisions, nine sales force time allocation decisions, one sales force salary decision, one sales force commission decision, twenty-seven advertising media decisions, nine advertising content decisions, three quality-improvement R&D decisions, and three cost-reduction R&D decisions. Successful planning, implementation, and control of their respective marketing programs require that each company constantly monitor trends in its own and competitive decision variables and resulting performance.

## MULTIPLE REGRESSION ANALYSIS DATA MATRICES PACKAGE

The Web-based Multiple Regression Analysis Data Matrices Package is accessible online to competing participant teams in the marketing simulation COMPETE. Following the simulation run for each decision period, the competing participant teams log in to the COMPETE Online Decision Entry System (CODES) website (Palia and Mak 2001, Palia et al 2000). Their login is validated against a database of participating teams for each industry, and they have access to their decisions and printouts (results) for all prior decision periods.

First, they prepare a target folder "Regress" in the root directory of their computer hard drive. The path should read "C:\Regress". Next, they download and save the Excel version of the Results for periods 1 to 11 in the "C:\Regress" directory.

Then, they rename the downloaded results files as 1.xls, 2.xls, … where 1,2,… correspond to period numbers. Later, they download and save the Matrices.xls file to the same "C:\Regress" directory.

In order to generate the nine SBU data matrices automatically in Windows Explorer, they open the "Regress" folder on the C:\ drive, double-click the "Matrices.xls" file, and click OK approximately a dozen times on the dialog screens that appear in order to update the "Matrices.xls" workbook with their team's data. The tabs for each SBU Multiple Regression Data Matrix are at the bottom of the EXCEL screen.

Finally, they make a duplicate copy of the lower (unprotected) data matrix on each worksheet with the Copy – Paste Special – Select Paste Values option for subsequent (a) descriptive statistics, (b) correlation analysis, and (c) multiple regression analysis. Despite lack of observations at the end of competition (only 11 periods), and consequent low degrees of freedom (only 8 with two predictor variables using an unrestricted model), the multiple regression models generated, in some instances, have remarkably good fit and strong partial explanatory power of the predictor variables selected. The predictor variables are selected from the data matrix that consists of the independent variables used by the COMPETE engine. In addition, the underlying assumptions of multiple regression analysis are satisfied. Consequently, the forecasts, in some instances, prove to be remarkably accurate.

## CONCLUSION

The Online Multiple Regression Analysis Data Matrices provides competing participant teams in the COMPETE simulation with the opportunity to learn and apply multiple regression analysis in order to forecast the sales of their strategic business units accurately. This package facilitates the integration of computers, the Internet and the World Wide Web into the marketing curriculum.

## REFERENCES


Enrick, N. L. (1969) *Market and Sales Forecasting: A Quantitative Approach*, San Francisco, CA: Chandler.

Faria, A. J., Nulsen, Jr., R. O., & Roussos, D. S. (1994), *COMPETE: A Dynamic Marketing Simulation*, 4th ed. Burr Ridge, IL: Irwin.

Freund, J. E. & Williams, F. J. (1977), *Elementary Business Statistics: The Modern Approach*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall.

Kotler, P. (2003), *Marketing Management*, 11th ed. Upper Saddle River, NJ: Prentice-Hall.

Lilien, G. L. & Rangaswamy, A. (2003), *Marketing Engineering: Computer-Assisted Marketing Analysis and Planning*, 2nd ed. Upper Saddle River, NJ: Prentice-Hall.

Makridakis, S., Wheelwright, S. C., & McGee, V. E. (1983), *Forecasting: Methods and Applications*. New York, NY: Wiley.

McCarthy, E. J. & Perreault, Jr., W. D. (1984), *Basic Marketing*, 8th ed. Homewood, IL: Irwin.

_____ and _____ (1987), *Basic Marketing*, 9th ed. Homewood, IL: Irwin.

Neter, J. & Wasserman, W. (1974), *Applied Linear Statistical Models*, Homewood, IL: Irwin.

Palia, A. P., Mak W. K., & Roussos, D. S. (2000), "Facilitating Learning in the New Millennium With The COMPETE Online Decision Entry System (CODES)," in: Page, D. & Snyder, L. T. eds. *Developments in Business Simulation and Experiential Learning*, Vol. 27, Proceedings of the Twenty-seventh Annual Conference of the Association for Business Simulation and Experiential Learning, pp. 248-249.

_____ & Mak W. K. (2001), "An Online Evaluation of The COMPETE Online Decision Entry System (CODES)," in: Pittenger, K. S. & Vaughan, M. J., eds. *Developments in Business Simulation and Experiential Learning*, Vol 28, Proceedings of the Twenty-eighth Annual Conference for Business Simulation and Experiential Learning.

Perreault, Jr., W. D., Jr. & McCarthy, E. J. (1996), *Basic Marketing: A Global-Managerial Approach*, Chicago, IL: Irwin.

Pfaffenberger, R. C. & Patterson, J. H. (1977), *Statistical Methods: For Business and Economics,* Homewood, IL: Irwin.

Tatsuoka, M. M. (1971), *Multivariate Analysis: Techniques for Educational and Psychological Research,* New York, NY: Wiley.

Willis, R. E. (1987), *A Guide to Forecasting for Planners and Managers,* Englewood Cliffs, NJ: Prentice-Hall.