

Assignment 2 – Regression Models

Submission Guidelines

Submit your answers as a PDF file. You do not need to be overly concerned with formatting for this deliverable. Just ensure your responses are clear, readable, and address the full elements of each question (you can use screenshots of your results). Please put your name in the document and also in the filename.

Context

Many businesses use sales promotions to increase the demand or visibility of a product or service. These promotions often require increased expenditures (such as advertising) or loss of revenue (such as discounts). Firms must determine the effectiveness of such temporary promotions in generating additional demand. Regression based models using historical time series data are the primary tool for evaluating the incremental impact of short-term promotions.

Suppose you are hired by Kraft to conduct one such analysis. In the file labeled “Kraft.xlsx”, you are provided historical sales data for Kraft cheese along with marketing mix activities. See the sheet “data”. Specifically, the data includes the following variables:

- Week: week #
- Price Paid: net price paid by customers, including discounts
- Shelf Price: regular price without discount
- Feature: coded as AA to indicate advertisement in the local newspaper
- Units Sold = units sold in that week

Question 1: Kraft Promotion Analysis – Linear Regression (8 points)

Kraft is considering running a promotion of a 20% discount along with a feature ad. However, a potential concern for the management is a significant drop in sales the week following the promotion. In the industry jargon this is often referred to as “**Trough**”. If the trough following the promotion is large enough, the promotion may be rendered ineffective. Your objective is to develop a sales response model and evaluate the effectiveness of the promotion using historical data.

- Please create the following variables using the existing variables.
 - 1) % Discount = $(\text{Shelf Price} - \text{Price Paid}) / (\text{Shelf Price})$
 - 2) Feature Dummy = 1 when product is advertised, 0 otherwise
 - 3) Trough = Lag (% Discount) – this variable takes the value of previous week % Discount. It is used to capture the impact of promotion in previous week on sales in the current week.
- Estimate the following regression model. Please show the parameter estimates

Regression MODEL:

$$\text{Units Sold} = a + b * \text{ShelfPrice} + c * \% \text{Discount} + d * \text{Feature Dummy} + e * \text{Trough}$$

- How would you interpret each parameter, a, b, c, d and e? Which variable is the most important?
- Please assess the model fit and perform residual diagnostics.

Question 2: Kraft Promotion Analysis – Semi-log Regression (8 points)

Again, use the “data” sheet in “Kraft.xlsx”.

- a. Show the distribution of Unit Sold. Is it normally distributed?
- b. Apply the log transformation to “Unit Sold”. Run the following regression and show the parameter estimates

Regression MODEL:

$$\ln(\text{Unit Sold}) = a + b * \text{ShelfPrice} + c * \%Discount + d * \text{Feature Dummy} + e * \text{Trough}$$

- c. How would you interpret each parameter, a, b, c, d and e?
- d. Compare this model with the previous model. Which one fits better?

Question 3: Kraft Promotion Analysis –Trough (4 points)

Use the sheet “Promotion Analysis” in the file “Kraft.xlsx”. Using the parameter estimates from the semi-log model, compute the following and fill up the blanks in the sheet:

- a. “Baseline” sales & revenues (i.e., units sold in the absence of any promotion)
- b. Sales & Revenues for the week of promotion (20% Discount, Feature Ad) and Sales & Revenues in the week following promotion. What can you conclude from these results?