ONLINE CASE 3.10 ELWOOD FOOD COOPERATIVE
supplemental material to the text of
Modern Marketing Research: Concepts, Methods, and Cases
by Feinberg, Kinnear, and Taylor

Joyce Lauchner was the general manager of the Elwood Food Cooperative (EFC). She had become concerned of late that she had lost touch with the buying patterns of the cooperative members. EFC just seemed so big now compared with the early days. She hoped to make use of some data that were available to her to increase her understanding of the members' purchasing habits to better plan the mix and quantity of goods that EFC carried.

Background of EFC
EFC was founded in 1991 by Lauchner and a small group of volunteers. Located in an old warehouse on the northwest side of the town of Elwood, it had grown from ten original members in January 1991 to 500 members in September 2011. Elwood was a small town located in central Indiana, about 30 miles northeast of Indianapolis. EFC drew membership from a number of communities around Elwood, including Marion and Muncie.

The objective of EFC was to provide high-quality food products at prices below those available at local supermarkets. To accomplish this, EFC used shipping cartons as shelves, required shoppers to mark their own prices on goods, carried only the best-selling brands, and generally did not offer the "luxuries" associated with traditional supermarkets. To shop at EFC, one had to be a member. The membership fee was $45 per year. Any profits earned by EFC in a year were returned to the members as credits against purchases. Lauchner thought that the members bought most of their food at EFC.

Lauchner's Concerns
In the early days of EFC, Lauchner had prided herself on knowing all its members. She had spent a great deal of time in the store and felt she knew what people were buying and how much they were spending. As the membership grew, her administrative duties kept her in her office much more. She no longer knew all the members, nor did she have a good feel for their expenditure patterns. She wanted to develop a better understanding of these aspects of her business and thought perhaps some of the data that had already been collected on the membership might provide answers.

The Available Data
In June 2011 a questionnaire had been used to collect data on the membership. During the month, all members had come into the cooperative at least once. Thus, data were available on all members. The data consisted of demographic characteristics of the members plus their weekly food expenditures.

The data were available on the cards that had been filled out by the members at the time of the interview. Lauchner had these cards in a filing cabinet in her office. A description of the contents of the cards is presented in the next section. Actual card values are tabulated in Table C3.10.1. (Table C3.10.1, Demographic and Purchase Data is available online at ModernMarketingResearch.com as online_case_3.10_tables.xls.)

As a first step in understanding the membership, Lauchner wanted to know their average weekly expenditure on food. Because time was short, she wanted to do this without having to look at all 500 cards. However, she also wanted the average she calculated to be an accurate one. She wondered how she could make an accurate calculation.

EFC Data
Explanation of Items A–K in Table C3.10.1 (the variables)
A = household identification number (1–500)
B = weekly food expenditure, actual (e.g., $89.50)
C = number of persons in household, actual (1–9)
D = annual income of household, actual (e.g., $42,000)
E = education of head of household coded into five categories (1–5)
F = age of head of household, actual (e.g., 38)
G = weekly food expenditure, coded into seven categories (1–7)
H = any children younger than 6 years old in household, actual (1–2)
I = any children 6–18 years old in household, actual (1–2)
J = annual income of household, coded into six categories (1–6)
K = age of head of household, coded into seven categories (1–7).

Category Definitions for Variables
Variable C: Number of persons in household
1 = one person
2 = two persons
3 = three persons
4 = four persons
5 = five persons
6 = six persons
7 = seven persons
8 = eight persons
9 = nine or more persons

Variable E: Education of head of household
1 = less than grade 8
2 = grades 9–11
3 = high school graduate
4 = some college
5 = college graduate

Variable G: Weekly food expenditures
1 = less than $30
2 = $30–49.99
3 = $50–69.99
4 = $70–89.99
5 = $ 90–109.99
6 = $110–129.99
7 = $130 or greater

Category Definitions for Variables (continued)

Variable *H*: Any children younger than 6 years old in household
1 = no
2 = yes

Variable *I*: Any children 6–18 years old in household
1 = no
2 = yes

Variable *J*: Annual income of household
1 = less than $10,000
2 = $10,000–14,999
3 = $15,000–24,999
4 = $25,000–49,999
5 = $50,000–75,000
6 = $75,000 or greater

Variable *K*: Age of head of household
1 = less than 25
2 = 25–34
3 = 35–44
4 = 45–54
5 = 55–64
6 = 65–74
7 = 75 or older

As a second step, Lauchner wanted to analyze the demographic data and determine whether there was a relationship between spending and demographics. The data presented in Table C3.10.1 was subjected to correlation and regression analysis with the following three objectives in mind:

1. Determination of the strength of association between families' food spending and demographic characteristics
2. Determination of a function by which a family's spending can be estimated from its demographic data
3. Determination of the statistical "confidence" in the earlier tests.

The variables included in this analysis were as follows:

1. Expenditures: Weekly food ($)
2. Persons: Number in household
3. Income: Annual total of household ($1000s)
4. Education: Of head of household (five levels)
5. Age: Of head of household (years)
6. Children <6: 0 or 1
7. Children 6–18: 0 or 1.

Assume for the purposes of interpreting the results of this analysis, presented in Tables C3.10.2 and C3.10.3, that the 500 households used were randomly selected from a much larger population of households.

Table C3.10.2 Correlation Results

r	Expenditures	Persons	Income	Education	Age Children <6	Children <6	Children 6–18
Expenditures	—						
Persons	0.432	—					
Income	0.374	0.162	—				
Education	0.231	0.104	0.473	—			
Age	0.068	0.017	0.135	0.030	—		
Children <6	0.167	0.563	0.009	0.103	0.325	—	
Children 6–18	0.398	0.699	0.195	0.106	0.082	0.248	—

Table C3.10.3 Regression Results. Six-Variable Regression: Weekly Food Expenditures ($)

r	r^2	Adjusted r^2		Standard Error		n	
0.544	0.296	0.288		42.360		500	
Source of Variation	**Sum of Squares**	**df**	**Mean-Squares**	**F Test Statistic**	**p-Value One Tailed**		
Regression	372089.1	6	62014.8	34.561	0.0000		
Residual	884630.86	493	1794.38				
Total	1256720	499					
Model	**b**	**Std. Error**	**Std. Coeff. Beta**	**Test Statistic**	**p-Value Two Tailed**	**B Lower 95% CL**	**B Upper 95% CL**
Constant	33.086	11.878		2.785	0.0056	9.748	56.424
Persons	8.431	1.708	0.319	4.937	0.0000	5.075	11.786
Income ($1000s)	0.527	0.088	0.265	5.994	0.0000	0.354	0.700
Education	2.657	1.815	0.063	1.464	0.1438	–0.909	6.222
Age	0.062	0.166	0.015	0.374	0.7084	–0.264	0.388
Children <6	–4.480	5.328	–0.043	–0.841	0.4009	–14.949	5.989
Children 6–18	12.699	5.515	0.126	2.303	0.0217	1.863	23.535

Case Questions

1 Did Lauchner formulate any specific hypotheses to test? If so, what were they? If not, what hypotheses might the cooperative wish to test? How would you state them in statistical language?

2 How can Lauchner validate the income and spending data that she has collected? Consider both statistical and non-statistical types of validation.

3 What are the available alternative sampling procedures for estimating average weekly food expenditures? Explain how each of these sampling procedures would be carried out, and discuss the pros and cons of each. Are there any other variables whose averages it would be important to estimate?

4 Using the data file for all 500 households, calculate all the descriptive sample metrics that you believe might help Lauchner with the current situation. Do you believe these metrics should be considered representative of the overall EFC customer base? Explain your answer. If not, how could more representative metrics be collected?

5 Considering Lauchner's analysis objectives, what kind of analysis would you perform to address her information needs? Why? Use the full data file supplementing this case to conduct the recommended analyses. Develop a complete report (including statement of objectives, justification and description of the employed analyses, findings and conclusions) that you would submit to Lauchner. You should consider whether to limit your report to only the information needs suggested explicitly or to include additional analyses that seem relevant.

6 Evaluate the use of the correlation and regression analysis in this case. Do you believe they provide the adequate answers to Lauchner's questions? Do you believe that the relationships here are all linear? How should the nominal variables be included? Has a nominal variable been included incorrectly and, if so, how could you correct this potential mistake? Can such nominal variables be "correlated" with the other variables?

7 For the correlation results:
   a What interpretation can you give to the results shown? What can you say about "significance" here?
   b What assumptions underlie your interpretations? How could these assumptions be tested?
   c Is there a single "best" predictor of expenditures among the other six variables? What is it? Is this an appropriate interpretation?
   d There appears to be a positive correlation between "children <6" and "children 6–18," yet one correlates positively with age, the other negatively. Has an error been made? If not, how would you interpret this finding?

8 For the regression results:

a What interpretation can you give to these results? What assumptions underlie your interpretations? How could these assumptions be tested?

b If you were to rerun this regression, what variables would you leave out, and in what order? What criterion would you use to determine whether a variable should be *permanently* omitted?

c "Children 6–18" has the largest coefficient (other than "Constant," which we always ignore) but not the largest standardized coefficient. Is that reasonable? What does it mean?

d Do you think that the R^2 is somewhat low for a managerial application such as this? Does the modest R^2 value imply that our independent variables are poor predictors of food expenditures?

e Formulate and estimate a regression model—other than the one already presented—for weekly household food expenditures (Variable B). Can you come up with a "best" such model? What would the important independent variables be? Would multicollinearity be a problem? Can you treat expenditure as an interval-scaled variable? Defend your choices based on regression theory and your estimates.

f If you wanted to model the *nominal* version of expenditure (Variable G) instead, what sort of regression model might you use?

g Are there any other important (regression) relationships that Lauchner might wish to examine? If so, provide the estimates and interpretations.

9 Convert the "education" variable (five levels) into four dummy variables and repeat the regression. What can be concluded from the obtained results?