5. *Creating Effective Tables*

PROBLEM SET

1. Write a title for table 5A.

TABLE 5A.

Year	Median age (years)
1960	
1970	
1980	
1990	
2000	

Source: US Census of Population, various dates.

2. Answer the following questions for tables 5.2 through 5.7 in *Writing about Multivariate Analysis, 2nd Edition*.

   a. Who is described by the data?
   b. To what date or dates do the data pertain?
   c. Where were the data collected?
   d. What are the units of measurement? Are they the same for all cells in the table?
   e. Where in the table are the units of measurement defined?
   f. Does the table use footnotes? If so, why? If not, are any needed?
   g. Are panels used within the table? If so, why? If not, would the addition of panels improve the clarity of the table?

3. Table 5B needs several footnotes to be complete. What information would those footnotes provide?

TABLE 5B. Estimated OLS coefficients and standard errors from a model of BMI by demographic factors and health behaviors, Dietville, 2003

	Coefficient	Standard error
Intercept	19.03**	1.27
Age (years)		
Female		
Income level		
Poor		
Near poor		
Nonpoor		

	Coefficient	Standard error
Smoking		
None		
<1 pack/day		
1+ packs/day		
Exercise (days/week)		
<1		
1–2		
3+		
R^2	0.28	
F-statistic	4.21*	

4. What is missing from table 5C?

TABLE 5C. Results of an OLS model of log(poverty rate)

State median wage	−0.174	0.043
State median wage, squared	0.006	0.002
Log(state – federal EITC)	0.023	0.015
Log(state – federal minimum wage)	−0.015	0.011
Log(max state AFDC/FSP benefit)	0.543	0.194

5. Design a table for each of the following topics. Provide complete labeling and notes, show column spanner and panels if pertinent, and indicate what principle(s) you would use to organize items within the rows and/or columns, following the guidelines in chapters 5 and 6 of *Writing about Multivariate Analysis, 2nd Edition*.

   a. Age (years), gender, race, and educational attainment composition of a study sample.
   b. Bivariate measures of association between height (cm), weight (kg), percentage body fat, systolic blood pressure (millimeters of mercury [mm Hg]), and resting pulse (beats per minute).
   c. Results of logistic regression models of chances of high school graduation in the United States in 1998, stratified by gender and residence (urban versus rural). The key independent variables are mother's and father's educational attainment and occupation. Other control variables include race, family income, and number of siblings. Report effect size as odds ratios; statistical significance with *z*-statistics and symbols.
   d. Projected number of people receiving college degrees by region of the country from 2010 to 2025 under three different scenarios about rates of college attendance and completion.
   e. Net effects of an interaction between tercile of a student's own high school class rank and their mother's educational attainment (<HS, =HS, >HS) on the student's first-year college grade point average (GPA). Results are based on an OLS regression controlling for gender, race, and family income, using data from the high school classes of 1995 through 2000. Report results of inferential statistical tests using symbols, with the highest tercile of each independent variable as the reference category.

6. A journal for which you are writing an article allows no more than two tables, but your current draft has three. Combine tables 5D and 5E below into one table of 18 or fewer rows.

TABLE 5D. Number of wildfires by month, United States, 1998–2000

Month	1998	1999	2000	30-year average[a]
January				
February				
March				
April				
May				
June				
July				
August				
September				
October				
November				
December				
Total				

[a] 1970–1999.

TABLE 5E. Number of acres consumed by wildfire, by month, United States, 1998–2000

Month	1998	1999	2000	30-year average[a]
January				
February				
March				
April				
May				
June				
July				
August				
September				
October				
November				
December				
Total				

[a] 1970–1999.

7. There are at least seven things wrong with the labeling of table 5F. Identify and suggest ways to correct each error. Note: All numbers are correct.

TABLE 5F.1. Results of a logistic regression of political party preference, US, 2004

Variable	Odds ratio	Confidence interval	Wald chi-square
Age group 2	1.82	−0.015–3.83	4.13
Age group 3	2.01	−0.25–5.19	3.67
Race	0.53	−1.31–1.03	5.99
Proportion poor			
<10	1.26	−0.51–2.64	0.67
10–19	2.36	0.04–5.36	7.25
20–29			
>29	0.35	−2.02–0.93	7.69