Some Clarification about Cross-Sectional Designs

Your book is somewhat confusing about the basics of the cross-sectional design. This is an attempt at some clarification. The major purpose of cross-sectional designs is to compare how **two or more** existing groups compare to each other in regard to the outcome or dependent variable. Your book implies that you take one random, representative sample of the population and then sub-divide it into groups. I disagree with this implication. You should (virtually always) take ***independent, random, representative*** samples of each group at least for the independent variable of greatest interest (what your book calls a zero order relationship). That is, if you are primarily interested in comparing income of men and women, select a ***random, representative*** sample of each. Do not just start with one big sample.

The problem is that the number of samples you have to take multiplies rapidly as you add independent variables. For example, you may well think that educational level will also be very important. If you want to compare male college graduates and male high school graduates with female college graduates and female high school graduates, you are already taking four independent, representative samples. Whether someone works in the public or private sector could also be a big factor in determining income. If you now add working in the public sector versus working in the private sector, you end up with eight independent, random samples.

Male	High School	Public	Male	College	Public
Male	High School	Private	Male	College	Private
Female	High School	Public	Female	College	Public
Female	High School	Private	Female	College	Private

Now add another potentially important factor, years of work experience. Divide simply into those with 10 years or less and those with more than 10 years work experience. We get:

Male	High School	Public	10 Years or Less
Male	High School	Public	>10 Years
Male	High School	Private	10 Years or Less
Male	High School	Private	>10 Years
Male	College	Public	10 Years or Less
Male	College	Public	>10 Years
Male	College	Private	10 Years or Less
Male	College	Private	>10 Years
Female	High School	Public	10 Years or Less
Female	High School	Public	>10 Years
Female	High School	Private	10 Years or Less
Female	High School	Private	>10 Years
Female	College	Public	10 Years or Less
Female	College	Public	>10 Years
Female	College	Private	10 Years or Less
Female	College	Private	>10 Years

Taking independent samples for every group simply becomes impossible very rapidly. Therefore, at some point, you do have to start sub-dividing the samples. Nonetheless, the kinds of statistical analyses that you can perform and your ability to detect differences between groups will depend greatly on how you select the samples. Hence my recommendation that you always take independent samples based on the independent variable of greatest analytical interest, gender in this case.