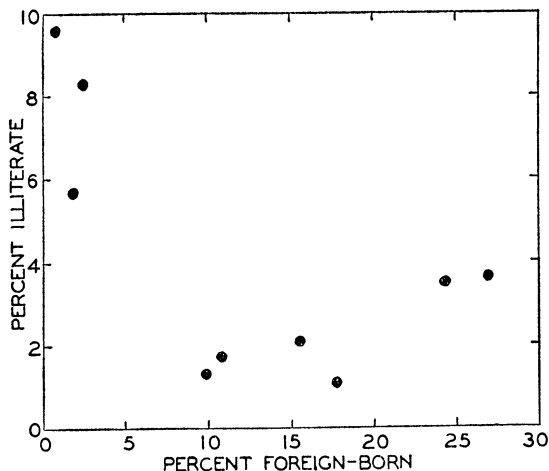


Elab 8.11

[f2008/f2008-17]

The Ecological Fallacy. Sometimes people use data in aggregated form to draw conclusions about individuals. For example, in 1950 W.S. Robinson described the correlation between immigration and illiteracy done in two different ways.[?] In the first, the unit of analysis is individual US states as shown in the figure — the plot shows the fraction of people in each state who are illiterate versus the fraction of people who are foreign born. The correlation is negative, meaning that states with higher foreign-born populations have *less* illiteracy.



Robinson’s second analysis involves the same data, but takes the unit of analysis as an individual person. The table gives the number of people who are illiterate and who are foreign born in the states included in the scatter plot.

TABLE 3. THE INDIVIDUAL CORRELATION BETWEEN NATIVITY AND ILLITERACY FOR THE UNITED STATES, 1930
(for the population 10 years old and over)

	Foreign Born	Native Born	Total
Illiterate	1,304	2,614	3,918
Literate	11,913	81,441	93,354
Total	13,217	84,055	97,272

The data in the table leads to a different conclusion than the analysis of states: the foreign born people are *more* likely to be illiterate.

This conflict between the results of the analyses, analogous to Simpson's paradox, is called the **ecological fallacy**. (The word "ecological" is rooted in the Greek word *oikos* for house — think of the choice between studying individuals or the groups of individuals in their houses.)

The ecological fallacy is not a paradox; it isn't a question of what is the correct unit of analysis. If you want to study the characteristic of individuals, your unit of analysis should be individuals. If you want to study groups, your unit of analysis should be those groups. It's a fallacy to study groups when your interest is in individuals.

One way to think about the difference between Robinson's conclusions with groups (the states) and the very different conclusions with individuals, is the factors that create the groups. Give an explanation, in everyday terms, why the immigrants that Robinson studied might tend to be clustered in states with low illiteracy rates, even if the immigrants themselves had high rates of illiteracy.