

Diversity Measurement

To avoid confusion, we use the term '*diversity index*' throughout this website. In statistical circles, however, the diversity index is also known as the *entropy index*, which is symbolized by E and can be formally defined as follows:

$$E = \sum_{r=1}^R p_r \ln \left(\frac{1}{p_r} \right)$$

where p_r refers to racial-ethnic group r 's proportion of the population in a given geographic unit and R indicates the total number of groups under consideration. The maximum value of E (the natural log of R) occurs when all groups are identical in size. Because there is no fixed upper bound, a population consisting of more equal-sized groups will produce a higher E score than one consisting of fewer equal-sized groups. An E of 0 (complete homogeneity) means that the population comprises a single group. Dividing E by its maximum value standardizes it to a 0-1 range. In the diversity data available here, we have multiplied the standardized scores by 100 so that 0 represents the lowest level of diversity and 100 the highest.

It is important to note that the diversity (entropy) index only measures the *magnitude* of diversity; the *structure* of diversity in a population—the specific racial-ethnic groups present—is ignored. The distinction between magnitude and structure becomes clear when one realizes that a community in which half of the residents are white and half are Asian would receive the same exact diversity index score as a community with a population in which half are black and half are Hispanic, despite potential differences between the two communities in socioeconomic status, intergroup relations, and other characteristics. To address this issue, most investigators supplement the index with group percentages or typologies that capture the structural dimension of diversity.