

Unholey Maps: How to Deal with Holes in your Data Sets

Aileen Buckley

ABSTRACT: Many data sets are incomplete, causing problems when performing spatial analyses or when mapping. Sometimes the data cannot be collected, for example, when access is denied to a location or when data collection methods are not sufficient for the task. Other times the data is collected but its quality is questionable or the method of collection is suspect. This happens, for example, when respondents may refuse or forget to answer a survey question. Still other times the data is collected but not shared, as in the case of data suppression for purposes of privacy. When data sets are incomplete, errors or biases in spatial analyses can occur, and maps can appear “holey” and uninformative. There are a variety of approaches that can be taken to help to mitigate these situations. In this paper, we explore a number of statistical approaches that compensate for missing data or, in some cases, fill in estimates for missing data. We also explore methods that cartographers have used to visually convey the lack of data. This paper examines these approaches in light of their relative strengths and weaknesses. We end with a set of general recommendations for performing analyses and making maps when your data sets are incomplete.

KEYWORDS: Map, cartography, data, missing, suppression, statistics, analysis

Aileen Buckley, Cartographer, Environmental Systems Research Institute (ESRI), Redlands, CA 92373