

## 4.1: Why Summarize Data?

---

When we summarize data, we are necessarily throwing away information, and one might plausibly object to this. As an example, let's go back to the PURE study that we discussed in Chapter 1. Are we not supposed to believe that all of the details about each individual matter, beyond those that are summarized in the dataset? What about the specific details of how the data were collected, such as the time of day or the mood of the participant? All of these details are lost when we summarize the data.

We summarize data in general because it provides us with a way to *generalize* - that is, to make general statements that extend beyond specific observations. The importance of generalization was highlighted by the writer Jorge Luis Borges in his short story "Funes the Memorious", which describes an individual who loses the ability to forget. Borges focuses in on the relation between generalization (i.e. throwing away data) and thinking: "To think is to forget a difference, to generalize, to abstract. In the overly replete world of Funes, there were nothing but details."

Psychologists have long studied all of the ways in which generalization is central to thinking. One example is categorization: We are able to easily recognize different examples of the category of "birds" even though the individual examples may be very different in their surface features (such as an ostrich, a robin, and a chicken). Importantly, generalization lets us make predictions about these individuals – in the case of birds, we can predict that they can fly and eat worms, and that they probably can't drive a car or speak English. These predictions won't always be right, but they are often good enough to be useful in the world.

---

This page titled [4.1: Why Summarize Data?](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Russell A. Poldrack](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.