

Aiden Martinez

Dr. Bradley

CODES 123

2/17/24

Reflection #2

The term “Data Censing” is typically used to describe the process of streamlining and correcting errors within datasets. The use of the term censing may be misleading, as it could potentially imply that certain data is unnecessary and needs to be removed; skewing datasets and harming research. If we were to change the verbiage of this term to indexing or categorizing, we both clarify the process to the layperson and change the schema of the process within the research community. Another concern of researchers is that of unintended bias. These biases that are unknown to the researchers themselves can alter datasets in severe ways. For example, a researcher grows up in a predominantly white community may have implicit biases against people of color, as they have had little exposure to people of color outside of popular media. This could unintentionally skew data and promote harmful stereotypes. So how do we prevent implicit biases and skewed datasets?

One strategy that is commonly used when categorizing data, one method that is commonly used is identifying groups. For example, if a dataset has a varied age pool, categorizing data by the age of the respondent can show trends or changes over time. By using demographics to categorize and compare datasets, we can identify extraneous data and how said data relates to the overall responses. To ensure fairness in research environments, it is important to have a diverse research community; coming from different economic, racial, sexual, and age groups. The greater the diversity of the researchers, the more a group can ensure that data collection is

accurate and fair for all respondents. Diverse research communities also promote new ideas and viewpoints when addressing potential issues within the research itself.