Data cleaning is described as “the process of preparing data for analysis by removing or modifying data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted” (“What Is Data Cleaning?”). When data cleaning is spoken through on the articles that were assigned, data cleaning is spoken down on.

It is argued that data shouldn’t be cleaned because the data never truly gets cleaned but only becomes messier. In the article *Against Cleaning*, the authors state that “what became evident was that cleaning up or correcting values was misleading- and even unproductive- way to think about how to make the data more useful for our own questions” (Rawson and Muñoz 281). From these findings, it became clear that sometimes cleaning data is not useful at all, on the contrary, cleaning data only makes things worse. Examples of what could become worse include, excluding voices that need to be seen within the data set, getting rid of a certain bias that is present in the data set to make the creator’s point valid, or fill in missing spots that would work best not filled in.

This exclusion of voices can be seen in our wicked problem at the Missouri Botanical Gardens with the African- American, African Diaspora, and Indigenous knowledges all being excluded by the westernized voices present in the herbarium with things such as plant naming.

Data cleaning can also be a good thing though. Typically data cleaning is described as “the process of preparing data for analysis by removing or modifying data that is incorrect, incomplete, irrelevant, duplicated, or improperly formatted” (“What Is Data Cleaning?”). There are multiple arguments over whether data should be cleaned or not, but as long as there is an ethical implication within the conversation when cleaning the data then data cleaning should be accepted.