

Data set Evaluation:

The data set used in this project largely satisfies the FAIR principles. Here we will focus on ACCESSIBILITY. This data was obtained from a published paper in an open-access journal. It was contained in a PDF file as part of the appendix of the paper. In order to use the data in R, we had to transform the file into a CSV file, since R was unable to read the data from the original PDF file. The data set contains the following metadata: names of the variables, units, and footnotes. The data file contains clarifications of particular data as well as more details about the collected data, so that researchers can reuse and understand the data. Additionally, since it is published in an open access Journal, anyone can have access to the data and to the publication associated to it. Finally, the publication contains information about the corresponding author, which makes it possible to find the data set even if the publication and its data are removed or hidden.

One suggestion I have about the accessibility of the data is that it would be more convenient to have access to both a PDF file and a CSV file so that a researcher be able to read the data in several ways. Also, I would suggest putting more explicit descriptions in for the variable, since for a data science person not related to the field of research, it could be challenging to realize what some technical names mean.

References:

GO FAIR. (n.d.). FAIR Principles. GO FAIR. Retrieved October 26, 2023, from <https://www.go-fair.org/fair-principles/>