Jackie Leger

LIS 545 B

2023-01-29

Updated: 2023-02-26

Term Project – Data & Metadata Profile

Word Count: 876 words

Data Set - [**Mortality of women of fertile age between 2006 and 2019: causes and trends**](https://figshare.com/articles/dataset/Mortality_of_women_of_fertile_age_between_2006_and_2019_causes_and_trends/21971443)

For this assignment, I have chosen the data for mortality of women of fertile age between 2006 and 2019. The data from this case study comes from Brazil with a focus on women of reproductive age or WRA, or women between the ages of 10 to 49 years old. With data such as this that focuses on health and reproduction, the key stakeholders are likely to be in the medical field; doctors, physicians, pediatricians, and obstetricians, to name a few. Women of reproductive age in Brazil are also key stakeholders for this data as it pertains their own health and well-being before, during, and after childbirth. Other researchers in the medical field focused on the same topic can also be key stakeholders as they could use this data to research ways to help prevent the many causes of WRA mortality rates.

The data in question is brought to us by Silmara Albert, Katrini Martinelli, Eliana Zandonade, and Edson Neto. These four researchers have presented their data in seven data files in the form of pictures of graphs that represent the data as well as excel files that break down the details of the data. Neither of which require any special software to view the data. Within these graphs and excel files we can see that the data presented is that of the main causes of death for the WRA. Due to the nature of the data and with the CC by 4.0 license the data does not have restrictions, but requirements for use of it. These requirements include, but are not limited to, appropriate credit given, providing a link to the CC by 4.0 license, and the indication of any changes that are made. (284; 466 words)

As previously mentioned, the data in the selected research is the mortality rate of women of reproductive age. For this dataset, the metadata are the causes of death for women within that age range of 10 to 49 years old. This can be found in the data files; both excel and pictures of the graphs. In the abstract presented by Albert et al. the six leading causes of death are as follows; “neoplasms (24.34), diseases of the circulatory system (20.15), external causes (18.69), infectious and parasitic diseases (8.79), and respiratory system diseases (6.37)” (2023). In the first table of data the authors list out various health disorders and diseases in the women such as blood disorders, endocrine, nutritional and metabolic diseases, and circulatory system diseases, to name a few. They also present these with the observed rates as well as the standardized rates that these health problems are generally found in WRA. The metadata presented is fairly comprehensive as it considers all possible variables when studying the causes of death. Due to the nature of the dataset and the metadata it is possible that the metadata standard being used is the RDF Data Cube Vocabulary. As this metadata standard states that it, “provides a means to publish multi-dimensional data, such as statistics, on the web in such a way that it can be linked to related data sets and concepts using the W3C RDF (Resource Description Framework) standard (*List of Metadata Standards | DCC, n.d.*)”. I say this because if we look at the possible chosen repository there are multiple datasets that center around the same exact data in various years. Meaning this particular dataset can be linked back to those other datasets and vice versa.

Although the data and metadata are comprehensive, one improvement I may suggest is to have it available in English as well. This, of course, isn’t the only language that may benefit from such information, but it would be a good starting point, especially considering the fact that the title of the data set is in English already. The inclusion of an indicator of which metadata standard it follows could also assist users in finding the data more easily. This data and metadata may also be useful for both biology majors and health science majors in understanding human biology where it pertains to childbirth. That is, of course, not to say that other majors such as library and science information majors cannot benefit from this data. However, people in this field may be less familiar with this data than those who in the previously mentioned science majors. Due to this lack of knowledge it may be more difficult to repurpose the data for other uses. One way to assist with this could be to explain the purpose behind the research and the goal of the research, as well as how it may assist other countries with WRA.

The new additions and extra resources for the data and metadata could also be beneficial to better understanding the corresponding research paper provided for those unfamiliar with the data. Published through SciELO Brazil, the paper with the same name as the data set is readable online, and even translatable through Google’s translation option. However, due to how newly published the research paper and dataset are, no results were found that cite or otherwise reference the use of this dataset. With the topic being so widely discussed already, it is only a matter of time before other researchers begin to utilize this data for their own research purposes.

References

Albert, S. B. Z., Martinelli, K. G., Zandonade, E., & Santos Neto, E. T. do. (2023). *Mortality of women of fertile age between 2006 and 2019: causes and trends* [Dataset]. SciElo. <https://figshare.com/articles/dataset/Mortality_of_women_of_fertile_age_between_2006_and_2019_causes_and_trends/21971443>

Albert, S. B. Z., Martinelli, K. G., Zandonade, E., & Santos Neto, E. T. do. (2023b). Mortality of women of childbearing age in Brazil from 2006 to 2019: causes and trends. *SciElo*. <https://www.scielo.br/j/rbepop/a/CP6YdG3RxTqzXDBmyCRJK9x/?lang=pt>

List of Metadata Standards | DCC. (n.d.). <https://www.dcc.ac.uk/guidance/standards/metadata/list?page=3>