Exercise 3: Assessing usability of data by analyzing definitions

Day 2: Wednesday, 3 August 2022, 11:00-14:00 (BKK Time)

1. Let’s say there was a recent DHS in the country of Poplandia. The DHS inquired about birth registration completeness to mothers in the survey. The DHS dataset can be disaggregated by various ethnicities, including the Mercurian Nomadic population. However, the DHS report methodology specifically states that the sample was not drawn in such a way to factor in the nomadic movements of this population and may not be representative. The response rate for Mercurian Nomads was very low (46%) i.e. households were not present (they had moved location) when enumerators went to go interview them. When you look at actual counts of mothers, they are in the 15-20 range and not published in the main DHS report.

   a. Can you use the DHS dataset to assess inequalities in birth registration in the Mercurian Nomadic population?

      The sample is not robust enough to factor in this population, and sample size under 30 is too low to draw strong conclusions on. However, more than likely this population is disadvantaged living in a remote area with inconsistent access to registry offices. This dataset might be useful when triangulating data. For example, if another survey focusing on the nomadic population, your assessment, and the DHS dataset all have similar results, they can help re-enforce your findings. They may have too many weaknesses to stand alone when it comes to large policy interventions, but are good to be used in conjunction with other sources.

   b. Bonus question: Can you think of any other factors to be careful of when it comes to mothers interviewed and DHS or MICS surveys?

      Many household surveys only interview MARRIED women or mothers. If a woman is a mother but not married, she may not be part of the sample. Further, women under age 15 or over age 49 may also be excluded. This can have implications for the registration of births to young, disadvantaged mothers.

2. Poplandia has done some more research into the data sources from the CRO and MOH to estimate death registration completeness. They are particularly interested in completeness by age. The CRO has a field for Date of birth and date of death but no variable for age of the deceased. The MOH has the age of the deceased coded in 5 year age group (e.g. 40-44 years), but no date of birth or date of death.
a. Can analysis of death registration completeness by age be done with this data? If not, what other considerations of proxies could be explored using these two data sources?

Yes, analysis can be done by age, but a new variable will need to be calculated from the civil registry data to calculate the age (group). It is preferable to have the DOB and DOD from MOH to calculate age and check data quality (or fill in missing values), but it is not strictly necessary.

3. As the inequality assessment has progressed, the civil registry office has provided death registration data for the entirety of the country, less the Eastern Province. However, the MOH is only able to provide data from its main hospital in the capital.

a. Can any analysis of death registration completeness by done? If so, what?

Yes, the data from the central hospital can be used as a proxy for deaths in the capital. This can still be disaggregated by any data available for sex, age, ethnicity etc. In this way, proceed with the analysis but explain your proxies/assumptions/caveats in your methodology and use caution and consideration in your interpretations.

b. If so, what are the caveats or things to be aware of?

However, it may not include deaths to those residing in the capital who died while traveling, and it may include those from outlying areas who traveled to the hospital for treatment and subsequently died. Some understanding or further investigation to whether the data is under or over-inflated to represent deaths in the capital may be needed.