How to know if you have the right data

Data Analysis
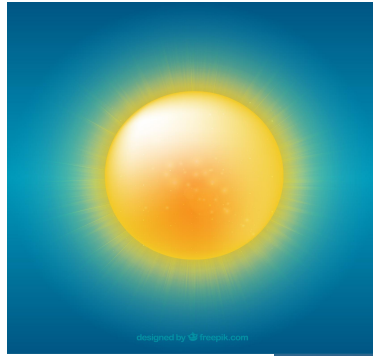
Number of observations is too small
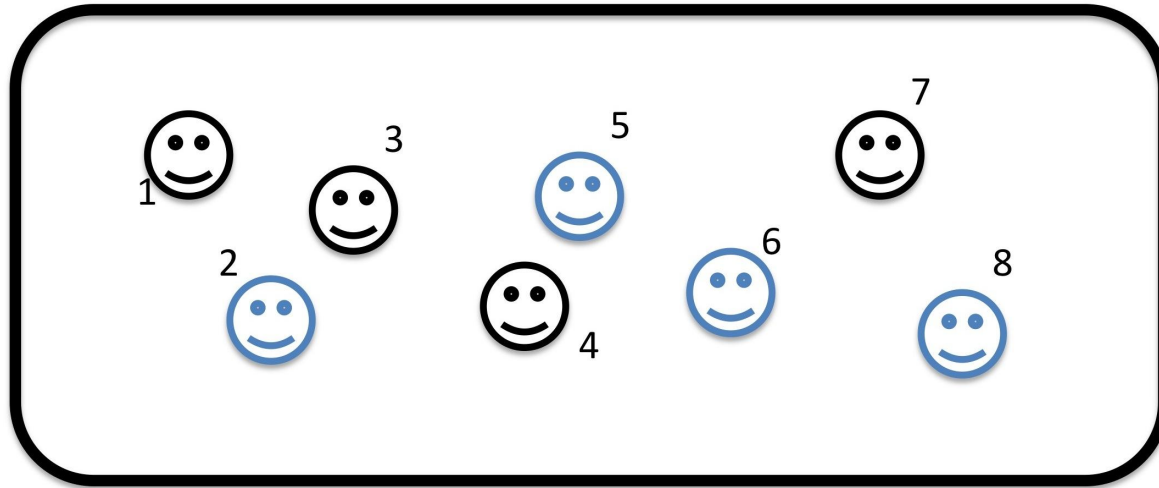
Dataset does not contain the exact variables you are looking for

Variables in the dataset are not collected in the same year

Dataset is not representative of the population that you are interested in
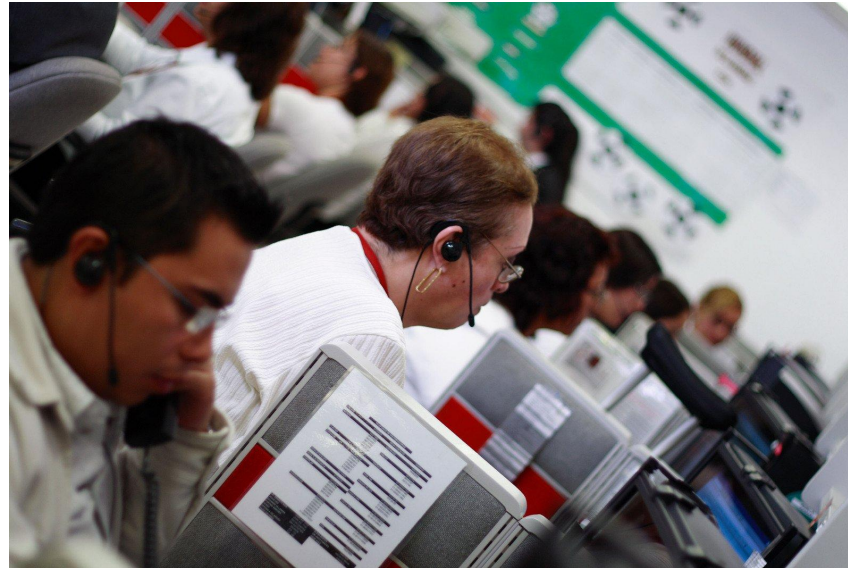
Dataset is not representative of the population that you are interested in

Dataset is not representative of the population that you are interested in

Examples of bad sampling:

- Surveying subscribers of a gun-related magazine for research on attitudes toward owning guns
- Surveying Facebook users for what TV shows Americans like

Best sampling practices:

- Always think about what your population is
- Collect data from a sample that is representative of your population
- If you have no choice but to work with a dataset that is not collected randomly and is biased, be careful not to generalize your results to the entire population

Some variables in the dataset is measured with error

Garbage in, garbage out!