

# 1 Methodology

For a dataset with  $n$  observations, each observation is given an equal probability of missing. The selection of data to eliminate can be written as  $\text{random\_choice}(\text{observationindexes})$  without replacement. In the code, `pandas.DataFrame.sa`