### Data coarsening

[*Data coarsening techniques include top/bottom coding, categorization, and grouping levels. The method top/bottom coding modifies the distribution of the variable by shrinking the tails. It can be decided to top code a variable, to bottom code a variable, or to do both (top/bottom code). A top coding consists of trimming a variable to some maximum value. For instance, a variable AGE can be top coded to 85 years old, which means all the people aged 85 or older are given the age 85 and all people younger than 85 keep their age on the public use file. Inversely, bottom coding consists of setting smaller values than a threshold to a specified minimum value. For instance, a variable AGE can be bottom coded to 5 years old, meaning that all children aged 5 or younger are assigned the age 5 while children older than 5 keep their age on the public use file. This method is used when there are too few respondents in the tails of the distribution of a sensitive variable. Categorization consists of recoding continuous or discrete variables in intervals. For instance, a variable AGE containing the age of the respondents (discrete variable) can be categorized in age groups, say, under 25 years old, more than 25 to 50 years old, and more than 50 years old. Grouping levels is done for ordinal and categorical variables and consists of merging levels of a variable to form broader categories. The idea is to combine levels with small frequency so that the resulting recoded variable will have more units in the smaller cells.*

*Provide discussion and tables showing initial risk analysis results that motivates top/bottom-coding, categorization and grouping levels. Details of the resulting data coarsening should be provided in an appendix.]*