5.2 Classify and code

This subprocess classifies and codes the input data. For example, automatic (or clerical) coding routines may assign numeric codes to text responses according to a pre-determined statistical classification to facilitate data capture and processing. Some questions have coded response categories on the questionnaires or administrative source of data, others are coded after collection using an automated process (which may apply machine learning techniques) or an interactive, manual process.

Quality Indicators:
- Compliance rate of classifications of input data to the pre-determined standard international classification and national versions of international classification schemes.
- All international or national classifications and breakdowns which are used for the data set are produced (i.e., although NACE Rev2 is introduced as an international classification, using the older version or using a different classification than the proposed classification).
- Compliance rate of coding of input data to the pre-determined standard coding scheme.
- Ratio between the number of values automatically coded and the total number of values submitted to coding.
- Extent to which quality control is used to manage the quality of automated and manual coding processes.
- Delay between expected and actual timing of adaptation of correspondence tables.
- Proportion of statistical units which cannot clearly be classified or mapped.

For details, see Quality Indicators for SQSPFM (version 2.0).