**Quality checks with R-scripts of data registered in FD2.**

Data collection made by Institute of Marine Research, Department of Aquatic resources (SLU Aqua), Swedish University of Agricultural Sciences.

Data editing is the application of routines to detect missing, invalid or inconsistent entries in data. Such routines highlight data records that are potentially in error which are evaluated and, if possible fixed, before data is used.

A set of checks for **missing values** are made with R-scripts on extracts from FD2. These include checks for, e.g., missing and duplicated trips, missing samples, missing total weight /sample weight values, missing length frequencies, missing lengths in length frequencies, missing specimens, missing biological data (various types of biological data).

**After missing values have been checked, data is checked for some clear errors (e.g.,** noticeably wrong haul coordinates, likely duplicated records**)** and **outliers. Outliers checks are** also made with R-scripts and include screening the data for potentially erroneous entries that may be highly impactful to final estimates such as too high/low landings or discards per haul, unusually high/low catch/sample weights, atypical values in trip characteristics (e.g., atypical number of days at sea, atypical positioning of fishing operations). Depending on the variable of interest, an outlier can be defined as an atypical observation relative to other data collected or relative to estimates derived from them. Example of outliers are estimated weights of samples (as obtained via length-weight relationship) being very different from registered sampled weights; atypical catch fraction values (as signalled by several types of box-plot analysis, e.g., by gear, etc); atypical lengths (as signalled by several types of box-plot analysis, e.g., by fraction, by size category, by gear, etc); atypical observations in scatterplots and boxplots between different biological variables (length, weight, age,...)

During the estimation **temporal consistency** data checks (e.g. variation of data with quarters/years) are made.

How the data checks are made is **documented** in the R-script and is stored locally.