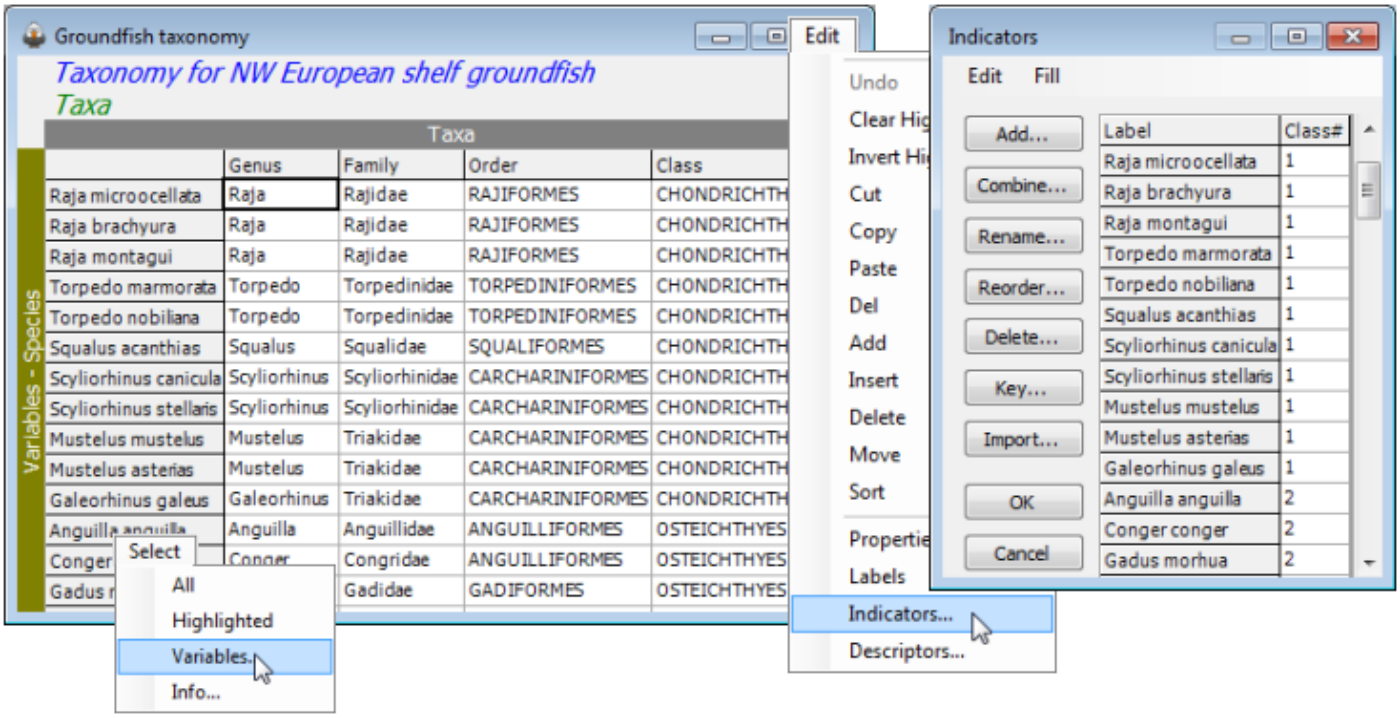
Variable information (aggregation files)

However, the full range of hierarchical indicators represented by a Linnaean classification (which species belong to which genera, which genera to families, families to orders, etc.) are usually also best held separately, as a different type of array – that of *variable information*. Mainly for historical reasons these are termed ‘aggregation files’ in PRIMER, since their initial use was for aggregating species abundances up to genus, family, order, … level information, to judge the extent of change to the interpretation of analyses under coarser identification of taxa (see Chapter 10 of the CiMC manual), and this binary file format is therefore denoted by *.agg. However, in PRIMER 7, arrays of variable information can be more general (and have other Type definitions than *Taxa*). Former aggregation file formats can be opened and PRIMER 7 outputs the full range of previous types, e.g. Save as type: PRIMER Var Info Files (*.agg) for PRIMER 7 (binary); PRIMER 6 or 5 aggregation files.agg (also binary); and simple Text (*.txt) or Excel (*.xls) (or *.xlsx) sheets. Examples using aggregation files will be seen later (Sections 5, 11, 15) though the simple rectangular format is seen here by opening Groundfish taxonomy.agg from the C:\Examples v7\Europe groundfish directory. Three ways in which it might be used are to: a) aggregate abundance to higher taxa with **Tools> Aggregate** (Section 11, and Chapter 10 of CiMC); b) compute biodiversity indices based on the relatedness of species in a single sample, e.g. with **Analyse>DIVERSE** (Section 15, Chapter 17); c) compute resemblance measures between two samples reflecting (higher) taxonomic relatedness of the species found there (Section 5, Chapter 17).

This new *variable information* sheet (below) permits the non-numeric entries which are essential for a variables $\times$ taxa ‘look-up’ table but also, and newly in PRIMER 7, will carry over several of the general features of sample $\times$ variables arrays, in that indicators defined on the variables can now be carried around with this array. This might permit the aggregation file to hold alternative names for single species, for example, with an indicator that can be used to select only the taxonomic revision relevant to the historic date of collection/identification of the species count matrix. Importantly, it also allows easy selection of aggregation file subsets, e.g. for testing taxonomic distinctness indices against differing ‘master lists’ by region, habitat or faunal group (Chapter 17 of CiMC). The simple indicator in the example below could be used to select only the Osteichthyes (Class# = 2) from the Variable information: Groundfish taxonomy, as well as from the data: Groundfish density(.pri).



Note the final entry on the Edit menu here. The concept of *Descriptors* is not particularly relevant to Variable information of type Taxa (they are potentially more relevant to other types of Variable information) but they are the third construction logically needed. Categories (or alternative labels) applied to Samples are termed *Factors*, when applied to Variables they are called *Indicators* and when applied to Variable information they are *Descriptors*.

Revision #9
Created 19 May 2024 23:14:26 by Arden
Updated 14 January 2025 22:10:15 by Abby Miller