**Data description for example data from:** [**https://github.com/MurraySAThompson/fish-feeding-guild-classifcation**](https://github.com/MurraySAThompson/fish-feeding-guild-classifcation)

This work was undertaken for the study "Fish functional groups of the North Atlantic and Arctic Oceans ", by Murray S.A. Thompson, Izaskun Preciado, Federico Maioli, Valerio Bartolino, Andrea Belgrano, Michele Casini, Pierre Cresson, Elena Eriksen, Gema Hernandez-Milian, Ingibjörg G. Jónsdóttir, Stefan Neuenfeldt, John K. Pinnegar, Stefán Ragnarsson, Sabine Schückel, Ulrike Schückel, Brian E. Smith, María Ángeles Torres, Thomas J. Webb, and Christopher P. Lynam (in review; <https://essd.copernicus.org/preprints/essd-2024-102/>)

**Data description**

File = ‘**example data for cluster analyses.csv**’

sample\_id = unique stomach sample identifier

pred\_id = abbreviated predator taxa and individual body-mass bin grouping

prey\_func = prey functional group

prey\_family = prey family

prey\_weight = prey wet weight in grams

prey\_count = prey count

ind\_prey\_weight = individual prey weight in g

direct\_observation = whether directly observed prey weights and counts (‘Y’) or modelled (‘NA’)

File = ‘**example samples for cluster analyses.csv**’

sample\_id = unique stomach sample identifier

data = data source. ‘Dapstom’ (An Integrated Database & Portal for Fish Stomach Records; Pinnegar, 2019), ‘ICES\_YOTS’ (ICES Year of the Stomach; Daan, 1981; ICES, 1997), and ‘USNWA’ (the Northeast US continental shelf; Smith & Link, 2010)

year = the year the sample was collected

latitude = sample location in decimal degrees latitude

longitude = sample location in decimal degrees longitude

haul\_id = unique haul identifier

pred\_taxa = predator taxa

pred\_weight\_g = predator wet weight in grams

pred\_length\_cm = predator length in cm

bin\_number = 20 equal size bins to categorise individual predator mass along a log10 transformed gradient from 0.1 micrograms to 190 tonnes

pred\_id = abbreviated predator taxa and individual body-mass bin grouping

n\_stomachs = number of stomach samples per unique stomach sample identifier (some stomach samples were pooled for species length classes)

direct\_observation = whether directly observed prey weights and counts (‘Y’) or modelled (‘NA’)

File = ‘**example\_survey\_data.csv**’

FileName = name of file downloaded from: <https://data.cefas.co.uk/view/21421> (Lynam & Ribeiro, 2022) where a full description of the data can be found.

HaulID = unique haul identifier

FishLength\_cm = fish length in cm

DensBiom\_kg\_Sqkm = biomass in kg wet weight per km2

DensAbund\_N\_Sqkm = abundance per km2

SciName = scientific name of observed taxa

Year = the year the sample was collected

SweptArea\_KM2 = swept area sampled in km2

Gear = survey gear used

Ship = ship used

Month = the month the sample was collected

Day = the day the sample was collected

TimeShot = the time the sample was collected

ShootLong\_degdec = sample location in decimal degrees longitude

ShootLat\_degdec = sample location in decimal degrees latitude

DensAbund\_N\_perhr = abundance per unit effort (in hours)

DensBiom\_kg\_perhr = biomass in kg wet weight per unit effort (in hours)

SurvStratum = survey stratum

Survey\_Acronym = survey acronym

File = ‘**example taxa.csv**’

taxa = organism taxa

worms\_aphia\_id = worms aphia identifier (see <https://www.marinespecies.org/aphia.php?p=webservice>)

functional\_group = adult functional group of organism following Webb & Vanhoorne (2020) using the “worrms” package (Chamberlain, 2019).

phylum = organism phylum

class = organism class

order = organism order

family = organism family

genus = organism genus

species = organism species

revised\_taxa = updated taxonomy

LWRa = a values for length-weight calculation

LWRb = b values for length-weight calculation

LWRSource = source of a and b values for length-weight calculation

**References**

Chamberlain, S. (2019). *worrms: World Register of Marine Species (WoRMS) Client* (R package version 0.4.0.).

Daan, N. (1981). *Data base report of the stomach sampling project 1981. ICES Cooperative Research Report, Vol. 164. 150 pp*.

ICES. (1997). *Database report of the stomach sampling project, 1991. ICES Cooperative Research Report, Vol. 219. 426 pp*. https://doi.org/https://doi.org/10.17895/ices.pub.4626

Lynam, C. P., & Ribeiro, J. (2022). *A data product derived from Northeast Atlantic groundfish data from scientific trawl surveys 1983-2020*. https://doi.org/https://doi.org/10.14466/CefasDataHub.126

Smith, B. E., & Link, J. S. J. (2010). The Trophic Dynamics of 50 Finfish and 2 Squid Species on the Northeast US Continental Shelf. *U.S. Dep. Commer. NOAA Technical Memorandum*, *NMFS-NE-21*(May), 1–29.

Webb, T. J., & Vanhoorne, B. (2020). Linking dimensions of data on global marine animal diversity: Dimensions of global marine diversity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *375*(1814), 20190445. https://doi.org/10.1098/rstb.2019.0445rstb20190445