This data publication contains data associated with the publication 'On the challenges of identifying benthic 'dominance' on Anthropocene coral reefs' in BioScience. This data publication includes nine .csv files. The first .csv file lists the sources and details of raw benthic composition data already in the public domain that were utilised in the manuscript. Relevant accession numbers, DOIs and URLs are provided in this file to access these underlying raw data. The second .csv file includes the reference details of 83 publications from which benthic composition data were derived. The third .csv file includes the data derived from the 83 publications (please note this is a subset of the full dataset used in the analyses, and the remaining raw data is already publicly available with full accession details listed in the first .csv file). This third .csv file contains data on the % coverage of 6 benthic categories (HardCoral, SoftCoral, Macroalgae, CTB [i.e. crustose coralline algae, turf, bare], Sand, Other) on coral reefs around the world from 1977-2018. For each data entry, the file also contains information on the biogeographic realm, the coordinates (latitude and longitude), the habitat, the depth (in metres) sampled, the year in which sampling occurred, and the method which was used to collect the benthic cover data in the field (line intercept transect [LIT], point intercept transect [PIT], chain intercept transect [CIT], photo quadrats along a transect [PQT], in-situ quadrats [QUAD], or analysis of individual frames from a video [VFA]). Note only data from PIT and PQT methods were used in the current manuscript. Data is based on mean site level benthic cover data each year, where a site is taken to be a unique set of GPS coordinates at a unique depth, in the same habitat.

The .csv files (4-6) contain data on which benthic category was dominant (i.e. most abundant) for each observation/site. File 4 contains dominance data across all 6 benthic categories, file 5 contains dominance data when the 'sand' category was excluded from calculations, while file 6 contains dominance data when both 'sand' and the 'CTB' category were excluded from calculations. Files 4-6 are based on data derived from PIT and PQT methods, across four major realms (Central Pacific, Indian Ocean, Indo-West Pacific, Western Atlantic). In all cases a '1' denotes the dominant category. Other metadata are as per the third .csv file, above.

The .csv files (7-9) contain data on which benthic category was dominant (i.e. most abundant) for each observation/site when algae were categorised at different levels. File 7 contains dominance data when all algae were pooled as a single 'algae' category, file 8 contains dominance data when algae were categorised as either 'macroalgae' or 'CTB', while file 9 contains dominance data when algae were categorised as either 'macroalgae', 'turf', 'CCA', 'rubble', or 'barematrix'. Files 7-9 are based on data derived from PIT and PQT methods, across three major realms (Central Pacific, Indo-West Pacific, Western Atlantic). In all cases a '1' denotes the dominant category. Other metadata are as per the third .csv file, above.