

Table. Column names (n=63), variable definition, variable nature and variable units for the database. These columns contain information on behavioral, morphological and ecological traits, as well as biogeographical, and conservation aspects relative to 2,198 reef fish species from the Eastern and Atlantic realms. The realms were defined according to Kulbicki et al. (2013).

Variable Name	Variable definition	Nature	Units
fb_code	This code is an identifier assigned to each species according to FishBase (Froese and Pauly 2020)		
Class	Taxonomic class		
Order	Taxonomic order		
Family	Taxonomic family		
Genus	Taxonomic genus		
Species	Taxonomic specific epithet		
Name	Scientific Name (Genus and specific epithet)		
Home_range	<p>Area of activity:</p> <ul style="list-style-type: none"> • Sedentary or Territorial species (sed): species that remain in a restricted area (less than a few 10 m²) for extended periods (at least a month). • Mobile species (mob): species that remain in area of more than 100 m² or traveling among different reef areas. • Very mobile (vmob): species which frequently change reefs or travel large distances on the same reef daily. 	Categorical (ordinal)	
Diel_activity	<p>Period of the day that species feed and breed:</p> <ul style="list-style-type: none"> • day: diurnal species - species active essentially during day time hours. • both: species active during both night and day or species with crepuscular behavior. • night: nocturnal species – species mainly active during the night. 	Categorical (nominal)	

Size_group	Gregariousness behavior: <ul style="list-style-type: none"> • solitary (sol): species that usually live alone. • pair: species usually seen in pair. • small groups (smallg): species that form small groups (3 to 20 individuals on average). • medium groups (medg): species that form groups of 21-50 individuals on average. • large groups (largeg): species that form large groups or schools (> 50 individuals on average). 	Categorical (ordinal)	
Level_water	Position in water column: <ul style="list-style-type: none"> • bottom: species staying on the bottom at all the times. • low: species that live slightly above the bottom but occasionally rest on the bottom. • high: species spending most of their active time high above the bottom (several meters). 	Categorical (ordinal)	
Occurrence_temperature	These data represent the averages temperature estimated to each species from temperature reported to each locality where these species occur. The temperature data were estimated from satellite data averaged between 2002 and 2009 from BIO-oracle (Tyberghein et al. 2012), [Mean=TempOccurrence_mean, Minimum=TempOccurrence_min, and Maximum=TempOccurrence_max]	Continuous	Degrees Celsius
Body_size_max	Maximum total length of species.	Continuous	Centimeters
Size_class	Maximum total length of species divided into ordered categories according to Mouillot et al. (2014). <ul style="list-style-type: none"> • s1: 0 – 7 cm • s2: 7.1 – 15 cm • s3: 15.1 – 30 cm • s4: 30.1 – 50 cm • s5: 50.1 – 80 cm 	Categorical (ordinal)	Centimeters

	<ul style="list-style-type: none"> • s6: > 80 cm 		
Body_shape	Morphological specialization for swimming. <ul style="list-style-type: none"> • box_shaped: globular form, as in Diodontidae. • compressed: body flattened laterally. • depressed: body flattened dorsoventrally. • eel_like: body elongated, but with snake-like shape and locomotion. • elongated: long in relation to length. • fusiform: body spindle-shaped, the most hydrodynamic form. 	Categorical (nominal)	
Aspect_ratio	An indicative value of species activity calculated from the squared height of caudal fin (h) divided by its surface (s) area (h^2/s).	Continuous	Nondimensional
Nphoto	Number of photos on which the aspect ratio metric was based.	Discrete	Nondimensional
Caudal_fin	Shape of caudal fin: <ul style="list-style-type: none"> • forked: caudal fin with a deeply concave or excavated hind margin. • lanceolated: spear shaped caudal fin. • pointed: when the dorsal and anal fins are joined to the caudal. • rounded: caudal fin with only one lobe and concave margin. • truncated: caudal fin with only one lobe and straight margin. • heterocercal: asymmetric fin. 	Categorical (nominal)	
Mouth_position	Anatomic jaw position: <ul style="list-style-type: none"> • superior: lower jaw is protruding upwards. • terminal: the upper and lower jaws are equal in length. • subterminal: upper jaw is protruding downwards. • inferior: mouth located in ventral position. • tubular: enlargement of mouth cavity. 	Categorical (nominal)	

	<ul style="list-style-type: none"> • elongated: long snout due to the increase in the bones premaxillary, maxillary and dental. 		
Trophic_level	Position in the food web based on diet composition and mean trophic level of food items	Continuous	Nondimensional
Diet	<p>Main items consumed by species:</p> <ul style="list-style-type: none"> • herbivore-detritivore (hd): feeds predominantly on the epilithic algal matrix (EAM). • macroalgal feeder (hm): feeds predominantly on macroalgae (>1cm high) or seagrass. • sessile invertebrates (is): feeds predominantly on sessile benthic invertebrates including corals, sponges and ascidians. • mobile benthic invertebrates (im): feeds predominantly on mobile benthic invertebrates including crustaceans, annelids, echinoderms and mollusks. • planktonic (pk): feeds predominantly on invertebrate material in the water column including zooplankton, and, detrital aggregates. • omnivore (om): feeds on a range of organisms and includes a large proportion of animal material and plant and/or detrital material. • piscivore (fc): feeds predominantly on fishes but may also include cephalopods. 	Categorical (nominal)	
Spawning	<p>The way that species deposit eggs in water: demersal: deposit their eggs directly on the substrate such as rocks.</p> <ul style="list-style-type: none"> • live: species without an external larval stage. pelagic: eggs free in the column water • attach: species that attach their eggs to objects that float in the water rather than the bottom substrate. 	Categorical (nominal)	

	<ul style="list-style-type: none"> • oral: species with parental care in which female or male keep the eggs in their mouth until complete development. 		
Depth_min	Minimal depth reported for the species.	Continuous	Meters
Depth_max	Maximal depth reported for the species.	Continuous	Meters
Depth_range	Difference between the maximum and minimum depth reported for the species	Continuous	Meters
Allometric_a	Constant allometric coefficient estimated from the relationship between size and weight	Continuous	Nondimensional
Allometric_b	Constant allometric coefficient estimated from the relationship between size and weight	Continuous	Nondimensional
Level_length_weight	Taxonomic level at which the information was compiled <ul style="list-style-type: none"> • family • genus • species 	Categorical	
PLD	Pelagic larval duration.	Discrete	Days
PLD_level	Taxonomic level at which the information was compiled <ul style="list-style-type: none"> • family • genus • species 	Categorical (nominal)	
Life_span	Maximum age expected according to growth curves reported in the FishBase	Discrete	Years
Geographic_range_index	This index was estimated from distribution data used to build range maps for each species, defined as the convex polygon shaping the area where each species is present (Parravicini et al. 2014). From these maps was sum the number of grid cell (5° x 5°; corresponding to approximately 550 x 550 km) where the each species has been recorded.	Discrete	Nondimensional
Range_extension	Classification on the geographical origin of species:	Categorical (nominal)	

- invasive: species are non-native to the ecosystem which can causes several alterations in the ecosystem and influence negatively native species.
- lessepsian: migratory species from the Red Sea to the Mediterranean.
- native: species with evolutionary history in the Atlantic or Eastern Pacific.
- transatlantic: species with restricted distribution in one part of the Atlantic but were recently reported in other localities.
- transpacific: species native from Pacific but were recently reported in Atlantic.

Latitude	Coordinate at which the species occurs. These data were compiled based on species checklists. [Lat_Max_North, Lat_Min_South]	Continuous	Degrees
Provinces	Presence or absence in provinces distributed on 20 columns named province_number (binary). These provinces were defined by Spalding et al. (2007)	Binary	0/1
Total_occupancy_province	Total number of provinces where one species occur	Discrete	
Ocurrence_Atlantic	Presence or absence of each species in the Atlantic Realm	Binary	0/1
Ocurrence_EasternPacific	Presence or absence of each species in the Eastern Pacific Realm	Binary	0/1
IUCN_status	Conservation status of species <ul style="list-style-type: none"> • NE: Not evaluated • DD: Data deficient • LC: Least concerned • NT: Near threatened • VU: Vulnerable • EN: Endangered • CR: Critically endangered 	Categorical (nominal)	
Vulnerability	Continuous values ranging from 0 to100, provided by FishBase and based on fish life-history traits. Values closer to 100 represent more	Continuous	Nondimensional

	vulnerable species whereas species with values closer to 0 are less vulnerables	
Price_category	<p>Price value based on ex-vessel price for the species. Ex-vessel prices are the prices that fishers receive directly for their catch, or the price at which the catch is sold when it first enters the supply chain.</p> <ul style="list-style-type: none"> • unknow • low • medium • high • very high 	Categorical (ordinal)