Package ‘BioWorldR’

May 6, 2026

Type Package

Title A Curated Collection of Biodiversity and Species Datasets and Utilities

Version 0.1.0

Maintainer Juan David Monroy <monroyjuandavid773@gmail.com>

Description Provides a curated collection of biodiversity and species-related datasets (birds, plants, reptiles, turtles, mammals, bees, marine data and related biological measurements), together with small utilities to load and explore them.
The package gathers data sourced from public repositories (including Kaggle and well-known ecological/biological R packages) and standardizes access for researchers, educators, and data analysts working on biodiversity, biogeography, ecology and comparative biology.
It aims to simplify reproducible workflows by packaging commonly used example datasets and metadata so they can be easily inspected, visualized, and used for teaching, testing, and prototyping analyses.

License GPL-3

URL https://github.com/Monroy31039/BioWorld, https://Monroy31039.github.io/BioWorld/

BugReports https://github.com/Monroy31039/BioWorld/issues

Encoding UTF-8

LazyData true

Suggests ggplot2, dplyr, testthat (>= 3.0.0), knitr, rmarkdown

RoxygenNote 7.3.2

Config/testthat/edition 3

VignetteBuilder knitr

NeedsCompilation no

Author Juan David Monroy [aut, cre]

Depends R (>= 3.5.0)

Repository CRAN

Date/Publication 2025-09-27 08:10:02 UTC

Contents

Apes_list . . . 2
Bees_data . . . 3
BioWorldR . . . 5
Birds_Peru . . . 5
Brain_animals . . . 6
Chimp_f . . . 7
Chimp_m . . . 7
Crab_morpho . . . 8
DogsLife_df . . . 9
Elephants_df . . . 10
Extinct_mammals_tbl_df . . . 10
Fish_metallo . . . 11
flwr_time . . . 12
Germ_lines_df . . . 13
Gorillas_df . . . 14
Gorilla_f . . . 15
Gorilla_m . . . 15
Orang_f . . . 16
Orang_m . . . 17
Plants_list . . . 17
Plants_occ . . . 18
Reptiles_df . . . 19
Savannah_trees . . . 20
Snakes_df . . . 20
sparrow_morph . . . 21
Tsd_turtle . . . 22
Turtles_list . . . 24

Index **25**

`Apes_list` *Great Ape Skull Landmark Data*

Description

This dataset, Apes_list, is a list containing landmark data of great ape skulls. It includes 8 landmarks in 2 dimensions for 167 individuals. The dataset consists of two elements: the landmark coordinates and the group factor indicating species. All values are preserved as in the original source, with no modifications.

Usage

```
data(Apes_list)
```

Format

A list with 2 elements:

x Array of landmark coordinates (numeric) with dimensions 8 landmarks × 2 dimensions × 167 individuals

group Factor indicating the species or group of each individual

Details

The dataset name has been kept as 'Apes_list' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the shapes package version 1.2.7

`Bees_data` *Bee Occurrence Records*

Description

The 'Bees_data' dataset contains occurrence records of bee species, including detailed taxonomy, geographic coordinates, collection metadata, and data quality flags. It is designed for biodiversity, ecology, and conservation studies involving bee populations.

Usage

```
data(Bees_data)
```

Format

A tibble with 105 observations and 124 variables:

database_id Unique database identifier (character)

scientificName Full scientific name (character)

family Taxonomic family (character)

subfamily Taxonomic subfamily (character)

genus Taxonomic genus (character)

subgenus Taxonomic subgenus (character)

subspecies Taxonomic subspecies (character)

species Species name (character)

specificEpithet Specific epithet (character)

infraspecificEpithet Infraspecific epithet (character)

acceptedNameUsage Accepted taxonomic name (character)

taxonRank Rank of the taxon (character)

scientificNameAuthorship Authority for the name (character)

decimalLatitude Geographic latitude (numeric)

decimalLongitude Geographic longitude (numeric)

stateProvince Administrative region (character)

continent Continent name (character)

locality Locality description (character)

eventDate Date of collection event (date/character)

year Year of collection (integer)

basisOfRecord Type of record (e.g., specimen, observation) (character)

occurrenceStatus Occurrence status (character)

recordedBy Collector(s) name(s) (character)

institutionCode Code of institution holding the record (character)

datasetName Name of the source dataset (character)

gbifID Global Biodiversity Information Facility record ID (character)

spatiallyValid Spatial validity flag (logical)

coordinateUncertaintyInMeters Uncertainty of coordinates (numeric)

hasGeospatialIssues Flag for geospatial issues (logical)

duplicateStatus Duplicate record flag (character)

.GBIFflags GBIF quality control flags (character)

... Additional metadata on taxonomy, collection, and validation flags (90+ fields)

Details

This dataset is rich in metadata and includes validation flags such as:

- Coordinate validity and uncertainty
- Taxonomic cleaning and synonym resolution
- Record duplication and licensing status
- Event and occurrence metadata

Source

Extracted from the **BeeBDC** package, version 1.3.0.

BioWorldR — *BioWorldR: A Curated Collection of Biodiversity and Species Datasets and Utilities*

Description

This package provides a curated collection of biodiversity and species-related datasets (birds, plants, reptiles, turtles, mammals, bees, marine data and related biological measurements), together with small utilities to load and explore them.

Details

BioWorldR: A Curated Collection of Biodiversity and Species Datasets and Utilities

A Curated Collection of Biodiversity and Species Datasets and Utilities

Author(s)

Maintainer: Juan David Monroy <monroyjuandavid773@gmail.com>

See Also

Useful links:

- https://github.com/Monroy31039/BioWorld

Birds_Peru — *Bird Species Recorded in Peru (Updated 2025 Version)*

Description

This dataset, Birds_Peru, is a tibble containing information on bird species recorded in Peru, based on the most recent taxonomic revisions by the South American Checklist Committee (SACC). The dataset includes 1,914 observations and 6 variables, covering taxonomic order, family, scientific names, common names in English and Spanish, and conservation or occurrence status. All values are preserved as in the original source, with no modifications.

Usage

```
data(Birds_Peru)
```

Format

A tibble with 1,914 observations and 6 variables:

order_name Taxonomic order (character string)

family_name Taxonomic family (character string)

scientific_name Scientific name of the species (character string)

english_name Common name in English (character string)

spanish_name Common name in Spanish (character string)

status Conservation or occurrence status (character string)

Details

The dataset name has been kept as 'Birds_Peru' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the avesperu package version 0.0.6

`Brain_animals` *Brain and Body Weight Data for Animals*

Description

This dataset, Brain_animals, is a tibble containing brain and body weight measurements for 28 animal species, useful for allometric and comparative studies. The dataset includes 28 observations and 3 variables. All values are preserved as in the original source, with no modifications.

Usage

```
data(Brain_animals)
```

Format

A tibble with 28 observations and 3 variables:

species Species name (character)

bodyweight Body weight of the species (numeric)

brainweight Brain weight of the species (numeric)

Details

The dataset name has been kept as 'Brain_animals' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified.

Source

Data taken from the BSDA package version 1.2.2

Chimp_f — *Female Chimpanzee Skull Landmark Data*

Description

This dataset, Chimp_f, is an array containing landmark data of female chimpanzee skulls. It includes 8 landmarks in 2 dimensions for 26 individuals. All values are preserved as in the original source, with no modifications.

Usage

```
data(Chimp_f)
```

Format

An array with dimensions $8 \times 2 \times 26$:

[1:8, 1:2, 1:26] Numeric values representing 8 landmarks in 2 dimensions for 26 female chimpanzee skulls

Details

The dataset name has been kept as 'Chimp_f' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the shapes package version 1.2.7

Chimp_m — *Male Chimpanzee Skull Landmark Data*

Description

This dataset, Chimp_m, is an array containing landmark data of male chimpanzee skulls. It includes 8 landmarks in 2 dimensions for 28 individuals. All values are preserved as in the original source, with no modifications.

Usage

```
data(Chimp_m)
```

Format

An array with dimensions $8 \times 2 \times 28$:

[1:8, 1:2, 1:28] Numeric values representing 8 landmarks in 2 dimensions for 28 male chimpanzee skulls

Details

The dataset name has been kept as 'Chimp_m' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the shapes package version 1.2.7

`Crab_morpho` *Morphological Data of Crabs*

Description

This dataset, Crab_morpho, is a data frame containing morphological measurements of crabs. The dataset includes 173 observations and 5 variables, covering sex, color, spine count, body width, and weight. All values are preserved as in the original source, with no modifications.

Usage

```
data(Crab_morpho)
```

Format

A data frame with 173 observations and 5 variables:

color Color score of the crab (integer)

spine Number of spines (integer)

weight Body weight (numeric)

width Body width (numeric)

satellites Sex of the crab (integer)

Details

The dataset name has been kept as 'Crab_morpho' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified.

Source

Data taken from the condvis package version 0.5-1

`DogsLife_df` *Dog Breed Life History Data*

Description

This dataset, DogsLife_df, is a data frame containing life history information for various dog breeds. The dataset includes 73 observations and 5 variables, covering breed, average weight, litter size, birth weight, and lifespan. All values are preserved as in the original source, with no modifications.

Usage

```
data(DogsLife_df)
```

Format

A data frame with 73 observations and 5 variables:

Breed Name of the dog breed (character)

Weight Average weight of the breed (numeric)

LitterSize Average litter size (numeric)

BirthWeight Average birth weight (numeric)

Lifespan Average lifespan in years (numeric)

Details

The dataset name has been kept as 'DogsLife_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the SRMData package version 1.0.2

Elephants_df — *Elephant Morphometric Measurements*

Description

This dataset, Elephants_df, is a data frame containing physical measurements of elephants. The dataset includes 1,470 observations and 5 variables, covering sex, age, chest circumference, height, and mass. All values are preserved as in the original source, with no modifications.

Usage

```
data(Elephants_df)
```

Format

A data frame with 1,470 observations and 5 variables:

Sex Sex of the elephant (character)

Age Age in years (numeric)

Chest Chest circumference (numeric)

Height Height (numeric)

Mass Mass (numeric)

Details

The dataset name has been kept as 'Elephants_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the SRMData package version 1.0.2

Extinct_mammals_tbl_df — *Extinct Mammals Data*

Description

This dataset, Extinct_mammals_tbl_df, is a data frame containing information on 85 extinct mammal species. It includes common and binomial names, taxonomic order, date of extinction, former geographic range, and a flag indicating the presence of a picture. The data were read from a CSV file and no modifications have been made.

Usage

```
data(Extinct_mammals_tbl_df)
```

Format

A data frame with 85 observations and 6 variables:

Common.name Common name of the species (character)

Binomial.name Scientific (binomial) name of the species (character)

Order Taxonomic order (character)

Date.of.extinction Date or year of extinction (character)

Former.range Former geographic range (character)

Picture Logical flag indicating presence of a picture (logical)

Details

The dataset name has been kept as 'Extinct_mammals_tbl_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'tbl_df' indicates that the dataset is a data frame/tibble. The original content has not been modified.

Source

Data taken from Kaggle: https://www.kaggle.com/datasets/battle11king/extinct-mammals

---

Fish_metallo *Metallogorgia melanotrichos Sampling Locations*

Description

This dataset, Fish_metallo, is a data frame containing sampling locations for the deep-sea octocoral species Metallogorgia melanotrichos. The dataset includes 38 observations and 3 variables, covering longitude, latitude, and depth in meters. All values are preserved as in the original source, with no modifications.

Usage

```
data(Fish_metallo)
```

Format

A data frame with 38 observations and 3 variables:

lon Longitude of sampling location (numeric)

lat Latitude of sampling location (numeric)

depth Depth of sampling location in meters (integer)

Details

The dataset name has been kept as 'Fish_metallo' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' is not used here because the dataset is identified by its function within BioWorldR. The original content has not been modified.

Source

Data taken from the marmap package version 1.0.12

`flwr_time` *Flowering Time Observations*

Description

This dataset, flwr_time, is a data frame containing flowering observations for two plant species (Willow and Skypilot), along with minimum temperature and altitude at the observation sites. The dataset includes 25 observations and 4 variables. All values are preserved as in the original source, with no modifications.

Usage

```
data(flwr_time)
```

Format

A data frame with 25 observations and 4 variables:

Willow Flowering count or observation for Willow (integer)

Skypilot Flowering count or observation for Skypilot (integer)

MinTemp Minimum temperature at site (numeric)

Altitude Altitude of observation site (numeric)

Details

The dataset name has been kept as 'flwr_time' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified.

Source

Data taken from the SRMData package version 1.0.2

`Germ_lines_df` *Germline Sequence Information for Various Species*

Description

This dataset, Germ_lines_df, is a data frame containing germline sequence information for multiple species. The dataset includes 7,285 observations and 8 variables, covering the nucleotide sequence, gene names, species, chain type, accession number, official species name, source, and strain. All values are preserved as in the original source, with no modifications.

Usage

```
data(Germ_lines_df)
```

Format

A data frame with 7,285 observations and 8 variables:

sequence Germline nucleotide sequence (factor)

names Gene name(s) (character string)

species Species code or name (character string)

chain Chain type (factor)

accession Accession number (factor)

species_name_official Official species name (character string)

source Source of the germline sequence (character string)

strain Strain or line information (character string)

Details

The dataset name has been kept as 'Germ_lines_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the VDJgermlines package version 0.1

`Gorillas_df` *Gorilla Chest-beating Rates and Physical Measurements*

Description

This dataset, Gorillas_df, is a data frame containing measurements related to chest-beating behavior in gorillas. The dataset includes 25 observations and 7 variables, covering physical measurements, focal time, age, sex, and number of chest beats. All values are preserved as in the original source, with no modifications.

Usage

```
data(Gorillas_df)
```

Format

A data frame with 25 observations and 7 variables:

BackBreadth Back breadth measurement (numeric)

ChestBeatRate Rate of chest beats (numeric)

FocalTime Focal observation time (numeric)

Male Sex of the individual (character)

NoChestBeats Number of chest beats observed (integer)

Age Age of the individual (numeric)

Age20 Indicator if age is above 20 (character)

Details

The dataset name has been kept as 'Gorillas_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the SRMData package version 1.0.2

Gorilla_f — *Female Gorilla Skull Landmark Data*

Description

This dataset, Gorilla_f, is an array containing landmark data of female gorilla skulls. It includes 8 landmarks in 2 dimensions for 30 individuals. All values are preserved as in the original source, with no modifications.

Usage

```
data(Gorilla_f)
```

Format

An array with dimensions $8 \times 2 \times 30$:

[1:8, 1:2, 1:30] Numeric values representing 8 landmarks in 2 dimensions for 30 female gorilla skulls

Details

The dataset name has been kept as 'Gorilla_f' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the shapes package version 1.2.7

Gorilla_m — *Male Gorilla Skull Landmark Data*

Description

This dataset, Gorilla_m, is an array containing landmark data of male gorilla skulls. It includes 8 landmarks in 2 dimensions for 29 individuals. All values are preserved as in the original source, with no modifications.

Usage

```
data(Gorilla_m)
```

Format

An array with dimensions $8 \times 2 \times 29$:

[1:8, 1:2, 1:29] Numeric values representing 8 landmarks in 2 dimensions for 29 male gorilla skulls

Details

The dataset name has been kept as 'Gorilla_m' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the shapes package version 1.2.7

`Orang_f` *Female Orangutan Skull Landmark Data*

Description

This dataset, Orang_f, is an array containing landmark data of female orangutan skulls. It includes 8 landmarks in 2 dimensions for 24 individuals. All values are preserved as in the original source, with no modifications.

Usage

```
data(Orang_f)
```

Format

An array with dimensions $8 \times 2 \times 24$:

[1:8, 1:2, 1:24] Numeric values representing 8 landmarks in 2 dimensions for 24 female orangutan skulls

Details

The dataset name has been kept as 'Orang_f' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the shapes package version 1.2.7

Orang_m — *Male Orangutan Skull Landmark Data*

Description

This dataset, Orang_m, is an array containing landmark data of male orangutan skulls. It includes 8 landmarks in 2 dimensions for 30 individuals. All values are preserved as in the original source, with no modifications.

Usage

```
data(Orang_m)
```

Format

An array with dimensions $8 \times 2 \times 30$:

[1:8, 1:2, 1:30] Numeric values representing 8 landmarks in 2 dimensions for 30 male orangutan skulls

Details

The dataset name has been kept as 'Orang_m' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the shapes package version 1.2.7

Plants_list — *Plant Species Occurrence in the US and Canada*

Description

This dataset, Plants_list, is a list containing data on plant species and the states in the United States and Canada where they occur. The list includes 33,721 plant species entries, each with its corresponding occurrence information. All values are preserved as in the original source, with no modifications.

Usage

```
data(Plants_list)
```

Format

A list with 33,721 elements:

abelia Species information (character string)

abelia_x_grandiflora Species information (character string)

abelmoschus Species information (character string)

... Additional species entries omitted for brevity

Details

The dataset name has been kept as 'Plants_list' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the eulerr package version 7.0.2

`Plants_occ` *Plant Occurrence Records*

Description

This dataset, Plants_occ, is a data frame containing occurrence records for 7 plant species, including species name and geographic coordinates. The dataset includes 1,658 observations and 3 variables. All values are preserved as in the original source, with no modifications.

Usage

```
data(Plants_occ)
```

Format

A data frame with 1,658 observations and 3 variables:

species Name of the plant species (character)

x Longitude coordinate (numeric)

y Latitude coordinate (numeric)

Details

The dataset name has been kept as 'Plants_occ' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified.

Source

Data taken from the florabr package version 1.3.0

`Reptiles_df` *Reptile Species Names and URLs*

Description

This dataset, Reptiles_df, is a data frame containing the valid names and URL addresses for all reptile species cataloged in The Reptile Database. The dataset includes 12,440 observations and 8 variables, covering taxonomic order, suborder, family, genus, species, year of description, author, and the URL reference. All values are preserved as in the original source, with no modifications.

Usage

```
data(Reptiles_df)
```

Format

A data frame with 12,440 observations and 8 variables:

order Taxonomic order (character string)

suborder Taxonomic suborder (character string)

family Taxonomic family (character string)

genus Genus name (character string)

species Species name (character string)

year Year of species description (character string)

author Author(s) of the species description (character string)

url URL reference for the species entry (character string)

Details

The dataset name has been kept as 'Reptiles_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the letsHerp package version 0.1.0

Savannah_trees — *Savannah River Tree Locations and Species*

Description

This dataset, Savannah_trees, is a data frame containing the locations and species classification of trees in a plot in the Savannah River, South Carolina, USA. The dataset includes 734 observations and 4 variables, covering spatial coordinates, tree vitality, and species. All values are preserved as in the original source, with no modifications.

Usage

```
data(Savannah_trees)
```

Format

A data frame with 734 observations and 4 variables:

x X-coordinate of the tree location (numeric)

y Y-coordinate of the tree location (numeric)

live Tree vitality indicator (integer)

sp Species classification (factor)

Details

The dataset name has been kept as 'Savannah_trees' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' is not used here because the dataset is identified by its function within BioWorldR. The original content has not been modified.

Source

Data taken from the nnspat package version 0.1.2

Snakes_df — *Snake Morphological Measurements*

Description

This dataset, Snakes_df, is a data frame containing morphological measurements for snakes, including diet (presence of crayfish), sex, snout-vent length (SVL), and teeth count. The dataset includes 200 observations and 4 variables. All values are preserved as in the original source, with no modifications.

Usage

```
data(Snakes_df)
```

Format

A data frame with 200 observations and 4 variables:

Crayfish Presence of crayfish in diet (character)

Sex Sex of the snake (character)

SVL Snout-vent length (numeric)

Teeth Number of teeth (integer)

Details

The dataset name has been kept as 'Snakes_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a data frame. The original content has not been modified.

Source

Data taken from the SRMData package version 1.0.2

---

sparrow_morph *Sparrow Morphological Data*

Description

The dataset sparrow_morph was extracted from Hermon Bumpus (1898) and records morphological variables in sparrows that survived or perished after a severe storm. It is one of the earliest examples of a natural selection dataset in ecological studies.

Usage

```
data(sparrow_morph)
```

Format

A data frame with 49 observations and 6 variables:

Survivorship Survival status of the sparrow (factor with 2 levels: survived or perished)

Total_length Total body length (numeric)

Alar_extent Wing (alar) extent (numeric)

L_beak_head Length from beak to head (numeric)

L_humerus Length of humerus (numeric)

L_keel_sternum Length of keel of sternum (numeric)

Source

Data extracted from Hermon Bumpus (1898), as reproduced in multiple ecological datasets.

`Tsd_turtle` *Turtle Temperature-Dependent Sex Determination Data*

Description

This dataset, Tsd_turtle, is a data frame containing extensive information on temperature-dependent sex determination (TSD) in reptiles. The dataset includes 2,738 observations and 48 variables, covering species, location, incubation temperatures, egg and hatchling measurements, sex ratios, clutch information, and references. All values are preserved as in the original source, with no modifications.

Usage

```
data(Tsd_turtle)
```

Format

A data frame with 2,738 observations and 48 variables:

Species Species name (factor)

Country Country of observation (factor)

Area Specific area or locality (character string)

Longitude Longitude of the location (numeric)

Latitude Latitude of the location (numeric)

Subspecies Subspecies information (character string)

RMU.2010 RMU code (2010 version) (factor)

RMU.2023 RMU code (2023 version) (factor)

Incubation.temperature.set Incubation temperature set experimentally (numeric)

Incubation.temperature.recorded Recorded incubation temperature (numeric)

Incubation.temperature.corrected Corrected incubation temperature (numeric)

Duplicated.data Indicator for duplicated data (logical)

Duplicate Duplicate identifier (character string)

Incubation.temperature.Constant Constant temperature indicator (logical)

Incubation.temperature.Accuracy Accuracy of incubation temperature (numeric)

Incubation.temperature.SD Standard deviation of incubation temperature (numeric)

Incubation.temperature.Amplitude Amplitude of incubation temperature (numeric)

2ndThird.Incubation.temperature.Amplitude Amplitude during 2nd/3rd incubation period (numeric)

Correction.factor Correction factor applied (numeric)

Egg.mass.mean Mean egg mass (numeric)

Egg.mass.sd Standard deviation of egg mass (numeric)

IP.min Minimum incubation period (numeric)

IP.max Maximum incubation period (numeric)

IP.mean Mean incubation period (numeric)

IP.SD Standard deviation of incubation period (numeric)

IP.SE Standard error of incubation period (numeric)

Length.hatchlings.mean Mean length of hatchlings (numeric)

Length.hatchlings.sd Standard deviation of hatchling length (numeric)

SCL.hatchlings.mean Mean straight carapace length of hatchlings (numeric)

SCL.hatchlings.sd Standard deviation of hatchling carapace length (numeric)

Mass.hatchlings.mean Mean hatchling mass (numeric)

Mass.hatchlings.sd Standard deviation of hatchling mass (numeric)

Total Total eggs recorded (numeric)

Hatched Number of hatched eggs (numeric)

NotHatched Number of unhatched eggs (numeric)

Undeveloped Number of undeveloped eggs (numeric)

Unidentified Number of unidentified eggs (numeric)

Intersexes Number of intersex hatchlings (numeric)

Males Number of male hatchlings (numeric)

Females Number of female hatchlings (numeric)

Sexed Number of sexed hatchlings (numeric)

Number.clutch Number of clutches observed (numeric)

Clutch Clutch identifier (factor)

Box Box or container identifier (character string)

Reference Reference for the data (character string)

Note Additional notes (character string)

Digital_Identifier Digital identifier (character string)

Version Dataset version (Date)

Details

The dataset name has been kept as 'Tsd_turtle' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The suffix 'df' is not used here because the dataset is identified by its function within BioWorldR. The original content has not been modified.

Source

Data taken from the embryogrowth package version 10.3

`Turtles_list` *Turtle Carapace and Clutch Size Data*

Description

This dataset, Turtles_list, is a list containing data to examine the effect of turtle carapace length on clutch size. The list includes three main elements: traits, phylo, and traits_info. All values are preserved as in the original source, with no modifications.

Usage

```
data(Turtles_list)
```

Format

A list with 3 elements:

traits A tibble with 240 observations and 17 variables, including species, morphometric traits (M1–M5), body measurements (B1–B2), life-history traits (LH1–LH2), and clutch size traits (CS1–CS2)

phylo A list of 4 elements containing phylogenetic information (phylo structure)

traits_info A data frame with 16 observations and 2 variables: trait_id and trait_name

Details

The dataset name has been kept as 'Turtles_list' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the BioWorldR package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

Source

Data taken from the phylosignalDB package version 0.2.2

Index

Apes_list, 2

Bees_data, 3
BioWorldR, 5
BioWorldR-package *(BioWorldR)*, 5
Birds_Peru, 5
Brain_animals, 6

Chimp_f, 7
Chimp_m, 7
Crab_morpho, 8

DogsLife_df, 9

Elephants_df, 10
Extinct_mammals_tbl_df, 10

Fish_metallo, 11
flwr_time, 12

Germ_lines_df, 13
Gorilla_f, 15
Gorilla_m, 15
Gorillas_df, 14

Orang_f, 16
Orang_m, 17

Plants_list, 17
Plants_occ, 18

Reptiles_df, 19

Savannah_trees, 20
Snakes_df, 20
sparrow_morph, 21

Tsd_turtle, 22
Turtles_list, 24