

# Package: birdsize (via r-universe)

October 28, 2024

**Title** Estimate Avian Body Size Distributions

**Version** 0.0.0.9000

**Date** 2023-02-27

**Description** Generate estimated body size distributions for populations or communities of birds, given either species ID or species' mean body size. Designed to work naturally with the North American Breeding Bird Survey, or with any dataset of bird species, abundance, and/or mean size data.

**License** MIT + file LICENSE

**URL** <https://github.com/diazrenata/birdsize>

**BugReports** <https://github.com/diazrenata/birdsize/issues>

**Depends** R (>= 2.10)

**Imports** stats, truncnorm

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

**VignetteBuilder** knitr

**Config/testthat/edition** 3

**Encoding** UTF-8

**LazyData** true

**Roxygen** list(markdown = TRUE)

**RoxygenNote** 7.2.3

**Repository** <https://ropensci.r-universe.dev>

**RemoteUrl** <https://github.com/ropensci/birdsize>

**RemoteRef** main

**RemoteSha** 78d7840018d750a5917547e71843da953a3a89cb

## Contents

community_generate . . . . .	2
demo_route_clean . . . . .	3
demo_route_raw . . . . .	4
filter_bbs_survey . . . . .	5
individual_metabolic_rate . . . . .	6
known_species . . . . .	7
pop_generate . . . . .	7
species_define . . . . .	9
species_lookup . . . . .	10
toy_aou_community . . . . .	10
toy_size_community . . . . .	11
toy_species_name_community . . . . .	11

<b>Index</b>	<b>12</b>
--------------	-----------

---

community_generate	<i>Simulate individual measurements for many populations</i>
--------------------	--

---

### Description

For a community (i.e. a collection of populations of different species, or of the same species at different points in time or locations, etc), simulate individual-level size and metabolic rate measurements.

### Usage

```
community_generate(
  community_data_table,
  abundance_column_name = "speciestotal"
)
```

### Arguments

`community_data_table`  
dataframe containing at least one of `AOU`, `scientific_name`, or `mean_size` and a column for species abundances

`abundance_column_name`  
character, the name of the column with species abundances. Defaults to "speciestotal".

### Value

a dataframe one row per individual, all columns from `community_data_table`, and additional columns for species attributes.

Specifically:

- AOU: the AOU, if provided
- sim\_species\_id: the sim\_species\_id if provided
- genus: the genus associated with the AOU if provided, or the genus if provided
- species: the species associated with the AOU if provided, or the species if provided
- individual\_mass: the simulated body mass (in grams) for this individual
- individual\_bmr: the simulated basal metabolic rate for this individual
- mean\_size: the mean body mass for this species (i.e. the parameter used for simulation)
- sd\_size: the standard deviation of body mass for this species (i.e. the parameter used for simulation)
- abundance: the number of individuals simulated of this species (i.e. parameter used for simulation)
- sd\_method: the method for finding the standard deviation for body mass for this species
- scientific\_name: the scientific name

### Examples

```
demo_community <- community_generate(demo_route_clean)
head(demo_community)
```

---

demo_route_clean	<i>Cleaned data for a hypothetical Breeding Bird Survey route.</i>
------------------	--

---

### Description

This is the cleaned version of [demo\\_route\\_raw](#), a toy dataset for use in vignettes and function testing. It contains all the same column names as would be expected for a Breeding Bird survey route dataset downloaded, e.g. from ScienceBase or the Data Retriever (Pardieck et al. 2019). However, the actual data values are simulated data.

### Usage

```
demo_route_clean
```

### Format

A data frame with 27 rows and 15 variables:

**record\_id** inherited from data downloaded through the Data Retriever

**routedataid** inherited from Pardieck et al. 2018 (as are all following fields). Unique data identification number.

**countrynum** inherited. Three-digit numerical code for country. In these data, the toy countrynum is 900.

**statenum** inherited. Two-digit numerical code for state, province, or territory. For these data, the toy number is 99.

**route** inherited. Three-digit code identifying the route, unique within states. For this dataset, 1.

**rpId** inherited. Three-digit run protocol identifier. Here, set to 101.

**year** inherited. Four-digit year of the survey.

**AOU** inherited. Five-digit species identification number.

**count10** inherited. Total individuals of the species recorded on stops 1-10.

**count20** inherited. Total individuals of the species recorded on stops 11-20.

**count30** inherited. Total individuals of the species recorded on stops 21-30.

**count40** inherited. Total individuals of the species recorded on stops 31-40.

**count50** inherited. Total individuals of the species recorded on stops 40-50.

**stoptotal** inherited. Total number of stops (of 50), where the species was recorded.

**speciestotal** inherited. Total individuals of the species recorded across the entire run of the route (sum of stops).

## Details

Nearly all column names are inherited from Pardieck et al. (2019) and are further explained in the Files and Fields Definitions document included as part of the Breeding Bird Survey data release.

The cleaning process removes unidentified species and those poorly sampled by Breeding Bird Survey methods (using the [filter\\_bbs\\_survey](#) function) and, for these data, filters the year to 1994 for rapid testing.

## References

- Pardieck, K. L., Ziolkowski, D. J., Lutmerding, M., Aponte, V., & Hudson, M.-A. (2019). North American Breeding Bird Survey Dataset 1966—2018, version 2018.0. U.S. Geological Survey. <https://doi.org/10.5066/P9HE8XYJ>

---

demo\_route\_raw

*Raw data for a hypothetical Breeding Bird Survey route.*

---

## Description

A toy dataset for use in vignettes and function testing. It contains all the same column names as would be expected for a Breeding Bird survey route dataset downloaded, e.g. from ScienceBase or the Data Retriever (Pardieck et al. 2019). However, the actual data values are simulated data.

## Usage

demo\_route\_raw

## Format

A data frame with 1160 rows and 15 variables:

**record\_id** inherited from data downloaded through the Data Retriever

**routedataid** inherited from Pardieck et al. 2018 (as are all following fields). Unique data identification number.

**countrynum** inherited. Three-digit numerical code for country. In these data, the toy countrynum is 900.

**statenum** inherited. Two-digit numerical code for state, province, or territory. For these data, the toy number is 99.

**route** inherited. Three-digit code identifying the route, unique within states. For this dataset, 1.

**rpil** inherited. Three-digit run protocol identifier. Here, set to 101.

**year** inherited. Four-digit year of the survey.

**AOU** inherited. Five-digit species identification number.

**count10** inherited. Total individuals of the species recorded on stops 1-10.

**count20** inherited. Total individuals of the species recorded on stops 11-20.

**count30** inherited. Total individuals of the species recorded on stops 21-30.

**count40** inherited. Total individuals of the species recorded on stops 31-40.

**count50** inherited. Total individuals of the species recorded on stops 40-50.

**stoptotal** inherited. Total number of stops (of 50), where the species was recorded.

**speciestotal** inherited. Total individuals of the species recorded across the entire run of the route (sum of stops).

## Details

Nearly all column names are inherited from Pardieck et al. (2019) and are further explained in the Files and Fields Definitions document included as part of the Breeding Bird Survey data release.

## References

- Pardieck, K. L., Ziolkowski, D. J., Lutmerding, M., Aponte, V., & Hudson, M.-A. (2019). North American Breeding Bird Survey Dataset 1966—2018, version 2018.0. U.S. Geological Survey. <https://doi.org/10.5066/P9HE8XYJ>

---

filter\_bbs\_survey

*Clean raw Breeding Bird Survey survey data*

---

## Description

The raw data for the Breeding Bird Survey includes unidentified species and some species that are not well-sampled by the BBS methods. This function filters a dataframe to remove those species.

**Usage**

```
filter_bbs_survey(bbs_survey_data)
```

**Arguments**

```
bbs_survey_data
```

data frame with columns for species and AOU

**Value**

bbs\_survey\_data with unidentified species, nightbirds, waterbirds, non-targets removed

**Examples**

```
head(filter_bbs_survey(demo_route_raw))
```

---

```
individual_metabolic_rate
```

*Estimate individual-level BMR*

---

**Description**

Given an individual's body mass (in grams), use allometric scaling (Fristoe 2015) to estimate basal metabolic rate.

**Usage**

```
individual_metabolic_rate(mass)
```

**Arguments**

```
mass
```

mass in grams

**Value**

estimated basal metabolic rate

**References**

- Fristoe, T. S. (2015). Energy use by migrants and residents in North American breeding bird communities. *Global Ecology and Biogeography*, 24(4), 406–415. <https://doi.org/10.1111/geb.12262>

**Examples**

```
individual_metabolic_rate(10)
```

---

known_species	<i>List of species known to birdsize</i>
---------------	--

---

**Description**

A table of the AOU (Pardieck et al 2019), genus, and species of the 443 species with species-specific data built in to birdsize.

**Usage**

```
known_species
```

**Format**

A data frame with 443 rows and 6 variables:

**AOU** AOU used in Paradiack et al. (2019)

**genus** genus, from Paradiack et al. (2019)

**species** species, from Paradiack et al. (2019)

**References**

- Pardieck, K.L., Ziolkowski Jr., D.J., Lutmerding, M., Aponte, V., and Hudson, M-A.R., 2019, North American Breeding Bird Survey Dataset 1966 - 2018 (ver. 2018.0): U.S. Geological Survey, Patuxent Wildlife Research Center, <https://doi.org/10.5066/P9HE8XYJ>.

---

pop_generate	<i>Simulate body masses for a population</i>
--------------	--

---

**Description**

Draws body mass measurements for a population of birds (of all the same species) given the population size and either (1) the species AOU or (2) the mean and potentially standard deviation of body mass for that species.

**Usage**

```
pop_generate(  
  abundance = NA_integer_,  
  AOU = NA_integer_,  
  scientific_name = NA_character_,  
  mean_size = NA_real_,  
  sd_size = NA_real_,  
  sim_species_id = 1  
)
```

### Arguments

abundance	integer number of individuals to draw. <i>Required.</i>
AOU	the numeric AOU code used for this species in the Breeding Bird Survey
scientific_name	as "Genus species"
mean_size	numeric, mean body mass (in grams) for this species.
sd_size	numeric, standard deviation of body mass for this species.
sim_species_id	defaults AOU or 1

### Details

abundance is required, as well as *one of*: AOU, scientific\_name, or mean\_size.

### Value

a dataframe with abundance rows - one record per individual - and columns for species attributes.

Specifically:

- AOU: the AOU, if provided
- sim\_species\_id: the sim\_species\_id if provided
- scientific\_name: the scientific name if provided
- individual\_mass: the simulated body mass (in grams) for this individual
- individual\_bmr: the simulated basal metabolic rate for this individual
- mean\_size: the mean body mass for this species (i.e. the parameter used for simulation)
- sd\_size: the standard deviation of body mass for this species (i.e. the parameter used for simulation)
- abundance: the number of individuals simulated of this species (i.e. parameter used for simulation)
- sd\_method: the method for finding the standard deviation for body mass for this species

### Examples

```
pop_generate(abundance = 5, AOU = 2881)
pop_generate(abundance = 5, scientific_name = "Selasphorus calliope")
pop_generate(abundance = 5, mean_size = 20, sd_size = 3)
```



---

species_define	<i>Define a species</i>
----------------	-------------------------

---

### Description

Creates a list with taxonomic/identifying information and parameters for mean and standard deviation of body mass.

### Usage

```
species_define(  
  AOU = NA_integer_,  
  scientific_name = NA_character_,  
  mean_size = NA_real_,  
  sd_size = NA_real_,  
  sim_species_id = 1  
)
```

### Arguments

AOU	the numeric AOU code used for this species in the Breeding Bird Survey
scientific_name	the species' scientific name, as "Genus species"
mean_size	mean body size
sd_size	sd of body size
sim_species_id	identifier; if using taxonomic info, defaults to AOU. If not, defaults to 1. Supplying other values can be useful for simulation models.

### Details

The identifying information used depends on which parameters are provided, with the following order of preference: AOU > scientific name > user provided mean and sd > user provided mean and estimated sd.

### Value

list with species parameter information

### Examples

```
species_define(AOU = 2881)  
species_define(scientific_name = "Perdix perdix")  
species_define(mean_size = 400, sd_size = 30)  
species_define(mean_size = 400)
```

---

species\_lookup      *Species lookup*

---

**Description**

Given either AOU or scientific name, looks up a species' taxonomic information and mean and standard deviation of body size in [sd\\_table](#).

**Usage**

```
species_lookup(AOU = NA_integer_, scientific_name = NA_character_)
```

**Arguments**

AOU                    the numeric AOU code used for this species in the Breeding Bird Survey  
scientific\_name        the species' scientific name, as "Genus species"

**Value**

data frame with columns AOU, genus, species, mean\_mass, mean\_sd, contains\_estimates, scientific\_name

**Examples**

```
species_lookup(AOU = 2881)
species_lookup(scientific_name = "Selasphorus calliope")
```

---

toy\_aou\_community      *Toy data frame of abundances and AOU's (for vignettes)*

---

**Description**

This data table is a toy data frame for the vignettes. It has abundances and AOU codes for 5 species to make up a hypothetical community.

**Usage**

```
toy_aou_community
```

**Format**

A data frame with 5 rows and 2 variables:

**AOU** AOU

**abundance** Number of individuals to simulate masses for

---

toy\_size\_community      *Toy data frame of abundances and species mean sizes (for vignettes)*

---

**Description**

This data table is a toy data frame for the vignettes. It has abundances and mean body sizes for 5 species to make up a hypothetical community.

**Usage**

```
toy_size_community
```

**Format**

A data frame with 5 rows and 3 variables:

**mean\_size** Mean mass, in g

**abundance** Number of individuals to simulate masses for

**sim\_species\_id** ID

---

toy\_species\_name\_community  
*Toy data frame of abundances and species names (for vignettes)*

---

**Description**

This data table is a toy data frame for the vignettes. It has abundances and scientific names for 5 species to make up a hypothetical community.

**Usage**

```
toy_species_name_community
```

**Format**

A data frame with 5 rows and 2 variables:

**scientific\_name** Scientific name

**abundance** Number of individuals to simulate masses for

# Index

## \* datasets

- demo\_route\_clean, 3
- demo\_route\_raw, 4
- toy\_aou\_community, 10
- toy\_size\_community, 11
- toy\_species\_name\_community, 11

## \* export

- known\_species, 7

community\_generate, 2

demo\_route\_clean, 3

demo\_route\_raw, 3, 4

filter\_bbs\_survey, 4, 5

individual\_metabolic\_rate, 6

known\_species, 7

pop\_generate, 7

sd\_table, 10

species\_define, 9

species\_lookup, 10

toy\_aou\_community, 10

toy\_size\_community, 11

toy\_species\_name\_community, 11