Taxon Profiles Data Inclusion Criteria

Contact Species360 Support

This document explains what data is included and excluded from the analysis for Taxon Profiles. The source of the data included in Taxon Profiles is anonymous, and specific Institutions are not indicated. Taxon Profiles first iteration will include Weight reports, but Survival and Reproduction reports will be included in future iterations of this resource.

Topics

Data Rules for All analysis

All data included in Taxon Profiles are Species360 records extracted from ZIMS for Husbandry. Strict guidelines were followed to ensure accurate species level representation. This may involve excluding certain records and data points. See below for details on the exclusion criteria and any data transformation for this resource.

- Animals were excluded from analysis if:
  - Animal ever spent time in a private (local) collection
  - Animal is an incomplete accession
  - Taxonomy is ranked at family or higher
  - Birth date earlier than all sires' birth dates
  - Birth date earlier than all dams' birth dates
  - Death date earlier than animal birth date
  - Animal has no associated collection moves

- For animals included in analysis, specific data points are excluded if:
  - For collection, weight, length, and health status, exclude if the entry is dated before the animal's birth, lay, conception, or first transaction date
  - Contraception records entered before the animal birth date are excluded
  - The entry is dated in the future
  - Data is flagged as provisional
  - Value is flagged to be "excluded from norms" during data entry

Data Transformation

- All GANs and institution IDs are encrypted
- Obsolete taxonomy are mapped to the current taxonomy name
- Weight units of measurement are converted to kg
- Length unit of measurements are converted to cm

Data Updates

The data will be refreshed every three months. Updates will be staggered based on taxonomic class, so different taxa will update on a rolling schedule, the grouping and cadence is as follows:

- Week 0: Non-vertebrates
- Week 3: Fish
- Week 5: Amphibians
- Week 7: Reptiles
- Week 9: Birds
- Week 11: Mammals

Back to topics

Data Rules for Weights analysis

After initial data validation (described above), additional criteria specific to weight measurements are applied.

Data Set Selection

- Animal exclusion criteria:
  - Taxon (or below) has only been held by one member since 1990
  - Animal type is not an individual
  - Animal birth date estimate that is:
    - Undetermined
    - Indeterminate
    - Approximately Before
    - Approximately After
    - Approximate Variance is greater than or equal to 365 days
    - Range is greater than or equal to 365 days
  - Animal taxonomy rank is genus or higher
- Weight value exclusion criteria:
  - Measurement date is before 1 January 1990. This date was chosen to include modern care practices and more standardized data entry protocols within institutions.
  - Measurement date is after the animal's death date
  - Measurement date is before the animal's birth date
  - Carcass/dead weight
  - Organ weight
  - Egg weight
  - Antler weight
  - Estimated weight (includes medical weights)
  - Measurement date is:
    - Undetermined
    - Indeterminate
    - Approximately Before
    - Approximately After
    - Approximate Variance is greater than or equal to 31 days
    - Range is greater than or equal to 31 days
  - Animal has an abnormal health status
  - Measurement is:
    - Less than zero
    - Null
    - Has no unit of measurement

Methods:

- Once the dataset was defined, weights were sorted based on the age in months.

- The graph shows a line for the calculated mean weight for each month of age with a shaded area to represent one standard deviation above and below the mean.
- If animals were weighed multiple times within an age range, those weights were averaged so one individual is not overrepresented.
- Weight outliers were identified using Tukey's method (1.5 times the interquartile range to define upper and lower fences) for each age range.
- Weight outliers were further defined for the entire species data set using the median age-in-months to approximate an "age of maturity" for the species. Tukey's method was used to define the upper and lower fences for all weights from all animals older than the approximate age of maturity. Weights that fall above the upper fence for immature animals were removed from analysis. Weight that fall outside both upper and lower fences were removed from mature animals. This accommodates inclusion of weights from young, growing animals but still provides outlier detection for entries made with incorrect units of measurement for animals that are mature.
- Age outliers were identified using Tukey's method to define the upper fence for all ages represented in the data set. Ages greater than the upper fence were removed from the analysis.
- Institutional data: *coming soon!*
- Increased granularity for the first month of life: *coming soon!*

Back to topics

Data Rules for Reproduction analysis

Reproduction reports are coming soon!

Data Rules for Survival analysis

Survival reports are coming soon!

Back to topics

Revised 9th September 2024