

Assignment 8

Ellen Bledsoe

2024-03-12

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.4.4      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.1
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

2. Portal Data Review (25 points)

```
## Rows: 35549 Columns: 9
## -- Column specification -----
## Delimiter: ","
## chr (2): species_id, sex
## dbl (7): record_id, month, day, year, plot_id, hindfoot_length, weight
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 54 Columns: 4
## -- Column specification -----
## Delimiter: ","
## chr (4): species_id, genus, species, taxa
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 24 Columns: 2
## -- Column specification -----
## Delimiter: ","
## chr (1): plot_type
## dbl (1): plot_id
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.

## [1] "2a"

## # A tibble: 3,027 x 5
##   year month   day species_id weight
```

```
##      <dbl> <dbl> <dbl> <chr>      <dbl>
## 1 1977      8     19 D0          52
## 2 1977     10     17 D0          33
## 3 1977     10     17 D0          50
## 4 1977     10     17 D0          48
## 5 1977     10     17 D0          31
## 6 1977     10     18 D0          41
## 7 1977     11     12 D0          44
## 8 1977     11     12 D0          48
## 9 1977     11     14 D0          39
## 10 1977     12     10 D0         40
## # i 3,017 more rows
```

```
## [1] "2b"
```

```
## # A tibble: 5,150 x 3
##   year species_id hindfoot_length
##   <dbl> <chr>          <dbl>
## 1 1995 PP          23
## 2 1995 PP          22
## 3 1995 PP          22
## 4 1995 PP          21
## 5 1995 PP          21
## 6 1995 PP          20
## 7 1995 PP          22
## 8 1995 PP          24
## 9 1995 PP          22
## 10 1995 PP          22
## # i 5,140 more rows
```

```
## [1] "2c"
```

```
## 'summarise()' has grouped output by 'species_id'. You can override using the
## '.groups' argument.
```

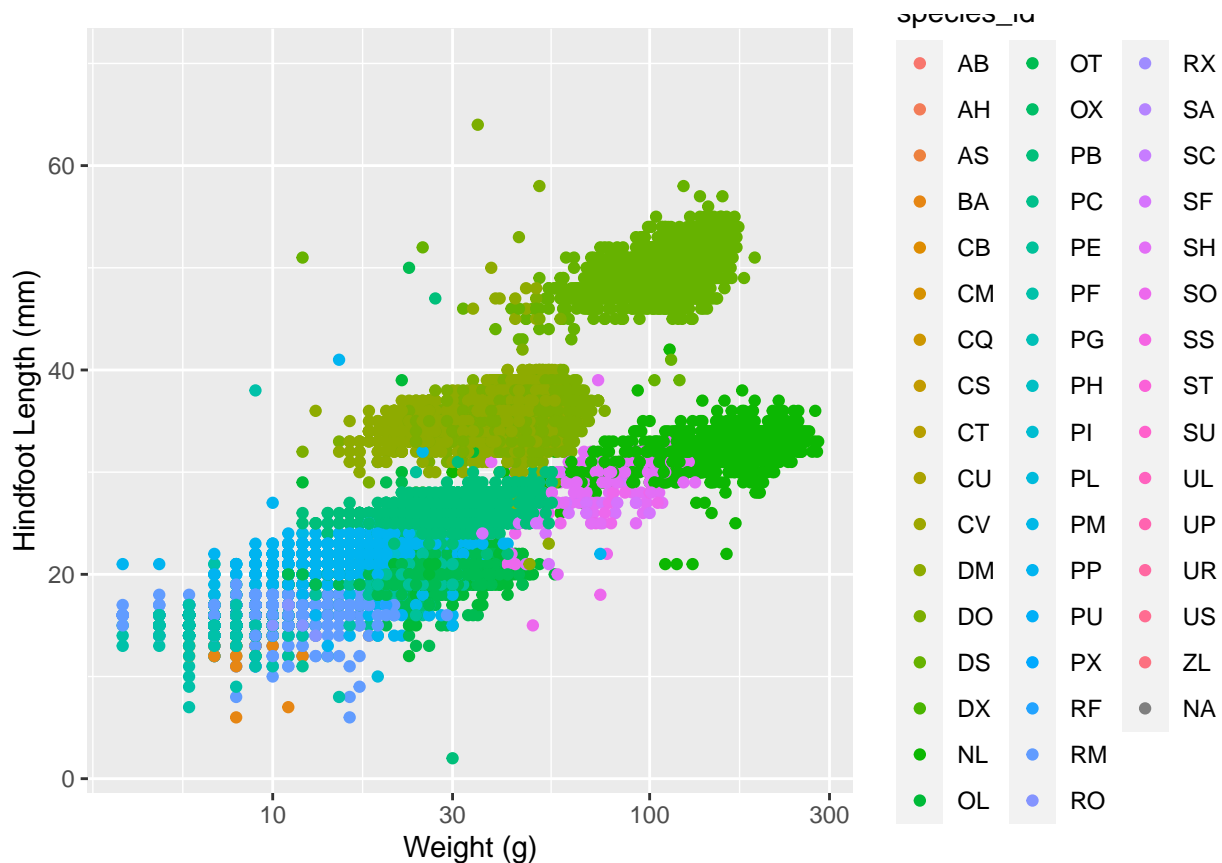
```
## # A tibble: 340 x 3
## # Groups:   species_id [25]
##   species_id year mean_hf
##   <chr>      <dbl> <dbl>
## 1 AH        1999    35
## 2 AH        2000    31
## 3 BA        1989    13
## 4 BA        1990   13.8
## 5 BA        1991   12.9
## 6 BA        1992    12
## 7 DM        1977   35.7
## 8 DM        1978   36.1
## 9 DM        1979   35.9
## 10 DM       1980   35.8
## # i 330 more rows
```

```
## [1] "2d"
```

```
## # A tibble: 16,167 x 5
##   year genus   species   weight plot_type
##   <dbl> <chr>   <chr>     <dbl> <chr>
## 1  1977 Dipodomys merriami      NA Control
## 2  1977 Dipodomys merriami      NA Rodent Exclosure
## 3  1977 Dipodomys merriami      NA Long-term Krat Exclosure
## 4  1977 Dipodomys merriami      NA Spectab exclosure
## 5  1977 Dipodomys merriami      NA Spectab exclosure
## 6  1977 Dipodomys spectabilis    NA Rodent Exclosure
## 7  1977 Dipodomys merriami      NA Rodent Exclosure
## 8  1977 Dipodomys merriami      NA Long-term Krat Exclosure
## 9  1977 Dipodomys merriami      NA Control
## 10 1977 Dipodomys merriami      NA Short-term Krat Exclosure
## # i 16,157 more rows
```

```
## [1] "2e"
```

```
## Warning: Removed 4811 rows containing missing values ('geom_point()').
```

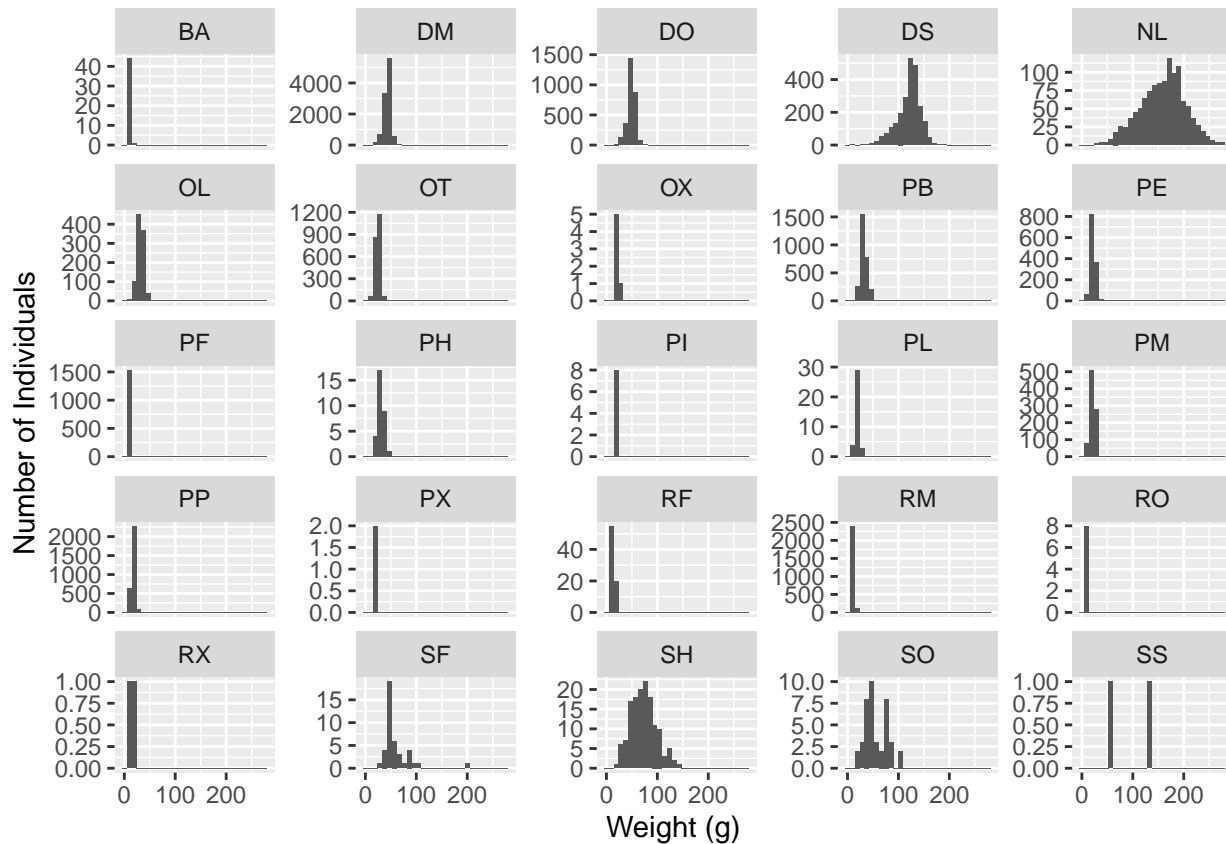


```
## [1] "2f"
```

```
## # A tibble: 32,283 x 9
##   record_id month   day   year plot_id species_id sex hindfoot_length weight
##   <dbl> <dbl> <dbl> <dbl> <dbl> <chr>   <chr>      <dbl> <dbl>
## 1      63     8    19  1977     3 DM      M          35    40
```

```
## 2      64      8      19 1977      7 DM      M      37      48
## 3      65      8      19 1977      4 DM      F      34      29
## 4      66      8      19 1977      4 DM      F      35      46
## 5      67      8      19 1977      7 DM      M      35      36
## 6      68      8      19 1977      8 DO      F      32      52
## 7      69      8      19 1977      2 PF      M      15      8
## 8      70      8      19 1977      3 OX      F      21      22
## 9      71      8      19 1977      7 DM      F      36      35
## 10     74      8      19 1977      8 PF      M      12      7
## # i 32,273 more rows
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```



```
## [1] "2g, optional"
```

```
## Joining with 'by = join_by(plot_id)'
```

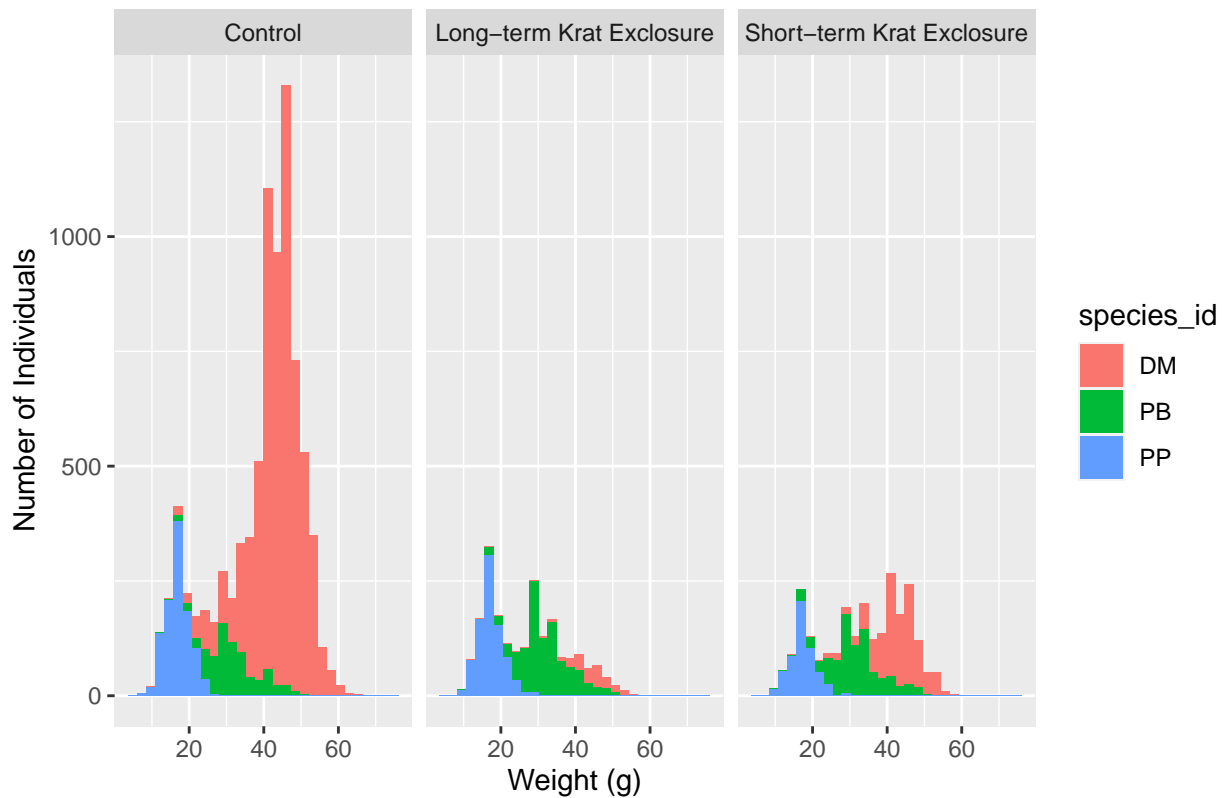
```
## # A tibble: 13,415 x 10
##   record_id month   day  year plot_id species_id sex  hindfoot_length weight
##       <dbl> <dbl> <dbl> <dbl>   <dbl>   <chr>      <chr>          <dbl>   <dbl>
## 1         3     7    16  1977     2 DM        F           37     NA
## 2         5     7    16  1977     3 DM        M           35     NA
## 3        13     7    16  1977     3 DM        M           35     NA
## 4        14     7    16  1977     8 DM       <NA>         NA     NA
## 5        15     7    16  1977     6 DM        F           36     NA
```

```
## 6      16      7      16 1977      4 DM      F      36      NA
## 7      18      7      16 1977      2 PP      M      22      NA
## 8      21      7      17 1977     14 DM      F      34      NA
## 9      23      7      17 1977     13 DM      M      36      NA
## 10     26      7      17 1977     15 DM      M      31      NA
## # i 13,405 more rows
## # i 1 more variable: plot_type <chr>
```

```
## 'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.
```

```
## Warning: Removed 438 rows containing non-finite values ('stat_bin()').
```

Size distribution comparison across treatments



3. Megafaunal Extinction (35 points)

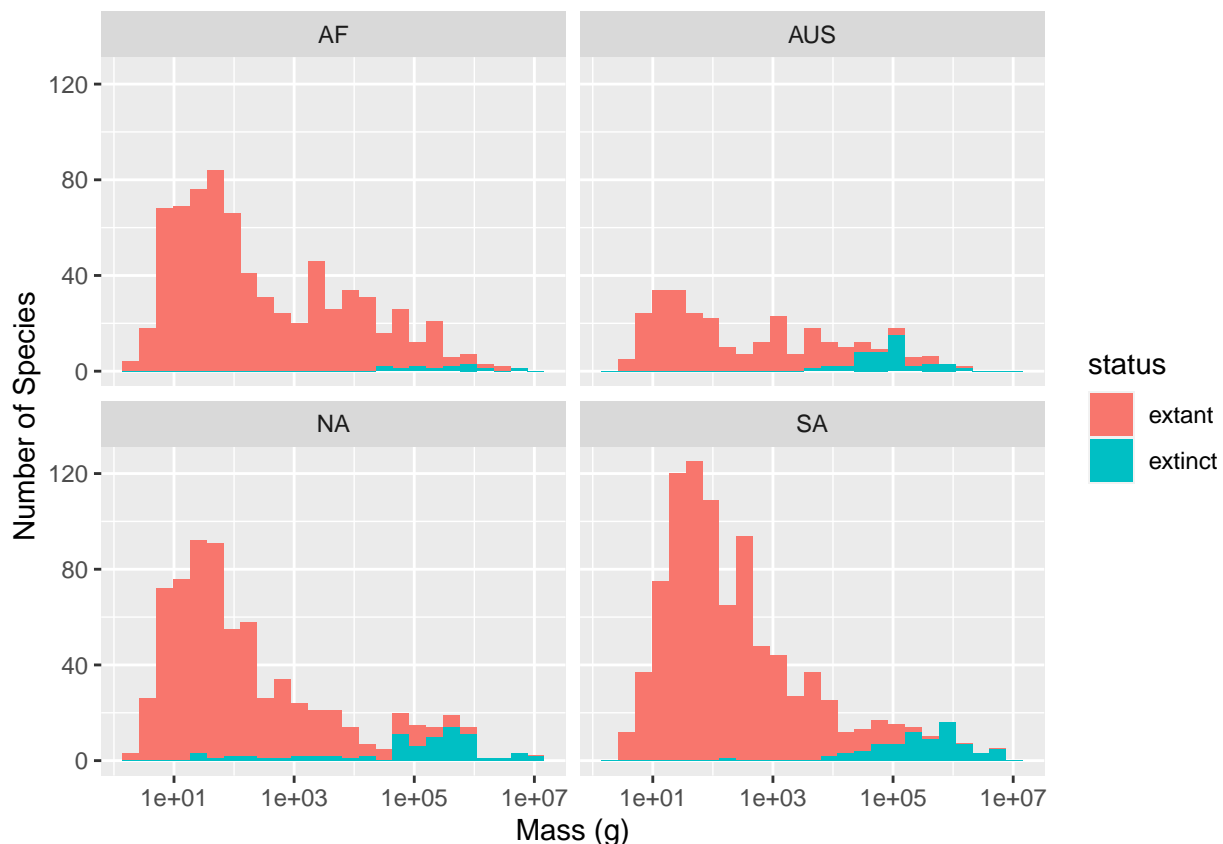
```
## [1] "3a"

## Rows: 5731 Columns: 8
## -- Column specification -----
## Delimiter: "\t"
## chr (7): continent, status, order, family, genus, species, reference
## dbl (1): mass
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

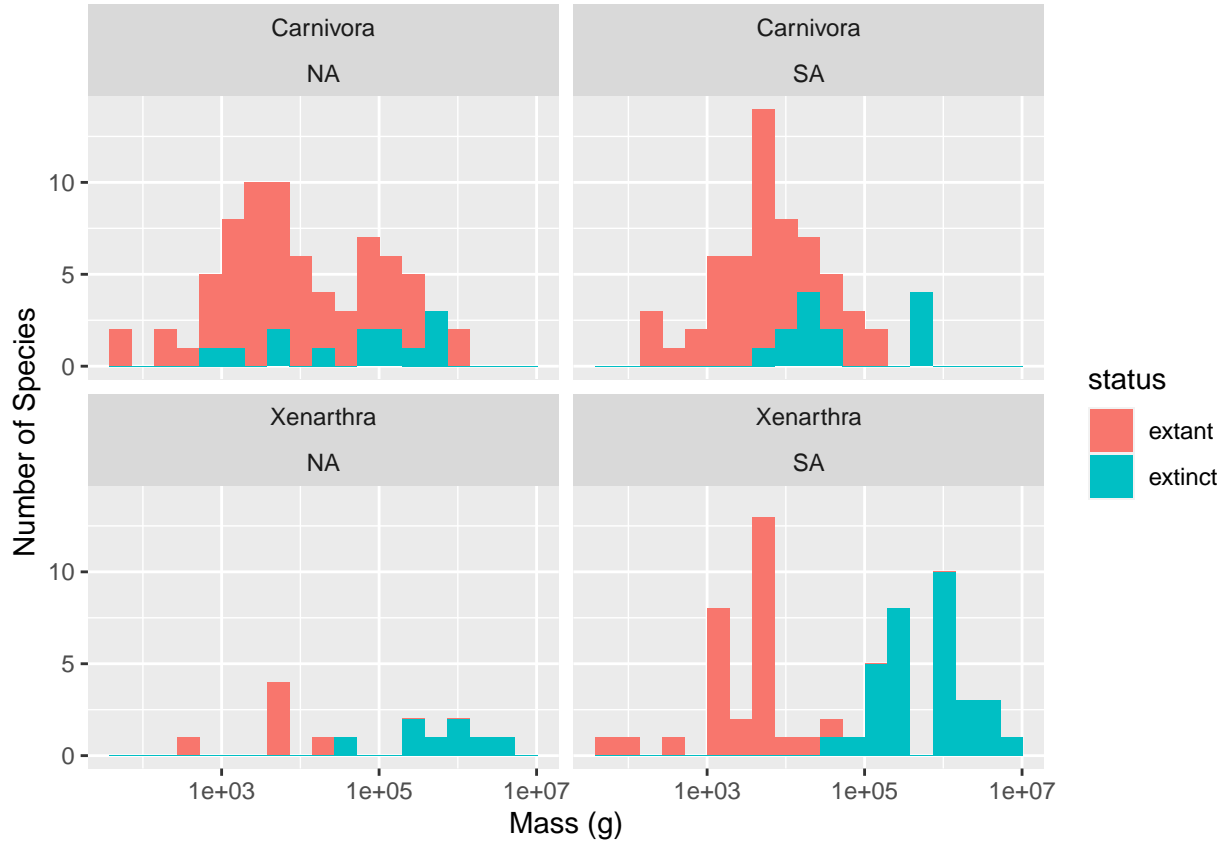
```
## spc_tbl_ [5,731 x 8] (S3: spec_tbl_df/tbl_df/tbl/data.frame)
## $ continent: chr [1:5731] "AF" "AF" "AF" "AF" ...
## $ status   : chr [1:5731] "extant" "extant" "extant" "extant" ...
## $ order    : chr [1:5731] "Artiodactyla" "Artiodactyla" "Artiodactyla" "Artiodactyla" ...
## $ family   : chr [1:5731] "Bovidae" "Bovidae" "Bovidae" "Bovidae" ...
## $ genus    : chr [1:5731] "Addax" "Aepyceros" "Alcelaphus" "Ammodorcas" ...
## $ species  : chr [1:5731] "nasomaculatus" "melampus" "buselaphus" "clarkei" ...
## $ mass     : num [1:5731] 70000 52500 171002 28050 48000 ...
## $ reference: chr [1:5731] "60" "63, 70" "63, 70" "60" ...
## - attr(*, "spec")=
## .. cols(
## ..   continent = col_character(),
## ..   status = col_character(),
## ..   order = col_character(),
## ..   family = col_character(),
## ..   genus = col_character(),
## ..   species = col_character(),
## ..   mass = col_double(),
## ..   reference = col_character()
## .. )
## - attr(*, "problems")=<externalptr>

## [1] "3b"

## Warning: Removed 417 rows containing non-finite values ('stat_bin()').
```

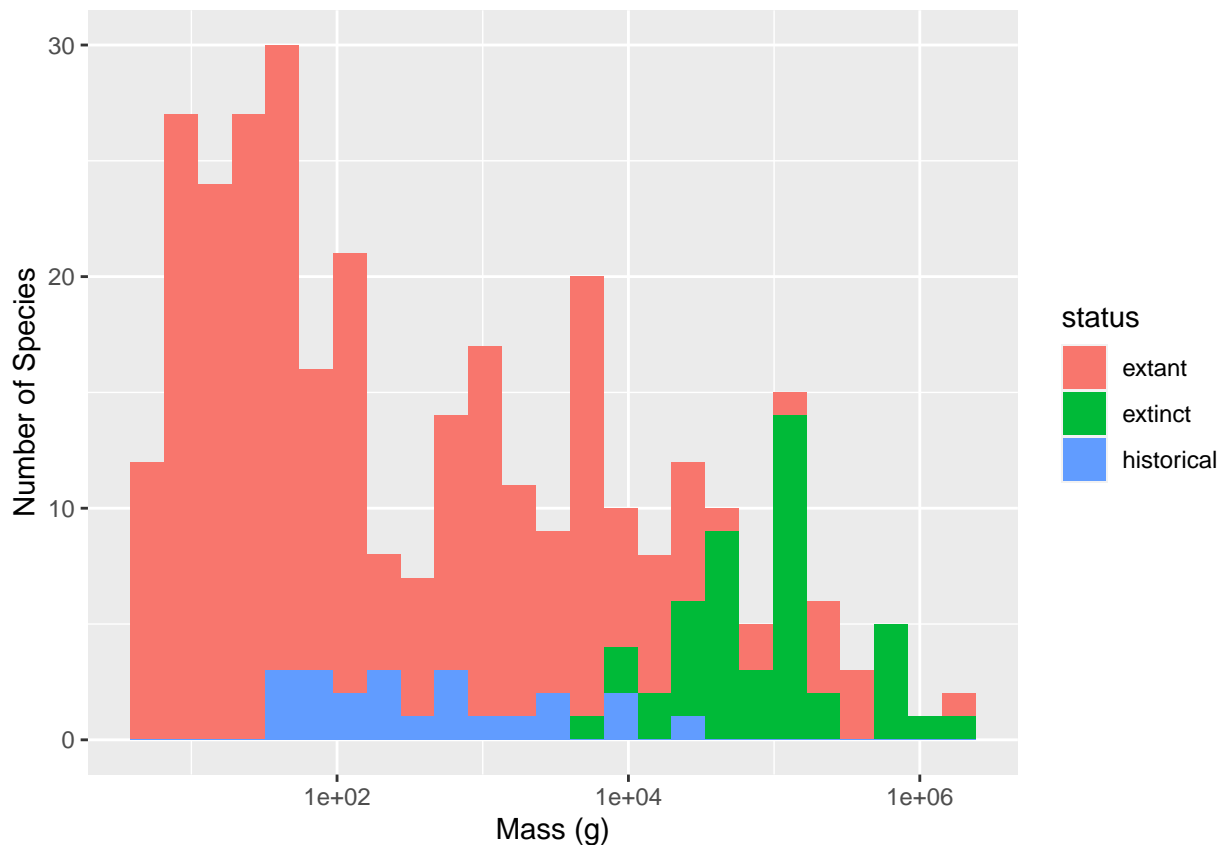


```
## [1] "3c"
```



```
## [1] "3d"
```

```
## Warning: Removed 9 rows containing non-finite values ('stat_bin()').
```



```
## [1] "3e, optional"
```

```
## # A tibble: 5 x 1
##   continent
##   <chr>
## 1 AF
## 2 AUS
## 3 Insular
## 4 NA
## 5 SA
```

```
## Joining with 'by = join_by(continent)'
```

```
## # A tibble: 3,091 x 8
##   continent status order      family genus      species      mass reference
##   <chr>      <chr> <chr>      <chr> <chr>      <chr>      <dbl> <chr>
## 1 AF        extant Artiodactyla Bovidae Addax      nasomacul~ 7.00e4 60
## 2 AF        extant Artiodactyla Bovidae Aepyceros melampus  5.25e4 63, 70
## 3 AF        extant Artiodactyla Bovidae Alcelaphus buselaphus 1.71e5 63, 70
## 4 AF        extant Artiodactyla Bovidae Ammodorcas clarkei  2.80e4 60
## 5 AF        extant Artiodactyla Bovidae Ammotragus lervia  4.80e4 75
## 6 AF        extant Artiodactyla Bovidae Antidorcas marsupial~ 3.90e4 60
## 7 AF        extinct Artiodactyla Bovidae Antidorcas bondi  3.4 e4 1
## 8 AF        extinct Artiodactyla Bovidae Antidorcas australis  4 e4 2
## 9 AF        extant Artiodactyla Bovidae Bos        taurus  9 e5 <NA>
## 10 AF       extant Artiodactyla Bovidae Capra      walie  1 e5 <NA>
## # i 3,081 more rows
```


4. Palmer Penguins (35 points)

```
## # A tibble: 6 x 8
##   species island   bill_length_mm bill_depth_mm flipper_length_mm body_mass_g
##   <fct>   <fct>           <dbl>         <dbl>           <int>       <int>
## 1 Adelie  Torgersen         39.1          18.7            181        3750
## 2 Adelie  Torgersen         39.5          17.4            186        3800
## 3 Adelie  Torgersen         40.3          18             195        3250
## 4 Adelie  Torgersen          NA           NA             NA          NA
## 5 Adelie  Torgersen         36.7          19.3            193        3450
## 6 Adelie  Torgersen         39.3          20.6            190        3650
## # i 2 more variables: sex <fct>, year <int>

## Rows: 152 Columns: 17
## -- Column specification -----
## Delimiter: ","
## chr  (9): studyName, Species, Region, Island, Stage, Individual ID, Clutch C...
## dbl  (7): Sample Number, Culmen Length (mm), Culmen Depth (mm), Flipper Leng...
## date (1): Date Egg
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 124 Columns: 17
## -- Column specification -----
## Delimiter: ","
## chr  (9): studyName, Species, Region, Island, Stage, Individual ID, Clutch C...
## dbl  (7): Sample Number, Culmen Length (mm), Culmen Depth (mm), Flipper Leng...
## date (1): Date Egg
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
## Rows: 68 Columns: 17
## -- Column specification -----
## Delimiter: ","
## chr  (9): studyName, Species, Region, Island, Stage, Individual ID, Clutch C...
## dbl  (7): Sample Number, Culmen Length (mm), Culmen Depth (mm), Flipper Leng...
## date (1): Date Egg
##
## i Use 'spec()' to retrieve the full column specification for this data.
## i Specify the column types or set 'show_col_types = FALSE' to quiet this message.
```

Result from the `setdiff()` function:

```
## # A tibble: 0 x 8
## # i 8 variables: species <chr>, island <chr>, bill_length_mm <dbl>,
## #   bill_depth_mm <dbl>, flipper_length_mm <dbl>, body_mass_g <dbl>, sex <chr>,
## #   year <dbl>
```