

# hw\_5.rmd

Andrew Mikolinski

2023-11-09

## Question 1

```
library("neonDivData")
data("data_plant", package = "neonDivData")
data_plant$genus <- sub("^([A-Za-z]+)\\s.*", "\\1", data_plant$taxon_name)
unique_genus_names <- sort(unique(data_plant$genus))
random_genus_sample <- sample(unique_genus_names, 100)
print(random_genus_sample)
```

```
## [1] "Malus" "Atriplex" "Mimosa"
## [4] "Hemizonella" "Sideritis" "Bowlesia"
## [7] "Cleome" "Saussurea" "Jatropha"
## [10] "Spigelia" "Machaonia" "Scleria"
## [13] "Helianthemum" "Sphenopholis" "Amphicarpaea"
## [16] "Phytolacca" "Bahia" "Pleuraphis"
## [19] "Phalacroseris" "Wyethia" "Baptisia"
## [22] "Stenotus" "Gratiola" "Streptanthus"
## [25] "Dalechampia" "Cynanchum" "Ammi"
## [28] "Vernicia" "Allotropa" "Bignonia"
## [31] "Labordia" "Tradescantia" "Schoenoplectus"
## [34] "Acanthospermum" "Krascheninnikovia" "Brickellia"
## [37] "Euonymus" "Urera" "Oryza"
## [40] "Monotropa" "Schismus" "Mollugo"
## [43] "Androstephium" "Baccharis" "Cyanthillium"
## [46] "Decodon" "Puccinellia" "Melothria"
## [49] "Ctenium" "Rhododon" "Vanilla"
## [52] "Aliciella" "Canella" "Loeflingia"
## [55] "Heterocodon" "Celastrus" "Guazuma"
## [58] "Crassula" "Chasmanthium" "Sorbus"
## [61] "Anemone" "Petradoria" "Allionia"
## [64] "Scirpus" "Zea" "Frasera"
## [67] "Actaea" "Nephrolepis" "Zanthoxylum"
## [70] "Micropus" "Chromolaena" "Symplocarpus"
## [73] "Yucca" "Halesia" "Argemone"
## [76] "Tilia" "Leucaena" "Belamcanda"
## [79] "Cleistes" "Psilostrophe" "Kallstroemia"
## [82] "Littorella" "Gymnocladus" "Olneya"
## [85] "Morelotia" "Logfia" "Hydrocotyle"
## [88] "Thelesperma" "Centaurium" "Desmanthus"
## [91] "Erysimum" "Pholistoma" "Spermolepis"
```

```
## [94] "Astragalus"          "Chimaphila"          "Agropyron"
## [97] "Leucopoa"            "Lepisorus"           "Petiveria"
## [100] "Stylomecon"
```

## Question 2

```
data_plant$taxon_name2 <- sub("^[A-Za-z]+\\s[A-Za-z]+).*", "\\1", data_plant$taxon_name)
unique_species_names <- sort(unique(data_plant$taxon_name2))
random_species_sample <- sample(unique_species_names, 100)
print(random_species_sample)
```

```
## [1] "Lathyrus brachycalyx"          "Proboscidea parviflora"
## [3] "Thuja occidentalis"           "Lespedeza frutescens"
## [5] "Carex stipata"                 "Seymeria pectinata"
## [7] "Clinopodium douglasii"        "Ampelopsis cordata"
## [9] "Eragrostis pectinacea"        "Panicum hallii"
## [11] "Heteropogon contortus"        "Cuscuta glomerata"
## [13] "Carex phaeocephala"           "Grindelia adenodonta"
## [15] "Lemna trisulca"               "Sansevieria spp"
## [17] "Ayeria microphylla"           "Cirsium palustre"
## [19] "Lonicera hispidula"           "Phacelia ramosissima"
## [21] "Fagus sp"                     "Rubus aboriginum"
## [23] "Croton argyranthemus"         "Ceanothus fendleri"
## [25] "Chrysanthamnus sp"            "Abutilon hirtum"
## [27] "Asclepias pedicellata"        "Aralia racemosa"
## [29] "Machaeranthera canescens"     "Guettarda krugii"
## [31] "Sideroxylon spp"              "Houstonia longifolia"
## [33] "Scirpus cyperinus"            "Carex bigelowii"
## [35] "Crepis spp"                   "Aristida ternipes"
## [37] "Trifolium spp"                "Viburnum rufidulum"
## [39] "Campanula rapunculoides"      "Clematis sp"
## [41] "Thelypteris noveboracensis"   "Orobanche uniflora"
## [43] "Erigeron modestus"           "Phacelia hydrophyloides"
## [45] "Aristida adscensionis"        "Lonicera spp"
## [47] "Rumex orbiculatus"            "Pennisetum villosum"
## [49] "Rhododendron catawbiense"     "Elymus albicans"
## [51] "Phoradendron californicum"    "Trachelospermum difforme"
## [53] "Equisetum fluviatile"         "Utricularia subulata"
## [55] "Eryngium yuccifolium"         "Tephrosia hispidula"
## [57] "Asplenium contiguum"          "Gaura coccinea"
## [59] "Muhlenbergia reverchonii"     "Sonchus asper"
## [61] "Urochloa reptans"             "Rumex hastatulus"
## [63] "Asclepias viridis"            "Cercocarpus montanus"
## [65] "Hemitomes congestum"          "Actaea spp"
## [67] "Zinnia acerosa"               "Sabatia macrophylla"
## [69] "Crataegus xuberis Sarg. (pro sp.)" "Cryptantha paysonii"
## [71] "Euonymus fortunei"            "Yeatesia viridiflora"
## [73] "Lescurea saxicola"            "Geum spp"
## [75] "Platanthera psychodes"        "Ctenium aromaticum"
## [77] "Paspalum laeve"               "Viguiera dentata"
## [79] "Sida spinosa"                 "Coccoloba sp"
## [81] "Atriplex canescens"           "Euphorbia dentata"
```

```
## [83] "Cirsium nuttallii"      "Amsinckia sp"
## [85] "Emilia fosbergii"      "Packera tomentosa"
## [87] "Fimbristylis autumnalis" "Veratrum parviflorum"
## [89] "Veronica chamaedrys"   "Malvastrum coromandelianum"
## [91] "Poa saltuensis"        "Anthraenantia villosa"
## [93] "Solidago flexicaulis"  "Physalis crassifolia"
## [95] "Amsonia sp"            "Leersia lenticularis"
## [97] "Psilostrophe tagetina" "Lathyrus polyphyllus"
## [99] "Salix niphoclada"      "Leucothrinax morrisii"
```

### Question 3

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
n_1 <- data_plant %>%
  filter(sample_area_m2 == 1) %>%
  group_by(siteID) %>%
  summarize(richness_1m2 = n_distinct(taxon_name2))
n_10 <- data_plant %>%
  filter(sample_area_m2 %in% c(1, 10)) %>%
  group_by(siteID) %>%
  summarize(richness_10m2 = n_distinct(taxon_name2))
n_100 <- data_plant %>%
  filter(sample_area_m2 %in% c(1, 10, 100)) %>%
  group_by(siteID) %>%
  summarize(richness_100m2 = n_distinct(taxon_name2))
n_all <- left_join(n_1, n_10) %>% left_join(n_100)
```

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

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

```
n_all
```

```
## # A tibble: 47 x 4
##   siteID richness_1m2 richness_10m2 richness_100m2
##   <chr>      <int>      <int>      <int>
## 1 ABBY         196         239         273
```

```
## 2 BARR      76      90      94
## 3 BART      81     108     133
## 4 BLAN     282     330     407
## 5 BONA      75      92     104
## 6 CLBJ     417     487     528
## 7 CPER     189     225     244
## 8 DCFS     234     275     309
## 9 DEJU     155     184     199
## 10 DELA    304     391     461
## # i 37 more rows
```

## Question 4

```
library(tidyr)
n_all_ldf <- n_all %>%
  pivot_longer(
    cols = starts_with("richness_"),
    names_to = "spatial_scale",
    values_to = "richness"
  )
n_all_ldf
```

```
## # A tibble: 141 x 3
##   siteID spatial_scale richness
##   <chr>   <chr>         <int>
## 1 ABBY   richness_1m2      196
## 2 ABBY   richness_10m2     239
## 3 ABBY   richness_100m2    273
## 4 BARR   richness_1m2       76
## 5 BARR   richness_10m2      90
## 6 BARR   richness_100m2     94
## 7 BART   richness_1m2       81
## 8 BART   richness_10m2     108
## 9 BART   richness_100m2    133
## 10 BLAN  richness_1m2     282
## # i 131 more rows
```

## Question 5

```
library(ggplot2)
ggplot(n_all_ldf, aes(x = spatial_scale, y = richness, group = siteID)) +
  geom_line() +
  geom_point() +
  labs(x = "Spatial Scale",
       y = "Richness")
```

