Homework 5

Nathaniel Haulk

10/28/2021

```
library("devtools")

#devtools::install_github("daijiang/neonDivData")

library("neonDivData")

library("tidyverse")
```

Question 1:

```
#sort(unique(data_plant$taxon_name))

## Pulls the genus for each plant
data_plant$genus = sub(" .*", "", data_plant$taxon_name)

## Selects a random number 100 times between 1 and the length of data_plant
i = sample(1:nrow(data_plant), 100, replace=FALSE)

## Records the genus at the random
genus_samp = data_plant$genus[i]

## Prints out 100 genus
print(genus_samp)
```

```
##   [1] "Dennstaedtia"   "Schedonorus"    "Carex"          "Viola"
##   [5] "Smilax"         "Smilax"         "Carex"          "Ranunculus"
##   [9] "Coreopsis"      "Cynodon"        "Polygonum"      "Petasites"
##  [13] "Carya"          "Hypochaeris"    "Viola"          "Festuca"
##  [17] "Quercus"        "Bouteloua"      "Andropogon"     "Dichanthium"
##  [21] "Pseudotsuga"    "Quercus"        "Cissus"         "Ostrya"
##  [25] "Dichanthelium"  "Ceratocephala"  "Quercus"        "Petasites"
##  [29] "Rubus"          "Microstegium"   "Parthenocissus" "Baccharis"
##  [33] "Callicarpa"     "Alnus"          "Picea"          "Quercus"
##  [37] "Symphyotrichum" "Diospyros"      "Symphoricarpos" "Bignonia"
##  [41] "Carex"          "Cercis"         "Parthenocissus" "Medeola"
##  [45] "Artemisia"      "Lycopodium"     "Smilax"         "Vernonia"
##  [49] "Polygonum"      "Carya"          "Prosopis"       "Cryptantha"
##  [53] "Amorpha"        "Carex"          "Solidago"       "Solanum"
```

```
## [57]  "Juncus"          "Elephantopus"   "Lindera"        "Pentagramma"
## [61]  "Yucca"           "Saxifraga"      "Tridens"        "Alliaria"
## [65]  "Muhlenbergia"    "Lespedeza"      "Ampelopsis"     "Tragopogon"
## [69]  "Aquilegia"       "Vitis"          "Dactylis"       "Galactia"
## [73]  "Alyssum"         "Acer"           "Solidago"       "Carex"
## [77]  "Galium"          "Liatris"        "Osmorhiza"      "Calla"
## [81]  "Conyza"          "Polystichum"    "Lotus"          "Viola"
## [85]  "Artemisia"       "Sabal"          "Portulaca"      "Tsuga"
## [89]  "Dasyochloa"      "Acer"           "Circaea"        "Carex"
## [93]  "Trientalis"      "Thelesperma"    "Stellaria"      "Aesculus"
## [97]  "Doellingeria"    "Prunus"         "Vaccinium"      "Ageratina"
```

Question 2:

```
data_plant$taxon_name2 = sub("(\\w\\s\\w+).*", "\\1", data_plant$taxon_name)

## Selects a random number 100 times between 1 and the length of data_plant
i = sample(1:nrow(data_plant), 100, replace=FALSE)

## Records the genus and species name at the random
taxon2_samp = data_plant$taxon_name2[i]

## Prints out data
print(taxon2_samp)
```

```
##   [1] "Acer negundo"              "Spermolepis echinata"
##   [3] "Apocynum androsaemifolium" "Cyperus sp"
##   [5] "Acer rubrum"               "Astragalus flexuosus"
##   [7] "Elymus lanceolatus"        "Sanicula canadensis"
##   [9] "Oxalis sp"                 "Eriogonum tomentosum"
##  [11] "Juniperus virginiana"      "Ipomopsis laxiflora"
##  [13] "Symphoricarpos orbiculatus" "Evolvulus nuttallianus"
##  [15] "Veronica officinalis"      "Hypericum sp"
##  [17] "Campsis radicans"          "Symphyotrichum ericoides"
##  [19] "Sorghastrum nutans"        "Chamaesyce revoluta"
##  [21] "Arundinaria gigantea"      "Vaccinium corymbosum"
##  [23] "Aristida beyrichiana"      "Acer saccharum"
##  [25] "Galeopsis tetrahit"        "Psoralidium tenuiflorum"
##  [27] "Osmorhiza longistylis"     "Dichelostemma capitatum"
##  [29] "Aristida purpurea"         "Prunus serotina"
##  [31] "Euphorbia pubentissima"    "Parthenocissus quinquefolia"
##  [33] "Aristolochia serpentaria"  "Frangula caroliniana"
##  [35] "Solidago simplex"          "Empetrum nigrum"
##  [37] "Lotus plebeius"            "Ligustrum sinense"
##  [39] "Toxicodendron radicans"    "Lupinus sp"
##  [41] "Primula angustifolia"      "Vaccinium angustifolium"
##  [43] "Logfia sp"                 "Cornus foemina"
##  [45] "Mimosa quadrivalvis"       "Gutierrezia sarothrae"
##  [47] "Maianthemum canadense"     "Hordeum pusillum"
##  [49] "Fraxinus americana"        "Microstegium vimineum"
##  [51] "Symphyotrichum ericoides"  "Corydalis flavula"
```

```
##  [53] "Gelsemium sp"              "Bouteloua curtipendula"
##  [55] "Populus tremuloides"       "Cassiope tetragona"
##  [57] "Cirsium undulatum"         "Maianthemum canadense"
##  [59] "Toxicodendron radicans"    "Portulaca umbraticola"
##  [61] "Cornus florida"            "Euphorbia esula"
##  [63] "Galium triflorum"          "Aralia nudicaulis"
##  [65] "Phegopteris connectilis"   "Osmorhiza berteroi"
##  [67] "Symphyotrichum ericoides"  "Tidestromia lanuginosa"
##  [69] "Symphyotrichum sp"         "Cirsium arvense"
##  [71] "Ipomoea hederacea"         "Ipomoea sp"
##  [73] "Crataegus sp"              "Tidestromia lanuginosa"
##  [75] "Carex sp"                  "Trepocarpus aethusae"
##  [77] "Calylophus hartwegii"      "Ailanthus altissima"
##  [79] "Dalea purpurea"            "Rhus typhina"
##  [81] "Cirsium canescens"         "Eriogonum effusum"
##  [83] "Crotalaria rotundifolia"   "Ruellia caroliniensis"
##  [85] "Mitchella repens"          "Boerhavia sp"
##  [87] "Vernonia angustifolia"     "Tsuga canadensis"
##  [89] "Rubus flagellaris"         "Cornus florida"
##  [91] "Agropyron sp"              "Ambrosia psilostachya"
##  [93] "Alliaria petiolata"        "Calystegia sepium"
##  [95] "Smilax rotundifolia"       "Elymus repens"
##  [97] "Potentilla fissa"          "Sorghum bicolor"
##  [99] "Chaenactis stevioides"     "Thalictrum dasycarpum"
```

Question 3

```r
## Creates data frame n_1 that counts the number of sample_areas with a size of 1 m2 in data_plant
n_1 = data_plant %>%
  group_by(siteID) %>%
  summarise(richness_1m2 = sum(sample_area_m2 == "1"))

## Creates data frame n_10 that counts the number of sample_areas with a size of 1 and 10 m2 in data_pl
n_10 = data_plant %>%
  group_by(siteID) %>%
  summarise(richness_10m2 = sum(sample_area_m2 == "1" | sample_area_m2 == "10"))

## Creates data frame n_100 that counts the number of sample_areas with a size of 1, 10, and 100 m2 in
n_100 = data_plant %>%
  group_by(siteID) %>%
  summarise(richness_100m2 = sum(sample_area_m2 == "1" | sample_area_m2 == "10"| sample_area_m2 == "100

## Joins n_1, n_10, n_100
n_all = left_join(n_1 ,n_10) %>%
  left_join(n_100)
```

```
## Joining, by = "siteID"
## Joining, by = "siteID"
```

```
## prints n_all
print(n_all)
```

```
## # A tibble: 47 x 4
##     siteID richness_1m2 richness_10m2 richness_100m2
##     <chr>          <int>         <int>          <int>
##  1 ABBY            7587         13603          16946
##  2 BARR            4615          6041           6616
##  3 BART            5875         10659          13354
##  4 BLAN            8803         15270          19798
##  5 BONA            4091          5925           6983
##  6 CLBJ           10226         15656          18328
##  7 CPER           23663         37876          44284
##  8 DCFS           12067         16649          19622
##  9 DEJU            8673         13000          14918
## 10 DELA           12012         22521          29326
## # ... with 37 more rows
```

Question 4

```
## creates n_all_longer that records site richness based off spatial scale.
n_all_longer = n_all %>%
  pivot_longer(!siteID, names_to = "spatial_scale", values_to = "richness")

## Prints dataframe
print(n_all_longer)
```

```
## # A tibble: 141 x 3
##     siteID spatial_scale  richness
##     <chr>  <chr>             <int>
##  1 ABBY   richness_1m2       7587
##  2 ABBY   richness_10m2     13603
##  3 ABBY   richness_100m2    16946
##  4 BARR   richness_1m2       4615
##  5 BARR   richness_10m2      6041
##  6 BARR   richness_100m2     6616
##  7 BART   richness_1m2       5875
##  8 BART   richness_10m2     10659
##  9 BART   richness_100m2    13354
## 10 BLAN   richness_1m2       8803
## # ... with 131 more rows
```

Question 5

```
## Graphs n_all_longer
ggplot(data = n_all_longer, aes(x = spatial_scale, y = richness,  group = siteID))+
  geom_point()+
  geom_line()
```