Homework5

Kirti Agrawal

02/11/2021

Answer1

```
library(neonDivData)
DT::datatable(head(neonDivData::data_plant, n = 100))
```



```
y = neonDivData::data_plant
library(dplyr, warn.conflicts = FALSE)
data_plant = select(y, siteID,
        starts_with("taxon"))
data_plant$x = data_plant$taxon_name
data_plant
```

```
## # A tibble: 915,539 x 5
##    siteID taxon_id taxon_name              taxon_rank x
##    <chr>  <chr>    <chr>                   <chr>      <chr>
## 1 BART    VILA11   Viburnum lantanoides Michx.  species    Viburnum lantanoides ~
## 2 BART    ACSA3    Acer saccharum Marshall     species    Acer saccharum Marsha~
## 3 BART    FAGR     Fagus grandifolia Ehrh.     species    Fagus grandifolia Ehr~
## 4 BART    FAGR     Fagus grandifolia Ehrh.     species    Fagus grandifolia Ehr~
## 5 BART    FAGR     Fagus grandifolia Ehrh.     species    Fagus grandifolia Ehr~
```

```
## 6 BART    FAGR     Fagus grandifolia Ehrh.     species     Fagus grandifolia Ehr~
## 7 BART    ACSA3    Acer saccharum Marshall      species     Acer saccharum Marsha~
## 8 BART    FAGR     Fagus grandifolia Ehrh.     species     Fagus grandifolia Ehr~
## 9 BART    ACPE     Acer pensylvanicum L.        species     Acer pensylvanicum L.
## 10 BART   ACSA3    Acer saccharum Marshall      species     Acer saccharum Marsha~
## # ... with 915,529 more rows
```

```
library(tidyr)
j = data_plant %>% separate(c(x), c("genus", "Col2", "Col3"), " ")
```

```
## Warning: Expected 3 pieces. Additional pieces discarded in 363131 rows [23,
## 29, 30, 34, 38, 42, 44, 50, 66, 68, 71, 74, 78, 79, 92, 93, 108, 110, 114,
## 121, ...].
```

```
## Warning: Expected 3 pieces. Missing pieces filled with `NA` in 138266 rows
## [20, 24, 27, 31, 35, 36, 37, 39, 41, 47, 49, 51, 57, 60, 63, 67, 73, 76, 83,
## 87, ...].
```

```
j
```

```
## # A tibble: 915,539 x 7
##     siteID taxon_id taxon_name                  taxon_rank genus    Col2     Col3
##     <chr>  <chr>    <chr>                       <chr>      <chr>    <chr>    <chr>
## 1  BART   VILA11   Viburnum lantanoides Michx. species    Viburnum lantan~ Mich~
## 2  BART   ACSA3    Acer saccharum Marshall      species    Acer     saccha~ Mars~
## 3  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus    grandi~ Ehrh.
## 4  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus    grandi~ Ehrh.
## 5  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus    grandi~ Ehrh.
## 6  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus    grandi~ Ehrh.
## 7  BART   ACSA3    Acer saccharum Marshall      species    Acer     saccha~ Mars~
## 8  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus    grandi~ Ehrh.
## 9  BART   ACPE     Acer pensylvanicum L.        species    Acer     pensyl~ L.
## 10 BART   ACSA3    Acer saccharum Marshall      species    Acer     saccha~ Mars~
## # ... with 915,529 more rows
```

```
data_plant <- j %>% select(-starts_with("Col"))
data_plant
```

```
## # A tibble: 915,539 x 5
##     siteID taxon_id taxon_name                  taxon_rank genus
##     <chr>  <chr>    <chr>                       <chr>      <chr>
## 1  BART   VILA11   Viburnum lantanoides Michx. species    Viburnum
## 2  BART   ACSA3    Acer saccharum Marshall      species    Acer
## 3  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus
## 4  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus
## 5  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus
## 6  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus
## 7  BART   ACSA3    Acer saccharum Marshall      species    Acer
## 8  BART   FAGR     Fagus grandifolia Ehrh.     species    Fagus
## 9  BART   ACPE     Acer pensylvanicum L.        species    Acer
## 10 BART   ACSA3    Acer saccharum Marshall      species    Acer
## # ... with 915,529 more rows
```

```
print(data_plant[sample(nrow(data_plant), 100), ])
```

```
## # A tibble: 100 x 5
##     siteID taxon_id taxon_name                                taxon_rank genus
##     <chr>  <chr>    <chr>                                     <chr>      <chr>
##  1 SRER    ARNE2    Argythamnia neomexicana Müll. Arg.        species    Argytham~
##  2 OAES    CAHAF    Calylophus hartwegii (Benth.) P.H. Rave~  subspecies Calyloph~
##  3 TALL    DICHA2   Dichanthelium sp.                         genus      Dichanth~
##  4 STEI    ORYZO    Oryzopsis sp.                             genus      Oryzopsis
##  5 ONAQ    BRTE     Bromus tectorum L.                        species    Bromus
##  6 LENO    PAQU2    Parthenocissus quinquefolia (L.) Planch.  species    Partheno~
##  7 JERC    PIAS2    Pityopsis aspera (Shuttlw. ex Small) Sm~  species    Pityopsis
##  8 MLBS    GABA     Gaylussacia baccata (Wangenh.) K. Koch    species    Gaylussa~
##  9 SRER    OPEN3    Opuntia engelmannii Salm-Dyck ex Engelm.  species    Opuntia
## 10 BLAN    DUIN     Duchesnea indica (Andrews) Focke          species    Duchesnea
## # ... with 90 more rows
```

Answer2

```
library(neonDivData)
y = neonDivData::data_plant
library(dplyr, warn.conflicts = FALSE)
data_plant = select(y, siteID,
               starts_with("taxon"), sample_area_m2,
        latitude, longitude, elevation)
data_plant$taxon_name2 = sub("(\\w+\\s+\\w+).*", "\\1", c(data_plant$taxon_name))
print(data_plant[sample(nrow(data_plant), 100), ])
```

```
## # A tibble: 100 x 9
##     siteID taxon_id taxon_name        taxon_rank sample_area_m2 latitude longitude
##     <chr>  <chr>    <chr>             <chr>      <chr>             <dbl>     <dbl>
##  1 SERC    CIAR2    Cinna arundinac~  species    10                 38.9     -76.6
##  2 TALL    SMGL     Smilax glauca W~  species    10                 32.9     -87.4
##  3 TALL    DEMA2    Desmodium maril~  species    100                33.0     -87.4
##  4 TALL    DICHA2   Dichanthelium s~  genus      1                  32.9     -87.4
##  5 SOAP    GAAP2    Galium aparine ~  species    1                  37.0    -119.
##  6 STEI    CLINT    Clintonia sp.     genus      10                 45.5     -89.6
##  7 JORN    LIAU4    Linum australe ~  species    100                32.6    -107.
##  8 DSNY    PORU5    Polygala rugeli~  species    100                28.1     -81.4
##  9 SOAP    BRTE     Bromus tectorum~  species    1                  37.0    -119.
## 10 DELA    CEOC2    Cephalanthus oc~  species    1                  32.5     -87.8
## # ... with 90 more rows, and 2 more variables: elevation <dbl>,
## #   taxon_name2 <chr>
```

Answer3

```
n_1  <- data_plant %>%
    group_by(siteID) %>%
    summarise(richness_1m2 = n_distinct(sample_area_m2 == "1"))
n_1
```

```
## # A tibble: 47 x 2
##    siteID richness_1m2
##    <chr>         <int>
##  1 ABBY              2
##  2 BARR              2
##  3 BART              2
##  4 BLAN              2
##  5 BONA              2
##  6 CLBJ              2
##  7 CPER              2
##  8 DCFS              2
##  9 DEJU              2
## 10 DELA              2
## # ... with 37 more rows
```

```r
n_10  <- data_plant %>%
    group_by(siteID) %>%
    summarise(richness_10m2 = n_distinct(sample_area_m2 %in% c("1", "10")))
n_10
```

```
## # A tibble: 47 x 2
##    siteID richness_10m2
##    <chr>          <int>
##  1 ABBY               2
##  2 BARR               2
##  3 BART               2
##  4 BLAN               2
##  5 BONA               2
##  6 CLBJ               2
##  7 CPER               2
##  8 DCFS               2
##  9 DEJU               2
## 10 DELA               2
## # ... with 37 more rows
```

```r
n_100  <- data_plant %>%
    group_by(siteID) %>%
    summarise(richness_100m2 = n_distinct(sample_area_m2 %in% c("1", "10", "100")))
n_100
```

```
## # A tibble: 47 x 2
##    siteID richness_100m2
##    <chr>           <int>
##  1 ABBY                1
##  2 BARR                1
##  3 BART                1
##  4 BLAN                1
##  5 BONA                1
##  6 CLBJ                1
##  7 CPER                1
##  8 DCFS                1
##  9 DEJU                1
## 10 DELA                1
## # ... with 37 more rows
```

```
n_all = dplyr::left_join(n_1, n_10) %>% left_join(n_100)
```

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

```
n_all
```

```
## # A tibble: 47 x 4
##    siteID richness_1m2 richness_10m2 richness_100m2
##    <chr>         <int>         <int>          <int>
##  1 ABBY              2             2              1
##  2 BARR              2             2              1
##  3 BART              2             2              1
##  4 BLAN              2             2              1
##  5 BONA              2             2              1
##  6 CLBJ              2             2              1
##  7 CPER              2             2              1
##  8 DCFS              2             2              1
##  9 DEJU              2             2              1
## 10 DELA              2             2              1
## # ... with 37 more rows
```
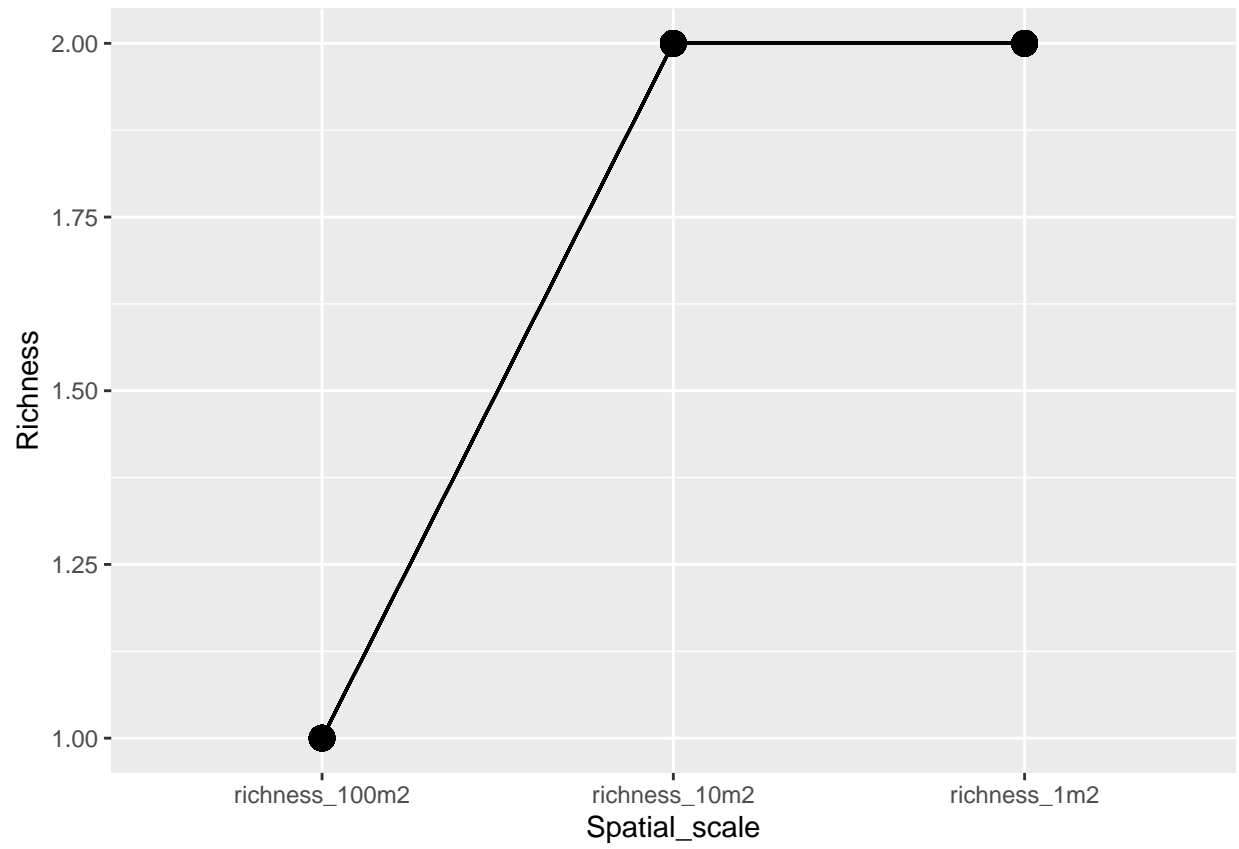
Answer4

```
library(tidyr)
n_all_long <- pivot_longer(n_all, cols=c(richness_1m2, richness_10m2, richness_100m2), names_to = "Spat:
print(n_all_long)
```

```
## # A tibble: 141 x 3
##    siteID Spatial_scale  Richness
##    <chr>  <chr>             <int>
##  1 ABBY   richness_1m2          2
##  2 ABBY   richness_10m2         2
##  3 ABBY   richness_100m2        1
##  4 BARR   richness_1m2          2
##  5 BARR   richness_10m2         2
##  6 BARR   richness_100m2        1
##  7 BART   richness_1m2          2
##  8 BART   richness_10m2         2
##  9 BART   richness_100m2        1
## 10 BLAN   richness_1m2          2
## # ... with 131 more rows
```

Answer5

```
library(ggplot2)
p <- ggplot(n_all_long, aes(Spatial_scale, Richness))
p + geom_point(size = 4) + geom_line(aes(group = siteID))
```

END