EmmaBeyer_A03_DataExploration.Rmd

Emma Beyer

Fall 2023

OVERVIEW

This exercise accompanies the lessons in Environmental Data Analytics on Data Exploration.

Directions

1. Rename this file `<FirstLast>_A03_DataExploration.Rmd` (replacing `<FirstLast>` with your first and last name).
2. Change "Student Name" on line 3 (above) with your name.
3. Work through the steps, **creating code and output** that fulfill each instruction.
4. Assign a useful **name to each code chunk** and include ample **comments** with your code.
5. Be sure to **answer the questions** in this assignment document.
6. When you have completed the assignment, **Knit** the text and code into a single PDF file.
7. After Knitting, submit the completed exercise (PDF file) to the dropbox in Sakai.

TIP: If your code extends past the page when knit, tidy your code by manually inserting line breaks.

TIP: If your code fails to knit, check that no `install.packages()` or `View()` commands exist in your code.

Set up your R session

1. Check your working directory, load necessary packages (tidyverse, lubridate), and upload two datasets: the ECOTOX neonicotinoid dataset (ECOTOX_Neonicotinoids_Insects_raw.csv) and the Niwot Ridge NEON dataset for litter and woody debris (NEON_NIWO_Litter_massdata_2018-08_raw.csv). Name these datasets "Neonics" and "Litter", respectively. Be sure to include the subcommand to read strings in as factors.

```
getwd()
```

```
## [1] "/home/guest/EDA/EDE_Fall2023"
```

```
#working directory set to "/home/guest/EDA/EDA_Fall2023"

#install.packages("tidyverse")
library(tidyverse)
#install.packages("lubridate")
```

```
library(lubridate)

Neonics <- read.csv("./Data/Raw/ECOTOX_Neonicotinoids_Insects_raw.csv",stringsAsFactors = T)
#uploaded ECOTOX neonicotinoid dataset
#View(Neonics)

Litter <- read.csv("./Data/Raw/NEON_NIWO_Litter_massdata_2018-08_raw.csv",stringsAsFactors = T)
#uploaded Niwot Ridge NEON dataset
#View(Litter)
```

Learn about your system

2. The neonicotinoid dataset was collected from the Environmental Protection Agency's ECOTOX Knowledgebase, a database for ecotoxicology research. Neonicotinoids are a class of insecticides used widely in agriculture. The dataset that has been pulled includes all studies published on insects. Why might we be interested in the ecotoxicology of neonicotinoids on insects? Feel free to do a brief internet search if you feel you need more background information.

   Answer: Neonicotinoids are the most common insecticide used in the world and are one of the most effective. It's know for being highly effective with insects, and less toxicity to mammals than previous insecticides.

3. The Niwot Ridge litter and woody debris dataset was collected from the National Ecological Observatory Network, which collectively includes 81 aquatic and terrestrial sites across 20 ecoclimatic domains. 32 of these sites sample forest litter and woody debris, and we will focus on the Niwot Ridge long-term ecological research (LTER) station in Colorado. Why might we be interested in studying litter and woody debris that falls to the ground in forests? Feel free to do a brief internet search if you feel you need more background information.

   Answer: More litter and woody debris could be a potential hazard during wildfire season. Areas with more litter would be at risk for a more intense wildfire.

4. How is litter and woody debris sampled as part of the NEON network? Read the NEON_Litterfall_UserGuide.pdf document to learn more. List three pieces of salient information about the sampling methods here:

   Answer: 1. Litter and fine woody debris are collected from elevated and ground traps. 2. All masses are reported at the spatial resolution of a single trap and the temporal resolution of a single collection event. 3. Collection events are separated into their respective functional groups: leaves, needles, twigs/branches, woody material, seeds, flowers, other, and mixed.

Obtain basic summaries of your data (Neonics)

5. What are the dimensions of the dataset?

```
dim(Neonics)
```

```
## [1] 4623   30
```

```
#shows 4623 rows and 30 columns of Neonics dataset
```

6. Using the `summary` function on the "Effect" column, determine the most common effects that are studied. Why might these effects specifically be of interest?

```
summary(Neonics$Effect)
```

```
##     Accumulation        Avoidance         Behavior     Biochemistry
##               12              102              360               11
##          Cell(s)      Development        Enzyme(s) Feeding behavior
##                9              136               62              255
##         Genetics           Growth        Histology       Hormone(s)
##               82               38                5                1
##    Immunological     Intoxication       Morphology        Mortality
##               16               12               22             1493
##       Physiology       Population     Reproduction
##                7             1803              197
```

```
#shows a summary of the Effects column of Neonics
```

Answer: The effects are important to see information about insect mortality, chemical accumulation, insect behaviors to the chemical, and phyical changes to the insects.

7. Using the `summary` function, determine the six most commonly studied species in the dataset (common name). What do these species have in common, and why might they be of interest over other insects? Feel free to do a brief internet search for more information if needed.[TIP: The `sort()` command can sort the output of the summary command...]

```
species_name_summary <- summary(Neonics$Species.Common.Name)
#shows a summary of the Species Common Name column of Neonics
species_name_summary
```

```
##                          Honey Bee                     Parasitic Wasp
##                                667                                285
##               Buff Tailed Bumblebee                Carniolan Honey Bee
##                                183                                152
##                          Bumble Bee                   Italian Honeybee
##                                140                                113
##                     Japanese Beetle                  Asian Lady Beetle
##                                 94                                 76
##                     Euonymus Scale                           Wireworm
##                                 75                                 69
##                  European Dark Bee                  Minute Pirate Bug
##                                 66                                 62
##                Asian Citrus Psyllid                      Parastic Wasp
##                                 60                                 58
##              Colorado Potato Beetle                   Parasitoid Wasp
##                                 57                                 51
##                 Erythrina Gall Wasp                     Beetle Order
##                                 49                                 47
```

```
##        Snout Beetle Family, Weevil          Sevenspotted Lady Beetle
##                                 47                                46
##                     True Bug Order            Buff-tailed Bumblebee
##                                 45                                39
##                       Aphid Family                  Cabbage Looper
##                                 38                                38
##                Sweetpotato Whitefly                   Braconid Wasp
##                                 37                                33
##                      Cotton Aphid                  Predatory Mite
##                                 33                                33
##              Ladybird Beetle Family                      Parasitoid
##                                 30                                30
##                      Scarab Beetle                   Spring Tiphia
##                                 29                                29
##                       Thrip Order              Ground Beetle Family
##                                 29                                27
##                 Rove Beetle Family                  Tobacco Aphid
##                                 27                                27
##                      Chalcid Wasp            Convergent Lady Beetle
##                                 25                                25
##                     Stingless Bee                 Spider/Mite Class
##                                 25                                24
##                Tobacco Flea Beetle                  Citrus Leafminer
##                                 24                                23
##                    Ladybird Beetle                       Mason Bee
##                                 23                                22
##                           Mosquito                  Argentine Ant
##                                 22                                21
##                             Beetle          Flatheaded Appletree Borer
##                                 21                                20
##                Horned Oak Gall Wasp                 Leaf Beetle Family
##                                 20                                20
##                  Potato Leafhopper          Tooth-necked Fungus Beetle
##                                 20                                20
##                       Codling Moth           Black-spotted Lady Beetle
##                                 19                                18
##                       Calico Scale                Fairyfly Parasitoid
##                                 18                                18
##                        Lady Beetle             Minute Parasitic Wasps
##                                 18                                18
##                          Mirid Bug                   Mulberry Pyralid
##                                 18                                18
##                           Silkworm                   Vedalia Beetle
##                                 18                                18
##               Araneoid Spider Order                        Bee Order
##                                 17                                17
##                     Egg Parasitoid                      Insect Class
##                                 17                                17
##            Moth And Butterfly Order       Oystershell Scale Parasitoid
##                                 17                                17
## Hemlock Woolly Adelgid Lady Beetle              Hemlock Wooly Adelgid
##                                 16                                16
##                               Mite                      Onion Thrip
##                                 16                                16
```

```
##              Western Flower Thrips                       Corn Earworm
##                                 15                                 14
##                  Green Peach Aphid                         House Fly
##                                 14                                 14
##                          Ox Beetle                 Red Scale Parasite
##                                 14                                 14
##                 Spined Soldier Bug               Armoured Scale Family
##                                 14                                 13
##                  Diamondback Moth                     Eulophid Wasp
##                                 13                                 13
##                 Monarch Butterfly                     Predatory Bug
##                                 13                                 13
##              Yellow Fever Mosquito                Braconid Parasitoid
##                                 13                                 12
##                       Common Thrip        Eastern Subterranean Termite
##                                 12                                 12
##                             Jassid                          Mite Order
##                                 12                                 12
##                          Pea Aphid                  Pond Wolf Spider
##                                 12                                 12
##           Spotless Ladybird Beetle             Glasshouse Potato Wasp
##                                 11                                 10
##                           Lacewing             Southern House Mosquito
##                                 10                                 10
##            Two Spotted Lady Beetle                       Ant Family
##                                 10                                  9
##                       Apple Maggot                          (Other)
##                                  9                                670
```

```
sorted_species <- sort(species_name_summary, decreasing = TRUE)
#sort the species in decreasing order
sorted_species
```

```
##                            (Other)                         Honey Bee
##                                670                               667
##                     Parasitic Wasp              Buff Tailed Bumblebee
##                                285                               183
##                  Carniolan Honey Bee                       Bumble Bee
##                                152                               140
##                    Italian Honeybee                  Japanese Beetle
##                                113                                94
##                  Asian Lady Beetle                    Euonymus Scale
##                                 76                                75
##                           Wireworm                 European Dark Bee
##                                 69                                66
##                  Minute Pirate Bug                Asian Citrus Psyllid
##                                 62                                60
##                      Parastic Wasp              Colorado Potato Beetle
##                                 58                                57
##                     Parasitoid Wasp                Erythrina Gall Wasp
##                                 51                                49
##                      Beetle Order         Snout Beetle Family, Weevil
##                                 47                                47
##           Sevenspotted Lady Beetle                    True Bug Order
```

```
##                             46                                45
##             Buff-tailed Bumblebee                       Aphid Family
##                             39                                38
##                   Cabbage Looper               Sweetpotato Whitefly
##                             38                                37
##                    Braconid Wasp                     Cotton Aphid
##                             33                                33
##                   Predatory Mite              Ladybird Beetle Family
##                             33                                30
##                      Parasitoid                      Scarab Beetle
##                             30                                29
##                   Spring Tiphia                      Thrip Order
##                             29                                29
##              Ground Beetle Family                 Rove Beetle Family
##                             27                                27
##                   Tobacco Aphid                     Chalcid Wasp
##                             27                                25
##             Convergent Lady Beetle                    Stingless Bee
##                             25                                25
##                 Spider/Mite Class                Tobacco Flea Beetle
##                             24                                24
##                  Citrus Leafminer                    Ladybird Beetle
##                             23                                23
##                       Mason Bee                          Mosquito
##                             22                                22
##                   Argentine Ant                             Beetle
##                             21                                21
##         Flatheaded Appletree Borer                Horned Oak Gall Wasp
##                             20                                20
##                Leaf Beetle Family                   Potato Leafhopper
##                             20                                20
##         Tooth-necked Fungus Beetle                      Codling Moth
##                             20                                19
##          Black-spotted Lady Beetle                      Calico Scale
##                             18                                18
##                Fairyfly Parasitoid                       Lady Beetle
##                             18                                18
##             Minute Parasitic Wasps                         Mirid Bug
##                             18                                18
##                  Mulberry Pyralid                          Silkworm
##                             18                                18
##                    Vedalia Beetle               Araneoid Spider Order
##                             18                                17
##                        Bee Order                    Egg Parasitoid
##                             17                                17
##                     Insect Class            Moth And Butterfly Order
##                             17                                17
##       Oystershell Scale Parasitoid Hemlock Woolly Adelgid Lady Beetle
##                             17                                16
##              Hemlock Wooly Adelgid                              Mite
##                             16                                16
##                      Onion Thrip               Western Flower Thrips
##                             16                                15
##                     Corn Earworm                  Green Peach Aphid
```

```
##                                14                                14
##                         House Fly                          Ox Beetle
##                                14                                14
##                Red Scale Parasite                 Spined Soldier Bug
##                                14                                14
##             Armoured Scale Family                   Diamondback Moth
##                                13                                13
##                     Eulophid Wasp                  Monarch Butterfly
##                                13                                13
##                     Predatory Bug              Yellow Fever Mosquito
##                                13                                13
##                Braconid Parasitoid                      Common Thrip
##                                12                                12
##       Eastern Subterranean Termite                            Jassid
##                                12                                12
##                          Mite Order                        Pea Aphid
##                                12                                12
##                   Pond Wolf Spider           Spotless Ladybird Beetle
##                                12                                11
##             Glasshouse Potato Wasp                         Lacewing
##                                10                                10
##            Southern House Mosquito            Two Spotted Lady Beetle
##                                10                                10
##                          Ant Family                      Apple Maggot
##                                 9                                 9
```

```
top_six_species <- head(sorted_species, 6)
#shows the six most researched species in the sorted list
top_six_species
```

```
##               (Other)             Honey Bee        Parasitic Wasp
##                   670                   667                   285
## Buff Tailed Bumblebee   Carniolan Honey Bee            Bumble Bee
##                   183                   152                   140
```

Answer: The top six species are all bees or wasps. These are important species for pollination, which is essential for commercial farming and plant cultivation.

8. Concentrations are always a numeric value. What is the class of `Conc.1..Author.` column in the dataset, and why is it not numeric?

```
class(Neonics$Conc.1..Author.)
```

```
## [1] "factor"
```

```
#shows the class of Con.1..Author as factor
```

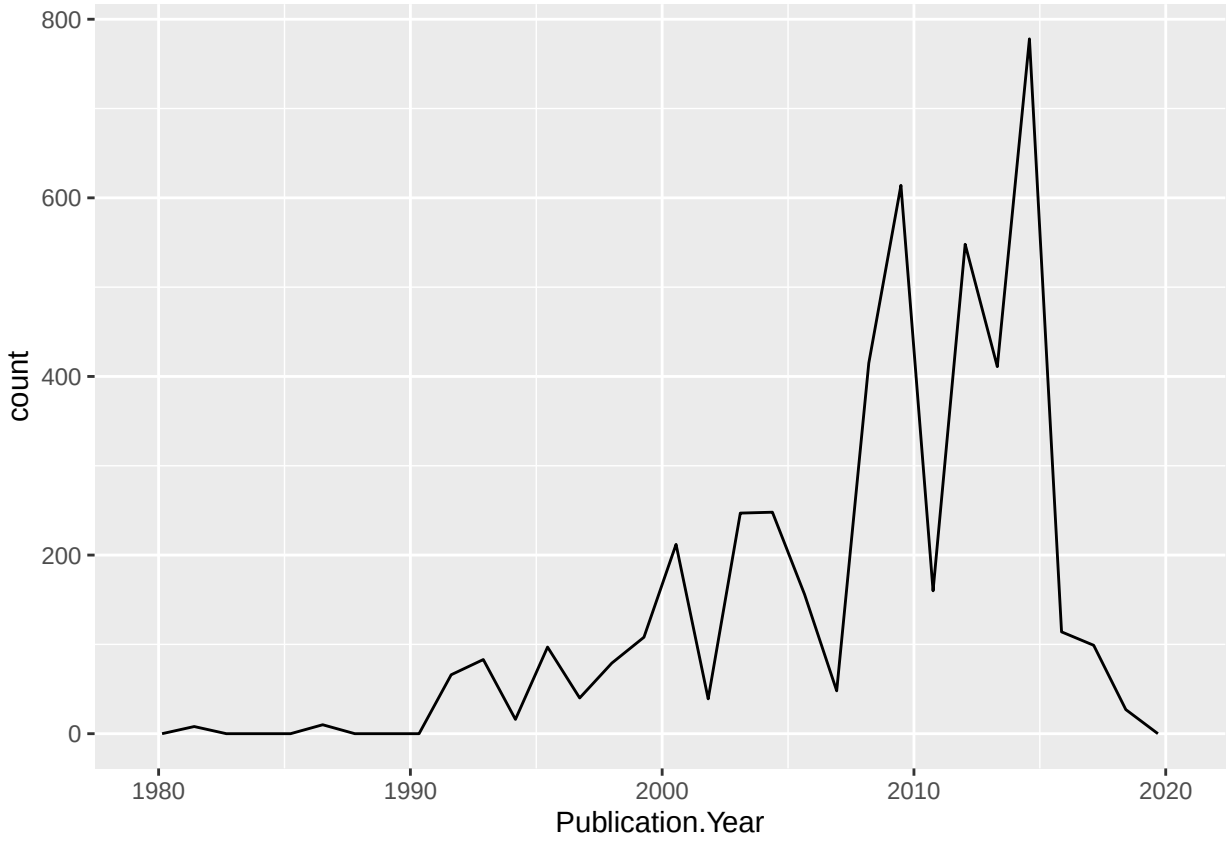
Answer: It's a factor, because it's referencing a thing not a number.

Explore your data graphically (Neonics)

9. Using `geom_freqpoly`, generate a plot of the number of studies conducted by publication year.

```
ggplot(Neonics) +
  geom_freqpoly(aes(x = Publication.Year))
```

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

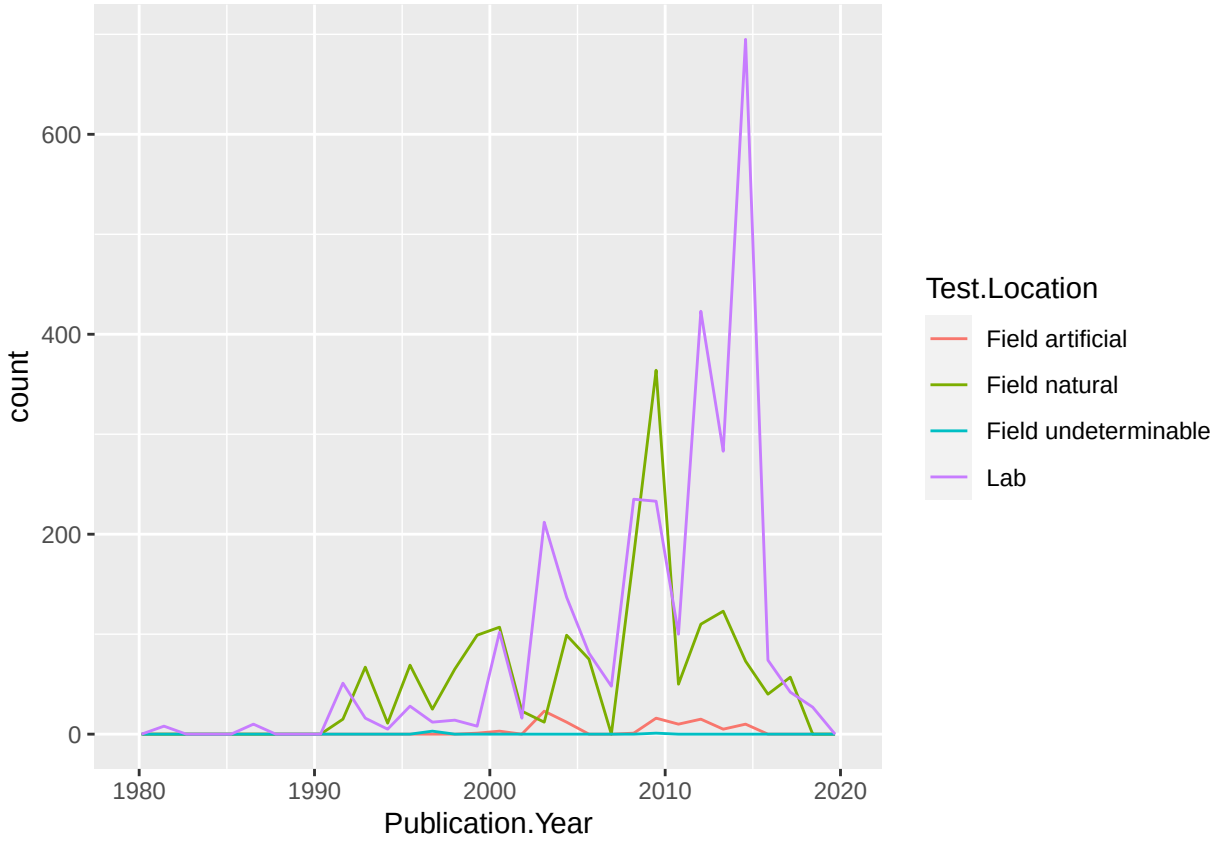



```
#line plot of publication years of studies
```

10. Reproduce the same graph but now add a color aesthetic so that different Test.Location are displayed as different colors.

```
ggplot(Neonics) +
  geom_freqpoly(aes(x = Publication.Year, color = Test.Location))
```

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




```
#line plot of publication and test locations in different colors
```

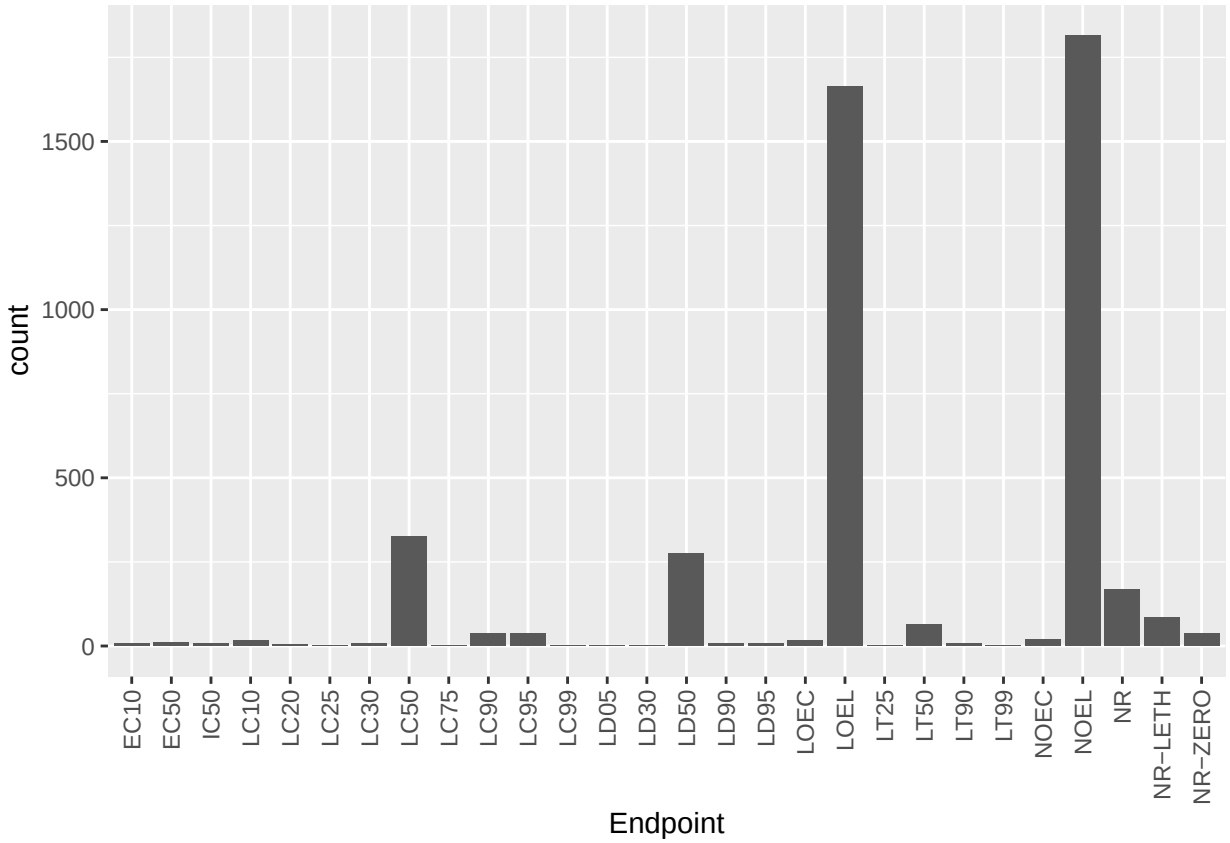
Interpret this graph. What are the most common test locations, and do they differ over time?

> Answer: There are more lab test locations starting after 2010 and more field natural test locations before 2010.

11. Create a bar graph of Endpoint counts. What are the two most common end points, and how are they defined? Consult the ECOTOX_CodeAppendix for more information.

[**TIP**: Add `theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))` to the end of your plot command to rotate and align the X-axis labels...]

```
ggplot(Neonics, aes(x = Endpoint)) +
  geom_bar() +
theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```




```
#bar graph of Endpoint counts and adjusting x-axis
```

Answer: NOEL was the most common, and LOEL was the second most common. NOEL stands for No-observable-effect-level and is terrestrial. LOEL stands for Lowest-observable-effect-level and is terrestrial.

Explore your data (Litter)

12. Determine the class of collectDate. Is it a date? If not, change to a date and confirm the new class of the variable. Using the `unique` function, determine which dates litter was sampled in August 2018.

```
class(Litter$collectDate)
```

```
## [1] "factor"
```

```
#class of collectDate is set to Factor

Litter$collectDate <- as.Date(Litter$collectDate, format = "%Y-%m-%d")
#changing the date format to date
class(Litter$collectDate)
```

```
## [1] "Date"
```

```
#proves that I changed the class to Date from Factor

unique(Litter$collectDate)
```

```
## [1] "2018-08-02" "2018-08-30"
```

```
#find dates that litter was samples in August 2018
```

13. Using the `unique` function, determine how many plots were sampled at Niwot Ridge. How is the information obtained from `unique` different from that obtained from `summary`?

```
unique(Litter$namedLocation)
```

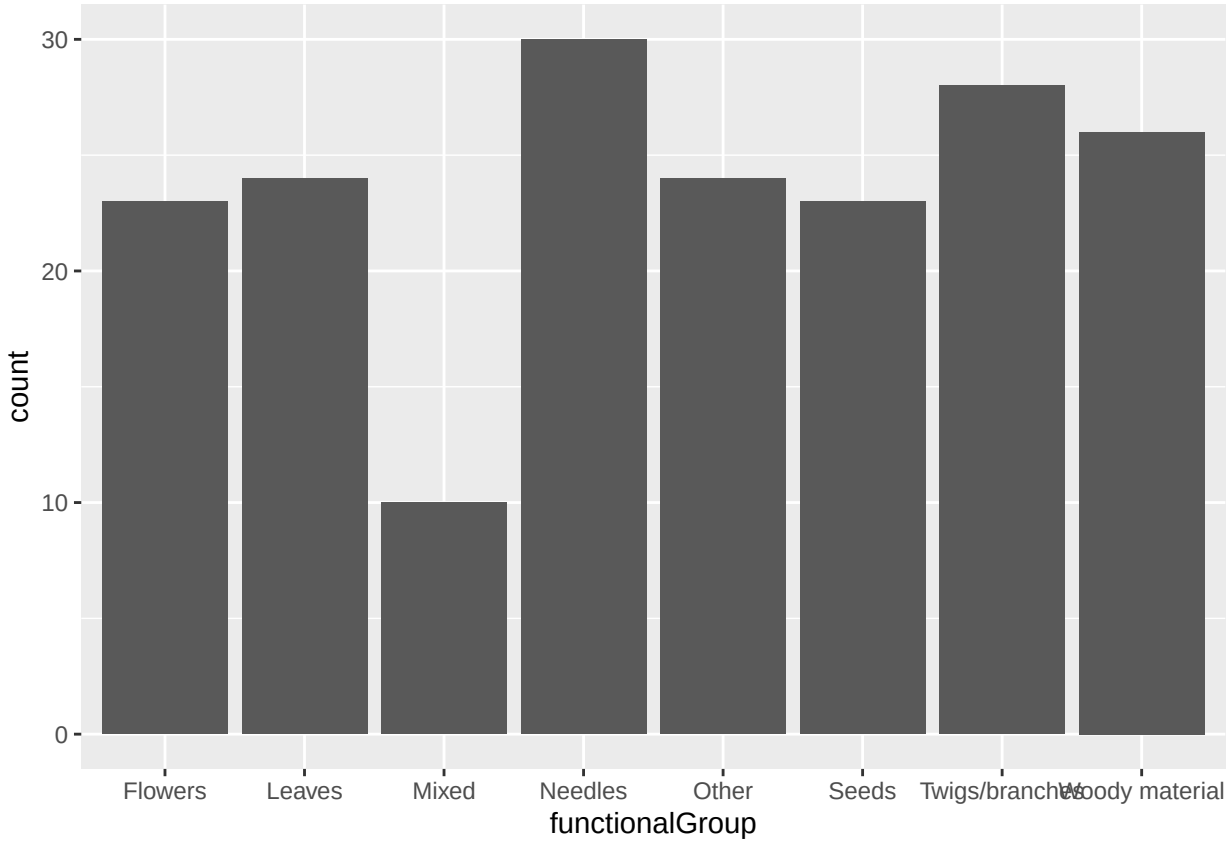
```
##  [1] NIWO_061.basePlot.ltr NIWO_064.basePlot.ltr NIWO_067.basePlot.ltr
##  [4] NIWO_040.basePlot.ltr NIWO_041.basePlot.ltr NIWO_063.basePlot.ltr
##  [7] NIWO_047.basePlot.ltr NIWO_051.basePlot.ltr NIWO_058.basePlot.ltr
## [10] NIWO_046.basePlot.ltr NIWO_062.basePlot.ltr NIWO_057.basePlot.ltr
## 12 Levels: NIWO_040.basePlot.ltr ... NIWO_067.basePlot.ltr
```

```
#find the unique values of locations in Niwot Ridge
```

Answer: The unique function shows how many unique values are in the column, while summary takes it a step further and show the unique values and a count each value in the data set.

14. Create a bar graph of functionalGroup counts. This shows you what type of litter is collected at the Niwot Ridge sites. Notice that litter types are fairly equally distributed across the Niwot Ridge sites.

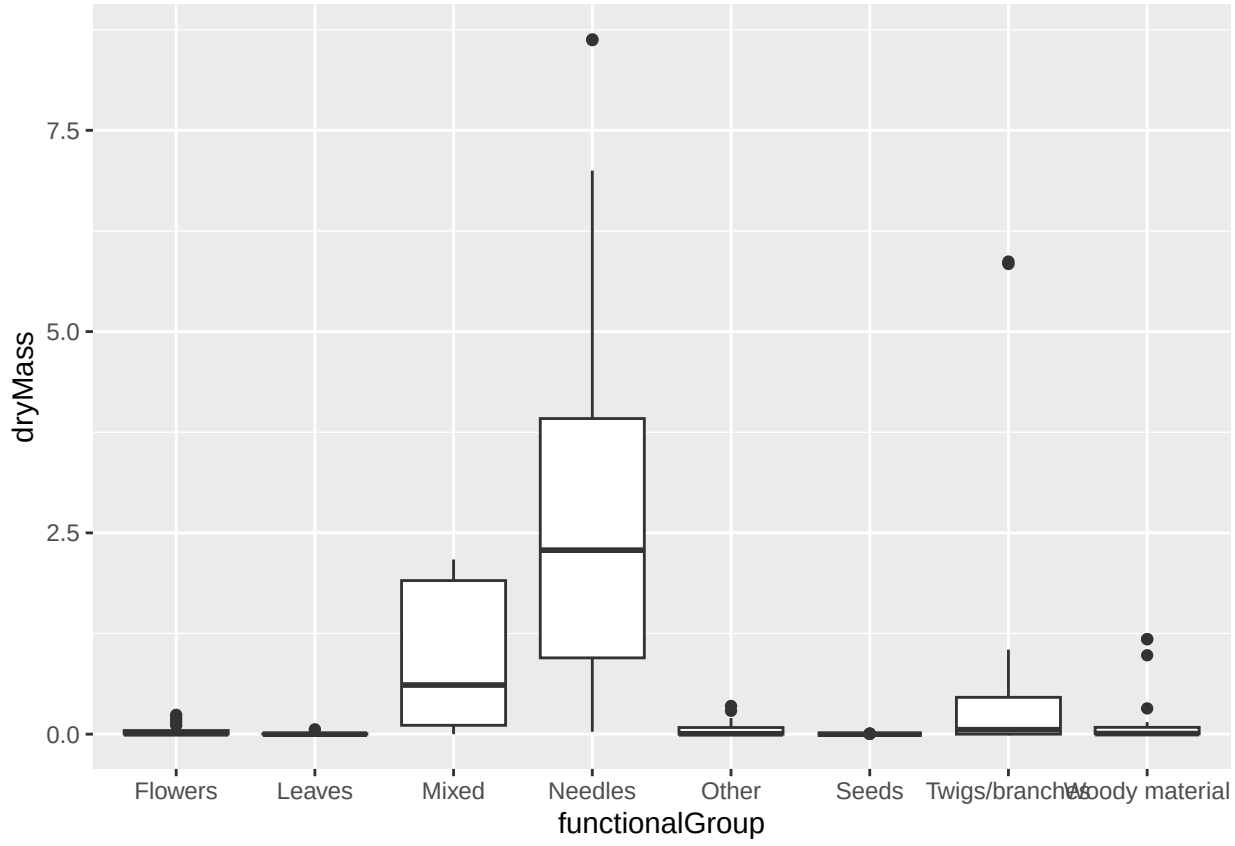
```
ggplot(Litter, aes(x = functionalGroup)) +
  geom_bar()
```




```
#bar graph of functionalGroup counts
```

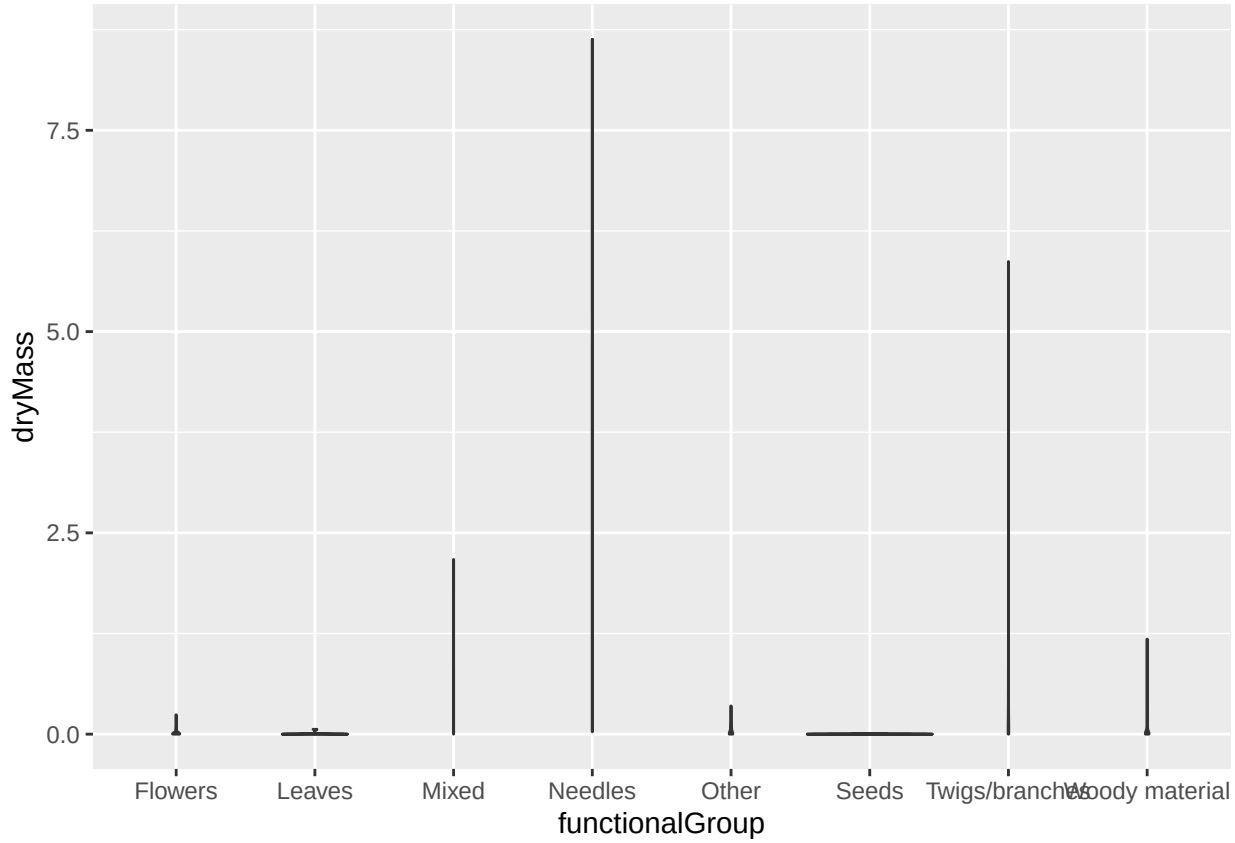
15. Using `geom_boxplot` and `geom_violin`, create a boxplot and a violin plot of dryMass by functionalGroup.

```
ggplot(Litter) +
  geom_boxplot(aes(x = functionalGroup, y = dryMass))
```




```
#box plot of functionalGroup (x) and dryMass (y)

ggplot(Litter) +
  geom_violin(aes(x = functionalGroup, y = dryMass),
              draw_quantiles = c(0.25, 0.50, 0.75))
```




```
#violin box plot of functionalGroup (x) and dryMass (y)
```

Why is the boxplot a more effective visualization option than the violin plot in this case?

> Answer:The box plot clearly shows the even distribution of the data, while the violin is too compacted. The violin show density well, and this data doesn't have well distributed density within the functional groups.

What type(s) of litter tend to have the highest biomass at these sites?

> Answer: Needles have the highest biomass.