FirstLast>_A03_DataExploration.Rmd

Desa Bolger

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() #checking WD
```

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

```
knitr::opts_chunk$set(echo = TRUE) #knitting, lubridate, and tidyverse, here, ggplot2 download
library(lubridate)
library(tidyverse)
library(here)
library(ggplot2)
```

```r
setwd(here())

#uploading two datasets
NeonicsFile <- here('Data','Raw','ECOTOX_Neonicotinoids_Insects_raw.csv')
print(NeonicsFile)
```

```
## [1] "/home/guest/EDE_Fall2023/Data/Raw/ECOTOX_Neonicotinoids_Insects_raw.csv"
```

```r
Neonics <- read.csv(
  file = here('Data','Raw','ECOTOX_Neonicotinoids_Insects_raw.csv'),
  stringsAsFactors = T)



LitterFile <- here('Data','Raw','NEON_NIWO_Litter_massdata_2018-08_raw.csv')
print(LitterFile)
```

```
## [1] "/home/guest/EDE_Fall2023/Data/Raw/NEON_NIWO_Litter_massdata_2018-08_raw.csv"
```

```r
Litter <- read.csv(
  file = here('Data','Raw','NEON_NIWO_Litter_massdata_2018-08_raw.csv'),
  stringsAsFactors = T)
```

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: It is likely that neonicotinoids can harm or kill unintented insects (not just pests for agriculture but other species needed for other ecosystems)

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: It can help explain how carbon moves through an ecosystem.

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. Debris is dried individually in each category and the mass is recorded. 2. Size must be a D $<$ 2 cm and a L $<$ 50 cm. 3. Traps were placed to get the debris.

Obtain basic summaries of your data (Neonics)

5. What are the dimensions of the dataset?

```r
dim(Neonics) #get dimensions
```

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

```r
#4623 by 30
```

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?

```r
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
```

```r
#Top 5: Population (1803), Mortality (1493), Behavior (360), Feeding behavior (255), Reproduction (197)

#ANSWER: If you see a change in population, death, behavior, or reproduction, it could signal that the
```

> Answer: Top 5: Population (1803), Mortality (1493), Behavior (360), Feeding behavior (255), Reproduction (197) If you see a change in population, death, behavior, or reproduction, it could signal that the insecticide has had a negative impact on a species.

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...]

```r
summary(Neonics$Species.Common.Name)
```

```
##                          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
```

```
#They are all bees/wasps! Other (670), Honey Bee (667), Parasitic Wasp (285),
#Buff Tailed Bumblebee (183), Carniolan Honey Bee (152), Bumble Bee (140),
#Italian Honeybee (113)

#They are all uncategorized or some sort of Bee/Wasp.  These may be of interest
#because they are less likely to be the main targets of the insecticides--
#perhaps studies are checking to see if there are unintended consequences to these species.


insect <- sort(summary(Neonics$Species.Common.Name), decreasing = TRUE)
insect
```

```
##                              (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
```

> Answer: They are all bees/wasps! Other (670), Honey Bee (667), Parasitic Wasp (285), Buff Tailed Bumblebee (183), Carniolan Honey Bee (152), Bumble Bee (140), Italian Honeybee (113)

#They are all uncategorized or some sort of Bee/Wasp. These may be of interest because they are less likely to be the main targets of the insecticides– perhaps studies are checking to see if there are unintended consequences to these species.

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?

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

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

```r
#It is a factor because some of the numbers have slashes at the end/ other
#various symbols, so they can't
#be classified specifically as a number or not
```

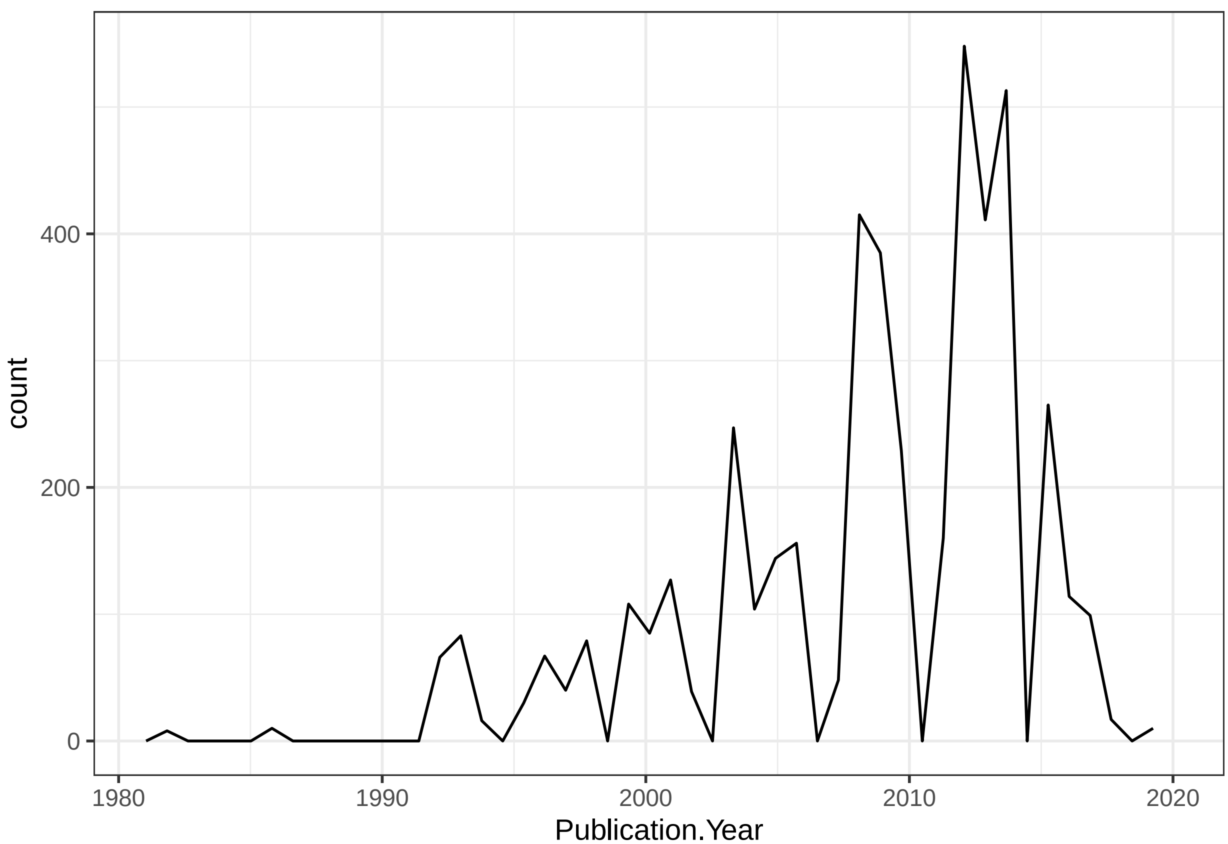
Answer: It is a factor because some of the numbers have slashes at the end/ other various symbols, so they can't be classified specifically as a number or not

Explore your data graphically (Neonics)

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

```r
ggplot(Neonics,
       aes(x = Publication.Year))+
      geom_freqpoly(bins = 50)+
  scale_x_continuous(limits = c(1981,2020))+ #changing x axis
  theme_bw()
```

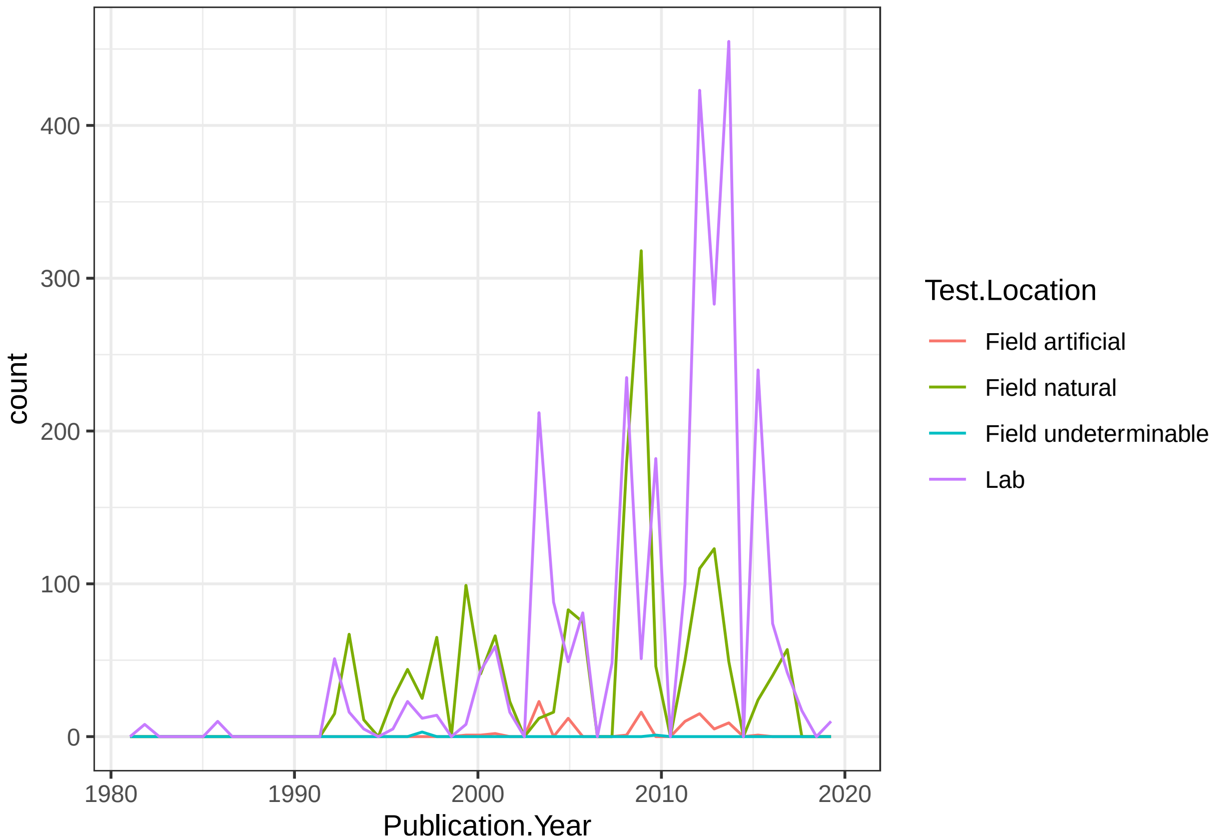
```
## Warning: Removed 3 rows containing missing values (`geom_path()`).
```



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

```
ggplot(Neonics,
       aes(x = Publication.Year, color = Test.Location))+ #change color by test location
      geom_freqpoly(bins = 50)+
  scale_x_continuous(limits = c(1981,2020))+
  theme_bw()
```

```
## Warning: Removed 12 rows containing missing values (`geom_path()`).
```



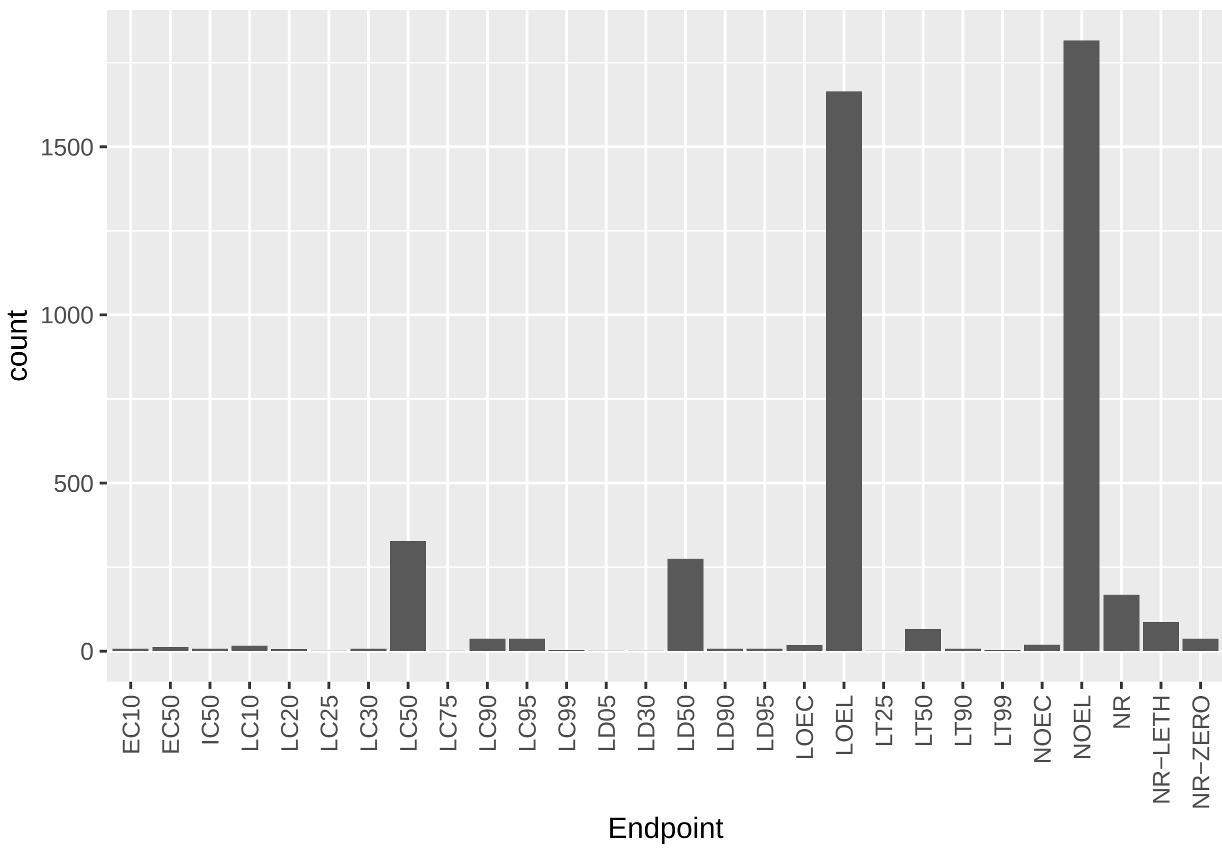
Interpret this graph. What are the most common test locations, and do they differ over time?

> Answer: The most common is the lab and Field natural, and they do seem to change over time. Lab peaks around 2014, and Field natural peaks around 2009.

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()+ #creates endpoints graph
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



Answer: The two most common endpoints are LOEL and NOEL, which according to the appendex are defined as LOEL (Lowest-observable-effect-level) and NOEL (No-observable-effect-level).

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.

```r
#Collection Date is a factor, not a date.

class(Litter$collectDate)
```

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

```r
#it is a factor, not a date.

Round2 <- unique(Litter$collectDate)
Round2
```

```
## [1] 2018-08-02 2018-08-30
## Levels: 2018-08-02 2018-08-30
```

```r
#August 2 and Aug 30 are sampling dates

#year month day conversion below

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

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

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`?

```r
summary(Litter$plotID)
```

```
## NIWO_040 NIWO_041 NIWO_046 NIWO_047 NIWO_051 NIWO_057 NIWO_058 NIWO_061
##       20       19       18       15       14        8       16       17
## NIWO_062 NIWO_063 NIWO_064 NIWO_067
##       14       14       16       17
```

```r
Special <- unique(Litter$plotID)
Special
```

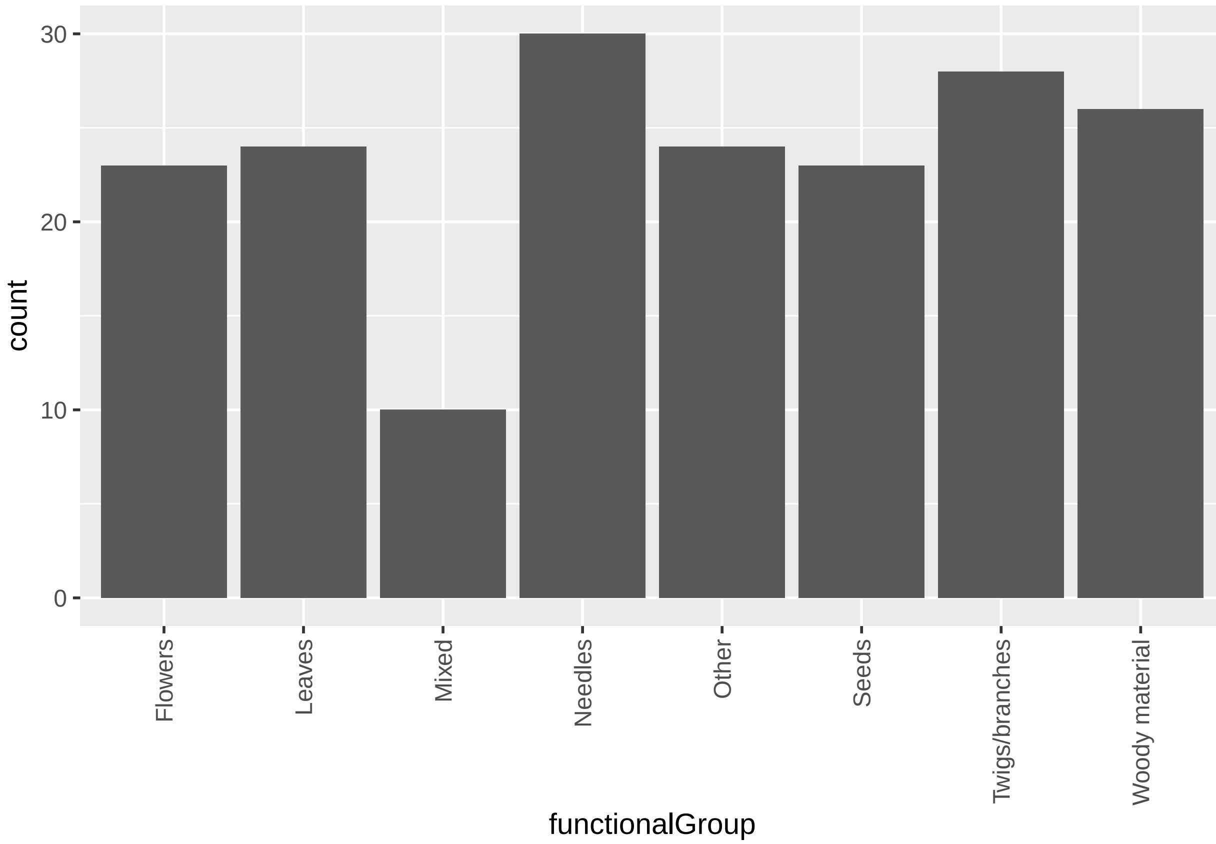
```
##  [1] NIWO_061 NIWO_064 NIWO_067 NIWO_040 NIWO_041 NIWO_063 NIWO_047 NIWO_051
##  [9] NIWO_058 NIWO_046 NIWO_062 NIWO_057
## 12 Levels: NIWO_040 NIWO_041 NIWO_046 NIWO_047 NIWO_051 NIWO_057 ... NIWO_067
```

```r
#Summary lists 12 plots sampled + their frequency. Unique  tells me the different types of samples and
```

Answer: Summary lists 12 plots sampled + their frequency. Unique tells me the different types of samples and the total number of different groups, but not their frequency

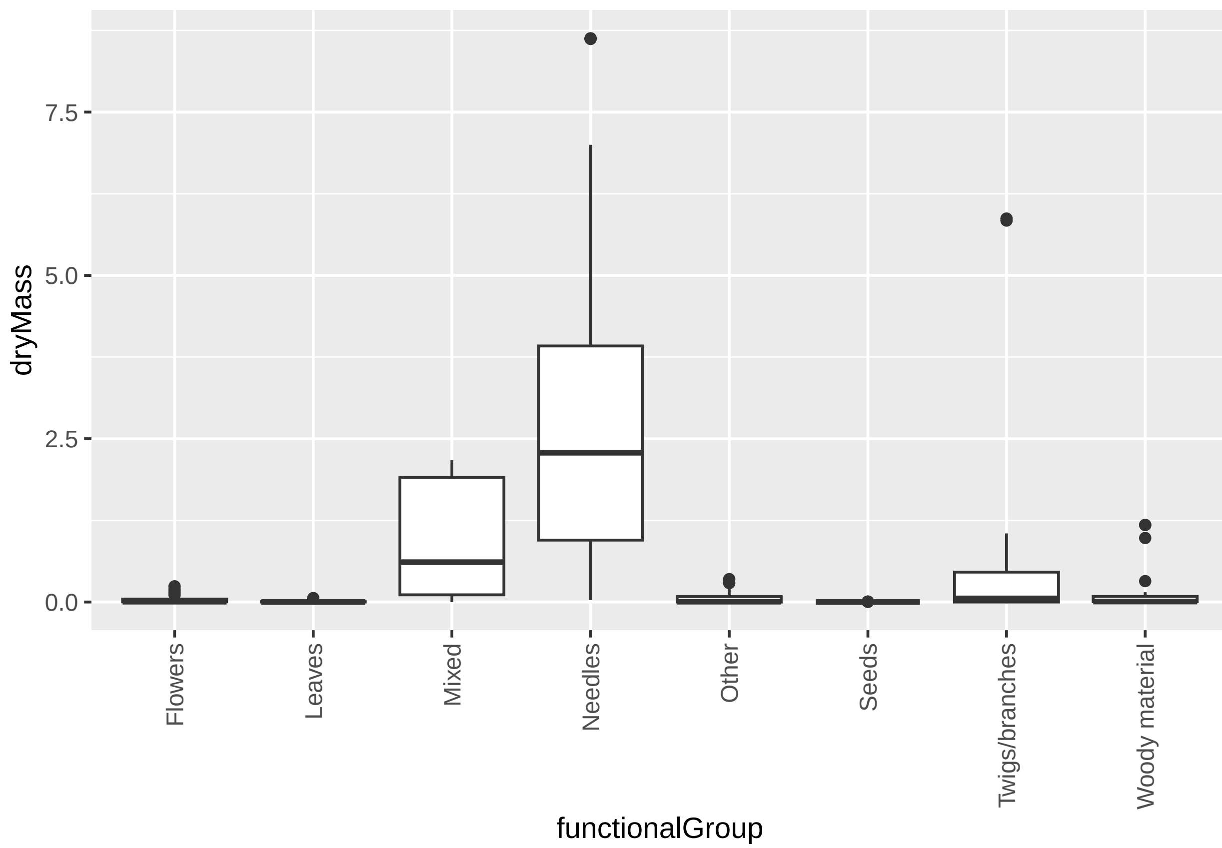
14. Create a bar graph of functionalGroup counts. This shows you what type of litter is collected at the Niwot Ridge sites. Notice that litter types are fairly equally distributed across the Niwot Ridge sites.

```r
ggplot(Litter,
       aes(x = functionalGroup))+
      geom_bar()+ #bar graph
theme(axis.text.x =element_text(angle = 90, vjust = 0.5, hjust=1))
```

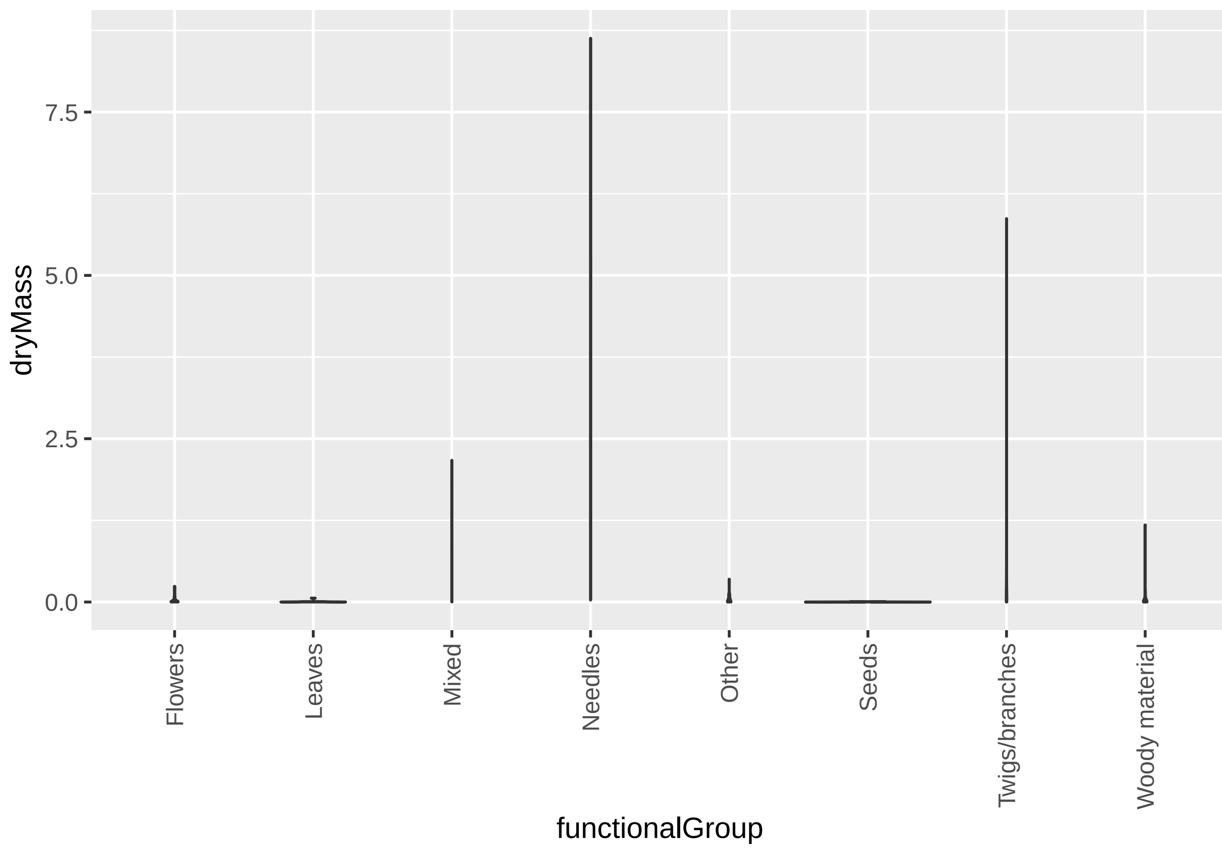


15. Using `geom_boxplot` and `geom_violin`, create a boxplot and a violin plot of dryMass by functionalGroup.

```
ggplot(Litter,
       aes(y= dryMass, x = functionalGroup))+
      geom_boxplot()+ #boxplot
  theme(axis.text.x =element_text(angle = 90, vjust = 0.5, hjust=1))
```




```r
ggplot(Litter,
       aes(y= dryMass, x = functionalGroup))+
      geom_violin()+ #violin
  theme(axis.text.x =element_text(angle = 90, vjust = 0.5, hjust=1))
```



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

> Answer: It shows the spread and outliers better. The violin does not have enough width to show a clear image.

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

> Answer: Needles, Mixed, and Twig branches have the highest mean dryMass.