Assignment 3: Data Exploration

Aislinn McLaughlin

OVERVIEW

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

Directions

1. Change "Student Name" on line 3 (above) with your name.
2. Work through the steps, **creating code and output** that fulfill each instruction.
3. Be sure to **answer the questions** in this assignment document.
4. When you have completed the assignment, **Knit** the text and code into a single PDF file.
5. After Knitting, submit the completed exercise (PDF file) to the dropbox in Sakai. Add your last name into the file name (e.g., "Salk_A03_DataExploration.Rmd") prior to submission.

The completed exercise is due on <>.

Set up your R session

1. Check your working directory, load necessary packages (tidyverse), 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.

```
getwd()
```

```
## [1] "/Users/Aislinn/Documents/GitHub/Environmental_Data_Analytics_2021"
```

```
setwd("~/Documents/GitHub/Environmental_Data_Analytics_2021")

library(tidyverse)

Neonics <- read.csv("./Data/Raw/ECOTOX_Neonicotinoids_Insects_raw.csv")
Litter <- read.csv("./Data/Raw/NEON_NIWO_Litter_massdata_2018-08_raw.csv")
```

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 ecotoxicologoy of neonicotinoids on insects? Feel free to do a brief internet search if you feel you need more background information.

   Answer: We want neonicotinoids to be as targeted as possible i.e. only kill the destructive insects. Ecotoxicology research will allow us to see what sort of effect they have on non-target insect populations such as bees, which are important for pollination, and ladybugs, which are important predators that may kill undesirable insects. On the other hand, this kind of research may also allow us to develop more effective pesticides to prevent harmful insects from destroying crops.

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: There are probably many reasons to study litter and woody debris. They can both be important habitat for insects and help return nutrients to the soil. Woody debris is also important for carbon capture - the longer it takes to decompose, the longer the carbon is kept out of the atmosphere.

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: *Sample masses are recorded to an accuracy of 0.01 grams.* Sampling sites must have woody specimens >2m tall. *Frequency of sample collection varies by trap type: ground traps are sampled once per year while elevated traps are sampled every 2 weeks or every 1-2 months.

Obtain basic summaries of your data (Neonics)

5. What are the dimensions of the dataset?

```
dim(Neonics)
```

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

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

```
sort(summary.factor(Neonics$Effect), decreasing = TRUE) #can't use summary() on chr
```

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

   Answer: Population and mortality are the most commonly studied effects. We want to know what effect the neonicotinoids are having on the size of the insect population and how they are influencing the death rate.

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.

```
sort(summary.factor(Neonics$Species.Common.Name), decreasing = TRUE)
```

```
##                           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                  Asiatic Honey Bee
##                                 9                                  9
##                Eulophid Parasitoid                    Lacewing Family
##                                 9                                  9
##                 Mealybug Destroyer             Alfalfa Leafcutter Bee
##                                 9                                  8
##                                Bee                          Bumblebee
##                                 8                                  8
##              Chilean Predatory Mite                    Dwarf Honey Bee
##                                 8                                  8
##            Neotropical Stingless Bee             Parasitic Wasp Family
##                                 8                                  8
##                 Spiralling Whitefly                 Beetle Mite Family
##                                 8                                  7
##                          Chinch Bug                Macedonian Honey Bee
##                                 7                                  7
```

```
##                      Moth                    Potato Tuberworm
##                         7                                   7
##       Russian Wheat Aphid                      Soldier Beetle
##                         7                                   7
## Southern One-Year Canegrub                Tarnished Plant Bug
##                         7                                   7
##           Ambrosia Beetle                          Aphid Wasp
##                         6                                   6
##         Black Vine Weevil                    Childers Canegrub
##                         6                                   6
##       Coconut Leaf Beetle     Elevenspotted Ladybird Beetle
##                         6                                   6
##             Encyrtid Wasp                    European Red Mite
##                         6                                   6
##             Fall Armyworm                           Fruit Fly
##                         6                                   6
##                 Hover Fly        Oblique Banded Leaf Roller
##                         6                                   6
##          Obscure Mealybug             Oribatid Mite Suborder
##                         6                                   6
##         Pistachio Psyllid             Redbay Ambrosia Beetle
##                         6                                   6
##       Silverleaf Whitefly                       Soybean Aphid
##                         6                                   6
##      Subterranean Termite                               Thrip
##                         6                                   6
##    Two-Spotted Spider Mite                        Apple Aphid
##                         6                                   5
##         Brown Planthopper                              Earwig
##                         5                                   5
##         Green June Beetle                        Hornfaced Bee
##                         5                                   5
##   Long Horned Beetle Family                      Plum Curculio
##                         5                                   5
##               Rove Beetle                       San Jose Scale
##                         5                                   5
##            Scelionid Wasp              Speckled Cutworm Moth
##                         5                                   5
##              Thrip Family                                 Ant
##                         5                                   4
##     Cabbage Seedpod Weevil              Common Green Lacewing
##                         4                                   4
##      Eucalyptus Gall Wasp              European Apple Sawfly
##                         4                                   4
##        European Honey Bee       European Tarnished Plant Bug
##                         4                                   4
##          Garden Symphylan                     Linyphiid Spider
##                         4                                   4
##              Onion Maggot                      Oriental Beetle
##                         4                                   4
##         Parsnip Seed Wasp                 Pea And Bean Weevil
##                         4                                   4
##               Pear Sucker               Red Imported Fire Ant
##                         4                                   4
```

```
##           Striped Cucumber Beetle                 Sugarcane Beetle
##                                 4                                4
##                              Wasp               Wolf Spider Family
##                                 4                                4
##           Yellow-faced Bumblebee              Ambrosia Bark Beetle
##                                 4                                3
##            Asian Ambrosia Beetle                    Beetle Family
##                                 3                                3
##                  Birch Leafminer                  Black Twig Borer
##                                 3                                3
##          Braconid Parasitoid Wasp              California Red Scale
##                                 3                                3
##             Crucifer Flea Beetle                           Cutworm
##                                 3                                3
##            Delphacid Planthopper          Egyptian Cotton Leafworm
##                                 3                                3
##               Encyrtid Parasitoid         Fly/Mosquito/Midge Order
##                                 3                                3
##     Formosan Subterranean Termite          Fruit-tree Pinhole Borer
##                                 3                                3
##            Green Rice Leafhopper                     Ground Beetle
##                                 3                                3
##                  Ichneumonid Wasp     Large-Jawed Orb Weaver Family
##                                 3                                3
##                  Leaf Cutting Ant         Mediterranean Fruit Fly
##                                 3                                3
##                  Minute Flour Bug                      Mite Family
##                                 3                                3
##                       Moth Family               Negatoria Canegrub
##                                 3                                3
##                 Sap Beetle Family               Scale Insect Order
##                                 3                                3
##              Scarab Beetle Family          Sheet-Web Weaver Family
##                                 3                                3
##                            Spider                   Sugarcane Grub
##                                 3                                3
##               Tenebrionid Beetle                Alfalfa Plant Bug
##                                 3                                2
##                        Alkali Bee                            Aphid
##                                 2                                2
##                      Assassin Bug                  Azalea Lace Bug
##                                 2                                2
##                      Banana Aphid                      Brown Scale
##                                 2                                2
##                    Brown Stinkbug                          Budworm
##                                 2                                2
##                     Cabbage Aphid                    Cabbage White
##                                 2                                2
##                    Cardamom Thrip                    Carrot Weevil
##                                 2                                2
##                 Celer Crab Spider                  Centipede Class
##                                 2                                2
##                   Citricola Scale                Clouded Plant Bug
##                                 2                                2
```

```
##             Coffee Bean Weevil                Cotton Fleahopper
##                              2                                2
##         Egyptian Alfalfa Weevil                  Engraver Beetle
##                              2                                2
##            Fig Longicorn Beetle       Glassy-winged Sharpshooter
##                              2                                2
##               Hawthorn Lace Bug             Hister Beetle Family
##                              2                                2
##           Jumping Spider Family                Lined Click Beetle
##                              2                                2
##               Maple Spider Mite                  Meshweaver Spider
##                              2                                2
##         Minute Pirate Bug Family                   Predaceous Fly
##                              2                                2
##            Pygmy Mangold Beetle                      Rose Sawfly
##                              2                                2
##            Serpentine Leafminer             Spider Mite Destroyer
##                              2                                2
##      Spotted Tentiform Leafminer                        Stink Bug
##                              2                                2
##              Tawny Mole Cricket          Tick/Chigger/Mite Order
##                              2                                2
##             Turf Running-spider                     Turnip Aphid
##                              2                                2
##             Western Bigeyed Bug               Western Damsel Bug
##                              2                                2
##               Western Plant Bug      White Apple Leafhopper Nymph
##                              2                                2
##         White-backed Planthopper           Whitemarked Fleahopper
##                              2                                2
##           Antlike Flower Beetle   Banded Soft-winged Flower Beetle
##                              1                                1
##           Banded Sunflower Moth                       Bee Family
##                              1                                1
##                    Beet Armyworm               Black Citrus Aphid
##                              1                                1
##              Blue Alfalfa Aphid                  Cabbage Root Fly
##                              1                                1
##              Cactus Lady Beetle                  Citrus Red Mite
##                              1                                1
##            Cottony Cushion Sale               Crapemyrtle Aphid
##                              1                                1
##                Damselbug Family             Ectoparasitoid Wasp
##                              1                                1
##             English Grain Aphid                        Fairyfly
##                              1                                1
##                     Flea Beetle                       Gall Midge
##                              1                                1
##   Grasshopper/Cricket/Locust Order              Greenhouse Whitefly
##                              1                                1
##        Grey Sunflower Seed Weevil          Harvestman Spider Order
##                              1                                1
##             Hawthorn Leaf Miner     Longtailed Fruit Fly Parasite
##                              1                                1
```

```
##                  Minute Lady Beetles               Painted Maple Aphid
##                                    1                                 1
##                        Pepper Weevil                Pine False Webworm
##                                    1                                 1
##                            Plant Bug                    Pollen Beetle
##                                    1                                 1
##                      Predacious Mite                    Predator Bug
##                                    1                                 1
##                Pseudocentipede Class            Pteromalid Wasp Family
##                                    1                                 1
##             Red Sunflower Seed Weevil             Rice Leaf Folder Moth
##                                    1                                 1
##                      Rose Grain Aphid               Scale Picnic Beetle
##                                    1                                 1
##                   Shiny Spider Beetle                Southern Army Worm
##                                    1                                 1
##                         Spirea Aphid      Spotted Sunflower Stem Weevil
##                                    1                                 1
##            Strawberry Blossom Weevil                   Sunflower Midge
##                                    1                                 1
##                       Sunflower Moth          Ten-spot Ladybird Beetle
##                                    1                                 1
##                        Tobacco Thrip          Twicestabbed Lady Beetle
##                                    1                                 1
##                          Wasp Family                            Weevil
##                                    1                                 1
##               Yellow Mealworm Beetle
##                                    1
```

Answer: The 6 most commonly studied species (in descending order) are: Honey Bee, Parasitic Wasp, Buff Tailed Bumblebee, Carniolan Honey Bee, Bumble Bee, Italian Honeybee. As I mentioned earlier, bees are important pollinators and we would want to know if neonicotinoids were adversely affecting populations and increasing their mortality. Fewer pollinators will have an adverse effect on crops.

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

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

```
## [1] "character"
```

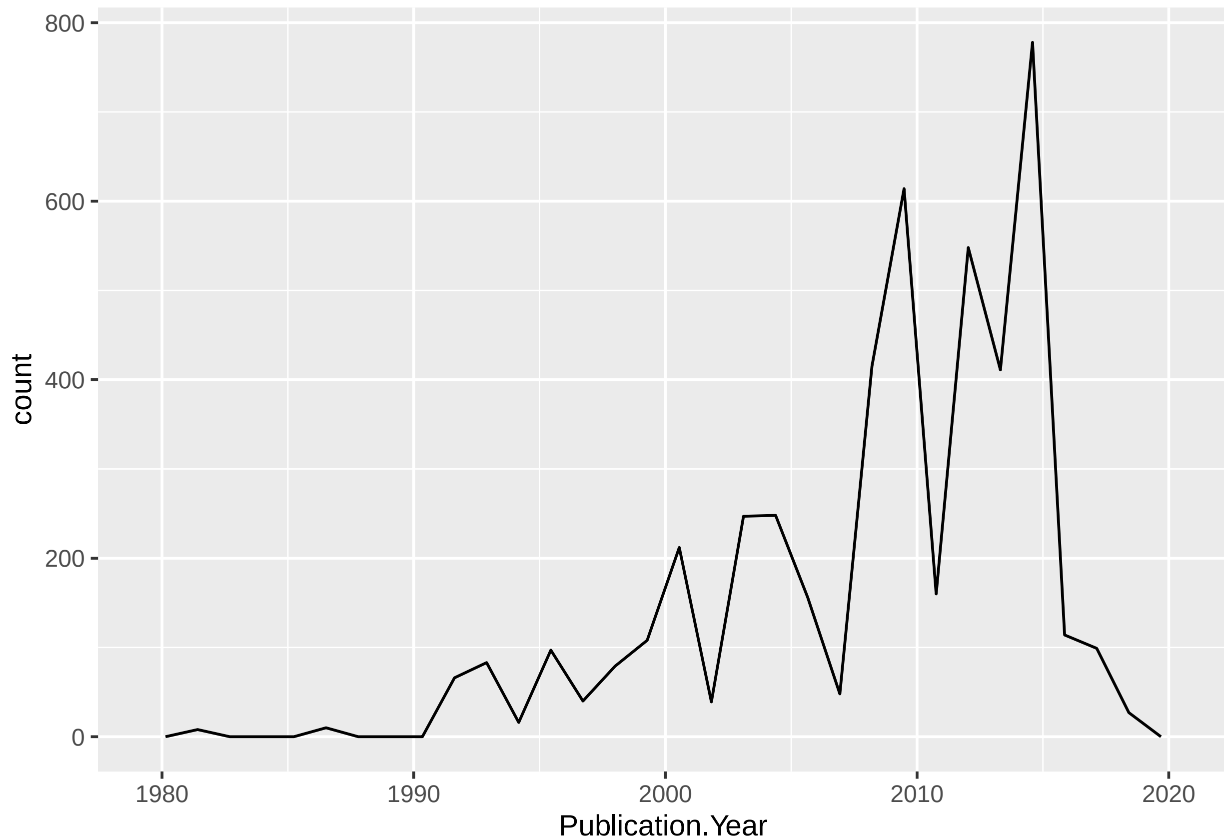
Answer: The class of the Conc.1..Author column is character. Simply by viewing the column, you can see that some values contain characters like '/' and '>', which is why the column is not considered numeric.

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))
```

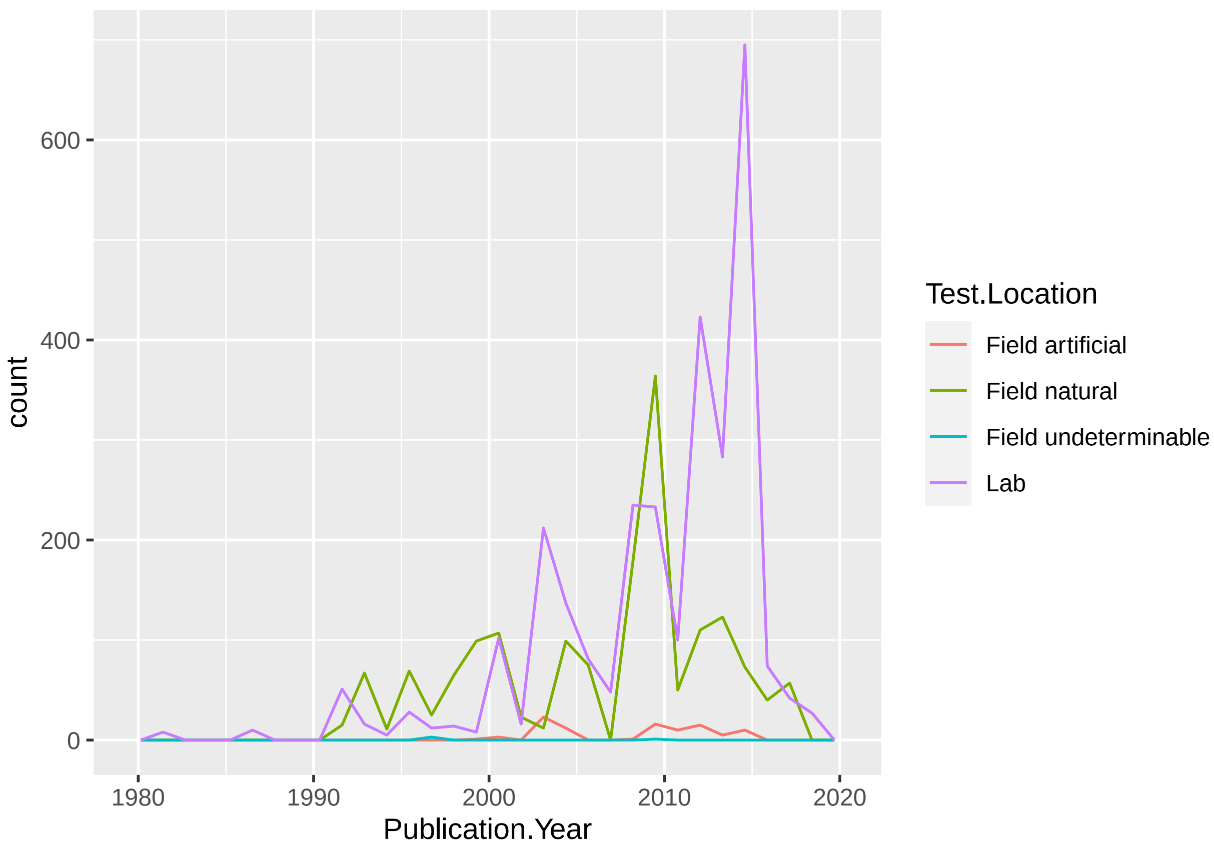
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



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`.
```

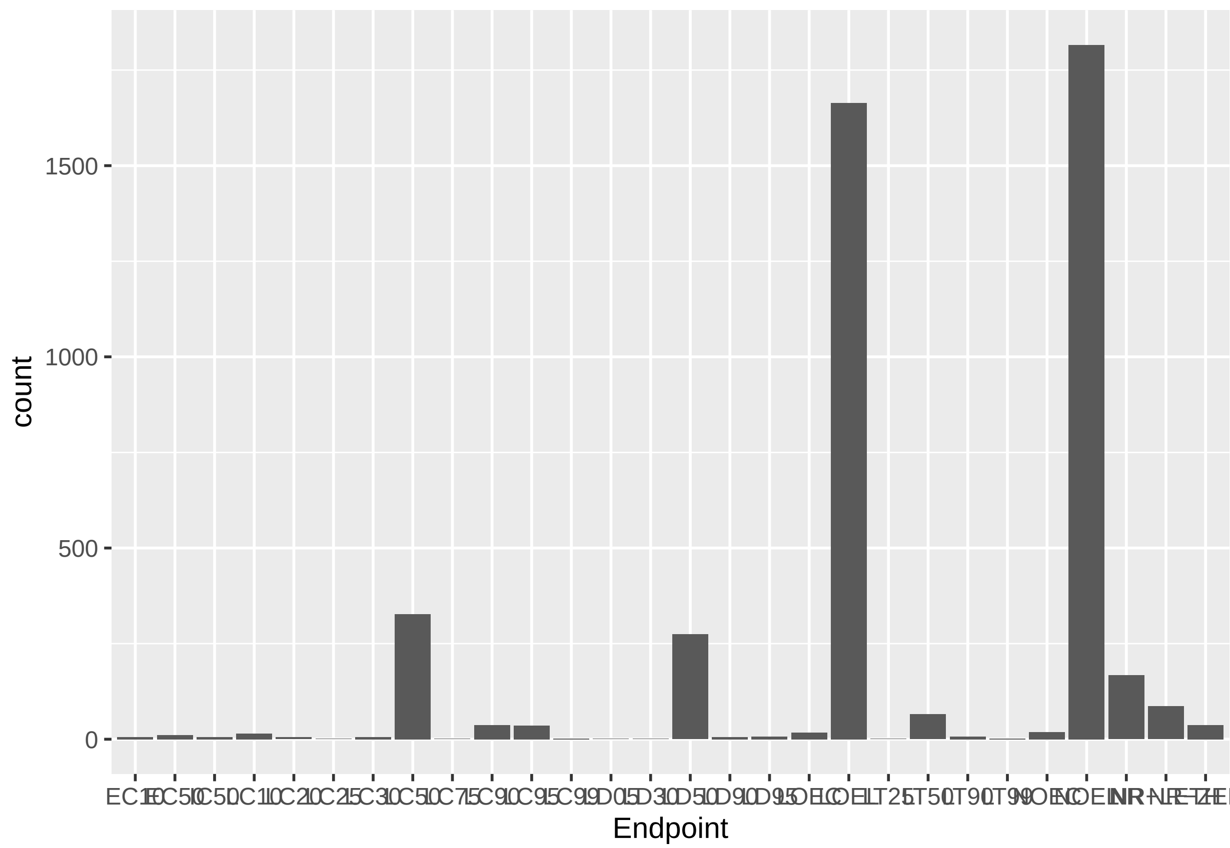


Interpret this graph. What are the most common test locations, and do they differ over time?

> Answer: From 1990-2010 the lab and the field (natural) were almost trading places as the most common test location. Many more lab tests were completed from around 2011-2016, and now lab and field (natural) test locations are back around the same level.

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.

```
ggplot(Neonics) +
  geom_bar(aes(x = Endpoint))
```



Answer: The two most common endpoints are NOEL and LOEL. I actually had to do a `sort(summary())` to figure that out because even after full screening the plot, my x-axis was so mushed together that I couldn't read the labels. There is probably a way to fix this in R by resizing the axis labels... NOEL is an acronym for "no-observable-effect-level" and indicates that the highest concentration did not produce effects significantly different from the responses for the control group. Used for terrestrial databases. LOEL is an acronym for "lowest-observable-effect-level" and indicates that the lowest dose produced effects that were significantly different from the responses for the control group.

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] "character"
```

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

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

```
unique(Litter$collectDate)
```

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

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`?

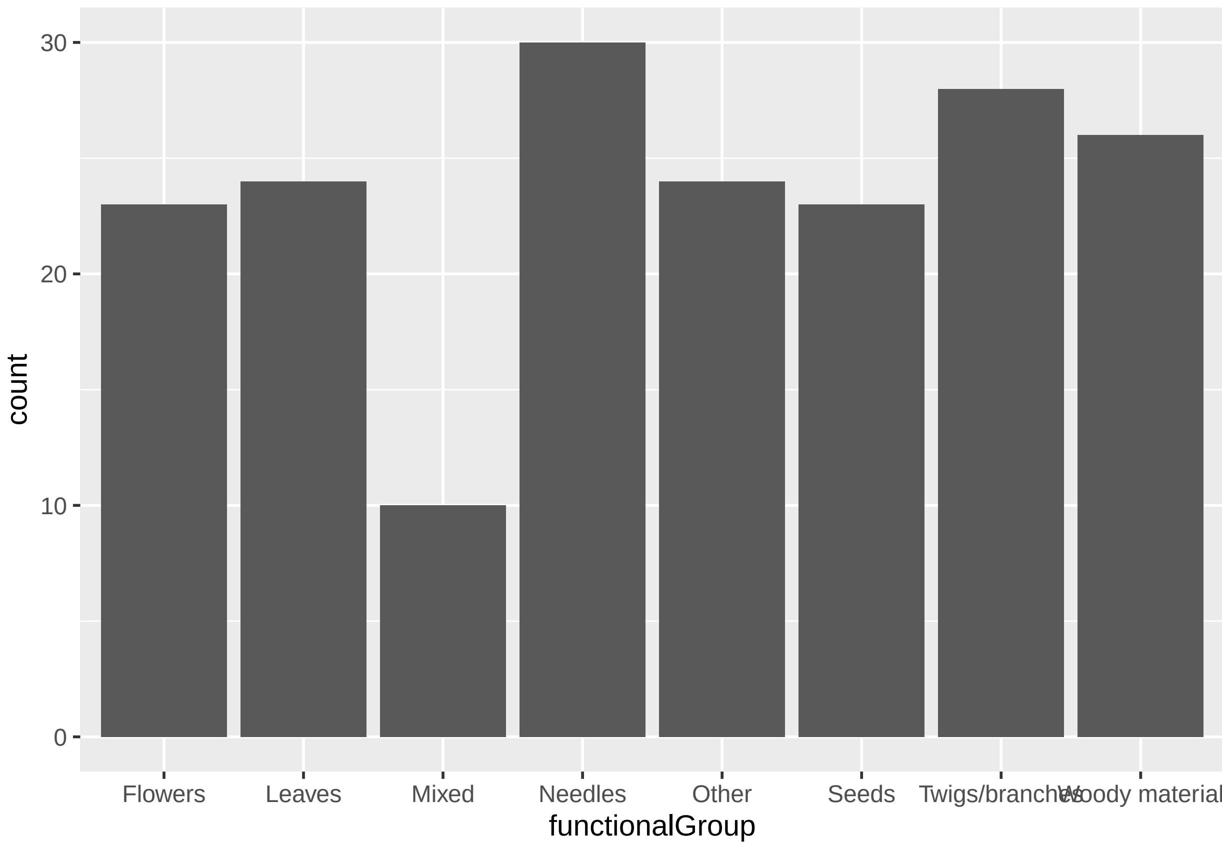
```
unique(Litter$plotID)
```

```
##  [1] "NIWO_061" "NIWO_064" "NIWO_067" "NIWO_040" "NIWO_041" "NIWO_063"
##  [7] "NIWO_047" "NIWO_051" "NIWO_058" "NIWO_046" "NIWO_062" "NIWO_057"
```

Answer: 12 plots were sampled at Niwot Ridge. `summary` gives you a count of how many samples were taken from each plot while `unique` simply tells you which plots were sampled without providing any information about the number of times they were sampled.

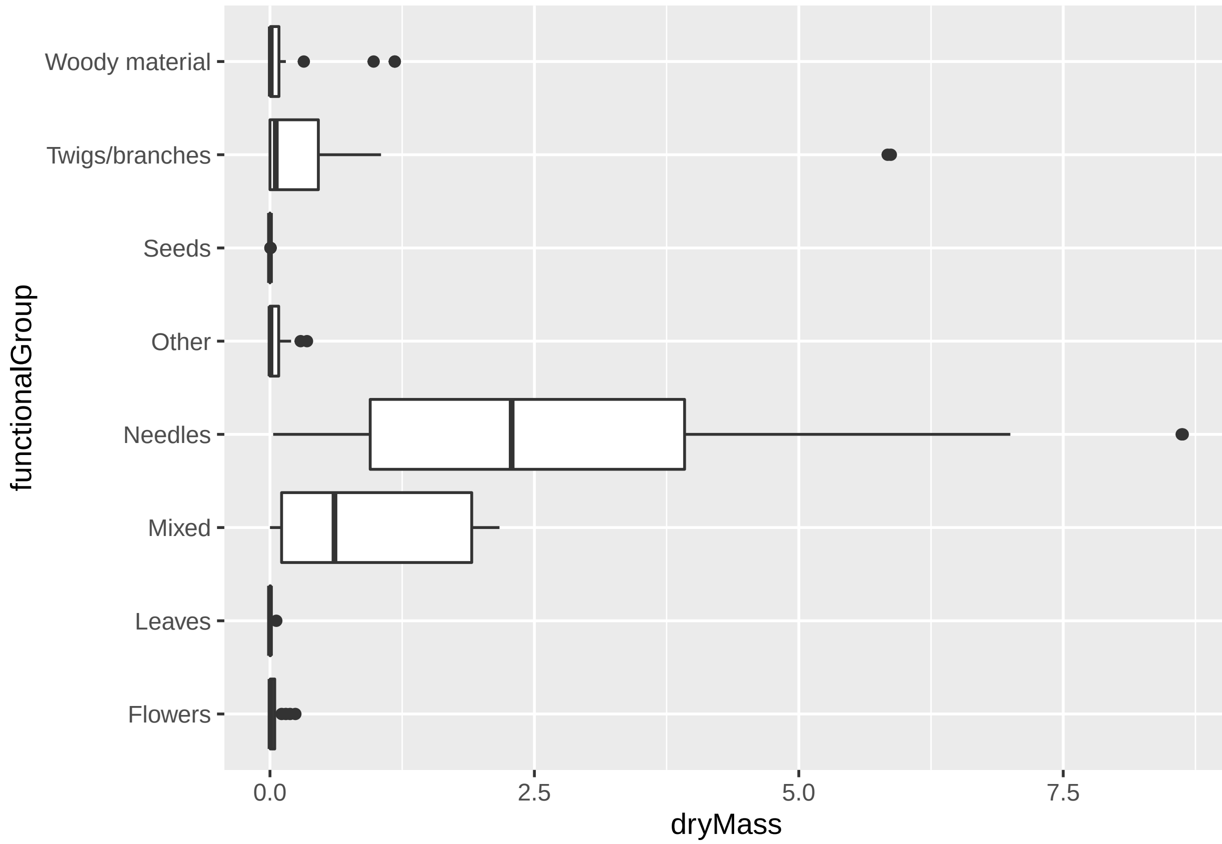
14. Create a bar graph of functionalGroup counts. This shows you what type of litter is collected at the Niwot Ridge sites. Notice that litter types are fairly equally distributed across the Niwot Ridge sites.

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

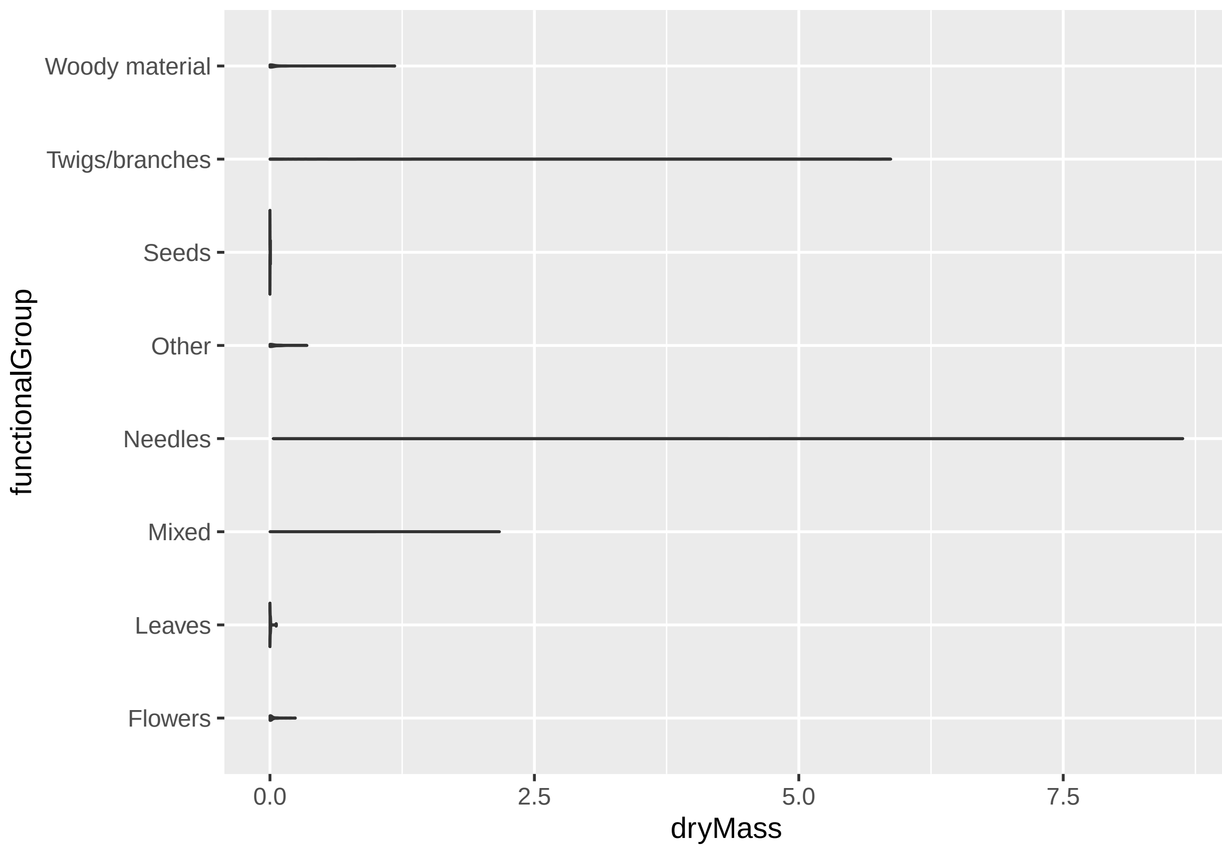


15. Using `geom_boxplot` and `geom_violin`, create a boxplot and a violin plot of dryMass by functionalGroup.

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




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



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

> Answer: The violin plot doesn't allow us to visualize the data well at all. While both boxplots and violin plots display distributions of continuous variables, the violin plot adds the `geom_density` function, which is useful for data that comes from an underlying smooth distribution but probably not as useful for approximately unimodal data.

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

> Answer: Needles are the type of litter with the highest biomass at these sites.