

Assignment 3: Data Exploration

Fiona Price

Fall 2024

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 Canvas.

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. Load necessary packages (tidyverse, lubridate, here), check your current working directory 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 sub-command to read strings in as factors.

```
#Load the dplyr, tidyverse, ggplot2, and here package
library(dplyr)
library(ggplot2)
library(here)
library(tidyverse)

#Upload the Neonics dataset
neonics <- read.csv(
  file = here("./Data/Raw/ECOTOX_Neonicotinoids_Insects_raw.csv"),
  stringsAsFactors = TRUE)
```

```
#Upload the Litter dataset
litter <- read.csv(
  file = here("./Data/Raw/NEON_NIWO_Litter_massdata_2018-08_raw.csv"),
  stringsAsFactors = TRUE)
```

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 were originally introduced as a less harmful insecticide. However, research now shows that they are harmful to insects. For example, they are very harmful to bees. The chemicals are absorbed by plants and are thus present in pollen and nectar. Additionally, they have long-lasting lifespans in plants. Overall, we are interested in the ecotoxicology of neonicotinoids on insects because they will affect insect populations, which will then have a cascading effect on ecosystems, as many of these insects are beneficial.

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: Litter and woody debris has many important ecological functions. First, litter and woody debris breaks down organic matter and returns nutrients back into the soil. Second, they provide habitat for organisms (including insects), helping to sustain biodiversity. Third, litter and woody debris helps retain moisture in the soil which is necessary for plant growth. They can also stabilize soils, helping to prevent erosion. Litter and woody debris are crucial for preserving biodiversity in ecosystems.

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 woody debris are collected from ground traps and elevated traps. These samples are then weighed. 2. Traps are placed in both targeted and randomized ways, depending on vegetation. Traps are only placed in tower plots. 3. Ground traps are sampled once per year, while elevated traps are sampled at varying frequencies depending on amount of vegetation.

Obtain basic summaries of your data (Neonics)

5. What are the dimensions of the dataset?

```
#Use the `dim` function to find the number of rows and columns in the dataset
dim(neonics)
```

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

```
#The dataset has 4623 rows and 30 columns.
```

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? [Tip: The `sort()` command is useful for listing the values in order of magnitude...]

```
#Use the `summary` function to look at an overview of the types of effects
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
```

```
#Create a table summarizing these effects
neonics_summary <- table(neonics$Effect)
```

```
#Now, sort the table using the `sort` function. Decreasing = TRUE indicates that
#the sort should be decreasing (so it starts with the most common effect).
sorted_neonics_summary <- sort(neonics_summary, decreasing = TRUE)
sorted_neonics_summary
```

```
##
##      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
```

```
#Population is the most common effect. Mortality is the second most common.
```

Answer: Population is the most common effect. Mortality is the second most common. These effects are specifically of interest because they reflect the impact that neonicotinoids have on insect population sizes. This data can be useful for predicting ecosystem dynamics.

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: Explore the help on the `summary()` function, in particular the `maxsum` argument...]

#Use the summary function to look at an overview of the common species names.
`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 Woolly 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

```
#Create a table summarizing these species.
neonics_species <- table(neonics$Species.Common.Name)
```

```
#Now, sort the table using the sort function. Decreasing = TRUE indicates that
#the sort should be decreasing (so it starts with the most common species).
sorted_neonics_species <- sort(neonics_species, decreasing = TRUE)
sorted_neonics_species
```

```
##
##              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 Woolly 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	Eleven-spotted 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	Horn-faced 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		

#The six most commonly studied species are the honey bee, parasitic wasp, buff tailed bumblebee, carniolan honey bee, bumble bee, and Italian honeybee.

#This is another way to do it using maxsum. Here, I'm indicating I want seven #levels to show (because the 7th will be (other)).

`summary(neonics$Species.Common.Name, maxsum = 7)`

##	Honey Bee	Parasitic Wasp	Buff Tailed Bumblebee
##	667	285	183
##	Carniolan Honey Bee	Bumble Bee	Italian Honeybee
##	152	140	113
##	(Other)		

```
##
```

```
3083
```

Answer: The six most commonly studied species are the honey bee, parasitic wasp, buff tailed bumblebee, carniolan honey bee, bumble bee, and Italian honeybee. Most of these species are pollinators; they are therefore of more interest because pollination is crucial for plant reproduction. Although parasitic wasps are not pollinators, they too are can indirectly contribute to plant reproduction.

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? [Tip: Viewing the dataframe may be helpful...]

```
#Use the `view` function to view the dataframe.
```

```
view(neonics)
```

```
#Use the `class` function to view the class of the `Conc.1..Author` column.
```

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

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

```
#The class is factor.
```

```
#Examine the specific categories.
```

```
summary(neonics$Conc.1..Author.)
```

```
##      0.37/      10/      NR/      NR      1      1023      0.40/      2/
##      208      127      108      94      82      80      69      63
##      10      0.053/      100      50/      0.5/      0.03      0.05/      0.45
##      62      59      56      51      45      44      43      43
##      0.1/      0.45/      1.0/      2.27/      50      0.125      500/      0.5
##      42      40      40      40      36      33      33      32
##      0.048/      0.15/      1/      48      25.0/      12/      0.027      2.4
##      30      30      30      30      28      27      26      26
##      0.2/      0.56/      100/      3      0.01/      1000/      3/      0.336
##      25      24      23      23      22      22      22      21
##      1.5/      0.05      1.5      2.60/      20.0/      6      6.80/      62.5/
##      21      20      20      20      20      20      20      20
##      0.005      0.4/      0.18/      0.3/      1000      40      0.00355/      0.1
##      18      18      17      17      17      17      16      16
##      0.4      150/      300      80/      0.053      0.24      0.28      125/
##      16      16      16      16      15      15      15      15
##      9      0.0001      0.0004/      0.084/      0.15      0.6      12.5/      144.0/
##      15      14      14      14      14      14      14      14
##      350/      40.0/      48/      56      84/      0.17/      125      14
##      14      14      14      14      14      13      13      13
##      16      17      0.047/      0.25/      0.28/      1.28/      1.81/      112
##      13      13      12      12      12      12      12      12
##      150      2.5/      25      60/      75/      0.02/      0.025/      0.29
##      12      12      12      12      12      11      11      11
##      37.5/      4/      5      (Other)
##      11      11      11      1817
```

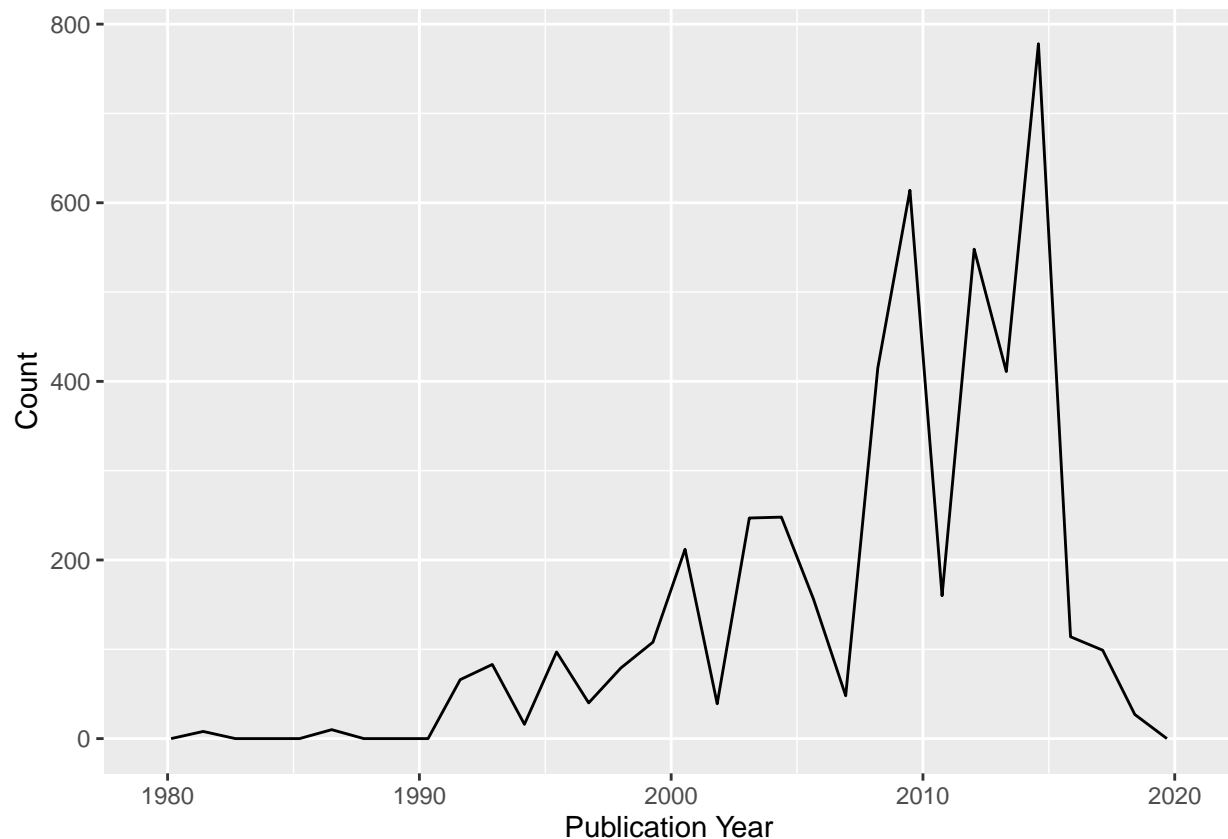
Answer: The class is factor. It is not numeric. Factors are categorical; because some of the data entries include “/” after the number, R reads this data as a factor rather than a number.

Explore your data graphically (Neonics)

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

```
ggplot(neonics, aes(Publication.Year)) + #Specify which data frame and which
#column to use along the x-axis
  geom_freqpoly() + #Specify the type of plot to be used
  labs(x = "Publication Year", y = "Count") #Change axis labels
```

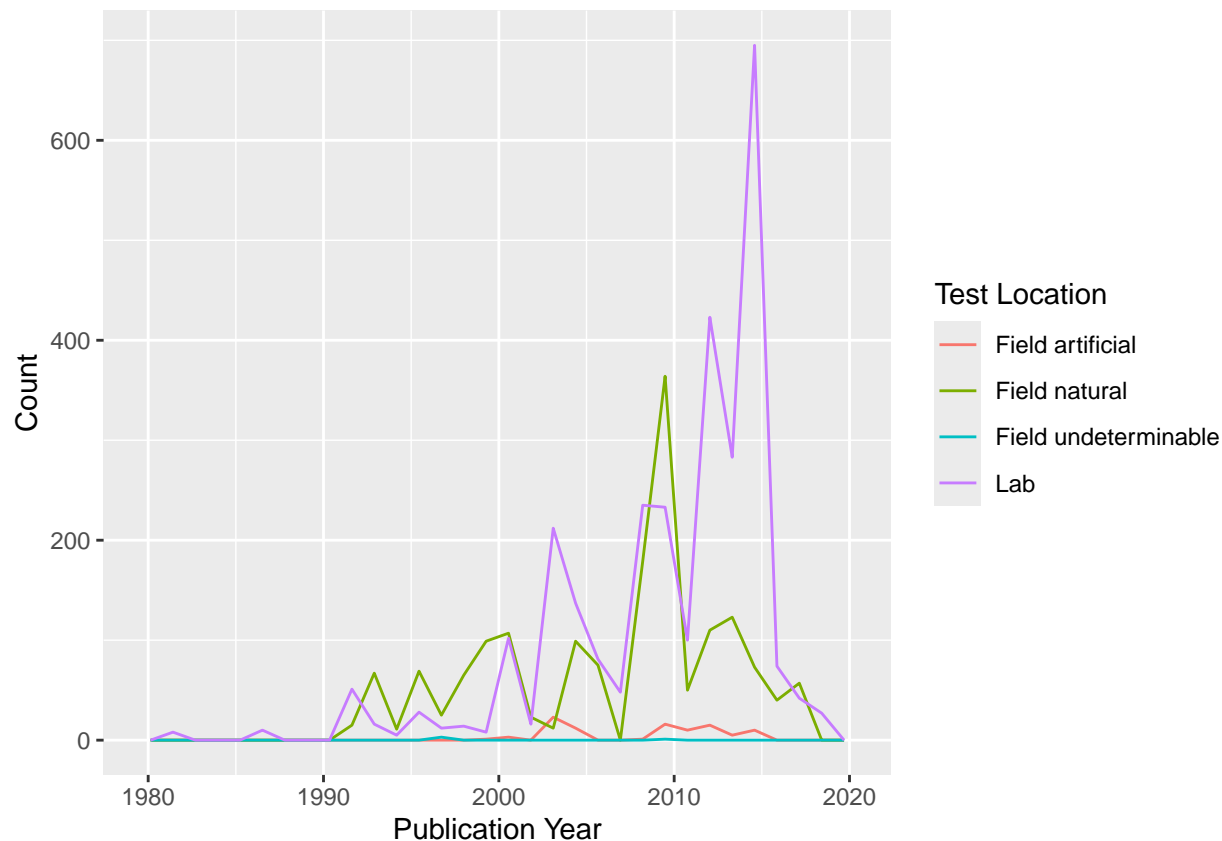
'stat_bin()' using 'bins = 30'. Pick better value with 'binwidth'.



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

```
ggplot(neonics, aes(Publication.Year, color = Test.Location)) + #Specify which
#data frame and which column to use along the x-axis and that color should
#vary based on test location
  geom_freqpoly() + #Specify the type of plot to be used
  labs(x = "Publication Year", y = "Count", color = "Test Location") #Change the
```

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



#axis labels and legend title

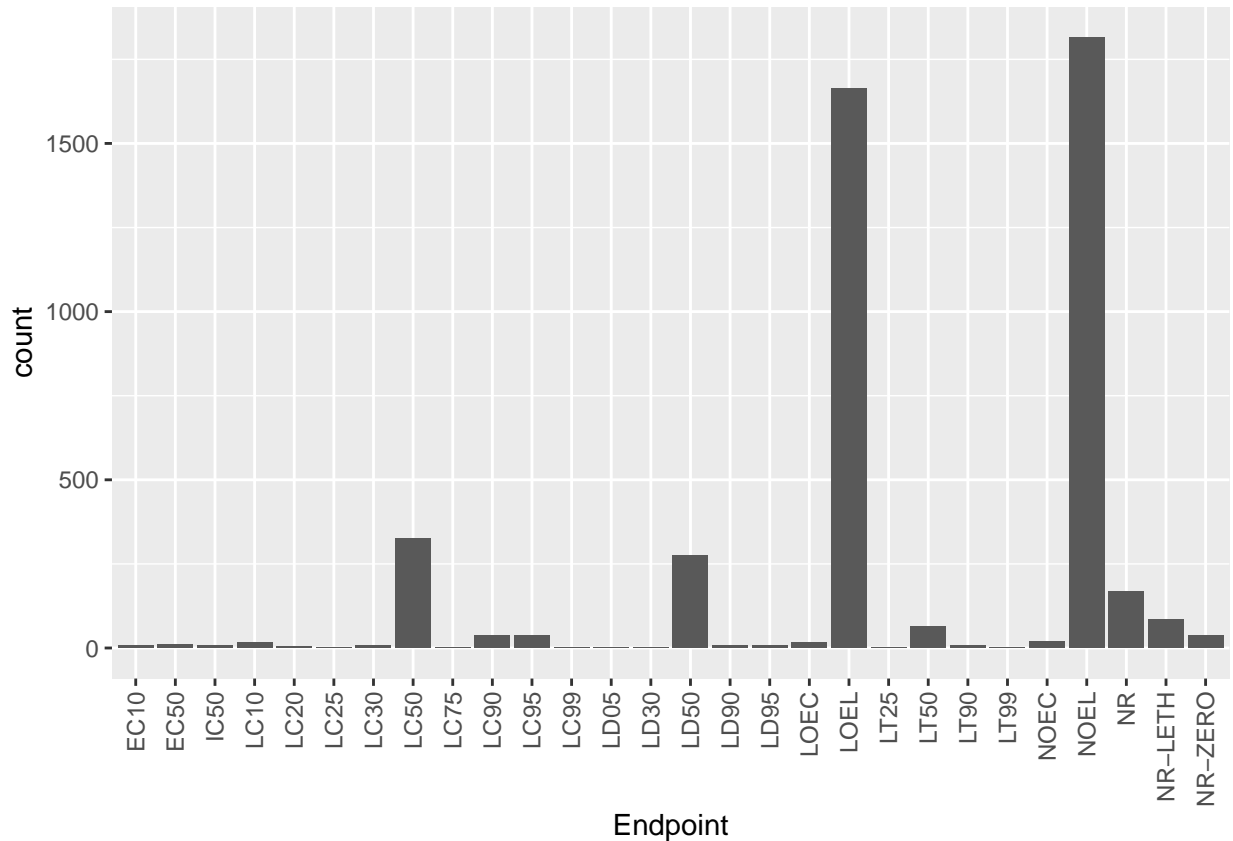
Interpret this graph. What are the most common test locations, and do they differ over time?

Answer: The most common test locations are the lab and natural field. Natural field peaked right before 2010 and then dropped; lab peaked in about 2015 and then dropped drastically. Both natural field and lab were far less common before the 2000s.

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(Endpoint)) + #Specify which
  #data frame and which column to use along the x-axis.
  geom_bar() + #Specify we want to use a bar graph.
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) #This
```



#rotates and aligns and the x-axis labels so we can see them clearly.

Answer: The two most common endpoints are LOEL and NOEL. LOEL is defined as the Lowest-observable-effect-level: lowest dose (concentration) producing effects that were significantly different (as reported by authors) from responses of controls (LOEAL/LOEC). NOEL is defined as no-observable-effect-level: highest dose (concentration) producing effects not significantly different from responses of controls according to author's reported statistical test (NOEAL/NOEC).

Explore your data (Litter)

- 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.

#Use the `class` function to determine the class of collectDate.

```
class(litter$collectDate)
```

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

#The class is factor,

#Change the class of collectDate to date using the `as.Date` function.

```
litter$collectDate <- as.Date(litter$collectDate)
```

#Confirm the new class of the variable.

```
class(litter$collectDate)
```

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

```
#The class is now date.
```

```
#Use the `unique` function to determine which dates litter were sampled in  
#August 2018.
```

```
unique(litter$collectDate)
```

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

```
#Sampling occurred on August 2nd and August 30th.
```

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

```
#Use `unique` to determine how many different plots were sampled at Niwot Ridge.
```

```
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
```

```
#There are 12 levels, so 12 different plots were sampled.
```

```
#Use `summary` to determine how many different plots were sampled at Niwot Ridge.
```

```
summary(litter$namedLocation)
```

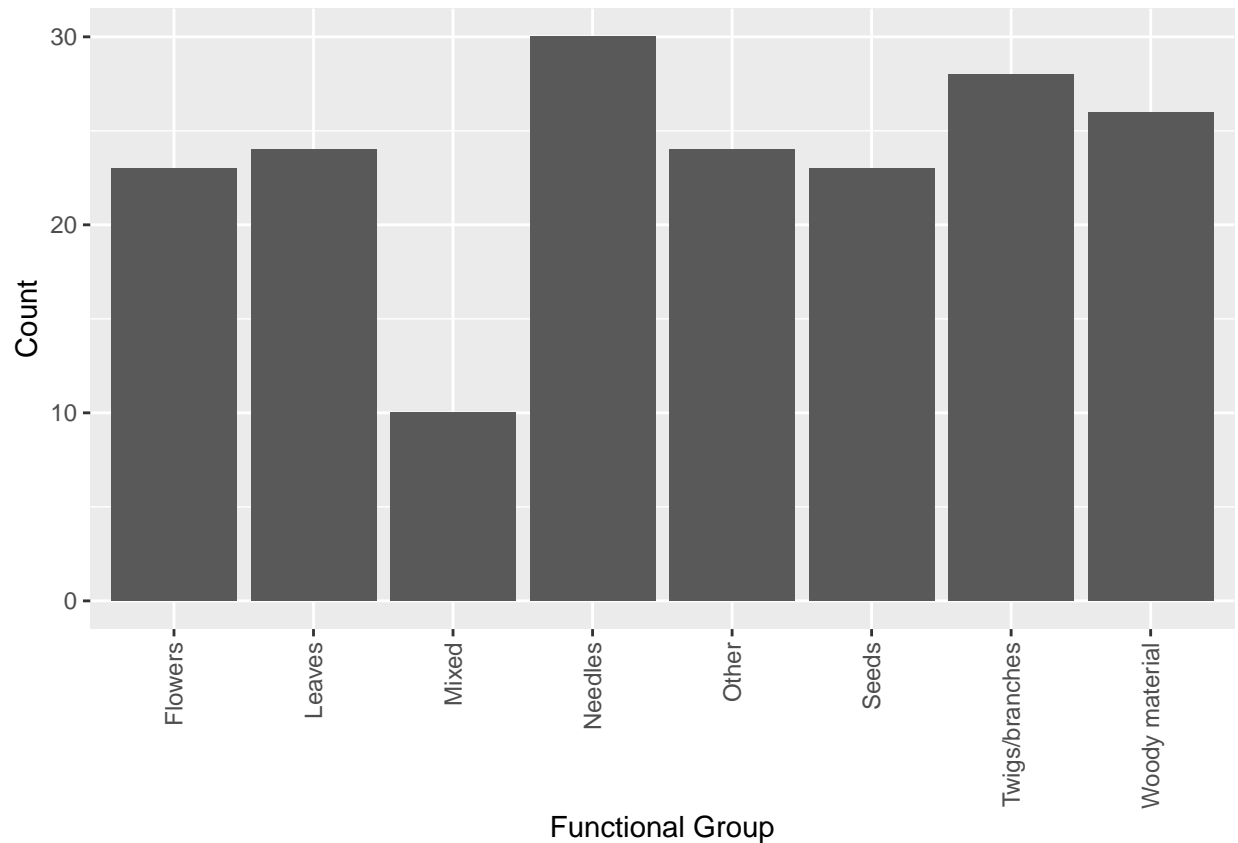
```
## NIWO_040.basePlot.ltr NIWO_041.basePlot.ltr NIWO_046.basePlot.ltr  
##                20                19                18  
## NIWO_047.basePlot.ltr NIWO_051.basePlot.ltr NIWO_057.basePlot.ltr  
##                15                14                8  
## NIWO_058.basePlot.ltr NIWO_061.basePlot.ltr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

```
#This returned the sample counts at each plot.
```

Answer: `unique` returns the number of levels of our data (12). `summary` returns the number of counts of each level. We can still find the number of levels using `summary` by manually counting the number of categories in the output.

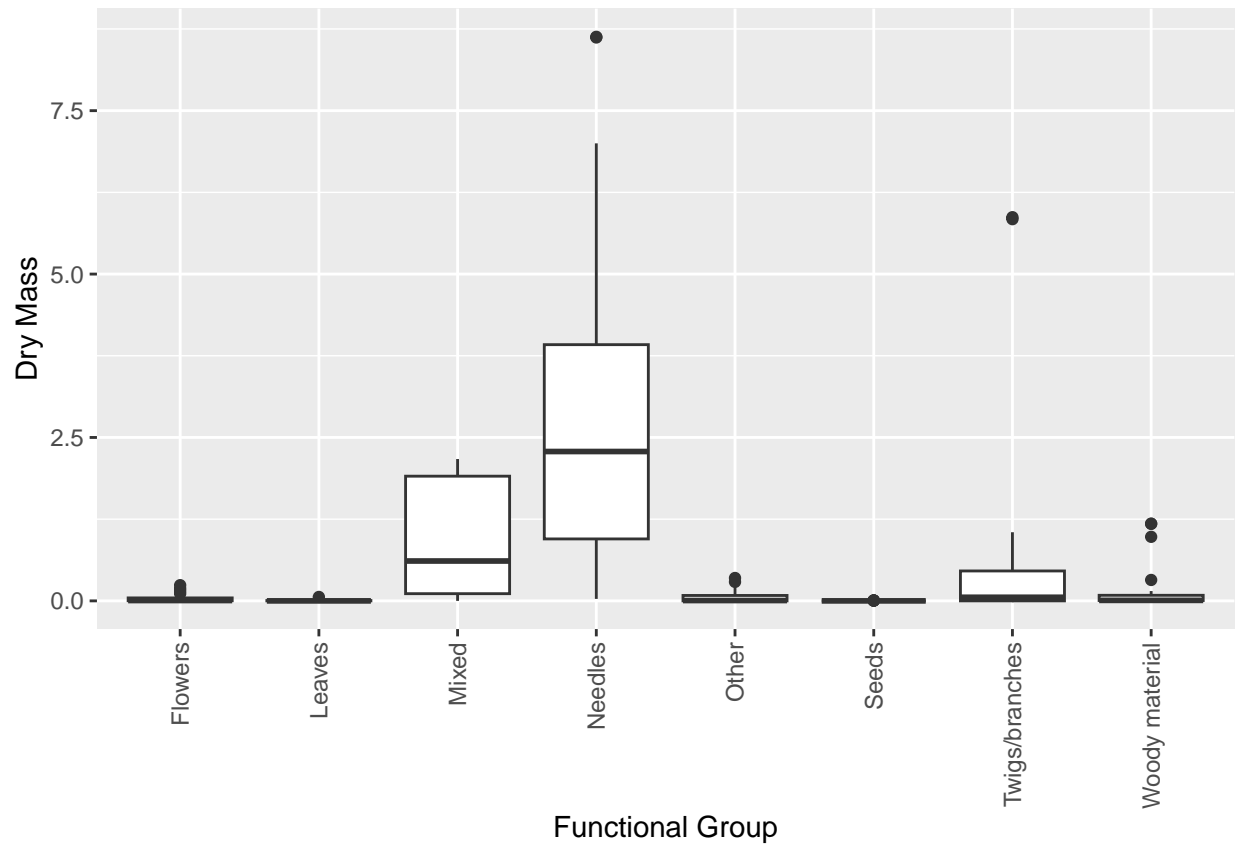
14. Create a bar graph of functionalGroup counts. This shows you what type of litter is collected at the Niwot Ridge sites. Notice that litter types are fairly equally distributed across the Niwot Ridge sites.

```
ggplot(litter, aes(functionalGroup)) + #Specify which  
#data frame and which column to use along the x-axis.  
geom_bar() + #Specify we want to use a bar graph.  
theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) + #This  
#rotates and aligns and the x-axis labels so we can see them clearly.  
labs(x = "Functional Group", y = "Count") #Change axis labels.
```

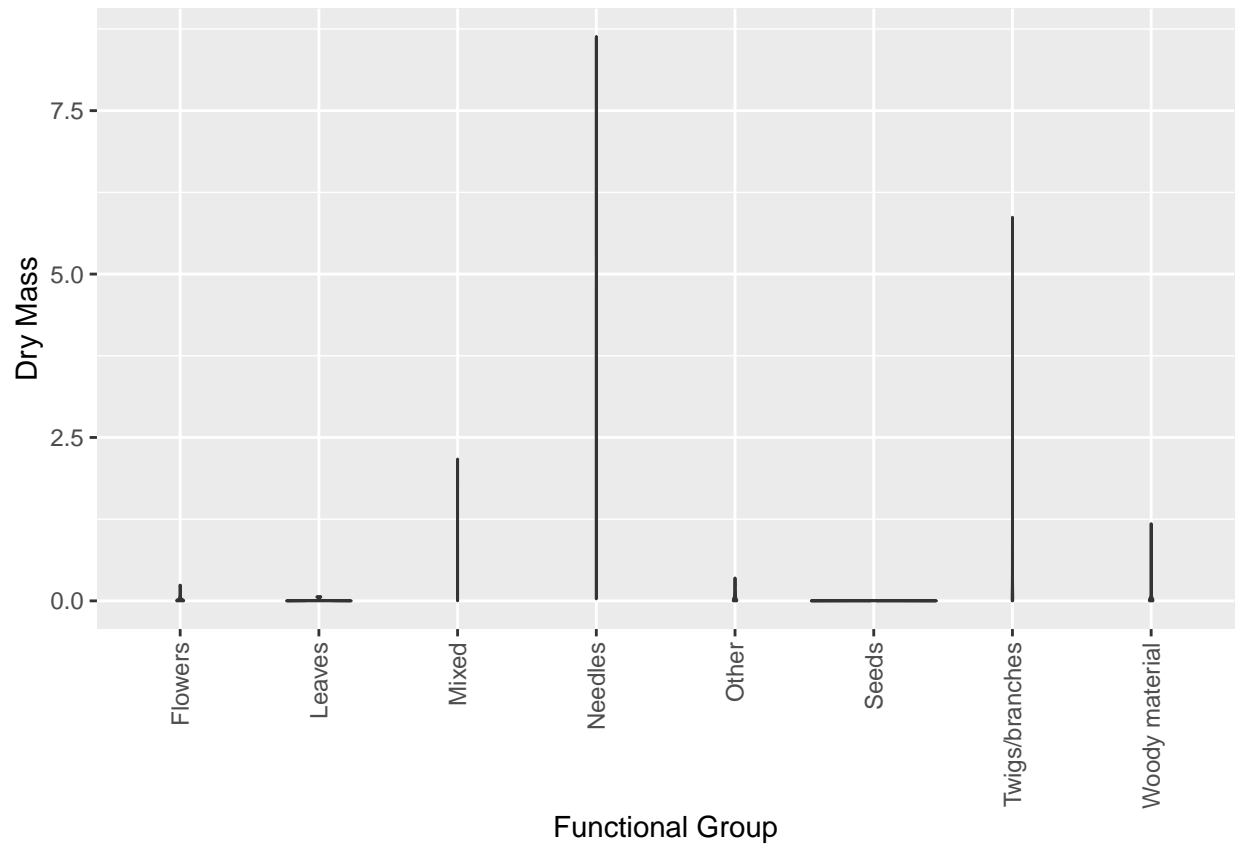



15. Using `geom_boxplot` and `geom_violin`, create a boxplot and a violin plot of `dryMass` by functional-Group.

```
#Start with boxplot:  
ggplot(litter) + #Specify which data frame to use.  
  geom_boxplot(aes(x = functionalGroup, y = dryMass)) + #Specify we want to use  
  #a boxplot with functional group on the x-axis and dry mass on the y-axis  
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) + #This  
  #rotates and aligns and the x-axis labels so we can see them clearly.  
  labs(x = "Functional Group", y = "Dry Mass") #Change axis labels.
```



```
#Now a violin plot:
ggplot(litter) + #Specify which data frame to use.
  geom_violin(aes(x = functionalGroup, y = dryMass)) + #Specify we want to use
  #a violin plot with functional group on the x-axis and dry mass on the y-axis
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1)) + #This
  #rotates and aligns and the x-axis labels so we can see them clearly.
  labs(x = "Functional Group", y = "Dry Mass") #Change axis labels.
```



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

Answer: The violin plot displays density distributions. However, the width #only changes with a continuous variable along the x-axis. Since functional group is categorical, the “violin” does not have a width.

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

Answer: Needles tend to have the highest biomass at these sites.