Assignment 3: Data Exploration

Andrew Brantley, Section #1

OVERVIEW

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

Directions

1. Change "Student Name, Section #" on line 3 (above) with your name and section number.
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., "FirstLast_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. **Be sure to add the `stringsAsFactors = TRUE` parameter to the function when reading in the CSV files.**

```r
#checking working directory
getwd()
```

```
## [1] "/Users/AndrewBrantley/Library/CloudStorage/Box-Box/Environmental Data Analytics/GithubRepos/Envi
```

```r
#loading necessary packages
library("tidyverse")

#uploading datasets
Neonics <- read.csv("../Data/Raw/ECOTOX_Neonicotinoids_Insects_raw.csv", stringsAsFactors = TRUE)

Litter <- read.csv("../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 ecotoxicologoy of neonicotinoids on insects? Feel free to do a brief internet search if you feel you need more background information.

   Answer: Understanding to what degree different types of insects are affected could be important in knowing if target pest insects are being eliminated or if non-target insects are being affected.

For example, if non-target pollinators are being affected by the insecticide this would be counter-productive to farmers working to produce more crop so ensuring that target species are the only ones being affected would be vital information for users of the insecticide.

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: Especially in Western states such as Colorado, forest litter that falls to the ground will often eventually become fuel for forest fires. Tracking the amount of litter being deposited in different areas of forest would be important information in modeling/predicting fire location, size, and intensity in future fire seasons as well as tracking current fires.

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: * Falling litter and woody debris are monitored through both elevated PVC traps and ground traps * The sampling takes place in areas with woody vegetation >2 meters tall * Trap placements can be either targeted or randomized depending on the vegetation within the plot

Obtain basic summaries of your data (Neonics)

5. What are the dimensions of the dataset?

```r
#dimensions of Neonics dataset
dim(Neonics)
```

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

```r
#other summaries
head(Neonics)
```

```
##   CAS.Number                                Chemical.Name
## 1   58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
## 2   58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
## 3   58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
## 4   58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
## 5   58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
## 6   58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
##                                              Chemical.Grade
## 1 Technical grade, technical product, technical formulation
## 2 Technical grade, technical product, technical formulation
## 3 Technical grade, technical product, technical formulation
## 4 Technical grade, technical product, technical formulation
## 5 Technical grade, technical product, technical formulation
## 6 Technical grade, technical product, technical formulation
##   Chemical.Analysis.Method Chemical.Purity Species.Scientific.Name
## 1               Unmeasured              99  Araecerus fasciculatus
## 2               Unmeasured              99  Araecerus fasciculatus
## 3               Unmeasured              95         Musca domestica
## 4               Unmeasured              95         Musca domestica
## 5               Unmeasured              95         Musca domestica
## 6               Unmeasured              95         Musca domestica
##   Species.Common.Name   Species.Group Organism.Lifestage Organism.Age
## 1  Coffee Bean Weevil Insects/Spiders              Adult           NR
```

```
## 2  Coffee Bean Weevil Insects/Spiders             Adult           NR
## 3           House Fly Insects/Spiders             Young           NR
## 4           House Fly Insects/Spiders             Young           NR
## 5           House Fly Insects/Spiders             Young           NR
## 6           House Fly Insects/Spiders             Adult            9
##   Organism.Age.Units    Exposure.Type   Media.Type Test.Location
## 1       Not reported Topical, general No substrate           Lab
## 2       Not reported Topical, general No substrate           Lab
## 3            Hour(s)             Food Filter paper           Lab
## 4            Hour(s)             Food Filter paper           Lab
## 5            Hour(s)             Food Filter paper           Lab
## 6             Day(s)             Food Filter paper           Lab
##   Number.of.Doses Conc.1.Type..Author. Conc.1..Author. Conc.1.Units..Author.
## 1              NR    Active ingredient           27.2             ug/g bdwt
## 2              NR    Active ingredient           19.7             ug/g bdwt
## 3              11    Active ingredient             47                  mg/L
## 4              11    Active ingredient             25                  mg/L
## 5              11    Active ingredient             13                  mg/L
## 6              11    Active ingredient            268                  mg/L
##      Effect Effect.Measurement Endpoint Response.Site Observed.Duration..Days.
## 1 Mortality          Mortality     LD50  Not reported                        1
## 2 Mortality          Mortality     LD50  Not reported                        1
## 3 Mortality          Mortality     LC50  Not reported                        1
## 4 Mortality          Mortality     LC50  Not reported                        1
## 5 Mortality          Mortality     LC50  Not reported                        1
## 6 Mortality          Mortality     LC50  Not reported                        1
##   Observed.Duration.Units..Days.
## 1                         Day(s)
## 2                         Day(s)
## 3                         Day(s)
## 4                         Day(s)
## 5                         Day(s)
## 6                         Day(s)
##                                                                   Author
## 1                                     Childers,C.C., and H.N. Nigg
## 2                                     Childers,C.C., and H.N. Nigg
## 3 Johnston,A.M., J. Lohr, J. Moes, K.R. Solomon, and E.R. Zaborski
## 4 Johnston,A.M., J. Lohr, J. Moes, K.R. Solomon, and E.R. Zaborski
## 5 Johnston,A.M., J. Lohr, J. Moes, K.R. Solomon, and E.R. Zaborski
## 6 Johnston,A.M., J. Lohr, J. Moes, K.R. Solomon, and E.R. Zaborski
##   Reference.Number
## 1           107388
## 2           107388
## 3           103312
## 4           103312
## 5           103312
## 6           103312
##
## 1                                                                       Contact Toxicity of Insectic
## 2                                                                       Contact Toxicity of Insectic
## 3 Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Suscep
## 4 Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Suscep
## 5 Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Suscep
## 6 Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Suscep
```

```
##                              Source Publication.Year
## 1   J. Econ. Entomol.75(3): 556-559             1982
## 2   J. Econ. Entomol.75(3): 556-559             1982
## 3 J. Econ. Entomol.79(6): 1439-1442             1986
## 4 J. Econ. Entomol.79(6): 1439-1442             1986
## 5 J. Econ. Entomol.79(6): 1439-1442             1986
## 6 J. Econ. Entomol.79(6): 1439-1442             1986
##
## 1 Purity: \xca NR - NR | Organism Age: \xca NR - NR Not reported | Conc 1 (Author): \xca Active ingre
## 2 Purity: \xca NR - NR | Organism Age: \xca NR - NR Not reported | Conc 1 (Author): \xca Active ingre
## 3              Purity: \xca NR - NR | Organism Age: \xca 24 - 48 Hour(s) | Conc 1 (Author): \xca Act
## 4              Purity: \xca NR - NR | Organism Age: \xca 24 - 48 Hour(s) | Conc 1 (Author): \xca Act
## 5               Purity: \xca NR - NR | Organism Age: \xca 24 - 48 Hour(s) | Conc 1 (Author): \xca Ac
## 6             Purity: \xca NR - NR | Organism Age: \xca NR - NR Day(s) | Conc 1 (Author): \xca Activ
```

```
tail(Neonics)
```

```
##      CAS.Number
## 4618  210880925
## 4619  210880925
## 4620  210880925
## 4621  210880925
## 4622  210880925
## 4623  210880925
##                                                                 Chemical.Name
## 4618 [C(E)]-N-[(2-Chloro-5-thiazolyl)methyl]-N'-methyl-N''-nitroguanidine
## 4619 [C(E)]-N-[(2-Chloro-5-thiazolyl)methyl]-N'-methyl-N''-nitroguanidine
## 4620 [C(E)]-N-[(2-Chloro-5-thiazolyl)methyl]-N'-methyl-N''-nitroguanidine
## 4621 [C(E)]-N-[(2-Chloro-5-thiazolyl)methyl]-N'-methyl-N''-nitroguanidine
## 4622 [C(E)]-N-[(2-Chloro-5-thiazolyl)methyl]-N'-methyl-N''-nitroguanidine
## 4623 [C(E)]-N-[(2-Chloro-5-thiazolyl)methyl]-N'-methyl-N''-nitroguanidine
##      Chemical.Grade Chemical.Analysis.Method Chemical.Purity
## 4618   Not reported              Unmeasured             NR
## 4619   Not reported              Unmeasured             NR
## 4620   Not reported              Unmeasured             50
## 4621   Not reported              Unmeasured             50
## 4622   Not reported            Not reported             NR
## 4623   Not reported              Unmeasured             50
##      Species.Scientific.Name    Species.Common.Name
## 4618     Tomarus subtropicus         Sugarcane Grub
## 4619     Tomarus subtropicus         Sugarcane Grub
## 4620   Trichogramma chilonis          Parastic Wasp
## 4621   Trichogramma chilonis          Parastic Wasp
## 4622     Xyleborus glabratus Redbay Ambrosia Beetle
## 4623          Zetzellia mali         Predatory Mite
##                              Species.Group Organism.Lifestage Organism.Age
## 4618                       Insects/Spiders             Instar           NR
## 4619                       Insects/Spiders              Larva            3
## 4620                       Insects/Spiders              Adult           NR
## 4621                       Insects/Spiders              Adult           NR
## 4622 Insects/Spiders; U.S. Invasive Species       Not reported           NR
## 4623                       Insects/Spiders       Not reported           NR
##      Organism.Age.Units              Exposure.Type      Media.Type
## 4618            Day(s)               Ground spray Artificial soil
## 4619            Instar               Ground spray    Not reported
```

```
## 4620      Not reported Environmental, unspecified    No substrate
## 4621      Not reported Environmental, unspecified    No substrate
## 4622      Not reported                      Spray    No substrate
## 4623      Not reported                      Spray    No substrate
##         Test.Location Number.of.Doses Conc.1.Type..Author. Conc.1..Author.
## 4618             Lab               2    Active ingredient           0.28/
## 4619 Field artificial              2    Active ingredient           0.28/
## 4620             Lab               7    Active ingredient          0.0113
## 4621             Lab               7    Active ingredient          3.4859
## 4622   Field natural               2          Formulation              6/
## 4623   Field natural               3          Formulation               6
##      Conc.1.Units..Author.     Effect Effect.Measurement Endpoint Response.Site
## 4618             AI kg/ha  Mortality          Mortality     LOEL  Not reported
## 4619             AI kg/ha  Mortality          Mortality     NOEL  Not reported
## 4620              AI mg/L  Mortality          Mortality     LC50  Not reported
## 4621              AI mg/L  Mortality          Mortality     LC95  Not reported
## 4622                   oz Population          Abundance     NOEL  Not reported
## 4623              oz/acre Population          Abundance     NOEL  Not reported
##      Observed.Duration..Days. Observed.Duration.Units..Days.
## 4618                       21                         Day(s)
## 4619                        7                         Day(s)
## 4620                       NR                         Day(s)
## 4621                       NR                         Day(s)
## 4622                       63                         Day(s)
## 4623                       72                         Day(s)
##
## 4618                                                                     Kostromyts
## 4619                                                                     Kostromyts
## 4620                                  Preetha,G., J. Stanley, S. Suresh, S. Kutta
## 4621                                  Preetha,G., J. Stanley, S. Suresh, S. Kutta
## 4622 Pena,J.E., J.H. Crane, J.L. Capinera, R.E. Duncan, P.E. Kendra, R.C. Ploetz, S. McLean, G. Brar
## 4623                                                                           Reis
##      Reference.Number
## 4618           165557
## 4619           165557
## 4620           150863
## 4621           150863
## 4622           156617
## 4623            93017
##
## 4618            Seasonal Phenology and Management of Tomarus subtropicus (Coleoptera: Scarabaeidae)
## 4619            Seasonal Phenology and Management of Tomarus subtropicus (Coleoptera: Scarabaeidae)
## 4620             Toxicity of Selected Insecticides to Trichogramma chilonis: Assessing Their Safety
## 4621             Toxicity of Selected Insecticides to Trichogramma chilonis: Assessing Their Safety
## 4622 Chemical Control of the Redbay Ambrosia Beetle, Xyleborus glabratus, and Other Scolytinae (Cole
## 4623                                      Evaluation of Seasonal Insecticide Programs Against New
##                            Source Publication.Year
## 4618 J. Econ. Entomol.101(6): 1847-1855            2008
## 4619 J. Econ. Entomol.101(6): 1847-1855            2008
## 4620      Phytoparasitica37(3): 209-215            2009
## 4621      Phytoparasitica37(3): 209-215            2009
## 4622        Fla. Entomol.94(4): 882-896            2011
## 4623      Arthropod Manag. Tests31:3 p.            2006
##
```

```
## 4618                         Purity: \xca NR - NR | Organism Age: \xca 1 - 3 Day(s) | Conc 1 (Author): \
## 4619                       Purity: \xca NR - NR | Organism Age: \xca NR - NR Instar | Conc 1 (Author): \
## 4620 Purity: \xca NR - NR | Organism Age: \xca NR - NR Not reported | Conc 1 (Author): \xca Active i
## 4621         Purity: \xca NR - NR | Organism Age: \xca NR - NR Not reported | Conc 1 (Author): \xca
## 4622                             Purity: \xca NR - NR | Organism Age: \xca NR - NR Not reported | Conc
## 4623                          Purity: \xca NR - NR | Organism Age: \xca NR - NR Not reported | Conc 1 (
```

```
str(Neonics)
```

```
## 'data.frame':    4623 obs. of  30 variables:
##  $ CAS.Number                      : int  58842209 58842209 58842209 58842209 58842209 58842209 5884
##  $ Chemical.Name                   : Factor w/ 9 levels "(1E)-N-[(6-Chloro-3-pyridinyl)methyl]-N-eth
##  $ Chemical.Grade                  : Factor w/ 9 levels "Analytical grade",..: 9 9 9 9 9 9 9 9 9 9 .
##  $ Chemical.Analysis.Method        : Factor w/ 5 levels "Measured","Not coded",..: 4 4 4 4 4 4 4 4 4
##  $ Chemical.Purity                 : Factor w/ 80 levels ">=98",">=99.0",..: 69 69 50 50 50 50 50 50
##  $ Species.Scientific.Name         : Factor w/ 398 levels "Acalolepta vastator",..: 69 69 248 248 24
##  $ Species.Common.Name             : Factor w/ 303 levels "Alfalfa Leafcutter Bee",..: 74 74 142 142
##  $ Species.Group                   : Factor w/ 4 levels "Insects/Spiders",..: 1 1 1 1 1 1 1 1 1 1 ..
##  $ Organism.Lifestage              : Factor w/ 20 levels "Adult","Cocoon",..: 1 1 19 19 19 1 19 1 1
##  $ Organism.Age                    : Factor w/ 39 levels "<=24","<=48",..: 39 39 39 39 39 36 39 36 3
##  $ Organism.Age.Units              : Factor w/ 11 levels "Day(s)","Days post-emergence",..: 9 9 4 4
##  $ Exposure.Type                   : Factor w/ 24 levels "Choice","Dermal",..: 23 23 11 11 11 11 11
##  $ Media.Type                      : Factor w/ 10 levels "Agar","Artificial soil",..: 7 7 3 3 3 3 3
##  $ Test.Location                   : Factor w/ 4 levels "Field artificial",..: 4 4 4 4 4 4 4 4 4 4 .
##  $ Number.of.Doses                 : Factor w/ 30 levels "' 4-5","' 4-7",..: 30 30 18 18 18 18 18 18
##  $ Conc.1.Type..Author.            : Factor w/ 3 levels "Active ingredient",..: 1 1 1 1 1 1 1 1 1 1
##  $ Conc.1..Author.                 : Factor w/ 1006 levels "<0.0004","<0.025",..: 639 510 813 622 44
##  $ Conc.1.Units..Author.           : Factor w/ 148 levels "%","% v/v","% w/v",..: 132 132 91 91 91 9
##  $ Effect                          : Factor w/ 19 levels "Accumulation",..: 16 16 16 16 16 16 16 16
##  $ Effect.Measurement              : Factor w/ 155 levels "Abundance","Accuracy of learned task, per
##  $ Endpoint                        : Factor w/ 28 levels "EC10","EC50",..: 15 15 8 8 8 8 8 8 8 8 ...
##  $ Response.Site                   : Factor w/ 19 levels "Abdomen","Brain",..: 14 14 14 14 14 14 14
##  $ Observed.Duration..Days.        : Factor w/ 361 levels "<.0002","<.0021",..: 145 145 145 145 145
##  $ Observed.Duration.Units..Days.  : Factor w/ 17 levels "Day(s)","Day(s) post-emergence",..: 1 1 1
##  $ Author                          : Factor w/ 433 levels "Abbott,V.A., J.L. Nadeau, H.A. Higo, and
##  $ Reference.Number                : int  107388 107388 103312 103312 103312 103312 103312 103312 10
##  $ Title                           : Factor w/ 458 levels "A Common Pesticide Decreases Foraging Suc
##  $ Source                          : Factor w/ 456 levels "Acta Hortic.1094:451-456",..: 295 295 296
##  $ Publication.Year                : int  1982 1982 1986 1986 1986 1986 1986 1986 1986 1986 ...
##  $ Summary.of.Additional.Parameters: Factor w/ 943 levels "Purity: \xca NC - NC | Organism Age: \xca
```

```
length(Neonics)
```

```
## [1] 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?

```
#summary of Effect column
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
```

Answer: The most commonly studied effects are mortality (1493) and population (1803). These are likely the most studied effects because they are the target result of these insecticides as they are meant to kill off insects that may be harming crops, lowering the poopulation.

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.

```
#summary of common 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 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
```

Answer: The six most commonly studied insects are as follows: Honey Bee, Parasitic Wasp, Buff Tailed Bumblebee, Carniolan Honey Bee, Bumble Bee, and Italian Honey Bee. These insects are largely responsible for pollination of plants around the globe making them of high interest as they are liekly not the target of these insecticide applications.

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

```r
#determining class of Conc.1..Author.
class(Neonics$Conc.1..Author.)
```

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

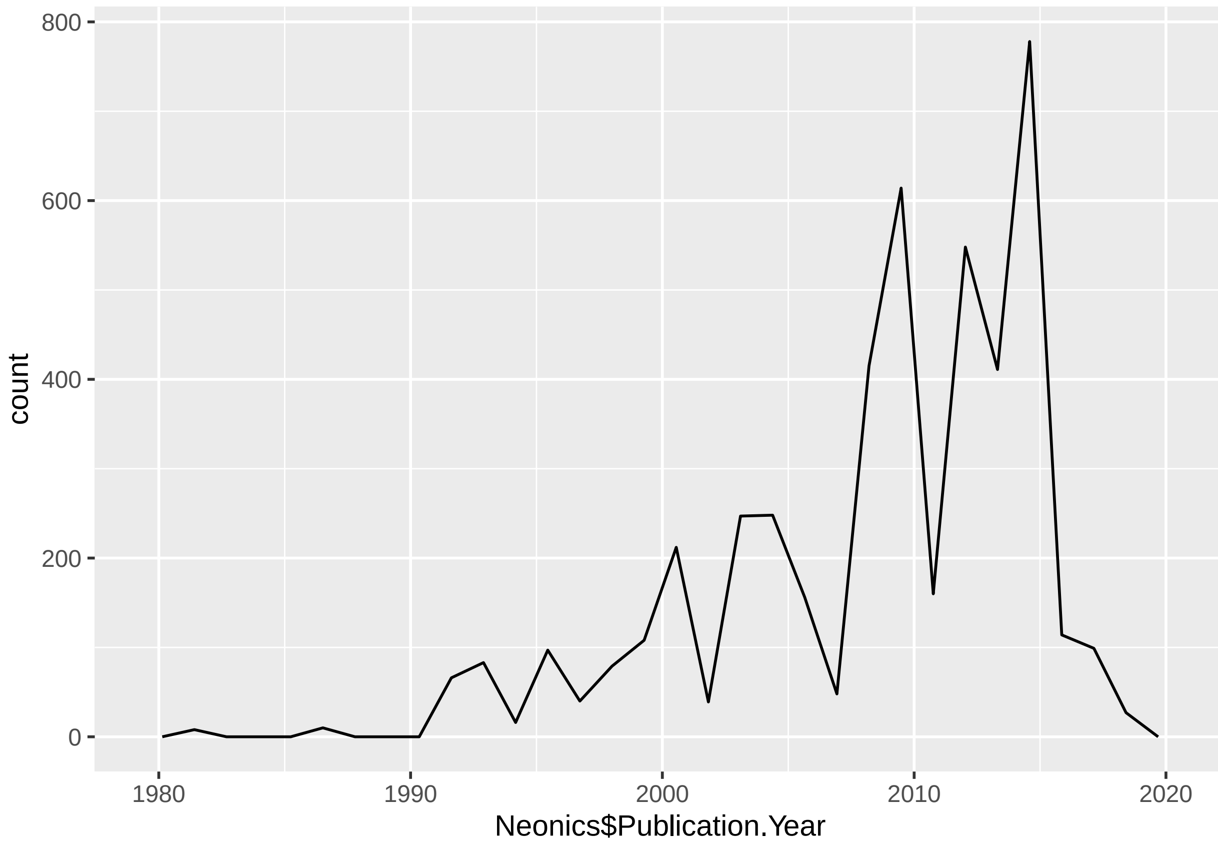
Answer: The class of Conc.1..Author. is factor and it is not numeric becuase of the argument we used when loading in the dataset "stringsAsFactors = TRUE". This column also features a variety of units which are paired with them in the next column over "Conc.Units.1..Author.".

Explore your data graphically (Neonics)

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

```r
#plot of number of studies each year (by publication)
ggplot(Neonics, aes(x=Neonics$Publication.Year)) +
  geom_freqpoly()
```

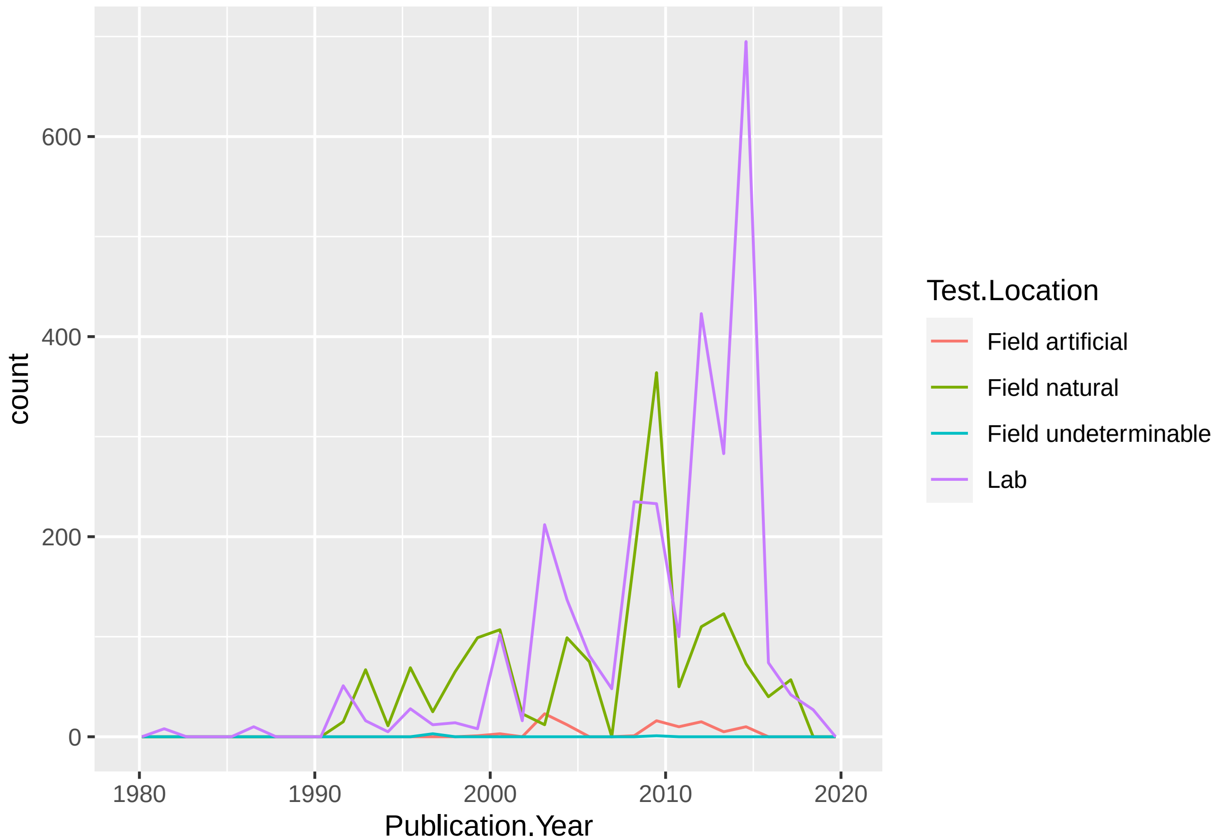
```
## Warning: Use of `Neonics$Publication.Year` is discouraged. Use
## `Publication.Year` instead.

## `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.

```
#plot of number of studies each year (by publication) with color based on Test.Location
ggplot(Neonics, aes(x=Publication.Year, color = Test.Location)) +
  geom_freqpoly()
```

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

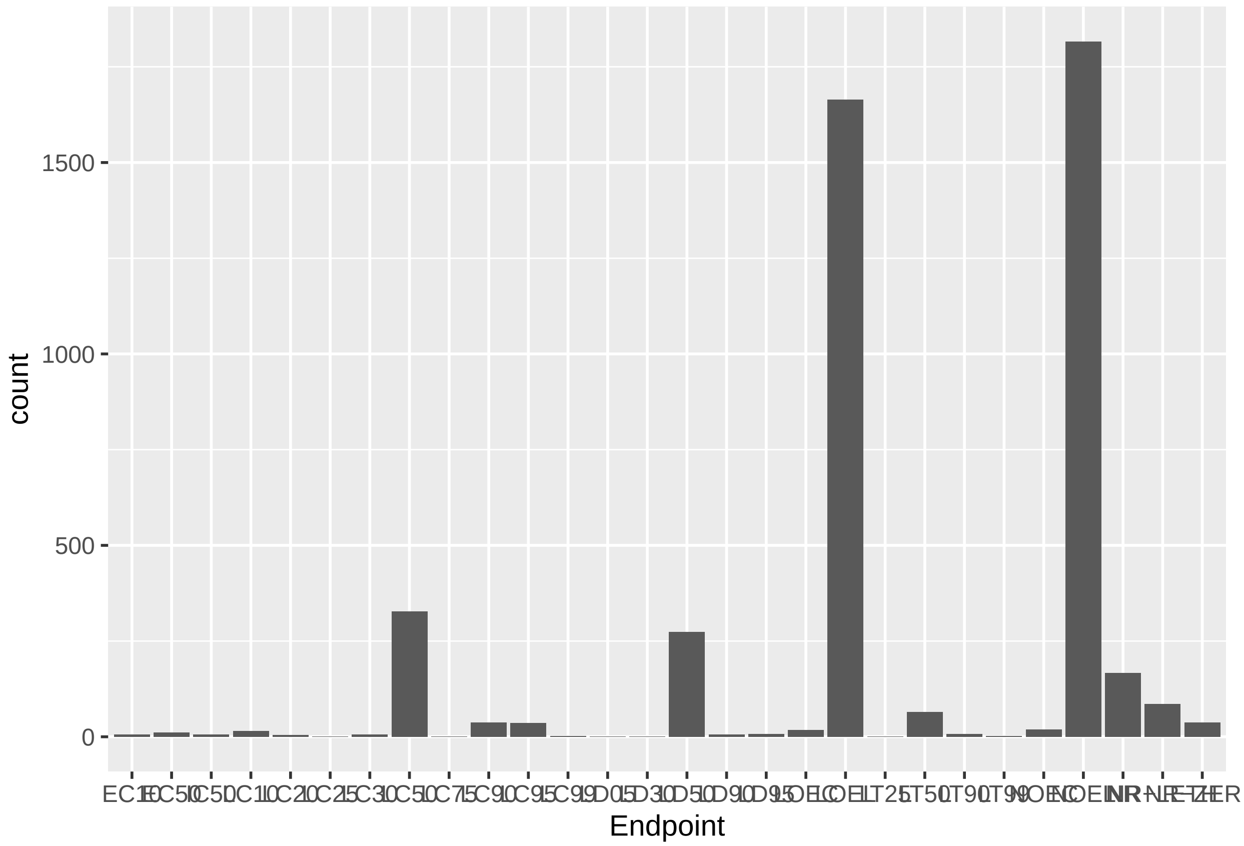


Interpret this graph. What are the most common test locations, and do they differ over time?

> Answer: The large majority of publications are now lab-based with field natural being the next most common. Prior to 2010 most publications in most years were actually field natural based with lab-based being the next most common.

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.

```
#barplot of end point counts
ggplot(Neonics, aes(x = Endpoint)) +
  geom_bar()
```



Answer: LOEL and NOEL are mcuh more common than any of the other endpoints. LOEL is the endpoint representing lowest-observable-effect-level and NOEL represents the endpoint of 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.

```
#checking class of collectDate
class(Litter$collectDate)
```

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

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

# confirming new class, confirmed to be Date
class(Litter$collectDate)
```

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

```
#which dates litter was sampled in August 2018, determined to be "2018-08-02" "2018-08-30"
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`?

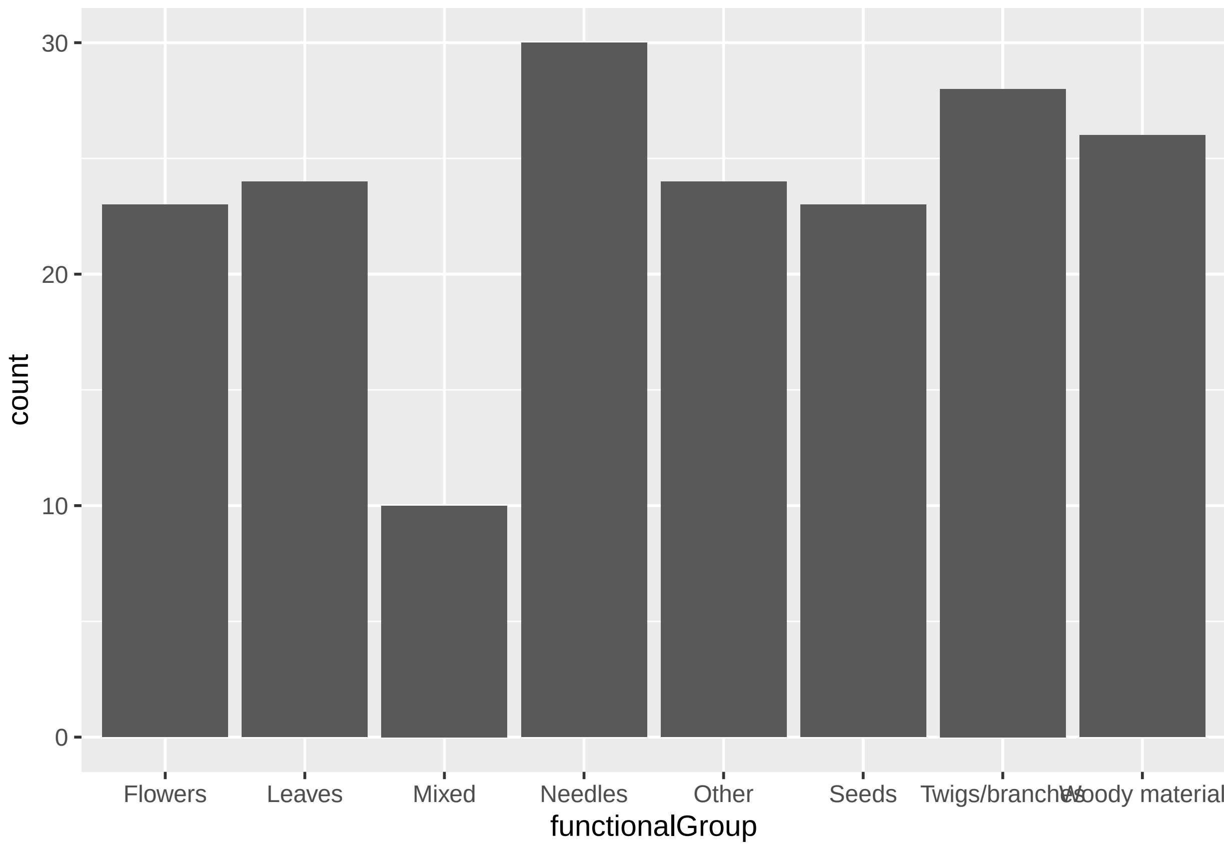
```r
#how many plots sampled at Niwot Ridge
unique(Litter$plotID)
```

```
##  [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
```

Answer: There are 12 different plots that were sampled at Niwot Ridge. Information obtained with the `unique` function returns each of the unique values in that column/vector while `summary` provides the all the unique values but also how many times each of them appear in the column/vector.

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
#bar graph of functionalGroup counts
ggplot(Litter, aes(functionalGroup)) +
  geom_bar()
```

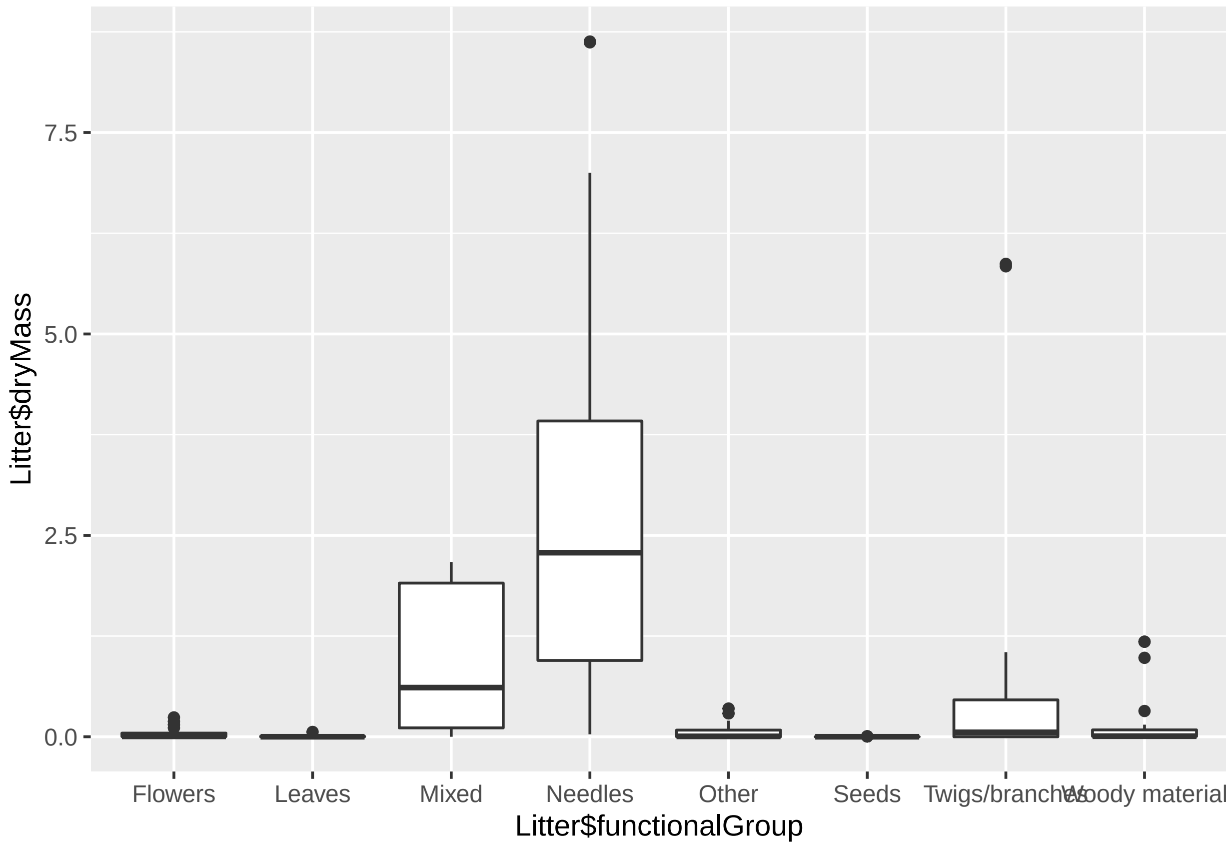


15. Using `geom_boxplot` and `geom_violin`, create a boxplot and a violin plot of dryMass by functionalGroup.

```r
#box plot of dryMass
ggplot(Litter, aes(x=Litter$functionalGroup, y=Litter$dryMass)) +
  geom_boxplot()
```

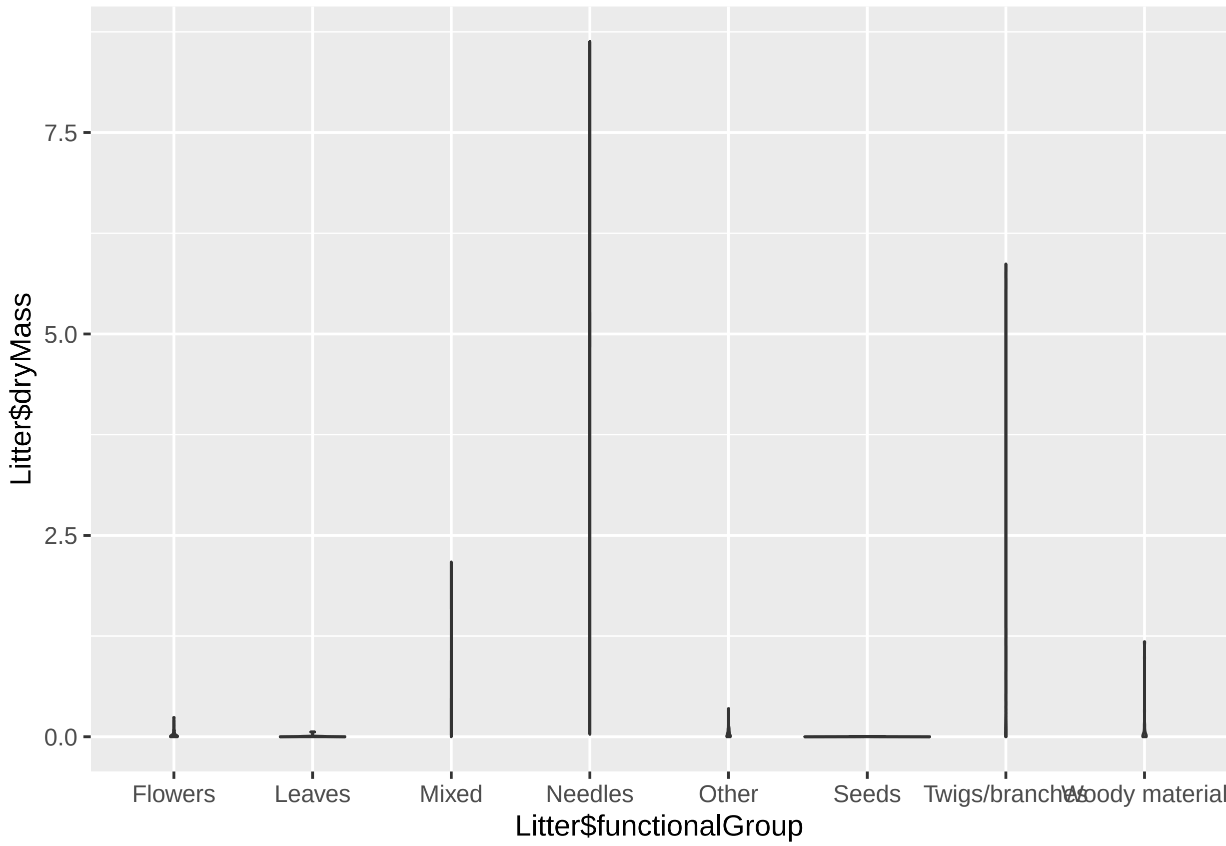
```
## Warning: Use of `Litter$functionalGroup` is discouraged. Use `functionalGroup`
## instead.
```

```
## Warning: Use of `Litter$dryMass` is discouraged. Use `dryMass` instead.
```




```
#violin plot of dryMass
ggplot(Litter, aes(x=Litter$functionalGroup, y=Litter$dryMass)) +
  geom_violin()
```

```
## Warning: Use of `Litter$functionalGroup` is discouraged. Use `functionalGroup`
## instead.

## Warning: Use of `Litter$dryMass` is discouraged. Use `dryMass` instead.
```



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

> Answer: The box plot more clearly shows the mean dryMass for each litter type as well as outliers that exist. On the other hand, the violin plots poorly represent spread within the data as they end up appearing as lines due to some dramatic outliers and the majority of the data being very low dryMass.

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

> Answer: The litter type with the highest biomass at these sites is needles with the next highest litter types being mixed and twigs/branches. All other litter types have fairly negligible dry masses.