Assignment 3: Data Exploration

Tasha Griffiths, Section #2

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.**

```
#check working directory location
getwd()
```

```
## [1] "C:/Users/Tasha Griffiths/Documents/Duke Year 1/Spring 22 Classes/Environmental Data Analytics/G
```

```
#install package if not on system
#install.packages("tidyverse")

#load package for session
library(tidyverse)
library(formatR)

#load two 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: By looking at the ecotoxicology of neonicotinoids on insects we can better understand the toxic effects on humans. We can also understand if the insecticides are impacting more species than just insects, or if they are effective in killing insects.

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: By understanding the litter in a forest, you can better understand impacts on soil health, tree growth and decline, and insect communities.

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: * Litter and woody debris are collected from both elevated and ground traps * Mass measurements are made and debris sorted * Locations of tower plots are selected randomly, in forest plots 40m X 40m * 1-4 trap pairs per plot. Spacing requirements are set between plots, roads, streams, and buildings.

Obtain basic summaries of your data (Neonics)

5. What are the dimensions of the dataset?

```
#basic summaries of dataset
class(Neonics)
```

```
## [1] "data.frame"
```

```
summary(Neonics)
```

```
##    CAS.Number
##  Min.   : 58842209
##  1st Qu.:138261413
##  Median :138261413
##  Mean   :147651982
##  3rd Qu.:153719234
##  Max.   :210880925
##
##                                                                                                Chemical.Name
```

```
## (2E)-1-[(6-Chloro-3-pyridinyl)methyl]-N-nitro-2-imidazolidinimine                           :2658
## 3-[(2-Chloro-5-thiazolyl)methyl]tetrahydro-5-methyl-N-nitro-4H-1,3,5-oxadiazin-4-imine: 686
## [C(E)]-N-[(2-Chloro-5-thiazolyl)methyl]-N’-methyl-N’’-nitroguanidine                        : 452
## (1E)-N-[(6-Chloro-3-pyridinyl)methyl]-N’-cyano-N-methylethanimidamide                        : 420
## N’’-Methyl-N-nitro-N’-[(tetrahydro-3-furanyl)methyl]guanidine                                : 218
## [N(Z)]-N-[3-[(6-Chloro-3-pyridinyl)methyl]-2-thiazolidinylidene]cyanamide                    : 128
## (Other)                                                                                      :  61
##                                                  Chemical.Grade
## Not reported                                            :3989
## Technical grade, technical product, technical formulation: 422
## Pestanal grade                                          :  93
## Not coded                                               :  53
## Commercial grade                                        :  27
## Analytical grade                                        :  15
## (Other)                                                 :  24
##                                                Chemical.Analysis.Method
## Measured                                                   : 230
## Not coded                                                  :  51
## Not reported                                               :   5
## Unmeasured                                                 :4321
## Unmeasured values (some measured values reported in article):  16
##
##
## Chemical.Purity                 Species.Scientific.Name
## NR     :2502    Apis mellifera               : 667
## 25     : 244    Bombus terrestris            : 183
## 50     : 200    Apis mellifera ssp. carnica  : 152
## 20     : 189    Bombus impatiens             : 140
## 70     : 112    Apis mellifera ssp. ligustica: 113
## 75     :  89    Popillia japonica            :  94
## (Other):1287    (Other)                      :3274
##            Species.Common.Name
## Honey Bee            : 667
## Parasitic Wasp       : 285
## Buff Tailed Bumblebee: 183
## Carniolan Honey Bee  : 152
## Bumble Bee           : 140
## Italian Honeybee     : 113
## (Other)              :3083
##                                                       Species.Group
## Insects/Spiders                                            :3569
## Insects/Spiders; Standard Test Species                     :  27
## Insects/Spiders; Standard Test Species; U.S. Invasive Species: 667
## Insects/Spiders; U.S. Invasive Species                     : 360
##
##
##
##   Organism.Lifestage  Organism.Age            Organism.Age.Units
## Not reported:2271     NR     :3851   Not reported        :3515
## Adult       :1222     2      : 111   Day(s)              : 327
## Larva       : 437     3      : 105   Instar              : 255
## Multiple    : 285     <24    :  81   Hour(s)             : 241
## Egg         : 128     4      :  81   Hours post-emergence:  99
## Pupa        :  69     1      :  59   Year(s)             :  64
```

```
##  (Other)     : 211     (Other): 335   (Other)             : 122
##                    Exposure.Type         Media.Type
##  Environmental, unspecified:1599   No substrate:2934
##  Food                      :1124   Not reported: 663
##  Spray                     : 393   Natural soil: 393
##  Topical, general          : 254   Litter      : 264
##  Ground granular           : 249   Filter paper: 230
##  Hand spray                : 210   Not coded   :  51
##  (Other)                   : 794   (Other)     :  88
##              Test.Location  Number.of.Doses        Conc.1.Type..Author.
##  Field artificial    :  96   2      :2441    Active ingredient:3161
##  Field natural       :1663   3      : 499    Formulation      :1420
##  Field undeterminable:   4   5      : 314    Not coded        :  42
##  Lab                 :2860   6      : 230
##                              4      : 221
##                              NR     : 217
##                              (Other): 701
##  Conc.1..Author. Conc.1.Units..Author.             Effect
##  0.37/  : 208    AI kg/ha  : 575       Population      :1803
##  10/    : 127    AI mg/L   : 298       Mortality       :1493
##  NR/    : 108    AI lb/acre: 277       Behavior        : 360
##  NR     :  94    AI g/ha   : 241       Feeding behavior: 255
##  1      :  82    ng/org    : 231       Reproduction    : 197
##  1023   :  80    ppm       : 180       Development     : 136
##  (Other):3924    (Other)   :2821       (Other)         : 379
##              Effect.Measurement    Endpoint                  Response.Site
##  Abundance             :1699     NOEL   :1816   Not reported          :4349
##  Mortality             :1294     LOEL   :1664   Midgut or midgut gland:  63
##  Survival              : 133     LC50   : 327   Not coded             :  51
##  Progeny counts/numbers: 120     LD50   : 274   Whole organism        :  41
##  Food consumption      : 103     NR     : 167   Hypopharyngeal gland  :  27
##  Emergence             :  98     NR-LETH:  86   Head                  :  23
##  (Other)               :1176     (Other): 289   (Other)               :  69
##  Observed.Duration..Days.       Observed.Duration.Units..Days.
##  1      : 713             Day(s)              :4394
##  2      : 383             Emergence           :  70
##  NR     : 355             Growing season      :  48
##  7      : 207             Day(s) post-hatch   :  20
##  3      : 183             Day(s) post-emergence:  17
##  0.0417 : 133             Tiller stage        :  15
##  (Other):2649             (Other)             :  59
##                                                                           Author
##  Peck,D.C.                                                                  : 208
##  Frank,S.D.                                                                 : 100
##  El Hassani,A.K., M. Dacher, V. Gary, M. Lambin, M. Gauthier, and C. Armengaud:  96
##  Williamson,S.M., S.J. Willis, and G.A. Wright                              :  93
##  Laurino,D., A. Manino, A. Patetta, and M. Porporato                        :  88
##  Scholer,J., and V. Krischik                                                :  82
##  (Other)                                                                    :3956
##  Reference.Number
##  Min.   :   344
##  1st Qu.:108459
##  Median :165559
##  Mean   :142189
```

```
##  3rd Qu.:168998
##  Max.   :180410
##
##
##  Long-Term Effects of Imidacloprid on the Abundance of Surface- and Soil-Active Nontarget Fauna in Tu
##  Reduced Risk Insecticides to Control Scale Insects and Protect Natural Enemies in the Production and
##  Effects of Sublethal Doses of Acetamiprid and Thiamethoxam on the Behavior of the Honeybee (Apis mel
##  Exposure to Neonicotinoids Influences the Motor Function of Adult Worker Honeybees
##  Toxicity of Neonicotinoid Insecticides on Different Honey Bee Genotypes
##  Chronic Exposure of Imidacloprid and Clothianidin Reduce Queen Survival, Foraging, and Nectar Storin
##  (Other)
##                                              Source     Publication.Year
##  Agric. For. Entomol.11(4): 405-419           : 200   Min.   :1982
##  Environ. Entomol.41(2): 377-386              : 100   1st Qu.:2005
##  Arch. Environ. Contam. Toxicol.54(4): 653-661:  96   Median :2010
##  Ecotoxicology23:1409-1418                    :  93   Mean   :2008
##  Bull. Insectol.66(1): 119-126                :  88   3rd Qu.:2013
##  PLoS One9(3): 14 p.                          :  82   Max.   :2019
##  (Other)                                      :3964
##
##  Purity: Ê NR - NR | Organism Age: Ê NR - NR Not reported | Conc 1 (Author): Ê Active ingredient NR/
##  Purity: Ê NR - NR | Organism Age: Ê NR - NR Not reported | Conc 1 (Author): Ê Active ingredient NR -
##  Purity: Ê NR - NR | Organism Age: Ê NR - NR Not reported | Conc 1 (Author): Ê Active ingredient NR -
##  Purity: Ê NR - NR | Organism Age: Ê NR - NR Not reported | Conc 1 (Author): Ê Active ingredient NR/
##  Purity: Ê NR - NR | Organism Age: Ê NR - NR Not reported | Conc 1 (Author): Ê Active ingredient NR -
##  Purity: Ê NR - NR | Organism Age: Ê NR - NR Not reported | Conc 1 (Author): Ê Formulation NR - NR ml
##  (Other)
```

```r
#to find dimensions of dataset
length(Neonics)
```

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

```r
dim(Neonics)
```

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

```r
#to view dataset
colnames(Neonics)
```

```
##  [1] "CAS.Number"                       "Chemical.Name"
##  [3] "Chemical.Grade"                   "Chemical.Analysis.Method"
##  [5] "Chemical.Purity"                  "Species.Scientific.Name"
##  [7] "Species.Common.Name"              "Species.Group"
##  [9] "Organism.Lifestage"               "Organism.Age"
## [11] "Organism.Age.Units"               "Exposure.Type"
## [13] "Media.Type"                       "Test.Location"
## [15] "Number.of.Doses"                  "Conc.1.Type..Author."
## [17] "Conc.1..Author."                  "Conc.1.Units..Author."
## [19] "Effect"                           "Effect.Measurement"
## [21] "Endpoint"                         "Response.Site"
## [23] "Observed.Duration..Days."         "Observed.Duration.Units..Days."
```

```
## [25] "Author"                          "Reference.Number"
## [27] "Title"                           "Source"
## [29] "Publication.Year"                "Summary.of.Additional.Parameters"
```

```
str(Neonics)
```

```
## 'data.frame':    4623 obs. of  30 variables:
##  $ CAS.Number                      : int  58842209 58842209 58842209 58842209 58842209 58842209 58842
##  $ Chemical.Name                   : Factor w/ 9 levels "(1E)-N-[(6-Chloro-3-pyridinyl)methyl]-N'-cy
##  $ Chemical.Grade                  : Factor w/ 9 levels "Analytical grade",..: 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
##  $ 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 "~10","~24","~7",..: 39 39 39 39 39 36 39 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
##  $ Conc.1..Author.                 : Factor w/ 1006 levels "~10","~30/","~40/",..: 639 510 813 622 4
##  $ 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 "~.1458","~10",..: 145 145 145 145 145 145
##  $ 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: Ê NC - NC | Organism Age: Ê NC -
```

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

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

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

Answer: Most common effects for the insecticide data are population- which can be due to researchers wanted to see if numbers of a particular insect are lowered by the chemical. Next is mortality, which can be to see if the chemical succeeded in killing the subject.

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(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: 5 of the 6 of the most common insects studied are in the bee family which is an important group of pollinators. This group may be of interest because farmers may not want their insecticides to harm bee species who are important in pollinating their crops. The other common insect is the parasitic wasp, which is another helpful insect for farmers since they kill many other kinds of insect pests.

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] "factor"
```

```
summary(Neonics$conc.1..Author.)
```

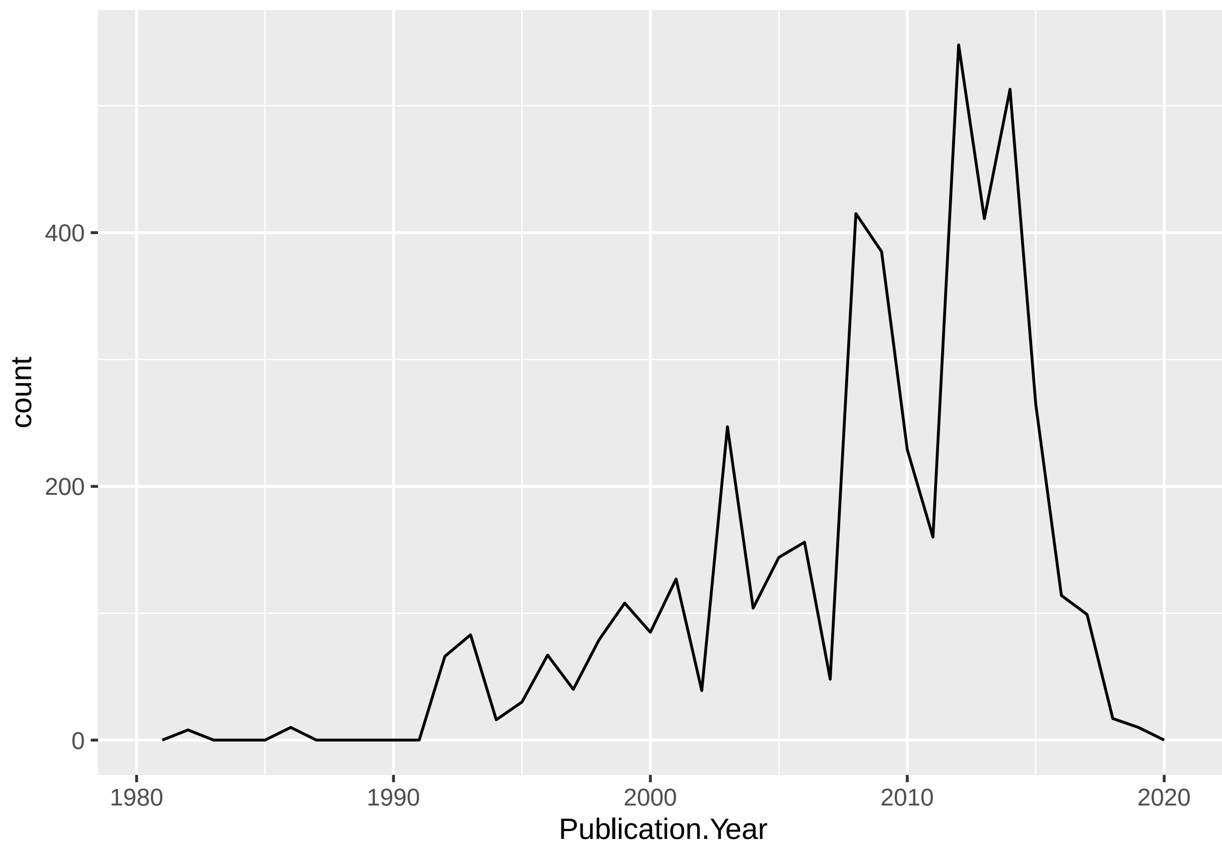
```
## Length  Class   Mode
##      0   NULL   NULL
```

Answer: Its not numeric because the data imported had both numbers, characters, and other text that was then imported with stringsasfactors which then switched the class to 'factor'.

Explore your data graphically (Neonics)

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

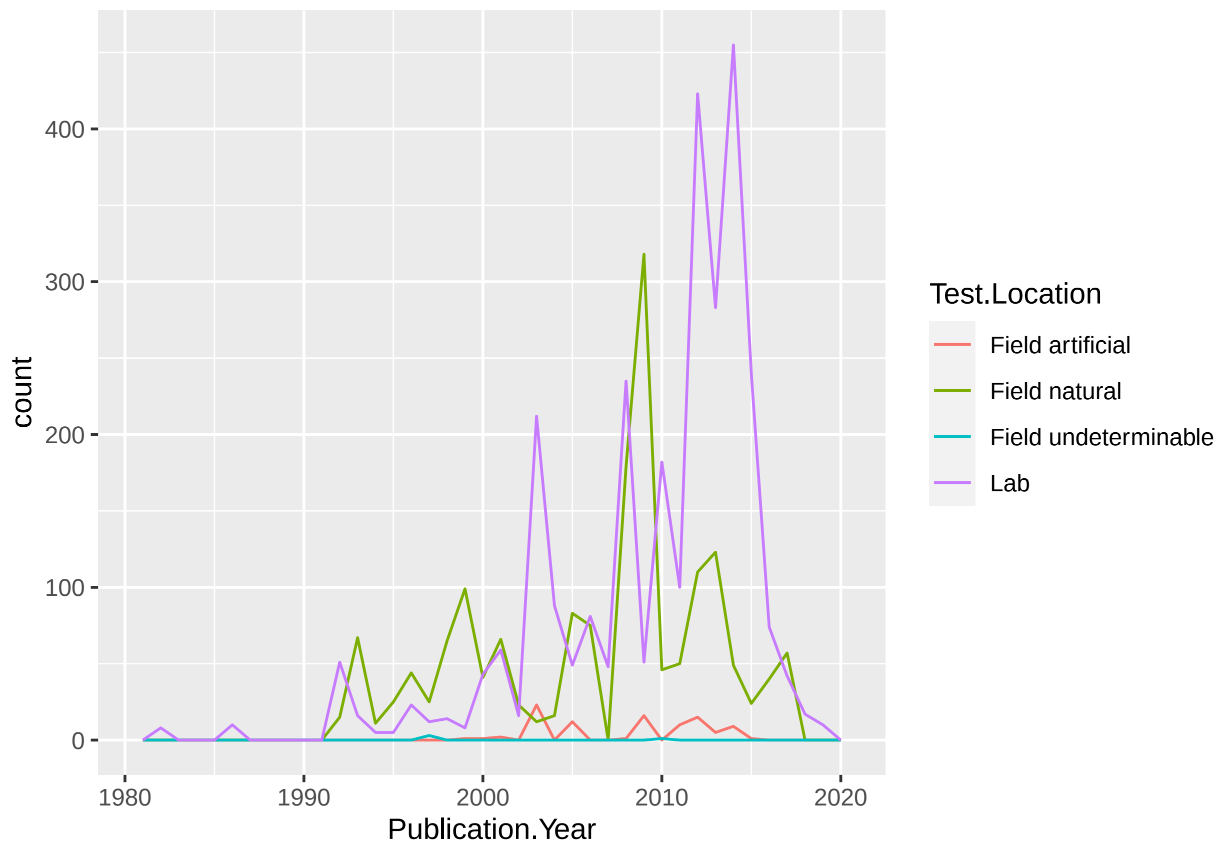
```
ggplot(Neonics) +
  geom_freqpoly(aes(x = Publication.Year), bins = 38)
```




```
#using 38 for bins since there has been 38 years of data collected in the study.
```

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), bins = 38)
```




```
#color added as a separate aestheic in the chain argument.
```

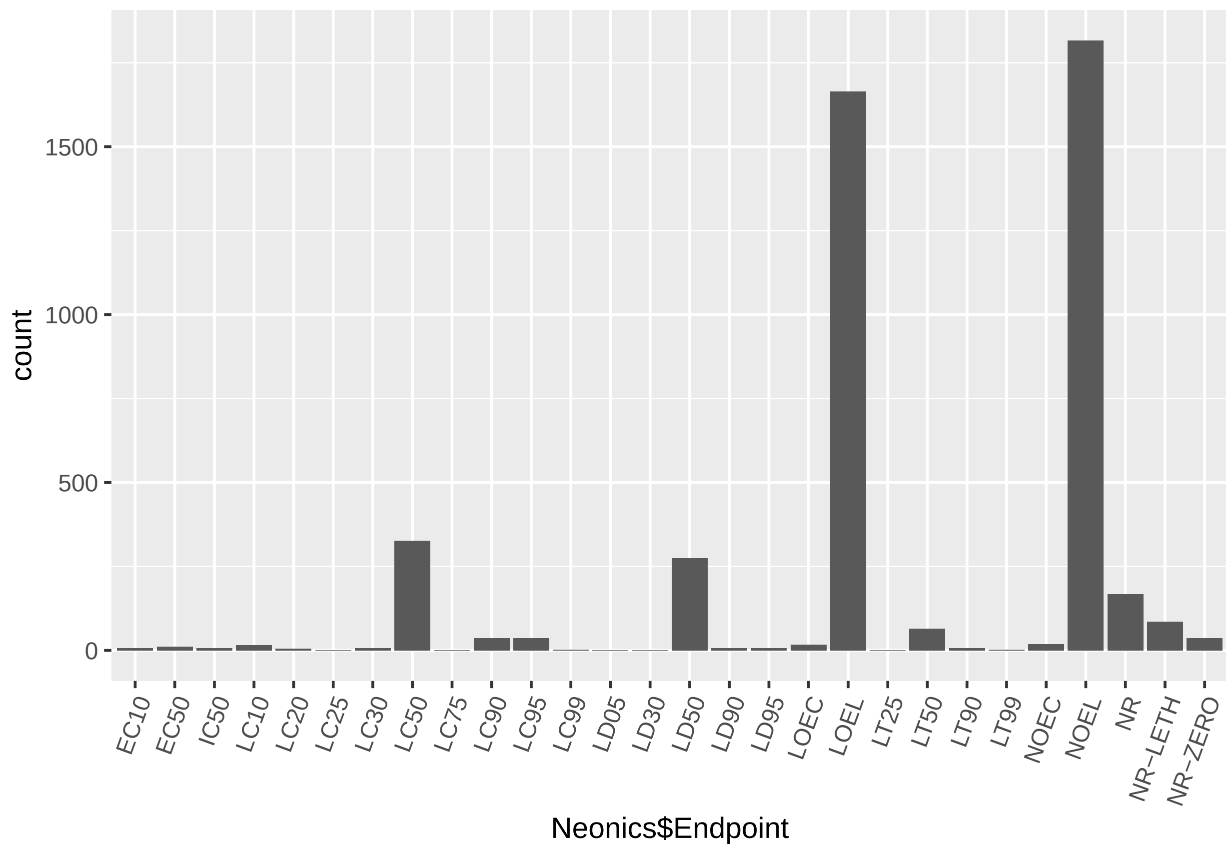
Interpret this graph. What are the most common test locations, and do they differ over time?

> Answer: Most common test location is the lab, which makes sense since it is a more controled location where you can handle a lot of samples at one time. Next frequent is the natural field. You can see that natural field was higher than lab from 1990 through 2000, and then lab had a major peak - I imagine this difference is due to advances in technology over that 10 year time period, and over the next 20 years technology increases would lead to more and more lab based samples.

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, aes(x = Neonics$Endpoint)) +
  geom_bar() +
theme(axis.text.x = element_text(angle = 70, hjust = 1))
```

```
## Warning: Use of ‘Neonics$Endpoint‘ is discouraged. Use ‘Endpoint‘ instead.
```




```r
#largest is NOEC, second is LOEC
```

> Answer: Most common endpoint is NOEC, which means 'No observable effect concentrations' used in aquatic systems and the next most common is LOEC which means 'No-observable-effect-level: highest dose (concentration) producing effects not significantly different from responses of controls according to author's reported statistical test' in terrestrial. Both of these endpoints refer to no-effects of the insecticide.

Explore your data (Litter)

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

```r
class(Litter$collectDate) #its a factor
```

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

```r
tail(Litter$collectDate) #used to see what the current year,month,date format is
```

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

```r
#old code
#as.character(Litter$collectDate)
#convert to date
#Litter$collectDate <- as.Date(Litter$collectDate, format = "%Y/%m/%d")
#class(Litter$collectDate) #conversion worked class is now a date

#install lubridate
library(lubridate)
```

```
##
## Attaching package: 'lubridate'

## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```r
#convert to date with year in first
Litter$collectDate <- ymd(Litter$collectDate)
class(Litter$collectDate)
```

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

```r
#which dates were sampled
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
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
```

```r
summary(Litter$namedLocation)
```

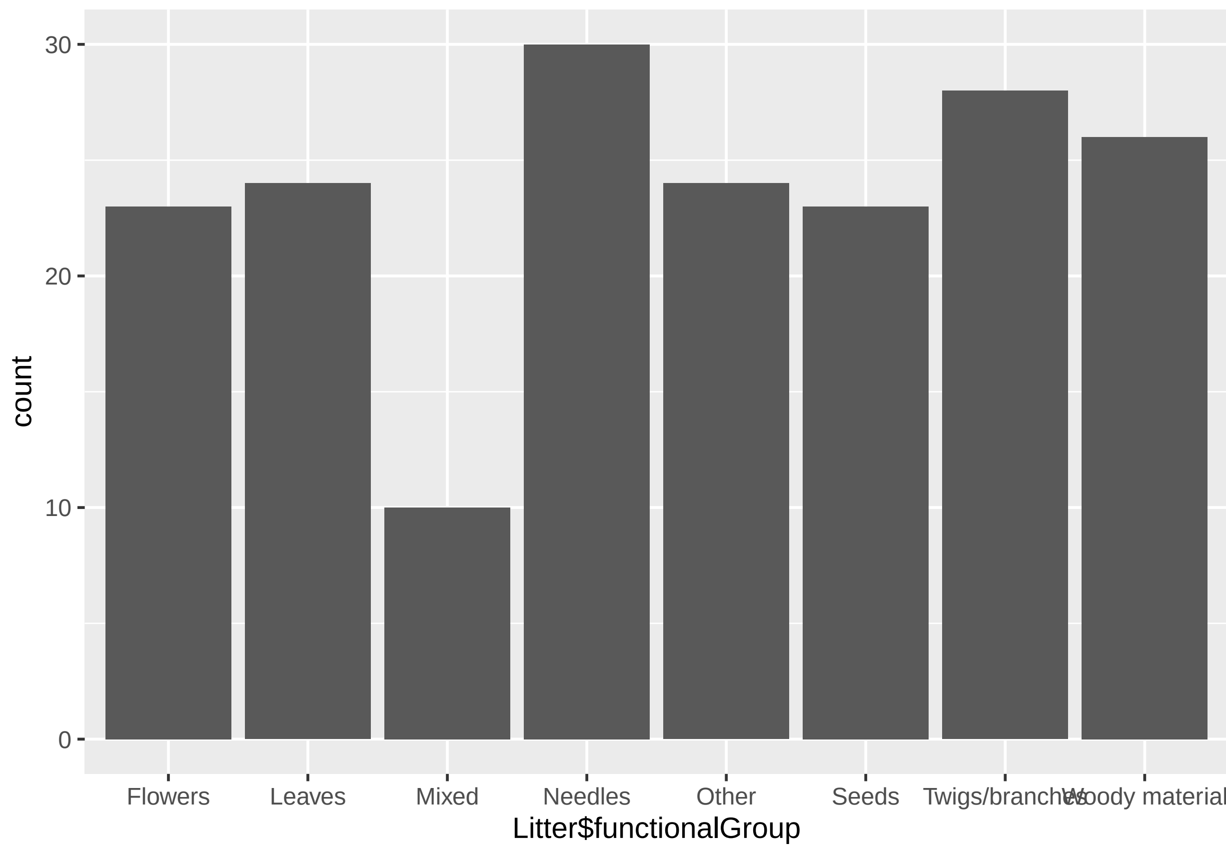
```
## NIWO_040.basePlot.ltr NIWO_041.basePlot.ltr NIWO_046.basePlot.ltr
##                    20                    19                    18
## NIWO_047.basePlot.ltr NIWO_051.basePlot.ltr NIWO_057.basePlot.ltr
##                    15                    14                     8
## NIWO_058.basePlot.ltr 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
```

Answer: There are 12 unique plots on Niwot Ridge. Summary is different from unique because summary 'summarizes' the number of times a unique name was used (provides statistics on those values) within the column while unique just pulls how many unqiue (individual) names are within the column, not how often they are used.

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

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

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



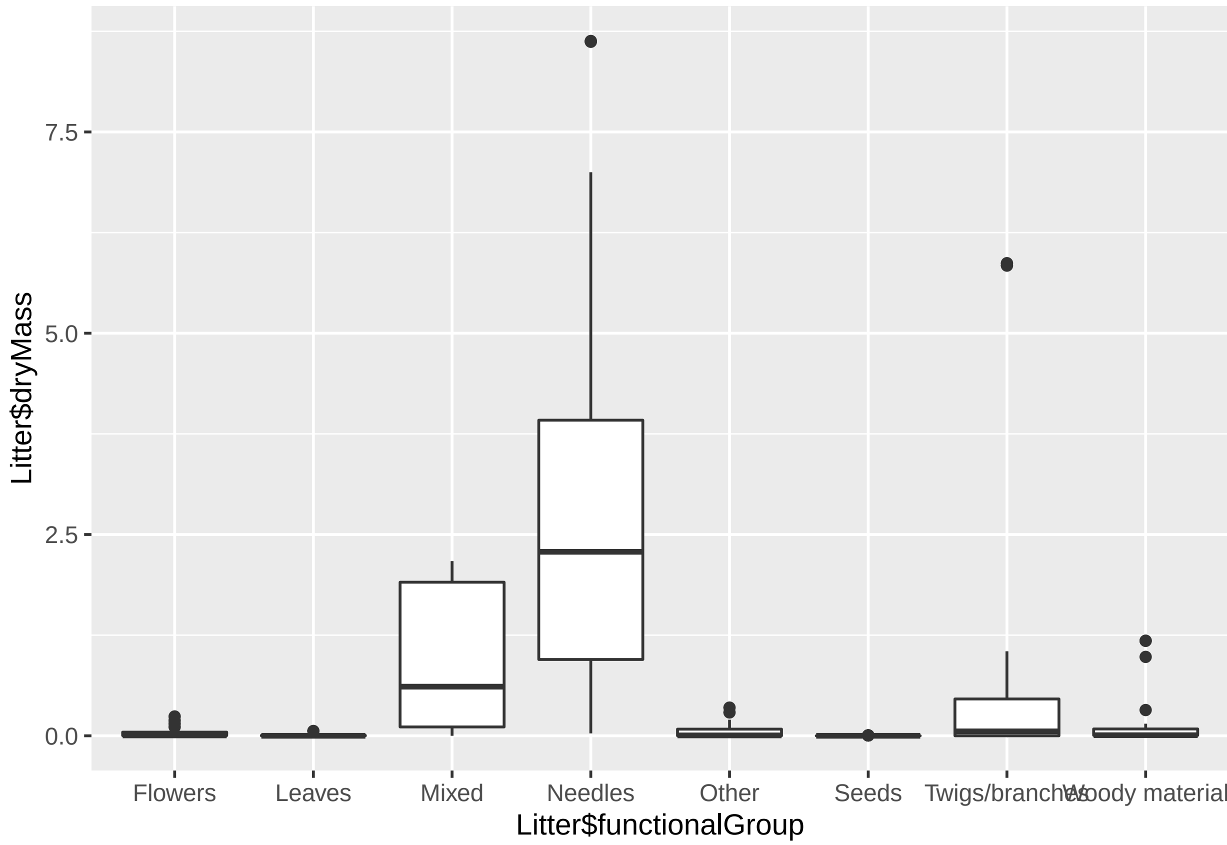
15. Using `geom_boxplot` and `geom_violin`, create a boxplot and a violin plot of dryMass by functionalGroup.

```
ggplot(Litter) +
  geom_boxplot(aes(x = Litter$functionalGroup, y = Litter$dryMass), bins = 50)
```

```
## Warning: Ignoring unknown parameters: bins
```

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

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

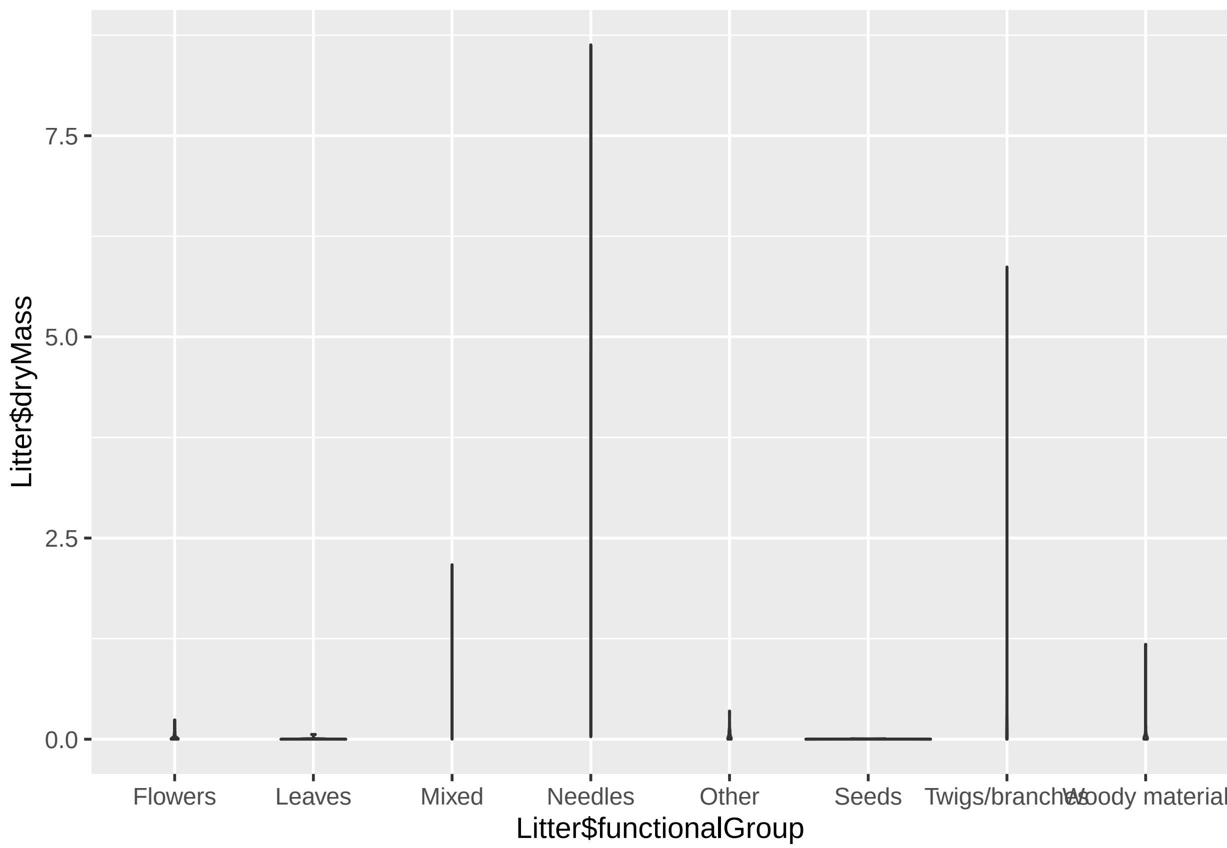



```
ggplot(Litter) +
  geom_violin(aes(x = Litter$functionalGroup, y = Litter$dryMass), bins = 50)
```

```
## Warning: Ignoring unknown parameters: bins

## 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: Violin polots are useful when seeing if there are a lot of repetitions around a certain value, however in this case there is essentailly no repitiion so the violin plots are too thin to be useful. This means that the range of distrubution among quartiles is easier to determine from the boxplot.

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

> Answer: Needles becuase there is more range and a wider box for that functional group.