Assignment 3: Data Exploration

Mandy Hooks

OVERVIEW

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

Directions

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

The completed exercise is due on <>.

Set up your R session

1. Check your working directory, load necessary packages (tidyverse), and upload two datasets: the ECOTOX neonicotinoid dataset (ECOTOX_Neonicotinoids_Insects_raw.csv) and the Niwot Ridge NEON dataset for litter and woody debris (NEON_NIWO_Litter_massdata_2018-08_raw.csv). Name these datasets "Neonics" and "Litter", respectively.

```
getwd()
```

```
## [1] "/Users/mandyhooks/Environmental_Data_Analytics_2021/Assignments"
```

```
library(tinytex)
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.3     v purrr   0.3.4
## v tibble  3.0.6     v dplyr   1.0.4
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.1
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
library(knitr)
library(utf8)

neonics<-read.csv("/Users/mandyhooks/Environmental_Data_Analytics_2021/Data/Raw/ECOTOX_Neonicotinoids_I
head(neonics,10)
```

```
##    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
## 7    58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
## 8    58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
## 9    58842209 Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine
## 10   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
## 7  Technical grade, technical product, technical formulation
## 8  Technical grade, technical product, technical formulation
## 9  Technical grade, technical product, technical formulation
## 10 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
## 7                Unmeasured              95         Musca domestica
## 8                Unmeasured              95         Musca domestica
## 9                Unmeasured              95         Musca domestica
## 10               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
## 7            House Fly Insects/Spiders              Young           NR
## 8            House Fly Insects/Spiders              Adult            9
## 9            House Fly Insects/Spiders              Adult            9
## 10           House Fly Insects/Spiders              Young           NR
##    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
## 7             Hour(s)             Food Filter paper           Lab
## 8              Day(s)             Food Filter paper           Lab
## 9              Day(s)             Food Filter paper           Lab
## 10            Hour(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
## 7               11    Active ingredient             170                   mg/L
## 8               11    Active ingredient              28                   mg/L
## 9               11    Active ingredient              48                   mg/L
## 10              11    Active ingredient              40                   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
## 7  Mortality          Mortality     LC50  Not reported                        1
## 8  Mortality          Mortality     LC50  Not reported                        1
## 9  Mortality          Mortality     LC50  Not reported                        1
## 10 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)
## 7                         Day(s)
## 8                         Day(s)
## 9                         Day(s)
## 10                        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
## 7  Johnston,A.M., J. Lohr, J. Moes, K.R. Solomon, and E.R. Zaborski
## 8  Johnston,A.M., J. Lohr, J. Moes, K.R. Solomon, and E.R. Zaborski
## 9  Johnston,A.M., J. Lohr, J. Moes, K.R. Solomon, and E.R. Zaborski
## 10 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
## 7            103312
## 8            103312
## 9            103312
```

```
## 10          103312
##
## 1                                                                    Contact Toxicity of Insecti
## 2                                                                    Contact Toxicity of Insecti
## 3  Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Susce
## 4  Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Susce
## 5  Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Susce
## 6  Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Susce
## 7  Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Susce
## 8  Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Susce
## 9  Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Susce
## 10 Toxicity of Synergized and Unsynergized Nitromethylene Heterocycle Insecticide (SD 35651) to Susce
##                              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
## 7  J. Econ. Entomol.79(6): 1439-1442             1986
## 8  J. Econ. Entomol.79(6): 1439-1442             1986
## 9  J. Econ. Entomol.79(6): 1439-1442             1986
## 10 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 ingr
## 2  Purity: \xca NR - NR | Organism Age: \xca NR - NR Not reported | Conc 1 (Author): \xca Active ingr
## 3                Purity: \xca NR - NR | Organism Age: \xca 24 - 48 Hour(s) | Conc 1 (Author): \xca A
## 4                Purity: \xca NR - NR | Organism Age: \xca 24 - 48 Hour(s) | Conc 1 (Author): \xca A
## 5                 Purity: \xca NR - NR | Organism Age: \xca 24 - 48 Hour(s) | Conc 1 (Author): \xca
## 6               Purity: \xca NR - NR | Organism Age: \xca NR - NR Day(s) | Conc 1 (Author): \xca Acti
## 7              Purity: \xca NR - NR | Organism Age: \xca 24 - 48 Hour(s) | Conc 1 (Author): \xca Acti
## 8                 Purity: \xca NR - NR | Organism Age: \xca NR - NR Day(s) | Conc 1 (Author): \xca A
## 9                Purity: \xca NR - NR | Organism Age: \xca NR - NR Day(s) | Conc 1 (Author): \xca Act
## 10               Purity: \xca NR - NR | Organism Age: \xca 24 - 48 Hour(s) | Conc 1 (Author): \xca A
```

```
str(neonics)
```

```
## 'data.frame':    4623 obs. of  30 variables:
##  $ CAS.Number                     : int  58842209 58842209 58842209 58842209 58842209 58842209 5884
##  $ Chemical.Name                  : chr  "Tetrahydro-2-(nitromethylene)-2H-1,3-thiazine" "Tetrahydr
##  $ Chemical.Grade                 : chr  "Technical grade, technical product, technical formulation
##  $ Chemical.Analysis.Method       : chr  "Unmeasured" "Unmeasured" "Unmeasured" "Unmeasured" ...
##  $ Chemical.Purity                : chr  "99" "99" "95" "95" ...
##  $ Species.Scientific.Name        : chr  "Araecerus fasciculatus" "Araecerus fasciculatus" "Musca d
##  $ Species.Common.Name            : chr  "Coffee Bean Weevil" "Coffee Bean Weevil" "House Fly" "Hous
##  $ Species.Group                  : chr  "Insects/Spiders" "Insects/Spiders" "Insects/Spiders" "Inse
##  $ Organism.Lifestage             : chr  "Adult" "Adult" "Young" "Young" ...
##  $ Organism.Age                   : chr  "NR" "NR" "NR" "NR" ...
##  $ Organism.Age.Units             : chr  "Not reported" "Not reported" "Hour(s)" "Hour(s)" ...
##  $ Exposure.Type                  : chr  "Topical, general" "Topical, general" "Food" "Food" ...
##  $ Media.Type                     : chr  "No substrate" "No substrate" "Filter paper" "Filter paper
##  $ Test.Location                  : chr  "Lab" "Lab" "Lab" "Lab" ...
##  $ Number.of.Doses                : chr  "NR" "NR" "11" "11" ...
##  $ Conc.1.Type..Author.           : chr  "Active ingredient" "Active ingredient" "Active ingredient
##  $ Conc.1..Author.                : chr  "27.2" "19.7" "47" "25" ...
```

```
##  $ Conc.1.Units..Author.           : chr  "ug/g bdwt" "ug/g bdwt" "mg/L" "mg/L" ...
##  $ Effect                          : chr  "Mortality" "Mortality" "Mortality" "Mortality" ...
##  $ Effect.Measurement              : chr  "Mortality" "Mortality" "Mortality" "Mortality" ...
##  $ Endpoint                        : chr  "LD50" "LD50" "LC50" "LC50" ...
##  $ Response.Site                   : chr  "Not reported" "Not reported" "Not reported" "Not reported"
##  $ Observed.Duration..Days.        : chr  "1" "1" "1" "1" ...
##  $ Observed.Duration.Units..Days.  : chr  "Day(s)" "Day(s)" "Day(s)" "Day(s)" ...
##  $ Author                          : chr  "Childers,C.C., and H.N. Nigg" "Childers,C.C., and H.N. Nig
##  $ Reference.Number                : int  107388 107388 103312 103312 103312 103312 103312 103312 103
##  $ Title                           : chr  "Contact Toxicity of Insecticides to Adults of the Coffee
##  $ Source                          : chr  "J. Econ. Entomol.75(3): 556-559" "J. Econ. Entomol.75(3):
##  $ Publication.Year                : int  1982 1982 1986 1986 1986 1986 1986 1986 1986 1986 ...
##  $ Summary.of.Additional.Parameters: chr  "Purity: \xca NR - NR | Organism Age: \xca NR - NR Not rep
```

```
litter<-read.csv("/Users/mandyhooks/Environmental_Data_Analytics_2021/Data/Raw/NEON_NIWO_Litter_massdata
head(litter,10)
```

```
##                                    uid         namedLocation domainID siteID
## 1  7f065fec-bcb2-4af9-b742-8e520fab7f6e NIWO_061.basePlot.ltr      D13   NIWO
## 2  88df210b-1445-4c3f-b19e-5dabd9305c6e NIWO_061.basePlot.ltr      D13   NIWO
## 3  7f3c549c-1dfa-43bf-a485-c7c2bcb31fd6 NIWO_061.basePlot.ltr      D13   NIWO
## 4  97806ab5-42d2-49c0-8463-db48cd5eab12 NIWO_061.basePlot.ltr      D13   NIWO
## 5  9d7c89f5-85f8-47b6-b415-1ae208580e6f NIWO_061.basePlot.ltr      D13   NIWO
## 6  6ca7a3e8-4d9e-4062-91a0-845f23b5b925 NIWO_061.basePlot.ltr      D13   NIWO
## 7  a0f02718-2a8e-4f02-beaa-edac27ab1b74 NIWO_061.basePlot.ltr      D13   NIWO
## 8  500eb7f8-1881-4a10-bd41-cce84f3b3c47 NIWO_061.basePlot.ltr      D13   NIWO
## 9  aa0ce5fb-6c8f-42cb-a325-f8c6ab214cff NIWO_064.basePlot.ltr      D13   NIWO
## 10 a588a308-b670-4f07-8040-6980d6cfdb72 NIWO_064.basePlot.ltr      D13   NIWO
##      plotID       trapID  weighDate    setDate collectDate     ovenStartDate
## 1  NIWO_061 NIWO_061_169 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 2  NIWO_061 NIWO_061_169 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 3  NIWO_061 NIWO_061_169 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 4  NIWO_061 NIWO_061_169 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 5  NIWO_061 NIWO_061_169 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 6  NIWO_061 NIWO_061_169 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 7  NIWO_061 NIWO_061_169 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 8  NIWO_061 NIWO_061_169 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 9  NIWO_064 NIWO_064_103 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
## 10 NIWO_064 NIWO_064_103 2018-08-06 2018-07-05  2018-08-02 2018-08-02T21:00Z
##          ovenEndDate               fieldSampleID
## 1  2018-08-06T18:02Z NEON.LTR.NIWO061169.20180802
## 2  2018-08-06T18:02Z NEON.LTR.NIWO061169.20180802
## 3  2018-08-06T18:02Z NEON.LTR.NIWO061169.20180802
## 4  2018-08-06T18:02Z NEON.LTR.NIWO061169.20180802
## 5  2018-08-06T18:02Z NEON.LTR.NIWO061169.20180802
## 6  2018-08-06T18:02Z NEON.LTR.NIWO061169.20180802
## 7  2018-08-06T18:02Z NEON.LTR.NIWO061169.20180802
## 8  2018-08-06T18:02Z NEON.LTR.NIWO061169.20180802
## 9  2018-08-06T18:02Z NEON.LTR.NIWO064103.20180802
## 10 2018-08-06T18:02Z NEON.LTR.NIWO064103.20180802
##                        massSampleID samplingProtocolVersion functionalGroup
## 1  NEON.LTR.NIWO061169.20180802.TWI       NEON.DOC.001710vE  Twigs/branches
## 2  NEON.LTR.NIWO061169.20180802.SDS       NEON.DOC.001710vE           Seeds
## 3  NEON.LTR.NIWO061169.20180802.WDY       NEON.DOC.001710vE  Woody material
## 4  NEON.LTR.NIWO061169.20180802.FLR       NEON.DOC.001710vE         Flowers
```

```
## 5  NEON.LTR.NIWO061169.20180802.WDY       NEON.DOC.001710vE  Woody material
## 6  NEON.LTR.NIWO061169.20180802.NDL       NEON.DOC.001710vE         Needles
## 7  NEON.LTR.NIWO061169.20180802.OTH       NEON.DOC.001710vE           Other
## 8  NEON.LTR.NIWO061169.20180802.LVS       NEON.DOC.001710vE          Leaves
## 9  NEON.LTR.NIWO064103.20180802.FLR       NEON.DOC.001710vE         Flowers
## 10 NEON.LTR.NIWO064103.20180802.WDY       NEON.DOC.001710vE  Woody material
##    dryMass qaDryMass remarks                   measuredBy
## 1    0.400         N      NA kstyers@battelleecology.org
## 2    0.005         N      NA kstyers@battelleecology.org
## 3    0.040         Y      NA kstyers@battelleecology.org
## 4    0.005         N      NA kstyers@battelleecology.org
## 5    0.070         N      NA kstyers@battelleecology.org
## 6    1.000         N      NA kstyers@battelleecology.org
## 7    0.200         N      NA kstyers@battelleecology.org
## 8    0.005         N      NA kstyers@battelleecology.org
## 9    0.190         N      NA kstyers@battelleecology.org
## 10   1.180         Y      NA kstyers@battelleecology.org
```

```
str(litter)
```

```
## 'data.frame':    188 obs. of  19 variables:
##  $ uid                    : chr  "7f065fec-bcb2-4af9-b742-8e520fab7f6e" "88df210b-1445-4c3f-b19e-5da
##  $ namedLocation          : chr  "NIWO_061.basePlot.ltr" "NIWO_061.basePlot.ltr" "NIWO_061.basePlot.
##  $ domainID               : chr  "D13" "D13" "D13" "D13" ...
##  $ siteID                 : chr  "NIWO" "NIWO" "NIWO" "NIWO" ...
##  $ plotID                 : chr  "NIWO_061" "NIWO_061" "NIWO_061" "NIWO_061" ...
##  $ trapID                 : chr  "NIWO_061_169" "NIWO_061_169" "NIWO_061_169" "NIWO_061_169" ...
##  $ weighDate              : chr  "2018-08-06" "2018-08-06" "2018-08-06" "2018-08-06" ...
##  $ setDate                : chr  "2018-07-05" "2018-07-05" "2018-07-05" "2018-07-05" ...
##  $ collectDate            : chr  "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" ...
##  $ ovenStartDate          : chr  "2018-08-02T21:00Z" "2018-08-02T21:00Z" "2018-08-02T21:00Z" "2018-08
##  $ ovenEndDate            : chr  "2018-08-06T18:02Z" "2018-08-06T18:02Z" "2018-08-06T18:02Z" "2018-08
##  $ fieldSampleID          : chr  "NEON.LTR.NIWO061169.20180802" "NEON.LTR.NIWO061169.20180802" "NEON
##  $ massSampleID           : chr  "NEON.LTR.NIWO061169.20180802.TWI" "NEON.LTR.NIWO061169.20180802.SDS
##  $ samplingProtocolVersion: chr  "NEON.DOC.001710vE" "NEON.DOC.001710vE" "NEON.DOC.001710vE" "NEON.DO
##  $ functionalGroup        : chr  "Twigs/branches" "Seeds" "Woody material" "Flowers" ...
##  $ dryMass                : num  0.4 0.005 0.04 0.005 0.07 1 0.2 0.005 0.19 1.18 ...
##  $ qaDryMass              : chr  "N" "N" "Y" "N" ...
##  $ remarks                : logi  NA NA NA NA NA NA ...
##  $ measuredBy             : chr  "kstyers@battelleecology.org" "kstyers@battelleecology.org" "kstyers
```

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: According to what I read on the internet, these particularly effect bees. Bees are paramount to having a healthy ecosystem.

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: According to a USDA article I saw, this litter/debris is important in managing wildfire risks. This is especially important as we see wildfires increase in frequency and intensity.

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: *Sampled randomly within the 90% flux footprint of the primary and secondary airsheds* In sites with forested tower airsheds, the litter sampling is targeted to take place in 20 40m x 40m plots. *in sites with low-statured vegetation over the tower airsheds, liter sampling is targeted to take place in 4 40m x 40m tower plots (to accommodate co-located soil sampling) plus 26 20m x 20m plots

Obtain basic summaries of your data (Neonics)

5. What are the dimensions of the dataset?

```
dim(neonics)
```

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

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

```
sort(table(neonics$Effect), decreasing = T)
```

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

Answer: The most common effects of population and mortality are important in measuring the health and growth of a species. We want to know if a species is dying or growing, especially if they are bees

7. Using the `summary` function, determine the six most commonly studied species in the dataset (common name). What do these species have in common, and why might they be of interest over other insects? Feel free to do a brief internet search for more information if needed.

```
sort(table(neonics$Species.Common.Name), decreasing = T) [1:6]
```

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

```
summary(neonics$Species.Common.Name)
```

```
##    Length     Class      Mode
##      4623 character character
```

Answer: All of these species are bees, and as I mentioned earlier this is crucial for having a healthy ecosystem, especially for plants.

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

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

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

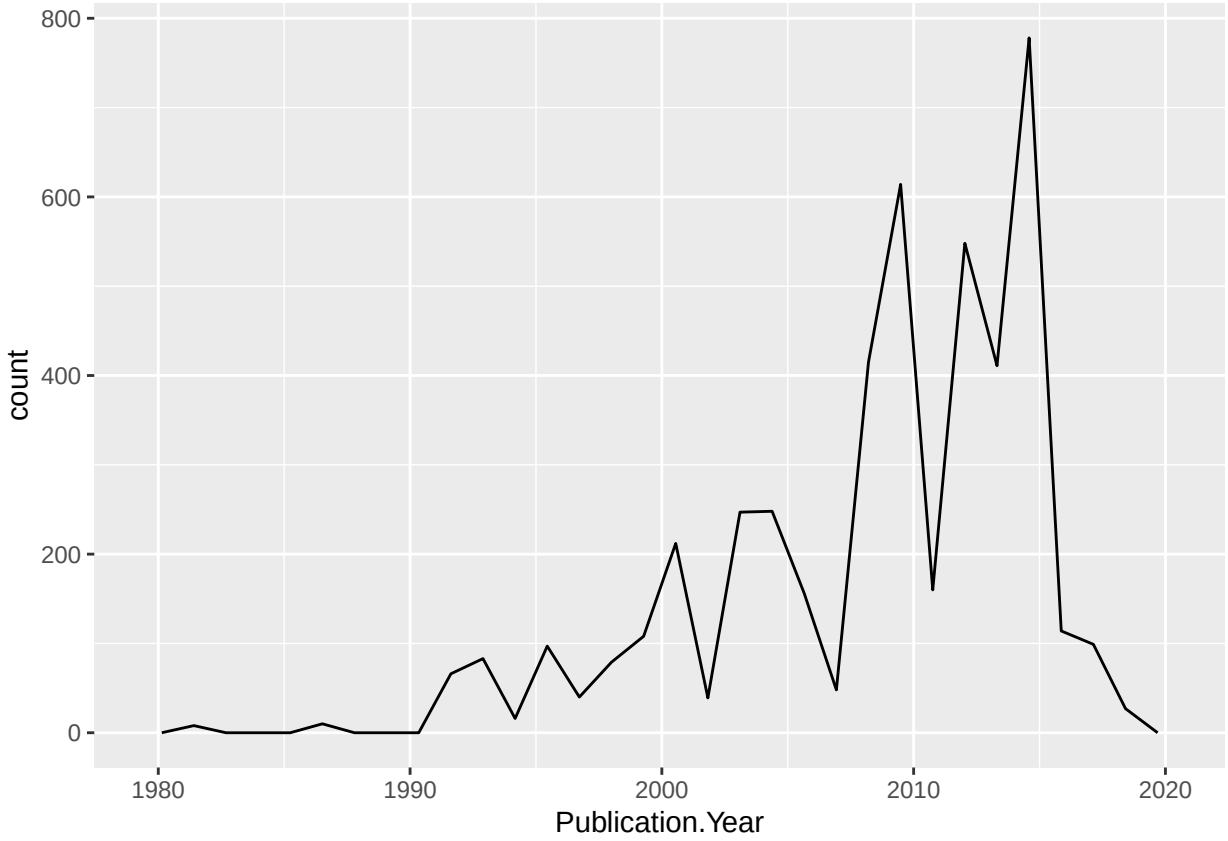
Answer: The CSV transfers the data over as a character, but we can change it to numeric using the as.numeric function

Explore your data graphically (Neonics)

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

```
ggplot(neonics) +
  geom_freqpoly(aes(x=Publication.Year))
```

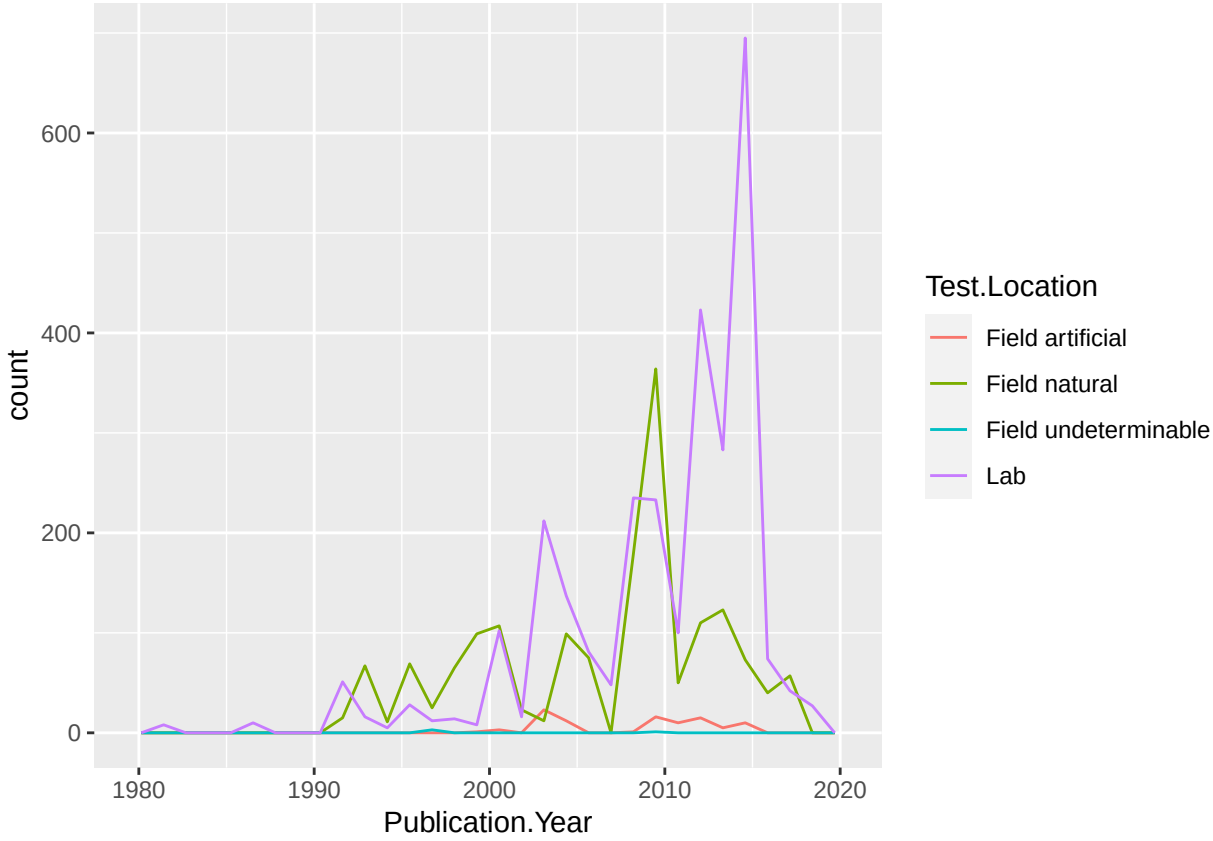
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



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

```
ggplot(neonics) +
  geom_freqpoly(aes(x=Publication.Year, color=Test.Location))
```

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

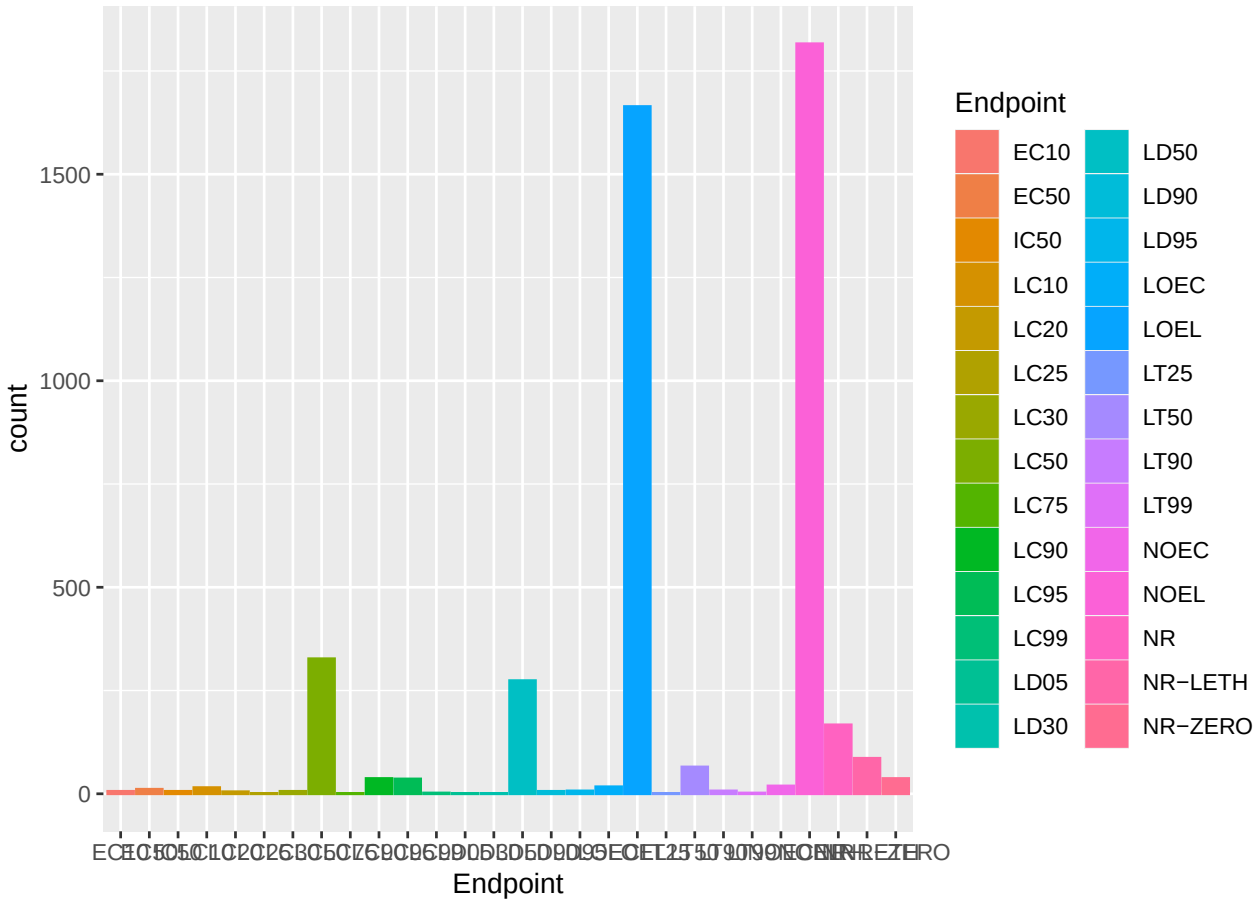


Interpret this graph. What are the most common test locations, and do they differ over time?

Answer: Lab has almost always been the predominant test location, field undeterminable has always been present but never a common one, and then overtime field natural rose. Lab is by far the most common over time.

11. Create a bar graph of Endpoint counts. What are the two most common end points, and how are they defined? Consult the ECOTOX_CodeAppendix for more information.

```
ggplot(neonics)+
  geom_bar(aes(x=Endpoint, color=Endpoint, fill=Endpoint))
```



Answer: The most common by far are LOEL (Lowest-observable-effect-level) and NOEL (No-observable-effect-level). LOEL is defined as lowest dose (concentration) producing effects that were significantly different (as reported by authors) from responses of controls. Noel is defined as highest dose (concentration) producing effects not significantly different from responses of controls according to author's reported statistical test

Explore your data (Litter)

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

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

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

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

```
##   [1] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##   [6] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [11] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [16] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [21] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [26] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [31] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [36] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [41] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [46] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
```

```
##  [51] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [56] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [61] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [66] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [71] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [76] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [81] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [86] "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02" "2018-08-02"
##  [91] "2018-08-02" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
##  [96] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [101] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [106] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [111] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [116] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [121] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [126] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [131] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [136] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [141] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [146] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [151] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [156] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [161] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [166] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [171] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [176] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [181] "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30" "2018-08-30"
## [186] "2018-08-30" "2018-08-30" "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`?

```
unique(litter$plotID)
```

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

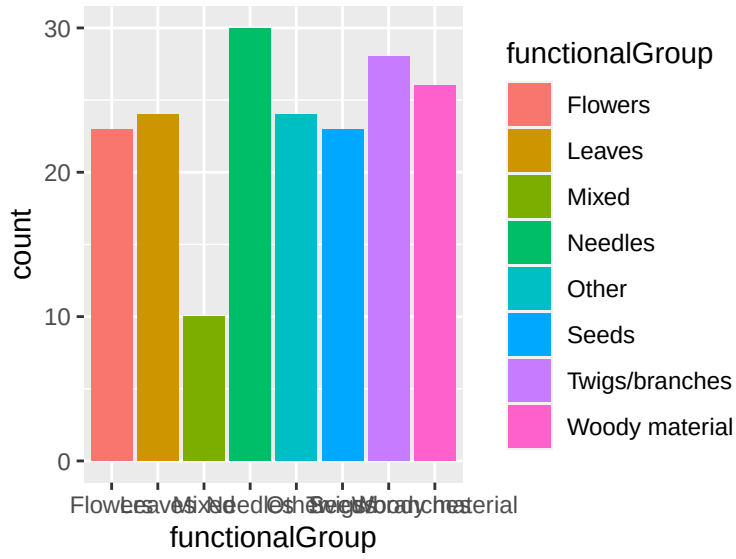
```
summary(litter$plotID)
```

```
##    Length     Class      Mode
##       188 character character
```

Answer: There are 12 plots saampled at Niwot Ridge. Unique allows me to see which ones and how many. Summary counts how many plots were ID'd and tells me their class and mode.

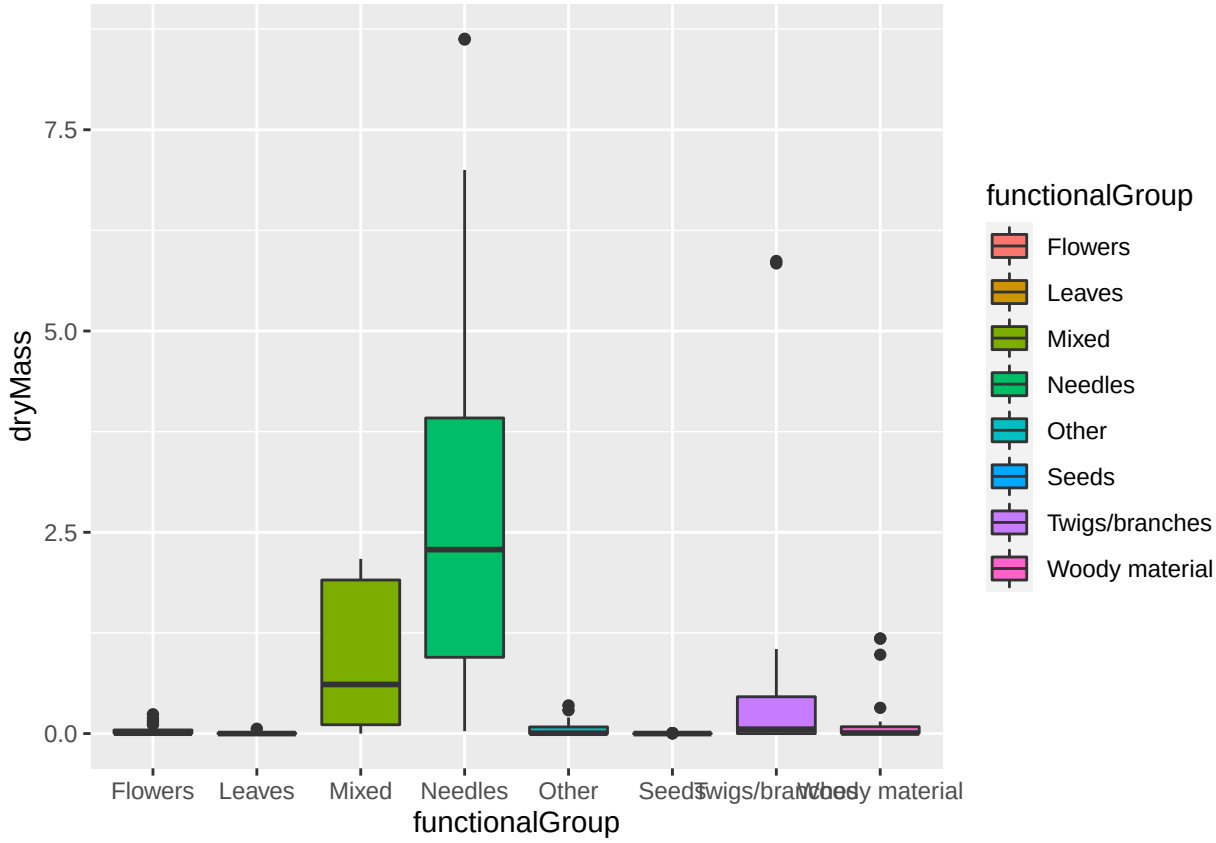
14. Create a bar graph of functionalGroup counts. This shows you what type of litter is collected at the Niwot Ridge sites. Notice that litter types are fairly equally distributed across the Niwot Ridge sites.

```
ggplot(litter)+
  geom_bar(aes(x=functionalGroup, fill=functionalGroup))
```

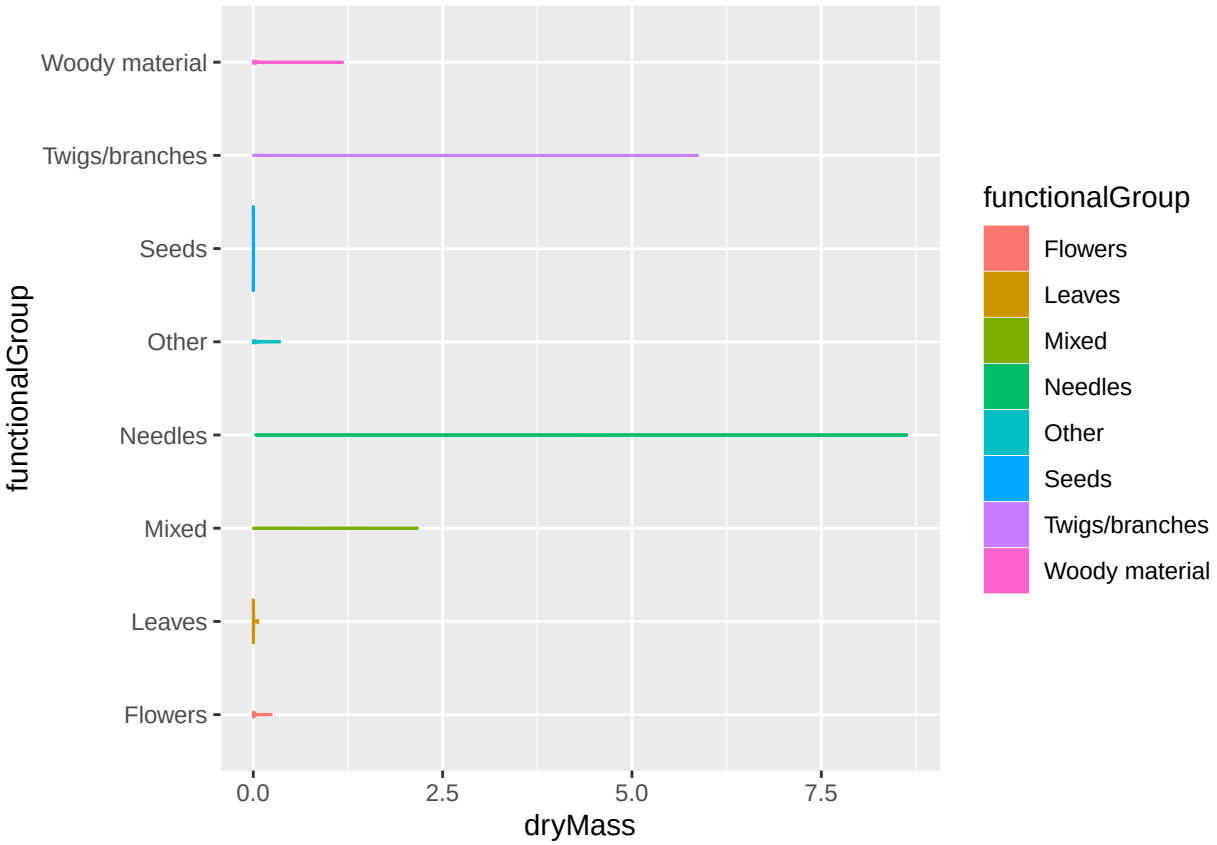


15. Using `geom_boxplot` and `geom_violin`, create a boxplot and a violin plot of dryMass by functionalGroup.

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




```
ggplot(litter)+
  geom_violin(aes(x=dryMass, y=functionalGroup,fill=functionalGroup, color=functionalGroup))
```



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

Answer: The boxplot allows the user to see the median, the skews, and outliers.The violin plot is hard to see and therefore hard to understand.

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

Answer: Needles are the highest in biomass, followed by mixed.