Assignment 3: Data Exploration

Nargis Taraki

Fall 2024

OVERVIEW

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

Directions

1. Rename this file `<FirstLast>_A03_DataExploration.Rmd` (replacing `<FirstLast>` with your first and last name).
2. Change "Student Name" on line 3 (above) with your name.
3. Work through the steps, **creating code and output** that fulfill each instruction.
4. Assign a useful **name to each code chunk** and include ample **comments** with your code.
5. Be sure to **answer the questions** in this assignment document.
6. When you have completed the assignment, **Knit** the text and code into a single PDF file.
7. After Knitting, submit the completed exercise (PDF file) to the dropbox in Canvas.

TIP: If your code extends past the page when knit, tidy your code by manually inserting line breaks.

TIP: If your code fails to knit, check that no `install.packages()` or `View()` commands exist in your code.

Set up your R session

1. Load necessary packages (tidyverse, lubridate, here), check your current working directory and upload two datasets: the ECOTOX neonicotinoid dataset (ECOTOX_Neonicotinoids_Insects_raw.csv) and the Niwot Ridge NEON dataset for litter and woody debris (NEON_NIWO_Litter_massdata_2018-08_raw.csv). Name these datasets "Neonics" and "Litter", respectively. Be sure toset include the subcommand to read strings in as factors.

```r
#Here I have loaded requierd packages.
library(tidyverse)
library(lubridate)
library(here)
# first I checkd my working Directory
getwd()
```

```
## [1] "/home/guest/EDE_Fall2024"
```

```r
list.files("Data")
```

```
## [1] "Metadata"  "Processed" "Raw"
```

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

   Answer:

Studying the ecotoxicology of neonicotinoids on insects is important because these widely used insecticides can significantly impact ecosystems. They are linked to declines in essential pollinators like bees and can harm beneficial insects as well. Insects play critical roles in pollination, decomposition, and food webs, so understanding how neonicotinoids affect them is vital for maintaining ecological balance and sustainable agriculture. Additionally, these chemicals can persist in the environment, leading to long-term effects on insect behavior and reproduction. Research in this area helps inform regulations and assess risks, ensuring that agricultural practices do not compromise insect biodiversity and ecosystem health.

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:

Studying forest litter and woody debris is important because they play a vital role in nutrient cycling, carbon storage, and soil health within forest ecosystems. As they decompose, they release essential nutrients back into the soil, support biodiversity by providing habitat for various organisms, and improve soil structure and water retention. Additionally, these materials influence fire behavior and water regulation in forests, making them crucial for understanding ecosystem resilience. Insights gained from this research also inform sustainable forest management practices, helping to maintain healthy and productive forests.

4. How is litter and woody debris sampled as part of the NEON network? Read the NEON_Litterfall_UserGuide.pdf document to learn more. List three pieces of salient information about the sampling methods here:

   Answer:

1.NEON uses litterbags placed randomly at study sites to collect falling leaves and organic matter. 2.Fine woody debris is sampled in 0.5m x 3m cells within randomly selected plots to ensure accurate results. 3. Coarse downed wood is measured using the line-intercept method, where random lines are drawn to measure larger pieces of wood.

Obtain basic summaries of your data (Neonics)

5. What are the dimensions of the dataset?

```r
dim(Neonics)
```

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

```r
##4623 are the number of rows inthe dataset,whereas 30 is the number of columns.
```

6. Using the **summary** function on the "Effect" column, determine the most common effects that are studied. Why might these effects specifically be of interest? [Tip: The `sort()` command is useful for listing the values in order of magnitude...]

```r
summary_effects <- table(Neonics$Effect)
sort(summary_effects, decreasing = TRUE)
```

```
## 
##       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 frequently studied effect is Population, which focuses on how neonicotinoids affect the overall population sizes of insects. This could be due to neonicotinoids causing declines in insect populations, especially in pollinators like bees. Mortality is also a key focus because it's a direct measure of lethal effects. Other effects like Behavior, Feeding behavior, Reproduction, and Development provide insight into how sublethal doses of neonicotinoids might alter insect activities, potentially leading to long-term ecological consequences.

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

```r
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"
```

```r
summary_species <- summary(Neonics$Species.Common.Name, maxsum = 6)
summary_species
```

```
##             Honey Bee        Parasitic Wasp Buff Tailed Bumblebee
##                   667                   285                   183
##    Carniolan Honey Bee            Bumble Bee               (Other)
##                   152                   140                  3196
```

Answer:

The six most commonly studied species in the dataset are Honey Bees (667 times), Parasitic Wasps (285 times), Buff Tailed Bumblebees (183 times), Carniolan Honey Bees (152 times), and Bumble Bees (140 times). In addition, there are 3196 observations of other less commonly studied species.

Researchers focus on these species, especially honey bees and bumblebees, because they play important roles in pollinating crops and plants. Neonicotinoids can harm these insects, which could lead to problems for farming and ecosystems. Parasitic wasps are also studied because they help control pests. These species are important to study because of their roles in the environment and how sensitive they are to pesticides.

8. Concentrations are always a numeric value. What is the class of `Conc.1..Author.` column in the dataset, and why is it not numeric? [Tip: Viewing the dataframe may be helpful…]

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

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

```r
#Here I am changing factor to character
unique(Neonics$Conc.1..Author.)
```

```
##    [1] 27.2      19.7      47        25        13        268       170
##    [8] 28        48        40        83        900       15.3      20.4
##   [15] 5         NR        ~10       65.56     635.4     239.1     NR/
##   [22] 80.66     1075.7    75.27     636.5     183.4     1267.9    144.0/
##   [29] 0.295     0.267     20.636    76.035    148.3     12/       12.8
##   [36] 43.59     37.83     17.32     14.6      122.4     22.59     36/
##   [43] 7.6       144       144/      0.013     200       0.01      0.002
##   [50] 0.001     96        5.1/      12        18        1.2       120
##   [57] 0.012/    240/      0.012     6.0/      0.491     3.743     6.95
##   [64] 20.6      13.3      22.4      695.5     228.8     34.3      721.6
##   [71] 4.92      60.1      18.5      462.9     32.8      133       0.025/
##   [78] 110.3     1000/     96/       0.007     176.5     1507      17.24
```

```
##  [85] 75.26     1034.8    56.73     461.2     376.3     2240.5    48.96
##  [92] 5.3/      222.65    480       2.3       60.7      20        1.82
##  [99] 40.0/     3/        75/       0.048/    0.25/     75        0.39/
## [106] 0.026/    0.05/     0.005     2.8       50/       10.0/     20/
## [113] 0.43      0.73      8.1       14.53     7.07      64.6      1.69
## [120] 537       0.6       0.1/      0.5/      1         6         1.0/
## [127] 1/        1.38      6.25      0.0600/   100/      80000     0.125/
## [134] 8000      8/        84        112       1.81/     22.2      >100.0
## [141] 100       10.37     1.81      4.24      9.94      2.5/      56
## [148] 263.44    0.18/     0.56      2.27/     1.051     0.49/     23.93
## [155] 0.00084/  12.02     70/       108.27    0.3/      5/        84/
## [162] 0.134     <5.00     <4.00     5.8       16.69     0.154     0.080/
## [169] 71.3      0.09      <10/      0.084/    0.72/     0.02/     80/
## [176] 0.155     0.154/    115/      134.80/   35.183    5.2       4
## [183] 87.1      0.013/    22        4.1       23.8      4.34      2.31
## [190] 0.091     93.21     0.017     0.53      675.4     0.05      0.32
## [197] 24.46     25.39     203.6     19.24     0.609     167.9     43.02
## [204] 1.37      383       85.8      85.8/     0.0429/   0.010725  42.9
## [211] 42.9/     52.6/     0.0037/   225       12.50/    218.89    1.7
## [218] 13.8      35        245.5     87.5      242       1.5/      0.375
## [225] 0.1       0.15      7.32      0.558     >6300     6830      8352
## [232] 6.135     0.00077   1.5       0.15/     0.0400/   0.06825/  138.21
## [239] 53        32/       28/       0.83      10        2         25.0/
## [246] 0.03      0.06/     17.5      125/      0.021     46.7      6.55
## [253] 0.053     1000      2.217     0.297     5.7       1.847     0.047/
## [260] 0.5       0.75/     0.75      0.56/     0.00053   156       0.16
## [267] 0.88      0.125     90        360/      2.4       1.3       8
## [274] 0.4/      326.69    620.21    1104      217       0.29/     4.375/
## [281] 0.031     0.047     14        0.4       0.035     0.053/    0.291
## [288] 0.233     0.113     0.246     500       250       125       0.0022/
## [295] 0.227     138.84    0.0439    >0.5      0.0039    >0.1      0.078
## [302] >81       74.9      ~41       37        50        104       57
## [309] 42        30.6      >20       61        0.0179    80.9      0.0671
## [316] 0.536     4.17      1400      1.53      0.112     0.74/     0.08/
## [323] 48/       0.24      400/      0.12      1.25/     1200/     1.16
## [330] 1.62      2.5       2/        1.25      0.7       0.50/     200/
## [337] 0.24/     10/       500/      3.49      0.16/     0.3       24
## [344] 45.9      70        3.7       11.7      2.2/      0.118     7/
## [351] <0.025    <0.5      0.0015    40/       0.0626    300       3
## [358] 0.00355/  30        15.1      6.6       6.5       191.044   99.063
## [365] 74.631    173.088   103.705   46.763    187.208   109.579   97.425
## [372] 99.82     34.37     29.79     170.52    85.47     65.14     83.97
## [379] 28.81     24.96     120.65    59.36     34.96     21.209    0.21
## [386] 2.16      0.21/     <1.5      24.33     6.7       4.8       5.4
## [393] 242.45    193.59    1.28/     6/        0.2/      0.37/     0.007/
## [400] 0.336     60/       5.6       9.6       30/       145.3     212
## [407] 286/      0.02      180       0.692     5.49      0.003     0.13/
## [414] 0.28/     0.00322   127       0.0192/   16        71        16/
## [421] 71/       0.4483    14/       59        0.04      0.2       1.28
## [428] 6.7/      0.064     15/       4/        0.01/     <2.5/     0.51
## [435] 13.9      23.6      16.2      32.4      4.4       >98.43    43.7
## [442] >98       1.44      98.43/    98.43     39.37     0.13      0.96
## [449] 0.48      0.17      0.18      0.29      0.23      150.5     1.00/
## [456] 0.005/    350/      26.63     35.813    0.1332    683.2     27.16
```

```
## [463] 6.57      5.53      5.35      5.71      5.79      4.76      4.45
## [470] 4.27      5.06      4.82      5.16      4.07      6.83      13.66
## [477] 120/      >=13.66   3.42      447.82    60        0.000048  2.86
## [484] 49.42     0.112/    136/      0.009     1.59      12.17     0.40/
## [491] 0.0005/   0.0005    21.4      16.4      0.36      150       0.081
## [498] 105       2558      132       3390      2051      7839      5.11
## [505] 1023      450       11.3      7.8       0.96/     0.33/     0.98
## [512] 3.3       23.37     208.9     13407     90/       62.5/     0.14/
## [519] 23.37/    3.6       0.063     0.102     0.105     0.14      0.071
## [526] 0.085     0.07      1.93      4.75      95.8      0.580/    0.699/
## [533] 0.06      21.5      5.18      ~40/      ~30/      801       2.63
## [540] 65.68     0.64      0.28      10.62     3.56      1.21      0.25
## [547] <8.79     190.2     364.07    30.3      >1000.00  1.4       0.45
## [554] 0.0375    1.8       800       0.0206/   39.32     46/       45.4/
## [561] 0.41      1.11      0.47      0.58      3.75      2.9       8.6
## [568] 0.44      4.6       0.022     0.9       0.061/    3.75/     1.3/
## [575] 0.0001    2.84      0.00071   4.66      0.018     0.00017   5.38
## [582] 23.54     5.38/     0.0056    0.014     0.028     0.037     0.051
## [589] 0.056     0.08      0.37      1.12      1.75      3.5       7
## [596] 0.34      31.42     0.011     0.0000073 0.9535    0.0011    0.033
## [603] 0.0035/   0.57      2.6       0.00005   0.0298    2.78      0.027
## [610] 0.024     0.004     1.1       2.1       365.6     85.3      0.008
## [617] 3.82      57.7      0.00007   276       0.6/      440       40.44
## [624] 12.5      0.0002    >0.01     0.026     13.3/     >0.220    40.7
## [631] 28.4      0.79      20.3      4.15      0.46      86.9      137/
## [638] 0.30/     1.35      0.0004/   12.429    150/      1500      31.3
## [645] 0.625     0.71      451       1.75/     8.0/      3.5/      17.92
## [652] 0.025     0.116/    0.594     3.392     4.552     3.392/    288
## [659] 0.8/      136       0.202     0.161     0.015     2.6/      0.225
## [666] 0.45/     5.25      0.49      3.2       0.0027    754.2     0.016
## [673] 1.04      9.5858    5737      0.088     50.28     95.48     1056.7
## [680] 311.9     2.008     2675      1.51      2.94      503.6     0.61
## [687] 15.09     4399      1.51/     13.45     15        995       0.138
## [694] 0.493     63.54     54.67     0.84      6.36      37.5/     0.72
## [701] 8.38      4.37      31.37     0.43/     0.93      1.58      4.93
## [708] 5.77      21.6      0.93/     232.37    750       56.35     0.35
## [715] 0.17/     0.52      17.6      112.5     0.946     0.302     241.3
## [722] 99.75     7.7       4.28      0.126/    0.123     3.313     2.462
## [729] 14.34     0.000047  0.000074  0.0000814 0.000101  3.21      4.51
## [736] >5.0      <0.088    0.0299    51.16     4.679     4.411     4.316
## [743] 0.124     0.0112    0.0373    0.0641    2.46      1.07      0.141
## [750] 3.8/      1.34/     3.62      10.3      7.5       0.0428    0.428
## [757] 0.1500/   51.32     <0.0004   0.481     3.8       1.43/     9
## [764] 9.07      8.86      5.73      5.56      5.46      5.64      5.36
## [771] 4.44      3.11      2.68      2.761     2.644     2.556     3.336
## [778] 3.018     2.936     4.546     4.383     3.151     4.32      3.9
## [785] 3.59      2.26      2.15      5.01      5.08      4.52      4.13
## [792] 3.68      2.48      2.44      1.99      1.65      1.64      3.53
## [799] 2.75      5.27      5.3       6.03      4.61      3.4       3.36
## [806] 3.38      3.31      3.09      12.5/     33.6      7.379     14.785
## [813] 5.101     7.379/    1.48/     8.03      3.18      2.38      25/
## [820] 0.0812/   0.003/    296.62    80        0.745     45.37     0.26/
## [827] 20.0/     0.26      0.42      0.075     0.63/     0.397     0.137
## [834] 0.114/    1.01      0.04/     5.1       2.17      249.23/   >2500.00
```

```
##  [841] 382.31    150.46    81.09     37.01     124.03    48.2      251.3
##  [848] 788.5     251.3/    43.4      1.6/      3.23      1.67      0.38
##  [855] 32        0.2175    94        0.494     0.00027/  0.87      0.145
##  [862] 0.52/     41.94     1.092     2.403     2.947     2.403/    0.150/
##  [869] 0.0014    0.1039    1.72      250/      3.41      1.86      13.35
##  [876] 21.19     0.09/     7.14      12.57     24.54     41.6      168/
##  [883] 0.000069/ 0.000006  0.91      1.45      0.023     0.032     0.0076
##  [890] 0.061     0.0006    1.29      2.0/      7.5/      2.89      0.0032
##  [897] 0.0013    0.0063    0.0125    1.20/     13.75     31.16     9.49
##  [904] 28.2      65.37     21.06     6.80/     2.60/     12.0/     600
##  [911] 0.567/    0.470/    13.9/     10.5/     2.24/     52.2      0.95
##  [918] 0.07/     136.25/   26.025/   52.05/    272.5/    545/      104/
##  [925] 26/       104.1/    105/      0.0002/   0.077     4.485     2.967
##  [932] 2.667     0.00368   0.0218    2.844     2.689     2.608     0.00739
##  [939] 26.9      9.5       1.08      2.18/     8.51      2.99      0.0095
##  [946] 0.0009    1.09/     4.671     4.514     3.885     3.789     3.747
##  [953] 4.627     4.507     4.369     1.24      2.82      2.79      5.37
##  [960] 5.07      4.83      2.85      2.61      2.2       2.19      4.53
##  [967] 3.12      2.96      4.71      4.64      4.29      21.418    21/
##  [974] 6.76      6.27      6.13      4.08      3.28      3.03      3.0/
##  [981] 0.00039   17        39        76        76/       5.28      0.647
##  [988] 0.46/     0.056/    0.225/    0.31      4.0/      7.72      9.89
##  [995] 15.63     30.7      5.0/      4.39/     40.019    1.12/     0.00008
## [1002] 0.0231    0.76/     224.05/   0.0113    3.4859
## 1006 Levels: <0.0004 <0.025 <0.088 <0.5 <1.5 <10/ <2.5/ <4.00 <5.00 ... NR/
```

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

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

```r
summary(Neonics$Conc.1..Author.)
```

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

```
##       13        13        12        12        12        12        12        12
##      150      2.5/        25       60/       75/     0.02/    0.025/      0.29
##       12        12        12        12        12        11        11        11
##    37.5/        4/         5   (Other)
##       11        11        11      1817
```

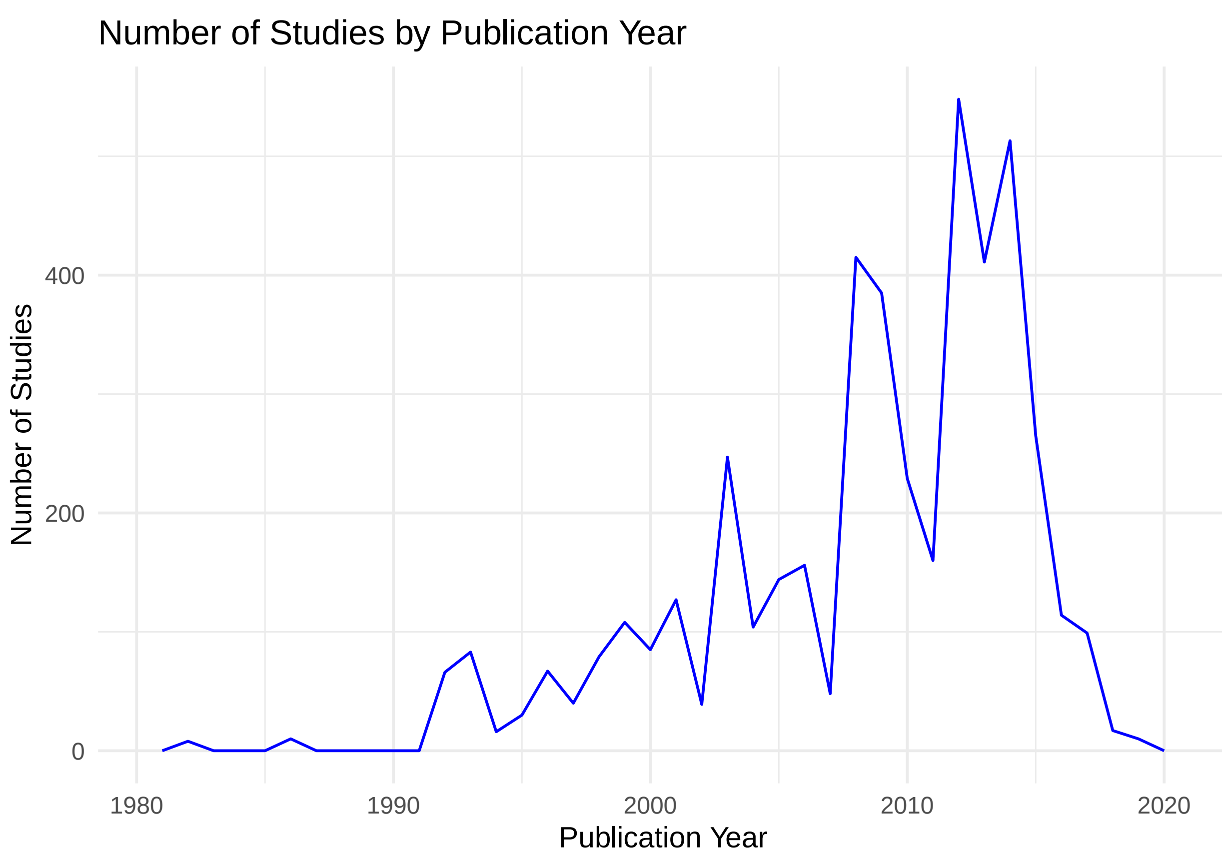
Answer:

The class of the Conc.1..Author. column is now numeric after conversion. It was originally a factor because it contained non-numeric characters or inconsistencies. After cleaning and converting the data, it is now suitable for numerical analysis.

Explore your data graphically (Neonics)

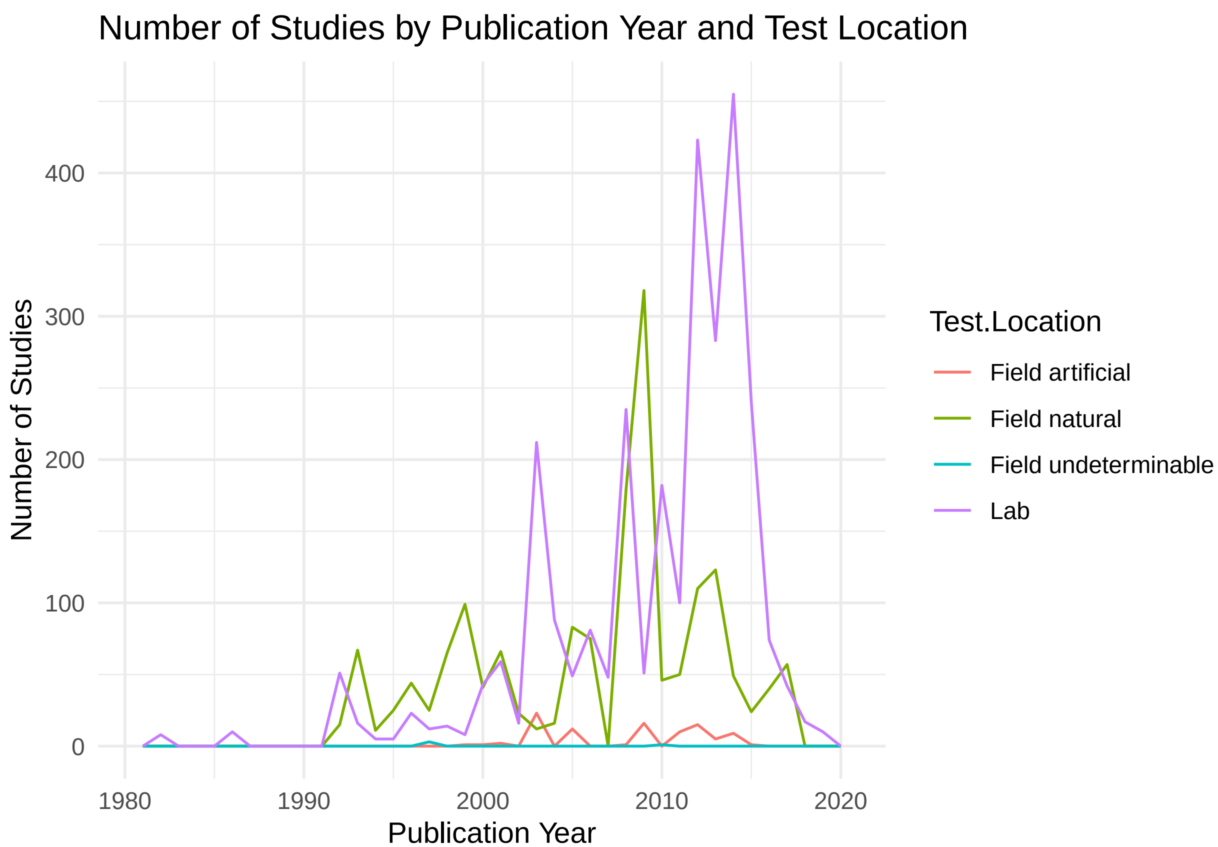
9. Using `geom_freqpoly`, generate a plot of the number of studies conducted by publication year.

```
ggplot(Neonics, aes(x = Publication.Year)) +
  geom_freqpoly(binwidth = 1, color = "blue") +
  labs(title = "Number of Studies by Publication Year",
       x = "Publication Year",
       y = "Number of Studies") +
  theme_minimal()
```



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

```r
ggplot(Neonics, aes(x = Publication.Year, color = Test.Location)) +
  geom_freqpoly(binwidth = 1) +
  labs(title = "Number of Studies by Publication Year and Test Location",
       x = "Publication Year",
       y = "Number of Studies") +
  theme_minimal()
```



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

> Answer:

The most common test location is the Lab, where the number of studies has increased significantly over the years. This trend suggests a growing emphasis on controlled laboratory experiments, which are essential for isolating specific effects of neonicotinoids.

The second most common test location is Field Natural, which peaked just before 2010. However, after reaching this peak, the number of studies in natural settings began to decline. This decline may indicate a shift in research focus or potential challenges in conducting field studies, such as environmental variability or funding issues.

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

[**TIP**: Add `theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))` to the end of your plot command to rotate and align the X-axis labels…]

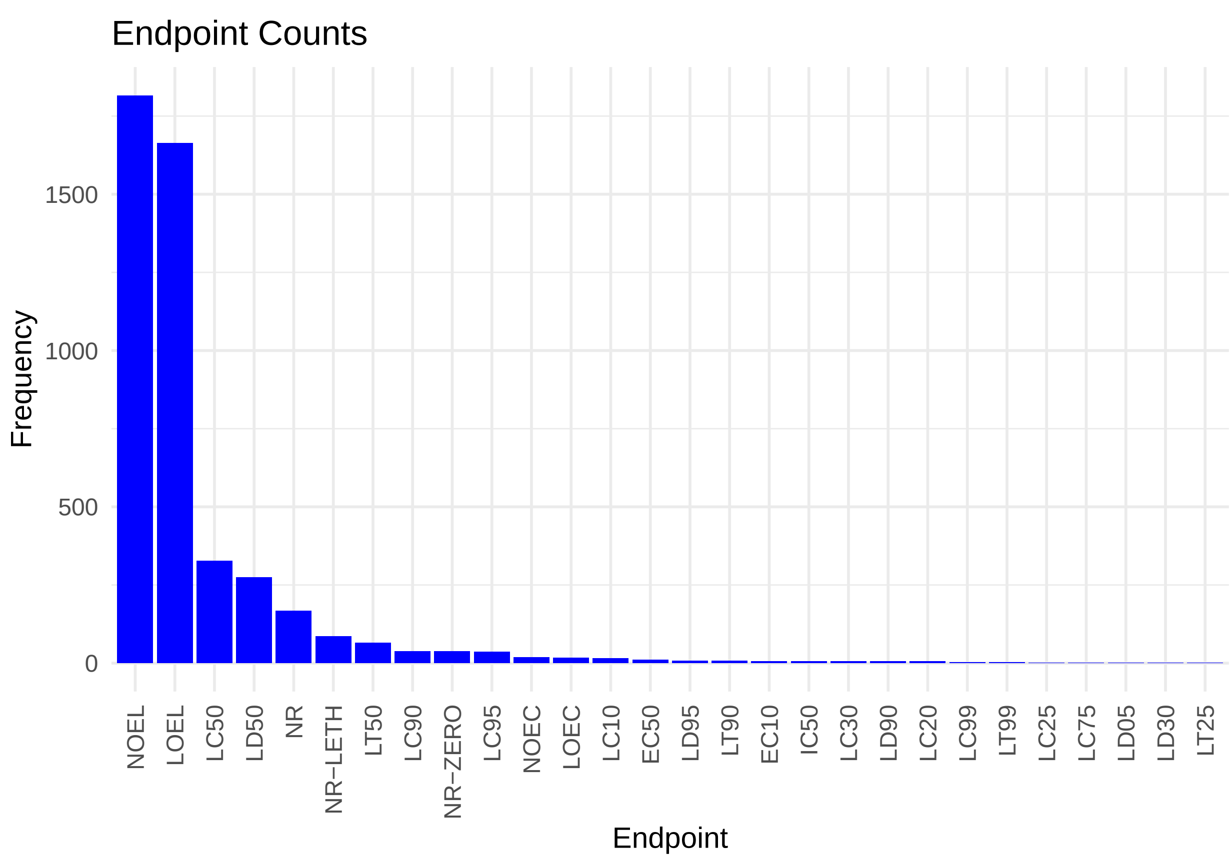
```r
#here I Load the necessary library
library(dplyr)
library(ggplot2)

# Then I Calculate the counts of each unique endpoint
endpoint_counts <- table(Neonics$Endpoint)

# later I Convert the counts to a dataframe for plotting
endpoint_counts_df <- data.frame(
  Endpoint = names(endpoint_counts),
  Frequency = as.numeric(endpoint_counts)
)

#Then I  Sort the dataframe by frequency in descending order
endpoint_counts_df <- endpoint_counts_df %>%
  arrange(desc(Frequency))

# At the end I Created the bar graph
ggplot(data = endpoint_counts_df, aes(x = reorder(Endpoint, -Frequency), y = Frequency)) +
  geom_bar(stat = "identity", fill = "blue") +
  labs(title = "Endpoint Counts",
       x = "Endpoint",
       y = "Frequency") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1))
```



Answer:

The two most common endpoints in the dataset are NOEL (No Observed Effect Level) and LOEL (Lowest Observed Effect Level). NOEL refers to the highest dose of a substance where no harmful effects are seen in the organisms, indicating a safe level of exposure. In contrast, LOEL is the lowest dose at which harmful effects are first observed, signaling that the substance is causing harm. Together, these endpoints help researchers understand safe and harmful exposure levels for substances.

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.

```
# I Check the class of the collectDate column
class(Litter$collectDate)
```

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

```
# Here I converted to the Date
Litter$collectDate <- as.Date(Litter$collectDate, format = "%Y-%m-%d") # Adjust format as needed

# Confirm the new class
class(Litter$collectDate)
```

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

```
# Here I use the unique function to find dates sampled in August 2018
august_dates <- unique(Litter$collectDate[Litter$collectDate >= "2018-08-01" & Litter$collectDate < "20
august_dates
```

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

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

```
# I am running this code to find unique plots sampled at Niwot Ridge
unique_plots <- unique(Litter$PlotID)  # Replace 'PlotID' with the actual column name for plot identifi
num_unique_plots <- length(unique_plots)
num_unique_plots
```

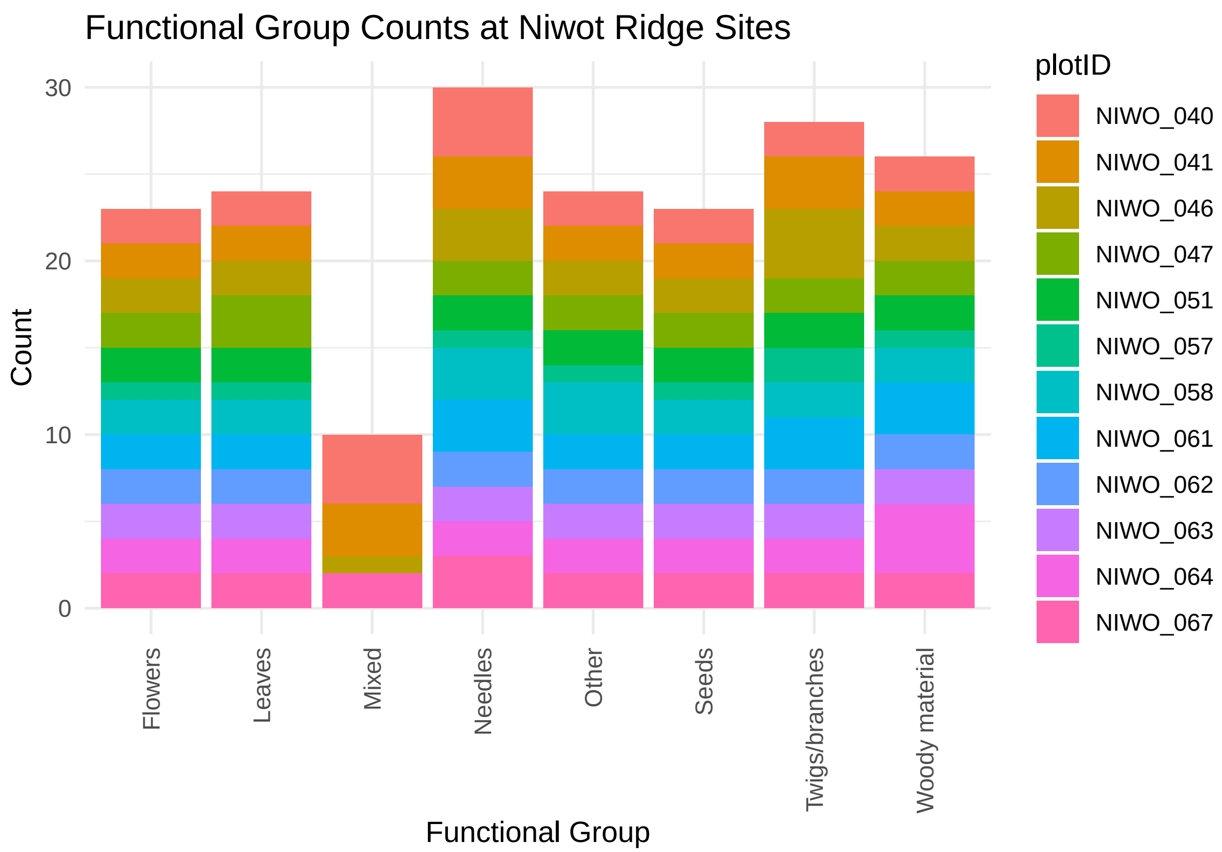
```
## [1] 0
```

Answer:

Using the unique function, we determined how many plots were sampled at Niwot Ridge. However, the result was 0, indicating that there are no unique plots found in the dataset under the specified column. The unique function lists distinct values, while the summary function provides a broader overview of statistics related to the variable, such as counts, means, and quartiles.

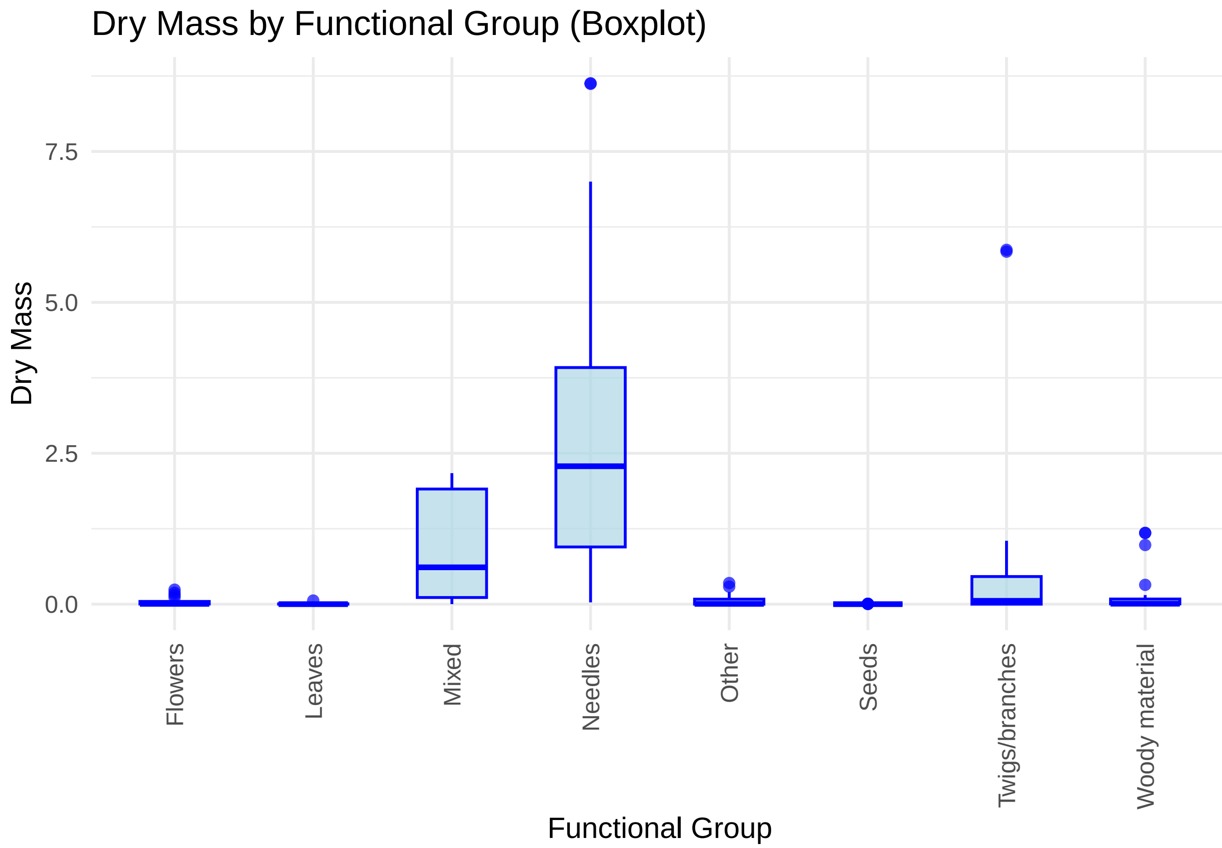
14. Create a bar graph of functionalGroup counts. This shows you what type of litter is collected at the Niwot Ridge sites. Notice that litter types are fairly equally distributed across the Niwot Ridge sites.

```r
ggplot(Litter, aes(x = functionalGroup, fill = plotID)) +
geom_bar() +
labs(title = "Functional Group Counts at Niwot Ridge Sites",
x = "Functional Group",
y = "Count") +
theme_minimal() +
theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1))
```



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

```r
# Boxplot for dryMass by functionalGroup
boxplot_plot <- ggplot(Litter, aes(x = functionalGroup, y = dryMass)) +
geom_boxplot(fill = "lightblue", color = "blue", alpha = 0.7, width = 0.5) +
labs(title = "Dry Mass by Functional Group (Boxplot)",
x = "Functional Group",
y = "Dry Mass") +
theme_minimal() +
theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust = 1))
print(boxplot_plot)
```



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

> Answer:

The boxplot and violin plot for dry mass by functional group reveal important insights about the distribution of dry mass values. The boxplot is more effective in this case because it clearly shows the central tendency with a median line, the spread of data through the interquartile range, and highlights any outliers, allowing for a comprehensive view of the data. In contrast, the violin plot mainly indicates the median without providing the same level of detail regarding data distribution. The functional groups that tend to have the highest biomass at these sites, as indicated by the median dry mass, are the Needles and Twigs/branches, which show the highest median values in the boxplot.

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

> Answer:

The types of litter that tend to have the highest biomass at these sites are Needles and Twigs/branches. This conclusion is based on the observation that these groups exhibit the highest median dry mass, as indicated by the horizontal line in the boxplot.