

# Data Wrangling

Katie Clouse

2025-03-19

## Installing packages

```
## Warning: package 'tidyverse' was built under R version 4.4.2
```

```
## Warning: package 'dplyr' was built under R version 4.4.2
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.5.1      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.1
## v purrr      1.0.2
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
## 'data.frame':   287 obs. of  16 variables:
## $ SampleID      : chr  "Corn2017LeafObjective2Collection1T1R1CAH2" "Corn2017LeafObjective2Collection1T1R1CAH2" ...
## $ Crop          : chr  "Corn" "Corn" "Corn" "Corn" ...
## $ Objective      : chr  "Objective 2" "Objective 2" "Objective 2" "Objective 2" ...
## $ Collection     : int   1 1 1 1 1 1 1 1 1 1 ...
## $ Compartment    : chr  "Leaf" "Leaf" "Leaf" "Leaf" ...
## $ DateSampled    : chr  "6/26/17" "6/26/17" "6/26/17" "6/26/17" ...
## $ GrowthStage    : chr  "V6" "V6" "V6" "V6" ...
## $ Treatment      : chr  "Conv." "Conv." "Conv." "Conv." ...
## $ Rep            : chr  "R1" "R1" "R1" "R1" ...
## $ Sample         : chr  "A" "B" "C" "A" ...
## $ Fungicide       : chr  "C" "C" "C" "F" ...
## $ Target_organism: chr  "Fungi" "Fungi" "Fungi" "Fungi" ...
## $ Location       : chr  "Kellogg Biological Station" "Kellogg Biological Station" "Kellogg Biological Station" ...
## $ Experiment      : chr  "LTER" "LTER" "LTER" "LTER" ...
## $ Year           : int   2017 2017 2017 2017 2017 2017 2017 2017 2017 2017 ...
## $ richness       : int    9 6 5 7 4 2 3 8 4 4 ...
```

**select() function** Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

filter()

```
##                               SampleID Crop Compartment DateSampled
## 1 Corn2017LeafObjective2Collection1T1R1CAH2 Corn          Leaf    6/26/17
## 2 Corn2017LeafObjective2Collection1T1R1CBA3 Corn          Leaf    6/26/17
## 3 Corn2017LeafObjective2Collection1T1R1CCB3 Corn          Leaf    6/26/17
## 4 Corn2017LeafObjective2Collection1T1R1FAC3 Corn          Leaf    6/26/17
## 5 Corn2017LeafObjective2Collection1T1R1FBD3 Corn          Leaf    6/26/17
## 6 Corn2017LeafObjective2Collection1T1R1FCE3 Corn          Leaf    6/26/17
##   GrowthStage Treatment Rep Sample Fungicide richness
## 1           V6      Conv. R1      A          C          9
## 2           V6      Conv. R1      B          C          6
## 3           V6      Conv. R1      C          C          5
## 4           V6      Conv. R1      A          F          7
## 5           V6      Conv. R1      B          F          4
## 6           V6      Conv. R1      C          F          2
```

```
##                               SampleID Crop Compartment DateSampled
## 1 Corn2017LeafObjective2Collection1T1R1CAH2 Corn          Leaf    6/26/17
## 2 Corn2017LeafObjective2Collection1T1R1CBA3 Corn          Leaf    6/26/17
## 3 Corn2017LeafObjective2Collection1T1R1CCB3 Corn          Leaf    6/26/17
## 4 Corn2017LeafObjective2Collection1T1R2CAF3 Corn          Leaf    6/26/17
## 5 Corn2017LeafObjective2Collection1T1R2CBG3 Corn          Leaf    6/26/17
## 6 Corn2017LeafObjective2Collection1T1R2CCH3 Corn          Leaf    6/26/17
##   GrowthStage Treatment Rep Sample Fungicide richness
## 1           V6      Conv. R1      A          C          9
## 2           V6      Conv. R1      B          C          6
## 3           V6      Conv. R1      C          C          5
## 4           V6      Conv. R2      A          C          3
## 5           V6      Conv. R2      B          C          8
## 6           V6      Conv. R2      C          C          4
```

```
##                               SampleID Crop Compartment DateSampled
## 1 Corn2017LeafObjective2Collection1T1R1CAH2 Corn          Leaf    6/26/17
## 2 Corn2017LeafObjective2Collection1T1R1CBA3 Corn          Leaf    6/26/17
## 3 Corn2017LeafObjective2Collection1T1R1FAC3 Corn          Leaf    6/26/17
## 4 Corn2017LeafObjective2Collection1T1R1FBD3 Corn          Leaf    6/26/17
## 5 Corn2017LeafObjective2Collection1T1R2CAF3 Corn          Leaf    6/26/17
## 6 Corn2017LeafObjective2Collection1T1R2CBG3 Corn          Leaf    6/26/17
##   GrowthStage Treatment Rep Sample Fungicide richness
## 1           V6      Conv. R1      A          C          9
## 2           V6      Conv. R1      B          C          6
## 3           V6      Conv. R1      A          F          7
## 4           V6      Conv. R1      B          F          4
## 5           V6      Conv. R2      A          C          3
## 6           V6      Conv. R2      B          C          8
```

mutate()

```
##                               SampleID Crop Compartment DateSampled
## 1 Corn2017LeafObjective2Collection1T1R1CAH2 Corn          Leaf    6/26/17
## 2 Corn2017LeafObjective2Collection1T1R1CBA3 Corn          Leaf    6/26/17
## 3 Corn2017LeafObjective2Collection1T1R1CCB3 Corn          Leaf    6/26/17
```

```
## 4 Corn2017LeafObjective2Collection1T1R1FAC3 Corn      Leaf      6/26/17
## 5 Corn2017LeafObjective2Collection1T1R1FBD3 Corn      Leaf      6/26/17
## 6 Corn2017LeafObjective2Collection1T1R1FCE3 Corn      Leaf      6/26/17
##      GrowthStage Treatment Rep Sample Fungicide richness    logRich
## 1          V6      Conv.  R1      A          C          9 2.1972246
## 2          V6      Conv.  R1      B          C          6 1.7917595
## 3          V6      Conv.  R1      C          C          5 1.6094379
## 4          V6      Conv.  R1      A          F          7 1.9459101
## 5          V6      Conv.  R1      B          F          4 1.3862944
## 6          V6      Conv.  R1      C          F          2 0.6931472
```

```
##                                     SampleID Crop Compartment DateSampled
## 1 Corn2017LeafObjective2Collection1T1R1CAH2 Corn      Leaf      6/26/17
## 2 Corn2017LeafObjective2Collection1T1R1CBA3 Corn      Leaf      6/26/17
## 3 Corn2017LeafObjective2Collection1T1R1CCB3 Corn      Leaf      6/26/17
## 4 Corn2017LeafObjective2Collection1T1R1FAC3 Corn      Leaf      6/26/17
## 5 Corn2017LeafObjective2Collection1T1R1FBD3 Corn      Leaf      6/26/17
## 6 Corn2017LeafObjective2Collection1T1R1FCE3 Corn      Leaf      6/26/17
```

```
##      GrowthStage Treatment Rep Sample Fungicide richness
## 1          V6      Conv.  R1      A          C          9
## 2          V6      Conv.  R1      B          C          6
## 3          V6      Conv.  R1      C          C          5
## 4          V6      Conv.  R1      A          F          7
## 5          V6      Conv.  R1      B          F          4
## 6          V6      Conv.  R1      C          F          2
```

```
##                                     logRich.SampleID logRich.Crop logRich.Compartment
## 1 Corn2017LeafObjective2Collection1T1R1CAH2      Corn      Leaf
## 2 Corn2017LeafObjective2Collection1T1R1CBA3      Corn      Leaf
## 3 Corn2017LeafObjective2Collection1T1R1CCB3      Corn      Leaf
## 4 Corn2017LeafObjective2Collection1T1R1FAC3      Corn      Leaf
## 5 Corn2017LeafObjective2Collection1T1R1FBD3      Corn      Leaf
## 6 Corn2017LeafObjective2Collection1T1R1FCE3      Corn      Leaf
```

```
##      logRich.DateSampled logRich.GrowthStage logRich.Treatment logRich.Rep
## 1          6/26/17          V6          Conv.      R1
## 2          6/26/17          V6          Conv.      R1
## 3          6/26/17          V6          Conv.      R1
## 4          6/26/17          V6          Conv.      R1
## 5          6/26/17          V6          Conv.      R1
## 6          6/26/17          V6          Conv.      R1
```

```
##      logRich.Sample logRich.Fungicide logRich.richness Crop_Treatment
## 1          A          C          9      Corn Conv.
## 2          B          C          6      Corn Conv.
## 3          C          C          5      Corn Conv.
## 4          A          F          7      Corn Conv.
## 5          B          F          4      Corn Conv.
## 6          C          F          2      Corn Conv.
```

the pipe%>% function

```
##                                     SampleID Crop Compartment DateSampled
## 1 Corn2017LeafObjective2Collection1T1R1CAH2 Corn      Leaf      6/26/17
## 2 Corn2017LeafObjective2Collection1T1R1CBA3 Corn      Leaf      6/26/17
## 3 Corn2017LeafObjective2Collection1T1R1CCB3 Corn      Leaf      6/26/17
```

```
## 4 Corn2017LeafObjective2Collection1T1R1FAC3 Corn      Leaf      6/26/17
## 5 Corn2017LeafObjective2Collection1T1R1FBD3 Corn      Leaf      6/26/17
## 6 Corn2017LeafObjective2Collection1T1R1FCE3 Corn      Leaf      6/26/17
##   GrowthStage Treatment Rep Sample Fungicide richness  logRich
## 1          V6      Conv.  R1      A          C          9 2.1972246
## 2          V6      Conv.  R1      B          C          6 1.7917595
## 3          V6      Conv.  R1      C          C          5 1.6094379
## 4          V6      Conv.  R1      A          F          7 1.9459101
## 5          V6      Conv.  R1      B          F          4 1.3862944
## 6          V6      Conv.  R1      C          F          2 0.6931472
```

### summarise() function

```
##   Mean.rich
## 1  2.304395
```

### multiple summary stats

```
##   Mean.rich  n    sd.dev  std.err
## 1  2.304395 144 0.7024667 0.0585389
```

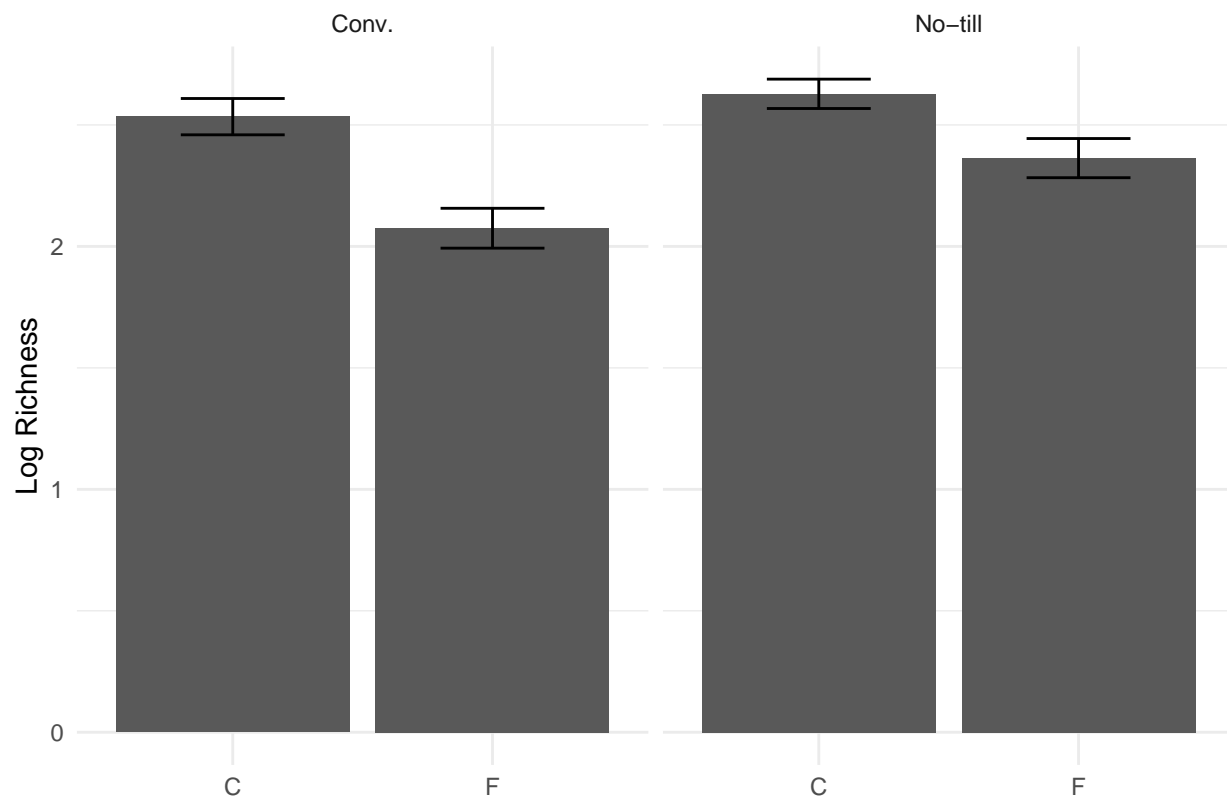
### group\_by() function

```
## 'summarise()' has grouped output by 'Treatment'. You can override using the
## '.groups' argument.
```

```
## # A tibble: 4 x 6
## # Groups:   Treatment [2]
##   Treatment Fungicide Mean.rich    n sd.dev std.err
##   <chr>      <chr>      <dbl> <int> <dbl>  <dbl>
## 1 Conv.      C          2.53   72  0.635  0.0748
## 2 Conv.      F          2.07   72  0.696  0.0820
## 3 No-till    C          2.63   72  0.513  0.0604
## 4 No-till    F          2.36   71  0.680  0.0807
```

### connecting to plotting

```
## 'summarise()' has grouped output by 'Treatment'. You can override using the
## '.groups' argument.
```



#### Joining

```
##                               SampleID Fungicide Crop Compartment
## 1 Corn2017LeafObjective2Collection1T1R1CAH2      C Corn      Leaf
## 2 Corn2017LeafObjective2Collection1T1R1CBA3      C Corn      Leaf
## 3 Corn2017LeafObjective2Collection1T1R1CCB3      C Corn      Leaf
## 4 Corn2017LeafObjective2Collection1T1R1FAC3      F Corn      Leaf
## 5 Corn2017LeafObjective2Collection1T1R1FBD3      F Corn      Leaf
## 6 Corn2017LeafObjective2Collection1T1R1FCE3      F Corn      Leaf
```

```
## GrowthStage Treatment Rep Sample
## 1          V6      Conv. R1      A
## 2          V6      Conv. R1      B
## 3          V6      Conv. R1      C
## 4          V6      Conv. R1      A
## 5          V6      Conv. R1      B
## 6          V6      Conv. R1      C
```

```
##                               SampleID richness
## 1 Corn2017LeafObjective2Collection1T1R1CAH2      9
## 2 Corn2017LeafObjective2Collection1T1R1CBA3      6
## 3 Corn2017LeafObjective2Collection1T1R1CCB3      5
## 4 Corn2017LeafObjective2Collection1T1R1FAC3      7
## 5 Corn2017LeafObjective2Collection1T1R1FBD3      4
## 6 Corn2017LeafObjective2Collection1T1R1FCE3      2
```

```
##                               SampleID Fungicide Crop Compartment
```

```
## 1 Corn2017LeafObjective2Collection1T1R1CAH2      C Corn      Leaf
## 2 Corn2017LeafObjective2Collection1T1R1CBA3      C Corn      Leaf
## 3 Corn2017LeafObjective2Collection1T1R1CCB3      C Corn      Leaf
## 4 Corn2017LeafObjective2Collection1T1R1FAC3      F Corn      Leaf
## 5 Corn2017LeafObjective2Collection1T1R1FBD3      F Corn      Leaf
## 6 Corn2017LeafObjective2Collection1T1R1FCE3      F Corn      Leaf
##   GrowthStage Treatment Rep Sample richness
## 1          V6      Conv. R1      A          9
## 2          V6      Conv. R1      B          6
## 3          V6      Conv. R1      C          5
## 4          V6      Conv. R1      A          7
## 5          V6      Conv. R1      B          4
## 6          V6      Conv. R1      C          2
```

## pivoting

```
## 'summarise()' has grouped output by 'Treatment'. You can override using the
## '.groups' argument.
```

```
## # A tibble: 4 x 3
## # Groups:   Treatment [2]
##   Treatment Fungicide Mean
##   <chr>      <chr>    <dbl>
## 1 Conv.      C        14.6
## 2 Conv.      F         9.75
## 3 No-till    C        15.4
## 4 No-till    F        13.1
```

## Wide format

```
## 'summarise()' has grouped output by 'Treatment'. You can override using the
## '.groups' argument.
```

```
## # A tibble: 2 x 3
## # Groups:   Treatment [2]
##   Treatment      C      F
##   <chr>    <dbl> <dbl>
## 1 Conv.    14.6  9.75
## 2 No-till  15.4 13.1
```

Now we can take the differences between the fungicide and control group now

```
## 'summarise()' has grouped output by 'Treatment'. You can override using the
## '.groups' argument.
```

```
## # A tibble: 2 x 4
## # Groups:   Treatment [2]
##   Treatment      C      F diff.fungicide
##   <chr>    <dbl> <dbl>         <dbl>
## 1 Conv.    14.6  9.75           4.89
## 2 No-till  15.4 13.1           2.32
```

If we want to plot this information, here is the code

```
## 'summarise()' has grouped output by 'Treatment'. You can override using the  
## '.groups' argument.
```

