Week 6 Assignment

Ellen Bledsoe

2024-02-20

Week 6 Assignment

Assignment Exercises

Set-up

Load the packages we will need. You can either load all of them individually (`readr`, `dplyr`, `tidyr`, `ggplot2`) or load the `tidyverse` package.

```
## -- 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   3.4.4     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 error
```

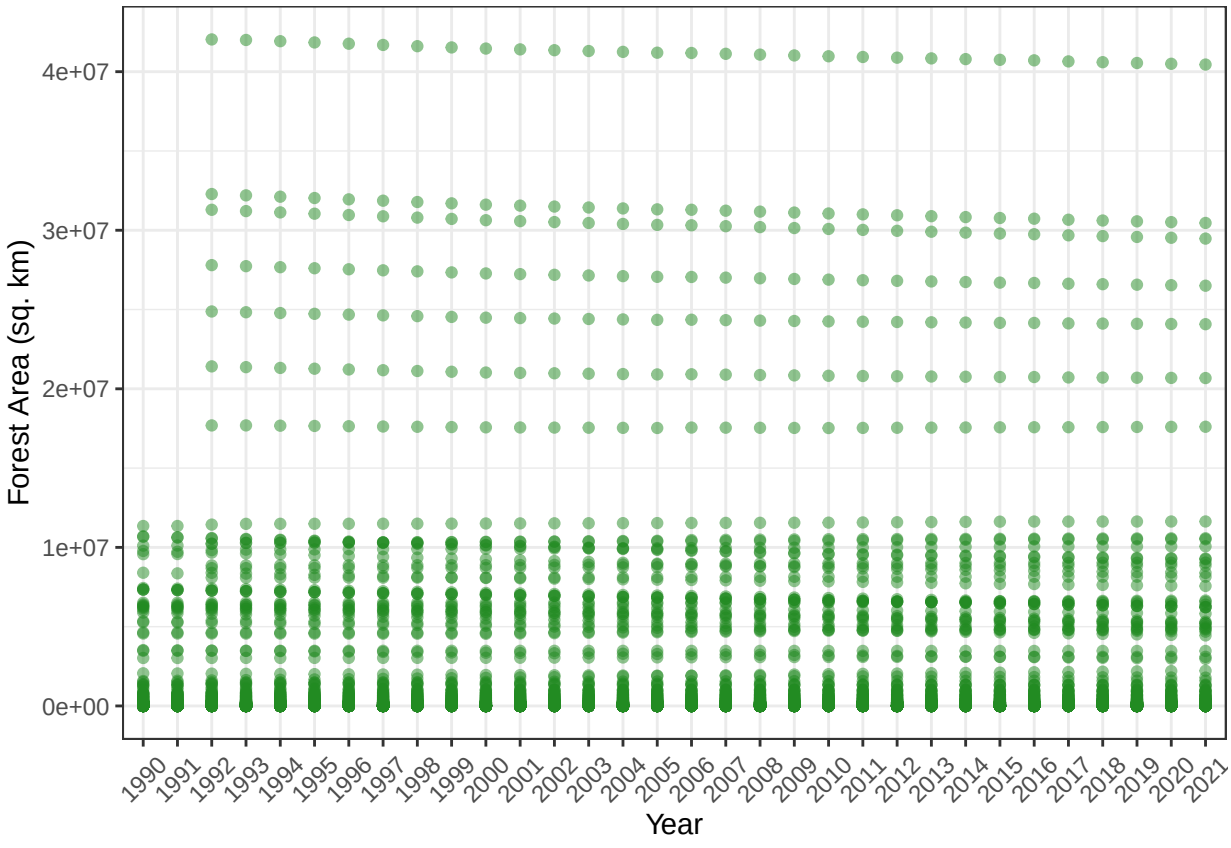
1. Forest Area per Country (15 pts)

```
## Rows: 266 Columns: 35
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (2): Country Name, Country Code
## dbl (32): 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, ...
## lgl  (1): 2022
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
## # A tibble: 8,778 x 4
##    `Country Name` `Country Code` Year  ForestArea_sqkm
##    <chr>          <chr>          <chr>           <dbl>
##  1 Aruba          ABW            1990              4.2
##  2 Aruba          ABW            1991              4.2
##  3 Aruba          ABW            1992              4.2
##  4 Aruba          ABW            1993              4.2
```

```
##  5 Aruba          ABW            1994              4.2
##  6 Aruba          ABW            1995              4.2
##  7 Aruba          ABW            1996              4.2
##  8 Aruba          ABW            1997              4.2
##  9 Aruba          ABW            1998              4.2
## 10 Aruba          ABW            1999              4.2
## # i 8,768 more rows
```

```
## # A tibble: 8,176 x 4
##    `Country Name` `Country Code` Year  ForestArea_sqkm
##    <chr>          <chr>          <chr>           <dbl>
##  1 Aruba          ABW            1990              4.2
##  2 Aruba          ABW            1991              4.2
##  3 Aruba          ABW            1992              4.2
##  4 Aruba          ABW            1993              4.2
##  5 Aruba          ABW            1994              4.2
##  6 Aruba          ABW            1995              4.2
##  7 Aruba          ABW            1996              4.2
##  8 Aruba          ABW            1997              4.2
##  9 Aruba          ABW            1998              4.2
## 10 Aruba          ABW            1999              4.2
## # i 8,166 more rows
```



2. OECD Data (10 pts)

```
## Rows: 127 Columns: 25
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (2): OECD_member, Country
## dbl (23): 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010, ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
## # A tibble: 6 x 25
##   OECD_member Country    `2000` `2001` `2002` `2003` `2004` `2005` `2006` `2007`
##   <chr>       <chr>       <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
## 1 OECD        Australia  3.77e5 3.77e5 4.00e5 4.00e5 4.02e5 4.06e5 4.12e5 4.17e5
## 2 OECD        Belgium    5.52e1 5.52e1 5.52e1 5.82e1 5.82e1 3.50e2 3.50e2 3.50e2
## 3 OECD        Canada     2.47e4 2.47e4 2.49e4 2.81e4 3.00e4 3.22e4 3.25e4 3.27e4
## 4 OECD        Chile      8.85e3 8.85e3 8.85e3 8.87e3 1.01e4 1.02e4 1.02e4 1.02e4
## 5 OECD        Colombia   2.94e4 2.94e4 2.94e4 2.94e4 2.94e4 6.09e4 6.09e4 6.09e4
## 6 OECD        Costa Rica 5.84e4 5.84e4 5.84e4 5.84e4 5.84e4 5.84e4 5.86e4 5.86e4
## # i 15 more variables: `2008` <dbl>, `2009` <dbl>, `2010` <dbl>, `2011` <dbl>,
## #   `2012` <dbl>, `2013` <dbl>, `2014` <dbl>, `2015` <dbl>, `2016` <dbl>,
## #   `2017` <dbl>, `2018` <dbl>, `2019` <dbl>, `2020` <dbl>, `2021` <dbl>,
## #   `2022` <dbl>
```

```
## # A tibble: 127 x 25
##    OECD_member Country   `2000` `2001` `2002` `2003` `2004` `2005` `2006` `2007`
##    <chr>       <chr>      <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
##  1 OECD        Australia 3.77e5 3.77e5 4.00e5 4.00e5 4.02e5 4.06e5 4.12e5 4.17e5
##  2 OECD        Belgium   5.52e1 5.52e1 5.52e1 5.82e1 5.82e1 3.50e2 3.50e2 3.50e2
##  3 OECD        Canada    2.47e4 2.47e4 2.49e4 2.81e4 3.00e4 3.22e4 3.25e4 3.27e4
##  4 OECD        Chile     8.85e3 8.85e3 8.85e3 8.87e3 1.01e4 1.02e4 1.02e4 1.02e4
##  5 OECD        Colombia  2.94e4 2.94e4 2.94e4 2.94e4 2.94e4 6.09e4 6.09e4 6.09e4
##  6 OECD        Costa Ri~ 5.84e4 5.84e4 5.84e4 5.84e4 5.84e4 5.84e4 5.86e4 5.86e4
##  7 OECD        Denmark   7.68e3 7.68e3 7.68e3 9.45e3 1.19e4 1.23e4 1.23e4 1.30e4
##  8 OECD        Estonia   5.81e2 5.81e2 5.81e2 5.81e2 6.47e3 6.53e3 6.53e3 6.54e3
##  9 OECD        Finland   7.17e3 7.22e3 7.22e3 7.22e3 7.25e3 7.45e3 7.46e3 7.46e3
## 10 OECD        France    7.88e4 7.88e4 7.88e4 7.89e4 7.89e4 8.09e4 8.12e4 8.47e4
## # i 117 more rows
## # i 15 more variables: `2008` <dbl>, `2009` <dbl>, `2010` <dbl>, `2011` <dbl>,
## #   `2012` <dbl>, `2013` <dbl>, `2014` <dbl>, `2015` <dbl>, `2016` <dbl>,
## #   `2017` <dbl>, `2018` <dbl>, `2019` <dbl>, `2020` <dbl>, `2021` <dbl>,
## #   `2022` <dbl>
```

```
## # A tibble: 2,921 x 4
##    OECD_member Country   Year  MarineProtectedArea_sqkm
##    <chr>       <chr>     <chr>                    <dbl>
##  1 OECD        Australia 2000                   376896.
##  2 OECD        Australia 2001                   377198.
##  3 OECD        Australia 2002                   399906.
##  4 OECD        Australia 2003                   399923
##  5 OECD        Australia 2004                   402052.
##  6 OECD        Australia 2005                   406364.
##  7 OECD        Australia 2006                   412438.
##  8 OECD        Australia 2007                   417116.
##  9 OECD        Australia 2008                   417560.
```

```
## 10 OECD        Australia 2009                   442165.
## # i 2,911 more rows
```

3. Santa Cruz Rodents Data Cleaning (20 pts)

```
## Rows: 51 Columns: 15
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr  (10): Site, Trap ID, Species, Status (R/N), Sex, Tail length, Hair samp...
## dbl   (4): Total Weight, Bag weight, Animal Weight, Hind foot length
## date  (1): Date
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

4. Remembering Joins (15 pts)

```
## New names:
## Rows: 80 Columns: 8
## -- Column specification
## -------------------------------------------------------- Delimiter: "," chr
## (4): Site, Trap Location, Type of Vegetation, Grouped_Veg dbl (4): ...1,
## Distance to Vegetation (m), Percent Veg Cover, Distance to Wa...
## i Use `spec()` to retrieve the full column specification for this data. i
## Specify the column types or set `show_col_types = FALSE` to quiet this message.
## * `` -> `...1`
```

```
## # A tibble: 80 x 8
##    RecordID Site     TrapID DistancetoVeg_m VegetationType PercentCover
##       <dbl> <chr>    <chr>            <dbl> <chr>                 <dbl>
##  1        1 Heritage 2A                 0   Bermuda grass            50
##  2        2 Heritage 2B                 0   Cheese bush              30
##  3        3 Heritage 2C                 5   Bermuda grass             0
##  4        4 Heritage 2D                 1   Salt cedar               20
##  5        5 Heritage 2E                 0   Bermuda grass            30
##  6        6 Heritage 2F                 0   Cockleburr               30
##  7        7 Heritage 2G                 0.5 Unknown grass            20
##  8        8 Heritage 2H                 0   Unknown grass            60
##  9        9 Heritage 2I                 0   Cheesebush               20
## 10       10 Heritage 2J                 0   Bermuda grass            50
## # i 70 more rows
## # i 2 more variables: DistancetoWater_m <dbl>, Grouped_Veg <chr>
```

```
## # A tibble: 80 x 3
##    Site     TrapID Grouped_Veg
##    <chr>    <chr>  <chr>
##  1 Heritage 2A     grass
##  2 Heritage 2B     shrubs
##  3 Heritage 2C     grass
##  4 Heritage 2D     shrubs
##  5 Heritage 2E     grass
##  6 Heritage 2F     forb
```

```
##  7 Heritage 2G     grass
##  8 Heritage 2H     grass
##  9 Heritage 2I     shrubs
## 10 Heritage 2J     grass
## # i 70 more rows
```

```
## Joining with `by = join_by(Site)`
```

```
## # A tibble: 80 x 17
##    Date       Site     `Trap ID` Species `Status (R/N)` Sex   `Total Weight`
##    <date>     <chr>    <chr>     <chr>   <chr>          <chr>          <dbl>
##  1 2022-11-14 Heritage 4C        SIOC    N              F                134
##  2 2022-11-14 Heritage 4C        SIOC    N              F                134
##  3 2022-11-14 Heritage 4C        SIOC    N              F                134
##  4 2022-11-14 Heritage 4C        SIOC    N              F                134
##  5 2022-11-14 Heritage 4C        SIOC    N              F                134
##  6 2022-11-14 Heritage 4C        SIOC    N              F                134
##  7 2022-11-14 Heritage 4C        SIOC    N              F                134
##  8 2022-11-14 Heritage 4C        SIOC    N              F                134
##  9 2022-11-14 Heritage 4C        SIOC    N              F                134
## 10 2022-11-14 Heritage 4C        SIOC    N              F                134
## # i 70 more rows
## # i 10 more variables: `Bag weight` <dbl>, `Animal Weight` <dbl>,
## #   `Hind foot length` <dbl>, `Tail length` <chr>, `Hair sample (Y/N)` <chr>,
## #   `Position (R/L)` <chr>, Handler <chr>, Notes <chr>, TrapID <chr>,
## #   Grouped_Veg <chr>
```

5. Santa Cruz Rodents Wrangling (20 pts)

```
## # A tibble: 51 x 17
##    Year  Month Day   Site  `Trap ID` Species `Status (R/N)` Sex   `Total Weight`
##    <chr> <chr> <chr> <chr> <chr>     <chr>   <chr>          <chr>          <dbl>
##  1 2022  11    14    Heri~ 4C        SIOC    N              F                134
##  2 2022  11    14    <NA>  4D        SIOC    N              M                136
##  3 2022  11    14    <NA>  4I        SIOC    N              <NA>              90
##  4 2022  11    14    <NA>  2H        REME    N              M                 38
##  5 2022  11    14    <NA>  4J        SIOC?   N              <NA>              NA
##  6 2022  11    14    <NA>  2F        REME    N              F                 22
##  7 2022  11    15    <NA>  4C        SIOC    R              <NA>              NA
##  8 2022  11    15    <NA>  4H        SIOC    N              F                 95
##  9 2022  11    15    <NA>  1H        REME    N              <NA>              26
## 10 2022  11    15    <NA>  1B        REME    N              F                 35
## # i 41 more rows
## # i 8 more variables: `Bag weight` <dbl>, `Animal Weight` <dbl>,
## #   `Hind foot length` <dbl>, `Tail length` <chr>, `Hair sample (Y/N)` <chr>,
## #   `Position (R/L)` <chr>, Handler <chr>, Notes <chr>
```

```
## # A tibble: 51 x 15
##    Date       Site     `Trap ID` Species `Status (R/N)` Sex   `Total Weight`
##    <chr>      <chr>    <chr>     <chr>   <chr>          <chr>          <dbl>
##  1 2022-11-14 Heritage 4C        SIOC    N              F                134
##  2 2022-11-14 <NA>     4D        SIOC    N              M                136
##  3 2022-11-14 <NA>     4I        SIOC    N              <NA>              90
```

```
##  4 2022-11-14 <NA>     2H        REME    N             M                 38
##  5 2022-11-14 <NA>     4J        SIOC?   N             <NA>              NA
##  6 2022-11-14 <NA>     2F        REME    N             F                 22
##  7 2022-11-15 <NA>     4C        SIOC    R             <NA>              NA
##  8 2022-11-15 <NA>     4H        SIOC    N             F                 95
##  9 2022-11-15 <NA>     1H        REME    N             <NA>              26
## 10 2022-11-15 <NA>     1B        REME    N             F                 35
## # i 41 more rows
## # i 8 more variables: 'Bag weight' <dbl>, 'Animal Weight' <dbl>,
## #   'Hind foot length' <dbl>, 'Tail length' <chr>, 'Hair sample (Y/N)' <chr>,
## #   'Position (R/L)' <chr>, Handler <chr>, Notes <chr>
```

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

```
## # A tibble: 10 x 3
## # Groups:   Site [3]
##    Site     Species Count
##    <chr>    <chr>   <int>
##  1 Drexel   CHPE        1
##  2 Heritage SIOC        1
##  3 <NA>     CHPE        2
##  4 <NA>     DIME        4
##  5 <NA>     DIME?       1
##  6 <NA>     NEAB        1
##  7 <NA>     PEER        5
##  8 <NA>     REME       10
##  9 <NA>     SIOC       25
## 10 <NA>     SIOC?       1
```

```
## # A tibble: 3 x 9
## # Groups:   Site [3]
##   Site      CHPE  SIOC  DIME 'DIME?'  NEAB  PEER  REME 'SIOC?'
##   <chr>    <int> <int> <int>   <int> <int> <int> <int>   <int>
## 1 Drexel       1     0     0       0     0     0     0       0
## 2 Heritage     0     1     0       0     0     0     0       0
## 3 <NA>         2    25     4       1     1     5    10       1
```

6. Mammals (20 pts)

The code chunk below has some made-up mammal data. Run the code chunk below to complete question 5.

```
##   site      genus   species density mass
## 1    1     Suncus  etruscus     6.2  4.2
## 2    1      Sorex  cinereus     5.2  5.0
## 3    2     Myotis nigricans    11.0  9.1
## 4    3 Notiosorex crawfordi     1.2  8.6
## 5    3     Suncus  etruscus     9.4  4.1
## 6    3     Myotis nigricans     9.6  8.7
```

```
## # A tibble: 12 x 5
##     site genus      species   measurement value
```

```
##    <dbl> <chr>      <chr>     <chr>       <dbl>
##  1     1 Suncus     etruscus  density       6.2
##  2     1 Suncus     etruscus  mass          4.2
##  3     1 Sorex      cinereus  density       5.2
##  4     1 Sorex      cinereus  mass          5
##  5     2 Myotis     nigricans density      11
##  6     2 Myotis     nigricans mass          9.1
##  7     3 Notiosorex crawfordi density       1.2
##  8     3 Notiosorex crawfordi mass          8.6
##  9     3 Suncus     etruscus  density       9.4
## 10     3 Suncus     etruscus  mass          4.1
## 11     3 Myotis     nigricans density       9.6
## 12     3 Myotis     nigricans mass          8.7
```

```
## # A tibble: 12 x 4
##     site taxon                measurement value
##    <dbl> <chr>                <chr>       <dbl>
##  1     1 Suncus etruscus      density       6.2
##  2     1 Suncus etruscus      mass          4.2
##  3     1 Sorex cinereus       density       5.2
##  4     1 Sorex cinereus       mass          5
##  5     2 Myotis nigricans     density      11
##  6     2 Myotis nigricans     mass          9.1
##  7     3 Notiosorex crawfordi density       1.2
##  8     3 Notiosorex crawfordi mass          8.6
##  9     3 Suncus etruscus      density       9.4
## 10     3 Suncus etruscus      mass          4.1
## 11     3 Myotis nigricans     density       9.6
## 12     3 Myotis nigricans     mass          8.7
```

```
## # A tibble: 6 x 4
##    site taxon                density  mass
##   <dbl> <chr>                  <dbl> <dbl>
## 1     1 Suncus etruscus          6.2   4.2
## 2     1 Sorex cinereus           5.2   5
## 3     2 Myotis nigricans        11     9.1
## 4     3 Notiosorex crawfordi     1.2   8.6
## 5     3 Suncus etruscus          9.4   4.1
## 6     3 Myotis nigricans         9.6   8.7
```