

functions-after.Rmd

Jai Jot Kaur

11/10/2022

If surveys.csv, species.csv, and plots.csv are not available in your workspace download them, and load them into R using read.csv().

```
surveys <- read.csv("../data-raw/surveys.csv")
species <- read.csv("../data-raw/species.csv")
plots <- read.csv("../data-raw/plots.csv")
```

Combine the surveys and species tables into a single data frame.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

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

```
combined <- inner_join(surveys, species, by = "species_id")
```

Write a function that: Takes three arguments - a data frame (the combined table created before), a genus name, and a species name, Uses functions from the dplyr package to produce a data frame with two columns: year and count, where count is the number of individuals (i.e., the number of rows) for the species indicated by genus and species in that year, Returns the resulting data frame. Note that this data frame is a time series by year.

```
time_series <- function(combined, genus_name = "Neotoma", species_name = "albigula") {
  combined %>%
    select(year, genus, species) %>%
    filter(genus == genus_name, species == species_name) %>%
    group_by(year) %>%
    summarise(count = n()) -> data_frame_result
  return(data_frame_result)
}
```

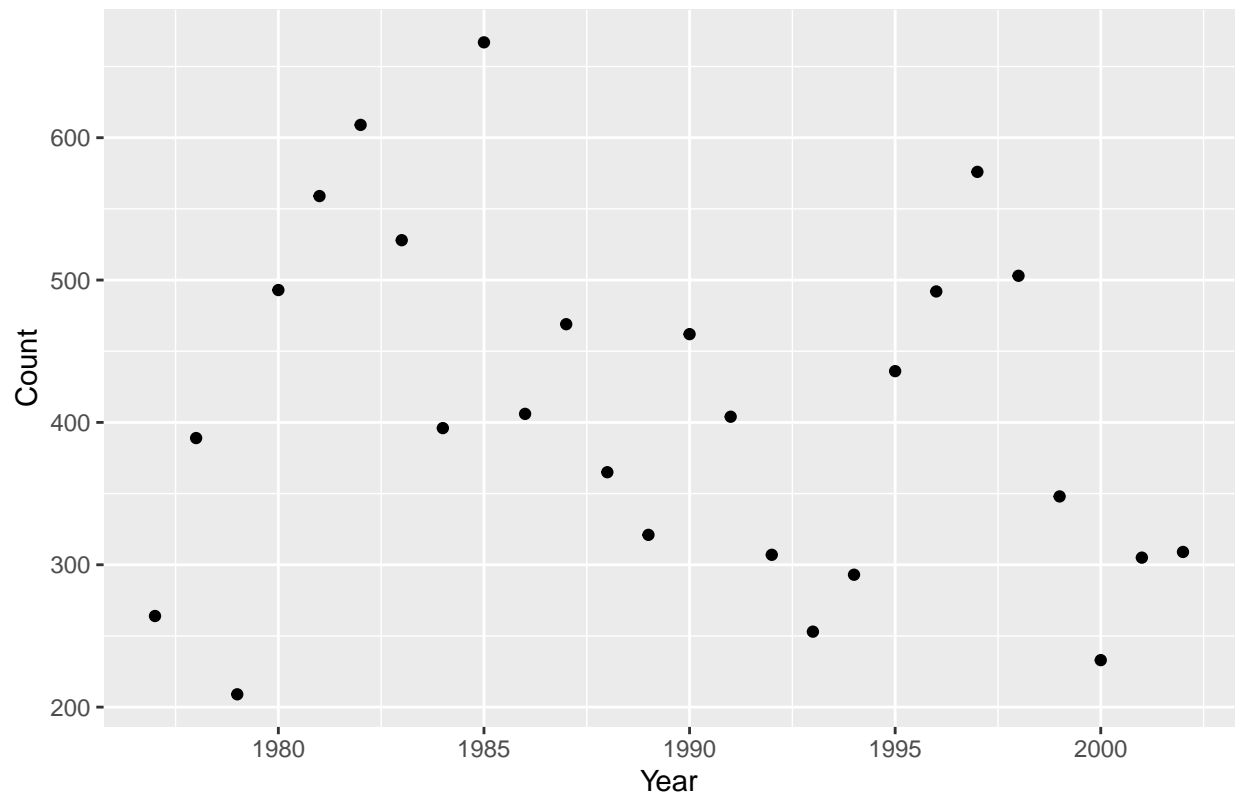
```
time_series(combined)
```

```
## # A tibble: 26 x 2
##   year count
##   <int> <int>
## 1  1977    31
## 2  1978    48
## 3  1979    30
## 4  1980    57
## 5  1981    63
## 6  1982   111
## 7  1983    98
## 8  1984    64
## 9  1985    45
## 10 1986    60
## # ... with 16 more rows
```

Use your function to get the data frame of time series by year for the genus *Dipodomys* and the species *merriami* and then make a graph of this time series using ggplot that has year on the x axis, count on the y axis, and displays the data as points.

```
Dipodomys_merriami <- time_series(combined, "Dipodomys", "merriami")
library("ggplot2")
ggplot(data = Dipodomys_merriami, mapping = aes(x = year, y = count)) +
  geom_point() +
  labs(x = "Year", y = "Count", title = "Time Series by Year for Dipodomys merriami")
```

Time Series by Year for *Dipodomys merriami*



Use your function to get the data frame time series by year for the genus *Chaetodipus* and species *penicillatus* and then make a graph of this time-series using ggplot that has year on the x axis, count on the y axis, and displays the data as blue points (with size = 1) connected by blue lines (with size = 2). Change the x axis label to Year and the y axis label to Number of Individuals.

```
Chaetodipus_penicillatus <- time_series(combined, "Chaetodipus", "penicillatus")
library("ggplot2")
ggplot(data = Chaetodipus_penicillatus, mapping = aes(x = year, y = count)) +
  geom_point(size = 1, color = "blue") +
  geom_line(size = 2, color = "blue") +
  labs(x = "Year", y = "Number of Individuals", title = "Time Series by Year for Chaetodipus penicillatus")
```

