

Creating functions in R

Marcos Padilla-Ruiz

2022-11-10

1. load data

```
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

plots <- read.csv(file = "../raw_data/plots.csv")
species <- read.csv(file = "../raw_data/species.csv")
surveys <- read.csv(file = "../raw_data/surveys.csv")
```

2. combine the survey and species data

```
combined <- inner_join(surveys, species)
```

```
## Joining, by = "species_id"
```

```
head(combined)
```

```
##   record_id month day year plot_id species_id sex hindfoot_length weight
## 1         1     7  16 1977      2         NL   M             32      NA
## 2         2     7  16 1977      3         NL   M             33      NA
## 3         3     7  16 1977      2         DM   F             37      NA
## 4         4     7  16 1977      7         DM   M             36      NA
## 5         5     7  16 1977      3         DM   M             35      NA
## 6         6     7  16 1977      1         PF   M             14      NA
##           genus species  taxa
## 1    Neotoma albigula Rodent
## 2    Neotoma albigula Rodent
## 3  Dipodomys merriami Rodent
## 4  Dipodomys merriami Rodent
## 5  Dipodomys merriami Rodent
## 6 Perognathus  flavus Rodent
```

3. write a function

```
combined1 <- function(combined, genus_name = "Neotoma", species_name = "albigula") {  
  combined %>%  
    select(year, genus, species) %>%  
    filter(genus == genus_name, species == species_name) %>%  
    group_by(year) %>%  
    summarize(count_by_year = n()) -> data_frame_result  
  
  return(data_frame_result)  
}
```

```
combined1(combined)
```

```
## # A tibble: 26 x 2  
##   year count_by_year  
##   <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
```

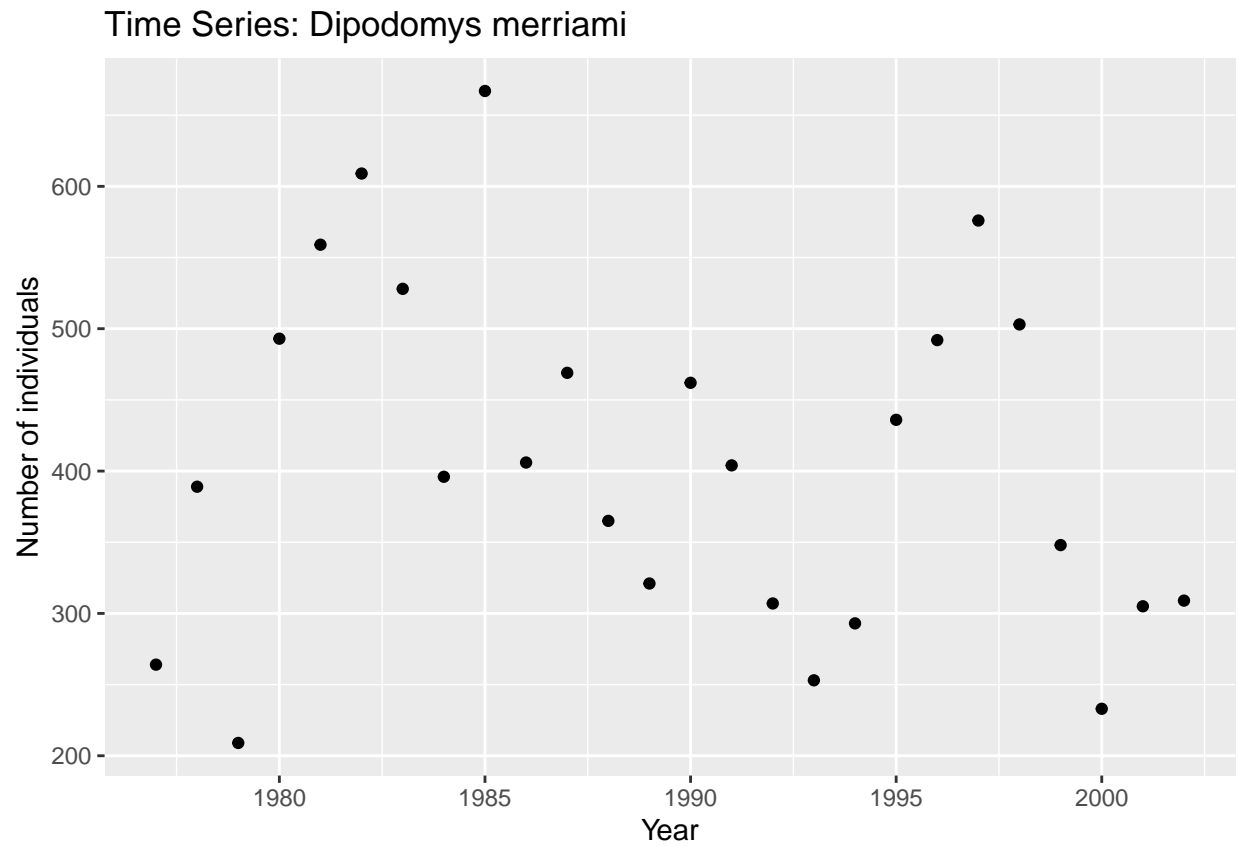
```
plot1 <- combined1(combined, genus_name = "Dipodomys", species_name = "merriami")  
plot2 <- combined1(combined, genus_name = "Chaetodipus", species_name = "penicillatus")
```

```
plot1
```

```
## # A tibble: 26 x 2  
##   year count_by_year  
##   <int>         <int>  
## 1  1977            264  
## 2  1978            389  
## 3  1979            209  
## 4  1980            493  
## 5  1981            559  
## 6  1982            609  
## 7  1983            528  
## 8  1984            396  
## 9  1985            667  
## 10 1986            406  
## # ... with 16 more rows
```

```
library(ggplot2)  
ggplot() +  
  geom_point(data = plot1,
```

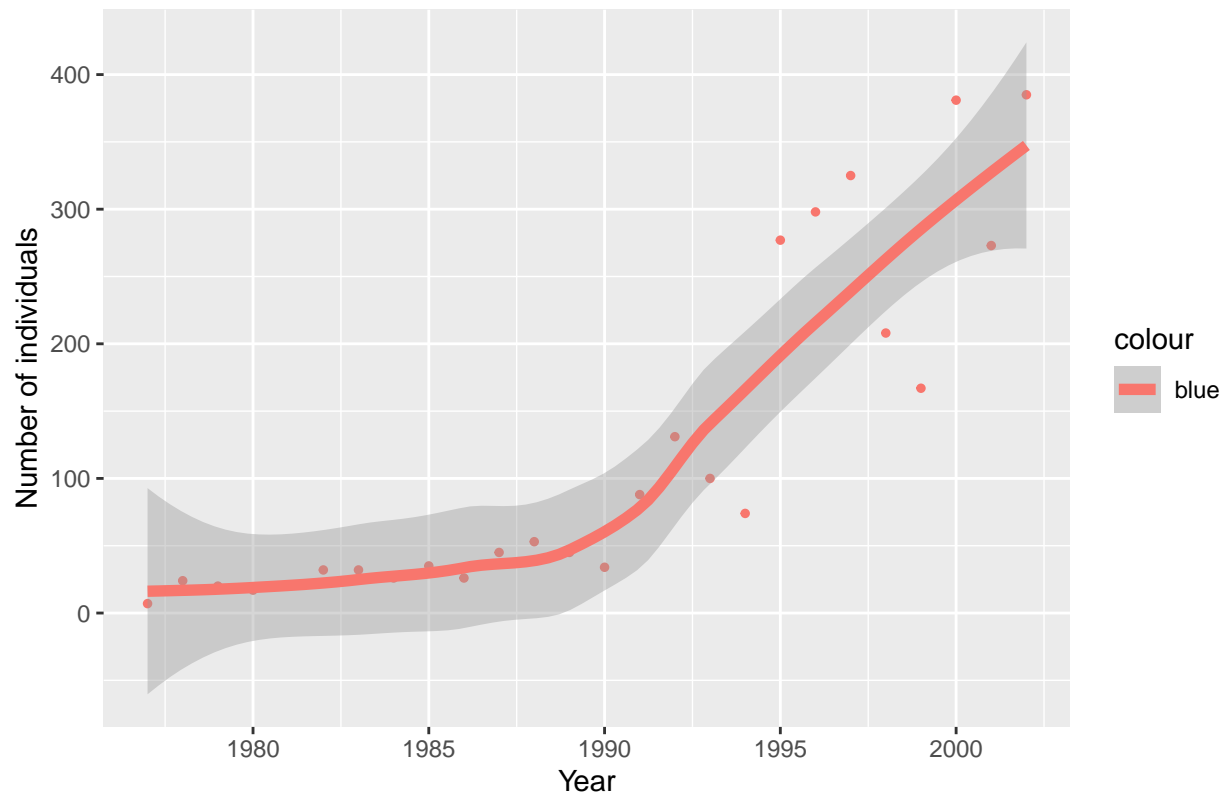
```
mapping = aes(x = year, y = count_by_year)) +
labs(title = "Time Series: Dipodomys merriami", x = "Year", y = "Number of individuals")
```



```
library(ggplot2)
ggplot() +
  geom_point(data = plot2,
            mapping = aes(x = year, y = count_by_year, color = "blue"), size = 1) +
  geom_smooth(data = plot2,
            mapping = aes(x = year, y = count_by_year, color = "blue"), size = 2) +
  labs(title = "Time Series: Chaetodipus penicillatus", x = "Year", y = "Number of individuals")

## 'geom_smooth()' using method = 'loess' and formula 'y ~ x'
```

Time Series: Chaetodipus penicillatus



not sure why the dots and lines are not blue