

class_24Oct

Vinay K L

2023-10-24

```
library(palmerpenguins)
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
```

```
palmerpenguins::penguins
```

```
## # A tibble: 344 x 8
##   species island bill_length_mm bill_depth_mm flipper_length_mm body_mass_g
##   <fct>   <fct>         <dbl>         <dbl>         <int>         <int>
## 1 Adelie  Torgersen         39.1          18.7          181          3750
## 2 Adelie  Torgersen         39.5          17.4          186          3800
## 3 Adelie  Torgersen         40.3           18          195          3250
## 4 Adelie  Torgersen          NA           NA           NA           NA
## 5 Adelie  Torgersen         36.7          19.3          193          3450
## 6 Adelie  Torgersen         39.3          20.6          190          3650
## 7 Adelie  Torgersen         38.9          17.8          181          3625
## 8 Adelie  Torgersen         39.2          19.6          195          4675
## 9 Adelie  Torgersen         34.1          18.1          193          3475
## 10 Adelie Torgersen         42           20.2          190          4250
## # i 334 more rows
## # i 2 more variables: sex <fct>, year <int>
```

For each species and island combination, find the average body mass and bill length.

```
penguins %>%
  group_by(species, island)%>%
  summarise(avg_bodymass = mean(body_mass_g, na.rm = TRUE),
            avg_billlength = mean(bill_length_mm, na.rm = TRUE)
  )
```

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

```
## # A tibble: 5 x 4
## # Groups:   species [3]
##   species island    avg_bodymass avg_billlength
##   <fct>    <fct>         <dbl>         <dbl>
## 1 Adelie  Biscoe           3710.           39.0
## 2 Adelie  Dream            3688.           38.5
## 3 Adelie  Torgersen        3706.           39.0
## 4 Chinstrap Dream      3733.           48.8
## 5 Gentoo  Biscoe           5076.           47.5
```

For the same species, do male penguins have larger body mass than females?

```
penguins %>%
  group_by(species, sex) %>%
  summarise(avg_bodymass = mean(body_mass_g, na.rm = TRUE)
  )
```

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

```
## # A tibble: 8 x 3
## # Groups:   species [3]
##   species sex    avg_bodymass
##   <fct>   <fct>         <dbl>
## 1 Adelie female       3369.
## 2 Adelie male       4043.
## 3 Adelie <NA>         3540
## 4 Chinstrap female    3527.
## 5 Chinstrap male     3939.
## 6 Gentoo female     4680.
## 7 Gentoo male     5485.
## 8 Gentoo <NA>       4588.
```