



faceting in ggplot2 with reference data layer

Gina Reynolds | 2024-02-07 | Image credit: Georgi Petrov, Upsplash

Small multiples (facets) are a powerful tool.

Small multiples (facets) are a powerful tool.

They break up and display data by category, so patterns within each category can be appreciated.

Small multiples (facets) are a powerful tool.

They break up and display data by category, so patterns within each category can be appreciated.

But a simple facet means that comparison to the whole becomes more of an exercise in mental gymnastics.

Small multiples (facets) are a powerful tool.

They break up and display data by category, so patterns within each category can be appreciated.

But a simple facet means that comparison to the whole becomes more of an exercise in mental gymnastics.

A solution is to present a reference layer of all the data beneath the featured category data. Let's look at the syntax in ggplot2.

```
library(tidyverse)
```

```
library(tidyverse)
```

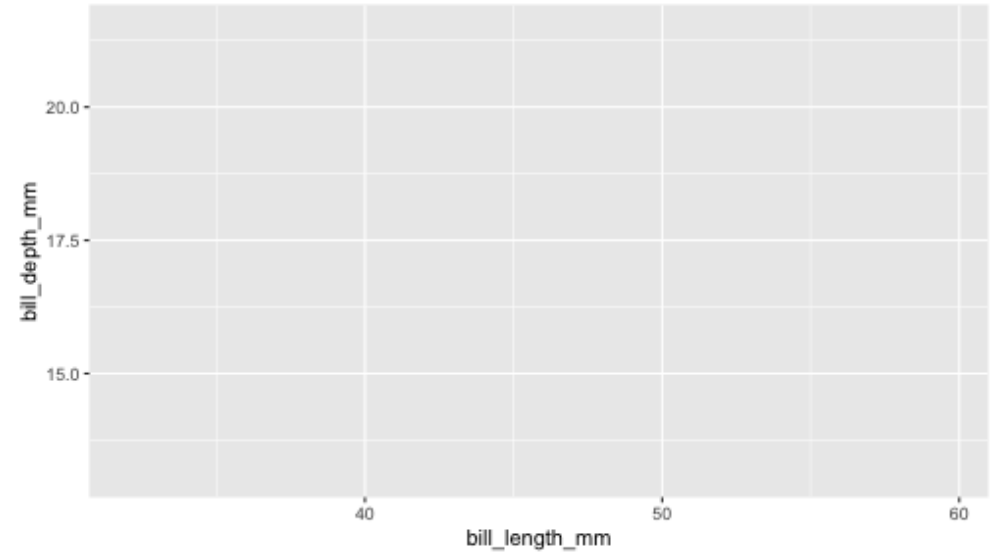
```
palmerpenguins::penguins
```

```
## # A tibble: 344 × 8
##   species island   bill_length_mm bill_depth_mm flipper_l
##   <fct>   <fct>         <dbl>         <dbl>
## 1 Adelie  Torgersen         39.1           18.7
## 2 Adelie  Torgersen         39.5           17.4
## 3 Adelie  Torgersen         40.3            18
## 4 Adelie  Torgersen          NA            NA
## 5 Adelie  Torgersen         36.7           19.3
## 6 Adelie  Torgersen         39.3           20.6
## 7 Adelie  Torgersen         38.9           17.8
## 8 Adelie  Torgersen         39.2           19.6
## 9 Adelie  Torgersen         34.1           18.1
## 10 Adelie Torgersen         42             20.2
## # i 334 more rows
## # i 2 more variables: sex <fct>, year <int>
```

```
library(tidyverse)
```

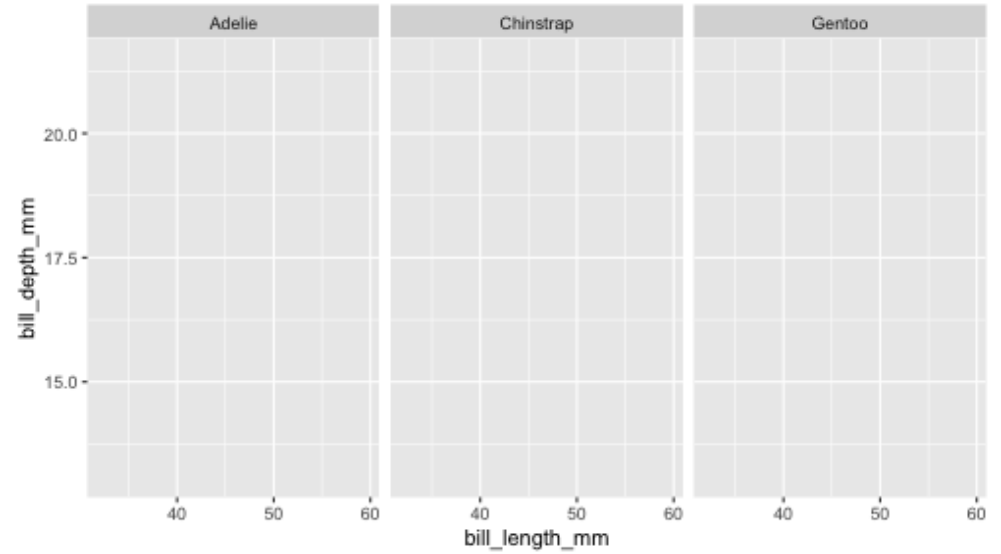
```
palmerpenguins::penguins |>
```

```
ggplot(aes(x = bill_length_mm, y = bill_depth_mm))
```



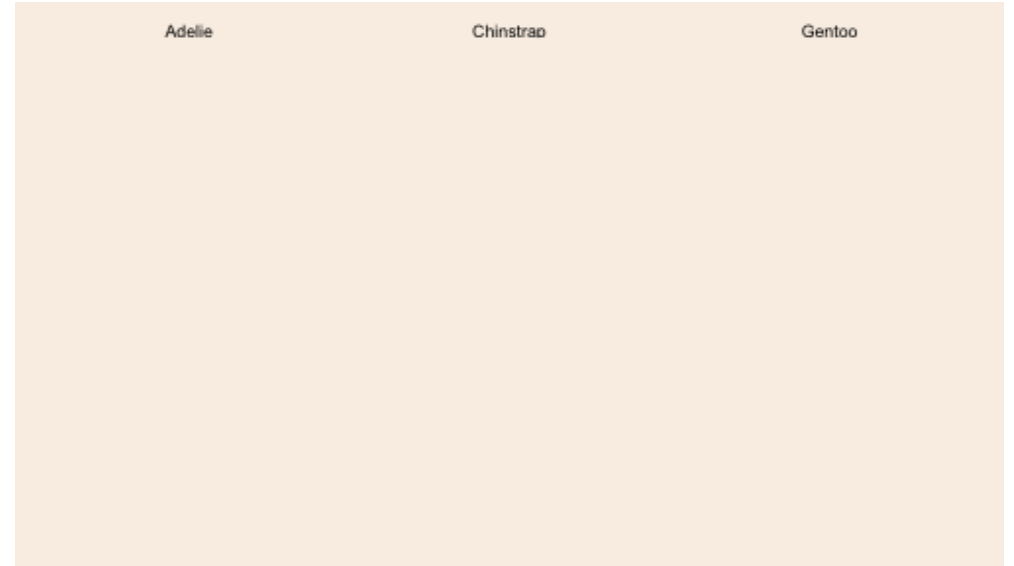

```
library(tidyverse)
```

```
palmerpenguins::penguins |>  
  ggplot(aes(x = bill_length_mm, y = bill_depth_mm))  
  facet_wrap(~ species)
```



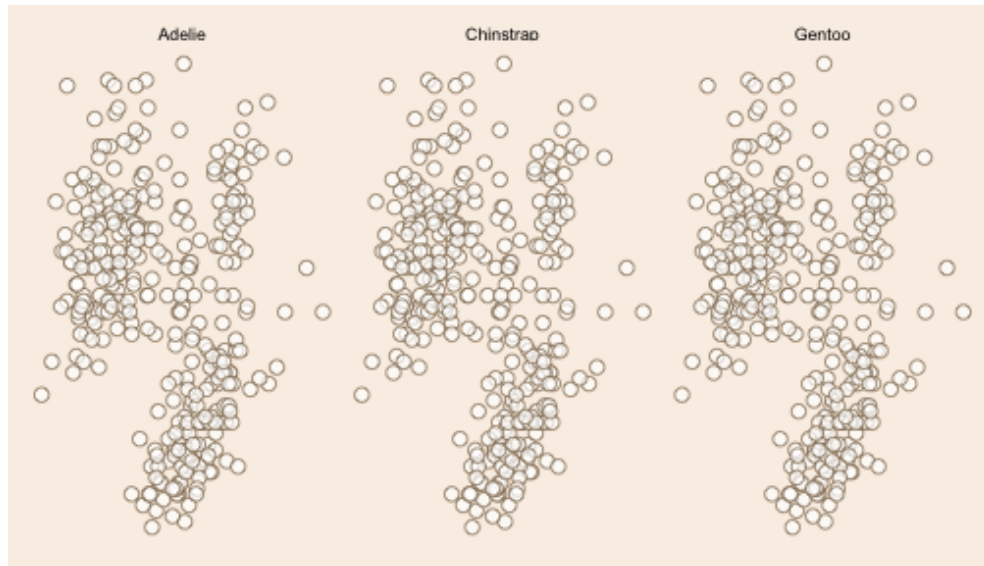
```
library(tidyverse)
```

```
palmerpenguins::penguins |>  
  ggplot(aes(x = bill_length_mm, y = bill_depth_mm))  
  facet_wrap(~ species) +  
  ggstamp::theme_void_fill("linen")
```



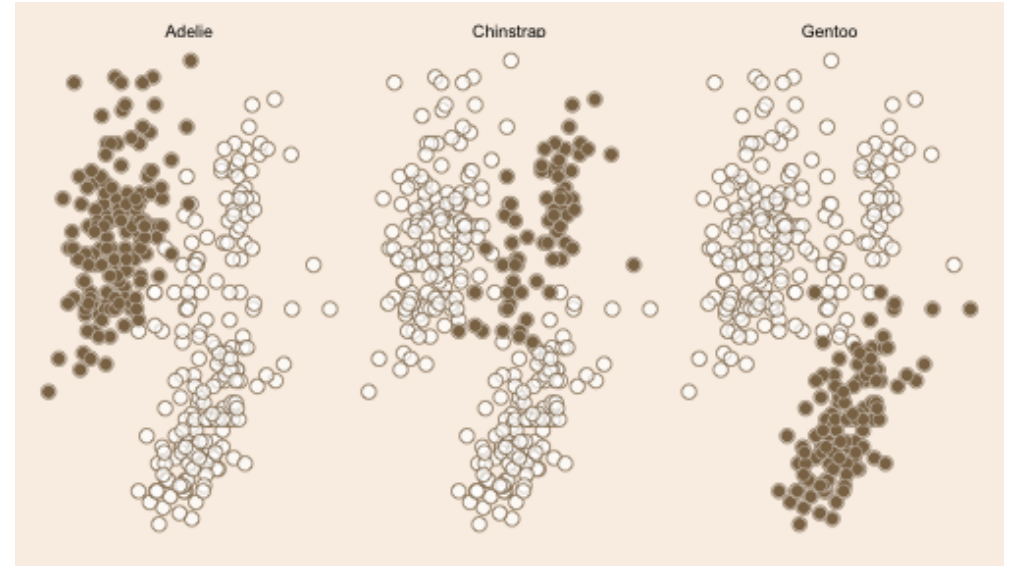
```
library(tidyverse)
```

```
palmerpenguins::penguins |>  
  ggplot(aes(x = bill_length_mm, y = bill_depth_mm))  
  facet_wrap(~ species) +  
  ggstamp::theme_void_fill("linen") +  
  # plot all data in each species facet  
  geom_point(data =  
    . %>% select(-species),  
    size = 3, shape = 21, stroke = .7,  
    color = "burlywood4", fill = "white",  
    alpha = .7)
```



```
library(tidyverse)
```

```
palmerpenguins::penguins |>  
  ggplot(aes(x = bill_length_mm, y = bill_depth_mm))  
  facet_wrap(~ species) +  
  ggstamp::theme_void_fill("linen") +  
  # plot all data in each species facet  
  geom_point(data =  
    . %>% select(-species),  
    size = 3, shape = 21, stroke = .7,  
    color = "burlywood4", fill = "white",  
    alpha = .7) +  
  # use global data and plot each species in its facet  
  geom_point(size = 3, shape = 21, stroke = .2,  
    color = "white", fill = "burlywood4")
```



Discussion

In the first layer, *all* data is displayed. This is because we've faceted by species, a variable we've remove via `. %>% select(-species)`

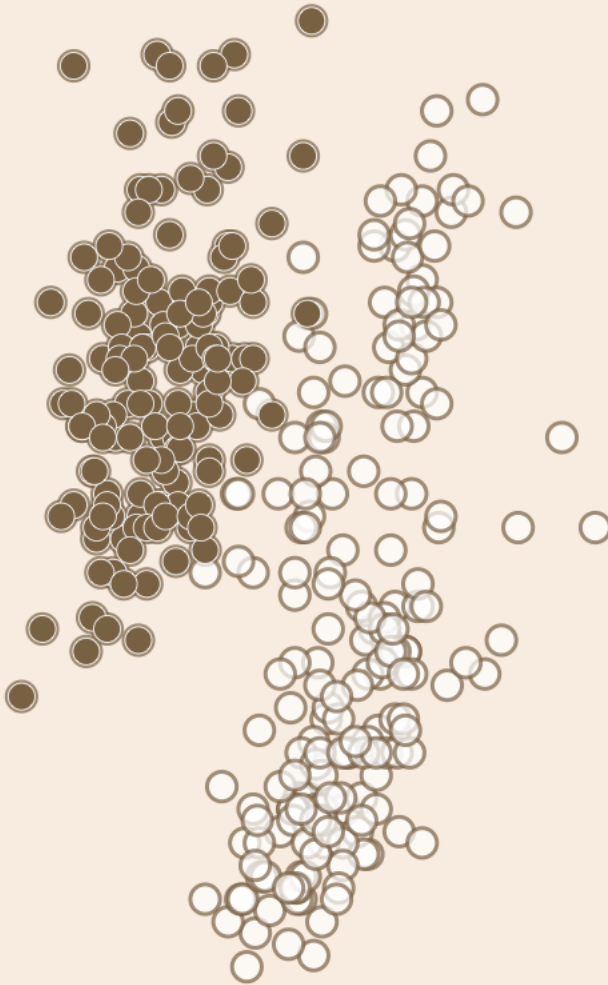
```
geom_point(  
  data = . %>% select(-species),  
  size = 3, shape = 21, stroke = .7,  
  color = "burlywood4", fill= "white",  
  alpha = .7  
)
```

The second layer, shows the category of interest, because the data it uses is the *globally declared data*, which contains the species faceting variable.

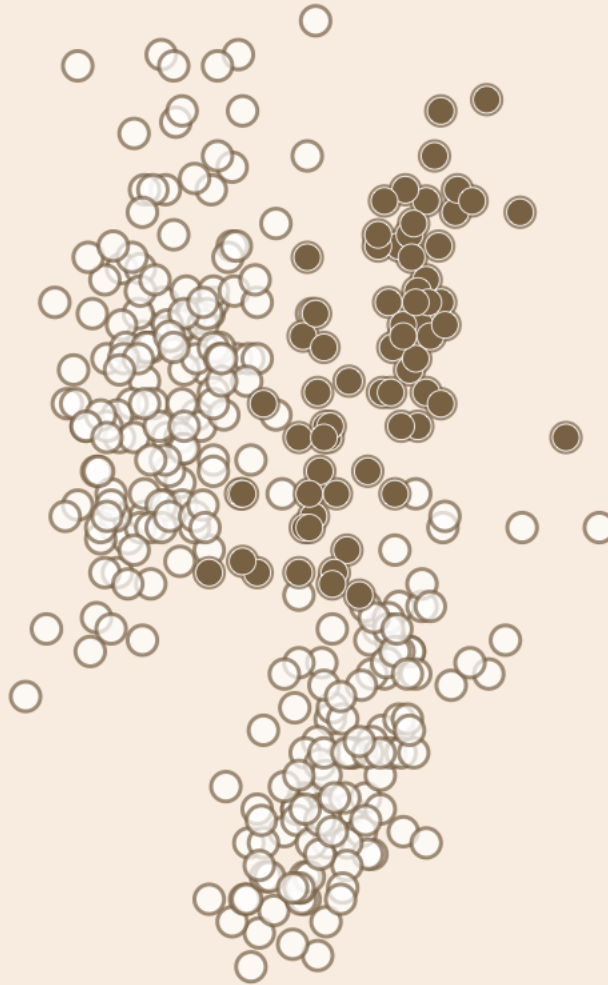
Colors selection draws our attention to this layer. A point with `shape = 21`, is a circle with a perimeter so a fill color and boarder color may be specified.

```
geom_point(  
  size = 3, shape = 21, stroke = .2,  
  color = "white", fill = "burlywood4"  
)
```

Adelie



Chinstrap



Gentoo

