

Tarea - Subplots con Facets

Jesus Mudarra Luján

2022-09-05

Ejercicio 1

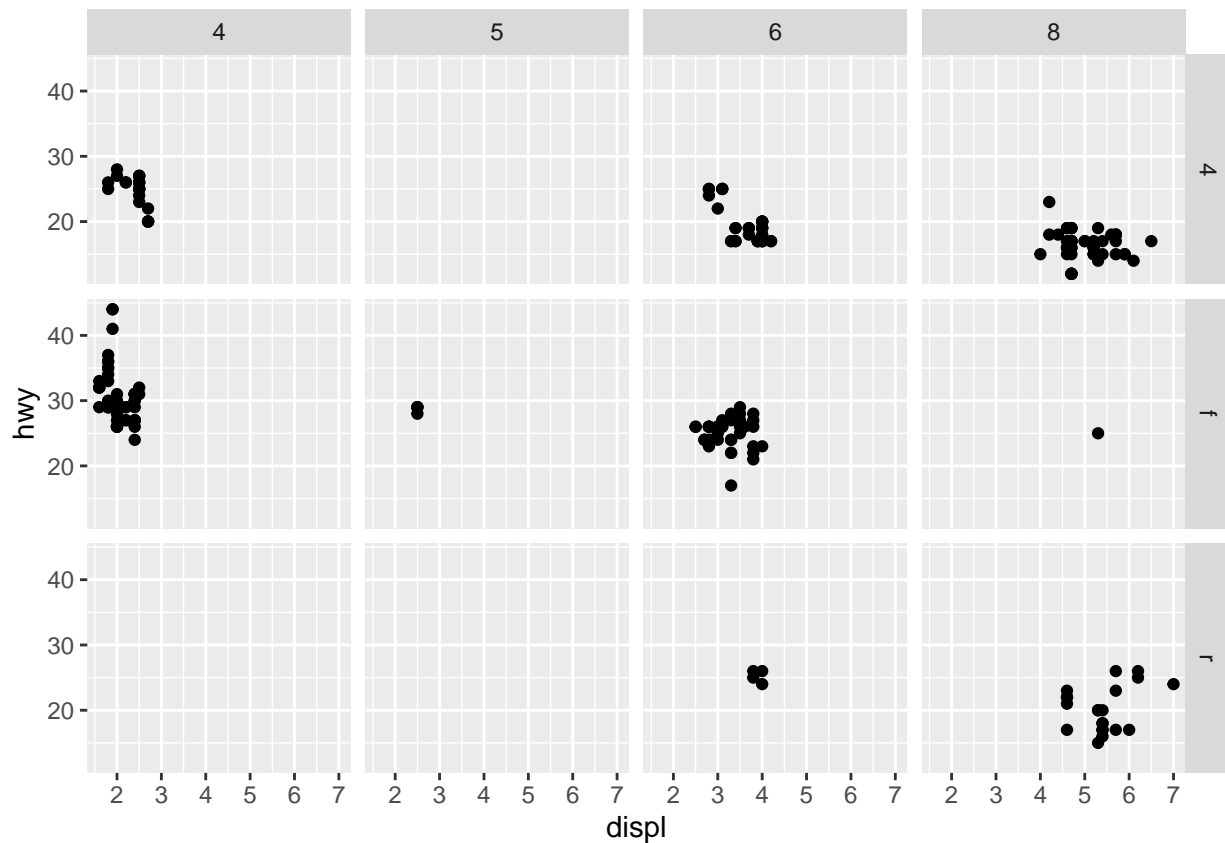
¿Qué ocurre si hacemos un facet de una variable continua?

Que para cada uno de los valores de la columna nos generará un gráfico.

Ejercicio 2

¿Qué significa si alguna celda queda vacía en el gráfico `facet_grid(drv~cyl)`?

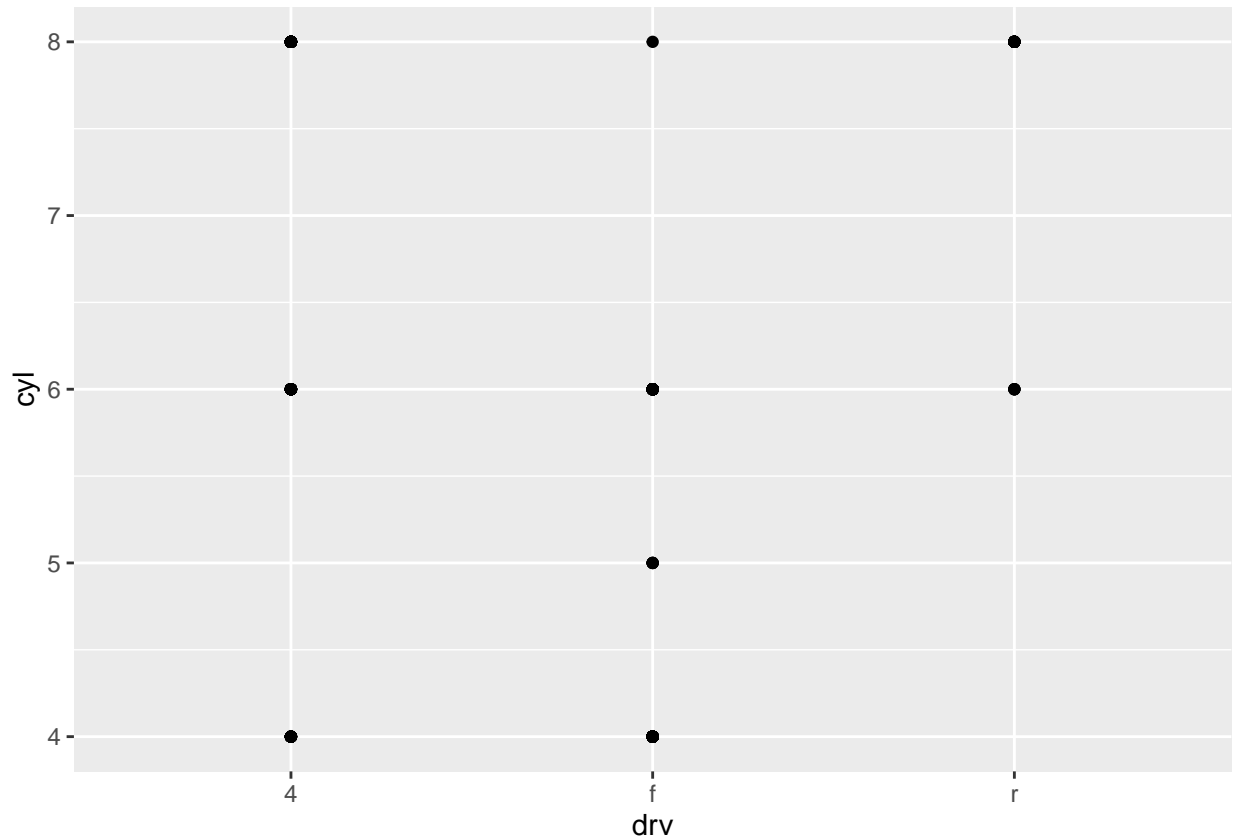
```
ggplot(data = mpg) +  
  geom_point(mapping = aes(x = displ, y = hwy)) +  
  facet_grid(drv ~ cyl)
```



Como se puede ver en el ejemplo anterior, significa que no existe ningún modelo de coche que tenga 5 cilindros y tenga tracción a las 4 ruedas.

¿Qué relación guardan esos huecos vacíos con el gráfico siguiente?

```
ggplot(data = mpg) +  
  geom_point(mapping = aes(x=drv, y = cyl))
```

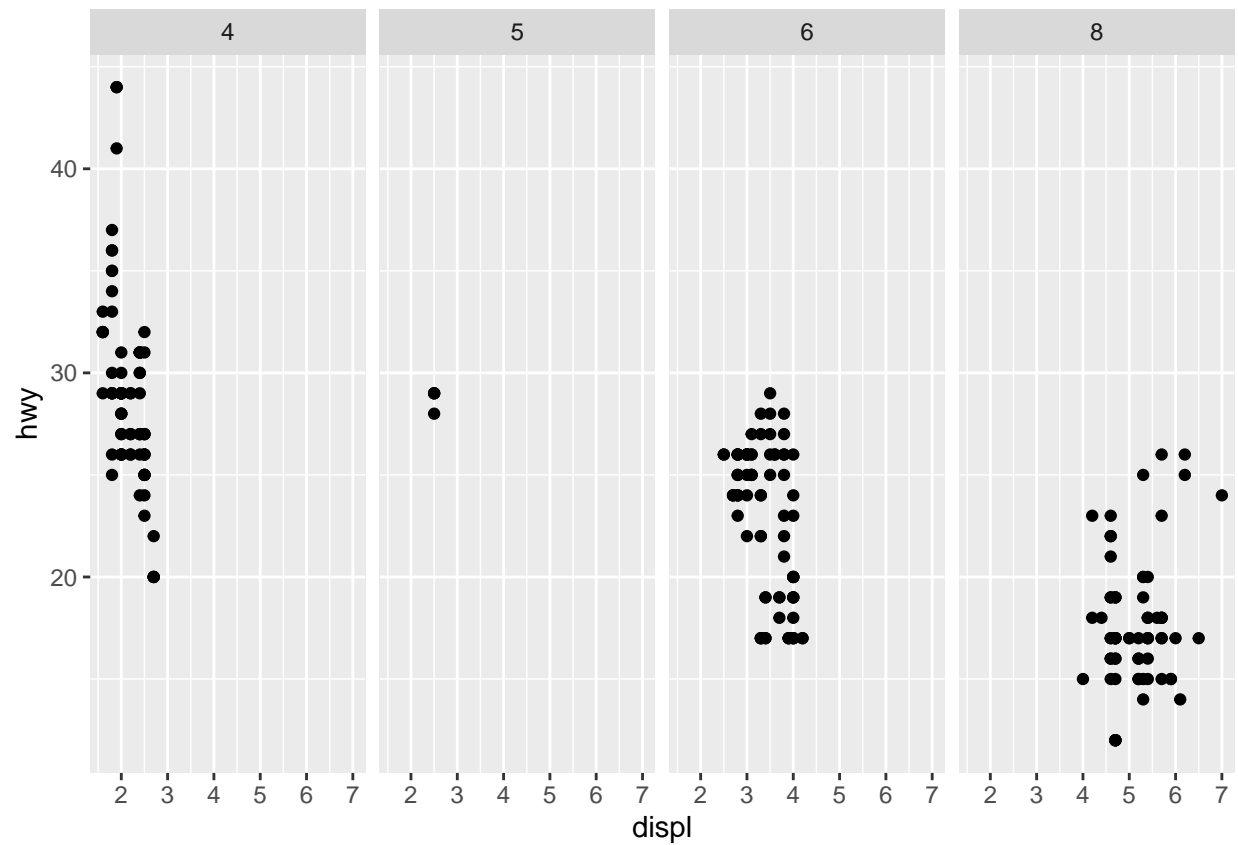


No hay coches de 7 cilindros. No hay coches de 4 cilindros que tenga tracción trasera. No hay coches de 5 cilindros que tenga tracción trasera ni a las 4 ruedas.

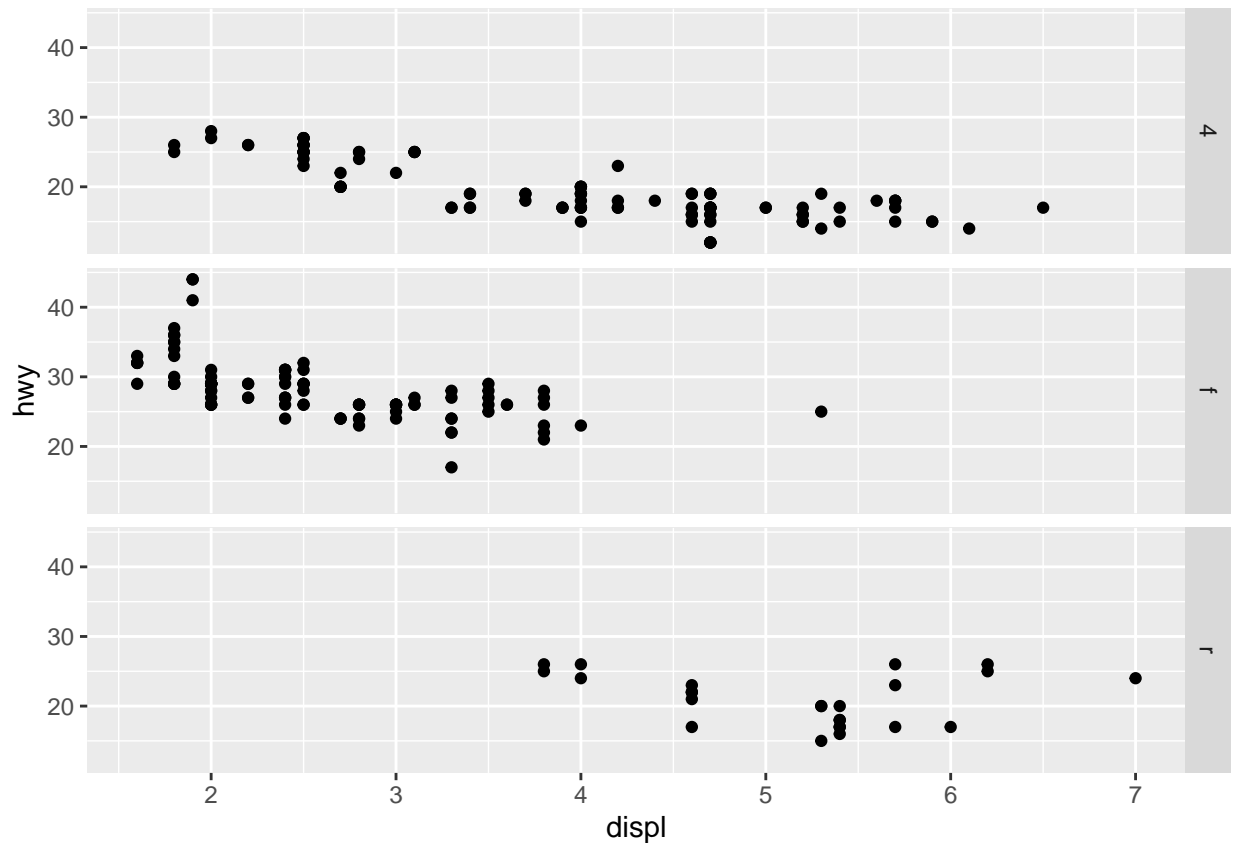
Ejercicio 3

¿Qué gráficos generan las siguientes dos instrucciones? ¿Qué hace el punto? ¿Qué diferencias hay de escribir la variable antes o después de la virgullita?

```
ggplot(data = mpg) +  
  geom_point(mapping = aes(x=displ, y = hwy)) +  
  facet_grid(.~cyl)
```



```
ggplot(data = mpg) +  
  geom_point(mapping = aes(x=displ, y = hwy)) +  
  facet_grid(drv~.)
```



Los gráficos representan la relación de los cilindros del motor con las millas de galón en carretera para cada tipo de tracción.

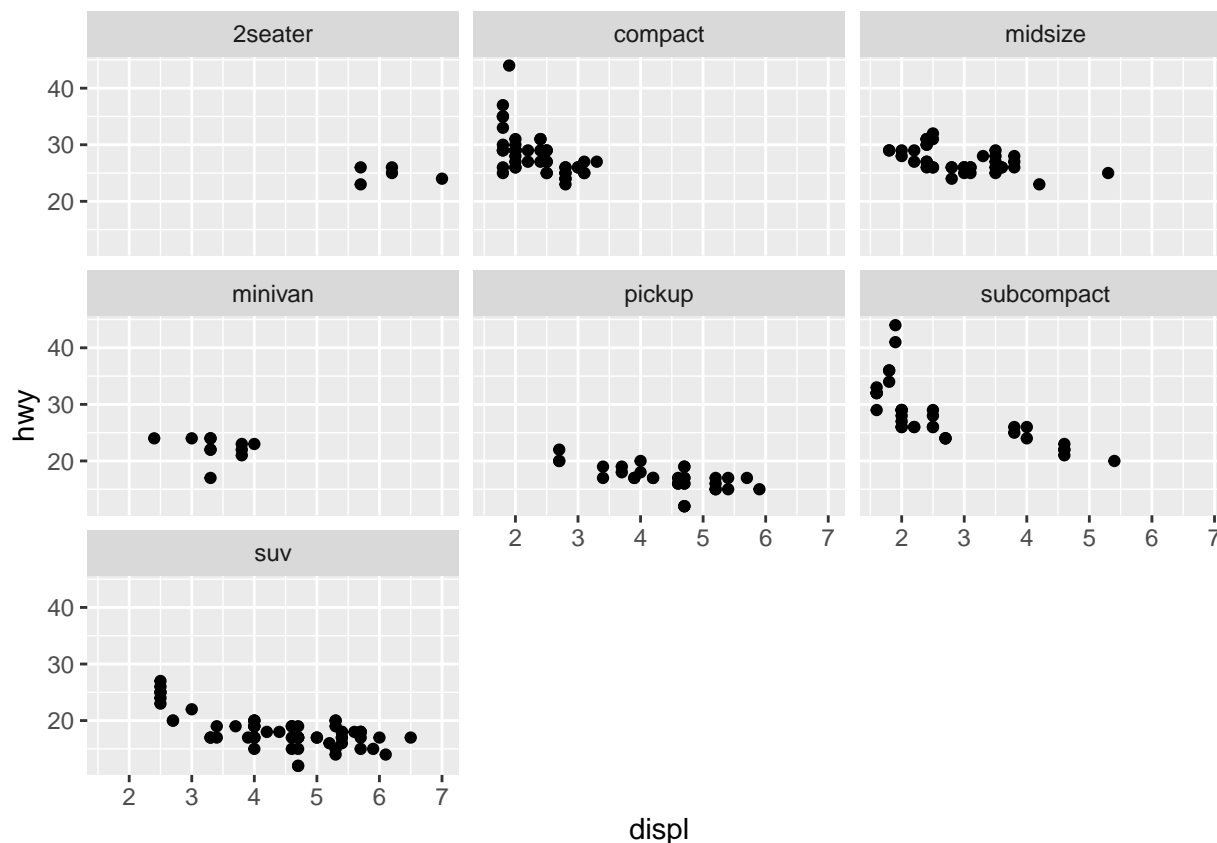
Utilizando el punto sería lo mismo que utilizar un `facet_wrap(~cyl)` o `facet_wrap(drv~)` ya que no se quiere hacer una comparación entre dos variables.

La diferencia entre escribir antes o después la tilde es la orientación de los subplots con la variable discreta.

Ejercicio 4

El primer facet que hemos pintado era el siguiente:

```
ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy)) +
  facet_wrap(~class, nrow = 3)
```



¿Qué ventajas crees que tiene usar facets en lugar de la estética del color? ¿Qué desventajas? ¿Qué cambiaría si tu dataset fuera mucho más grande?

- Ventajas: se puede visualizar por separado obteniendo gráfico con una cantidad menor de puntos.
- Desventajas: no queda toda la información representada en un mismo gráfico.
- Si el dataset fuera mucho más grande sería muy útil ya que utilizando la estética del color se terminarían solapando los puntos y sería muy complicado sacar conclusiones.

Ejercicio 5

Investiga la documentación de `?facet_wrap` y contesta a las siguientes preguntas:

- ¿Qué hace el parámetro `nrow`?

Indica el número de filas de subplots que se desea generar.

- ¿Y el parámetro `ncol`?

Indica el número de columnas de subplots que se desea generar.

- ¿Qué otras opciones sirven para controlar el layout de los paneles individuales?
 - scales: las escalas deben ser fijas (“fixed”, por defecto), libres (“free”) o libres en una dimensión (“free_x”, “free_y”)?

- shrink: si es TRUE, reducirá las escalas para ajustarse a la salida de las estadísticas, no a los datos brutos. Si es FALSE, será el rango de los datos brutos antes del resumen estadístico.
 - labeller: una función que toma un marco de datos de etiquetas y devuelve una lista o marco de datos de vectores de caracteres.
 - as.table: si es TRUE, el valor por defecto, las facetas se presentan como una tabla con los valores más altos en la parte inferior derecha. Si es FALSE, las facetas se presentan como un gráfico con el valor más alto en la parte superior derecha.
 - switch: por defecto, las etiquetas se muestran en la parte superior y derecha del gráfico. Si es “x”, las etiquetas superiores se mostrarán hacia abajo. Si es “y”, las etiquetas de la derecha se mostrarán a la izquierda. También puede establecerse en “ambos”.
 - drop: si es TRUE, el valor por defecto, todos los niveles de factor que no se utilicen en los datos se eliminarán automáticamente. Si es FALSE, se mostrarán todos los niveles de los factores, independientemente de que aparezcan o no en los datos.
 - dir: dirección: “h” para la horizontal, por defecto, o “v”, para la vertical.
 - strip.position: selecciona la posición donde colocar las etiquetas.
- ¿Por qué `facet_grid()` no tiene los parámetros de `nrow` ni de `ncol`?

Porque con `facet_grid()` se comparan las variables que se indican como variables por lo que no tendría sentido añadir `nrow` y `ncol`.

Ejercicio 6

Razona la siguiente afirmación:

Cuando representemos un facet con `facet_grid()` conviene poner la variable con más niveles únicos en las columnas.

Para que evitar que los subplots salgan más anchos de lo que deberían.