

Actividad 1 Solución

Ejercicio 1

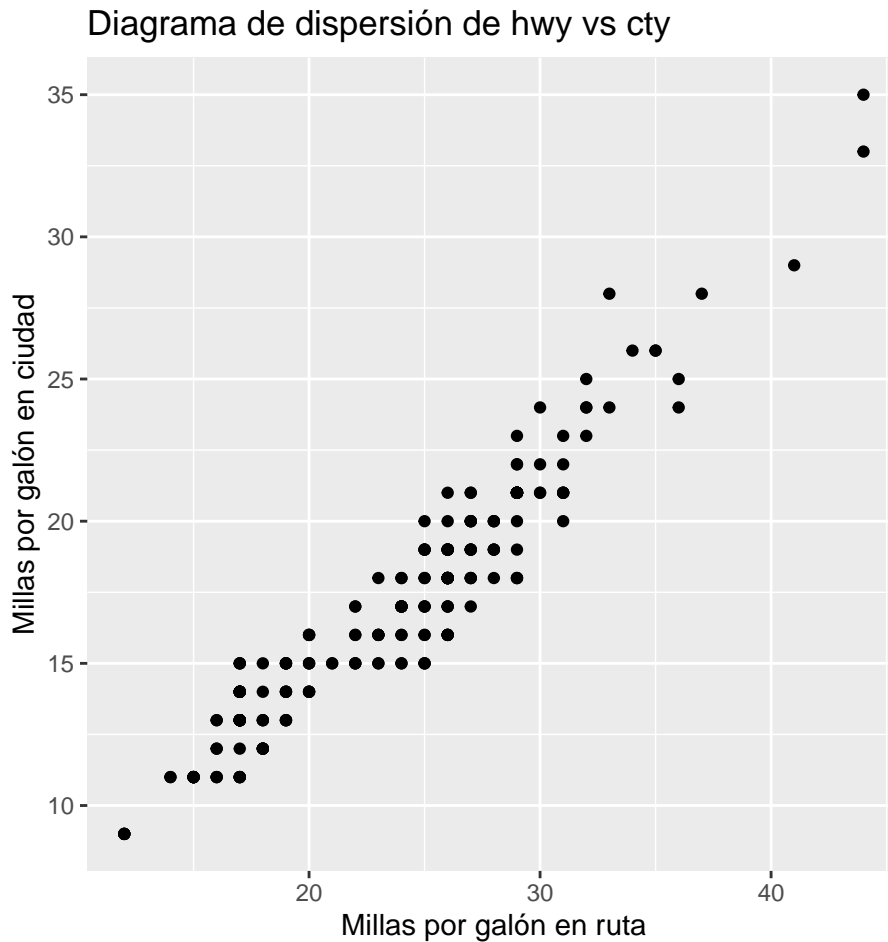
Con los datos `mpg` que se encuentran disponible en `ggplot2` hacer un gráfico con las siguientes características:

- Un diagrama de dispersión de `hwy` vs `cty`
- Incluir usando `labs` el nombre de los ejes, título informativo y nombre de la leyenda

Usar `?mpg` para ver la descripción de las variables.

```
library(ggplot2)

ggplot(data = mpg, aes(x = hwy, y = cty)) +
  geom_point() +
  theme(aspect.ratio = 1) +
  labs(x = "Millas por galón en ruta",
       y = "Millas por galón en ciudad",
       title = "Diagrama de dispersión de hwy vs cty")
```

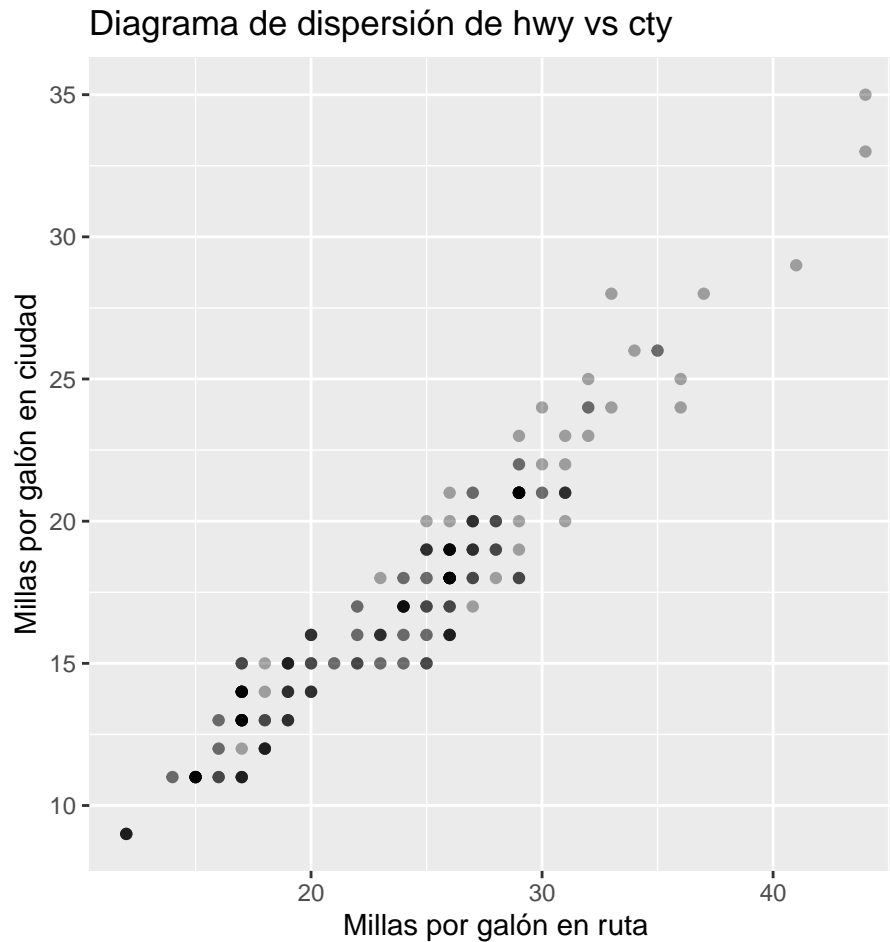


Ejercicio 2

En el gráfico anterior hay un problema de sobre ploteo, esto es hay más de una observación en algunas combinaciones de cty y hwy.

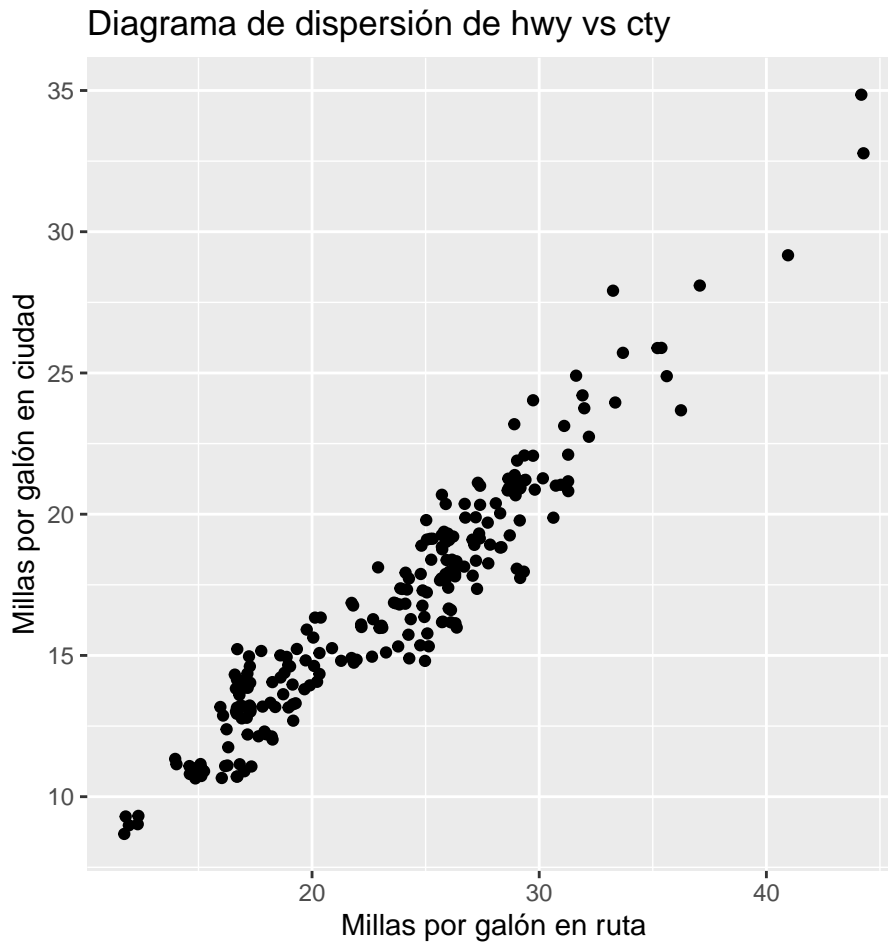
1. Hacer el gráfico anterior y usá transparencia (alpha) para ver el problema de sobre ploteo.

```
ggplot(data = mpg, aes(x = hwy, y = cty)) +  
  geom_point(alpha = 1/3) +  
  theme(aspect.ratio = 1) +  
  labs(x = "Millas por galón en ruta",  
       y = "Millas por galón en ciudad",  
       title = "Diagrama de dispersión de hwy vs cty")
```



2. Hacer el gráfico anterior y usá `geom_jitter()` para ver el problema de sobre ploteo.

```
ggplot(data = mpg, aes(x = hwy, y = cty)) +
  geom_jitter() +
  theme(aspect.ratio = 1) +
  labs(x = "Millas por galón en ruta",
       y = "Millas por galón en ciudad",
       title = "Diagrama de dispersión de hwy vs cty")
```



3. ¿Qué opción te resulta más apropiada y por qué ?

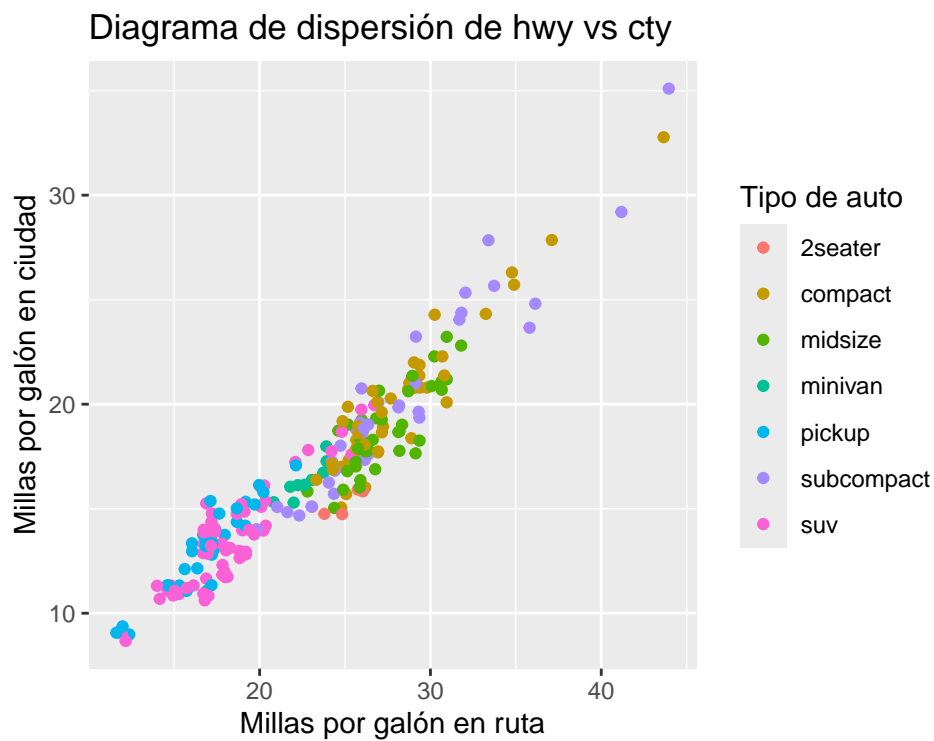
Usar transparencia en este caso es más apropiado ya que las variables son continuas y `geom_jitter()` cambia, aunque poco, sus valores originales. En el caso que tenga una variable continua y otra categórica se podría usar solamente una dirección para distorsionar con los argumentos `width` ó `height`, conviene distorsionar el eje que tiene la variable categórica para no cambiar la escala de la variable continua.

Ejercicio 3

Usá como base el gráfico del Ejercicio 2 con `geom_jitter`.

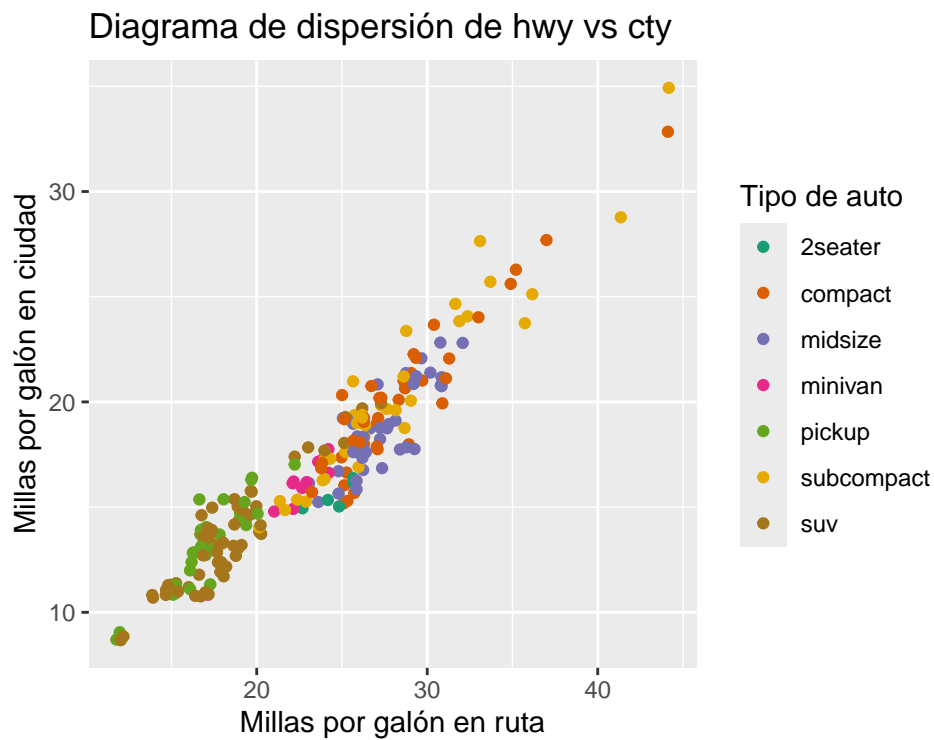
1. Coloreá usando la variable `class`

```
ggplot(data = mpg, aes(x = hwy, y = cty, color = class)) +
  geom_jitter() +
  theme(aspect.ratio = 1) +
  labs(
    x = "Millas por galón en ruta",
    y = "Millas por galón en ciudad",
    title = "Diagrama de dispersión de hwy vs cty",
    colour = "Tipo de auto"
  )
```



- Usá el gráfico anterior pero cambiando la paleta de colores a **Dark2**. Chequear la ayuda de `?scale_colour_brewer`. Para mirar otras posibles paletas visitar: <http://colorbrewer2.org/#type=sequential&scheme=BuGn&n=3>

```
ggplot(data = mpg, aes(x = hwy, y = cty, color = class)) +
  geom_jitter() +
  theme(aspect.ratio = 1) +
  labs(
    x = "Millas por galón en ruta",
    y = "Millas por galón en ciudad",
    title = "Diagrama de dispersión de hwy vs cty",
    colour = "Tipo de auto"
  ) +
  scale_colour_brewer(palette = "Dark2")
```



3. Usá la variable `class` para crear un gráfico para cada subconjunto, mirá `?facet_wrap`.

```
ggplot(data = mpg, aes(x = hwy, y = cty)) +
  geom_jitter() +
  facet_wrap(~class) +
  labs(
    x = "Millas por galón en ruta",
    y = "Millas por galón en ciudad",
    title = "Diagrama de dispersión de hwy vs cty",
    colour = "Tipo de auto"
  )
)
```

