Actividad Individual 6

Ángela Vieyto 5.487.839-8

Entrega 10/5/2021

Ejercicio 1

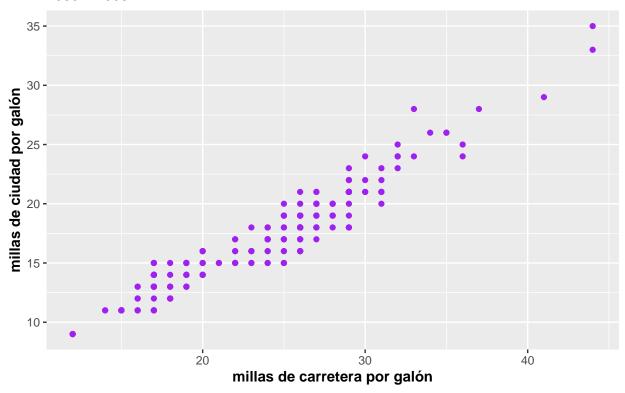
Con los datos mpg que se encuentran disponible en ggplot2 hacer un gráfico con las siguientes características:

- 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(mpg, aes(hwy, cty)) + geom_point(color = "purple") +
    ggtitle("Rendimiento del combustible en millas por galón") +
    labs(x = "millas de carretera por galón", y = "millas de ciudad por galón", subtitle = "1999 - 2008")
    theme (plot.title = element_text(size = rel(1.1), face = "bold")) +
    theme(axis.title = element_text(face = "bold"))
```

1999 - 2008



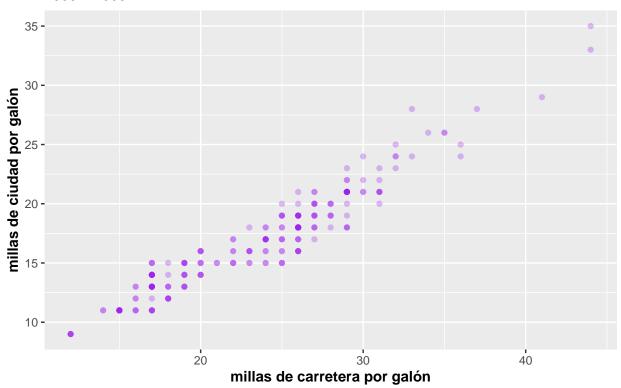
Ejercicio 2

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(mpg, aes(hwy, cty)) + geom_point(color = "purple", alpha = 0.3) +
    ggtitle("Rendimiento del combustible en millas por galón") +
    labs(x = "millas de carretera por galón", y = "millas de ciudad por galón", subtitle = "1999 - 2008")
    theme (plot.title = element_text(size = rel(1.1), face = "bold")) +
    theme(axis.title = element_text(face = "bold"))
```

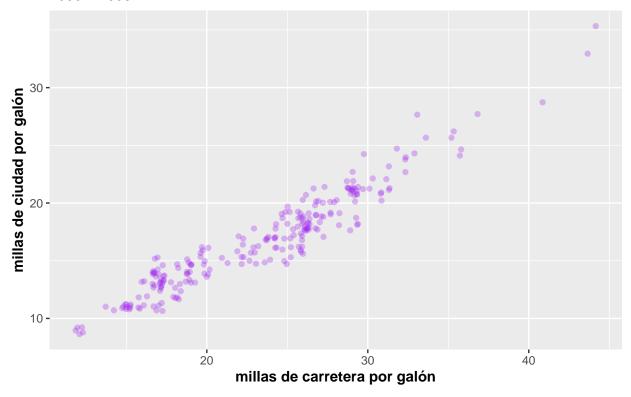
1999 - 2008



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

```
ggplot(mpg, aes(hwy, cty)) + geom_jitter(color = "purple", alpha = 0.3) +
    ggtitle("Rendimiento del combustible en millas por galón") +
    labs(x = "millas de carretera por galón", y = "millas de ciudad por galón", subtitle = "1999 - 2008")
    theme (plot.title = element_text(size = rel(1.1), face = "bold")) +
    theme(axis.title = element_text(face = "bold"))
```

1999 - 2008



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

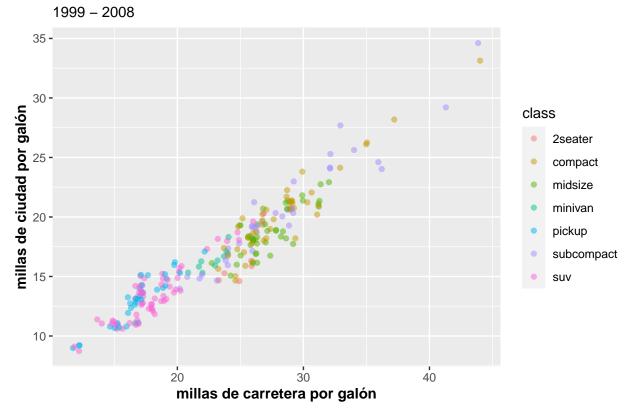
La opción <code>geom_jitter</code> facilita la visualización ya que atenúa el sobre ploteo generado por la gran cantidad de datos existentes. La forma en que lo hace es añadiendo ruido aleatorio a cada punto del diagrama, de forma de dispersarlos entre sí.

Ejercicio 3

Usá como base el gráfico del ejercicio 2 con 'geom_jitter.

1. Coloreá usando la variable class

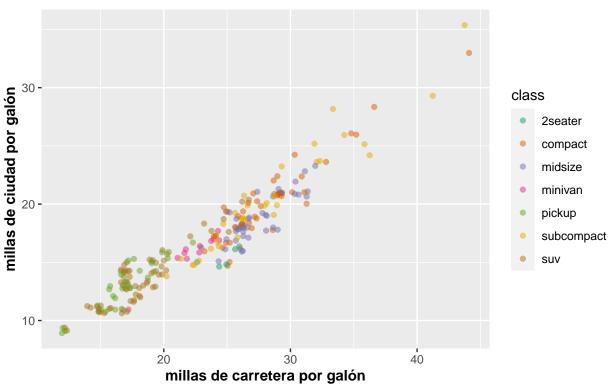
```
ggplot(mpg, aes(hwy, cty, color = class)) + geom_jitter(alpha = 0.5) +
    ggtitle("Rendimiento del combustible en millas por galón") +
    labs(x = "millas de carretera por galón", y = "millas de ciudad por galón", subtitle = "1999 - 2008")
    theme (plot.title = element_text(size = rel(1.1), face = "bold")) +
    theme(axis.title = element_text(face = "bold"))
```



2. 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(mpg, aes(hwy, cty, color = class)) + geom_jitter(alpha = 0.5) +
    scale_colour_brewer(palette = "Dark2") +
    ggtitle("Rendimiento del combustible en millas por galón") +
    labs(x = "millas de carretera por galón", y = "millas de ciudad por galón", subtitle = "1999 - 2008")
    theme (plot.title = element_text(size = rel(1.1), face = "bold")) +
    theme(axis.title = element_text(face = "bold"))
```

1999 - 2008



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

```
ggplot(mpg, aes(hwy, cty, color = class)) + geom_jitter(alpha = 0.5) +
  facet_wrap(vars(class)) +
  scale_colour_brewer(palette = "Dark2") +
  ggtitle("Rendimiento del combustible en millas por galón") +
  labs(x = "millas de carretera por galón", y = "millas de ciudad por galón", subtitle = "1999 - 2008")
  theme (plot.title = element_text(size = rel(1.1), face = "bold")) +
  theme(axis.title = element_text(face = "bold"))
```

1999 – 2008

