Near publication ready figures

Princy Parsana

PhD candidate, Computer Science @JHU

Battle Lab

ggplot: elegant visualization using grammar of graphics

- Author: Hadley Wickham

- Visualization package based on Lee Wilkinson's **G**rammar of **G**raphics

- Syntax allows one to think of graphs as a set of rules

- Flexible set of commands to create many small pieces of the graph

  - Legend, axes, colors, outline, etc.

- Two ways:

  - qplot

  - **ggplot**

Let's install ggplot2

```
install.packages("ggplot2")
install.packages("cowplot")
theme_set(theme_cowplot(font_size=12))
```

Fuel economy dataset

`head(mpg)`

manufacturer

model   model name

displ   engine displacement, in litres

year   year of manufacture

cyl   number of cylinders

trans   type of transmission

drv   f = front-wheel drive, r = rear wheel drive, 4 = 4wd

cty   city miles per gallon

hwy   highway miles per gallon

fl   fuel type

class   "type" of car

Fuel economy dataset

`head(mpg)`

manufacturer <chr>	model <chr>	displ <dbl>	year <int>	cyl <fctr>
audi	a4	1.8	1999	4
audi	a4	1.8	1999	4
audi	a4	2.0	2008	4
audi	a4	2.0	2008	4
audi	a4	2.8	1999	6
audi	a4	2.8	1999	6

manufacturer

model   model name

displ   engine displacement, in litres

year   year of manufacture

cyl   number of cylinders

trans   type of transmission

drv   f = front-wheel drive, r = rear wheel drive, 4 = 4wd

cty   city miles per gallon

hwy   highway miles per gallon

fl   fuel type
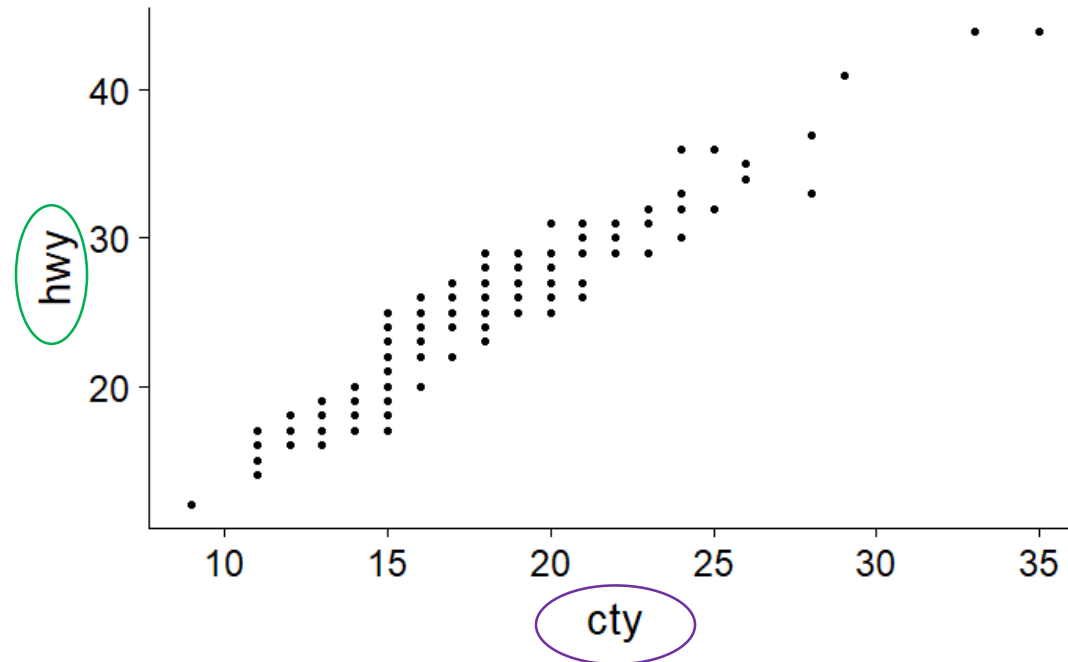
class   "type" of car

Simple scatter plot

```
plot_a <- ggplot(data = mpg, aes(x = cty, y = hwy)) +
    geom_point()
plot(plot_a)
```
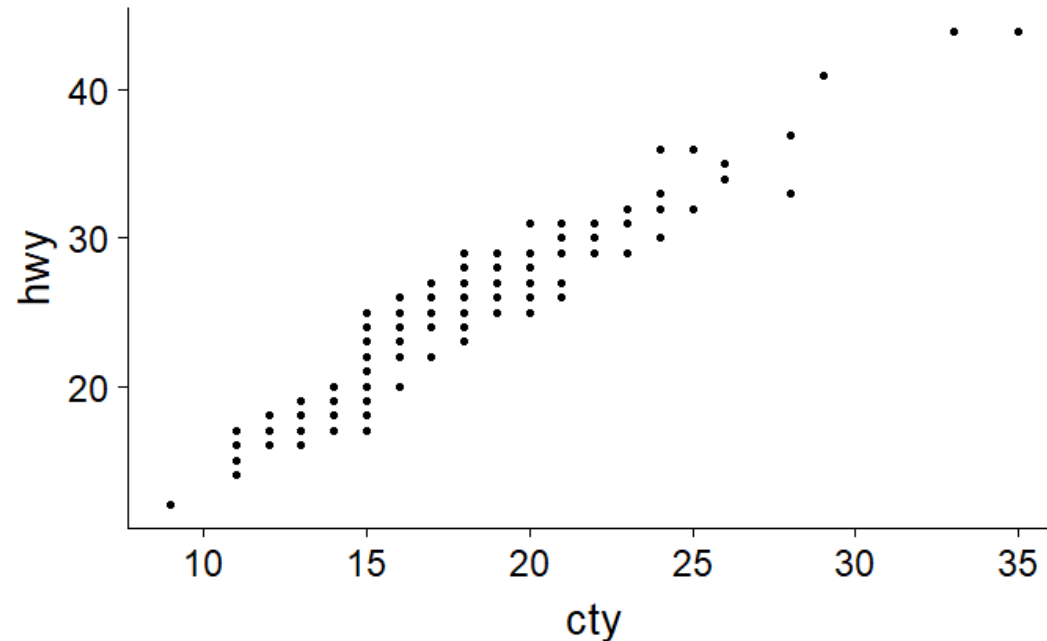
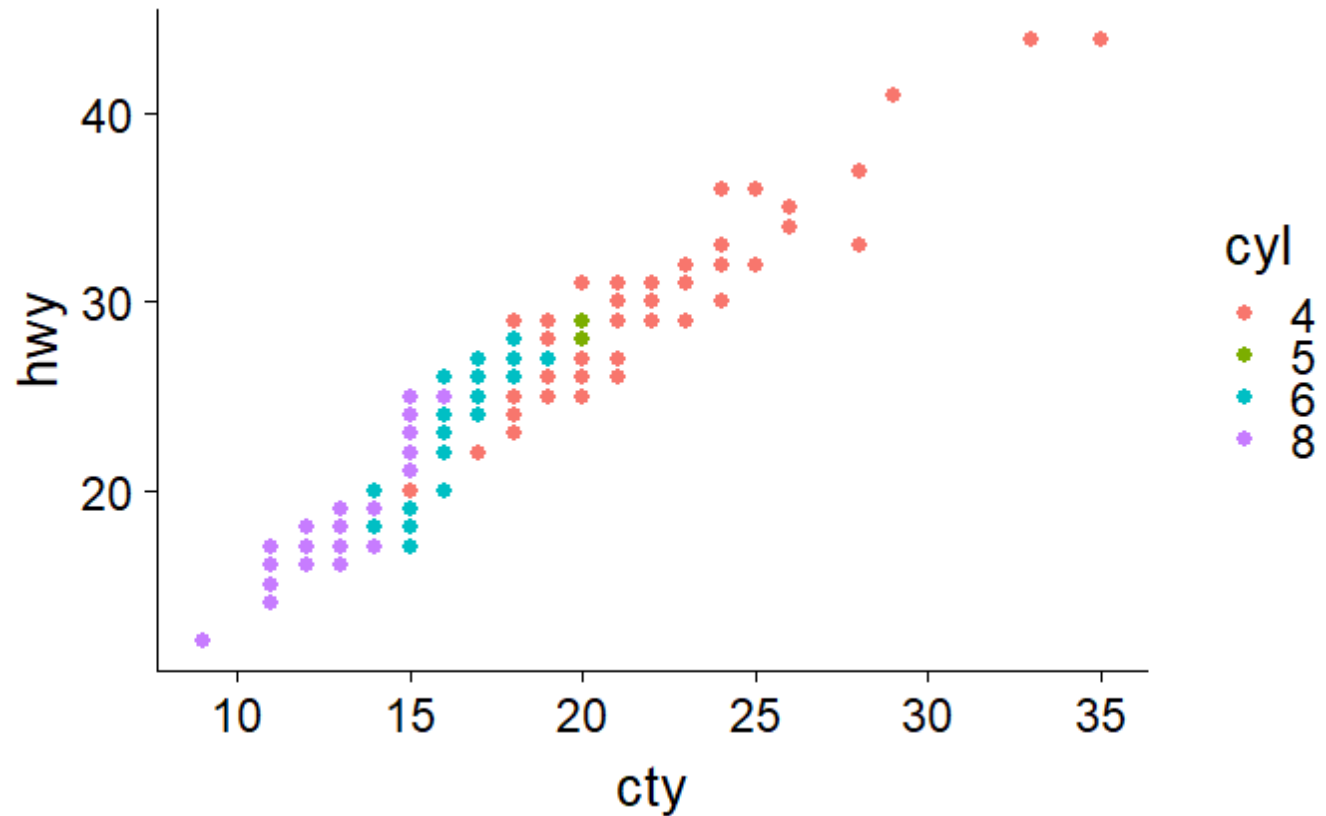# Aesthetics: visual characteristics to describe data

```
plot_a <- ggplot(data = mpg, aes(x = cty, y = hwy)) +
    geom_point()
plot(plot_a)
```
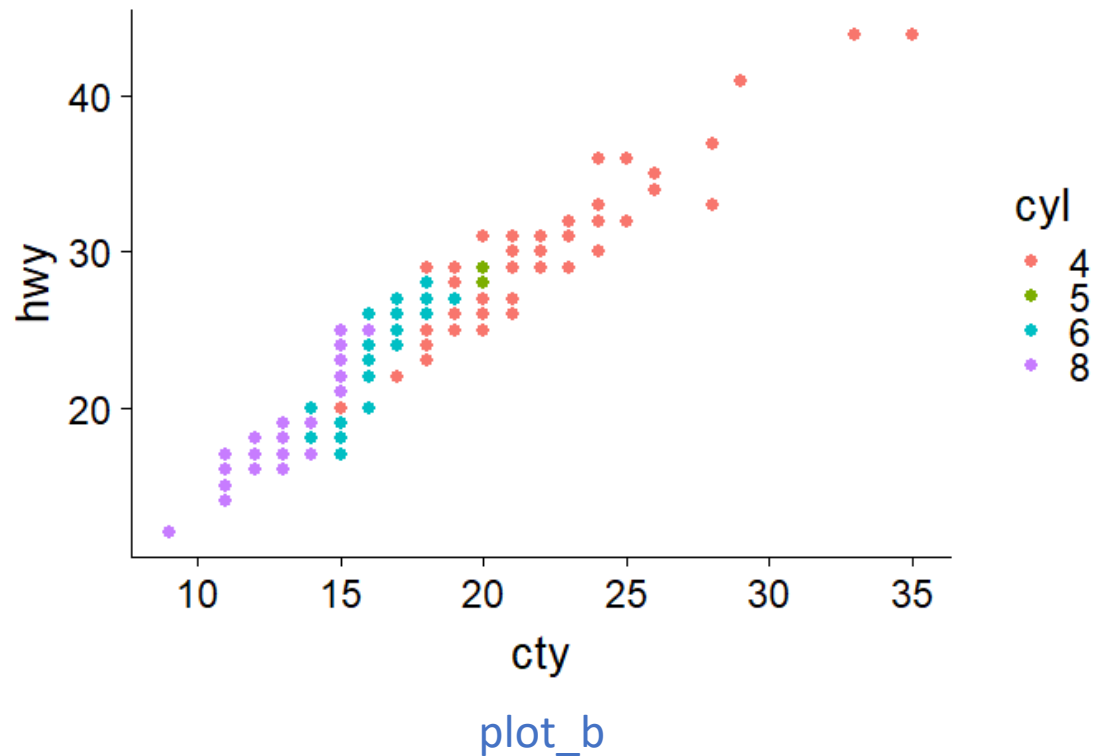
X and Y axes

```
plot_b = ggplot(mpg, aes(x = cty, y = hwy, col =
class)) +
  geom_point(size = 2.5)
plot(plot_b)
```
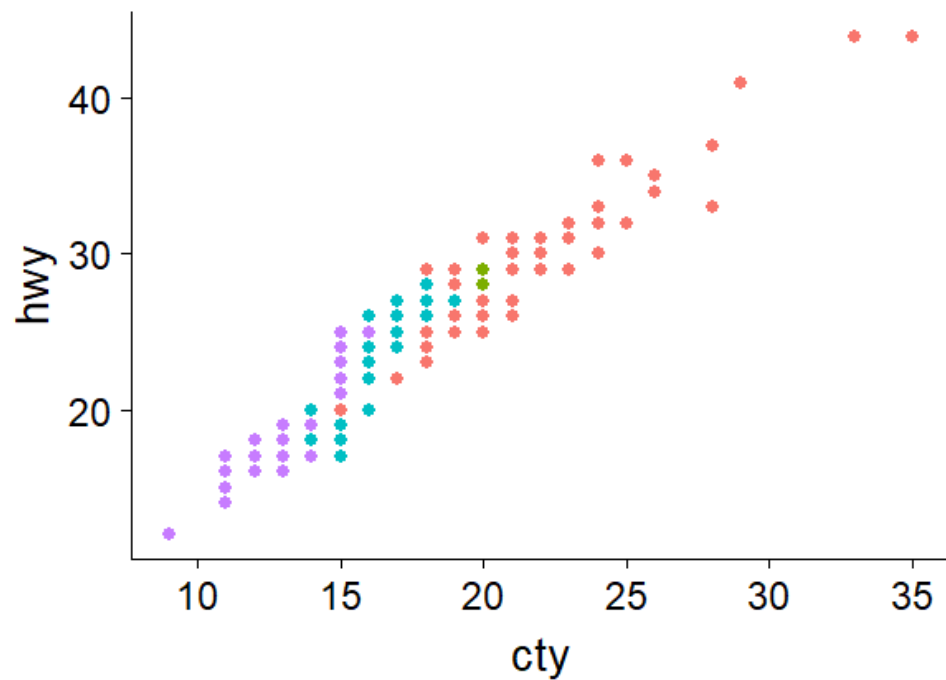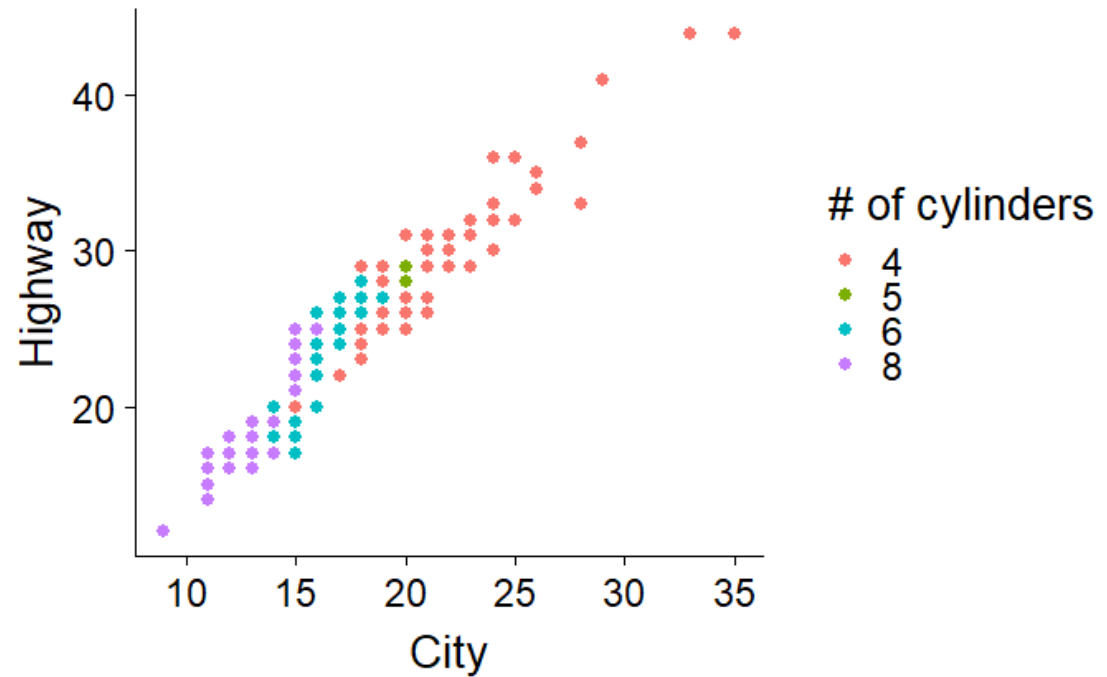
Color points using
aesthetic inputs

```
plot_c <- ggplot(mpg, aes(x = cty, y = hwy, col = class)) +
    geom_point(size = 2.5) +
    labs(color = "# of cylinders", x = "City", y = "Highway")
plot(plot_c)
```
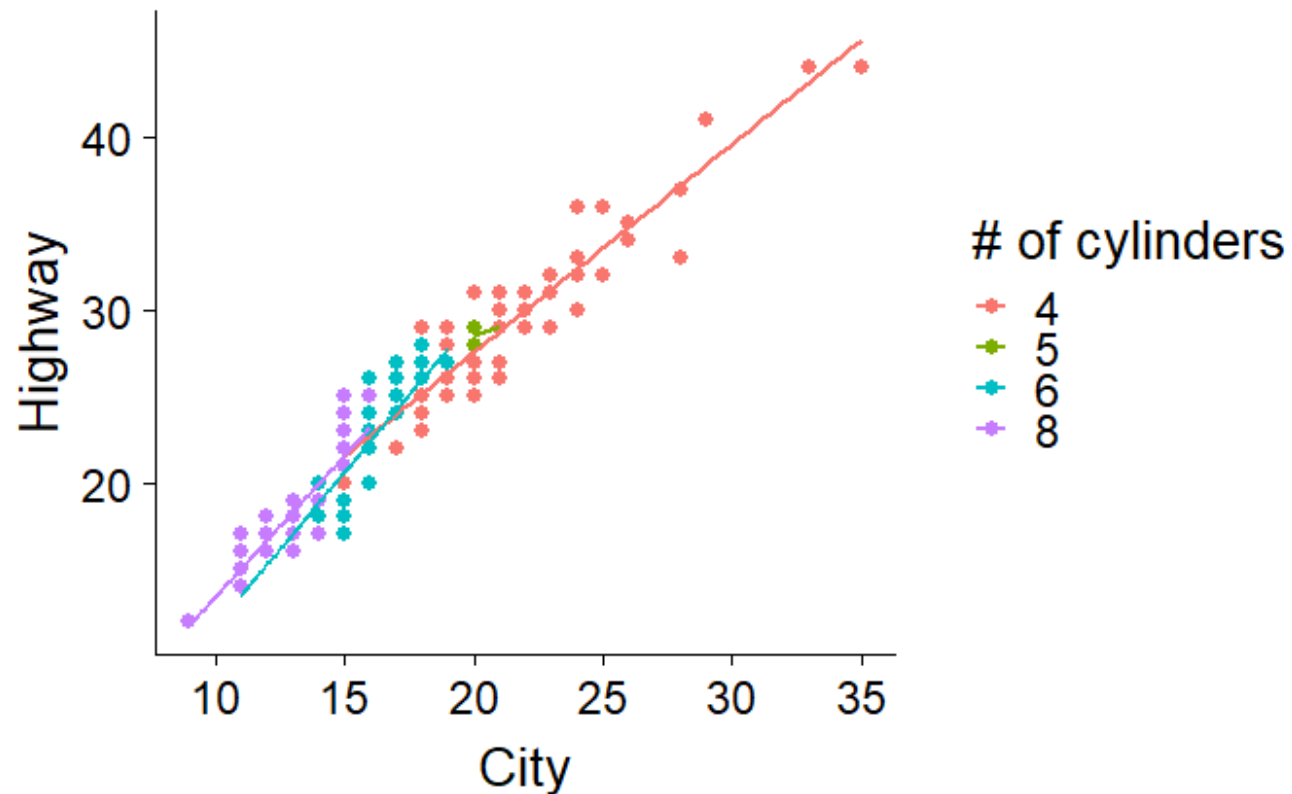

plot_b

```
plot_c <- ggplot(mpg, aes(x = cty, y = hwy, col = class)) +
    geom_point(size = 2.5) +
    labs(color = "# of cylinders", x = "City", y = "Highway")
plot(plot_c)
```


plot_b

plot_c

```
plot_d <- ggplot(mpg, aes(x = cty, y = hwy, col = cyl)) +
  geom_point(size = 2.5) + labs(color = "# of cylinders", x =
"City", y = "Highway") +
  geom_smooth(se = FALSE, method = "lm")

plot(plot_d)
```
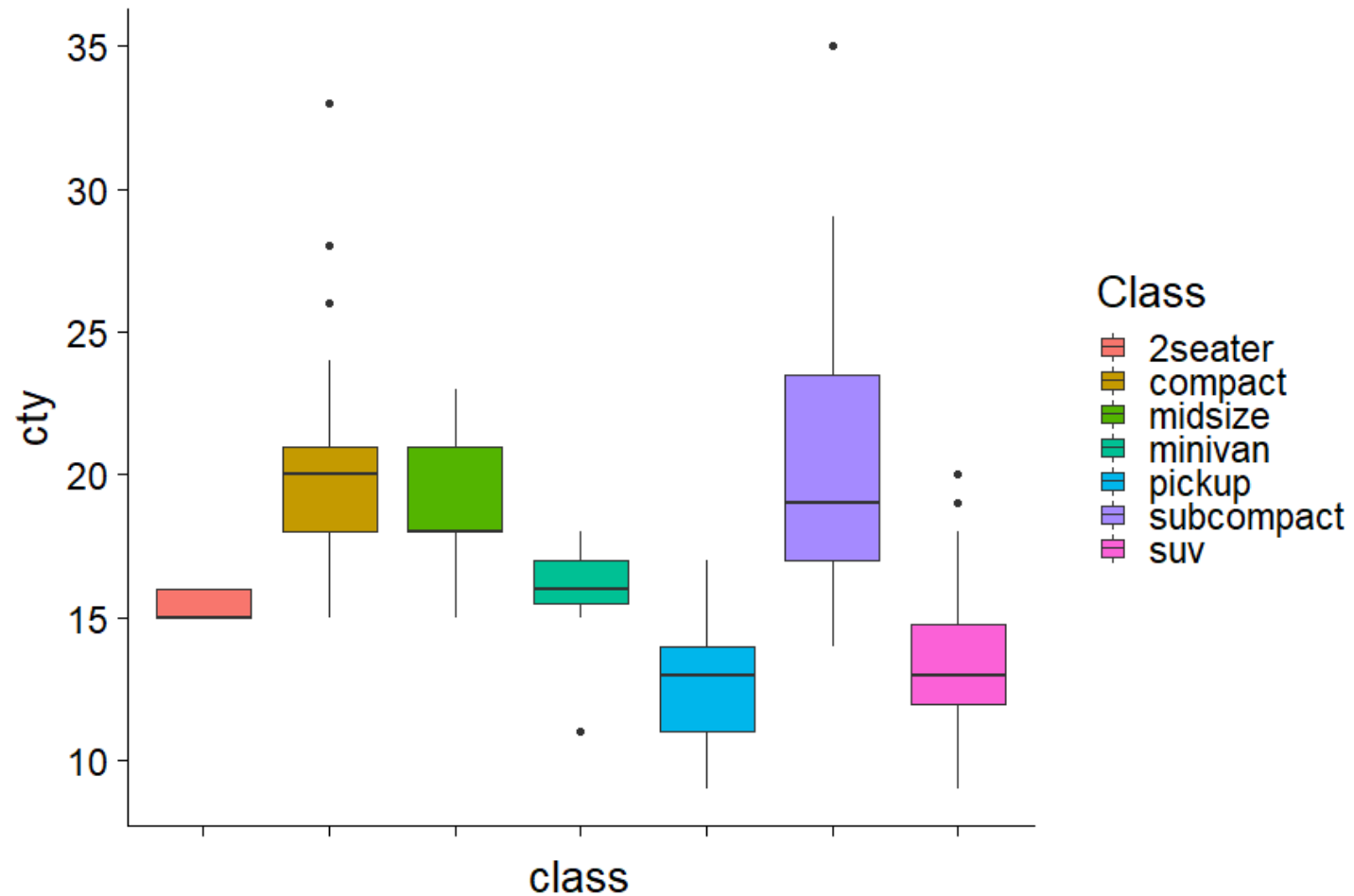
```
plot_e <- ggplot(mpg, aes(x = cty, y = hwy)) +
  geom_point(aes(col = class), size = 2.5) + labs(color = "Car
Type", x = "City", y = "Highway") +
  geom_smooth(se = FALSE, method = "lm")

plot(plot_e)
```
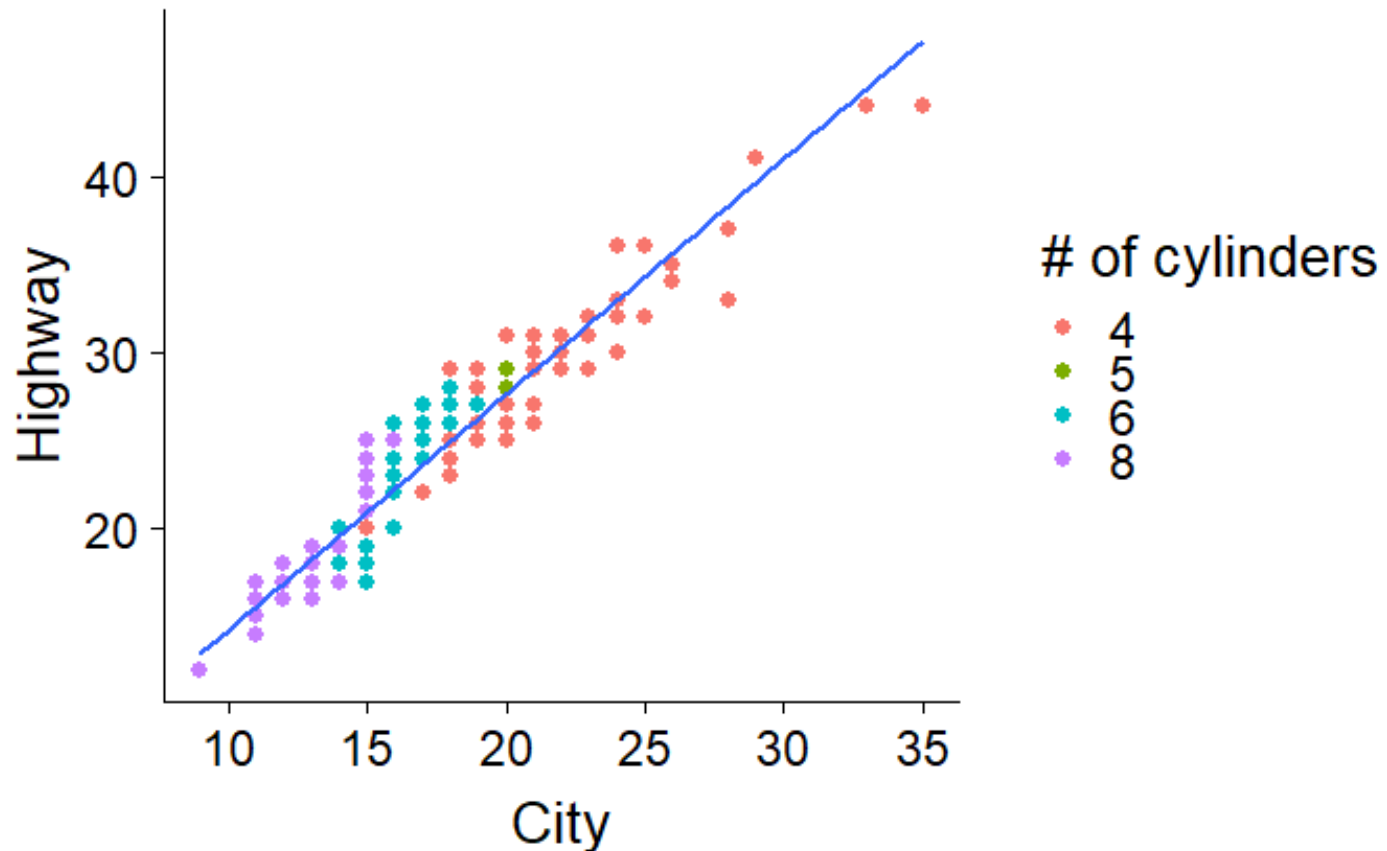
```
plot_f <- ggplot(mpg, aes(x = class, y = cty, fill = class))+
    geom_boxplot() +
    labs(fill = "Class")
plot(plot_f)
```
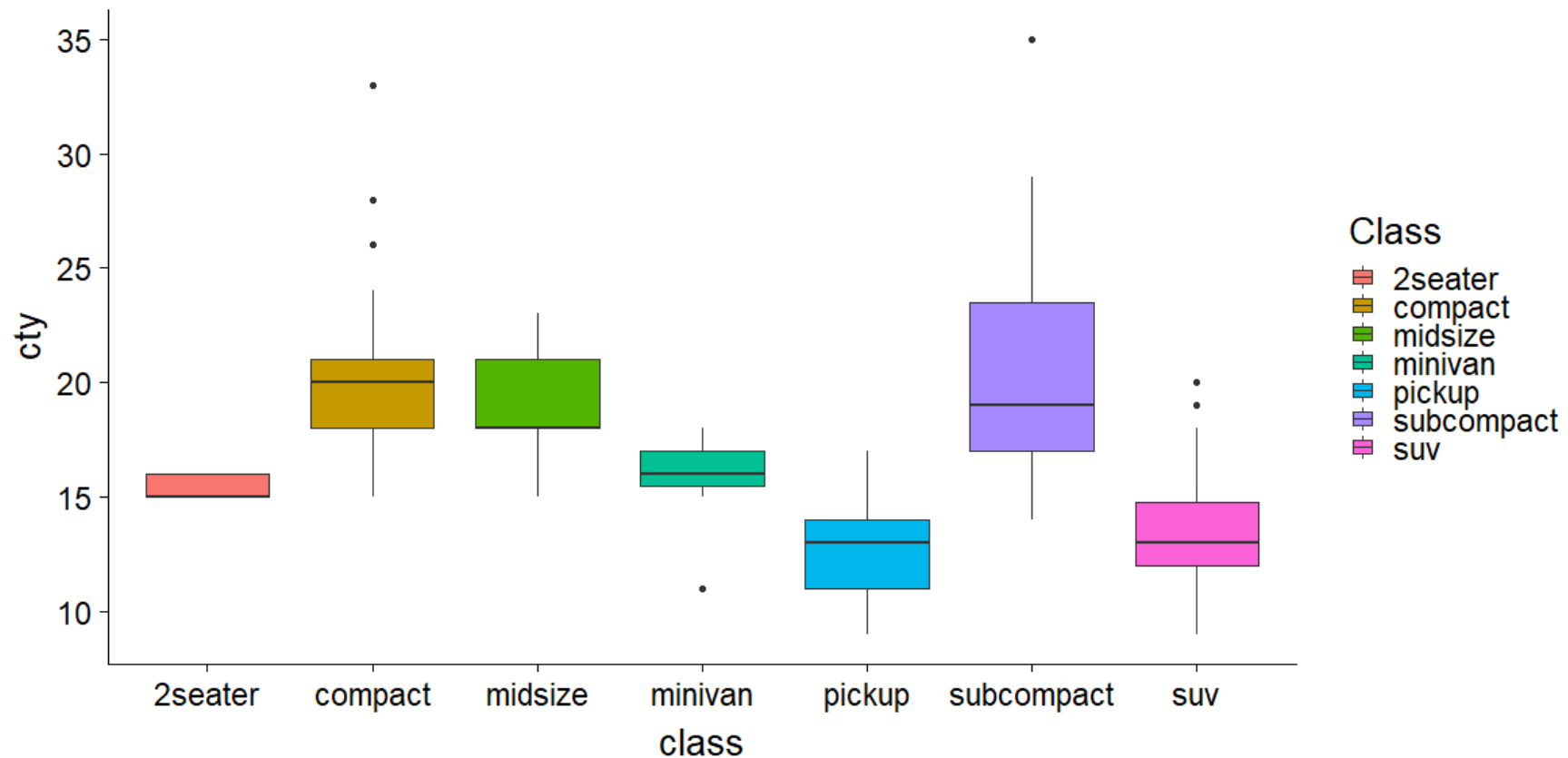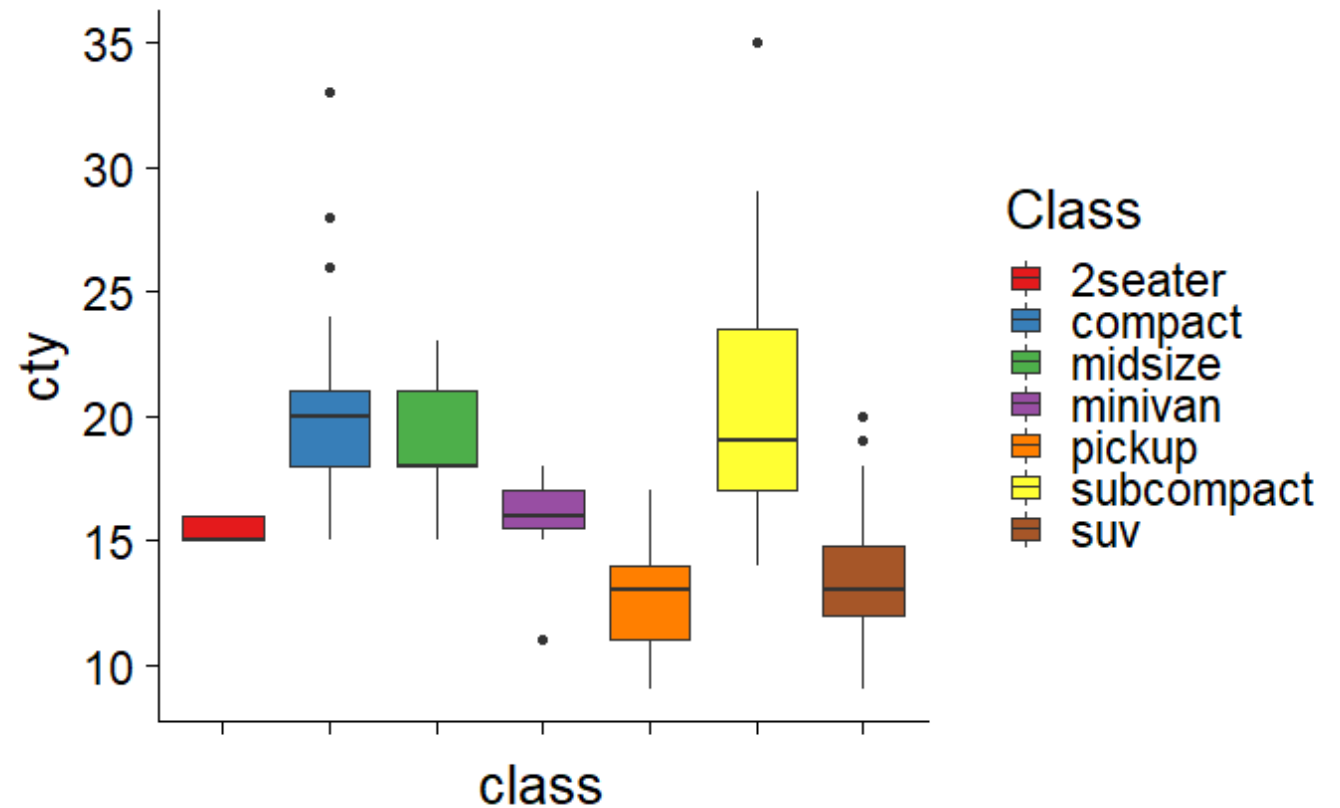
```
plot_g <- ggplot(mpg, aes(x = class, y = cty, fill = class))+
geom_boxplot() +
  labs(fill = "Class") + theme(axis.text.x = element_blank())
plot(plot_g)
```
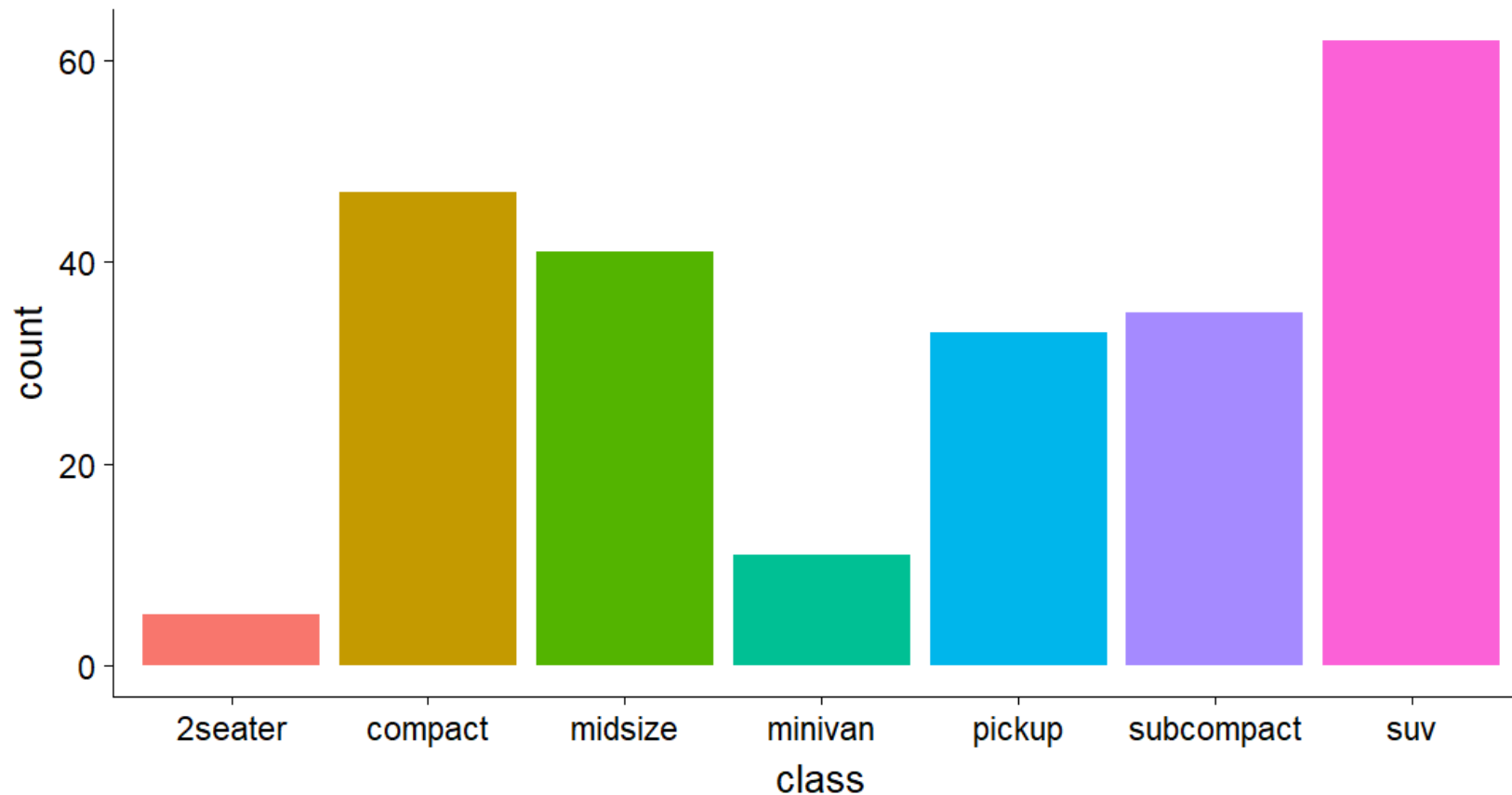
```
plot_h <- ggplot(mpg, aes(x = class, y = cty, fill = class))+
        geom_boxplot() + labs(fill = "Class") +
        scale_fill_brewer(palette="Set1") +
        theme(axis.text.x = element_blank())
plot(plot_h)
```

While coloring is flexible, it's a bit confusing

```
plot_i <- ggplot(mpg, aes(x = class, fill = class)) +
  geom_bar() +
  theme(legend.position = "none")

plot(plot_i)
```
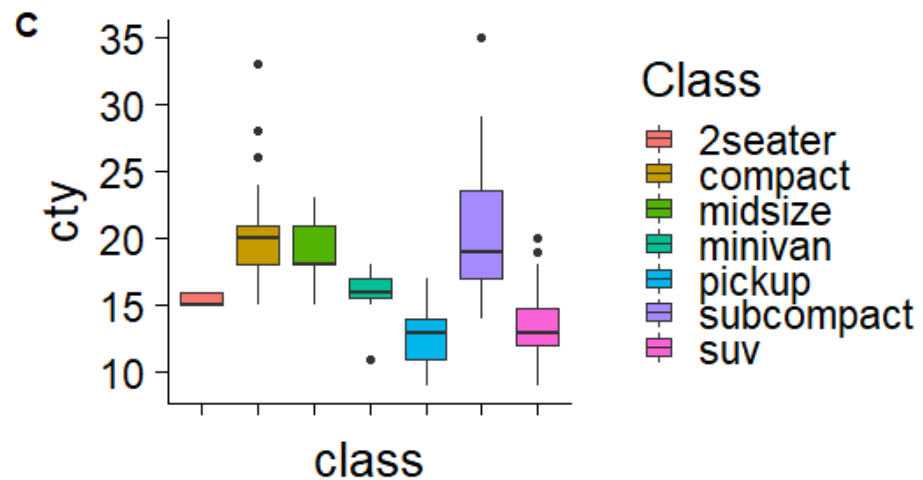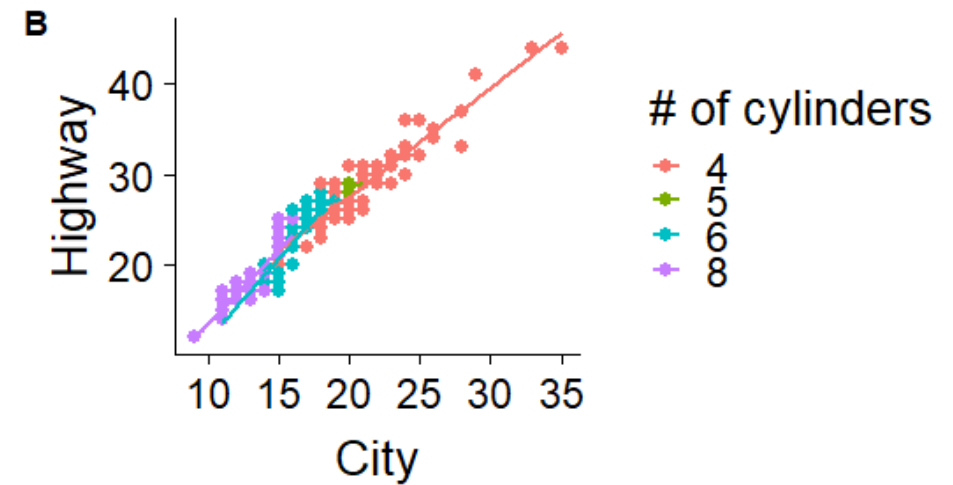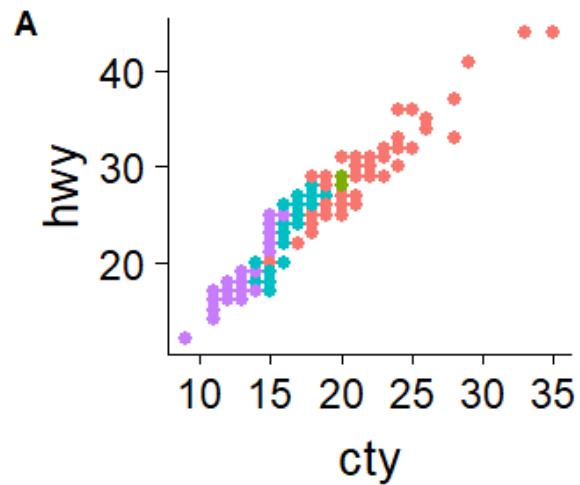
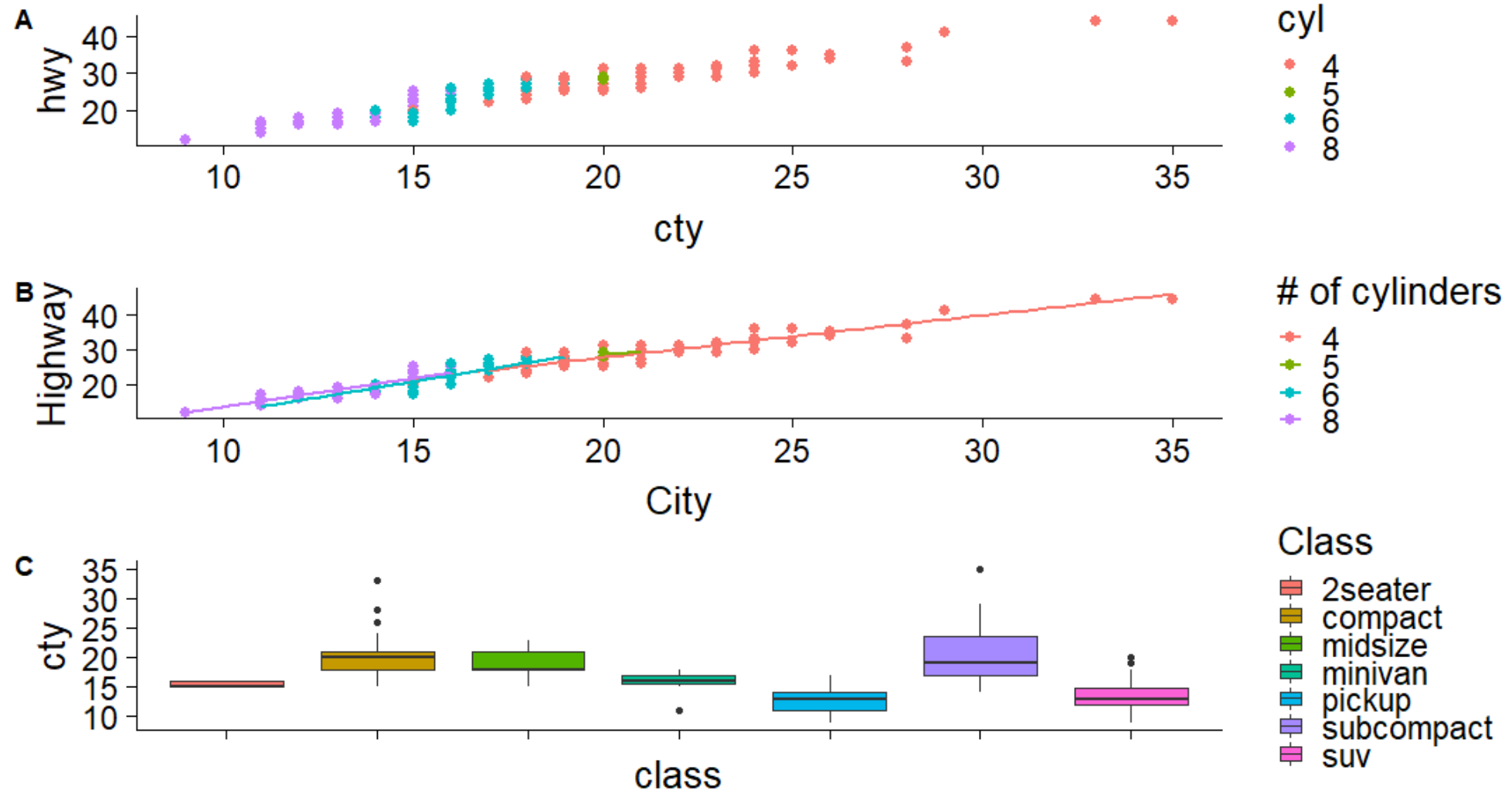# Cowplots – interface to put these together as figure

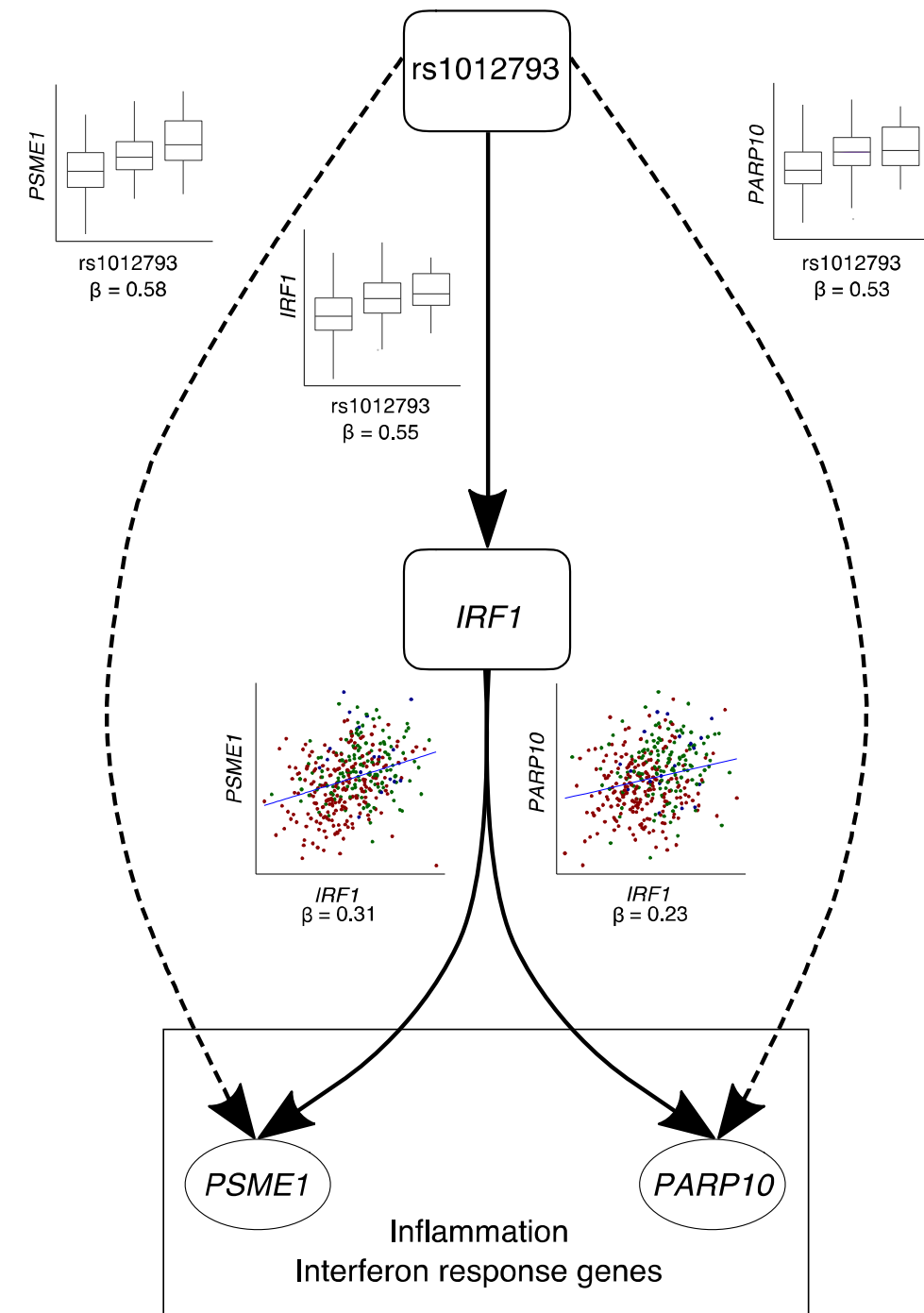So much easier than what we think!

Continue to learn

```
figure1 <- plot_grid(plot_b, plot_d, plot_g, labels = c("A", "B",
"C"), align = "v")
plot(figure1)
```

```
figure2 <- plot_grid(plot_b, plot_d, plot_g,
    labels = c("A", "B", "C"), align = "v", nrow = 3)
plot(figure2)
```

My motivations to automate plotting and figures!

GTEx Consortium. "Genetic effects on gene expression across human tissues." *Nature* (2017).

References

- ggplot2 - https://ggplot2.tidyverse.org/
- https://opr.princeton.edu/workshops/Downloads/2015Jan_ggplot2Koffman.pdf
- datavis.ca/courses/RGraphics/R-Graphics4.pdf