

```
In [1]: # load packages in the tidyverse library
library(tidyverse)

Registered S3 methods overwritten by 'ggplot2':
  method      from
[-quosures]   rlang
c.quosures    rlang
print.quosures rlang
Registered S3 method overwritten by 'rvest':
  method      from
read_xml.response xml2
-- Attaching packages -- tidyverse 1.2.1 --
✔ ggplot2 3.1.1      ✔ purrr  0.3.2
✔ tidble  2.1.1      ✔ dplyr  0.8.0.1
✔ tidyr  0.8.3       ✔ stringr 1.4.0
✔ readr   1.3.1      ✔ forcats 0.4.0
-- Conflicts --
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()     masks stats::lag()
```

Note: Use Anaconda to manage R packages. Running the line `install.packages("tidyverse")` may cause issues, and may require removing and re-creating your R environment.

```
In [14]: # data on fuel efficiency (from ggplot2 package)

mpg
#ggplot2::mpg
#print(mpg)
#print(mpg, n = 5)
#View(mpg)
```

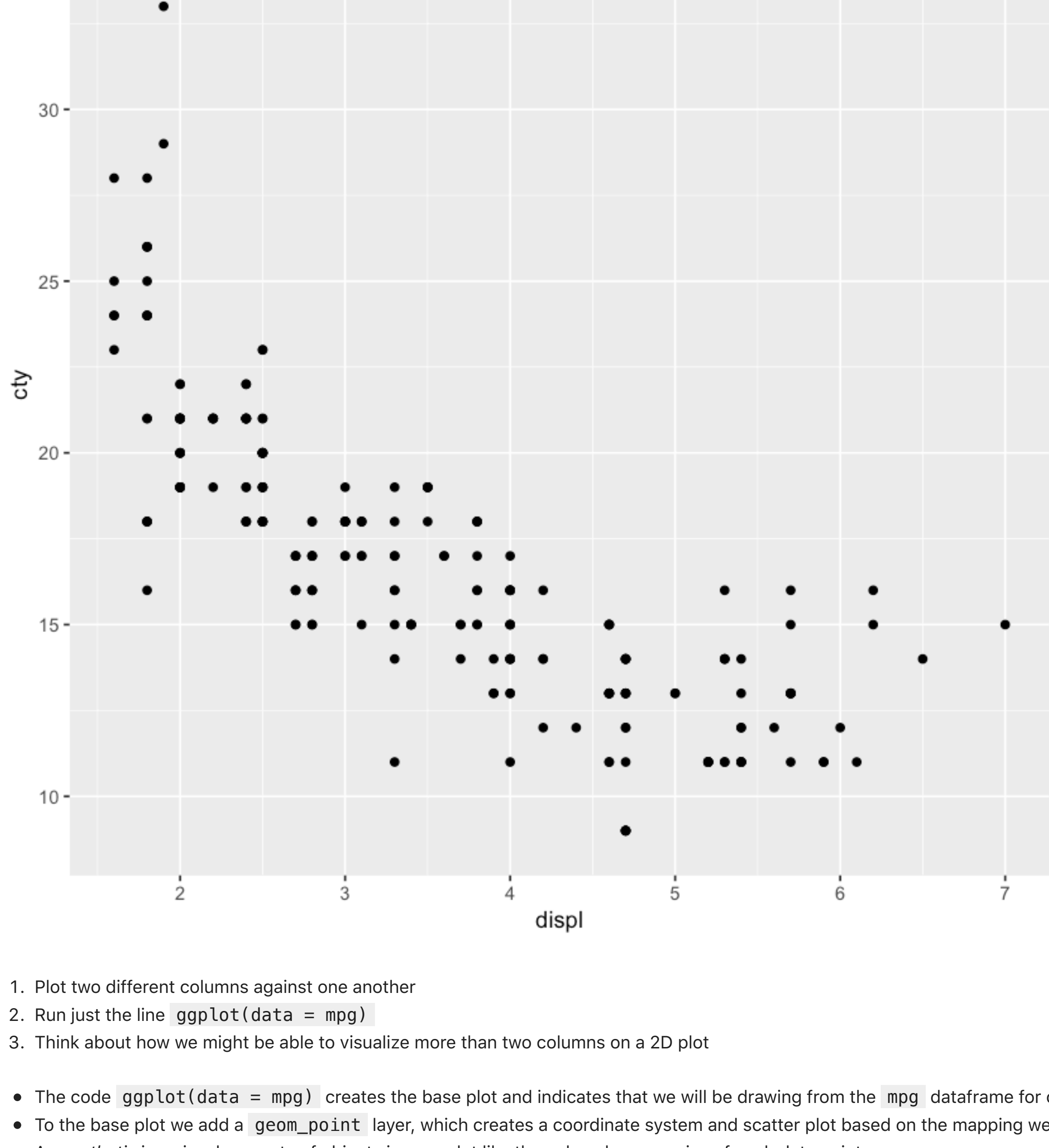
manufacturer	model	displ	year	cyl	trans	drv	cty	hwy	fl	class
audi	a4	1.8	1999	4	auto(l5)	f	18	29	p	compact
audi	a4	1.8	1999	4	manual(m5)	f	21	29	p	compact
audi	a4	2.0	2008	4	manual(m6)	f	20	31	p	compact
audi	a4	2.0	2008	4	auto(av)	f	21	30	p	compact
audi	a4	2.8	1999	6	auto(l5)	f	16	26	p	compact
audi	a4	2.8	1999	6	manual(m5)	f	18	26	p	compact
audi	a4	3.1	2008	6	auto(av)	f	18	27	p	compact
audi	a4 quattro	1.8	1999	4	manual(m5)	f	18	26	p	compact
audi	a4 quattro	1.8	1999	4	auto(l5)	f	16	25	p	compact
audi	a4 quattro	2.0	2008	4	manual(m6)	f	20	28	p	compact
audi	a4 quattro	2.0	2008	4	auto(s6)	f	19	27	p	compact
audi	a4 quattro	2.8	1999	6	auto(l5)	f	15	25	p	compact
audi	a4 quattro	2.8	1999	6	manual(m5)	f	17	25	p	compact
audi	a4 quattro	3.1	2008	6	auto(s6)	f	17	25	p	compact
audi	a4 quattro	3.1	2008	6	manual(m6)	f	15	25	p	compact
audi	a6 quattro	2.8	1999	6	auto(l5)	f	15	24	p	midsize
audi	a6 quattro	3.1	2008	6	auto(s6)	f	17	25	p	midsize
audi	a6 quattro	4.2	2008	8	auto(s6)	f	16	23	p	midsize
chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(4)	r	14	20	r	suv
chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(4)	r	11	15	e	suv
chevrolet	c1500 suburban 2wd	5.3	2008	8	auto(4)	r	14	20	r	suv
chevrolet	c1500 suburban 2wd	5.7	1999	8	auto(4)	r	13	17	r	suv
chevrolet	c1500 suburban 2wd	6.0	2008	8	auto(4)	r	12	17	r	suv
chevrolet	corvette	5.7	1999	8	manual(m6)	r	16	26	p	2seater
chevrolet	corvette	5.7	1999	8	auto(4)	r	15	23	p	2seater
chevrolet	corvette	6.2	2008	8	manual(m6)	r	16	26	p	2seater
chevrolet	corvette	6.2	2008	8	auto(s6)	r	15	25	p	2seater
chevrolet	corvette	7.0	2008	8	manual(m6)	r	15	24	p	2seater
chevrolet	K1500 tahoe 4wd	5.3	2008	8	auto(4)	f	14	19	r	suv
chevrolet	k1500 tahoe 4wd	5.3	2008	8	auto(4)	f	11	14	e	suv
...	...	...	...	...	...	...	...	...	...	...
toyota	toyota tacoma 4wd	3.4	1999	6	auto(4)	f	15	19	r	pickup
toyota	toyota tacoma 4wd	4.0	2008	6	manual(m6)	f	15	18	r	pickup
toyota	toyota tacoma 4wd	4.0	2008	6	auto(l5)	f	16	20	r	pickup
volkswagen	gti	2.0	1999	4	manual(m5)	f	21	29	r	compact
volkswagen	gti	2.0	1999	4	auto(4)	f	19	26	r	compact
volkswagen	gti	2.0	2008	4	manual(m6)	f	21	29	p	compact
volkswagen	gti	2.0	2008	4	auto(s6)	f	22	29	p	compact
volkswagen	gti	2.8	1999	6	manual(m5)	f	17	24	r	compact
volkswagen	jetta	1.9	1999	4	manual(m5)	f	33	44	d	compact
volkswagen	jetta	2.0	1999	4	manual(m5)	f	21	29	r	compact
volkswagen	jetta	2.0	1999	4	auto(4)	f	19	26	r	compact
volkswagen	jetta	2.0	2008	4	auto(s6)	f	22	29	p	compact
volkswagen	jetta	2.0	2008	4	manual(m6)	f	21	29	p	compact
volkswagen	jetta	2.5	2008	5	auto(s6)	f	21	29	r	compact
volkswagen	jetta	2.5	2008	5	manual(m5)	f	21	29	r	compact
volkswagen	jetta	2.8	1999	6	auto(4)	f	16	23	r	compact
volkswagen	jetta	2.8	1999	6	manual(m5)	f	17	24	r	compact
volkswagen	new beetle	1.9	1999	4	manual(m5)	f	35	44	d	subcompact
volkswagen	new beetle	1.9	1999	4	auto(4)	f	29	41	d	subcompact
volkswagen	new beetle	2.0	1999	4	manual(m5)	f	21	29	r	subcompact
volkswagen	new beetle	2.0	1999	4	auto(4)	f	19	26	r	subcompact
volkswagen	new beetle	2.5	2008	5	manual(m5)	f	20	28	r	subcompact
volkswagen	new beetle	2.5	2008	5	auto(s6)	f	20	29	r	subcompact
volkswagen	passat	1.8	1999	4	manual(m5)	f	21	29	p	midsize
volkswagen	passat	1.8	1999	4	auto(l5)	f	18	29	p	midsize
volkswagen	passat	2.0	2008	4	auto(s6)	f	19	28	p	midsize
volkswagen	passat	2.0	2008	4	manual(m6)	f	21	29	p	midsize
volkswagen	passat	2.8	1999	6	auto(l5)	f	16	26	p	midsize
volkswagen	passat	2.8	1999	6	manual(m5)	f	18	26	p	midsize
volkswagen	passat	3.6	2008	6	auto(s6)	f	17	26	p	midsize

```
In [11]: nrow(mpg)
11
ncol(mpg)
```

```
234
11
```

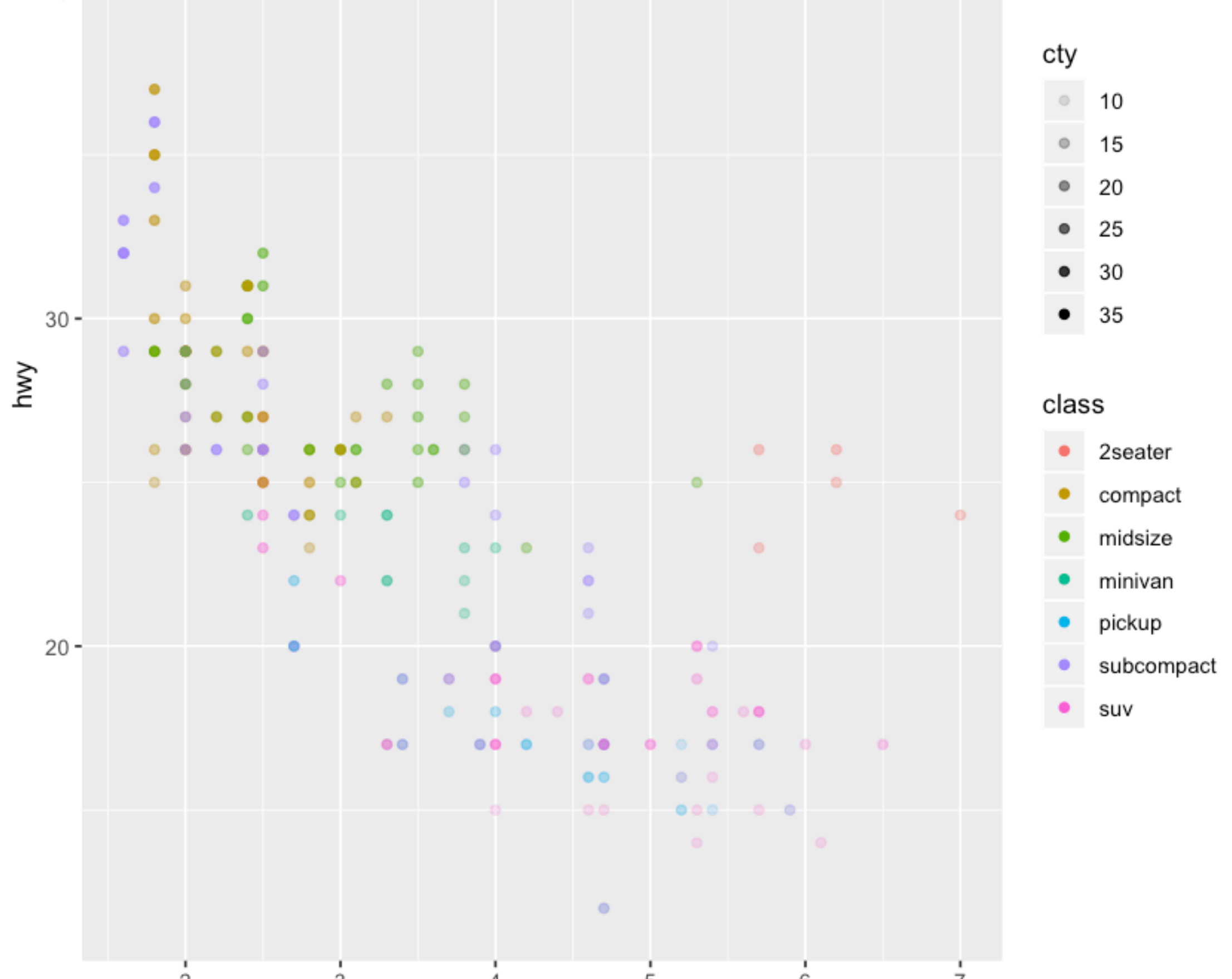
```
In [13]: # what are these columns?
?mpg
```

```
In [19]: # plot engine volume against highway mpg
ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = cty))
```



- Plot two different columns against one another
  - Run just the line `ggplot(data = mpg)`
  - Think about how we might be able to visualize more than two columns on a 2D plot
- The code `ggplot(data = mpg)` creates the base plot and indicates that we will be drawing from the `mpg` dataframe for our data.
  - To the base plot we add a `geom_point` layer, which creates a coordinate system and scatter plot based on the mapping we give it.
  - An *aesthetic* is a visual property of objects in your plot like the color, shape, or size of each data point.

```
In [24]: ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, alpha = cty, color = class))
```



- Try out the aesthetic options color, size, shape, and alpha. What do each do?
- How could we find out more about the available aesthetics for the `geom_point` layer?

```
In [25]: ?geom_point
```

Be careful with writing code across multiple lines!

```
In [ ]: # okay
ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, color = cty))

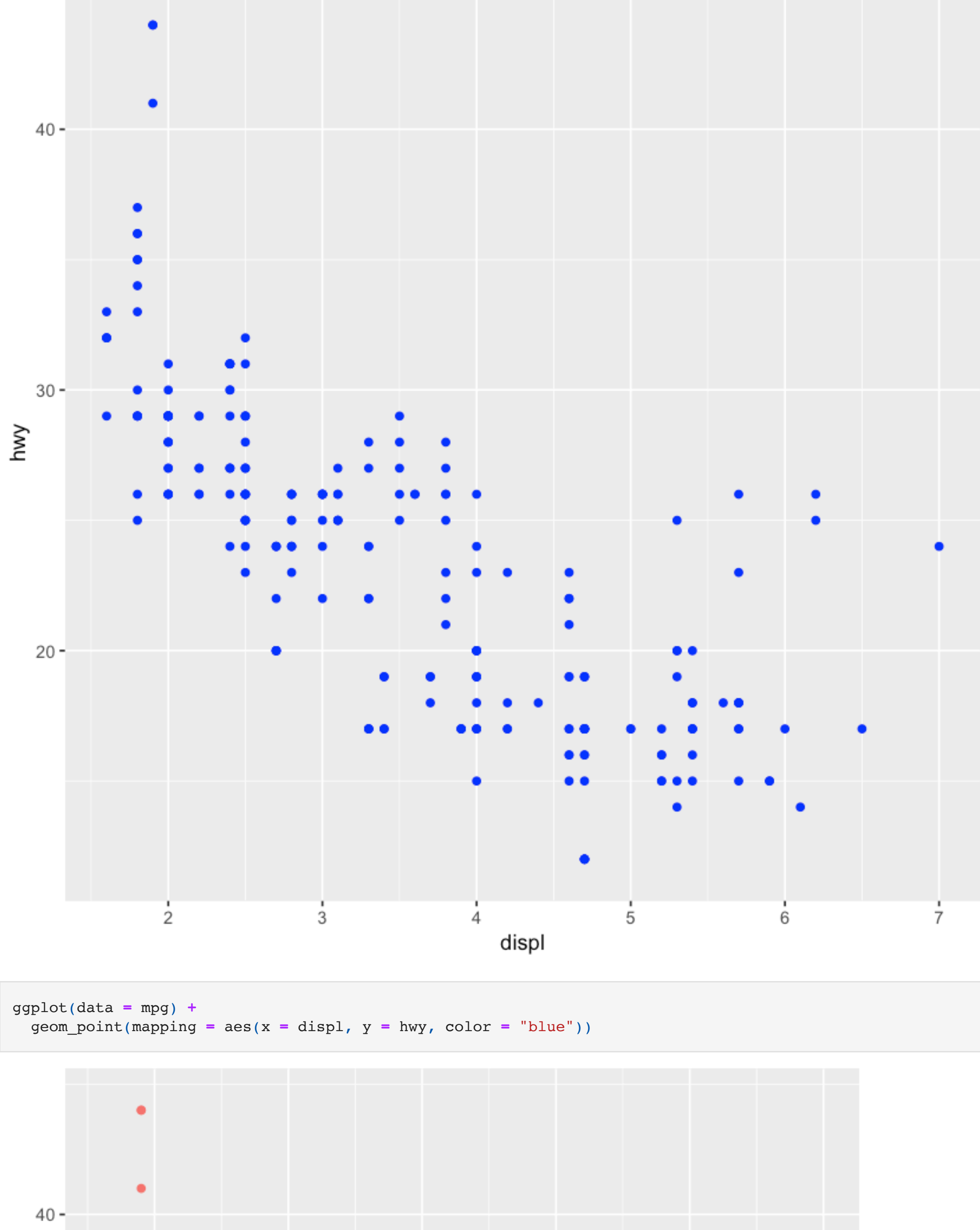
# not okay
ggplot(data = mpg)
+ geom_point(mapping = aes(x = displ, y = hwy, color = cty))

# okay
ggplot(data = mpg) +
  geom_point(
    mapping = aes(x = displ, y = hwy, color = cty))

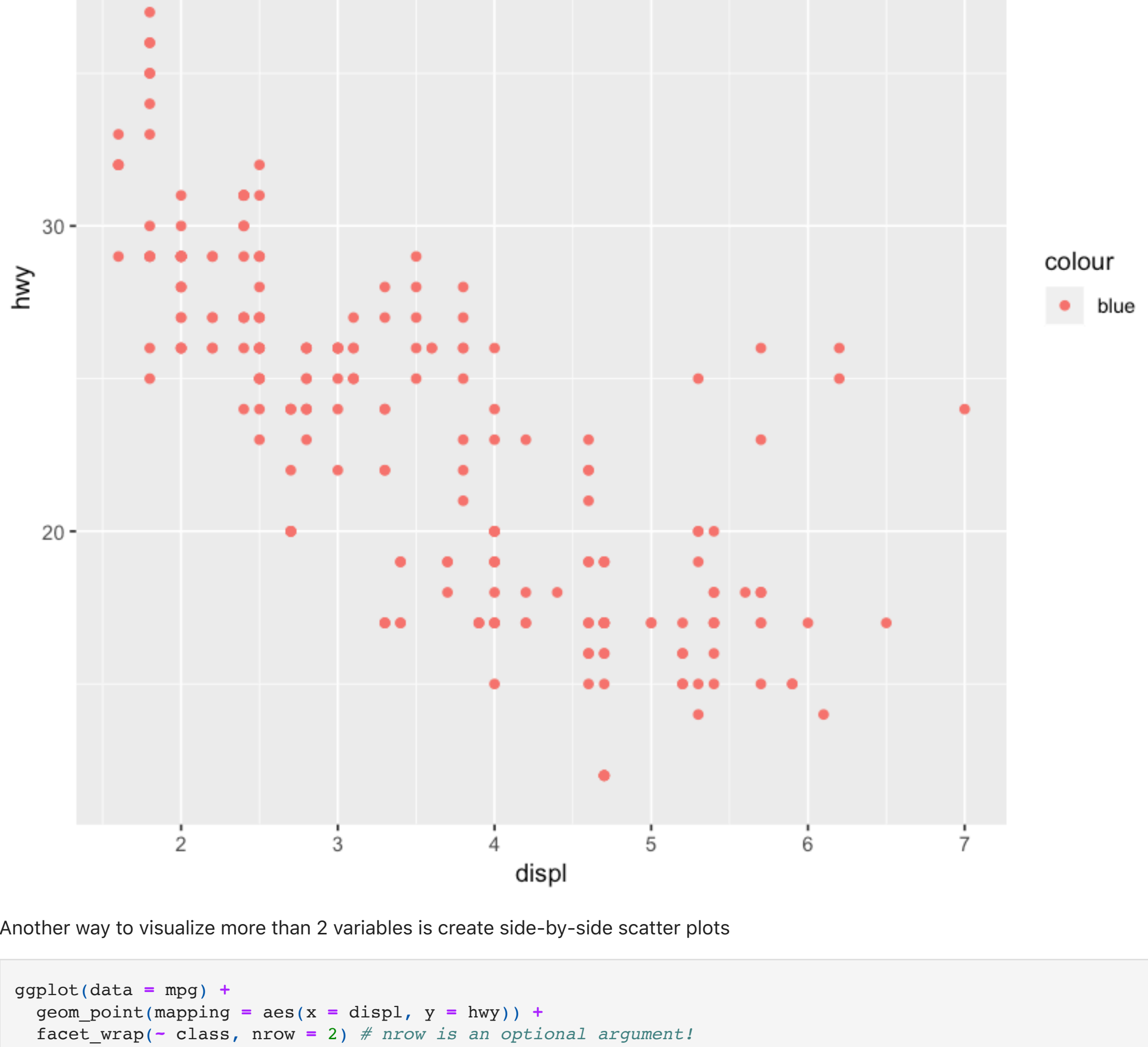
# not okay
ggplot(data = mpg) +
  geom_point
  (mapping = aes(x = displ, y = hwy, color = cty))
```

Run the following two blocks of code. What happened and why are they different?

```
In [28]: ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, color = "blue"))
```

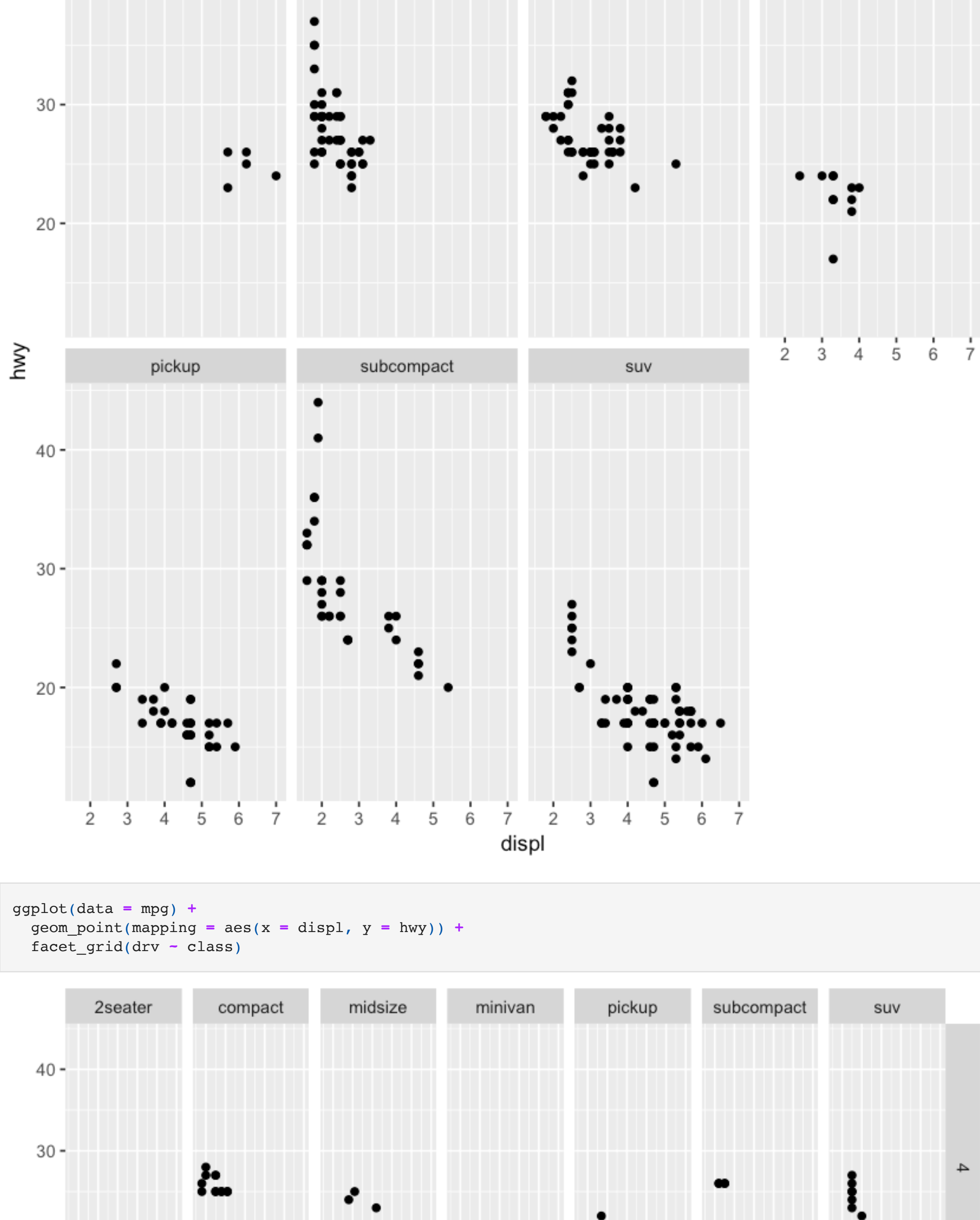


```
In [31]: ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, color = "blue"))
```

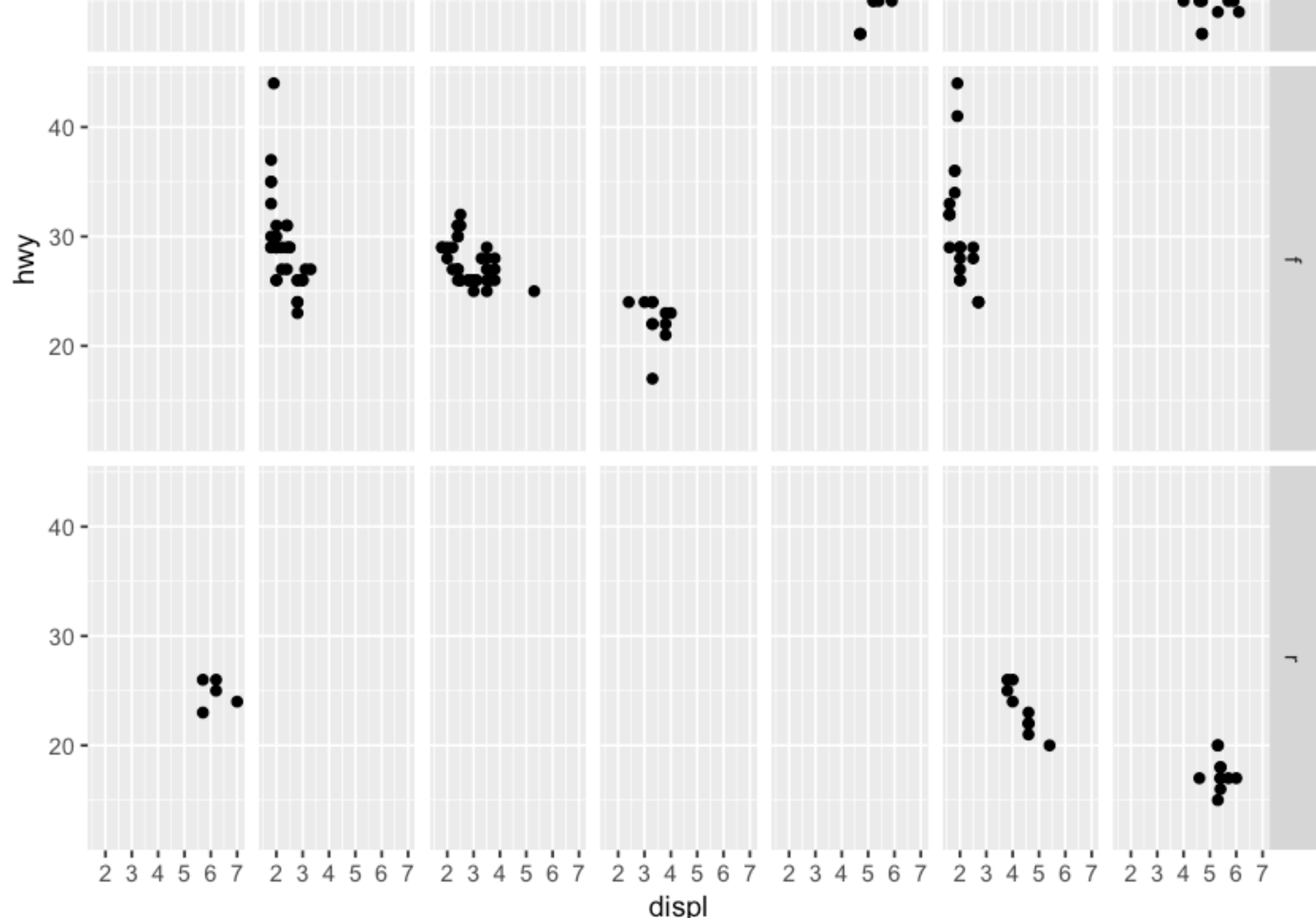


Another way to visualize more than 2 variables is create side-by-side scatter plots

```
In [35]: ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy)) +
  facet_wrap(~ class, nrow = 2) # nrow is an optional argument!
```



```
In [36]: ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy)) +
  facet_grid(drv ~ class)
```



- Try out a few different facet grids!
- Run the following two blocks of code. What happened?

```
In [ ]: ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy)) +
  facet_grid(. ~ class)
```

```
In [ ]: ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy)) +
  facet_grid(class ~ .)
```