

# R Notebook

import

```
library(tidyverse)
```

```
## — Attaching core tidyverse packages — tidyverse 2.0.0 —
## ✓ dplyr      1.1.3    ✓ readr      2.1.4
## ✓ forcats    1.0.0    ✓ stringr    1.5.0
## ✓ ggplot2    3.4.4    ✓ tibble     3.2.1
## ✓ lubridate  1.9.3    ✓ tidyr      1.3.0
## ✓ purrr      1.0.2
## — Conflicts — tidyverse_conflicts() —
## ✗ dplyr::filter() masks stats::filter()
## ✗ dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(vctrs)
```

```
##
## Attaching package: 'vctrs'
##
## The following object is masked from 'package:dplyr':
##
##   data_frame
##
## The following object is masked from 'package:tibble':
##
##   data_frame
```

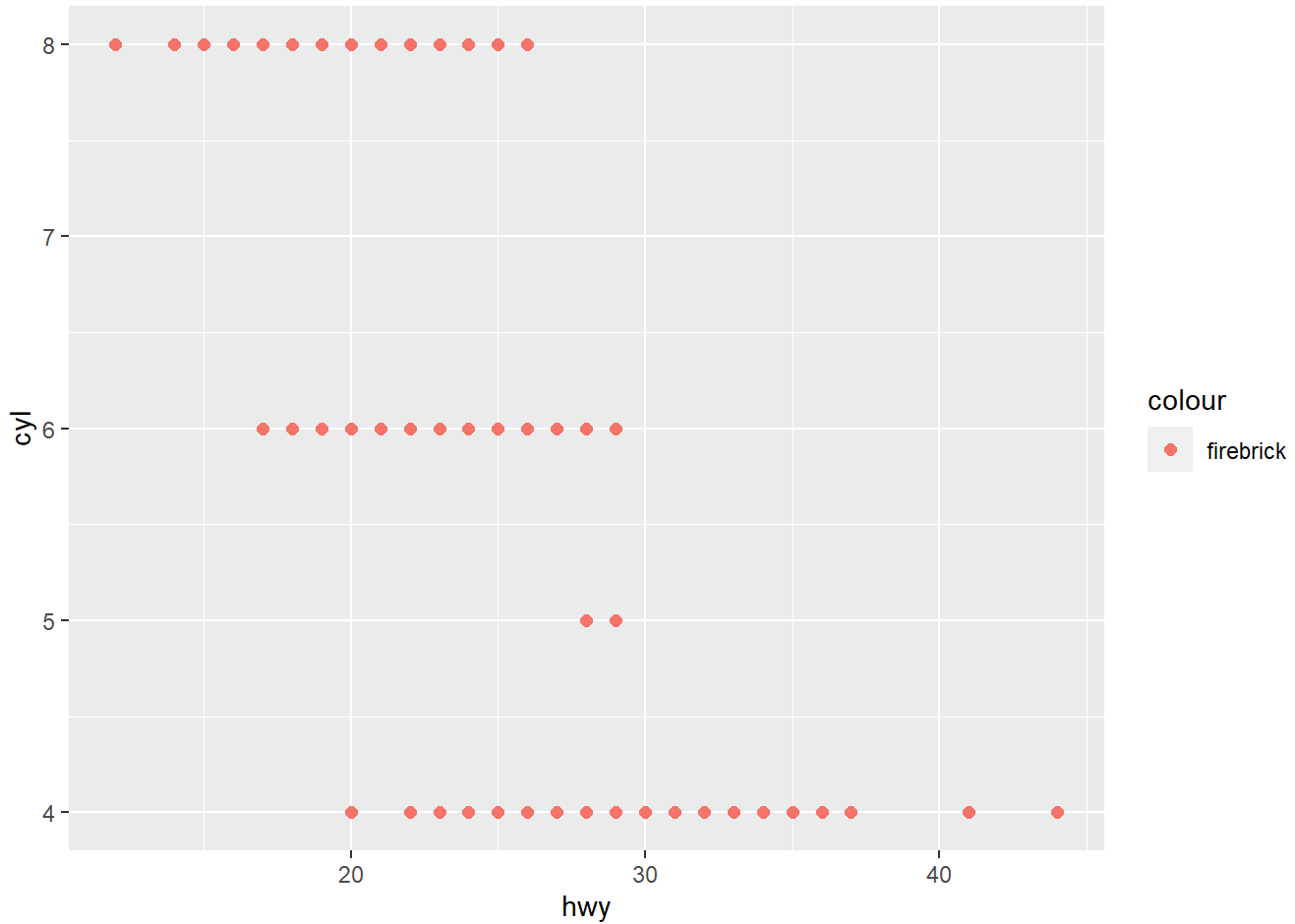
```
head(mpg)
```

manufacturer <chr>	model <chr>	displ <dbl>	year <int>	cyl <int>	trans <chr>	drv <chr>	cty <int>	hwy <int>	fl <chr>
audi	a4	1.8	1999	4	auto(l5)	f	18	29	p
audi	a4	1.8	1999	4	manual(m5)	f	21	29	p
audi	a4	2.0	2008	4	manual(m6)	f	20	31	p
audi	a4	2.0	2008	4	auto(av)	f	21	30	p
audi	a4	2.8	1999	6	auto(l5)	f	16	26	p
audi	a4	2.8	1999	6	manual(m5)	f	18	26	p

6 rows | 1-10 of 11 columns

## Exercise 12.1

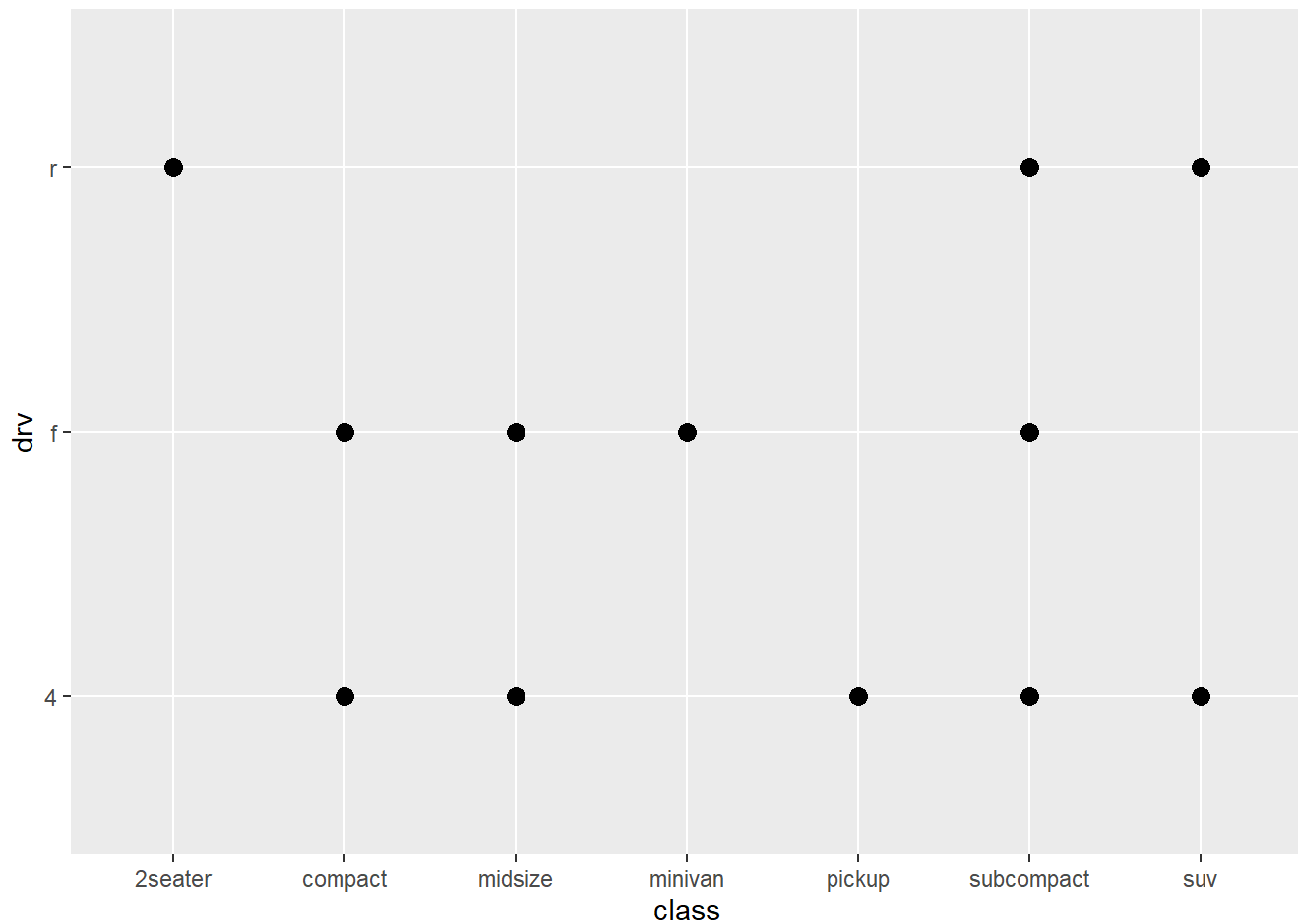
```
mpg |> ggplot(aes(  
  x = hwy,  
  y = cyl  
) + geom_point(size = 2, aes(color = 'firebrick'))
```



## Exercise 12.3

This plot is not useful because they are categorical data. They are not good for scatter plots

```
mpg |> ggplot(aes(  
  x = class,  
  y = drv  
) + geom_point(size = 3)
```



## Exercise 12.4

```
x <- seq(-6*pi, 6*pi, length.out = 100)
y <- sin(x)/x
df <- data.frame(x = x, y = sin(x)/x)
#df |> ggplot(
#  #data = df,
#  #aes(x = x, y = y)) +
#  #geom_line()
```

cars

	speed <dbl>	dist <dbl>
	4	2
	4	10
	7	4
	7	22
	8	16

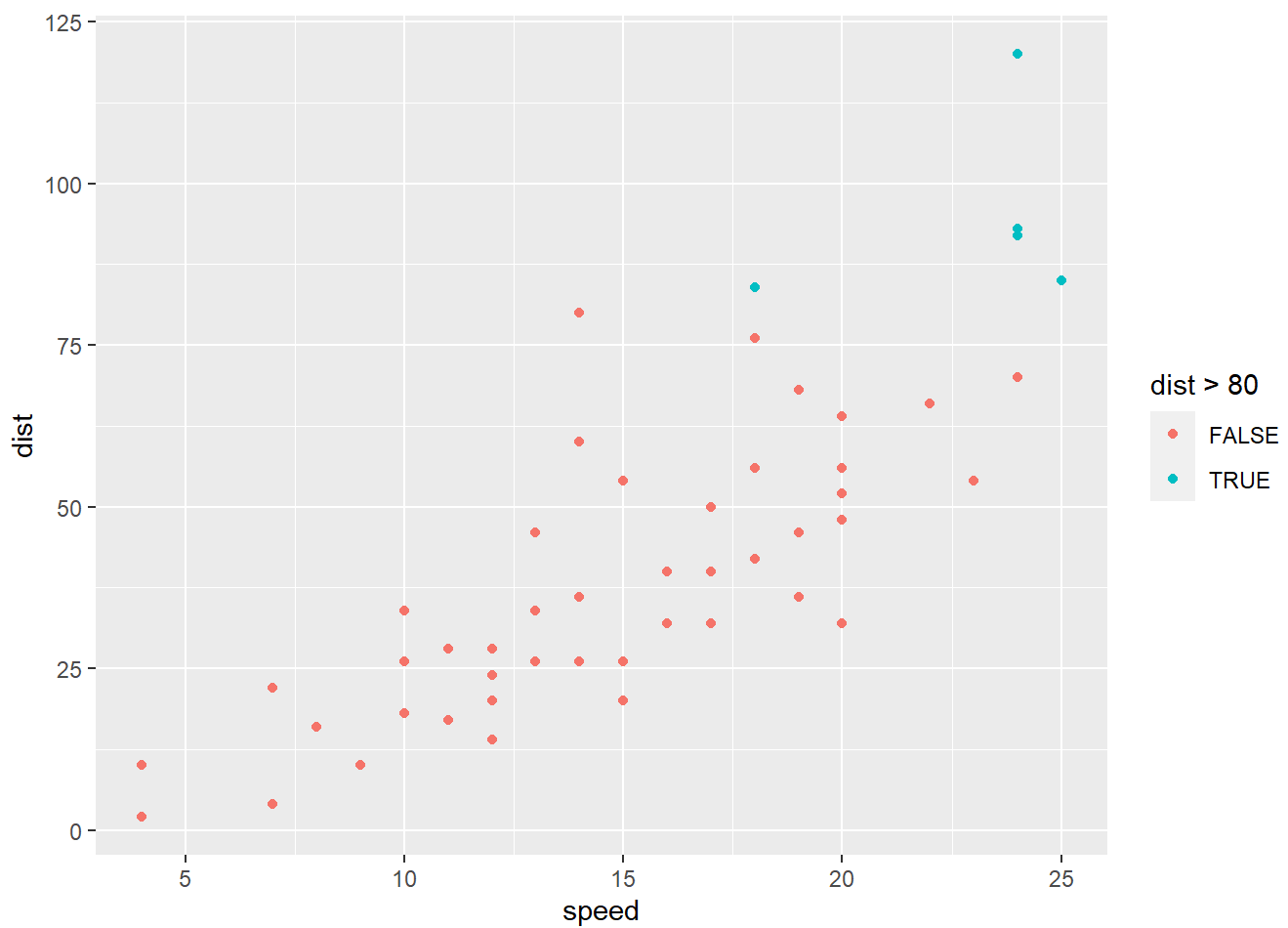
speed	dist
<dbl>	<dbl>
9	10
10	18
10	26
10	34
11	17

1-10 of 50 rows

Previous 1 2 3 4 5 Next

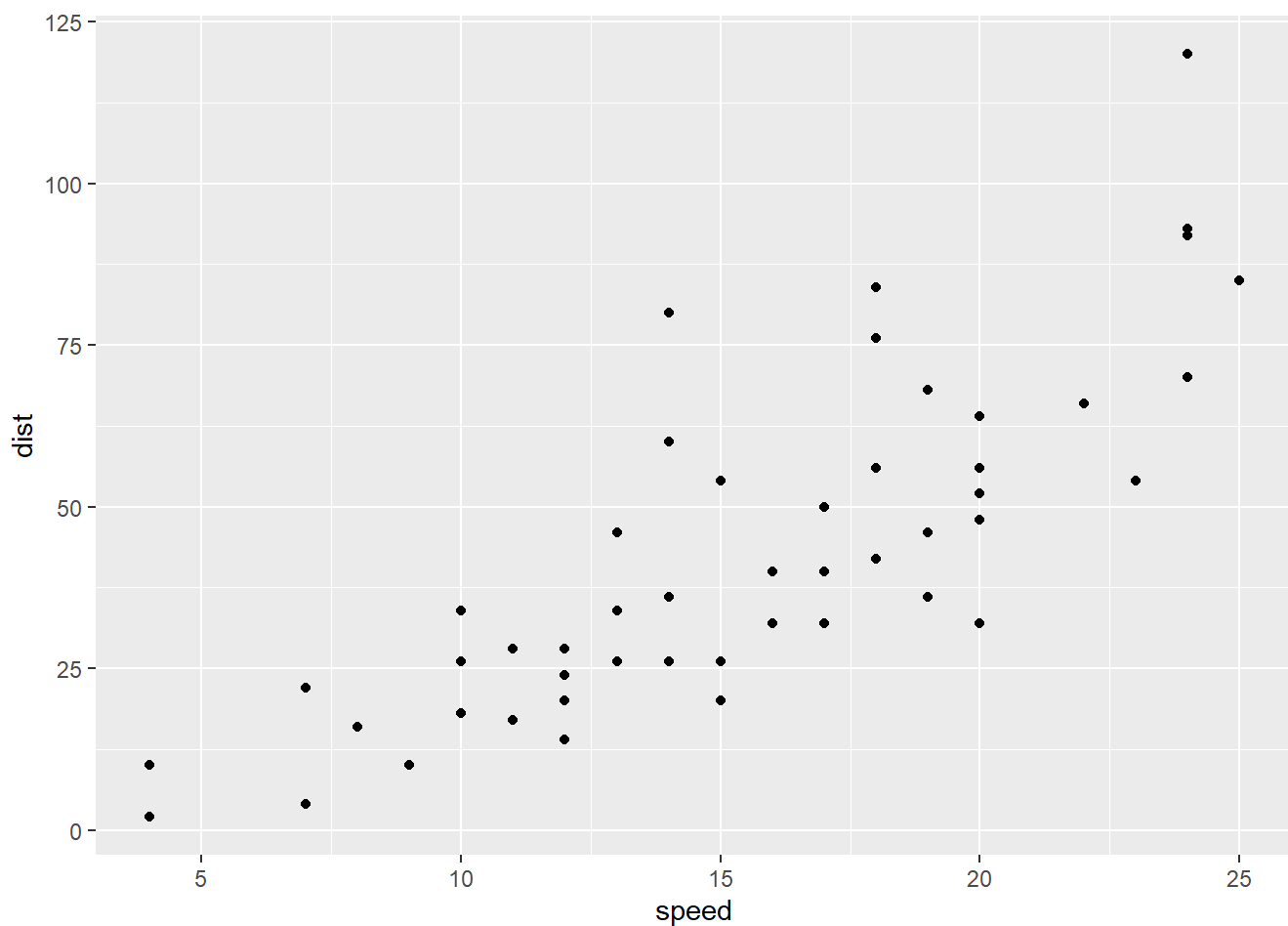
## Exercise 12.5

```
cars |>
  ggplot(mapping = aes(x = speed, y = dist))+
  geom_point(mapping = aes(color = dist > 80 ))
```



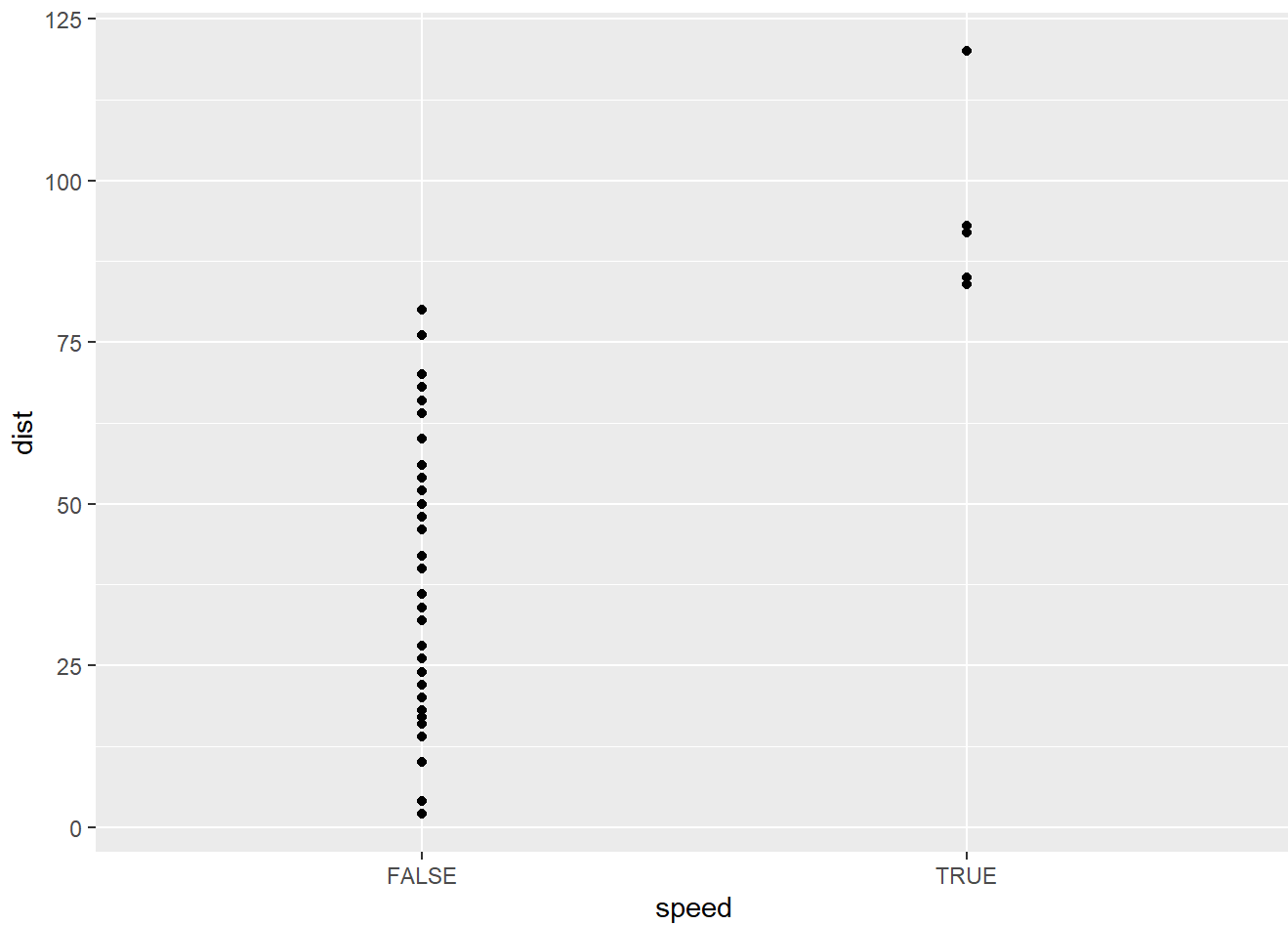
## Exercise 12.6

```
cars |>  
  ggplot(aes(x = speed, y = dist))+  
  geom_point()
```



## Exercise 12.7

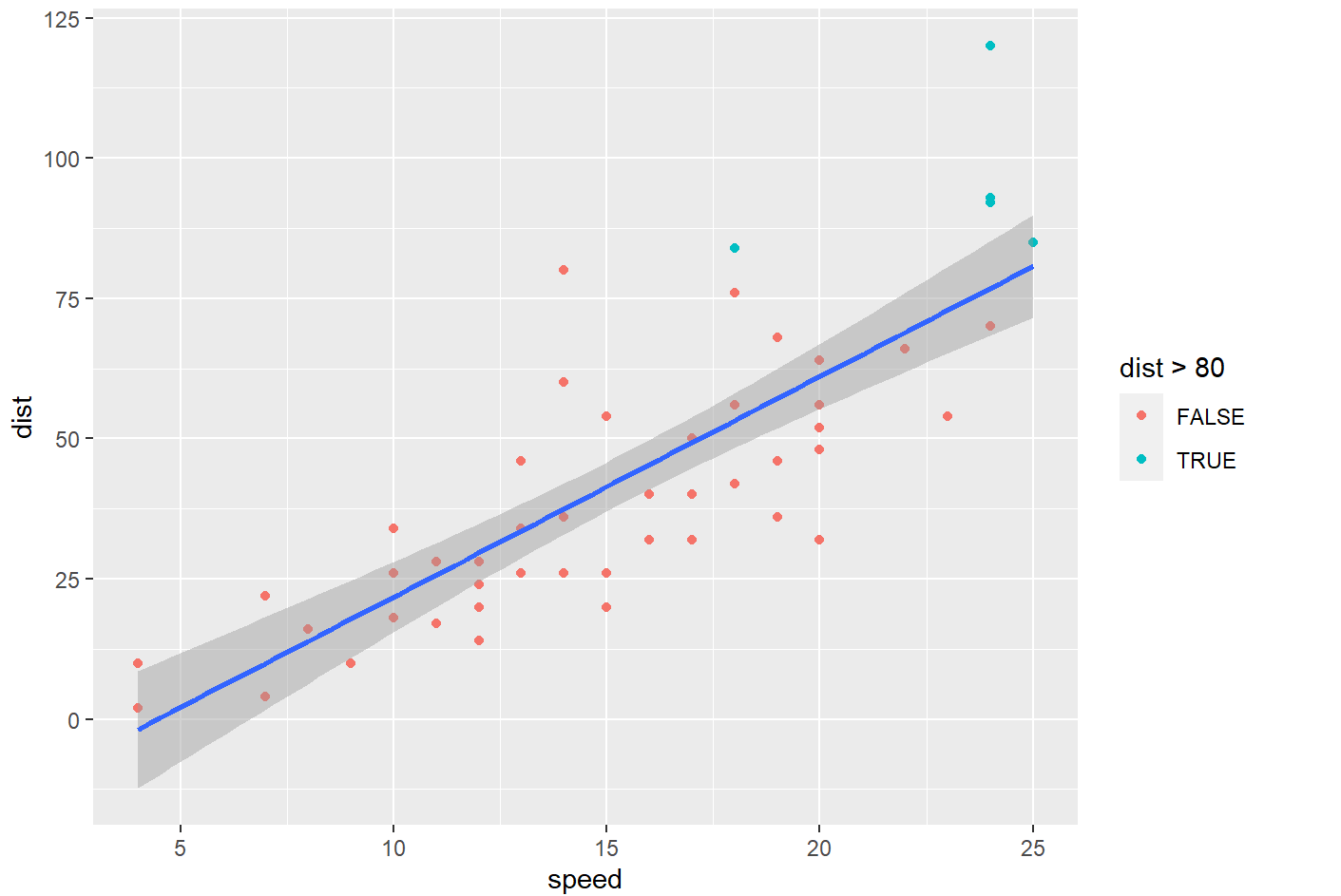
```
cars |>  
  ggplot(mapping = aes(x = speed, y = dist))+  
  geom_point(mapping = aes(dist > 80))+  
  scale_color_manual(values = c("black", "red"))
```



## Exercise 12.8

```
cars |>
  ggplot(aes(x = speed, y = dist))+
  geom_point(mapping = aes(color = dist > 80 )) +
  geom_smooth(method = 'lm')
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



cars

speed	dist
<dbl>	<dbl>
4	2
4	10
7	4
7	22
8	16
9	10
10	18
10	26
10	34
11	17

1-10 of 50 rows

Previous

1

2

3

4

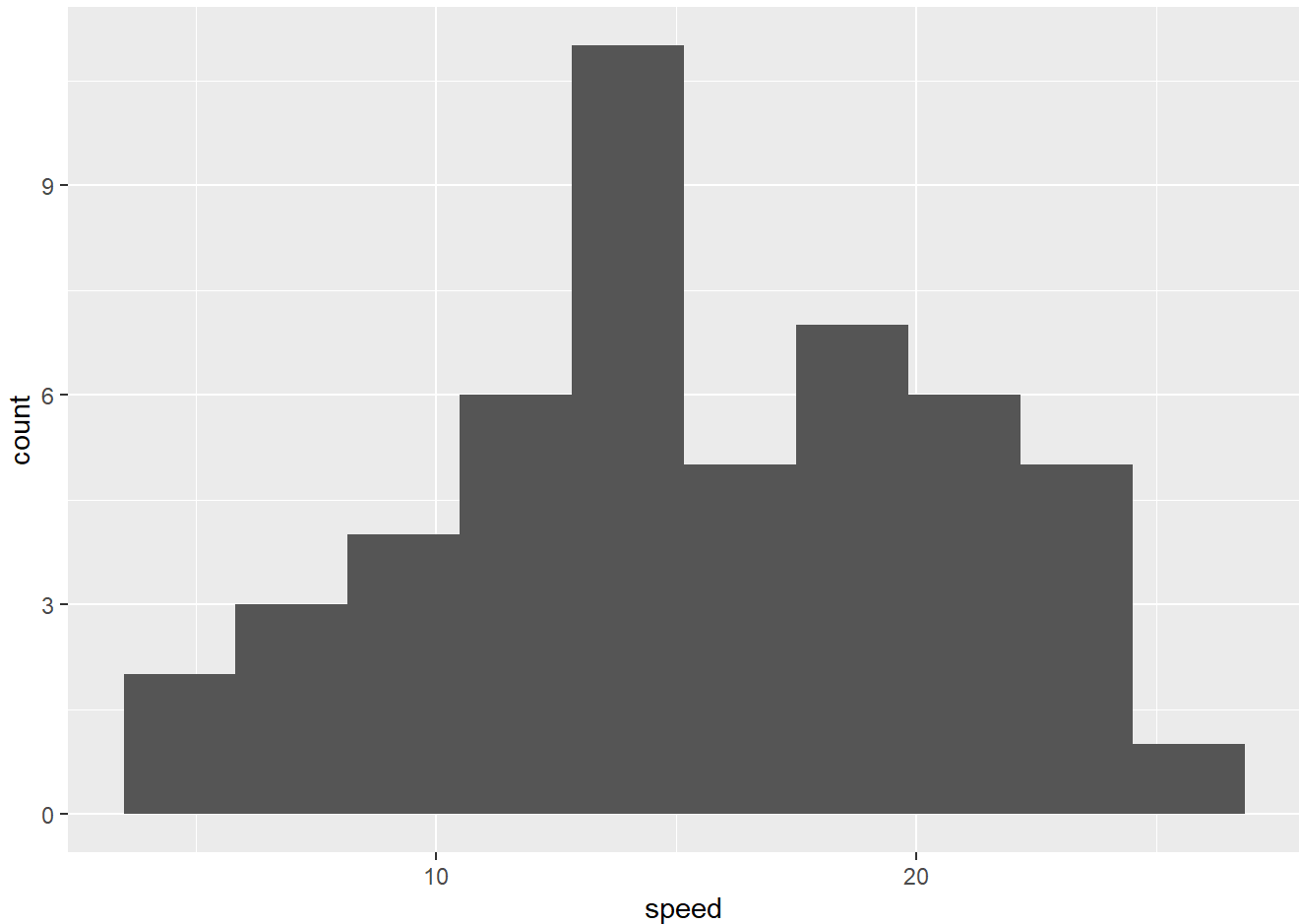
5

Next

## Exercise 12.9

Here, it's single aes mapping, so no y needed. You can add bins to cut points in vector x into evenly distributed groups

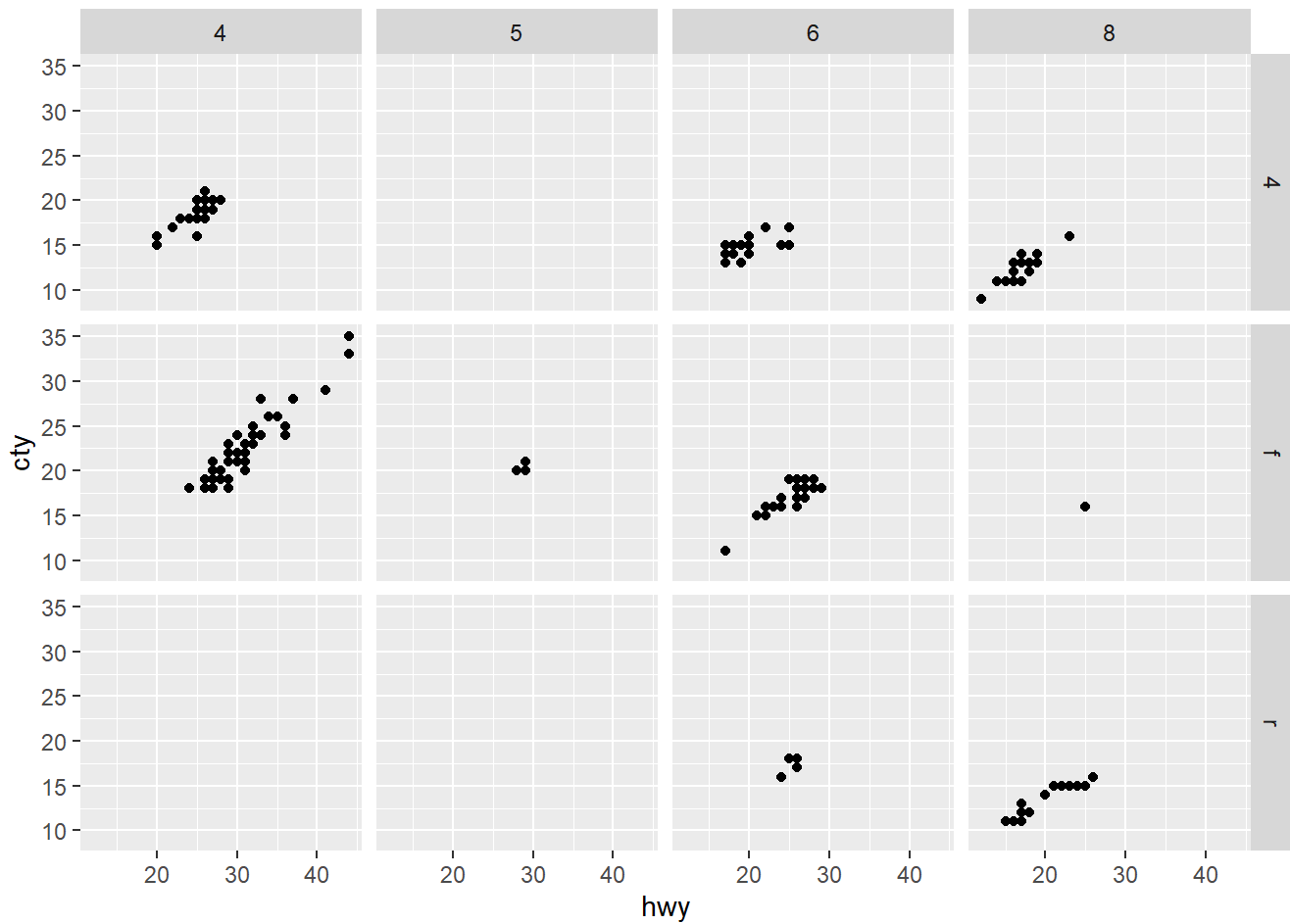
```
cars |>  
  ggplot(aes(  
    x = speed)) +  
  geom_histogram(bins = 10)
```



## Exercise 12.10

```
mpg <- mpg  
  
mpg |> ggplot() +  
  geom_point(aes(hwy, cty)) +  
  facet_grid(drv~cyl)
```





## Exercise 12.11

```
iris |>
  ggplot(aes(Sepal.Length, Sepal.Width, color = Species, shape = Species)) +
  geom_point() +
  geom_density2d() +
  ggtitle('IRIS') +
  theme_light()
```

