visualisation.R

dribr

2022-01-28

```
library(ggplot2)
```

```
## Warning in register(): Can't find generic `scale_type` in package ggplot2 to
## register S3 method.
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v tibble  3.1.6      v dplyr   1.0.7
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.1.1      v forcats 0.5.1
## v purrr   0.3.4
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
mpg
```
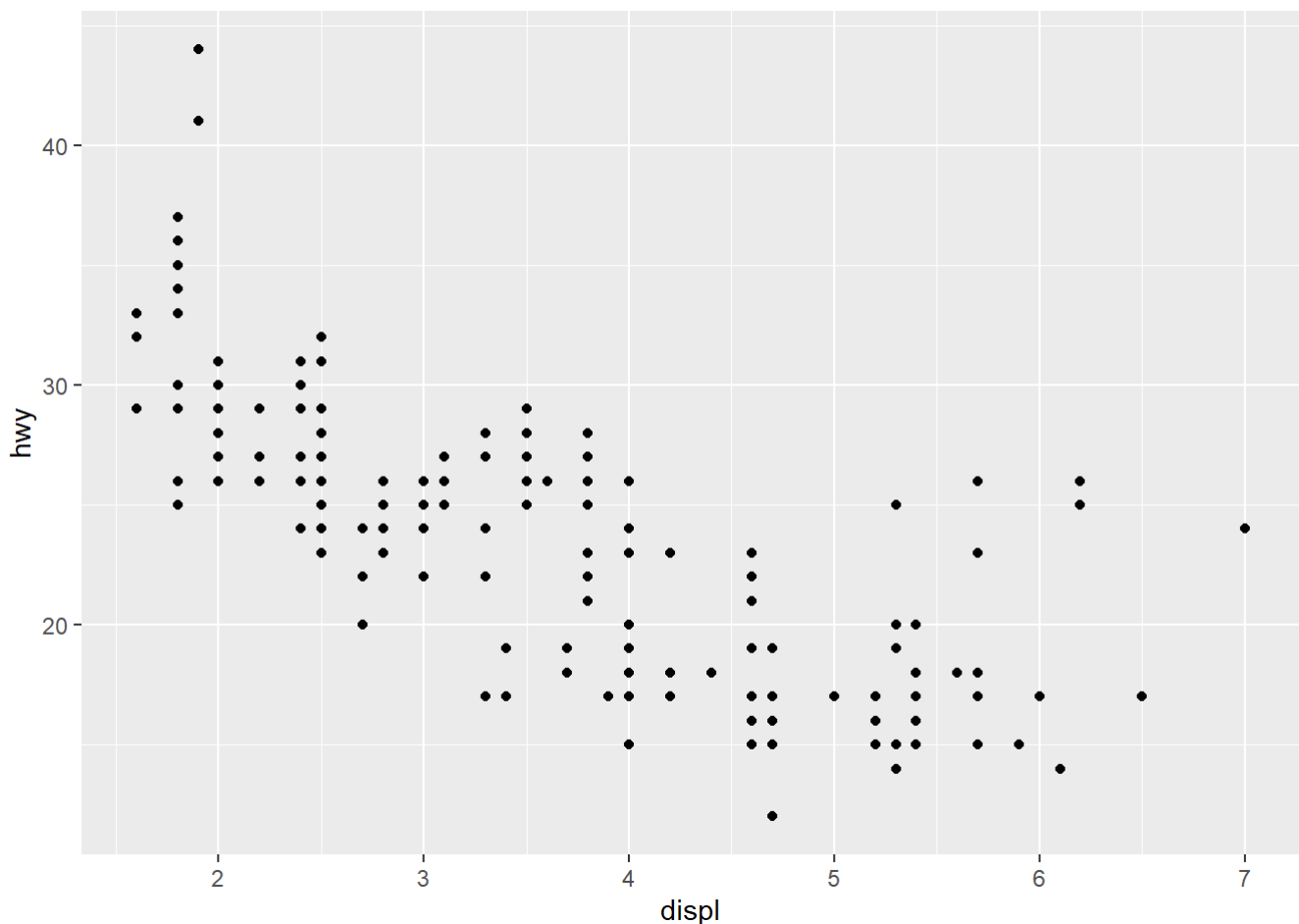
```
## # A tibble: 234 x 11
##    manufacturer model      displ  year   cyl trans drv     cty   hwy fl    class
##    <chr>        <chr>      <dbl> <int> <int> <chr> <chr> <int> <int> <chr> <chr>
##  1 audi         a4           1.8  1999     4 auto~ f        18    29 p     comp~
##  2 audi         a4           1.8  1999     4 manu~ f        21    29 p     comp~
##  3 audi         a4           2    2008     4 manu~ f        20    31 p     comp~
##  4 audi         a4           2    2008     4 auto~ f        21    30 p     comp~
##  5 audi         a4           2.8  1999     6 auto~ f        16    26 p     comp~
##  6 audi         a4           2.8  1999     6 manu~ f        18    26 p     comp~
##  7 audi         a4           3.1  2008     6 auto~ f        18    27 p     comp~
##  8 audi         a4 quattro   1.8  1999     4 manu~ 4        18    26 p     comp~
##  9 audi         a4 quattro   1.8  1999     4 auto~ 4        16    25 p     comp~
## 10 audi         a4 quattro   2    2008     4 manu~ 4        20    28 p     comp~
## # ... with 224 more rows
```

```
summary(mpg)
```
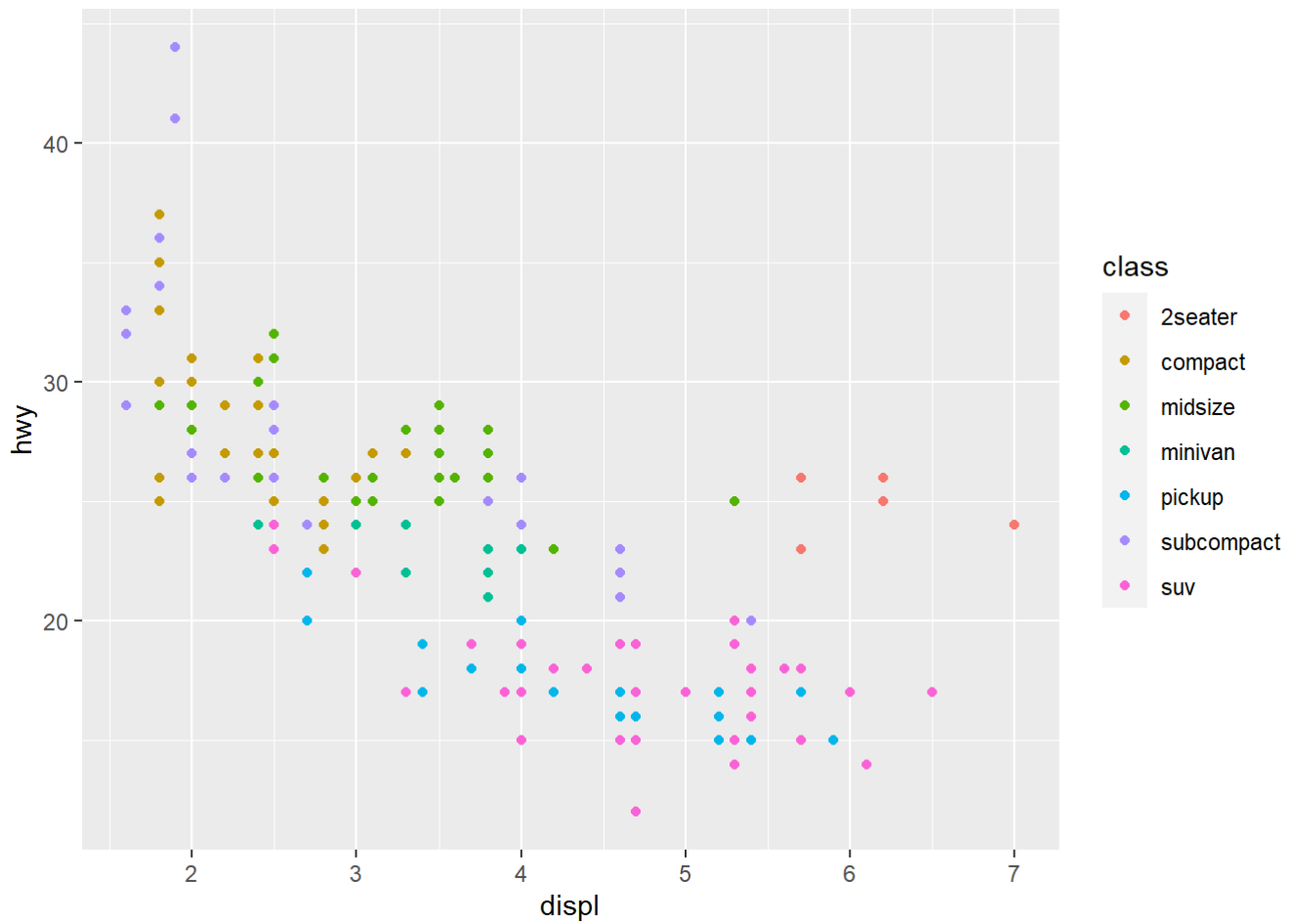
```
##   manufacturer        model              displ            year
##  Length:234         Length:234         Min.   :1.600   Min.   :1999
##  Class :character   Class :character   1st Qu.:2.400   1st Qu.:1999
##  Mode  :character   Mode  :character   Median :3.300   Median :2004
##                                        Mean   :3.472   Mean   :2004
##                                        3rd Qu.:4.600   3rd Qu.:2008
##                                        Max.   :7.000   Max.   :2008
##       cyl           trans              drv                 cty
##  Min.   :4.000   Length:234         Length:234         Min.   : 9.00
##  1st Qu.:4.000   Class :character   Class :character   1st Qu.:14.00
##  Median :6.000   Mode  :character   Mode  :character   Median :17.00
##  Mean   :5.889                                         Mean   :16.86
##  3rd Qu.:8.000                                         3rd Qu.:19.00
##  Max.   :8.000                                         Max.   :35.00
##       hwy            fl               class
##  Min.   :12.00   Length:234         Length:234
##  1st Qu.:18.00   Class :character   Class :character
##  Median :24.00   Mode  :character   Mode  :character
##  Mean   :23.44
##  3rd Qu.:27.00
##  Max.   :44.00
```

```
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy))
```
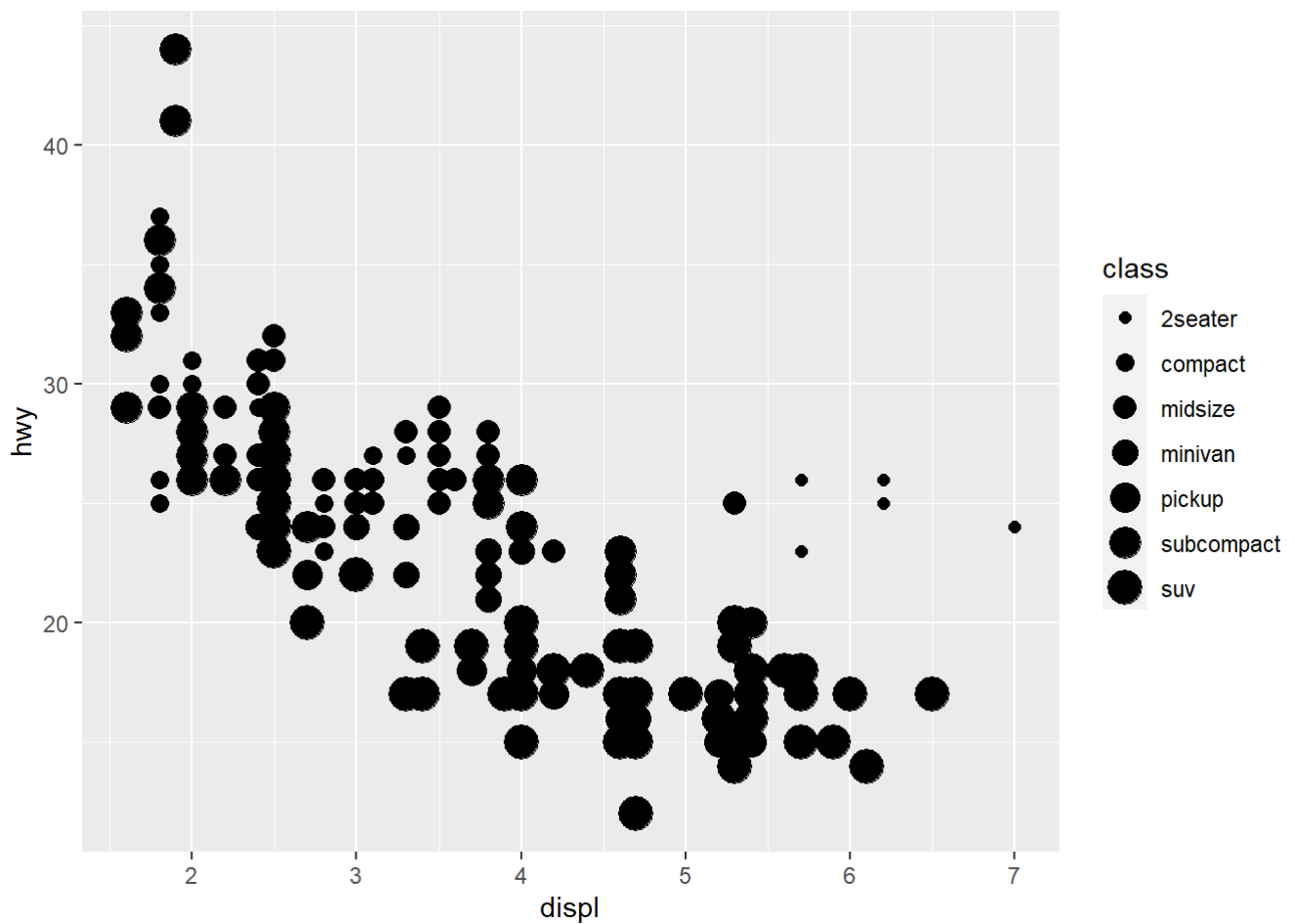


```
#The plot shows a negative relationship between engine size (displ) and fuel efficiency (hw
y).cars with big engines use more fuel.
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy, color=class))
```
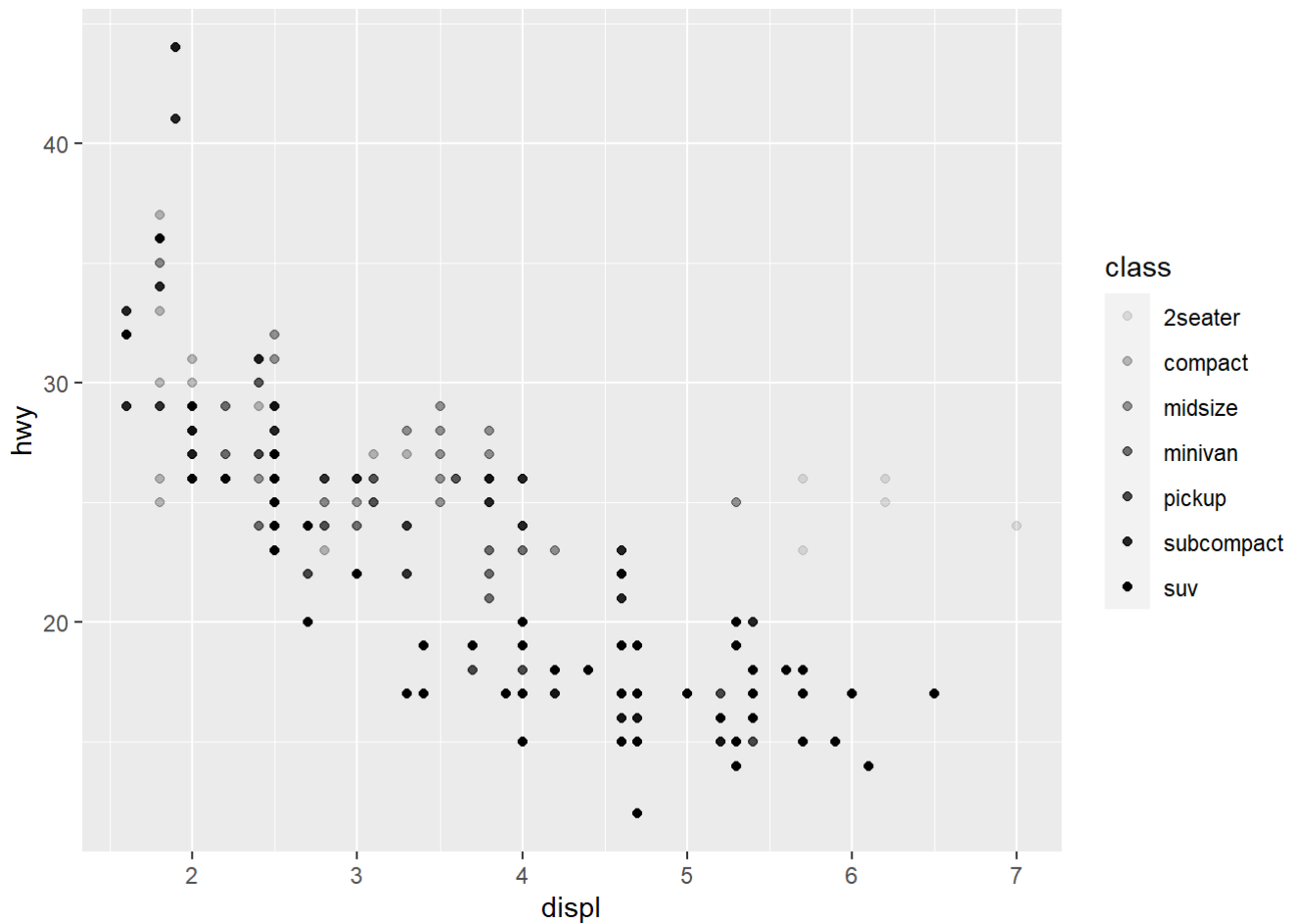
```
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy, size=class))
```

```
## Warning: Using size for a discrete variable is not advised.
```

```
#midsize and compact cars,improves their gas mileage
# Left
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy, alpha = class))
```
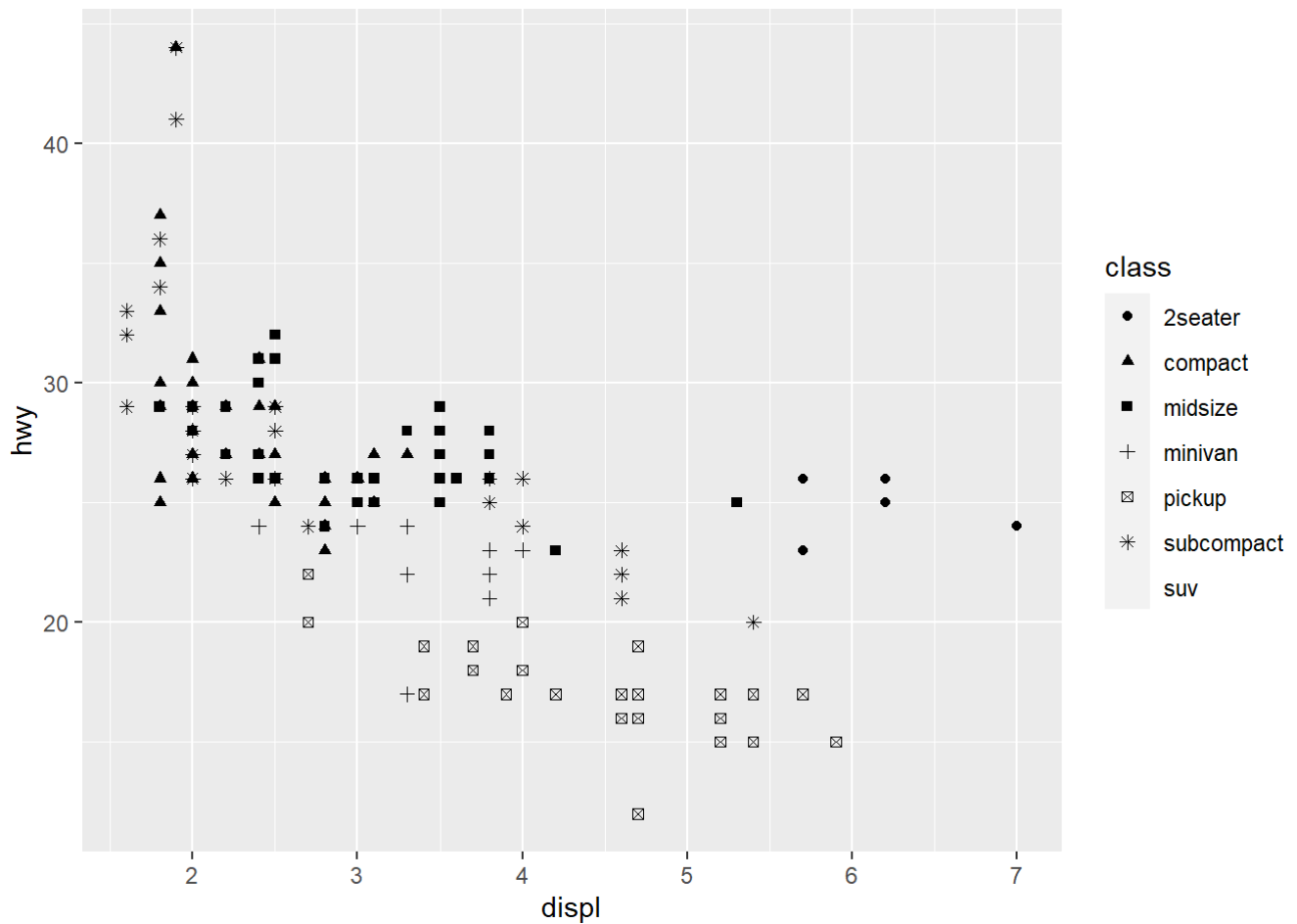
```
## Warning: Using alpha for a discrete variable is not advised.
```
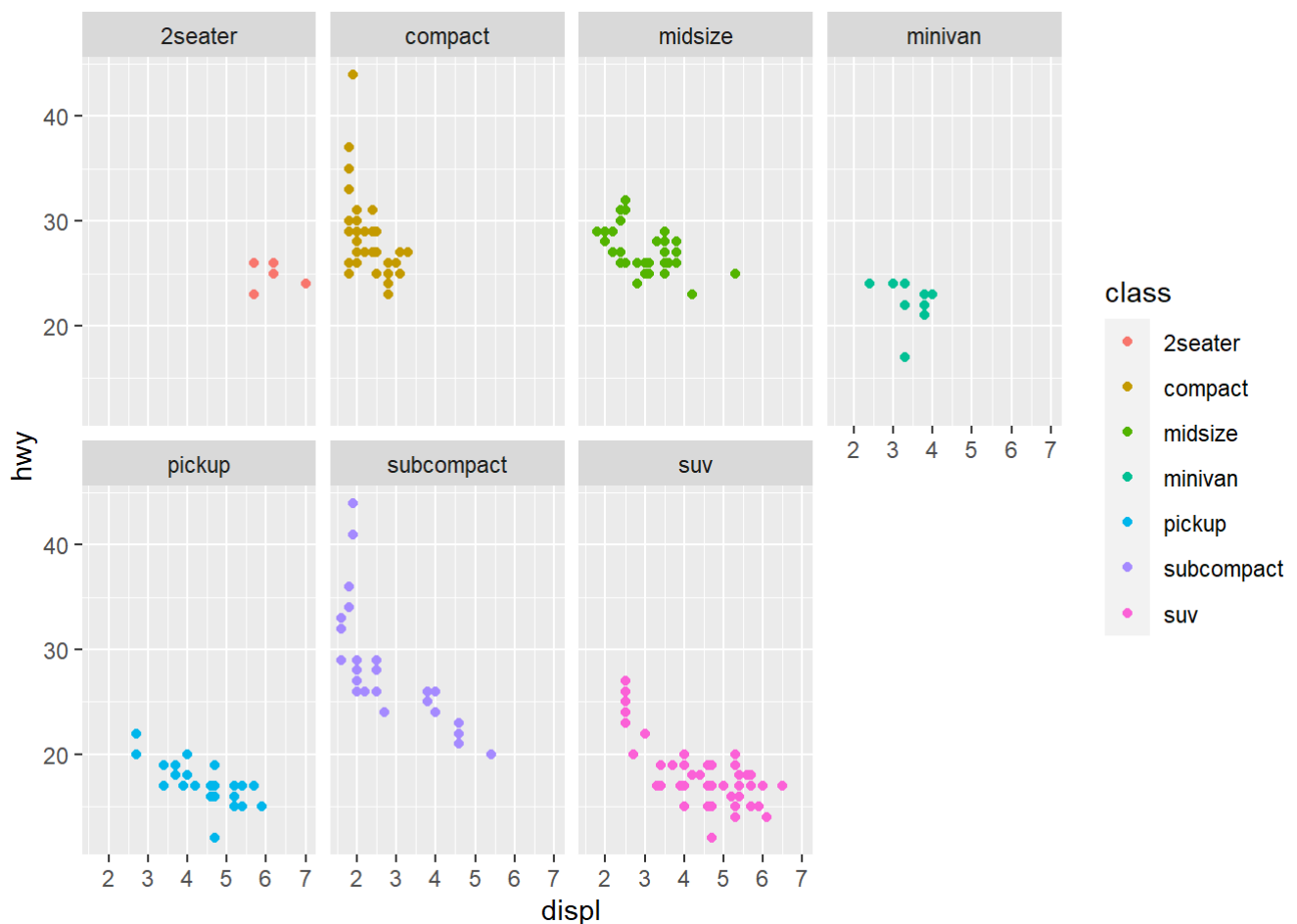
```
# Right
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy, shape = class))
```

```
## Warning: The shape palette can deal with a maximum of 6 discrete values because
## more than 6 becomes difficult to discriminate; you have 7. Consider
## specifying shapes manually if you must have them.
```

```
## Warning: Removed 62 rows containing missing values (geom_point).
```
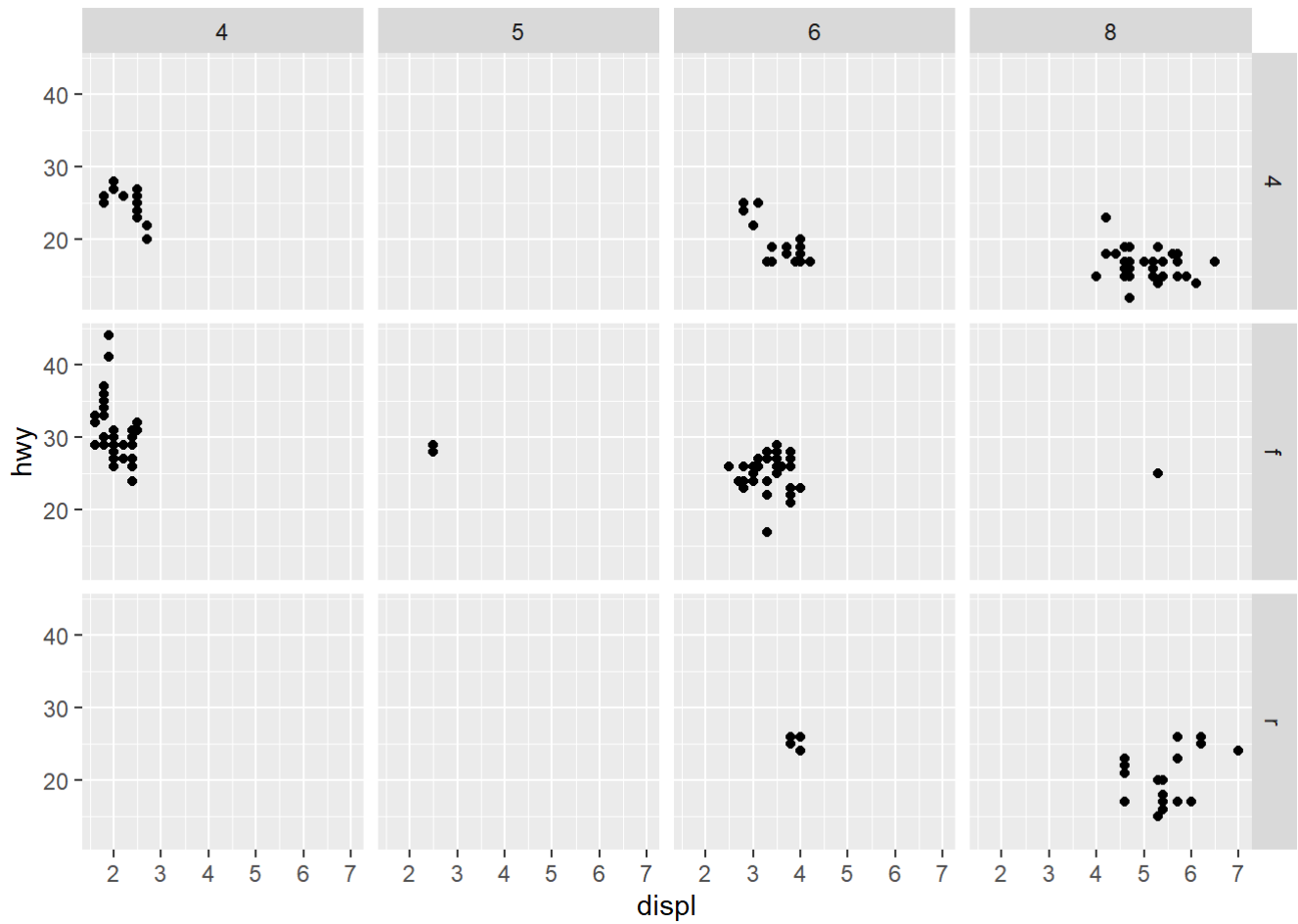
```
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy, color=class)) +  facet_wrap
(~ class, nrow = 2)
```
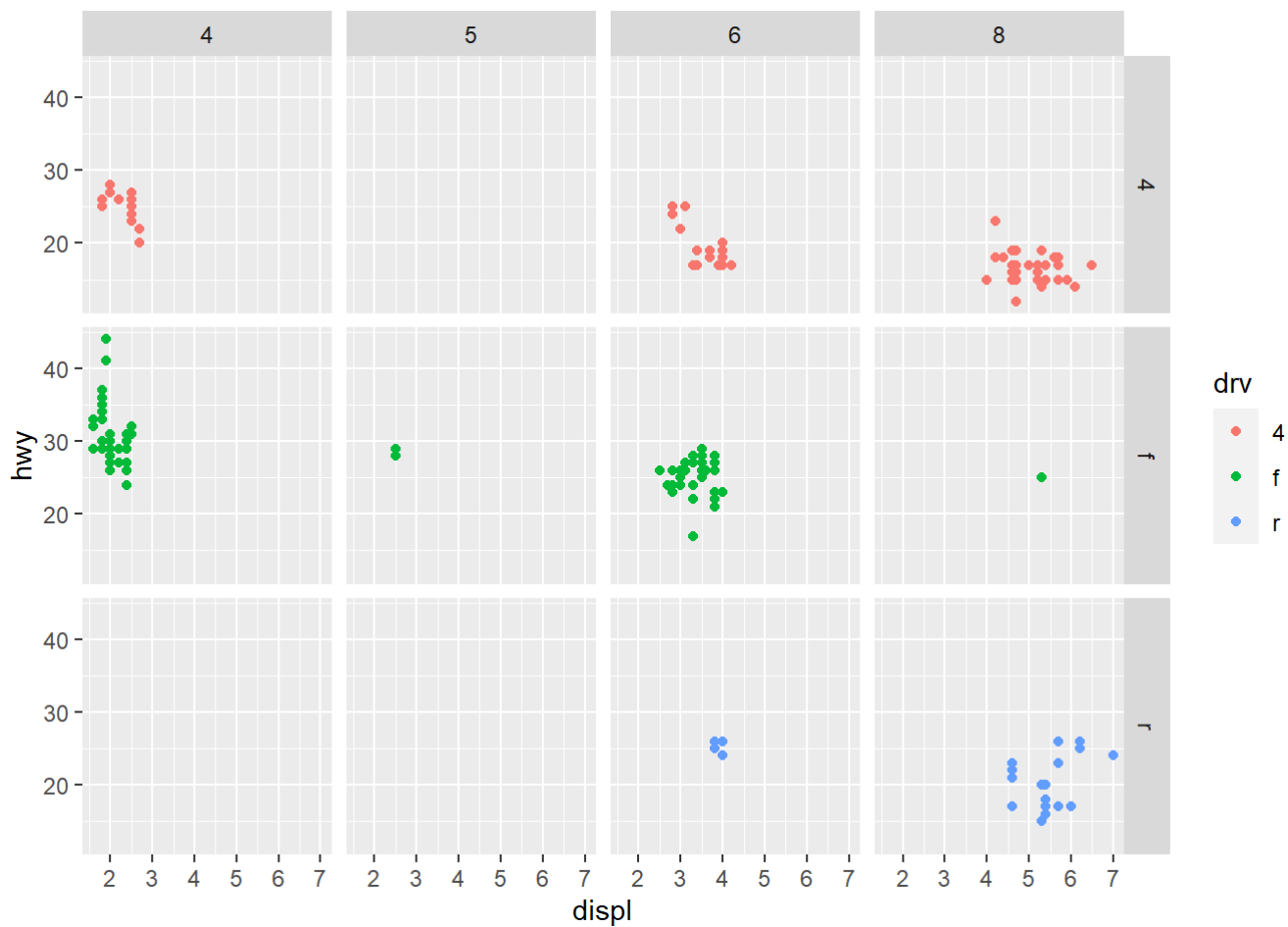
```
#shows higher the engine size lower the effiency
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy)) +facet_grid(drv ~ cyl)
```



```
ggplot(data = mpg) + geom_point(mapping = aes(x = displ, y = hwy, color=drv)) +facet_grid(drv
~ cyl)
```
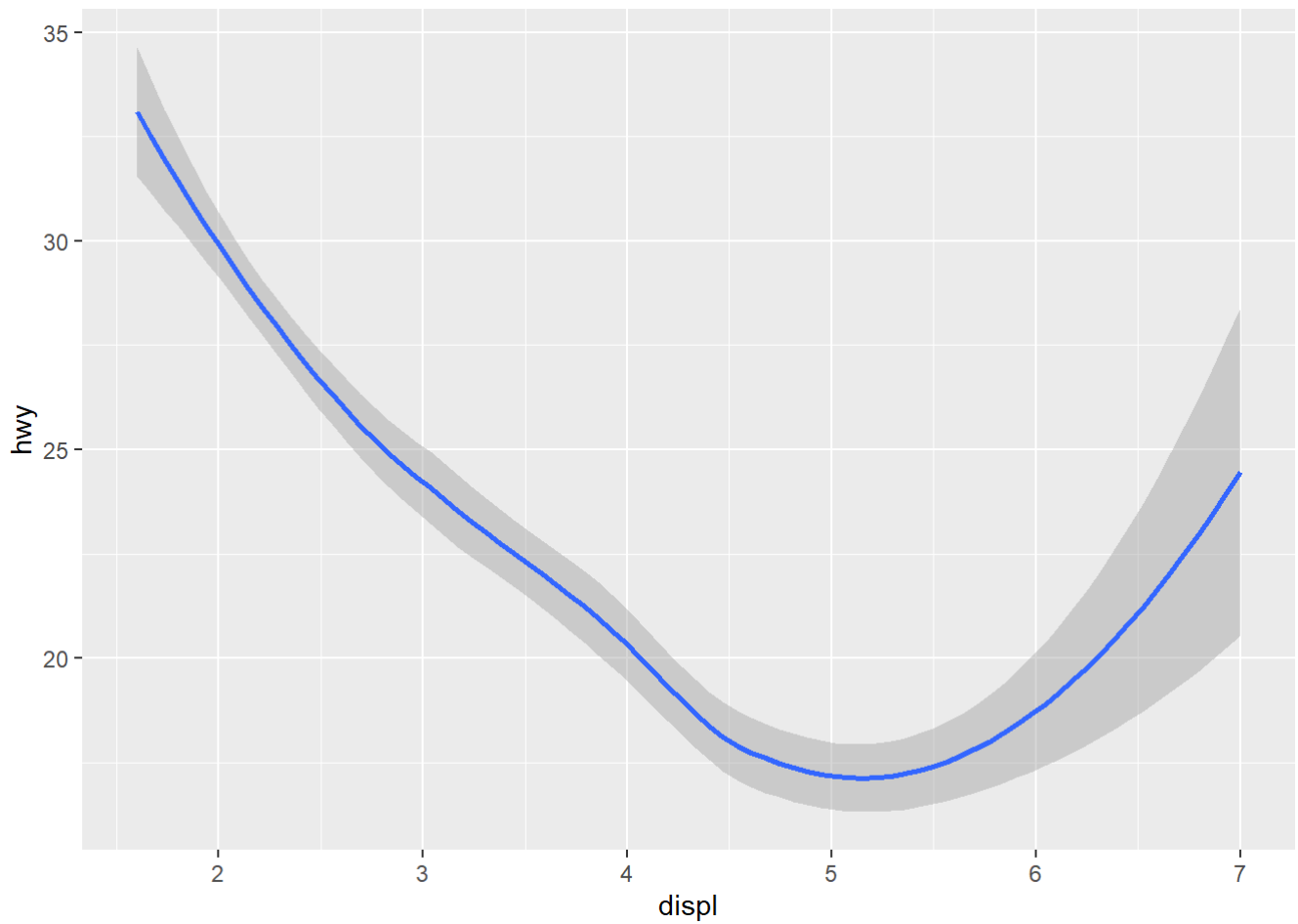
```
#the smooth geom is smooth line fitted to the data.
ggplot(data = mpg) +  geom_smooth(mapping = aes(x = displ, y = hwy))
```
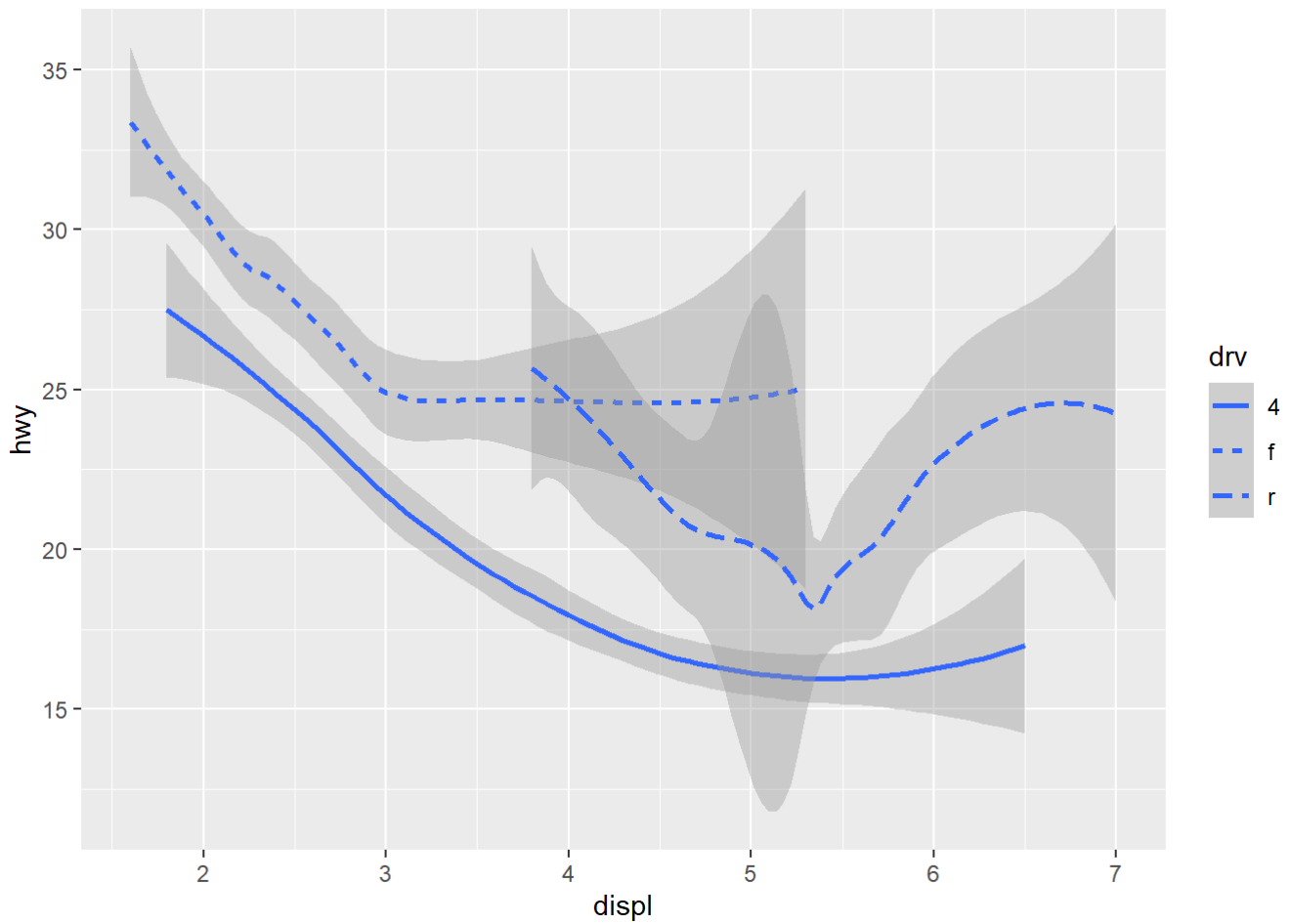
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

```
ggplot(data = mpg) + geom_smooth(mapping = aes(x = displ, y = hwy, linetype = drv))
```

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

#Here geom_smooth() separates the cars into three lines based on their drv value, which descr
ibes a car's drivetrain. One line describes all of the points with a 4 value, one line descri
bes all of the points with an f value, and one line describes all of the points with an r val
ue. Here, 4 stands for four-wheel drive, f for front-wheel drive, and r for rear-wheel drive.