Graphing Using ggplot Part-3

```
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

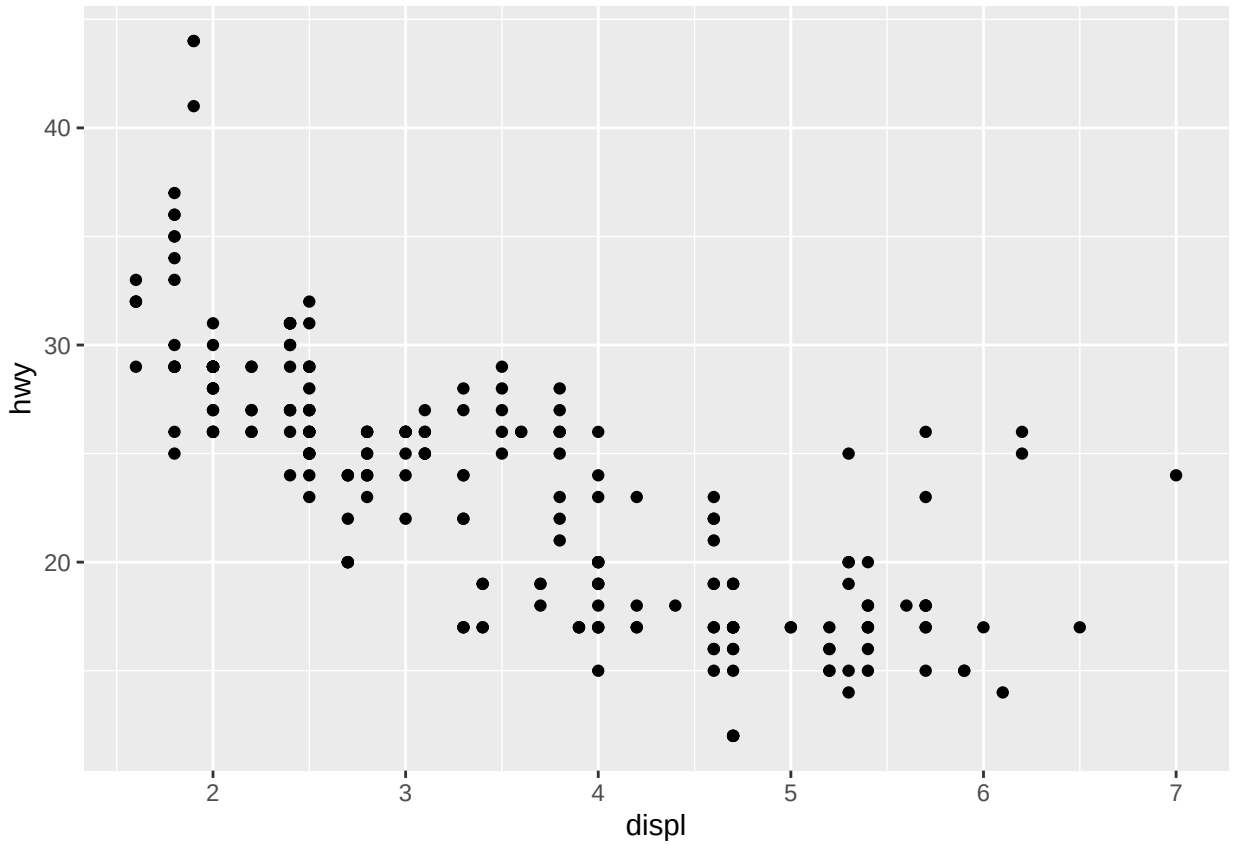
```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

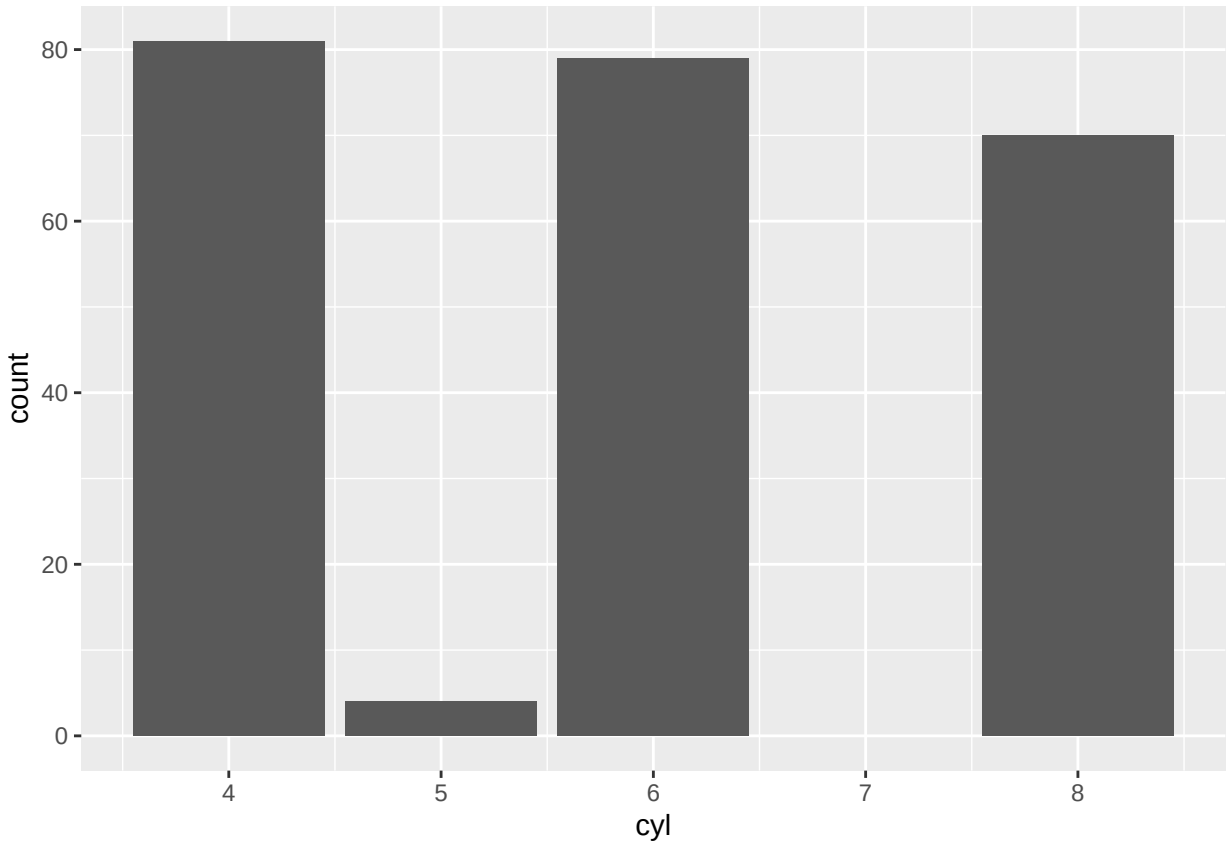
```
# Load the mpg dataset
data('mpg')
mpgData = mpg

# Print the first five rows (or samples) in the data frame
head(mpgData, 5)
```

```
## # A tibble: 5 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(l5)   f        18    29 p     compa~
## 2 audi         a4      1.8  1999     4 manual(m5) f        21    29 p     compa~
## 3 audi         a4      2    2008     4 manual(m6) f        20    31 p     compa~
## 4 audi         a4      2    2008     4 auto(av)   f        21    30 p     compa~
## 5 audi         a4      2.8  1999     6 auto(l5)   f        16    26 p     compa~
```

```
# Plot a scatter plot of mileage w.r.t. displacement
ggplot(data = mpgData) +
  geom_point(mapping = aes(x = displ, y = hwy))
```

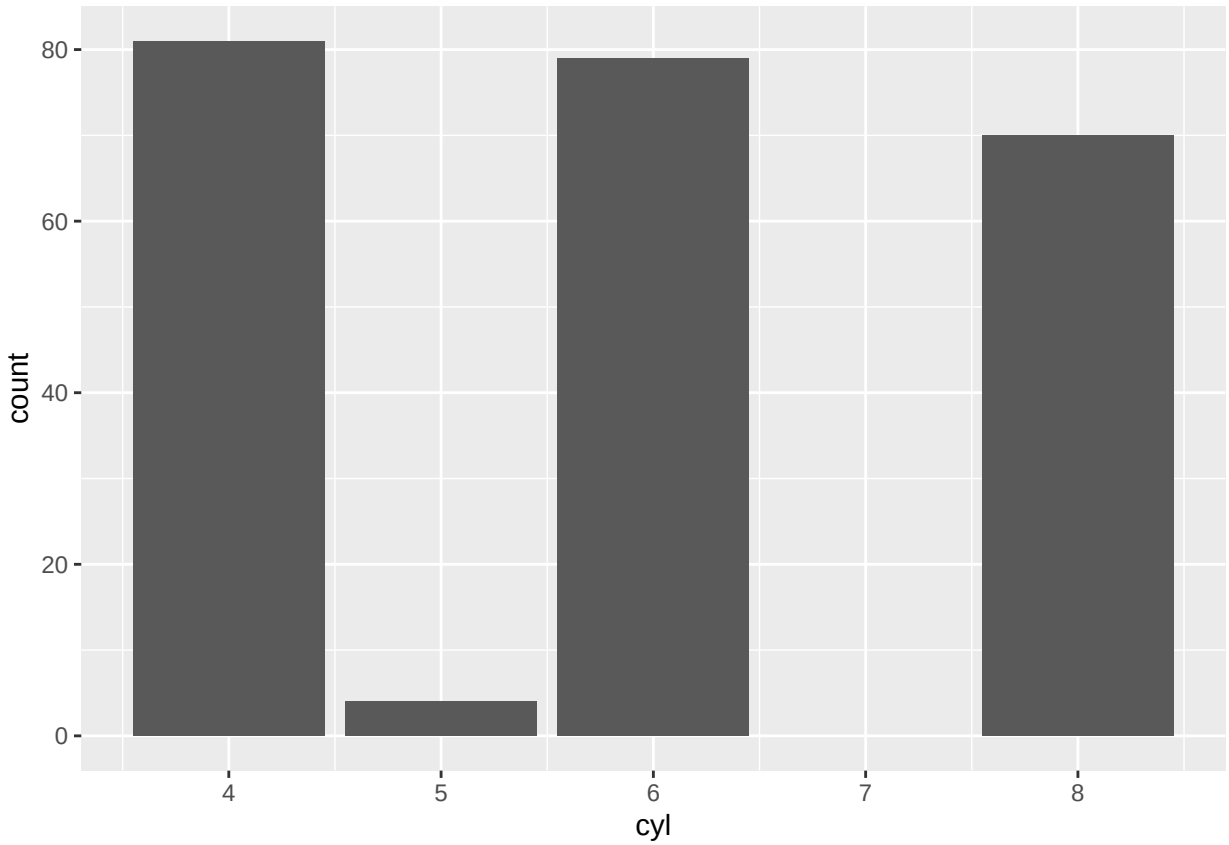



```
# Plot a bar chart w.r.t. number of cylinders
ggplot(data = mpgData) +
  geom_bar(mapping = aes(x = cyl))
```




```
#  Which stat geom_bar uses by default?
?geom_bar
```

```
# Using a stat explicitly
ggplot(data = mpgData) +
  stat_count(mapping = aes(x = cyl))
```




```
mpgData$cyl
```

```
##   [1] 4 4 4 4 6 6 6 4 4 4 4 6 6 6 6 6 6 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 4 4 6 6 6
##  [38] 4 6 6 6 6 6 6 6 6 6 6 6 6 6 6 8 8 8 8 8 6 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
##  [75] 8 8 8 6 6 6 6 8 8 6 6 8 8 8 8 8 6 6 6 6 8 8 8 8 8 4 4 4 4 4 4 4 4 4 4 4 4
## [112] 4 6 6 6 4 4 4 4 6 6 6 6 6 6 8 8 8 8 8 8 8 8 8 8 8 8 6 6 8 8 4 4 4 4 6 6 6
## [149] 6 6 6 6 6 8 6 6 6 6 8 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 6 6 6 8 4 4 4 4 6 6
## [186] 6 4 4 4 4 6 6 6 4 4 4 4 4 8 8 4 4 4 6 6 6 6 4 4 4 4 6 4 4 4 4 4 5 5 6 6 4
## [223] 4 4 4 5 5 4 4 4 4 6 6 6
```

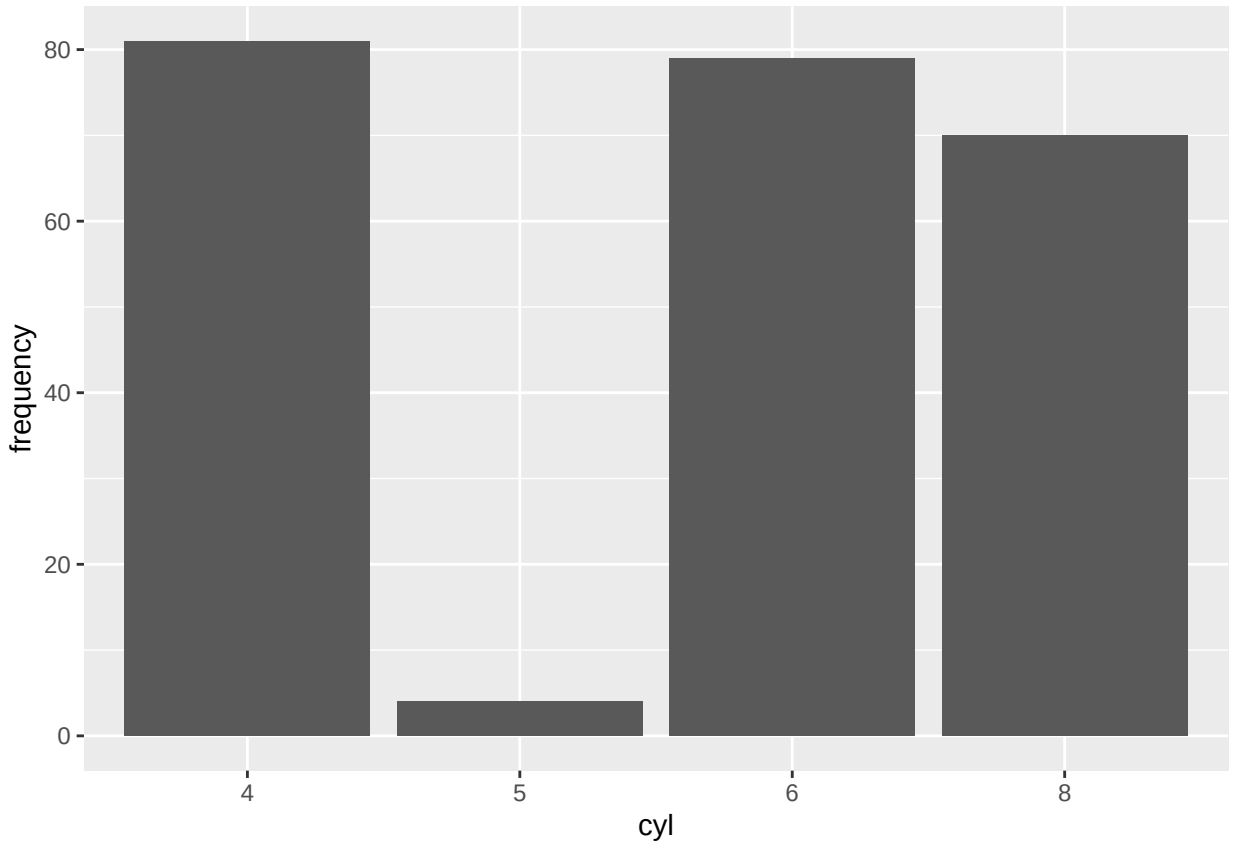
```
table(mpgData$cyl) # Shows the frequency table
```

```
##
##  4  5  6  8
## 81  4 79 70
```

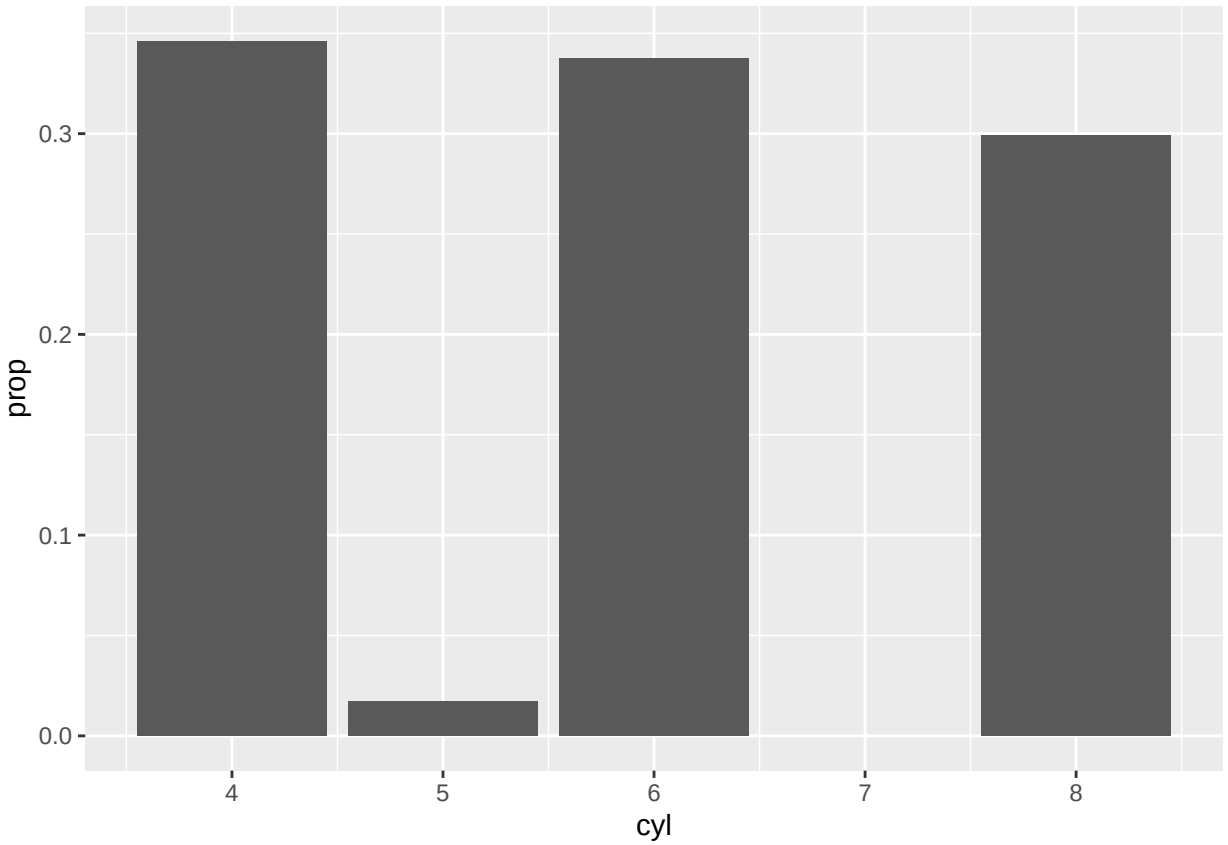
```
tempData = data.frame(table(mpgData$cyl))
colnames(tempData) = c('cyl', 'frequency')
tempData
```

```
##   cyl frequency
## 1   4        81
## 2   5         4
## 3   6        79
## 4   8        70
```

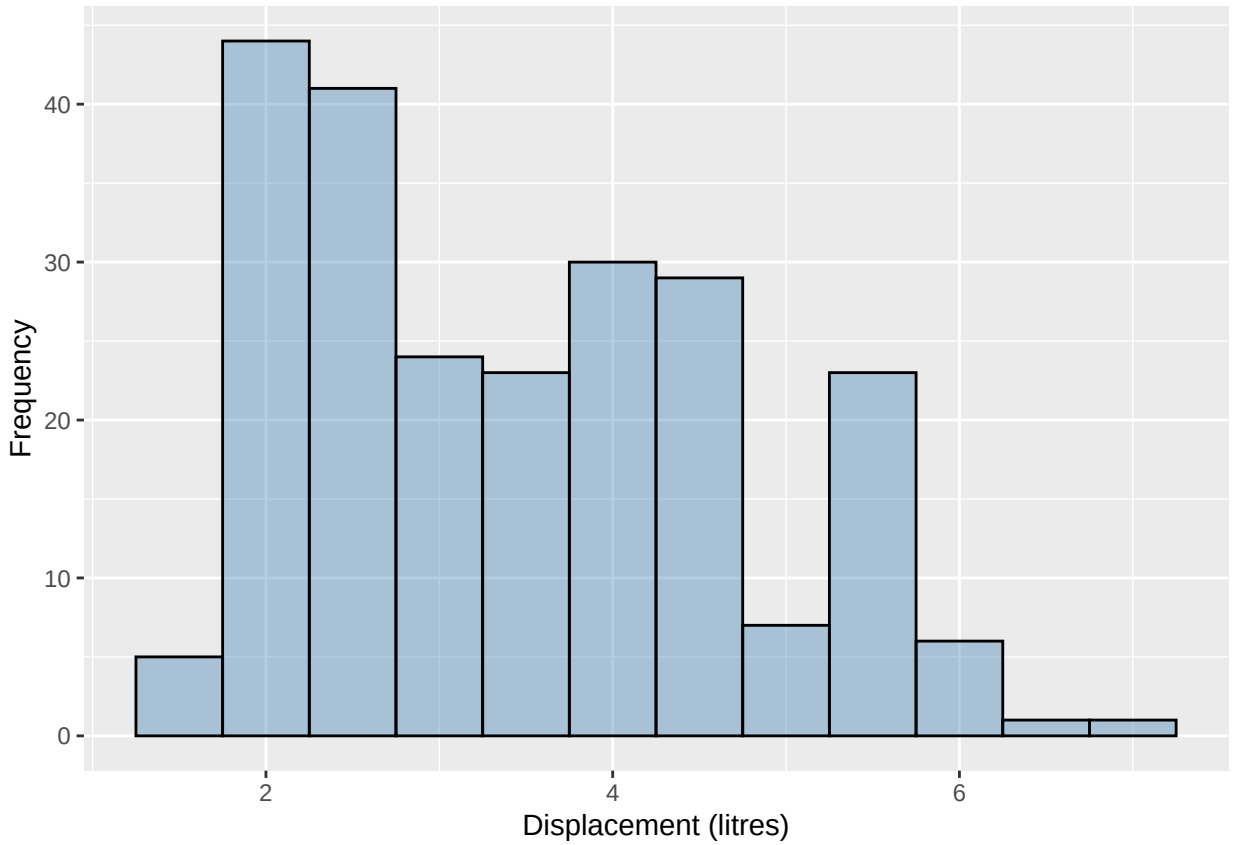
```
# Plot a bar chart w.r.t. number of cylinders directly
ggplot(data = tempData) +
  geom_bar(mapping = aes(x = cyl, y = frequency), stat = 'identity')
```




```
#  Plot a bar chart of proportion, rather than count w.r.t. number of cylinders
ggplot(data = mpgData) +
  geom_bar(mapping = aes(x = cyl, y = after_stat(prop))) # The dot-dot notation (`..prop..`) was deprec
```

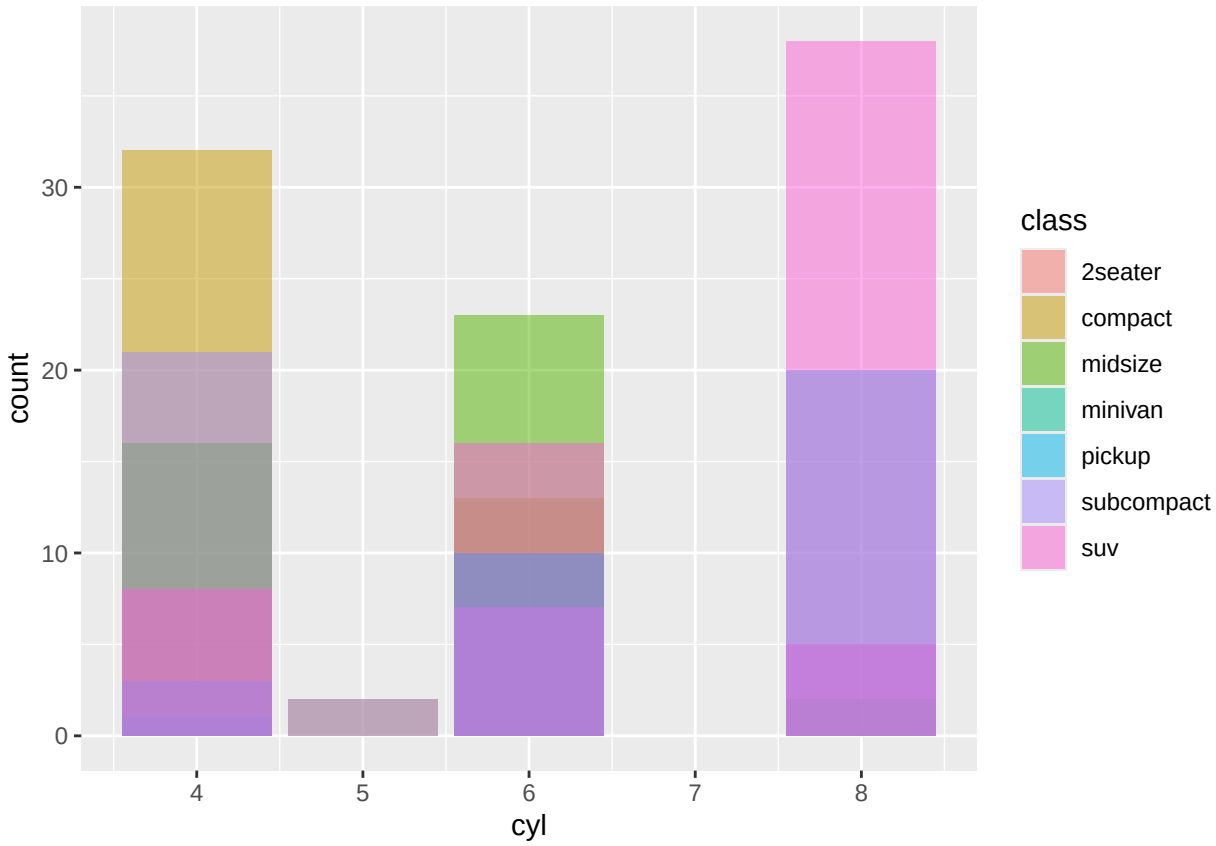



```
# Plot a histogram of engine displacement
ggplot(data = mpgData) +
  geom_histogram(aes(x = displ, y = after_stat(count)), binwidth = 0.5, color = 'black', fill = 'steelbl
  labs(x = 'Displacement (litres)', y = 'Frequency')
```




```
#The dot-dot notation (`..count..`) was deprecated in ggplot2 3.4.0.
```

```
# Plot a bar chart w.r.t. number of cylinders and class using color and fill
ggplot(data = mpgData) +
  geom_bar(mapping = aes(x = cyl, fill = class), position = 'identity', alpha = .5)
```


30
20
10
0
count
4
5
6
7
8
cyl
class
2seater
compact
midsize
minivan
pickup
subcompact
suv