

WORKSHEET -6

Mark Wryzel Macarobo BSIT-2A

12/23/2022

```
library(ggplot2) data(mpg) as.data.frame(data(mpg)) data(mpg) mpg  
str(mpg) library(dplyr) glimpse(mpg)
```

1.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.2
```

```
data(mpg)  
as.data.frame(data(mpg))
```

```
## data(mpg)  
## 1 mpg
```

```
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.2
```

```
##  
## 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
```

```
datampg <- glimpse(mpg)
```

```
## Rows: 234  
## Columns: 11  
## $ manufacturer <chr> "audi", "audi", "audi", "audi", "audi", "audi", "audi", "~  
## $ model <chr> "a4", "a4", "a4", "a4", "a4", "a4", "a4", "a4 quattro", "~  
## $ displ <dbl> 1.8, 1.8, 2.0, 2.0, 2.8, 2.8, 3.1, 1.8, 1.8, 2.0, 2.0, 2.~
```

```
## $ year      <int> 1999, 1999, 2008, 2008, 1999, 1999, 2008, 1999, 1999, 200~
## $ cyl       <int> 4, 4, 4, 4, 6, 6, 6, 4, 4, 4, 4, 6, 6, 6, 6, 6, 6, 8, 8, ~
## $ trans     <chr> "auto(l5)", "manual(m5)", "manual(m6)", "auto(av)", "auto~
## $ drv       <chr> "f", "f", "f", "f", "f", "f", "f", "f", "4", "4", "4", "4", "4~
## $ cty       <int> 18, 21, 20, 21, 16, 18, 18, 18, 16, 20, 19, 15, 17, 17, 1~
## $ hwy       <int> 29, 29, 31, 30, 26, 26, 27, 26, 25, 28, 27, 25, 25, 25, 2~
## $ fl        <chr> "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p~
## $ class     <chr> "compact", "compact", "compact", "compact", "compact", "c~
```

```
ncol(mpg)
```

```
## [1] 11
```

```
nrow(mpg)
```

```
## [1] 234
```

2.

```
total <- mpg %>%
  group_by(manufacturer) %>%
  tally(sort = TRUE)
```

2.a

```
datampg <- glimpse(mpg)
```

```
## Rows: 234
## Columns: 11
## $ manufacturer <chr> "audi", "audi", "audi", "audi", "audi", "audi", "audi", "~
## $ model        <chr> "a4", "a4", "a4", "a4", "a4", "a4", "a4", "a4 quattro", "~
## $ displ       <dbl> 1.8, 1.8, 2.0, 2.0, 2.8, 2.8, 3.1, 1.8, 1.8, 2.0, 2.0, 2.~
## $ year        <int> 1999, 1999, 2008, 2008, 1999, 1999, 2008, 1999, 1999, 200~
## $ cyl         <int> 4, 4, 4, 4, 6, 6, 6, 4, 4, 4, 4, 6, 6, 6, 6, 6, 8, 8, ~
## $ trans       <chr> "auto(l5)", "manual(m5)", "manual(m6)", "auto(av)", "auto~
## $ drv         <chr> "f", "f", "f", "f", "f", "f", "f", "f", "4", "4", "4", "4", "4~
## $ cty         <int> 18, 21, 20, 21, 16, 18, 18, 18, 16, 20, 19, 15, 17, 17, 1~
## $ hwy         <int> 29, 29, 31, 30, 26, 26, 27, 26, 25, 28, 27, 25, 25, 25, 2~
## $ fl          <chr> "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p", "p~
## $ class       <chr> "compact", "compact", "compact", "compact", "compact", "c~
```

```
unique <- datampg %>% group_by(manufacturer, model) %>%
  distinct() %>% count()
unique
```

```
## # A tibble: 38 x 3
## # Groups:   manufacturer, model [38]
##   manufacturer model      n
##   <chr>         <chr>    <int>
## 1 audi         a4          7
```

```
## 2 audi      a4 quattro      8
## 3 audi      a6 quattro      3
## 4 chevrolet c1500 suburban 2wd 4
## 5 chevrolet corvette       5
## 6 chevrolet k1500 tahoe 4wd  4
## 7 chevrolet malibu         5
## 8 dodge     caravan 2wd     9
## 9 dodge     dakota pickup 4wd 8
## 10 dodge    durango 4wd     6
## # ... with 28 more rows
```

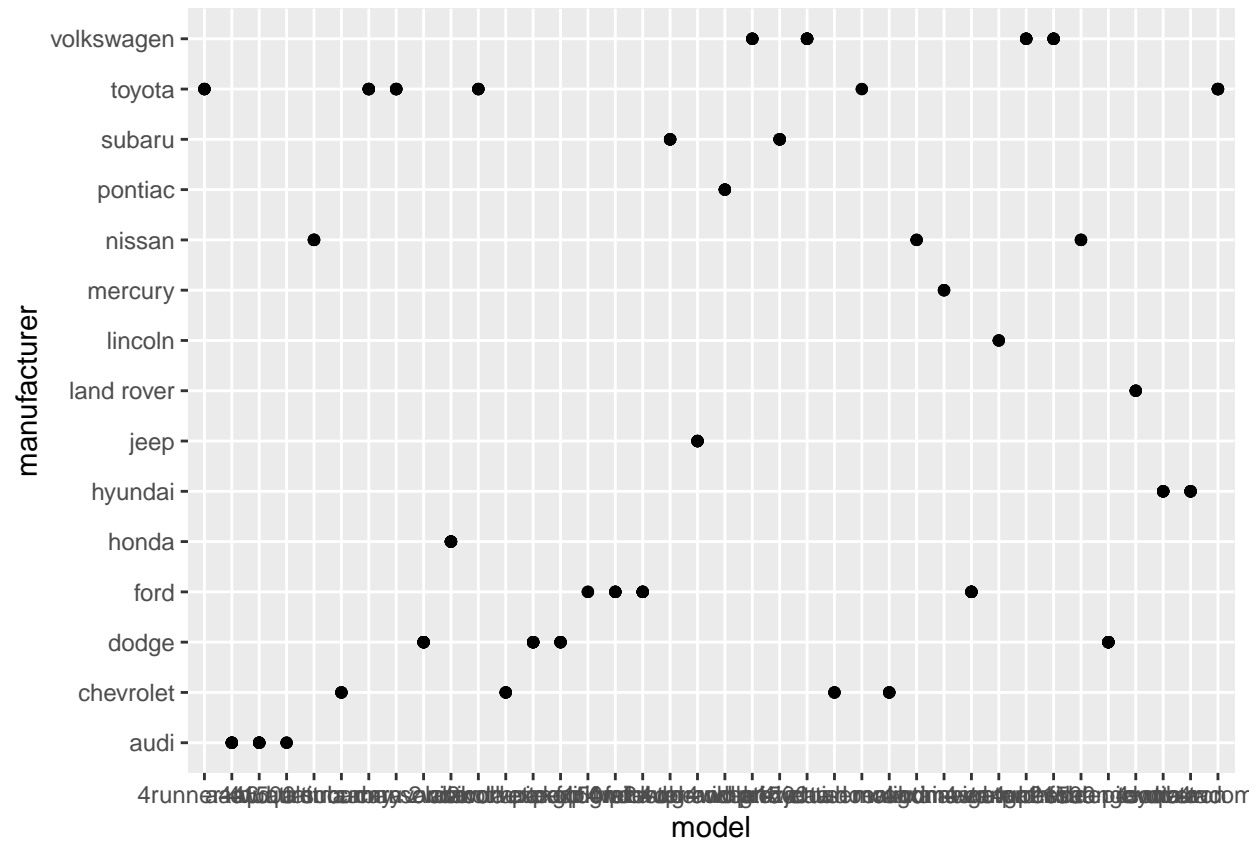
```
colnames(unique) <- c("Manufacturer", "Model", "Counts")
unique
```

```
## # A tibble: 38 x 3
## # Groups:   Manufacturer, Model [38]
##   Manufacturer Model      Counts
##   <chr>         <chr>      <int>
## 1 audi         a4           7
## 2 audi         a4 quattro    8
## 3 audi         a6 quattro    3
## 4 chevrolet    c1500 suburban 2wd 4
## 5 chevrolet    corvette      5
## 6 chevrolet    k1500 tahoe 4wd  4
## 7 chevrolet    malibu        5
## 8 dodge        caravan 2wd     9
## 9 dodge        dakota pickup 4wd 8
## 10 dodge       durango 4wd     6
## # ... with 28 more rows
```

2.b

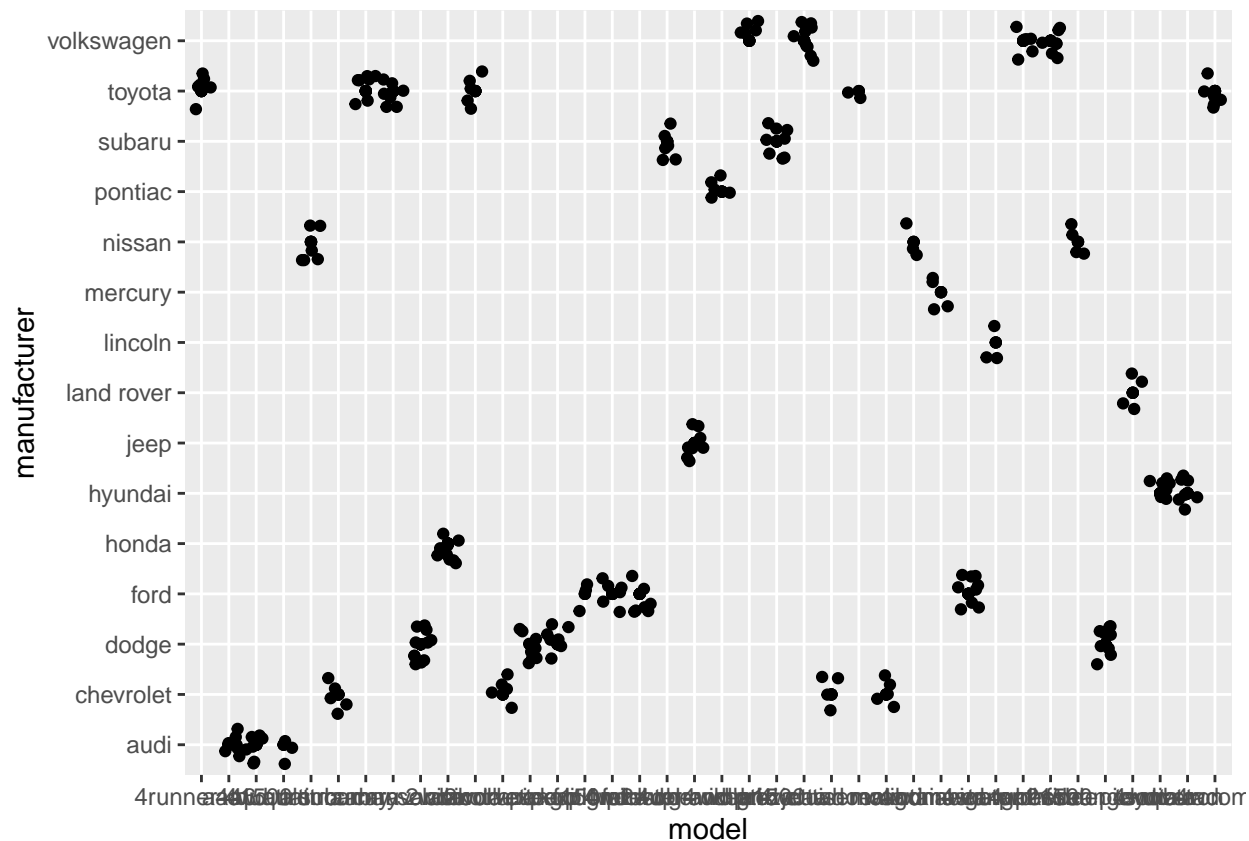
```
qplot(model, data = mpg, geom = "bar", fill = manufacturer)
```

```
## Warning: 'qplot()' was deprecated in ggplot2 3.4.0.
```

3.a

```
ggplot(mpg, aes(model, manufacturer)) + geom_point()
```

4.

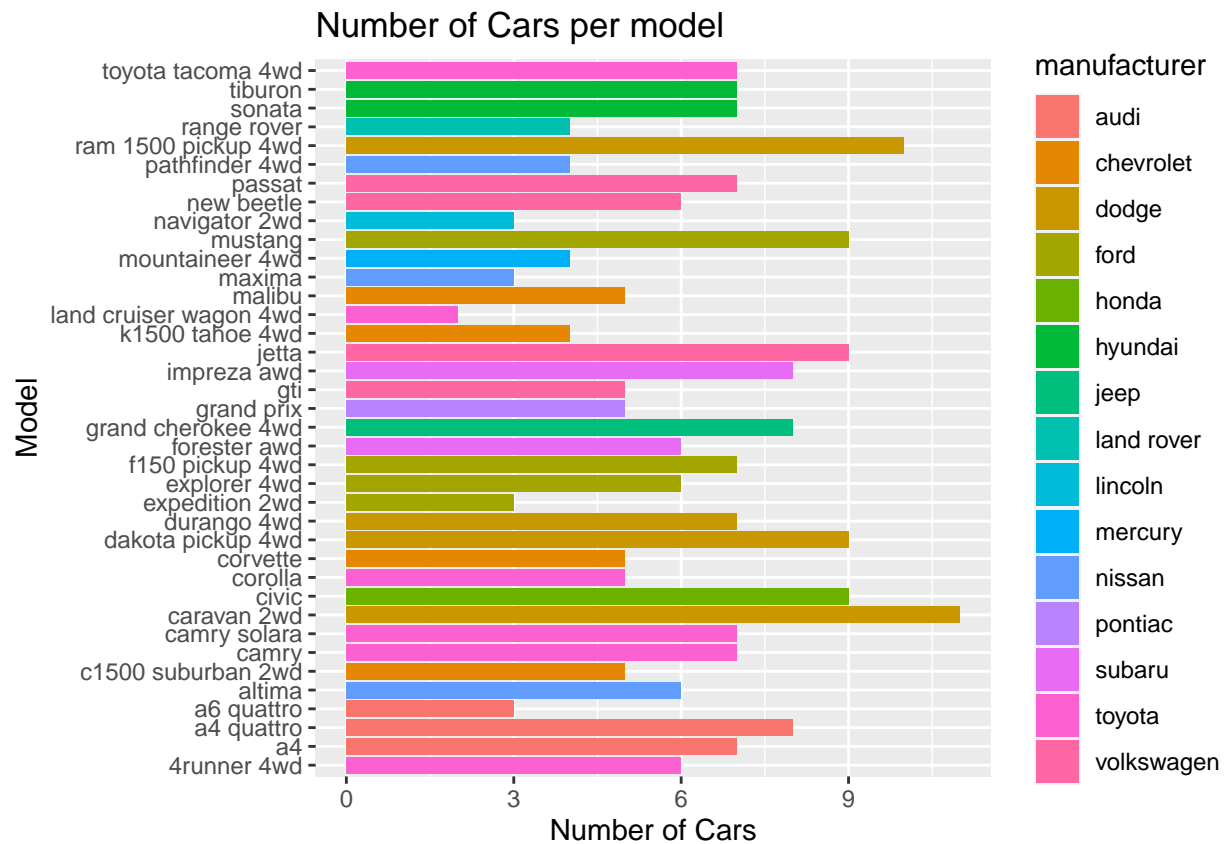
```
datampg <- unique %>% group_by(Model) %>% count()
datampg
```

```
## # A tibble: 38 x 2
## # Groups:   Model [38]
##   Model          n
##   <chr>      <int>
## 1 4runner 4wd         1
## 2 a4                 1
## 3 a4 quattro         1
## 4 a6 quattro         1
## 5 altima             1
## 6 c1500 suburban 2wd 1
## 7 camry             1
## 8 camry solara       1
## 9 caravan 2wd        1
## 10 civic             1
## # ... with 28 more rows
```

```
colnames(datampg) <- c("Model", "Counts")
```

4.a

```
qplot(model, data = mpg, main = "Number of Cars per model", xlab = "Model",
      ylab = "Number of Cars",
      geom = "bar", fill = manufacturer) + coord_flip()
```

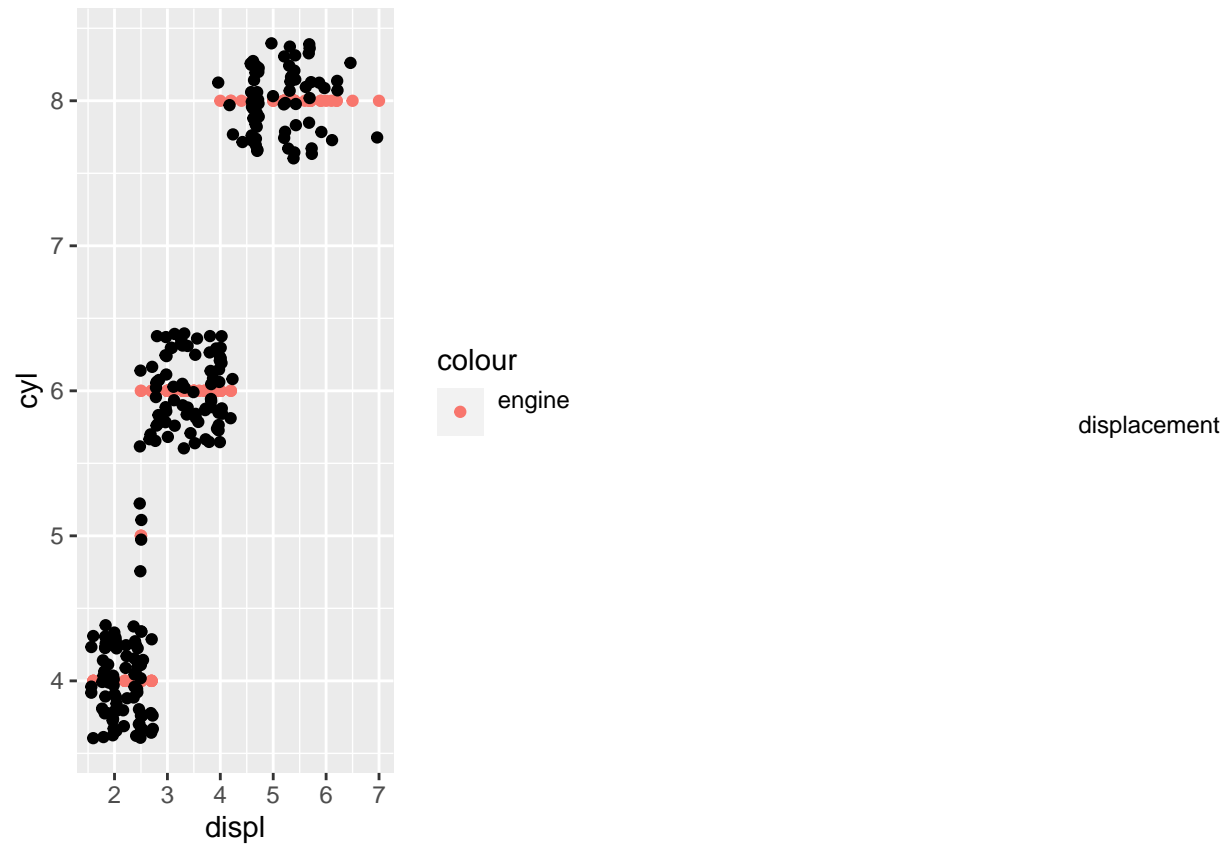


4.b

```
modelcars <- mpg %>%
  group_by(model) %>%
  tally(sort = TRUE)
```

5.a

```
ggplot(data = mpg, mapping = aes(x = displ, y = cyl, main = "Relationship between No of Cylinders and
                                Engine Displacement")) + geom_point(mapping = aes(colour = "engine
                                displacement"))
```

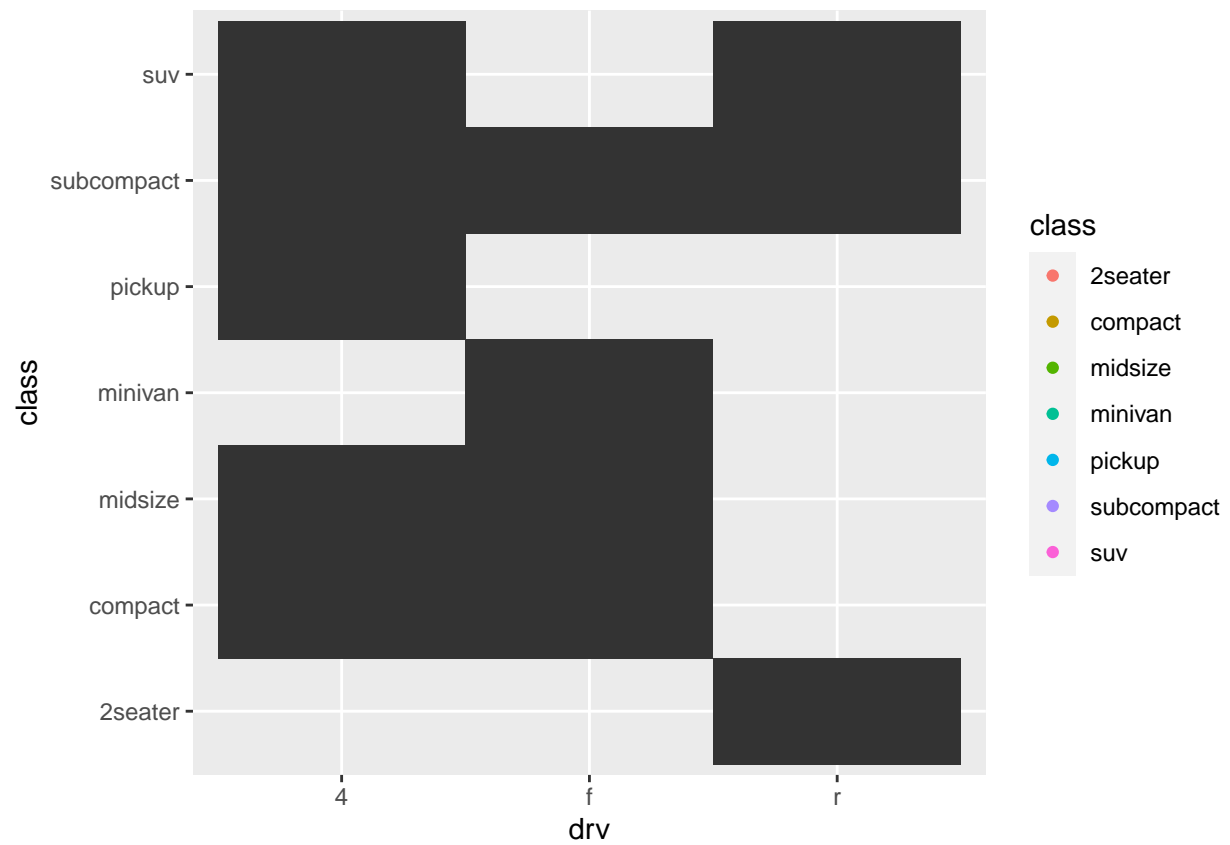



5.b

#Answer: The graph is jittered. The pink horizontal dots are the engine displacements.

6.a

```
ggplot(data = mpg, mapping = aes(x = drv, y = class)) + geom_point(mapping=aes(color=class)) +  
  geom_tile()
```

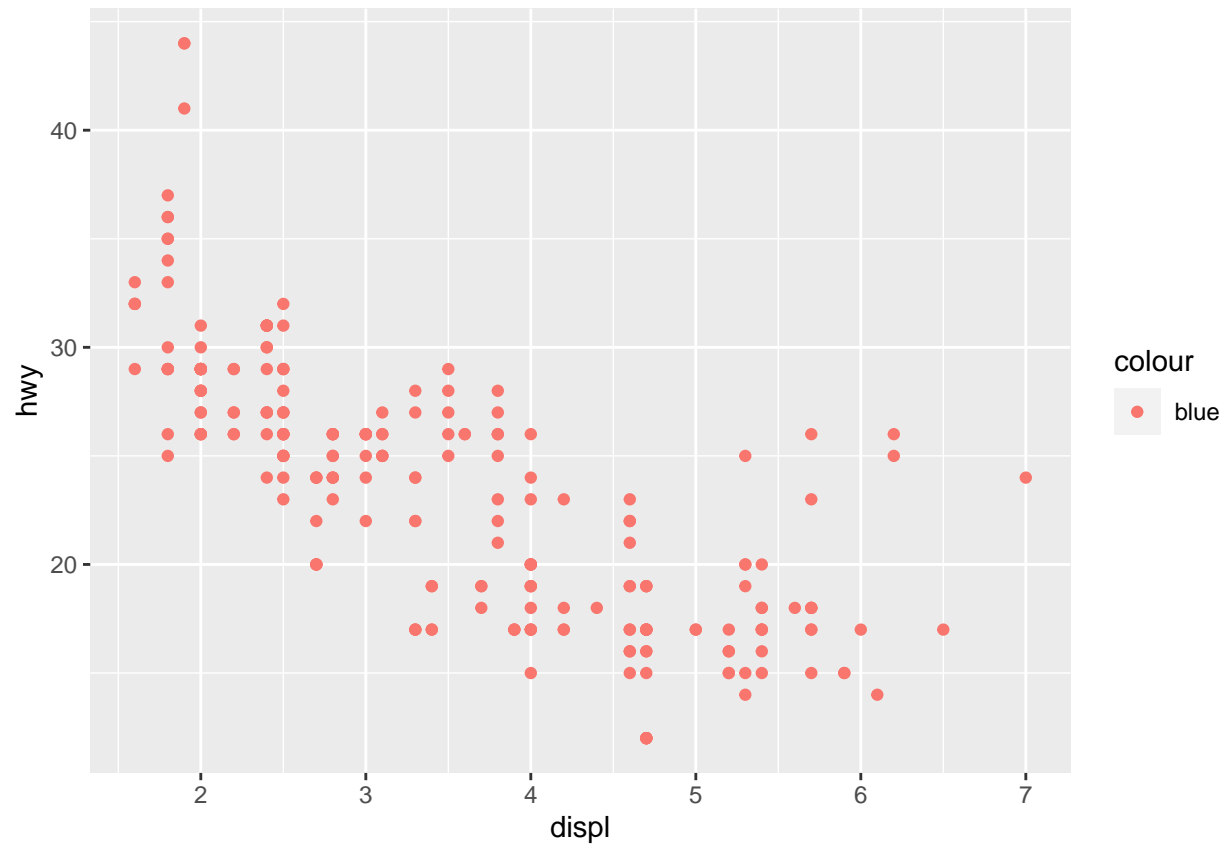


6.b

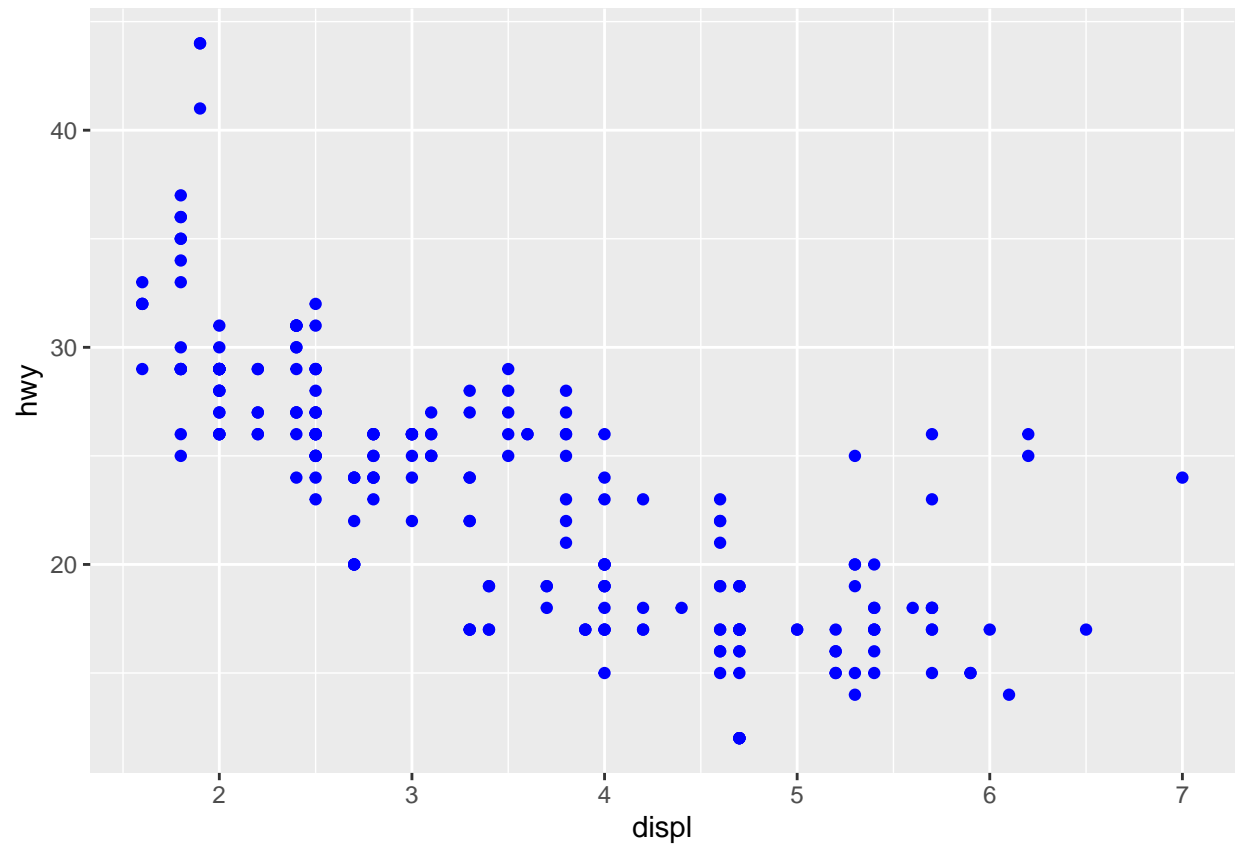
#Answer: Areas that are black are mapped using the geometric point graph. y object is class and x object

7.

```
#Code#1
ggplot(data = mpg) +
  geom_point(mapping = aes(x = displ, y = hwy, colour = "blue"))
```



```
#Code#2  
ggplot(data = mpg) +  
  geom_point(mapping = aes(x = displ, y = hwy), colour = "blue")
```



8.

```
?mpg
```

```
## starting httpd help server ... done
```

8.a

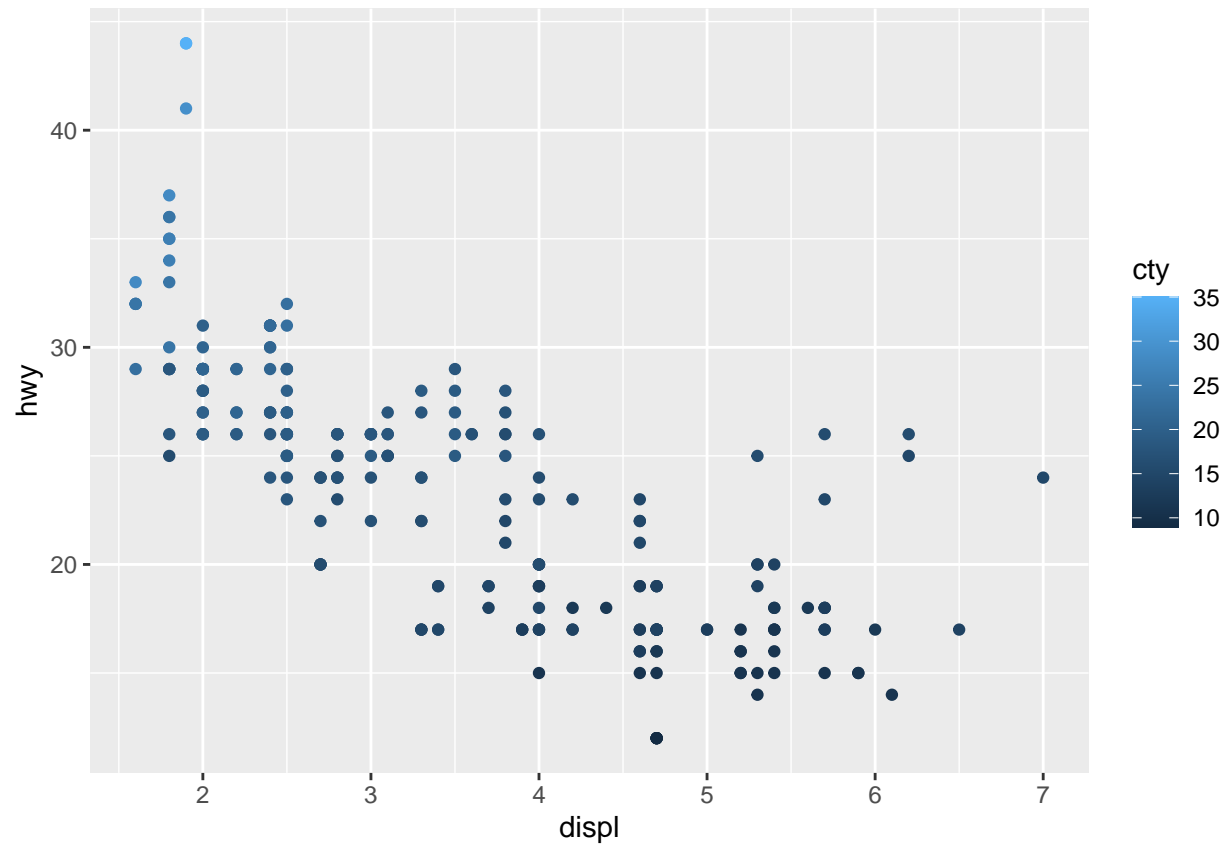
```
#Answer: manufacturer, model, trans, drv, fl, and class.
```

8.b

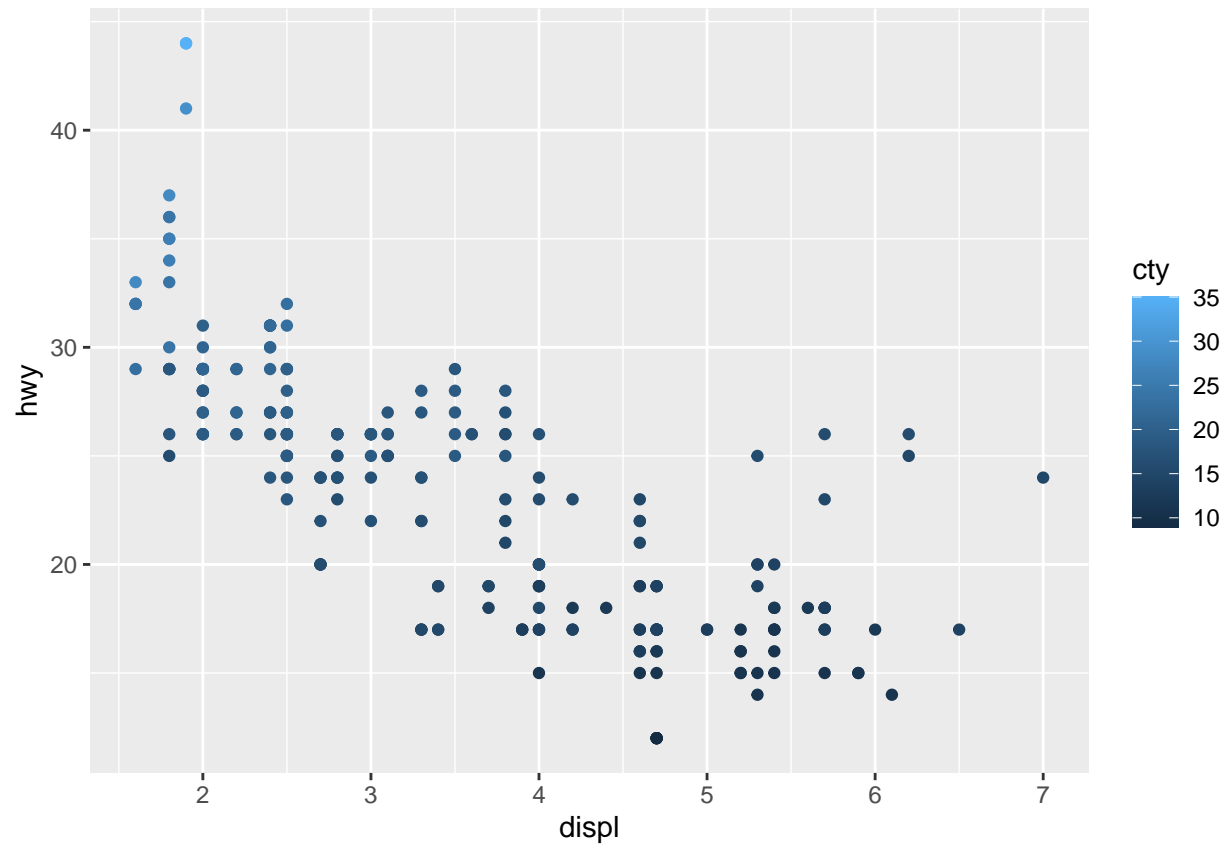
```
#Answer: They are called double or integers.
```

8.c

```
ggplot(mpg, aes(x = displ, y = hwy, colour = cty)) + geom_point()
```



```
ggplot(mpg, aes(x = displ, y = hwy, colour = cty)) + geom_point()
```



#Answer: The data monitors the cty by placing the cty in different hues of color blue.