Datacamp_Data Visualization with ggplot2 (Part 1)_Introduction

dizhen

2019/4/6/

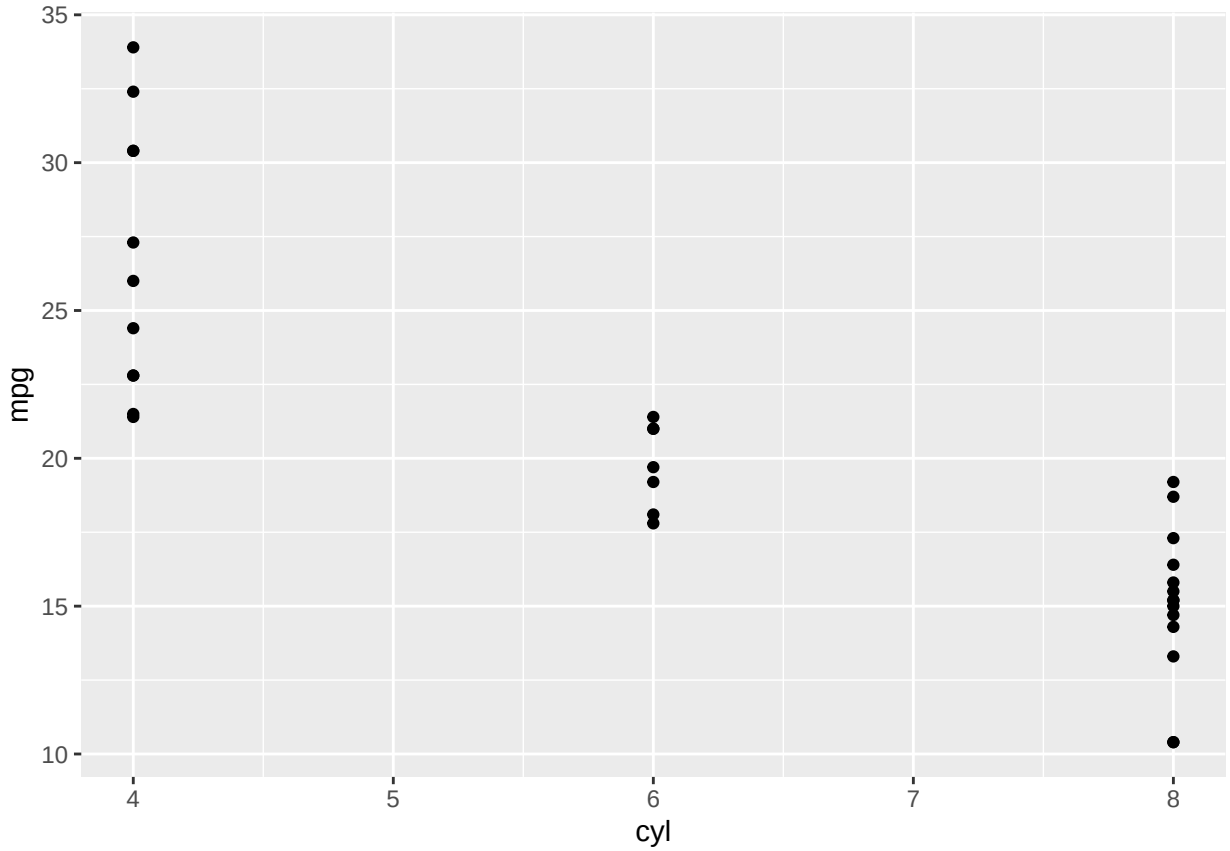
```
# Load the ggplot2 package
library(ggplot2)

# Explore the mtcars data frame with str()
str(mtcars)
```

```
## 'data.frame':    32 obs. of  11 variables:
##  $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
##  $ cyl : num  6 6 4 6 8 6 8 4 4 6 ...
##  $ disp: num  160 160 108 258 360 ...
##  $ hp  : num  110 110 93 110 175 105 245 62 95 123 ...
##  $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
##  $ wt  : num  2.62 2.88 2.32 3.21 3.44 ...
##  $ qsec: num  16.5 17 18.6 19.4 17 ...
##  $ vs  : num  0 0 1 1 0 1 0 1 1 1 ...
##  $ am  : num  1 1 1 0 0 0 0 0 0 0 ...
##  $ gear: num  4 4 4 3 3 3 3 4 4 4 ...
##  $ carb: num  4 4 1 1 2 1 4 2 2 4 ...
```

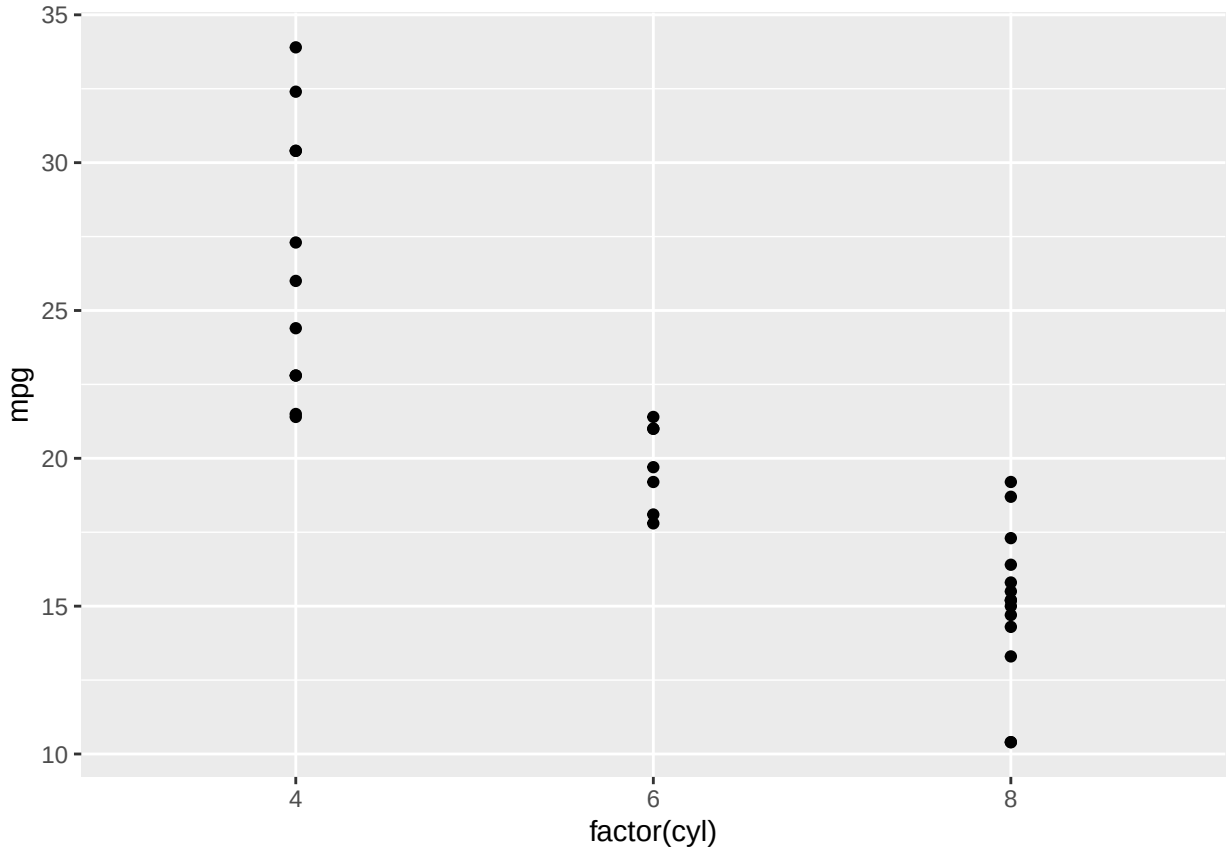
```
head(mtcars)
```

```
##                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

```
ggplot(mtcars, aes(x = cyl, y = mpg)) +
  geom_point()
```



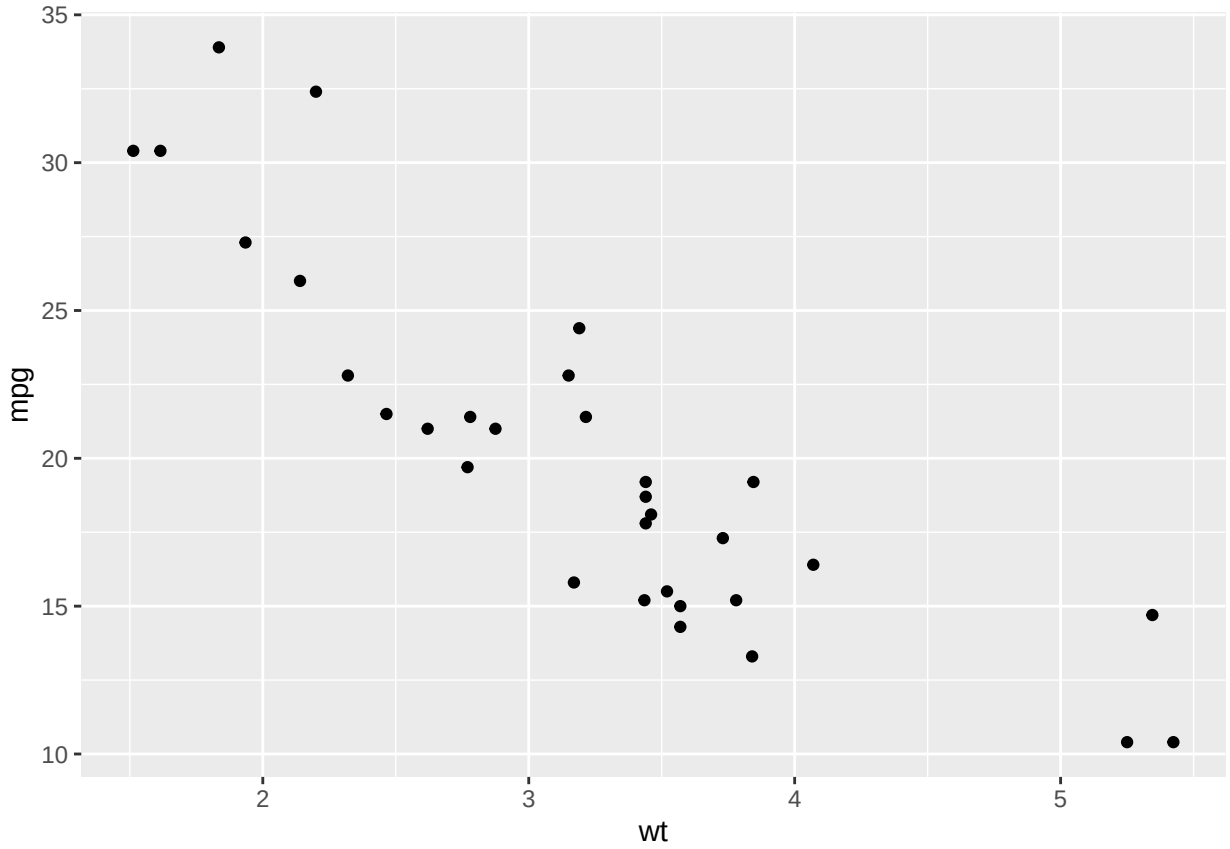

```
# cyl is treated as factor
ggplot(mtcars, aes(x = factor(cyl), y = mpg)) +
  geom_point()
```



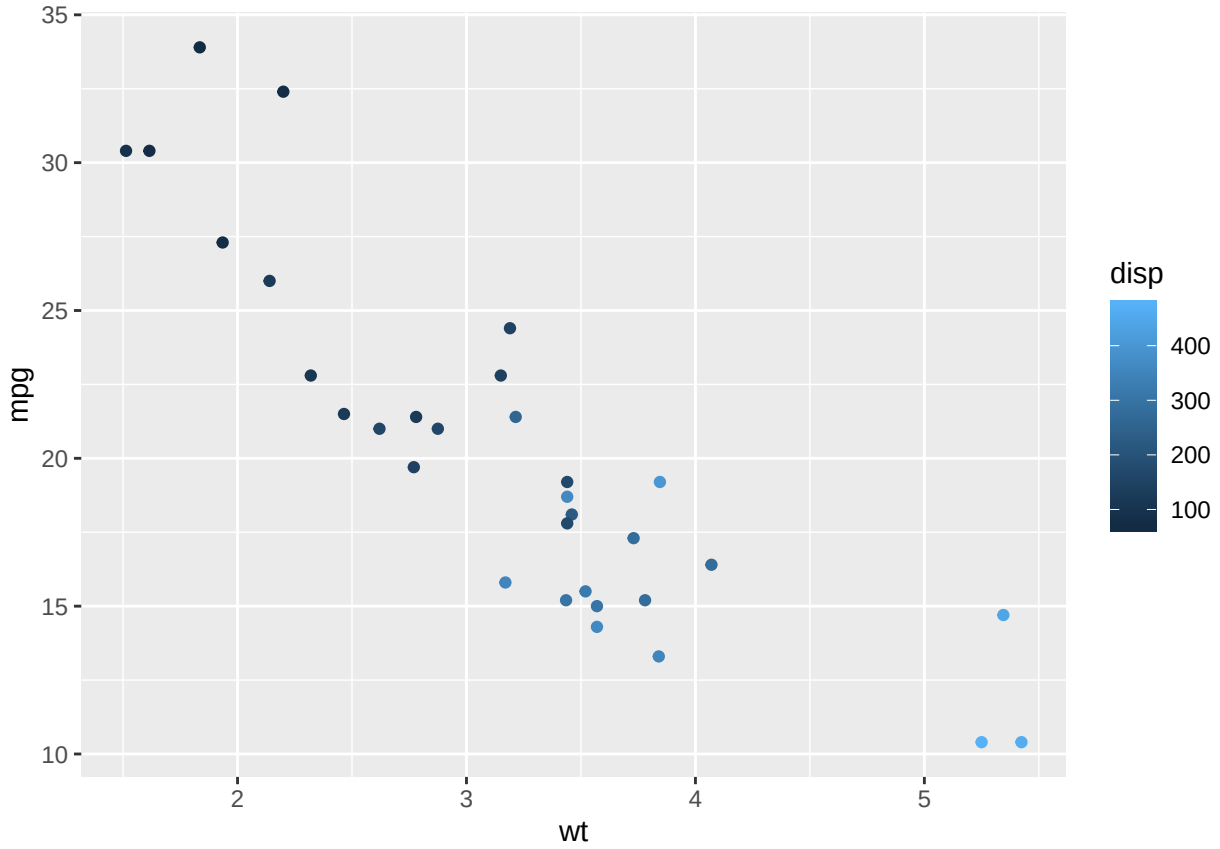
Essential Grammatic Elements:

1. Data: The dataset being plotted. (variables of interest)
2. Asethetics: The scales onto which we map our data. (x-axis, y-axis; colour, fill; size, labels; alpha, shape; line width, line type)
3. Geometries: The visual element used for out data. (point, line, histogram, bar, boxplot)
4. Facets: plotting small multiples (columns, rows)
5. Statistics: Representations of our data to aid understanding (binning, smoothing, descriptive, inferential)
6. Coordinates: The space on which the data will be plotted (cartesian, fixed, polar, limits)
7. Themes: All non-data ink. (non-data ink)

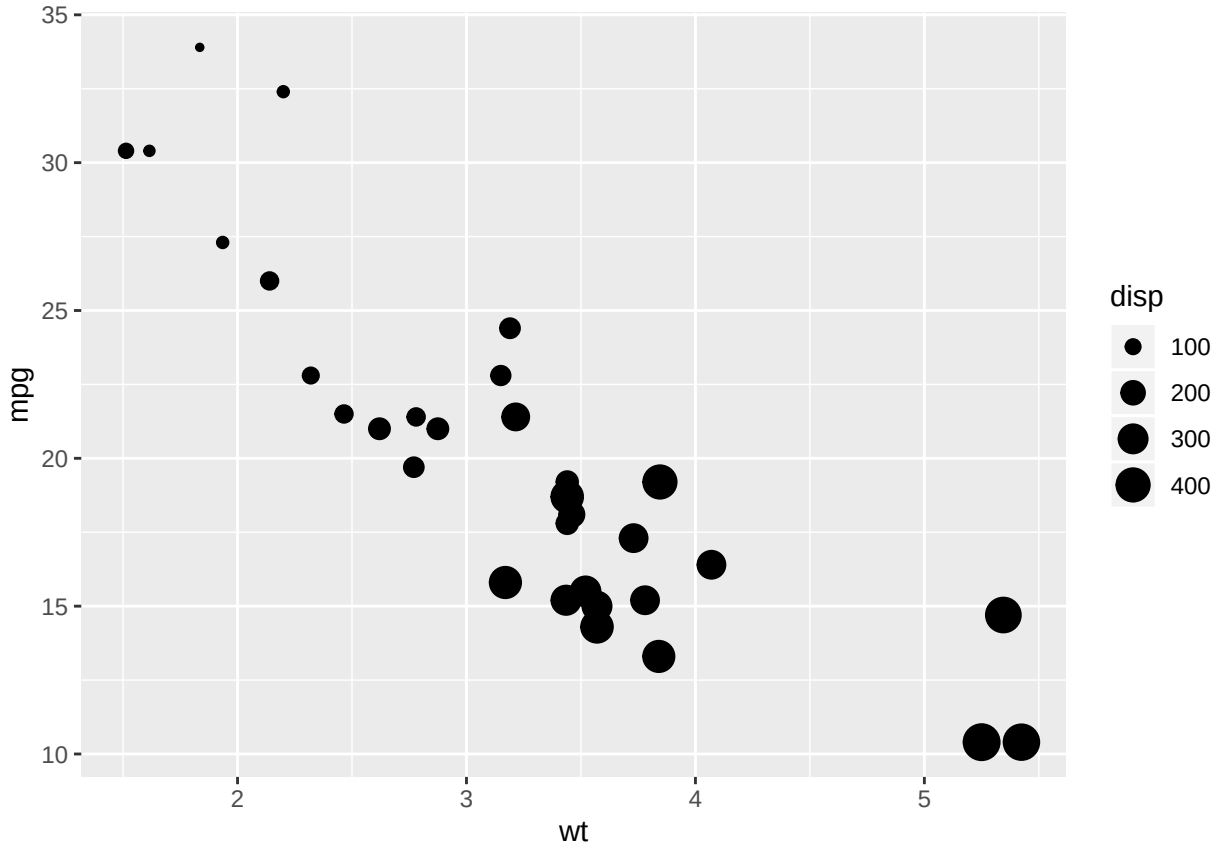
```
ggplot(mtcars, aes(x = wt, y = mpg)) +
  geom_point()
```




```
# color changes by disp (continuous)
ggplot(mtcars, aes(x = wt, y = mpg, color = disp)) +
  geom_point()
```




```
# size changes by disp (continuous)
ggplot(mtcars, aes(x = wt, y = mpg, size = disp)) +
  geom_point()
```



As "shape" only makes sense with categorical data, and *disp* is cotinuous. So the code below is incorrect.

```
ggplot(mtcars, aes(x = wt, y = mpg, shape = disp)) +
  geom_point()
```