class05.R

codyhargadon

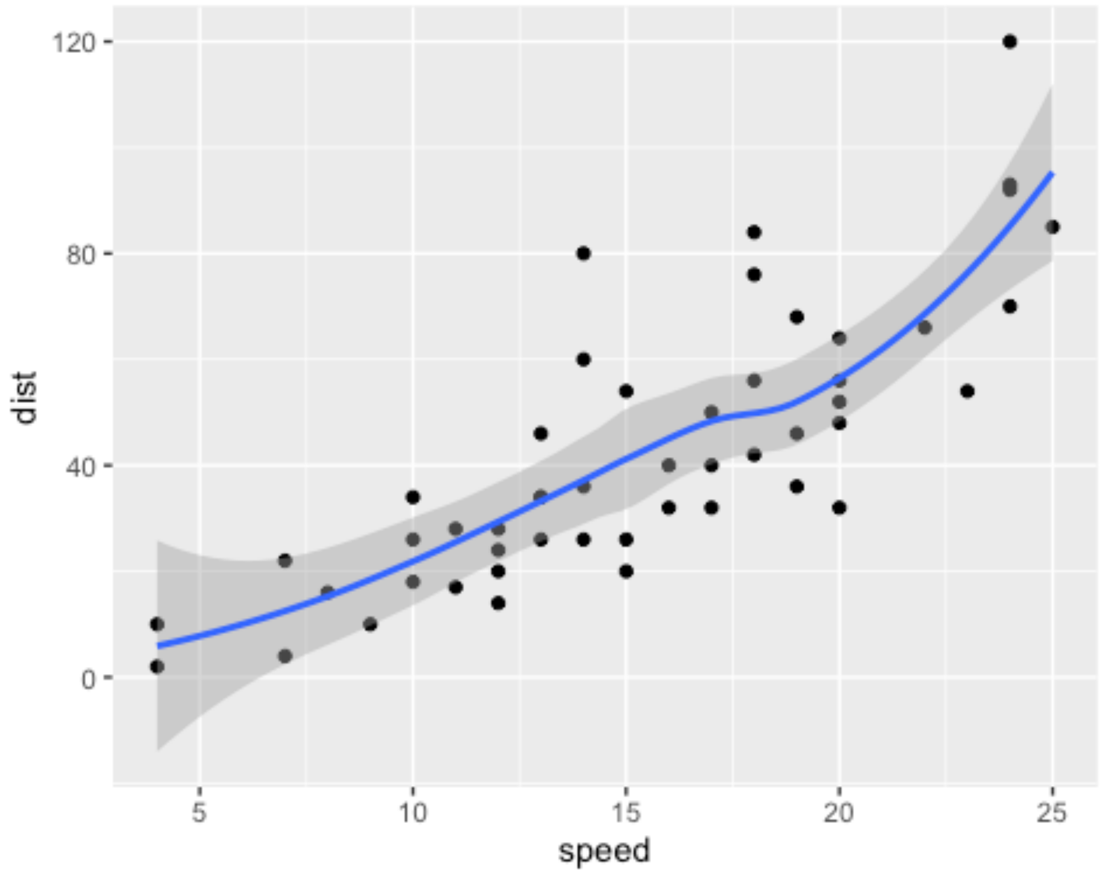
2021-10-12

```
# 'title: "Class 05 Data Visualization"
# 'author: "Alexis Cody Hargadon, A15479290"
# today's objective: learn how to use ggplot(), aes(), and geom_point()
#used plot type determined by type of variable:
  #continuous numeric: measured data, can have infinite values within
possible range
  #discrete categorical: observations can only exist at limited values, often
counts
library(ggplot2)
cars
```

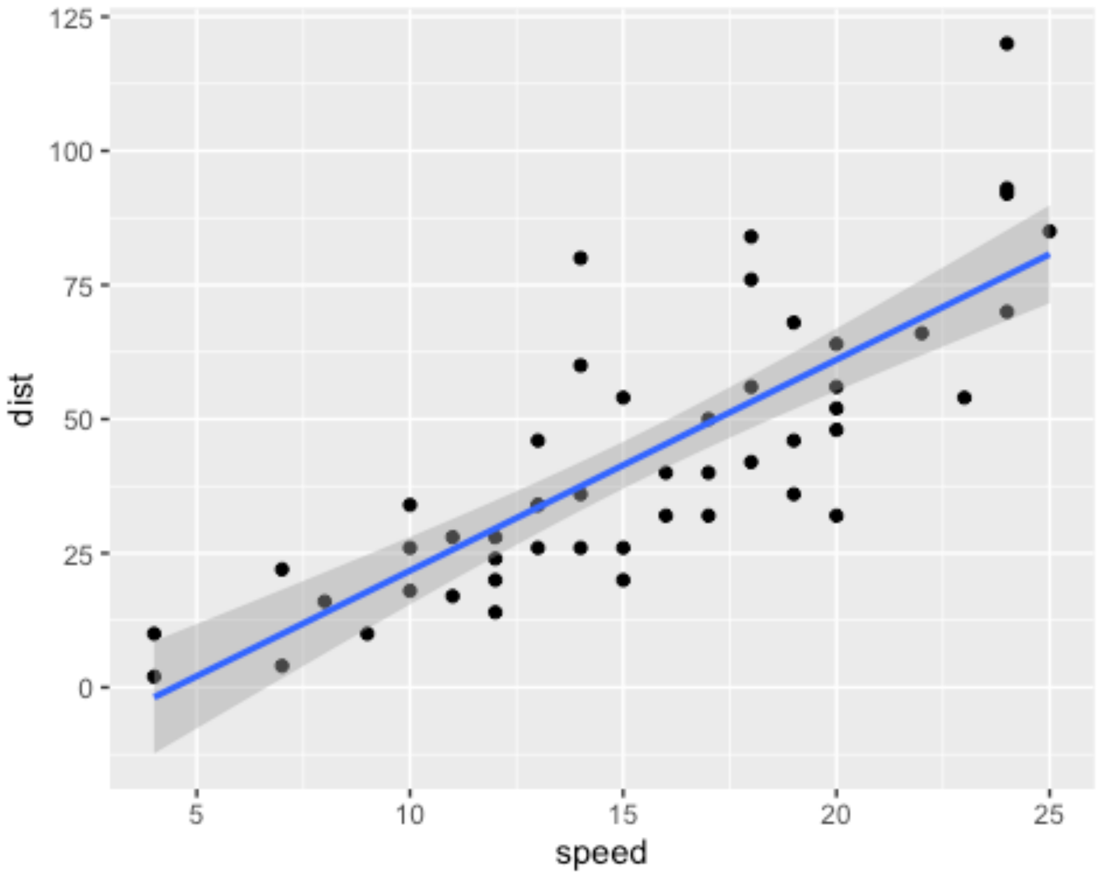
```
##    speed dist
## 1      4    2
## 2      4   10
## 3      7    4
## 4      7   22
## 5      8   16
## 6      9   10
## 7     10   18
## 8     10   26
## 9     10   34
## 10    11   17
## 11    11   28
## 12    12   14
## 13    12   20
## 14    12   24
## 15    12   28
## 16    13   26
## 17    13   34
## 18    13   34
## 19    13   46
## 20    14   26
## 21    14   36
## 22    14   60
## 23    14   80
## 24    15   20
## 25    15   26
## 26    15   54
## 27    16   32
## 28    16   40
## 29    17   32
## 30    17   40
## 31    17   50
## 32    18   42
## 33    18   56
## 34    18   76
## 35    18   84
## 36    19   36
## 37    19   46
## 38    19   68
## 39    20   32
## 40    20   48
## 41    20   52
## 42    20   56
## 43    20   64
## 44    22   66
## 45    23   54
## 46    24   70
## 47    24   92
## 48    24   93
## 49    24  120
## 50    25   85
```

```
#every ggplot had a data + aes + geom
ggplot(data=cars) +
  aes(x=speed, y= dist) +
  geom_point() +
  geom_smooth()
```

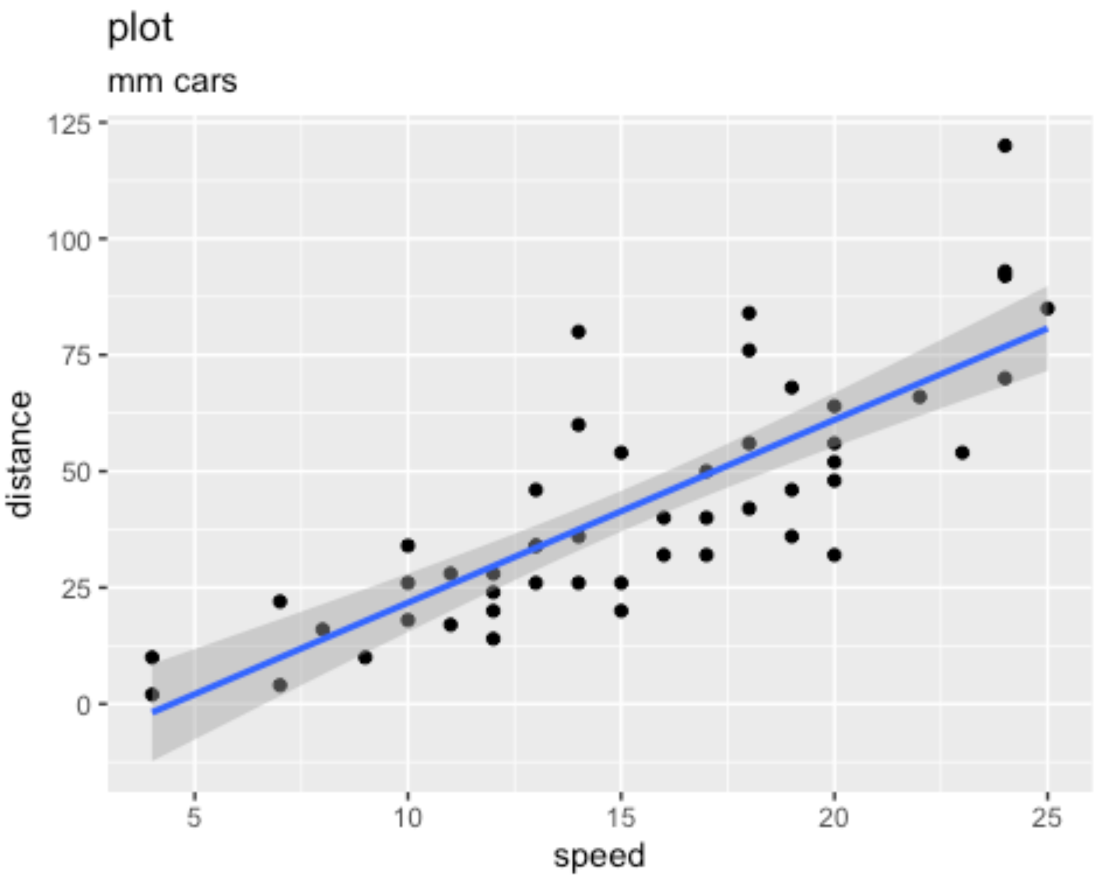
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



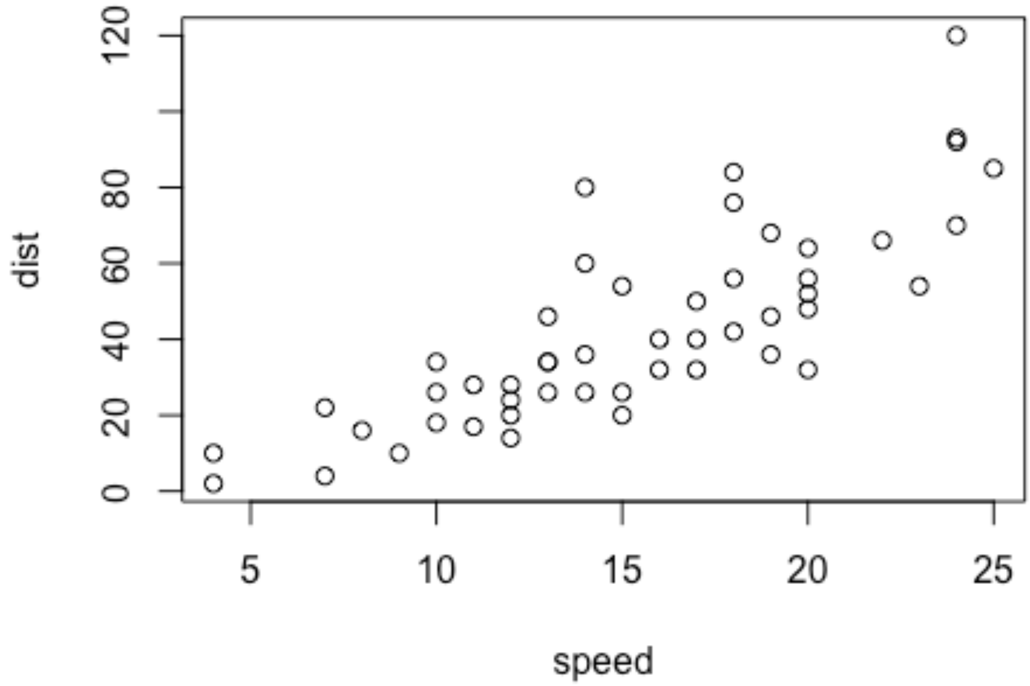
```
#change to a linear model
p <- ggplot(data=cars) +
  aes(x=speed, y= dist) +
  geom_point() +
  geom_smooth(method = "lm")
#can call p to make plot now
p

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



```
p +labs(title= "plot",
        subtitle="mm cars",
        x= "speed",
        y= "distance")

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

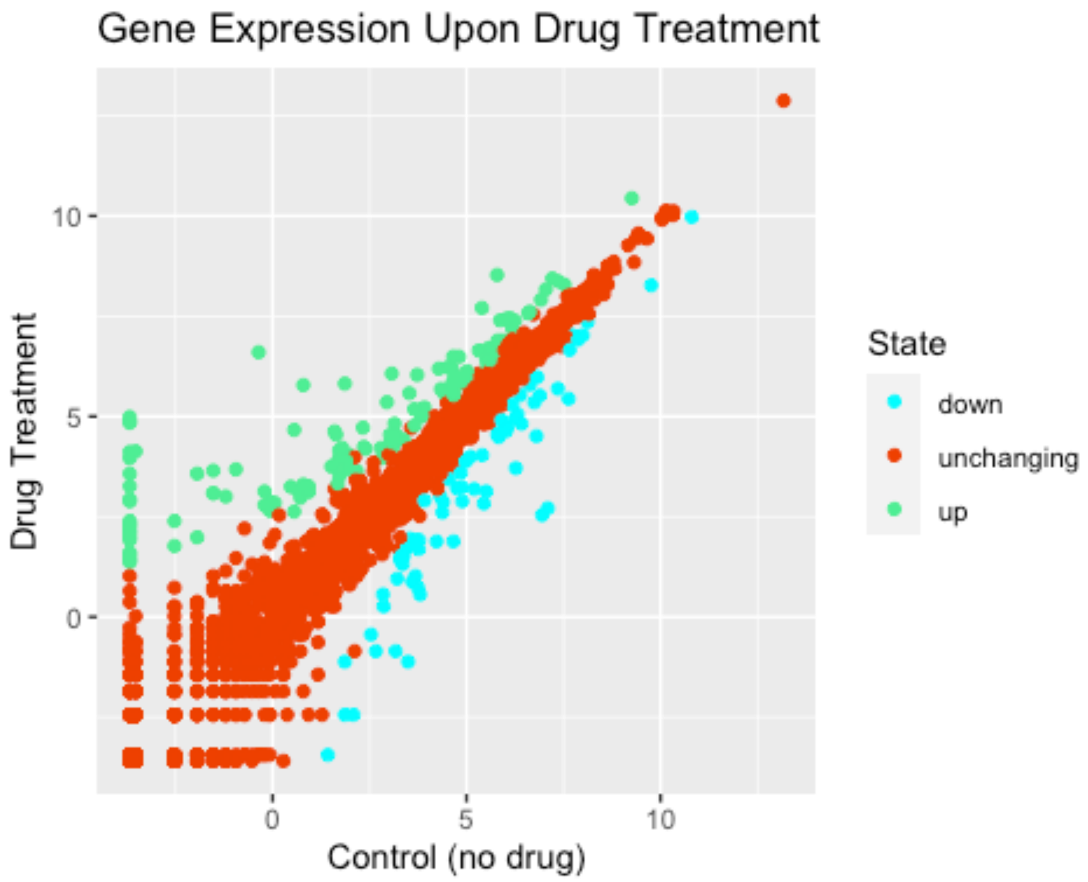


```
#base graphics is shorter
plot(cars)
```



```
url <- "https://bioboot.github.io/bimm143_S20/class-
material/up_down_expression.txt"
genes <- read.delim(url)
head(genes)

##          Gene Condition1 Condition2      State
## 1       A4GNT -3.6808610 -3.4401355 unchanging
## 2        AAAS  4.5479580  4.3864126 unchanging
## 3       AASDH  3.7190695  3.4787276 unchanging
## 4        AATF  5.0784720  5.0151916 unchanging
## 5        AATK  0.4711421  0.5598642 unchanging
## 6 AB015752.4 -3.6808610 -3.5921390 unchanging

nrow(genes)

## [1] 5196

colnames(genes)

## [1] "Gene"       "Condition1" "Condition2" "State"

ncol(genes)

## [1] 4
```

```
table(genes$State)

##
##       down unchanging         up
##         72       4997        127

#fraction of genes upregulated, answer with 2 significant figures
round( table(genes$State)/nrow(genes) *100, 2)

##
##       down unchanging         up
##       1.39      96.17       2.44

127/5196

## [1] 0.02444188

g <- ggplot(data=genes) +
  aes(x=Condition1, y=Condition2, col=State) +
  geom_point()
g+ scale_color_manual(values= c("cyan", "orangered2", "seagreen2")) +
labs(title= "Gene Expression Upon Drug Treatment", x="Control (no drug)",
y="Drug Treatment")
```


Gene Expression Upon Drug Treatment
Drug Treatment
Control (no drug)
State
down
unchanging
up