

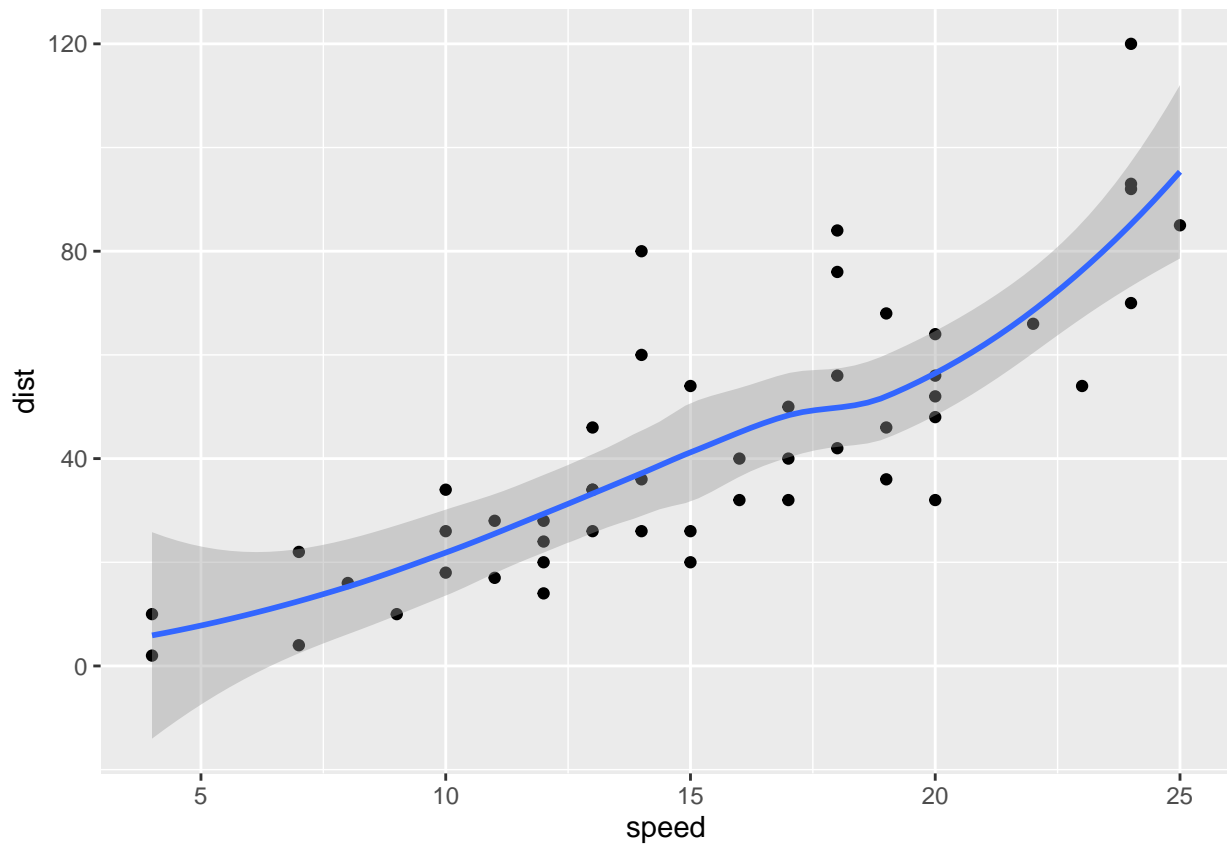
Class 5: Data Visualization

Natalie Avina (A15590695)

2021-10-12

```
#before using ggplot you need to load it up with library  
library(ggplot2)  
  
#ggplot has data + aes + geoms  
#geom_line is a trendline where geom_line treats it as consecutive data  
ggplot(data=cars)+ aes(x=speed, y=dist) + geom_point() + geom_smooth()
```

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

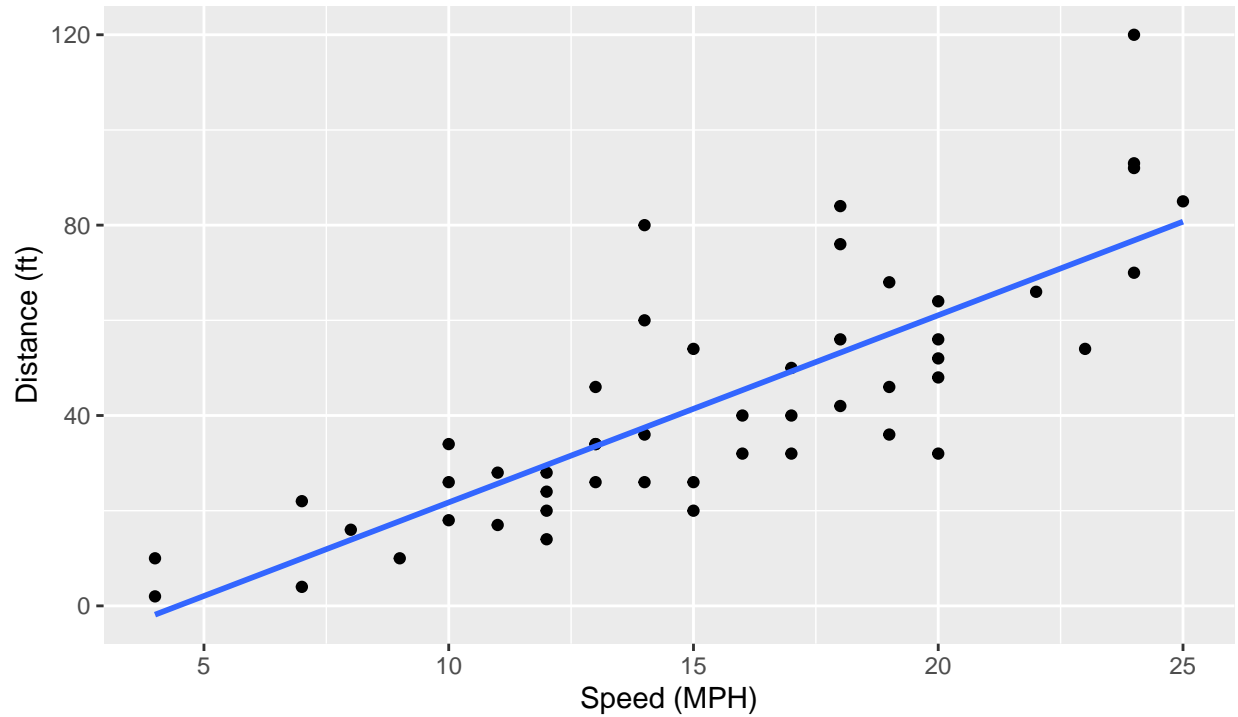


```
#change to a linear model and save as p to call it forward when needed  
p <- ggplot(data=cars)+ aes(x=speed, y=dist) + geom_point() + geom_smooth(method="lm", se=FALSE)  
p + labs(title="Speed and Stopping Distance of Cars", x= "Speed (MPH)", y= "Distance (ft)", subtitle =
```

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

Speed and Stopping Distance of Cars

relationship between speed and distance required for stopping



Dataset: cars

```
#base graphics is shorter: plot (cars)
## Warning: It is deprecated to specify `guide = FALSE` to remove a guide. Please
## use `guide = "none"` instead.
# to change size of points use size= parameter in aes
# to change color point use color=
# to change transparency use alpha=
```

```
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
```

```
#Q. how many genes?
nrow(genes)
```

```
## [1] 5196
```

```
#how to access State column  
table(genes$State)
```

```
##  
##      down  unchanged      up  
##      72      4997      127
```

```
#Q. How many genes are down regulated?  
table(genes$State)/nrow(genes)
```

```
##  
##      down  unchanged      up  
## 0.01385681 0.96170131 0.02444188
```

```
#Q. What % is up/down?  
perc<- table(genes$State)/nrow(genes)*100  
#round  
round(perc, 2)
```

```
##  
##      down  unchanged      up  
##      1.39      96.17      2.44
```

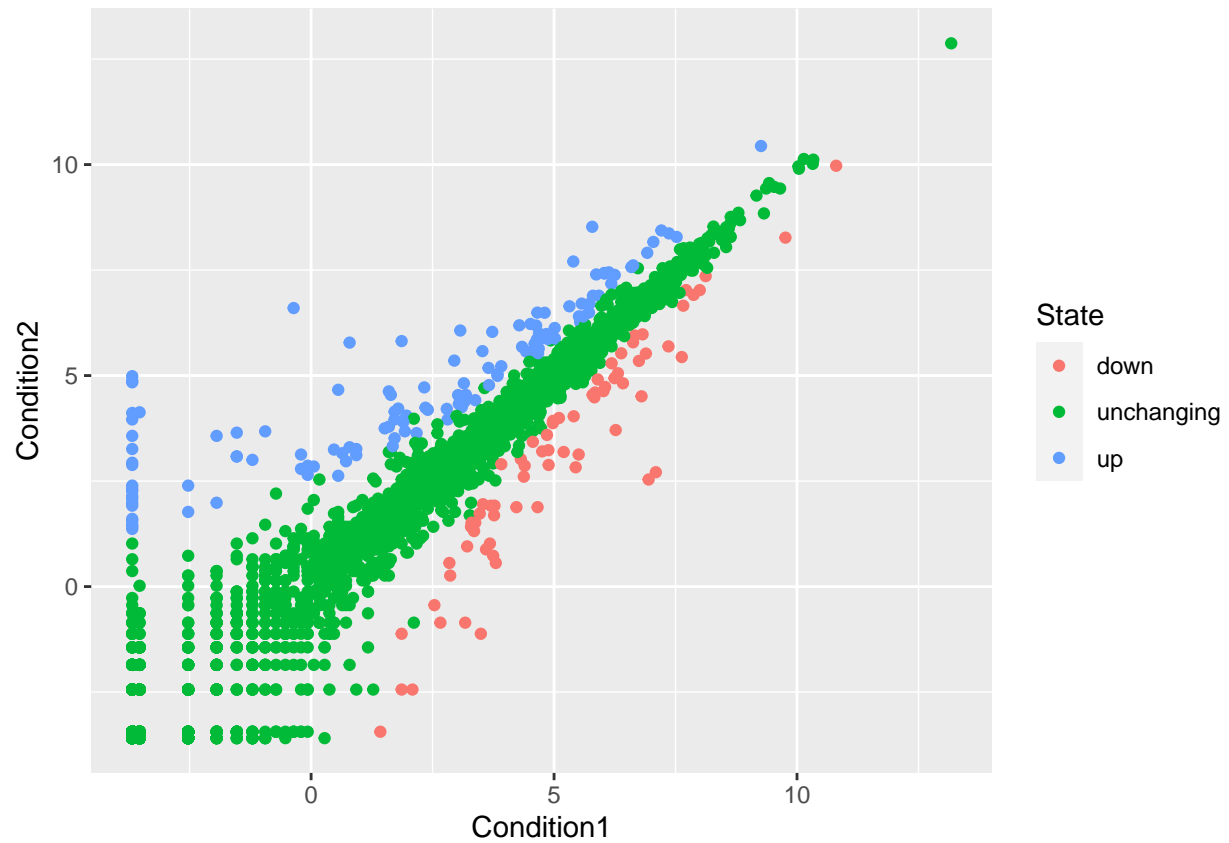
```
colnames(genes)
```

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

```
ncol(genes)
```

```
## [1] 4
```

```
ggplot(genes) +  
  aes(x=Condition1, y=Condition2, col=State)+  
  geom_point()
```



```
v <- ggplot(genes) +
  aes(x=Condition1, y=Condition2, col=State)+
  geom_point()
v + scale_color_manual(values = c("lightblue3", "darkseagreen3", "rosybrown3")) +
  labs(title="Gene Expression ", x= "Control (no drug)", y= "Drug Treatment")
```

