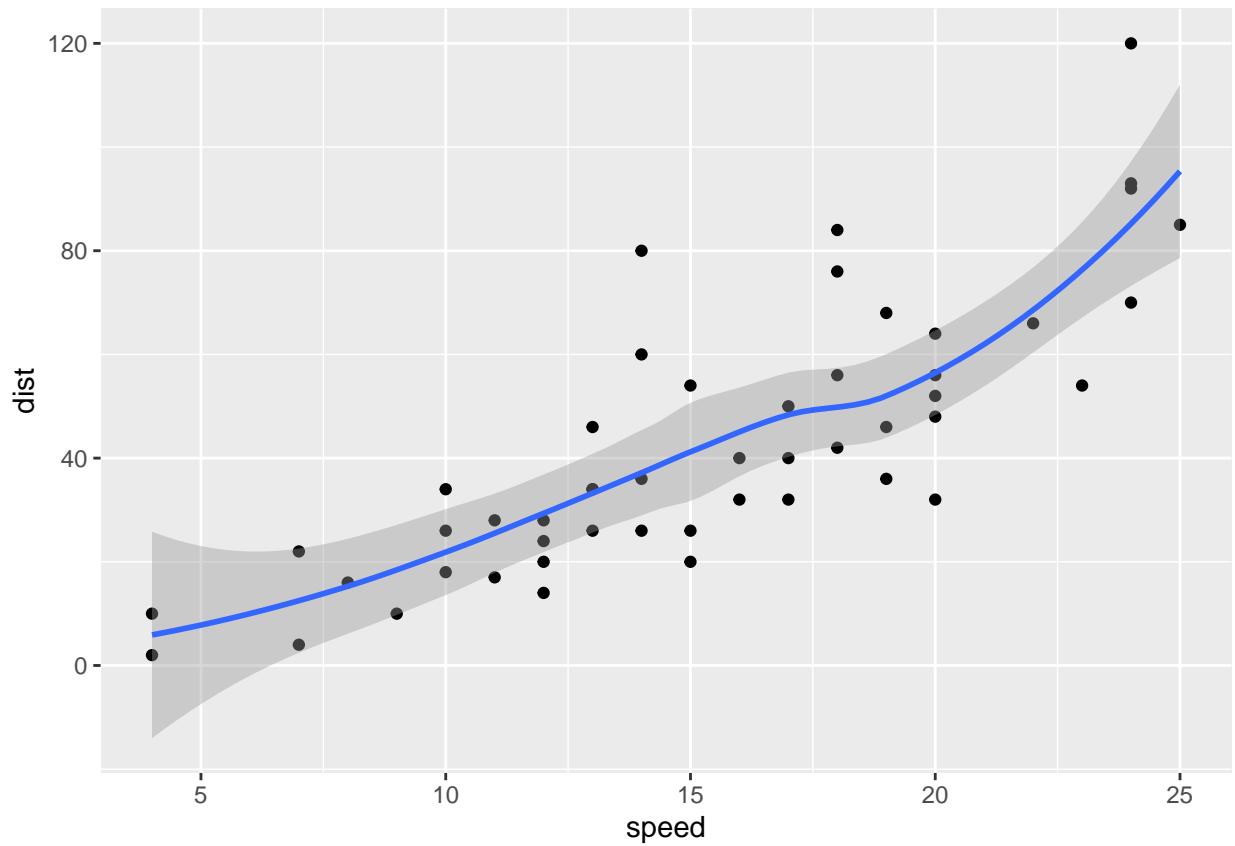


Class 05 Data Visualization

Camryn McCann A15437387

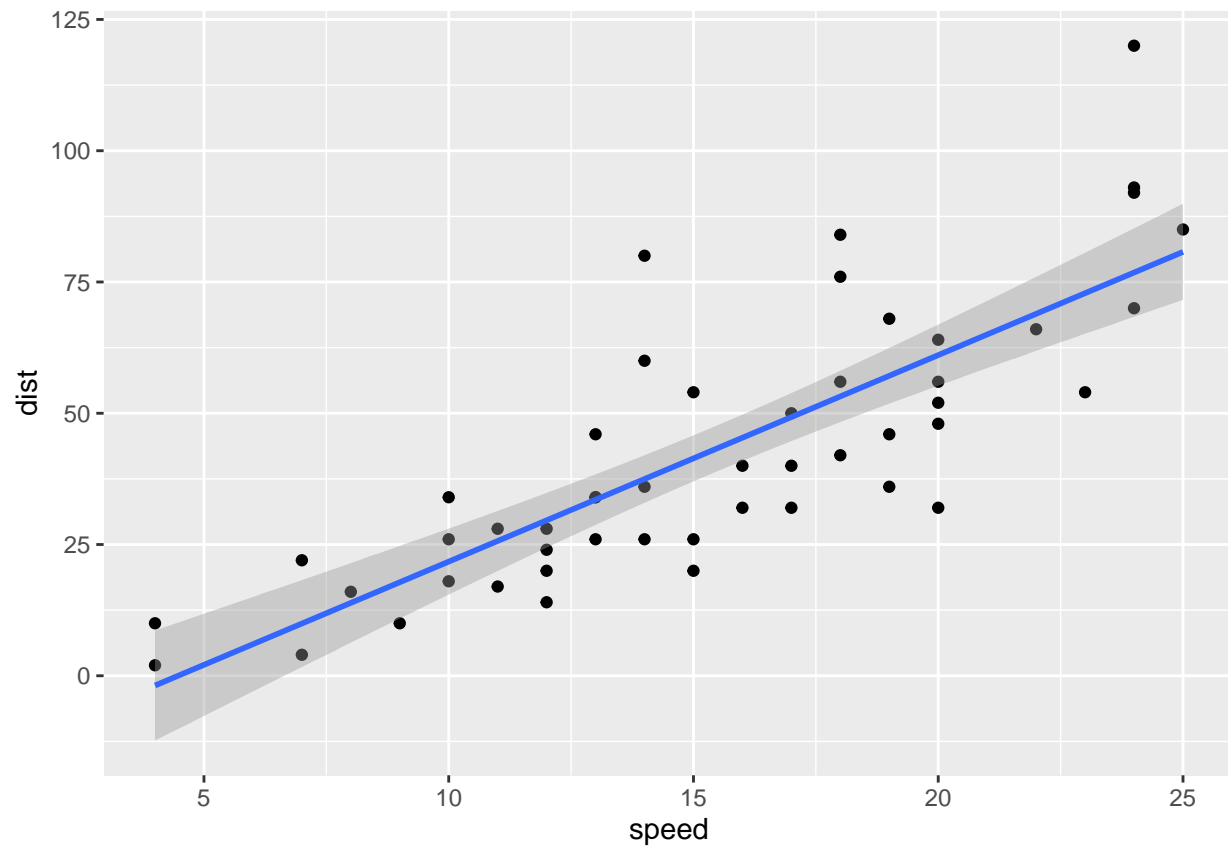
2021-10-12

```
# Class 05 Data Visualization  
  
# Must first call to load ggplot2  
library(ggplot2)  
  
# Every ggplot has a data + aes + geoms  
ggplot(data = cars) +  
  aes(x=speed, y=dist) +  
  geom_point(data = cars) +  
  geom_smooth(data = cars)  
  
## 'geom_smooth()' using method = 'loess' and formula 'y ~ x'
```



```
# Change to linear model
ggplot(data = cars) +
  aes(x=speed, y=dist) +
  geom_point(data = cars) +
  geom_smooth(method = "lm")
```

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

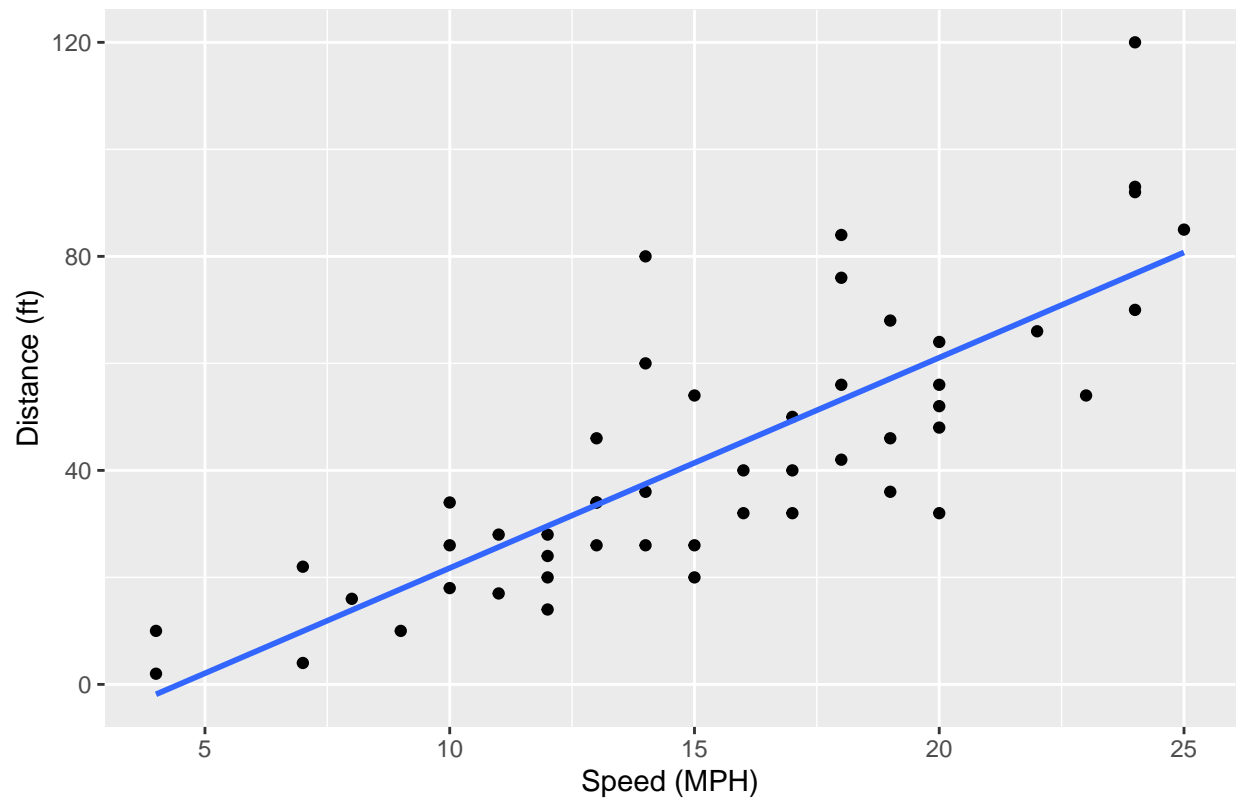


```
# Store plot as an object
p <- ggplot(data = cars) +
  aes(x=speed, y=dist) +
  geom_point(data = cars) +
  geom_smooth(method = "lm", se = FALSE)
```

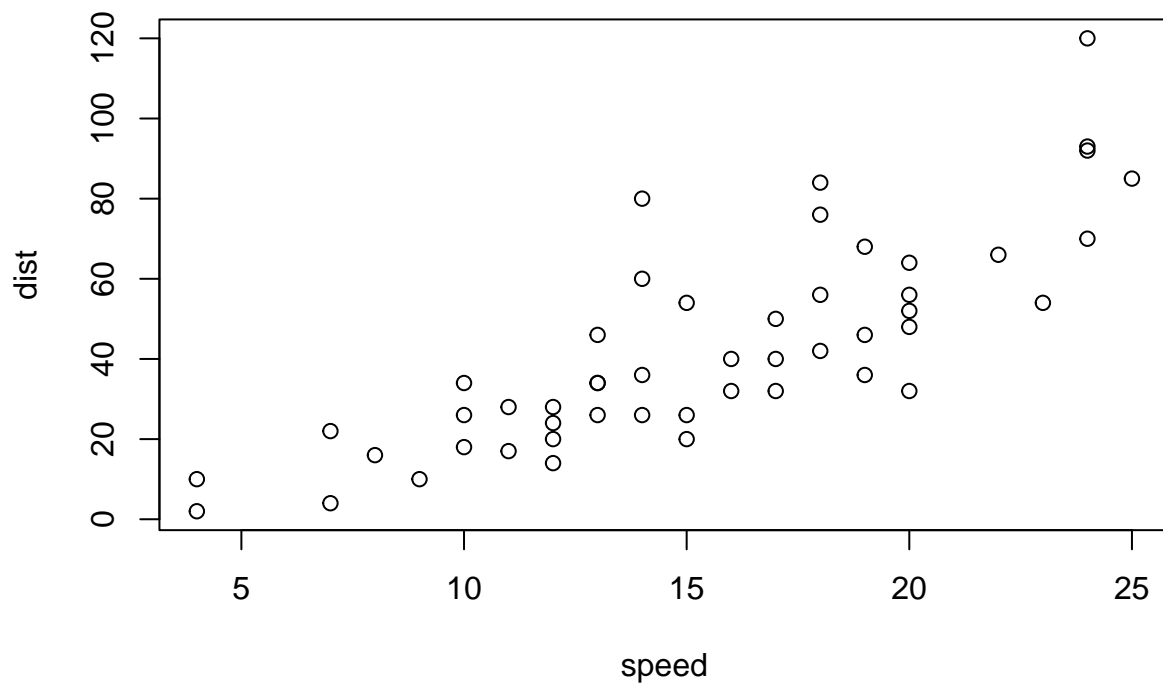
```
# Label the plot
p + labs(title="My Nice Plot", x = "Speed (MPH)", y = "Distance (ft)")
```

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

My Nice Plot



```
#Base graphics are shorter, for quick visualizations  
plot(cars)
```



```
# Access and read the data set
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
```

```
# Seeing how many
nrow(genes)
```

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

```
ncol(genes)
```

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

```
# Access one column + summarize its data
table(genes$State)
```

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

```
table(genes$State)/nrow(genes) * 100
```

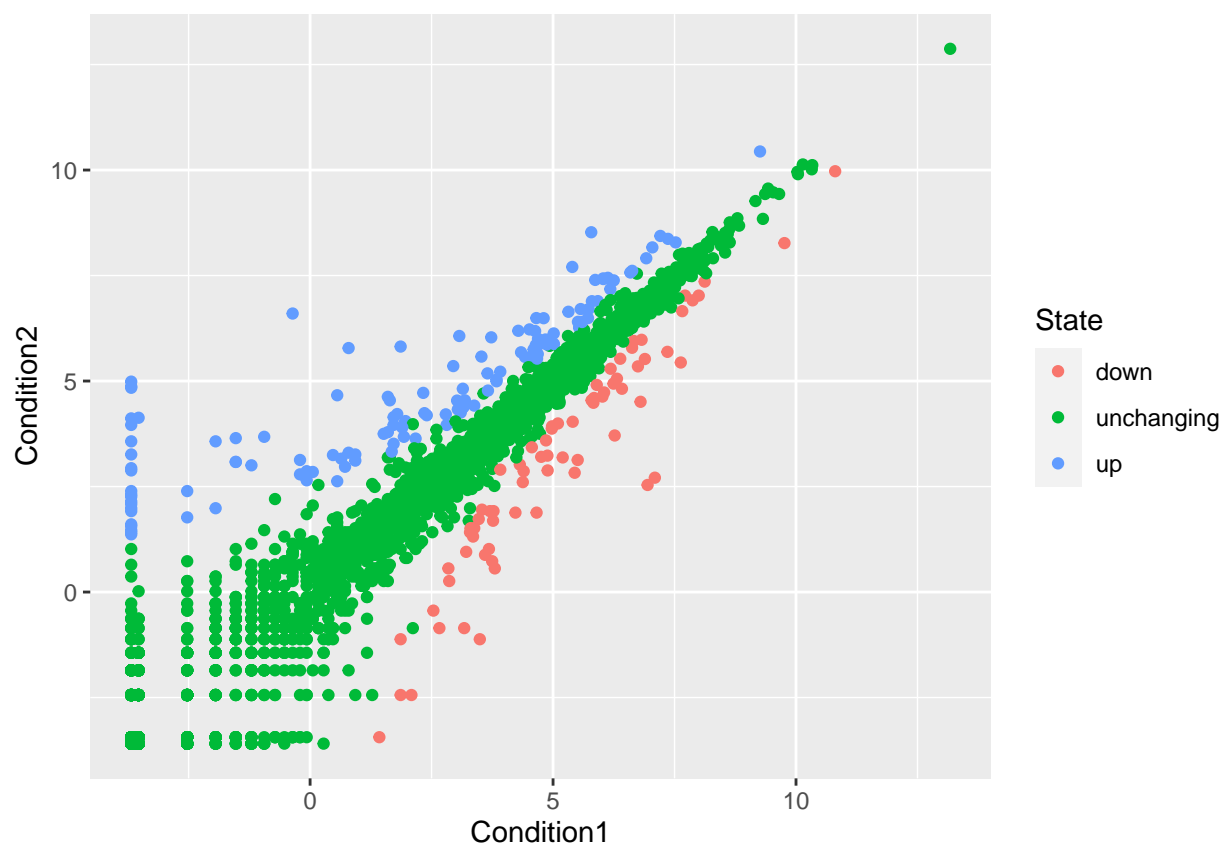
```
##
##      down  unchanged      up
## 1.385681 96.170131  2.444188
```

```
round (table(genes$State)/nrow(genes) * 100, 2)
```

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

```
# Back to plots
```

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



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

```
#Customize Plot
```

```
z + scale_color_manual(values = c("purple", "aquamarine", "green")) +
  labs(title = "Gene Expression Changes Upon Drug Treatment",
       x = "Control(no drug)", y = "Drug Treatment")
```

