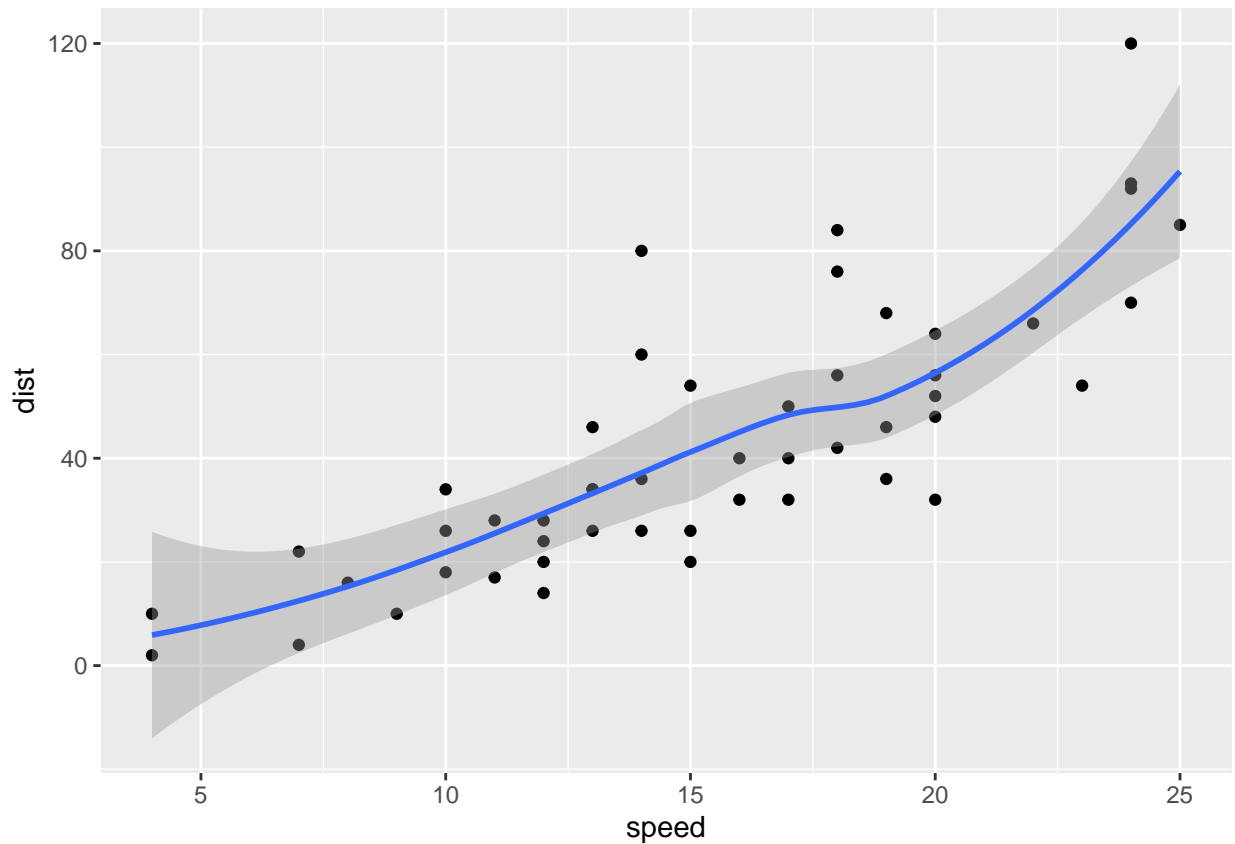


Class 5 Data Visualization

Tom Quach

2021-10-12

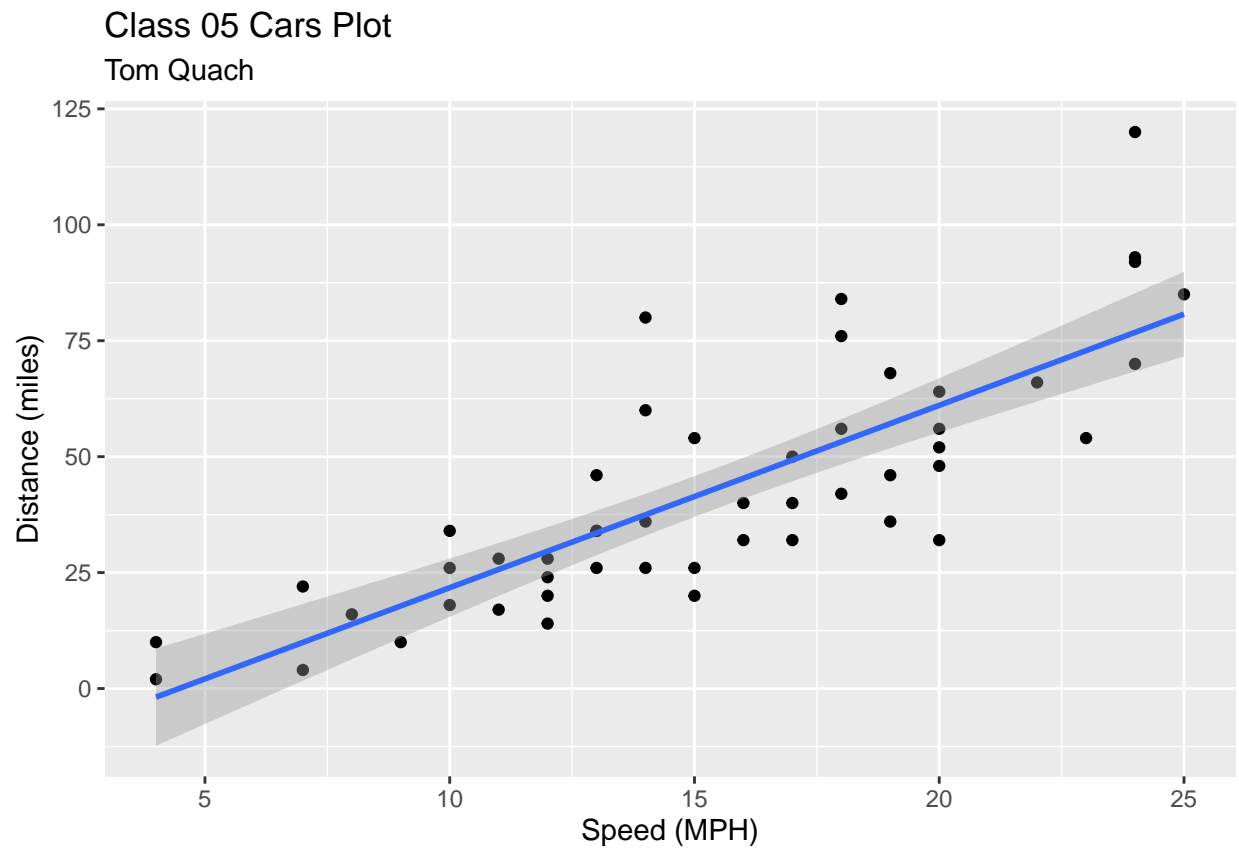
```
# Class 05 Data Visualiation  
  
# Lets start with a scatterplot  
  
# Before we can use it we need to load it up!  
library(ggplot2)  
  
# Every ggplot has a data + aes + geoms  
ggplot(data=cars) + aes(x=speed, y=dist) + geom_point() + geom_smooth()  
  
## 'geom_smooth()' using method = 'loess' and formula 'y ~ x'
```



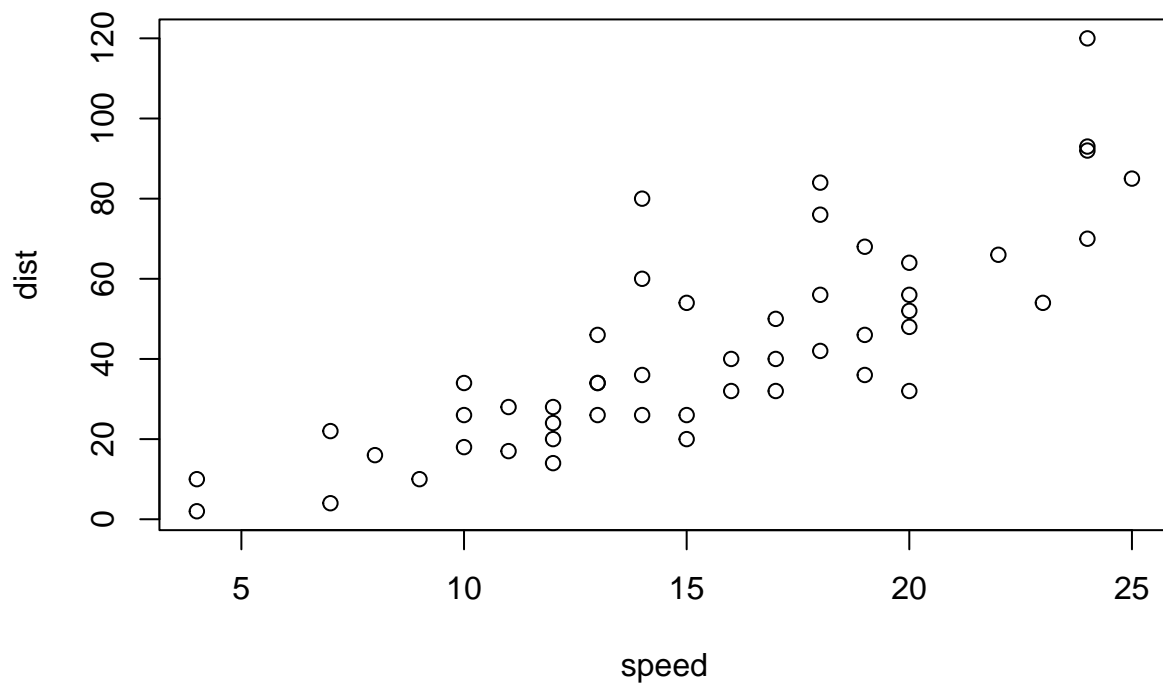
```
#Change to a linear model
p <- ggplot(data=cars) + aes(x=speed, y=dist) + geom_point() + geom_smooth(method="lm")

p + labs(title = "Class 05 Cars Plot", subtitle = "Tom Quach", x = "Speed (MPH)", y = "Distance (miles)")

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



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



```
#Getting the beginning data set for genes
```

```
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
```

```
#How many genes are in my data?
```

```
nrow(genes)
```

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

```
#How to access the State column
```

```
table(genes$State)
```

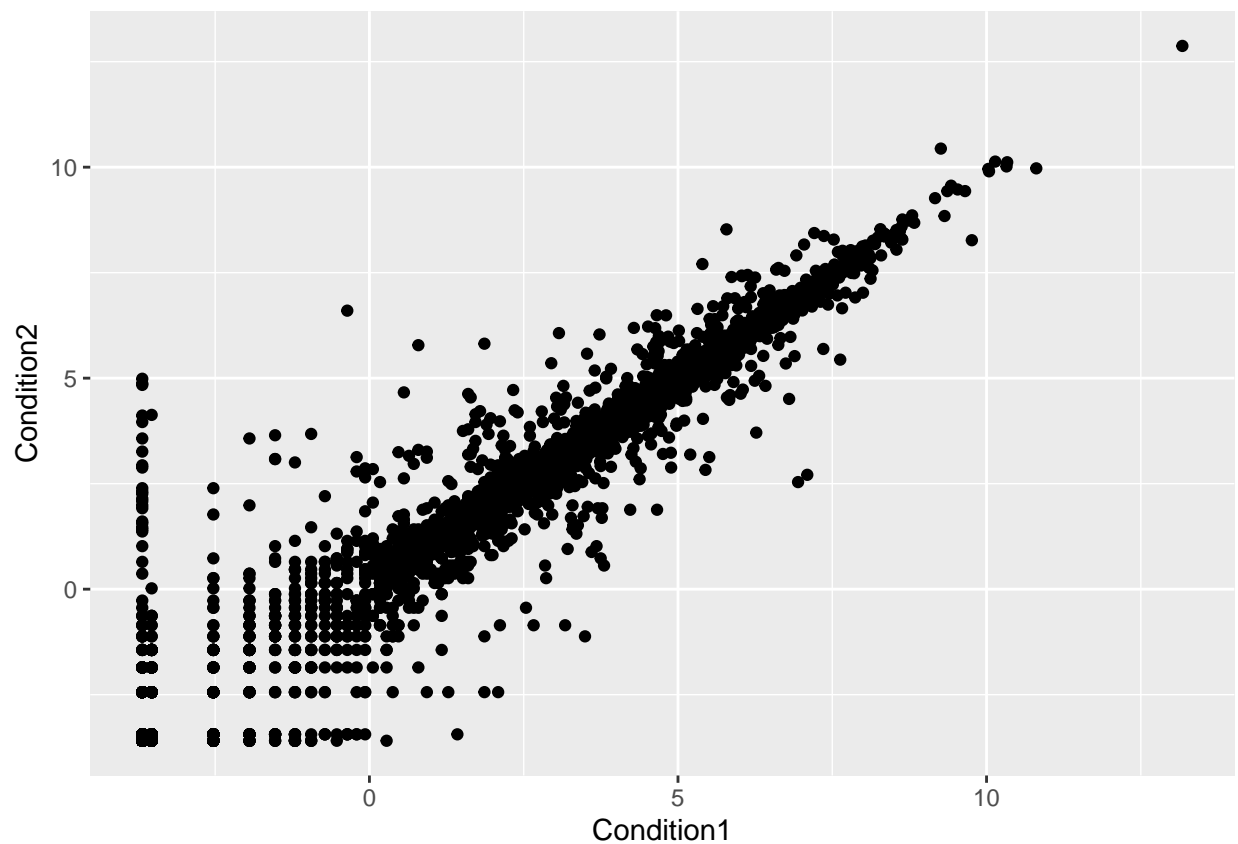
```
##
##      down unchanging      up
##      72      4997      127
```

```
#What percent are up/down
prec <- table(genes$State) / nrow(genes) * 100

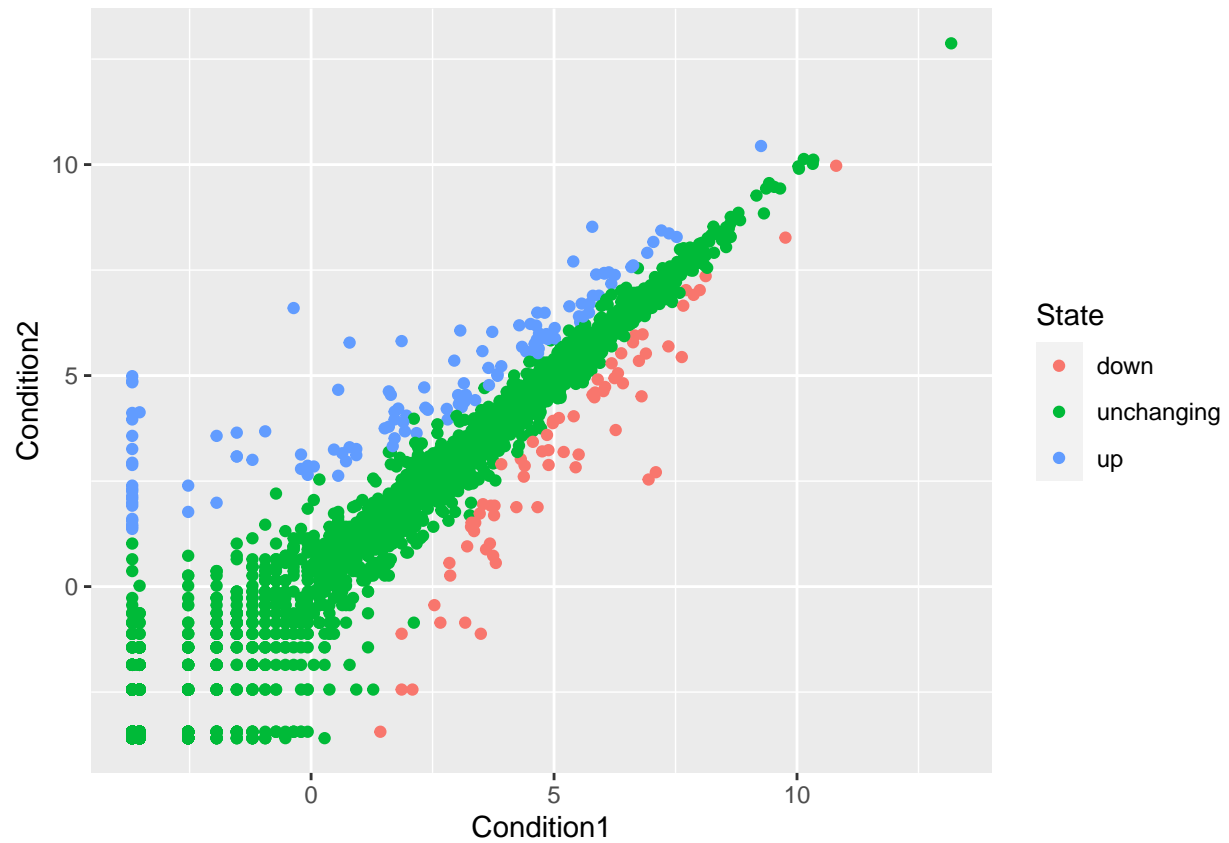
#Round percentage by 2 significant figures
round(prec, 2)
```

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

```
#Creating unnamed scatter plot. Condition 1 against condition 2 where each point is a gene
ggplot(data = genes) + aes(x=Condition1, y=Condition2) + geom_point()
```

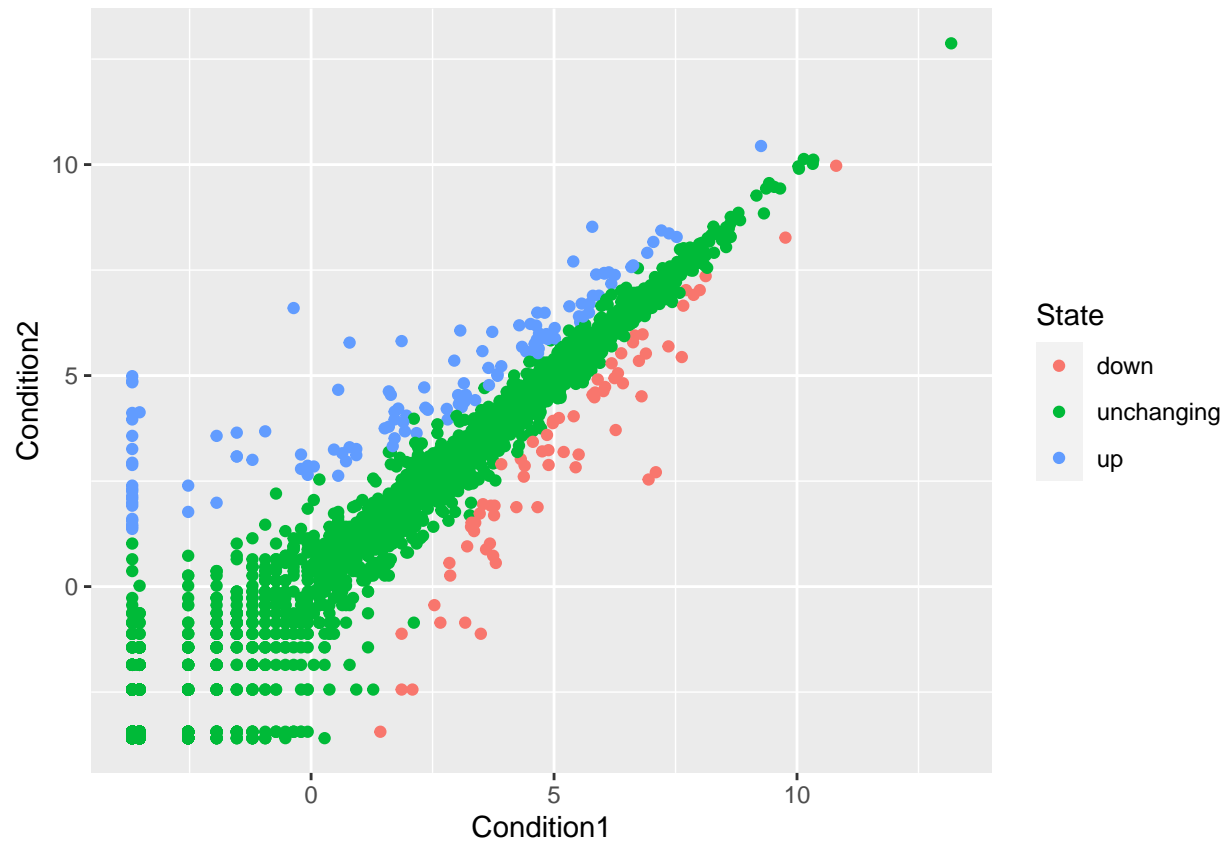


```
#Coloring the points so we can tell which one of the genes are down, unchanging, or up
ggplot(data = genes) + aes(x=Condition1, y=Condition2, color=State) + geom_point()
```

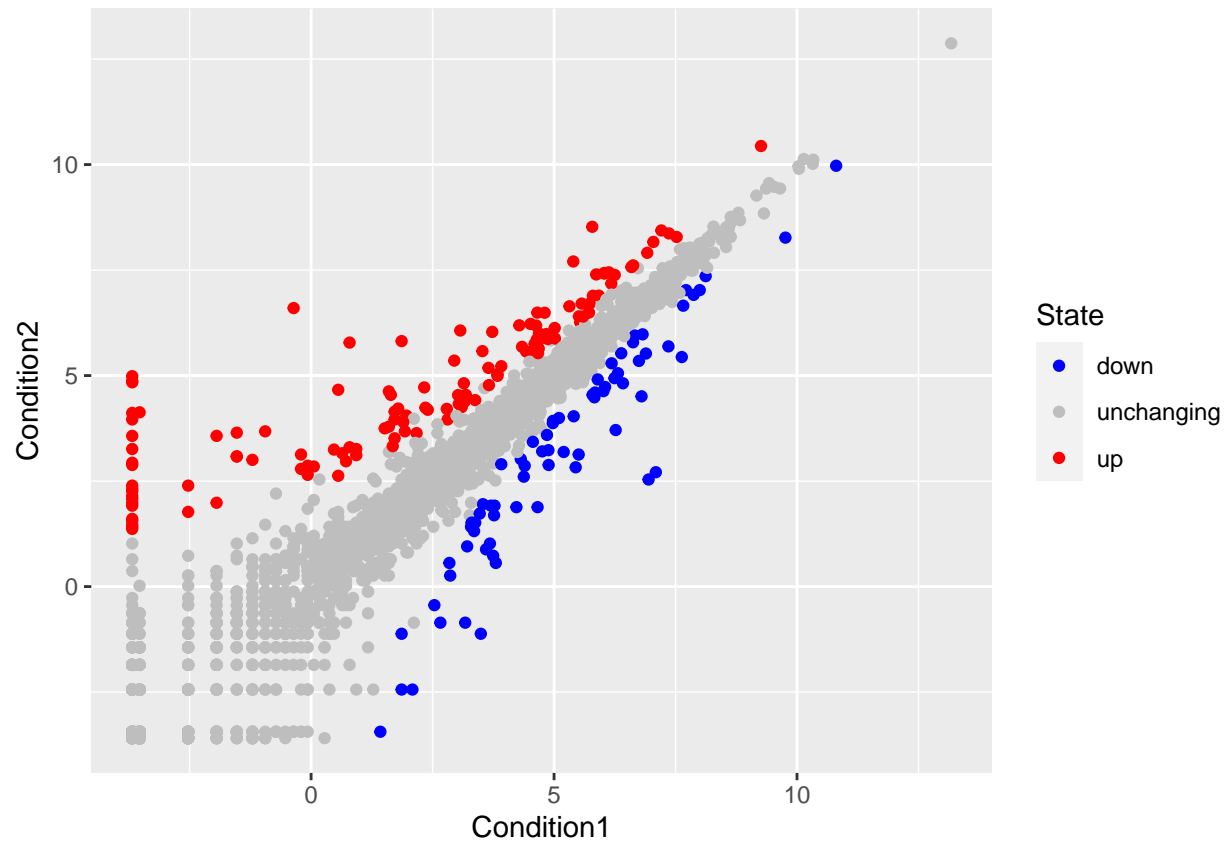


```
#Storing our plot code so we do not have to type it all out when needed
genesp <- ggplot(data = genes) + aes(x=Condition1, y=Condition2, color=State) + geom_point()

#Printing out genesp to see if it works
genesp
```

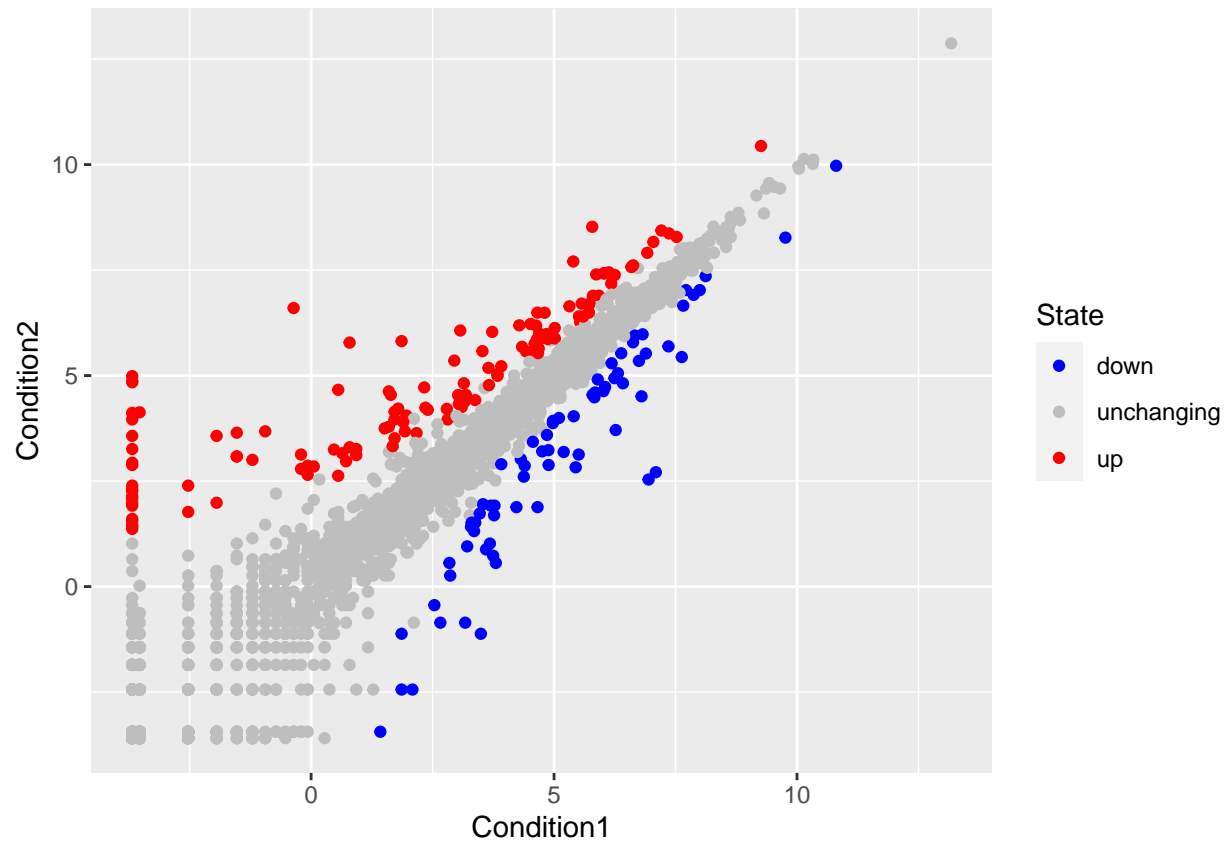


```
#Changing the color scheme of the points  
genesp + scale_color_manual(values=c("blue", "gray", "red"))
```



```
#Storing final plot
genespfinal <- genesp + scale_color_manual(values=c("blue", "gray", "red"))

#Print genespfinal to check if it works
genespfinal
```



```
#Naming the plot and axis
genespfinal + labs(title = "Gene Expression Changes Upon Drug Treatment", x = "Control (No Drug)", y =
```


Gene Expression Changes Upon Drug Treatment

