Lab Week 5 Data visualization

Sam Altshuler (PID: A59010373)
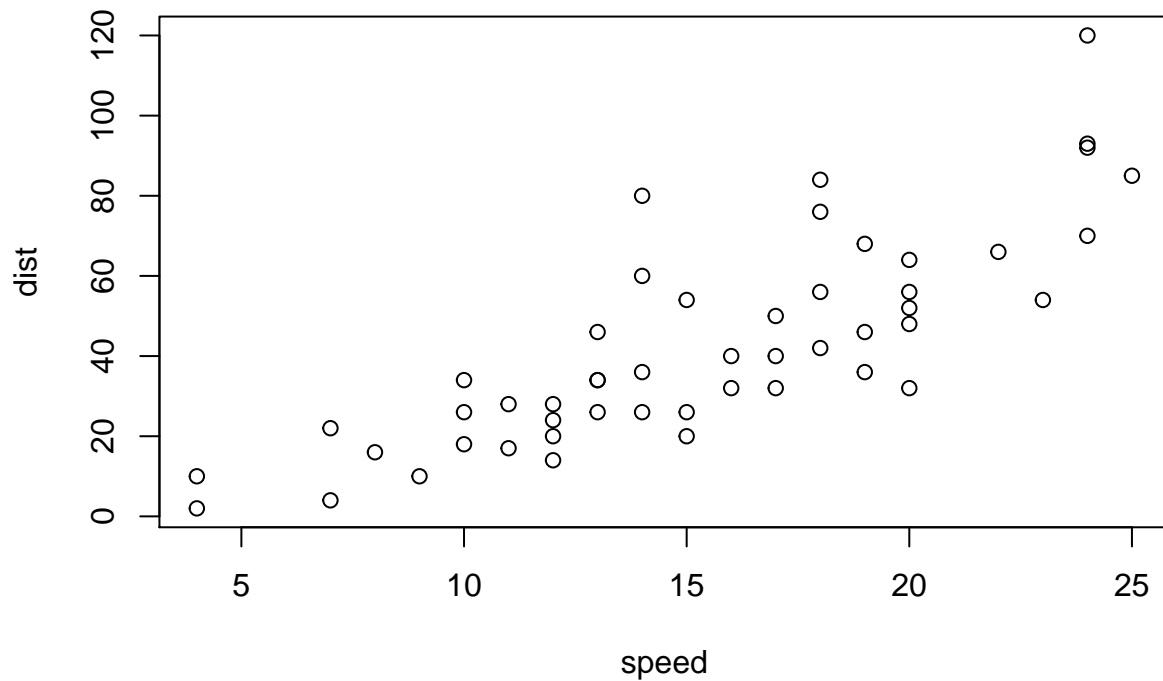
2/2/2022

```r
# load up ggplot, don't need to install every time
# install.packages("ggplot2")
library(ggplot2)
```
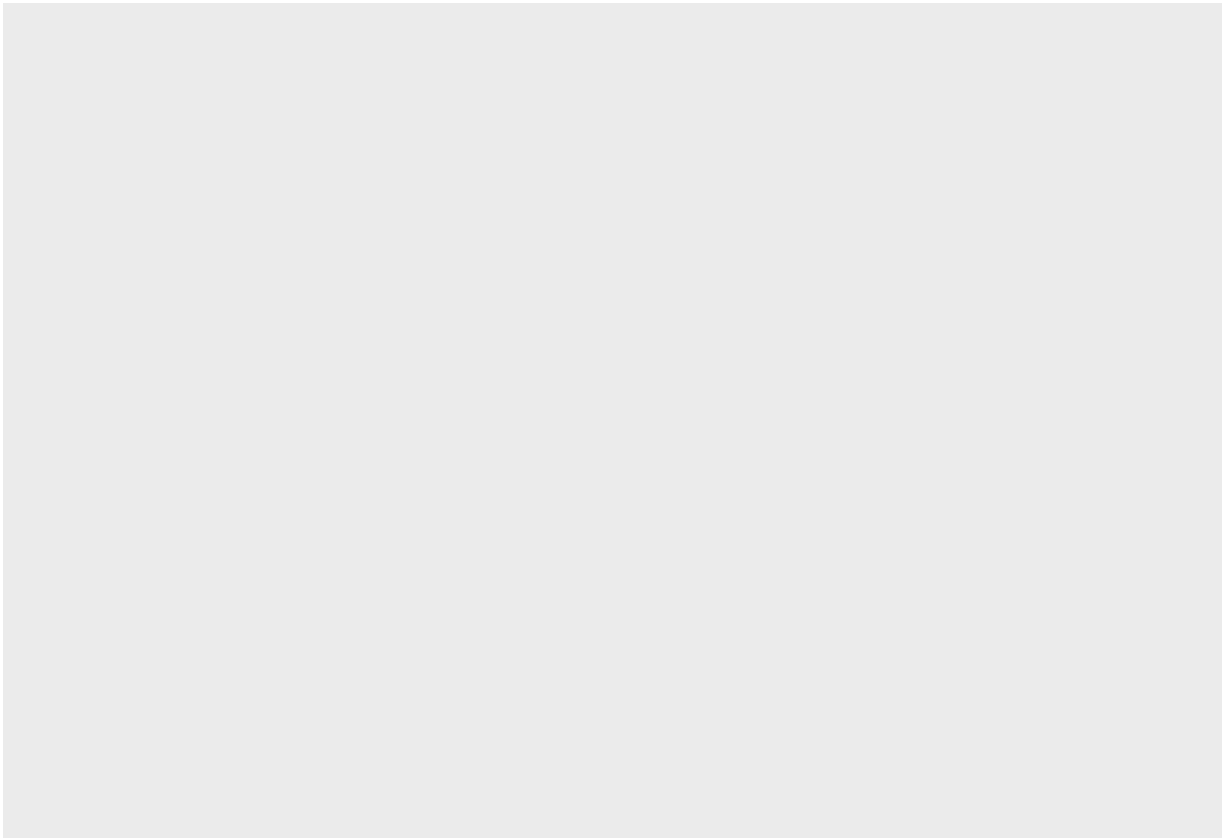
```
## Warning: package 'ggplot2' was built under R version 4.0.3
```

```r
# cars dataset already loaded in base R
View(cars)

# typical base R clock
plot(cars)
```
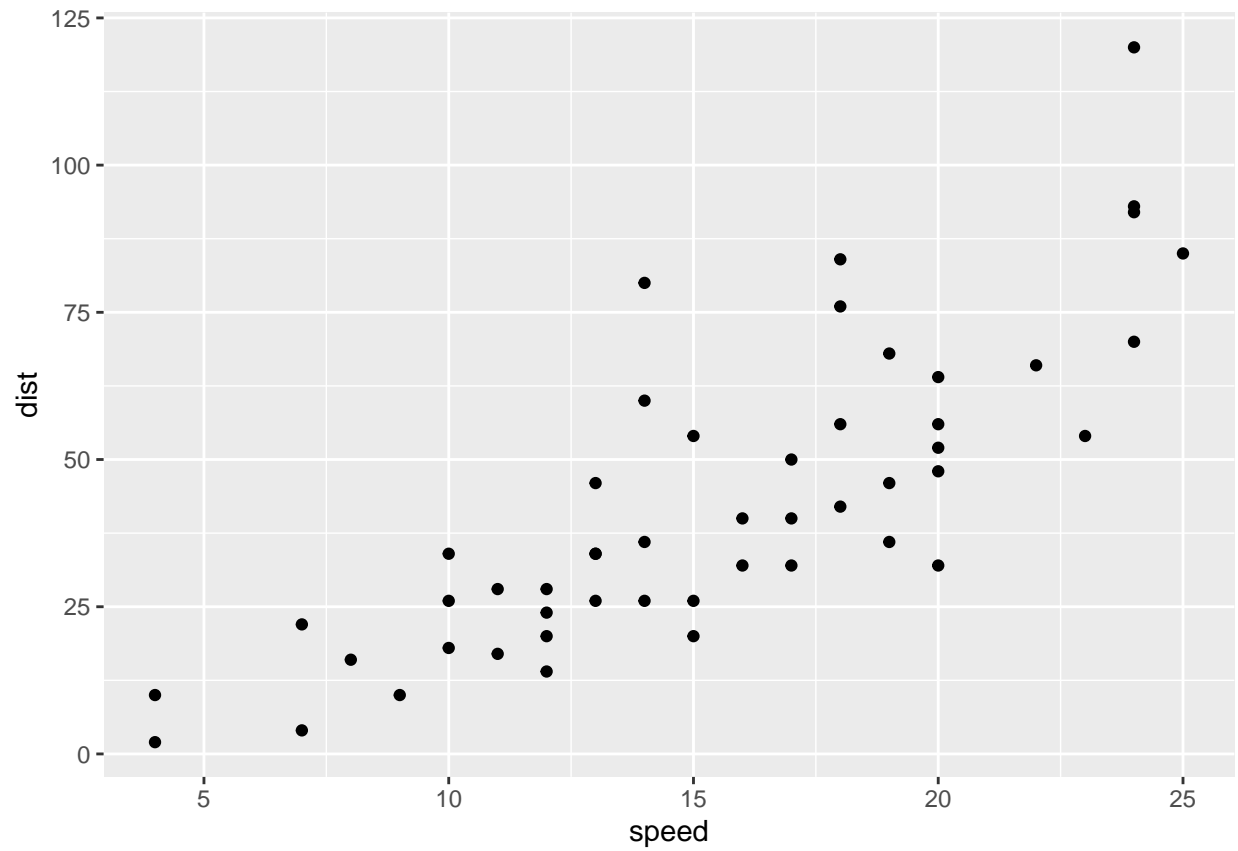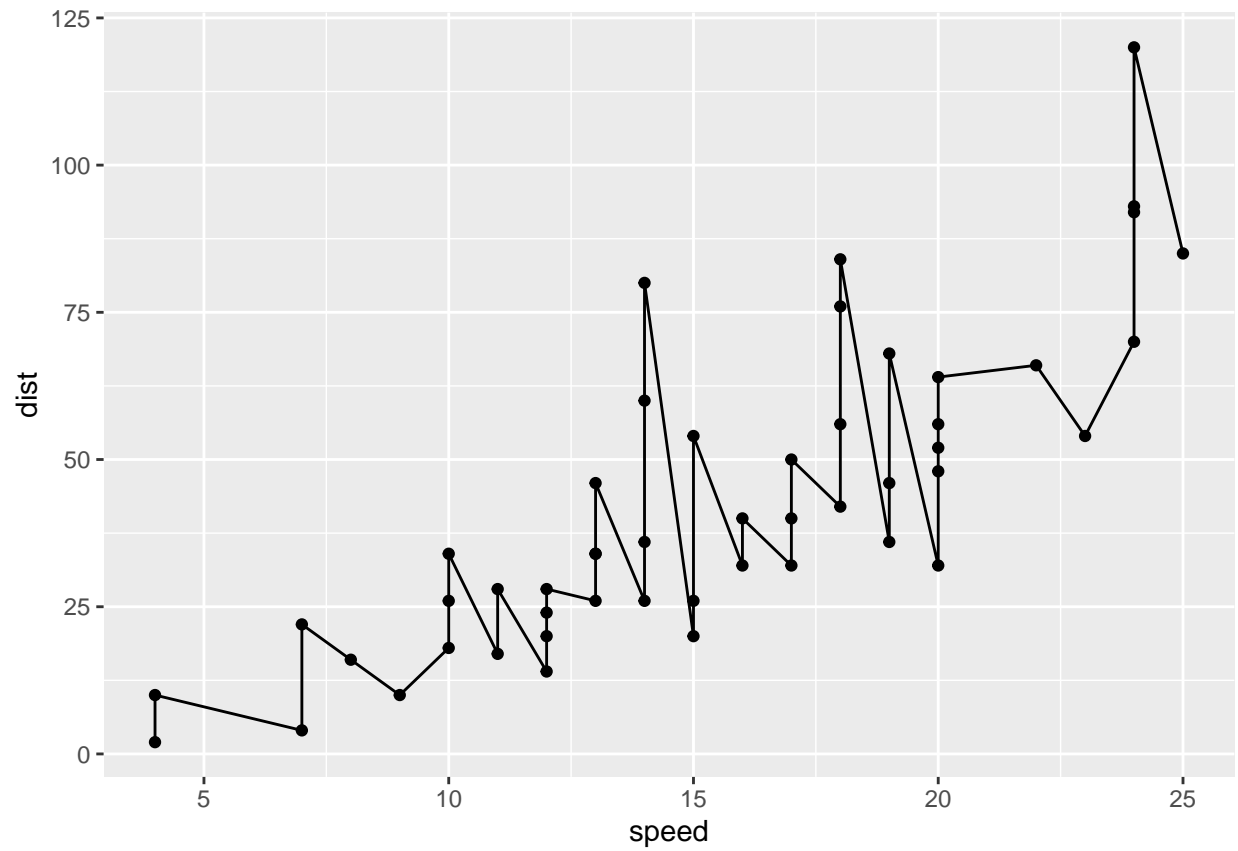
```r
# ggplot cars (no aesthetic choosen so blank square)
ggplot(cars)
```



```r
# add in x and y variables and make a scatter plot
p <- ggplot(cars, aes(x= speed, y = dist))+
  geom_point()
p
```
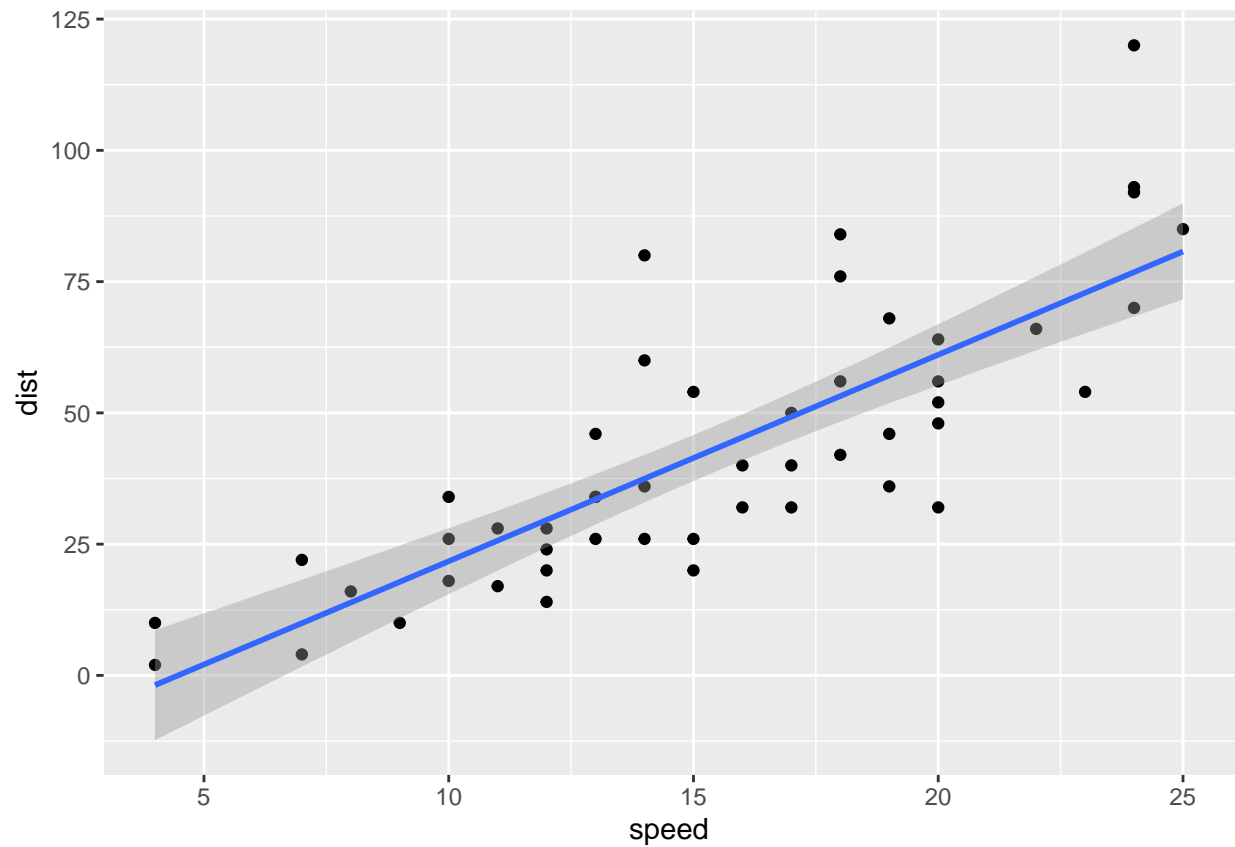
```
#add an unnecessary line
p+geom_line()
```

```r
# add a fitted line
p + geom_smooth(method = "lm")
```
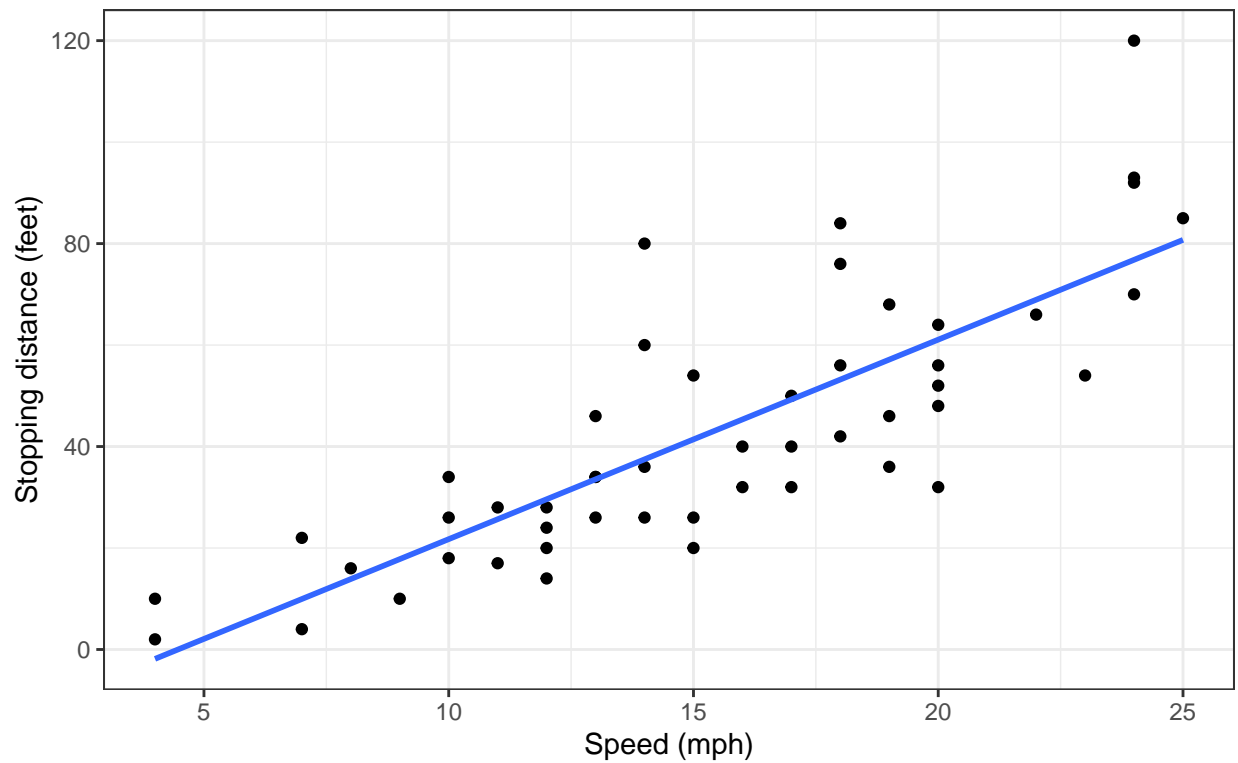
```
## `geom_smooth()` using formula 'y ~ x'
```

```r
# add in some labels/titles and black and white theme
p + labs(title = "Speed vs Stopping Distance of Cars",
         x = "Speed (mph)", y = "Stopping distance (feet)",
         caption = "Dataset: 'cars'") +
  theme_bw() +
  geom_smooth(method = "lm", se = F)
```

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

Speed vs Stopping Distance of Cars

Dataset: 'cars'

```
# add in some aesthetics witha new dataset
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
```

```
# Get some information about this dataset
nrow(genes)
```

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

```
ncol(genes)
```

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

```
colnames(genes)
```

```
## [1] "Gene"       "Condition1" "Condition2" "State"
```
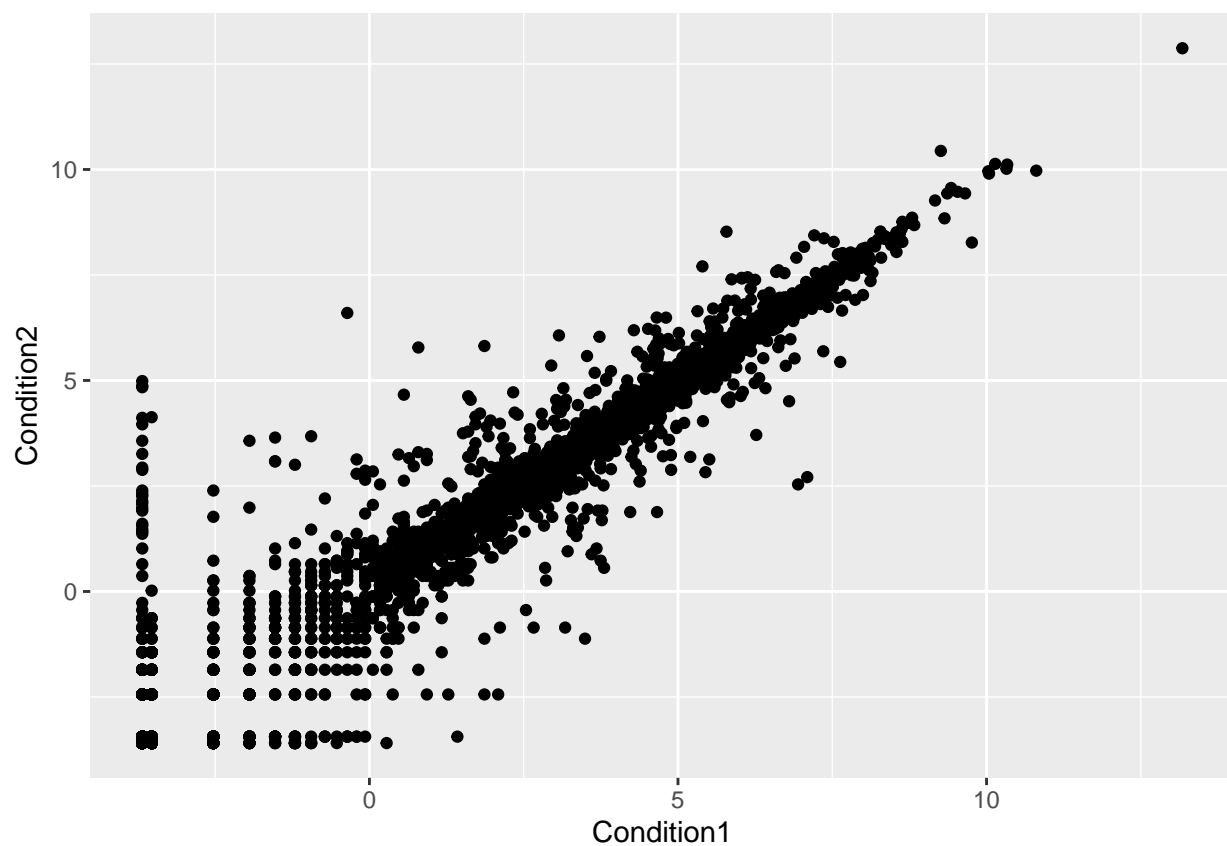
```
table(genes$State)
```

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

```
# percent of total genes in each state with two significant figures
round( table(genes$State)/nrow(genes) *100, 2)
```
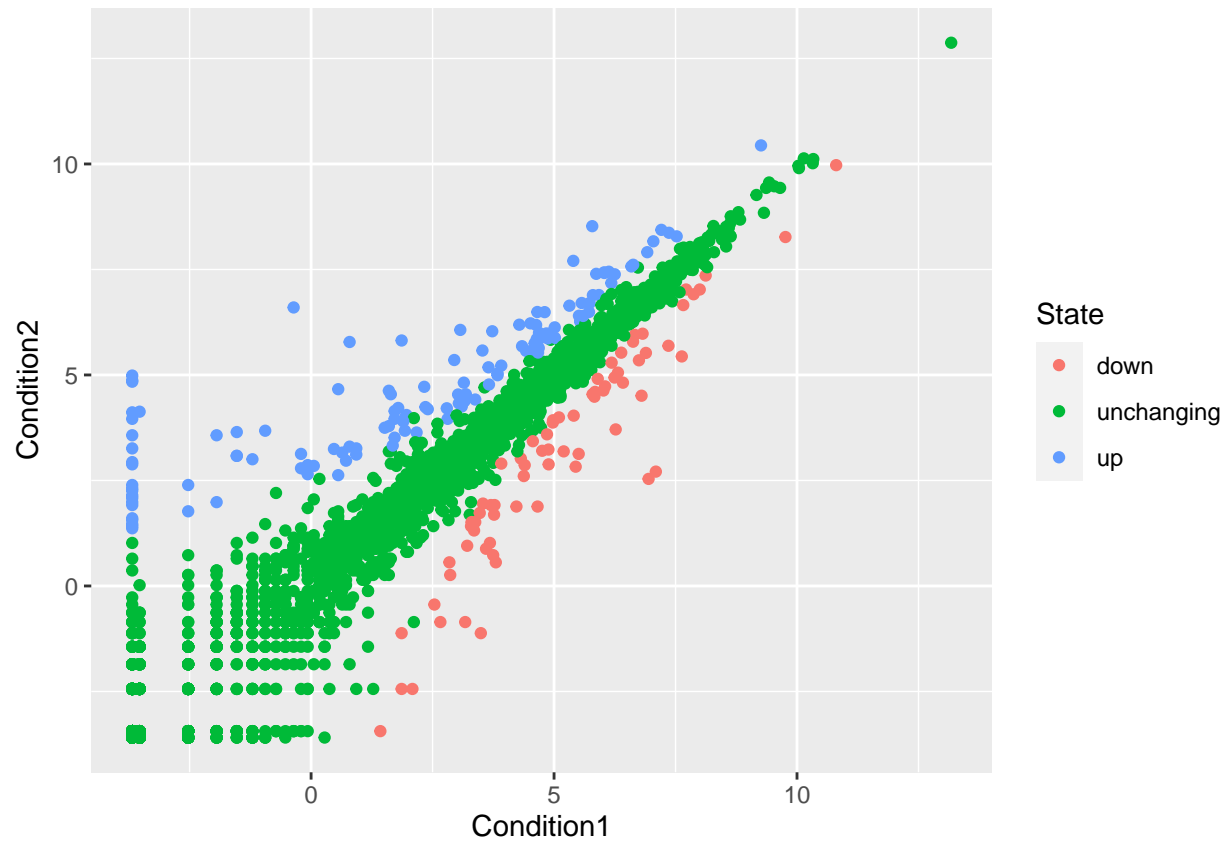
```
##
##       down unchanging         up
##       1.39      96.17       2.44
```

```
#plot the data from genes dataset
x <- ggplot(genes)+
  aes(x = Condition1, y = Condition2)+
  geom_point()
x
```

```
# color this plot by State (up, down, unchanging)
x <- ggplot(genes)+
  aes(x = Condition1, y = Condition2, col = State)+
  geom_point()
x
```



```
# Add some custom colors and labels
x <- x + scale_color_manual( values = c("blue", "gray", "red")) +
  labs(title = "Gene Expression Changes Upon Drug Treatment",
       x = "Control (no drug)", y = "Drug Treatment")
x
```

Gene Expression Changes Upon Drug Treatment