# week04 Data Visualization Lab

Muhammad Tariq (A15583449)

## Week 4 Data Visualization Lab

## Install the package ggplot2

```
options(repos = c(CRAN = "https://cloud.r-project.org"))
install.packages("ggplot2")
```

```
Installing package into 'C:/Users/Xain7/AppData/Local/R/win-library/4.4'
(as 'lib' is unspecified)
```

```
package 'ggplot2' successfully unpacked and MD5 sums checked
```

```
The downloaded binary packages are in
    C:\Users\Xain7\AppData\Local\Temp\RtmpUbQQtu\downloaded_packages
```
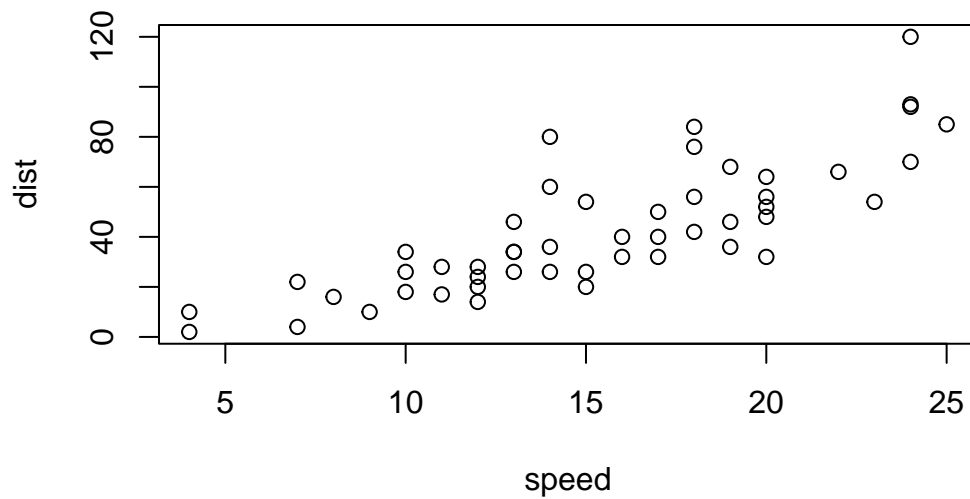
## Anytime I want to use this package I need to load it

```
library(ggplot2)
```

```
View(cars)
```
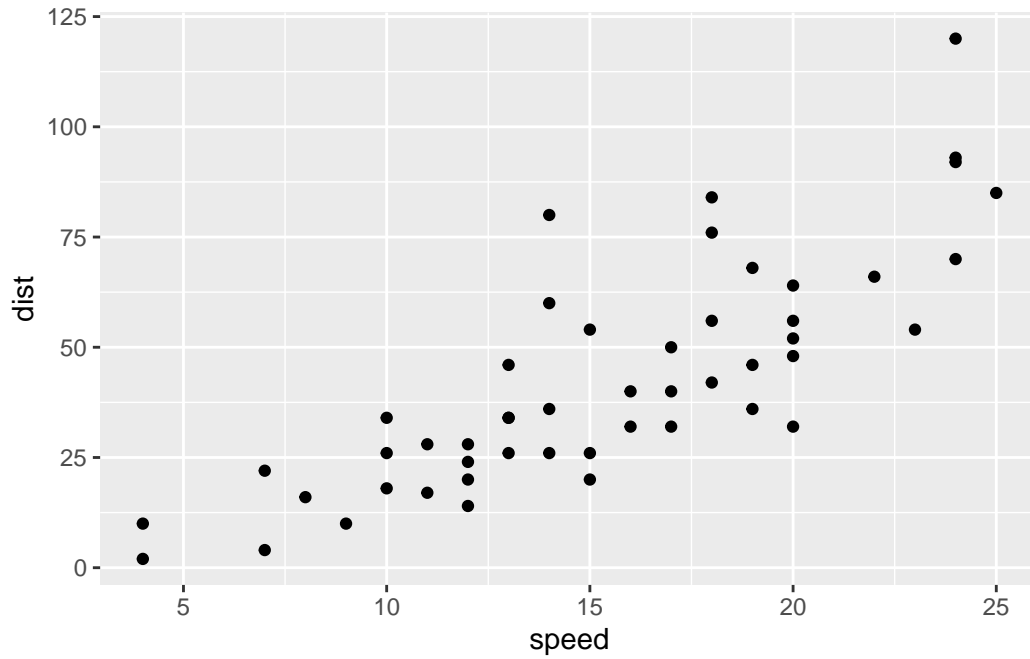
## A quick base R plot - this is not ggplot2

```
plot(cars)
```


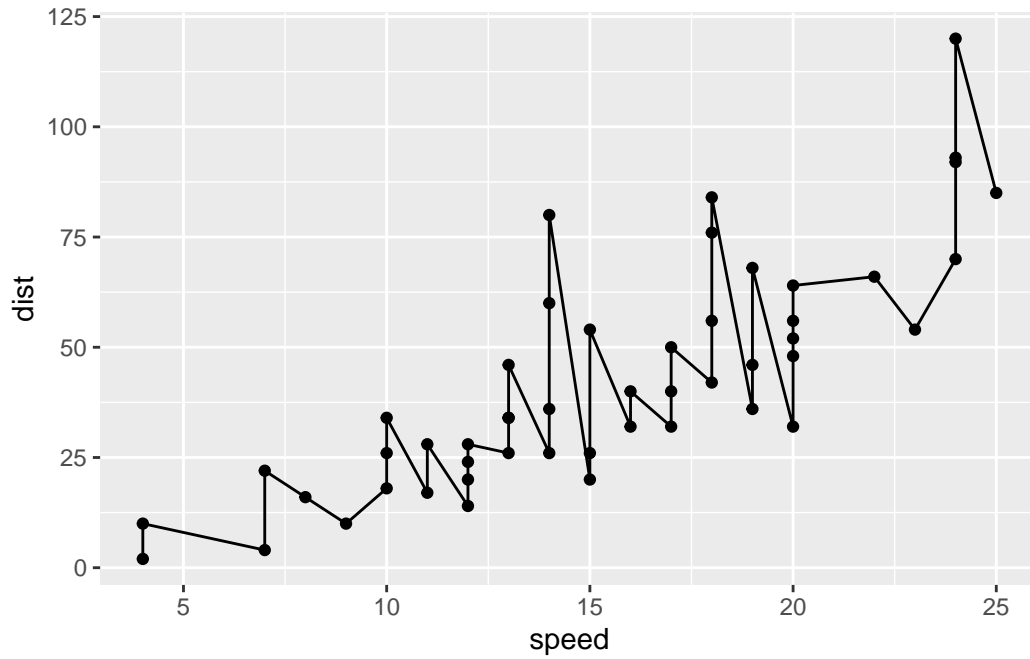
## Our first ggplot

**we need data + aes + geoms**

```
ggplot(data = cars) +
  aes(x=speed, y=dist) +
geom_point()
```

```
p <- ggplot(data = cars) +
  aes(x=speed, y=dist) +
  geom_point()
```
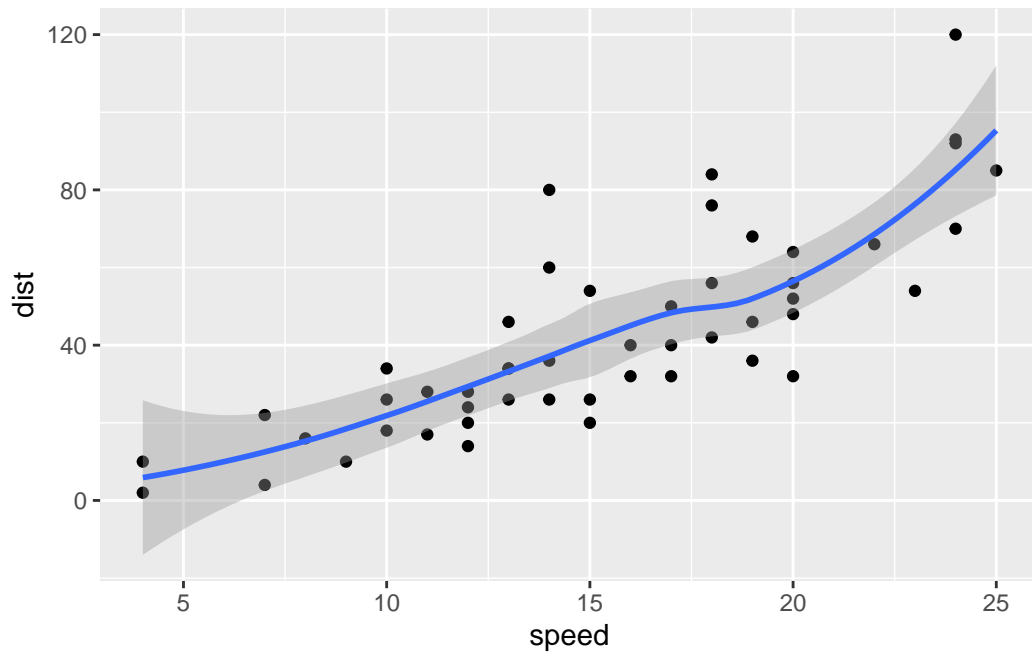
## Add a line geom with geom_line()

```
p + geom_line()
```
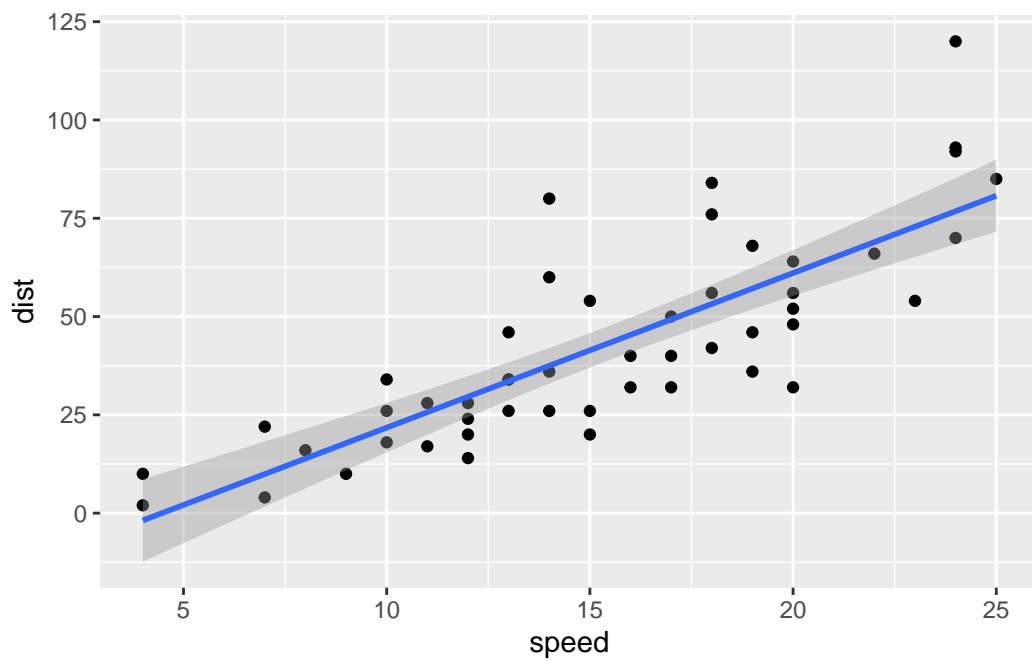
## Add a trend line close to the data

```
p + geom_smooth()
```

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

```
p + geom_smooth(method = "lm")
```

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

*#——#*

## Read in our durg expression data

```
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
```
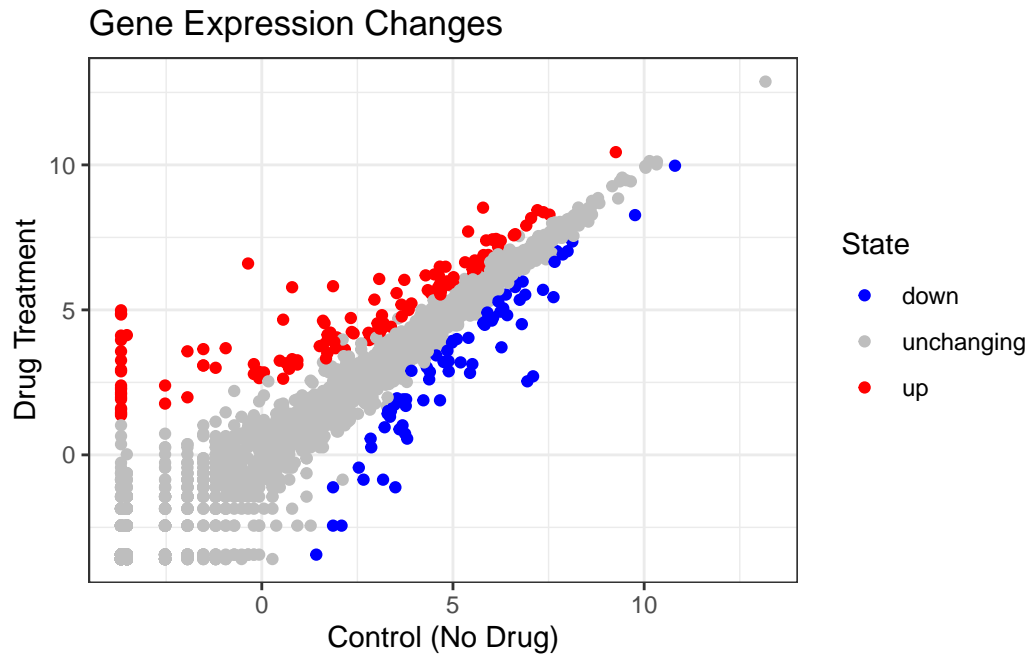
## Lets make a first plot attempt

```
g<- ggplot(data = genes) +
  aes(x=Condition1, y=Condition2, col=State) +
  geom_point()
g
```
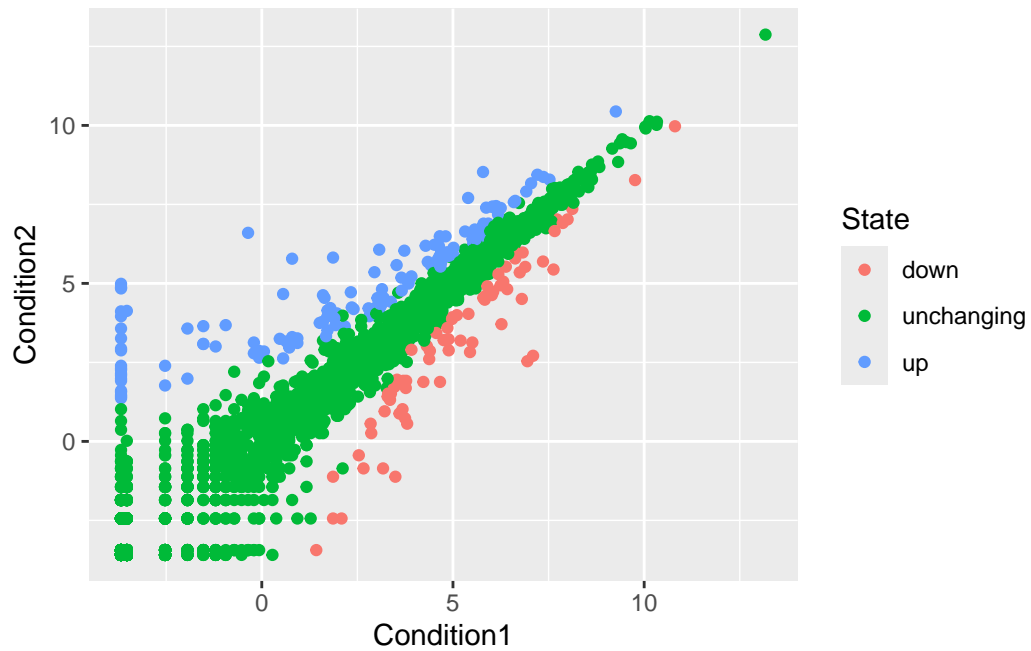
## Add some color

```
g + scale_color_manual(values = c("blue", "grey", "red")) +
  labs(title = "Gene Expression Changes",
       x="Control (No Drug)", y="Drug Treatment") +
  theme_bw()
```

Gene Expression Changes

## Q. How many genes are in this data set

```
nrow(genes)
```

```
[1] 5196
```

## Q. How many columns

```
ncol(genes)
```

```
[1] 4
```

#Q How many upregulated genes

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

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

## Q. What fraction of total genes are up regulated?

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

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

'