

# Class 5: Data Visualization with GG Plot

Candyd Lace R. Velasquez

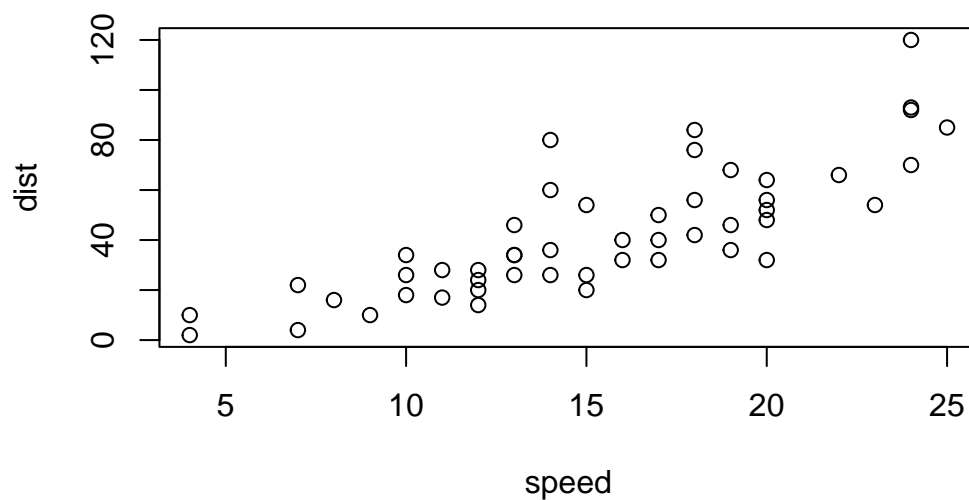
## Table of contents

<b>Our first plot</b>	<b>1</b>
A more interesting plot . . . . .	4

## Our first plot

R has base graphics

```
plot(cars)
```



How would I plot this with `ggplot2`? No, we need to install `ggplot` package first

To install any package use `install.packages()` function make sure to load up the package in the session you are using use `library()` function

```
#install.packages("ggplot2")
```

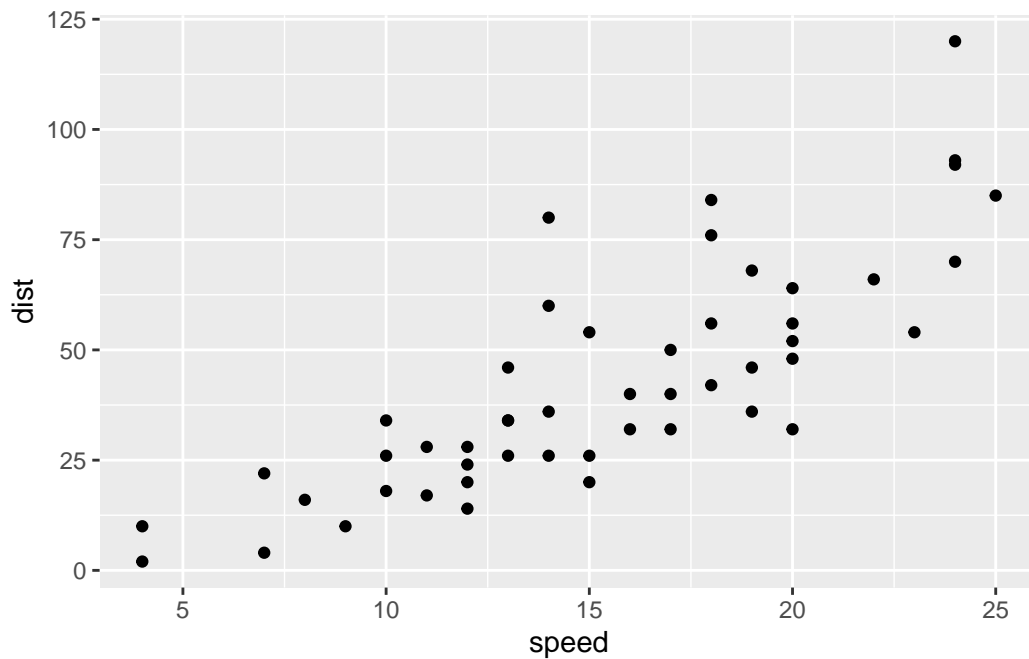
Every `ggplot` needs at least 3 layers :

-**Data**(ie the data.frame we have),

-**Aes** (the aesthetic mapping of our data to what we want to plot)

-**Geoms** (How we want to plot this stuff!)

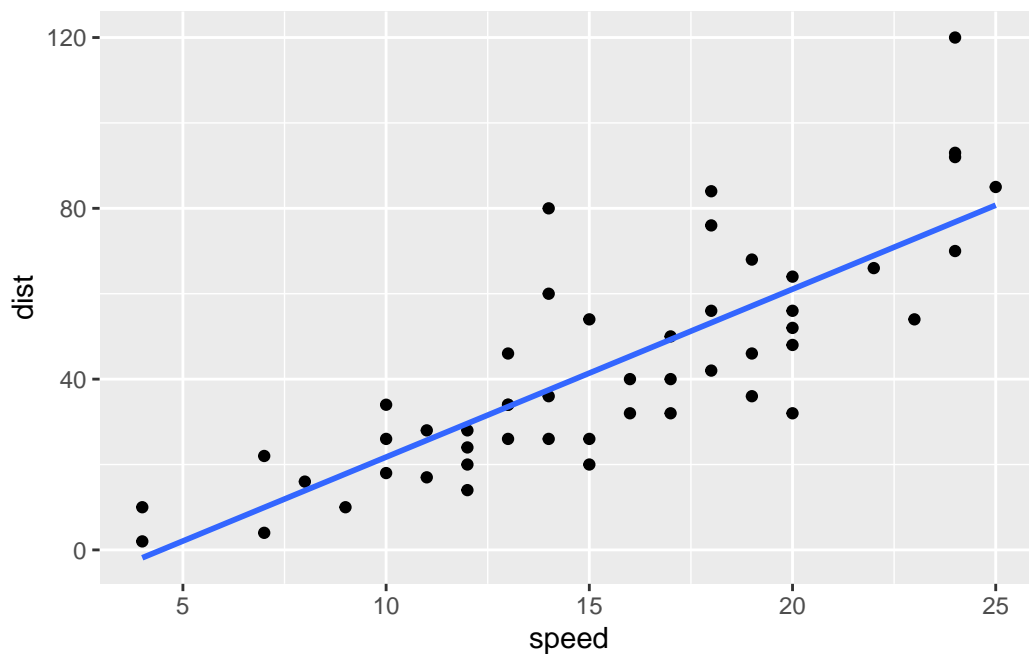
```
library(ggplot2)
ggplot(data=cars) +
  aes(x=speed, y=dist) +
  geom_point()
```



```
library(ggplot2)
```

```
ggplot(data=cars) +
  aes(x=speed, y=dist) +
  geom_point() +
  geom_smooth(method = lm, se = FALSE)
```

`geom\_smooth()` using formula 'y ~ x'



```
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

## A more interesting plot

First read

Q1. how many genes are in this dataset

```
nrow(genes)
```

```
[1] 5196
```

```
colnames(genes)
```

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

```
ncol(genes)
```

```
[1] 4
```

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

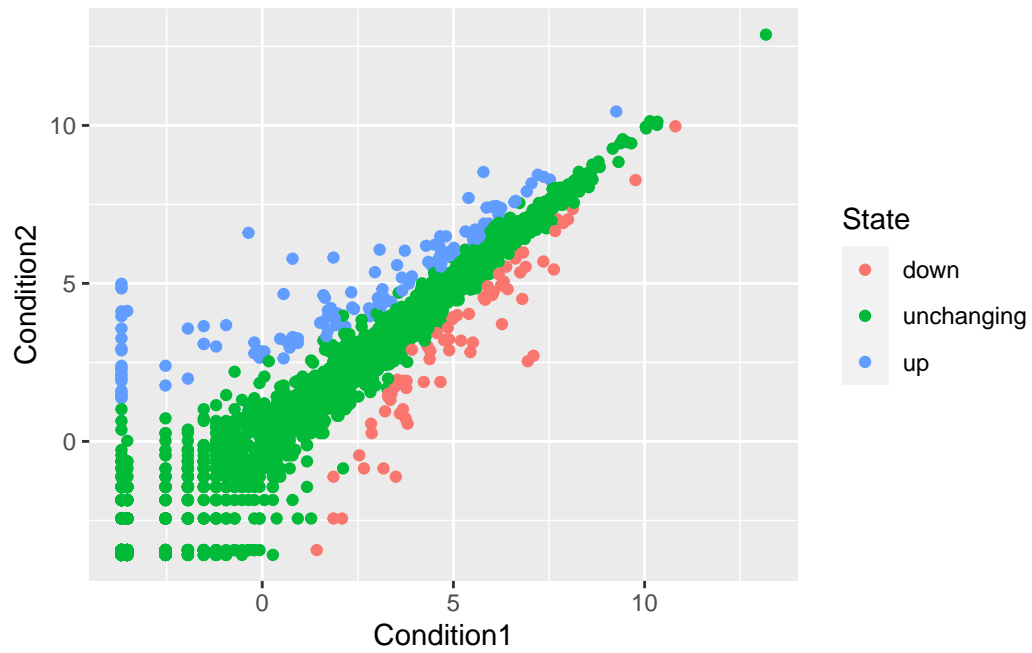
down	unchanging	up
72	4997	127

```
round(table(genes$State)/nrow(genes) * 100, 3)
```

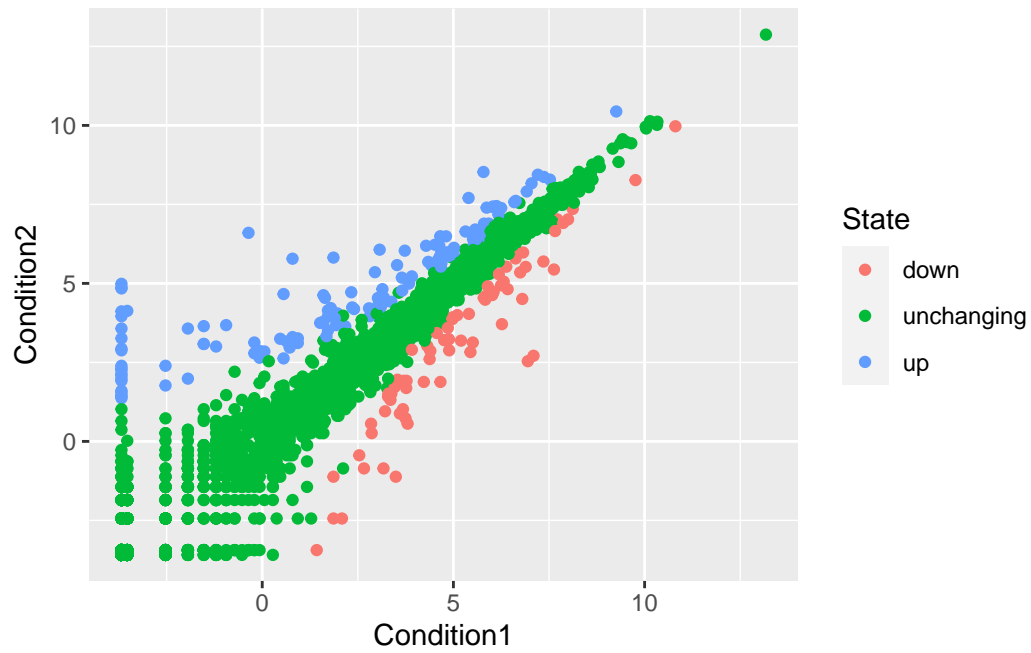
down	unchanging	up
1.386	96.170	2.444

There are 5196 in this data set

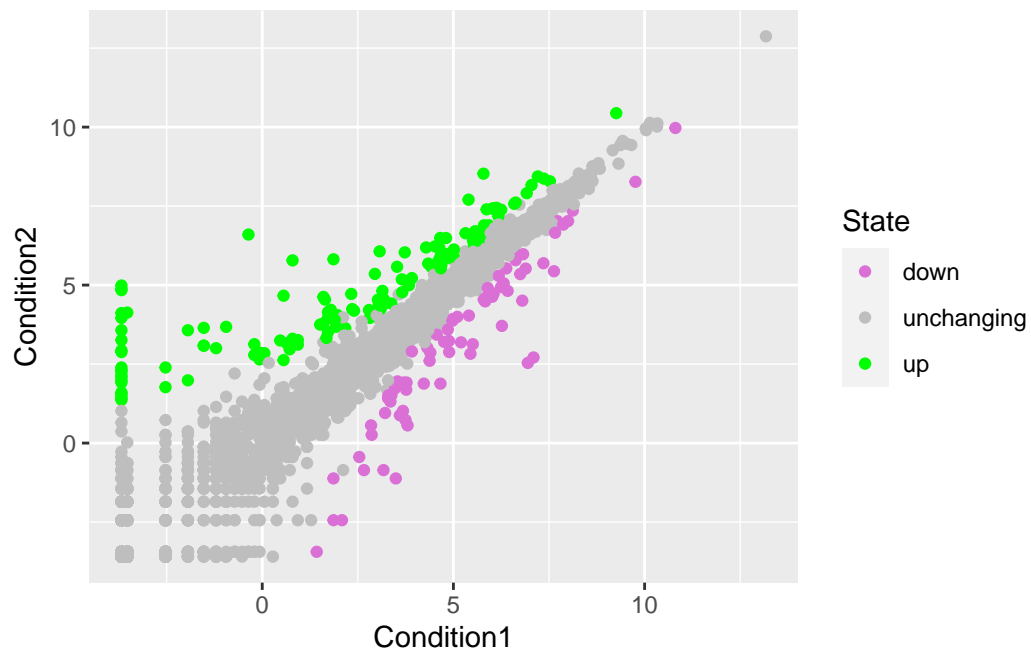
```
ggplot(data=genes) +  
  aes(x=Condition1, y=Condition2, col=State)+  
  geom_point()
```



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



```
p + scale_colour_manual( values=c("orchid","gray","green") )
```



```
p + scale_colour_manual(values=c("orchid", "gray", "green")) +
  labs(title="Gene Expression Changes Upon Drug Treatment",
       x="Control (no drug) ",
       y="Drug Treatment")
```

