

Class 5 - ggplot

Benjamin Lee

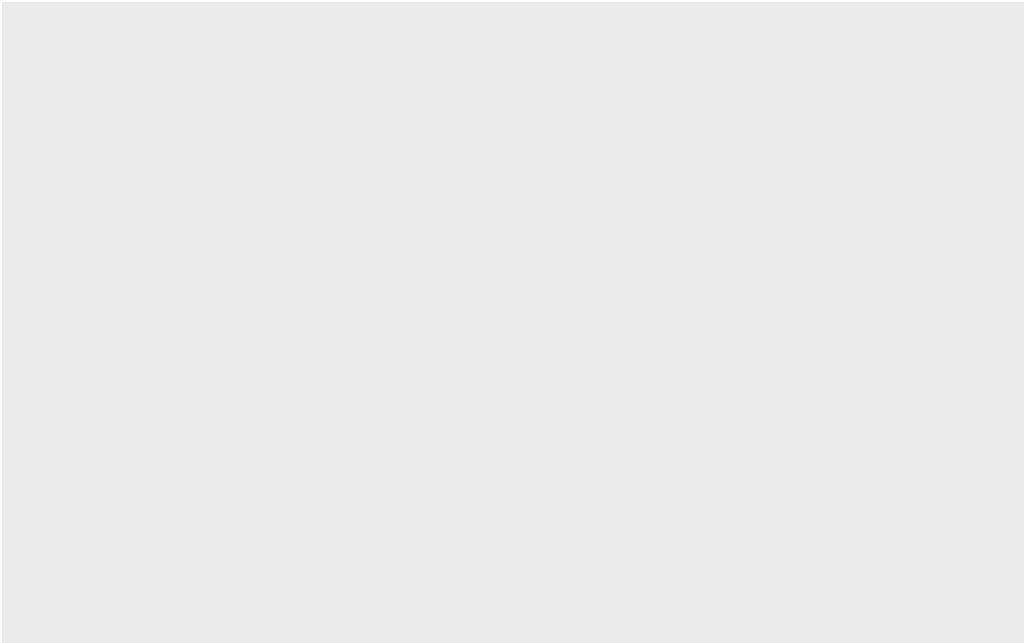
Our first ggplot

To use the ggplot2 package I first need to have it installed on my computer

To install any package we use the 'install.packages()' command.

Now can I use it? NO ! first we need to call 'library(ggplot2)'

```
library(ggplot2)  
ggplot()
```



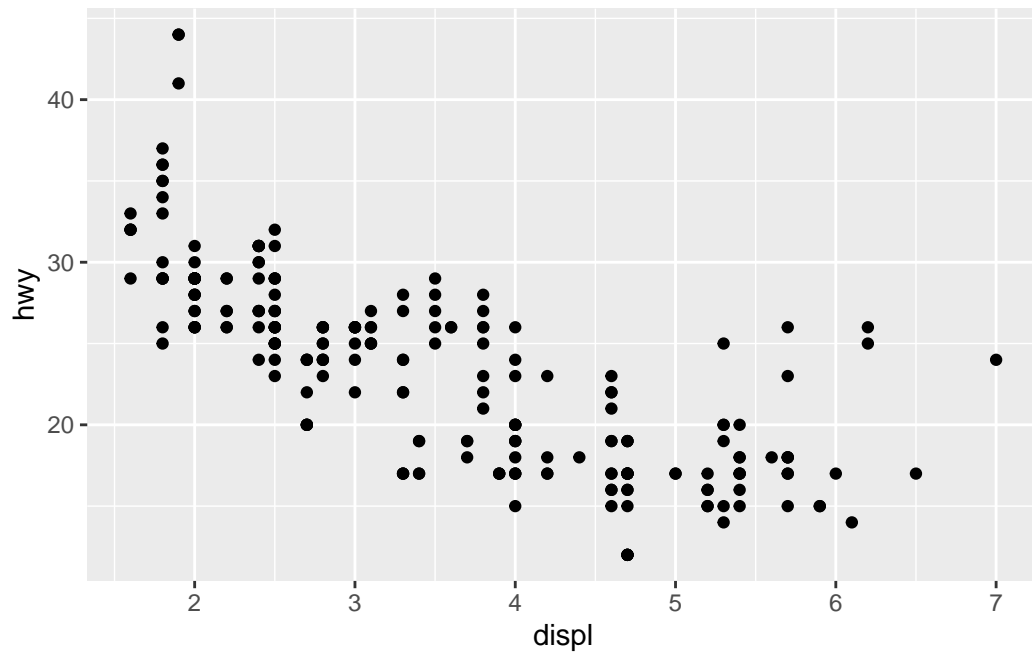
mpg

```
# A tibble: 234 x 11
  manufacturer model      displ  year   cyl trans drv     cty   hwy fl      class
  <chr>          <chr>    <dbl> <int> <int> <chr> <chr> <int> <int> <chr> <chr>
1 audi          a4         1.8  1999     4 auto~ f      18    29 p    comp~
2 audi          a4         1.8  1999     4 manu~ f      21    29 p    comp~
3 audi          a4         2    2008     4 manu~ f      20    31 p    comp~
4 audi          a4         2    2008     4 auto~ f      21    30 p    comp~
5 audi          a4         2.8  1999     6 auto~ f      16    26 p    comp~
6 audi          a4         2.8  1999     6 manu~ f      18    26 p    comp~
7 audi          a4         3.1  2008     6 auto~ f      18    27 p    comp~
8 audi          a4 quattro  1.8  1999     4 manu~ 4      18    26 p    comp~
9 audi          a4 quattro  1.8  1999     4 auto~ 4      16    25 p    comp~
10 audi         a4 quattro   2    2008     4 manu~ 4      20    28 p    comp~
# ... with 224 more rows
```

Our first plot of displ vs hwy All ggplot() graphs are made in the same way.

- data + aes + geoms

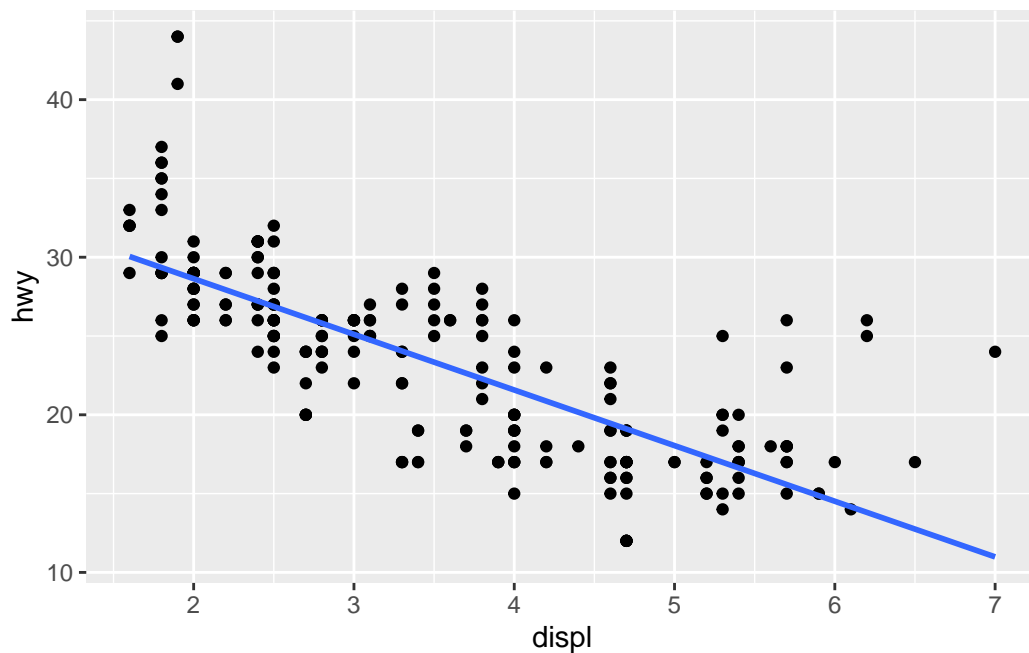
```
ggplot(mpg) +
  aes(x=displ, y=hwy) +
  geom_point()
```



I can add more layers:

```
ggplot(mpg) +  
  aes(x=displ, y=hwy) +  
  geom_point() +  
  geom_smooth(method = lm, se = FALSE)
```

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



Plot of gene expression data

First read the data from online.

```
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

Q. How many genes are in this dataset?

```
colnames(genes)
```

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

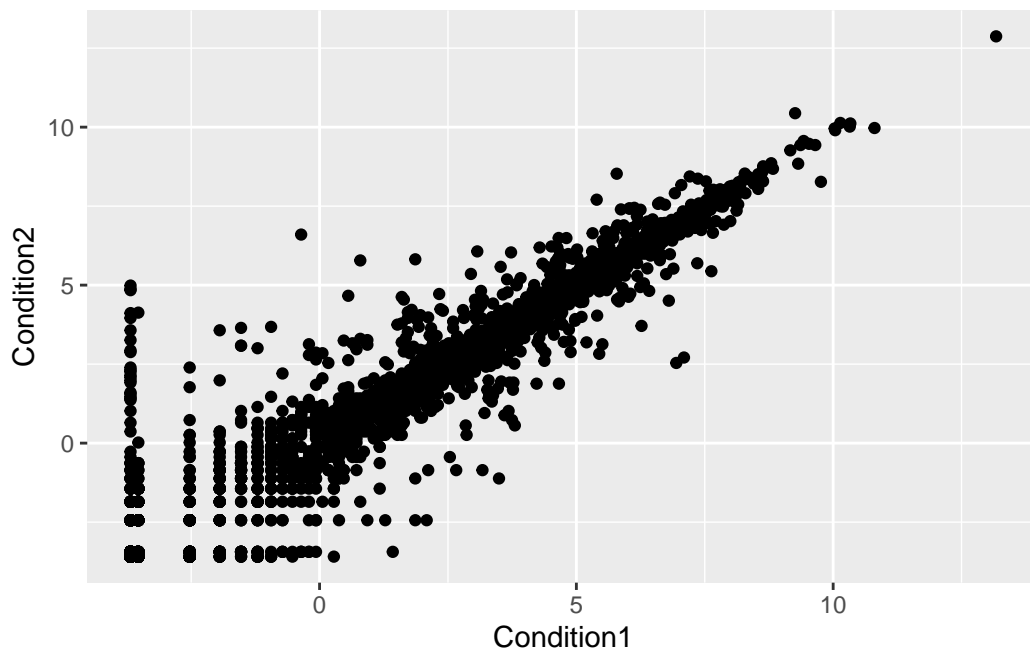
What are the colnames?

```
colnames(genes)
```

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

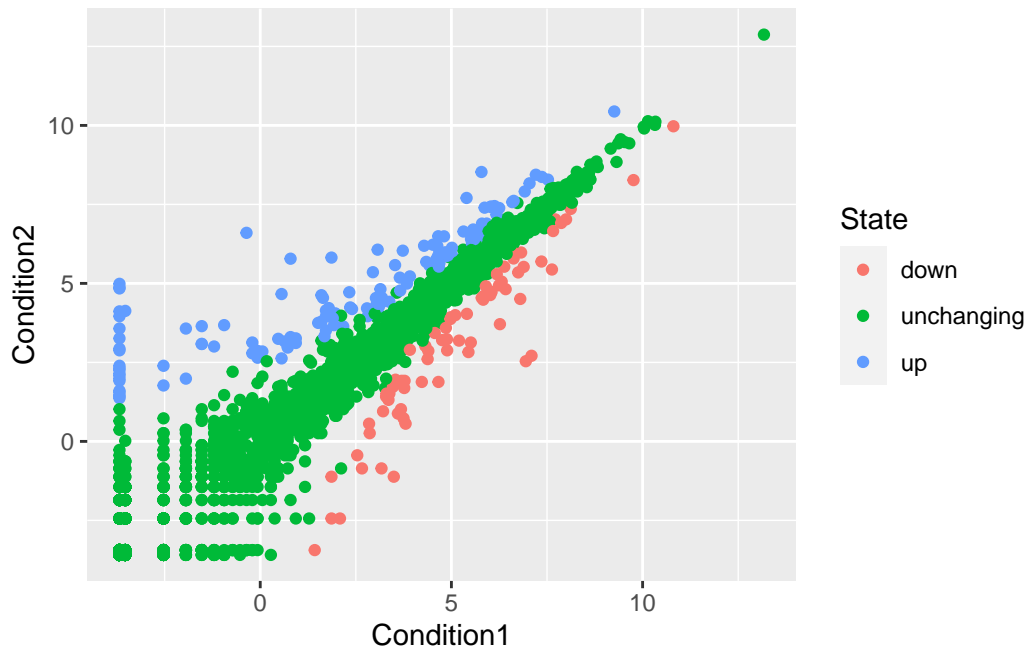
A first version of this data Condition1 vs Condition2

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



Let's add some color. To do this we can add another `aes()` mapping of color the 'State' column in our data.

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



Q. How many genes are upregulated and downregulated?

```
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

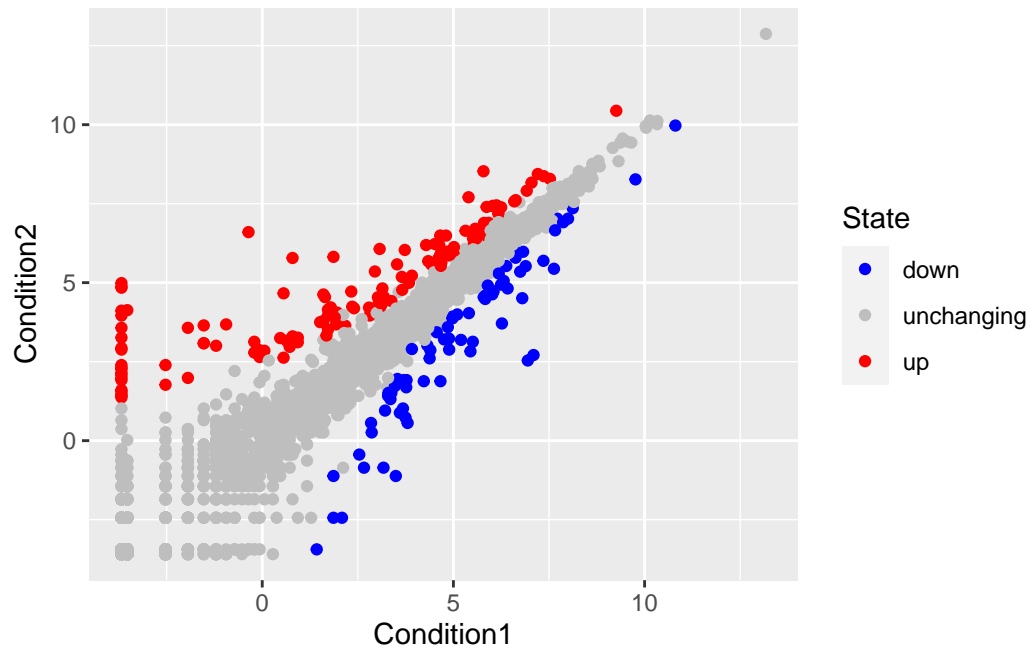
To get at just the State column

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

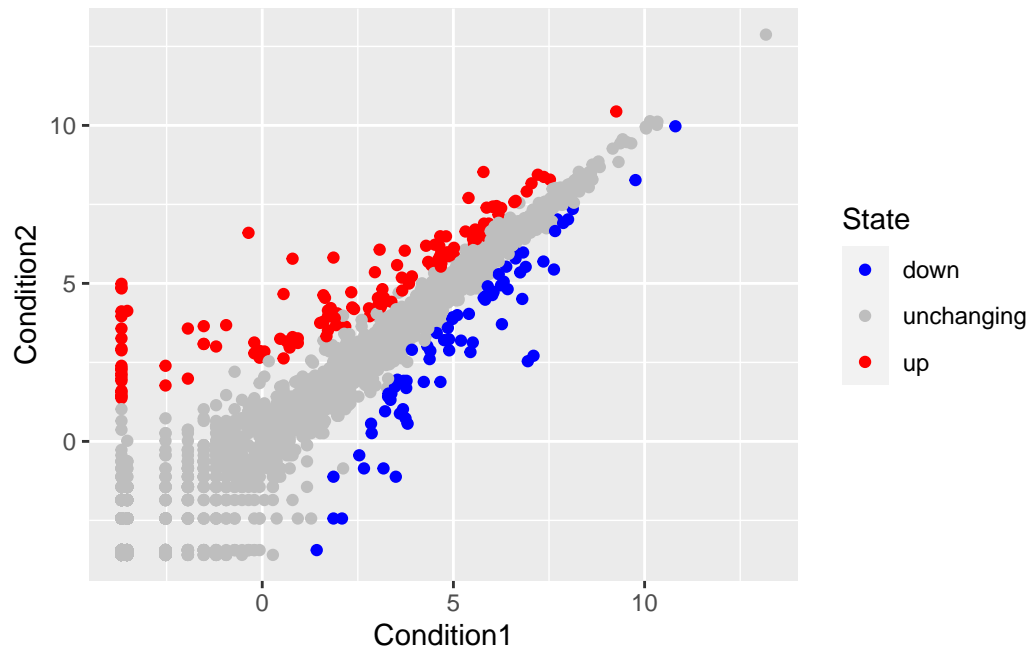
down	unchanging	up
72	4997	127

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

p + scale_colour_manual( values=c("blue","gray","red") )
```



```
ggplot(genes) +
  aes(x=Condition1, y=Condition2, col=State) +
  geom_point() +
  scale_colour_manual( values=c("blue","gray","red") )
```



Then just add to our object 'p'

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


Gene Expression Changes Upon Drug Treatment

