

Class 05 BIMM 143

AUTHOR

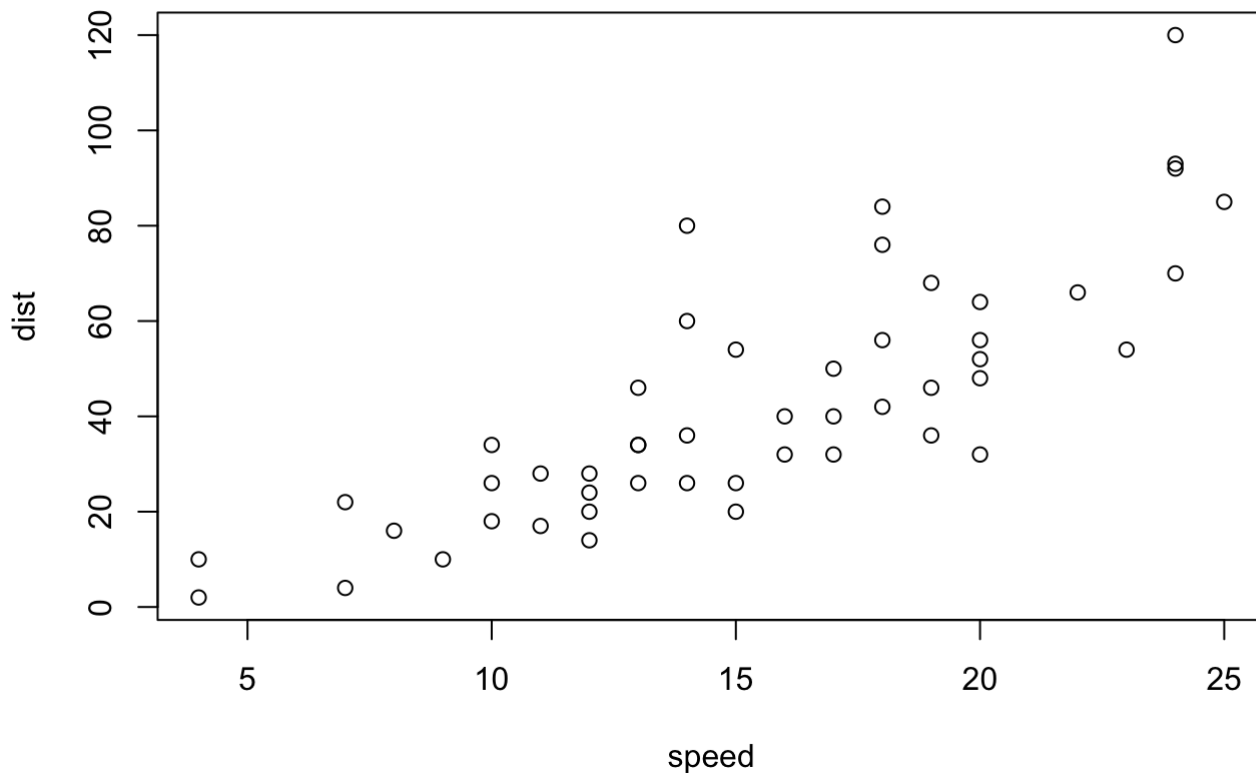
Arturo Agramont

PUBLISHED

April 19, 2023

We are going to start by generating the plot of class 04

```
plot(cars)
```



GGplot2

First we need to install the package. We do this by using the `install.packages` command.

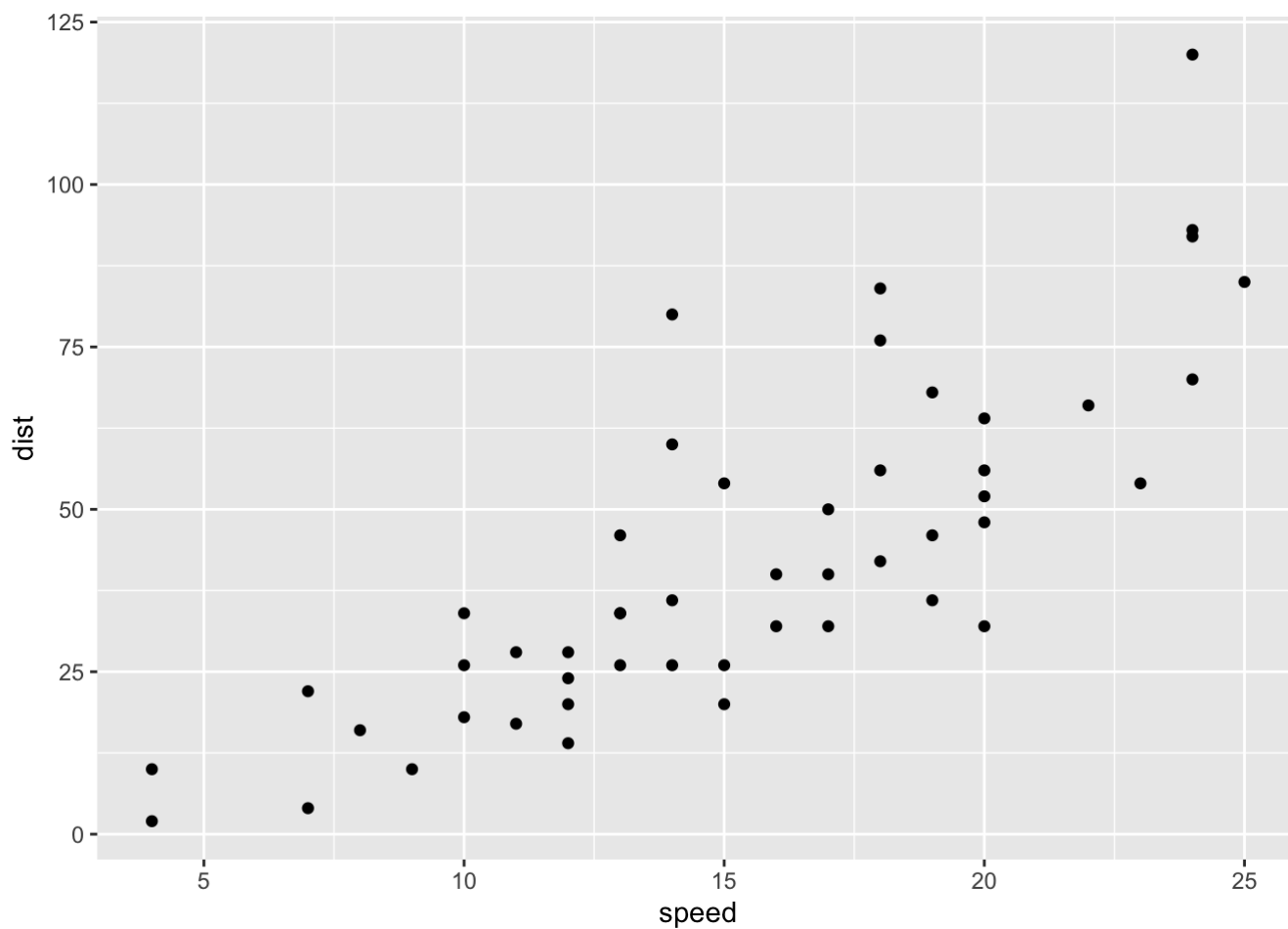
```
#install.packages("ggplot2")
```

After that we need to load the package

```
library("ggplot2")
```

We are going to build the plot of the cars data frame by using ggplot2.

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

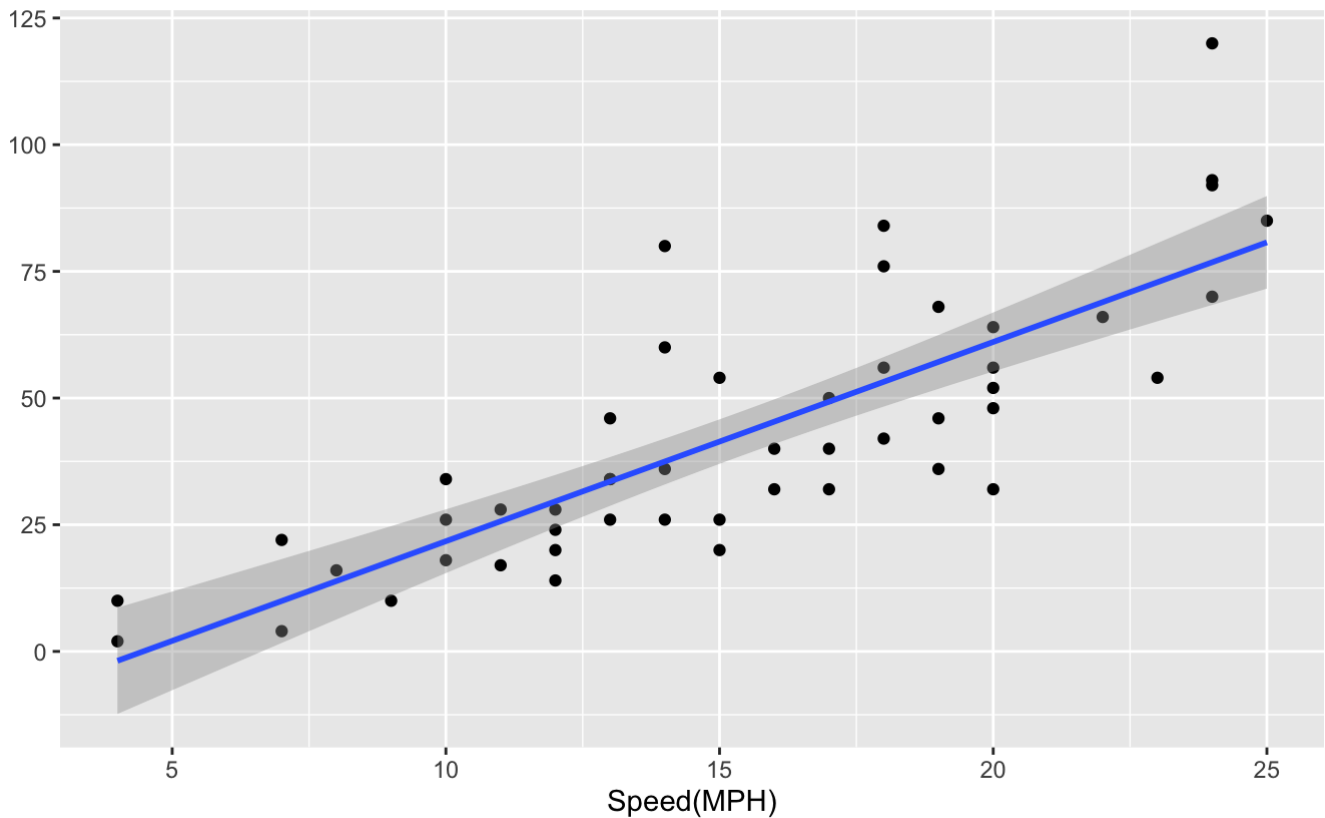


```
ggplot(data=cars) + aes(x=speed, y=dist) + geom_point() + geom_smooth(method = 'lm') + la
```

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

Distance Vs. Speed

A nice subtitle!



BIMM143

Q1. For which phases is data visualization important in our scientific workflows?

Communication of Results, Exploratory Data Analysis (EDA), and Detection of outliers

Q2. True or False? The ggplot2 package comes already installed with R?

False

Q. Which plot types are typically NOT used to compare distributions of numeric variables?

Network Graphs

Q. Which statement about data visualization with ggplot2 is incorrect?

ggplot2 is the only way to create plots in R.

Plotting Expression Data

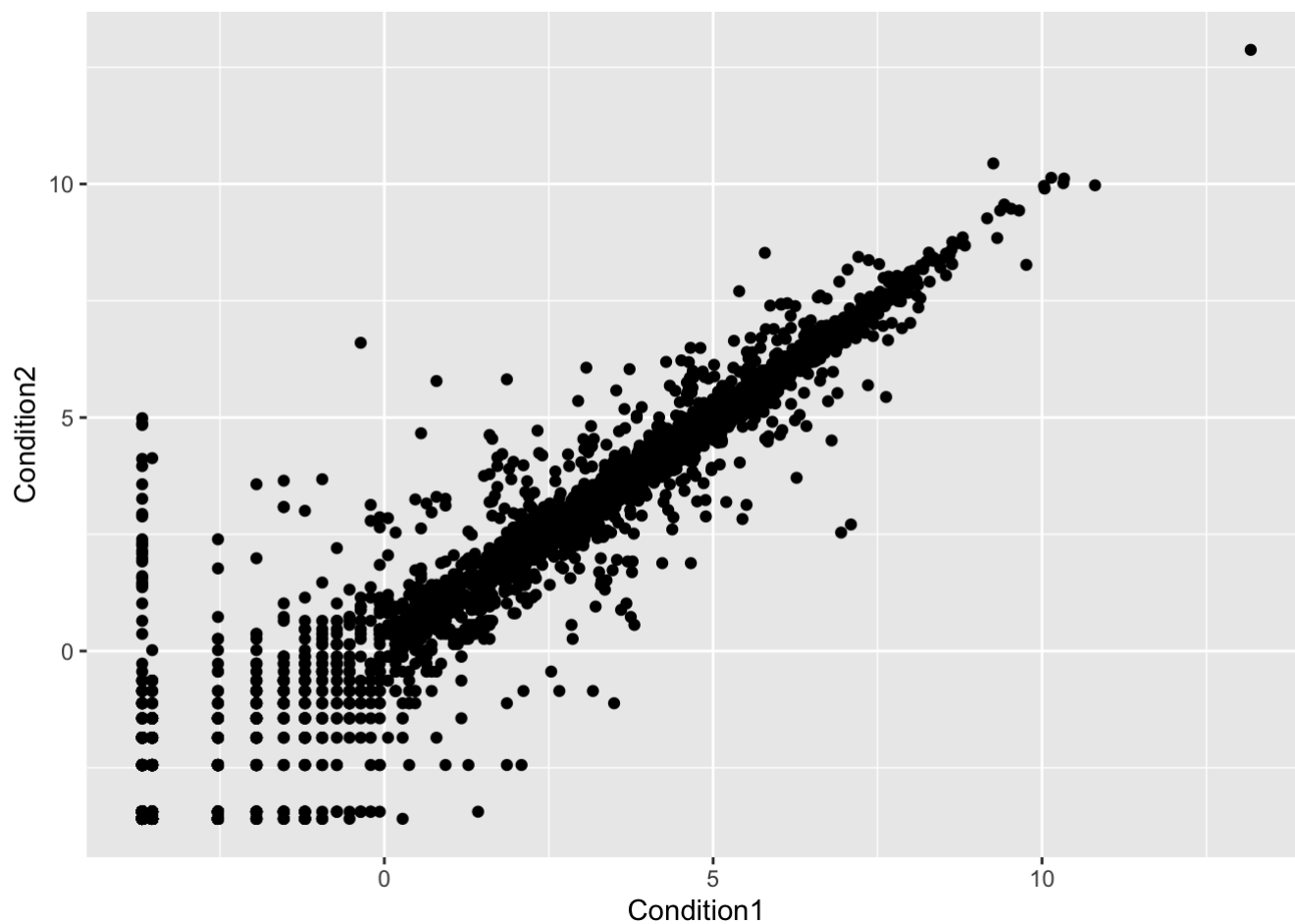
First we need to load the data from the URL.

```
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

Then make an initial plot using ggplot.

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



Finding the number of "up" regulated and the percentage

```
table(genes["State"])
```

State

down	unchanging	up
72	4997	127

```
round( table(genes["State"])/nrow(genes) *100, 2)
```

State

down	unchanging	up
72	4997	127

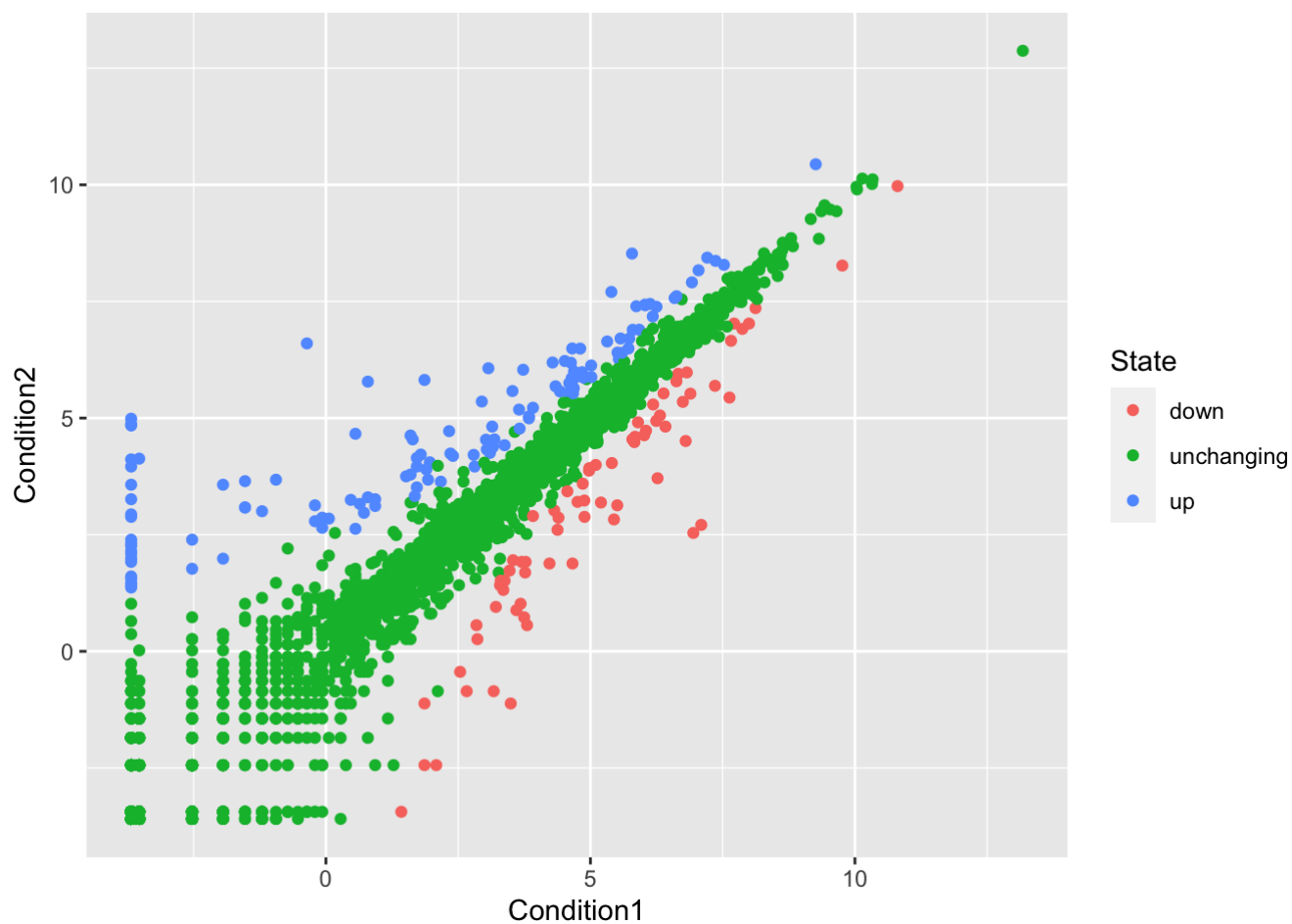
1.39

96.17

2.44

Adding Color and titles to the plot

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



```
p+ scale_color_manual( values = c("blue", "gray", "red")) + labs(title = "Gene expression
```

