

Class 5: Data Visualization with ggplot

Barry

Table of contents

Our first plot	1
A more interesting plot	5

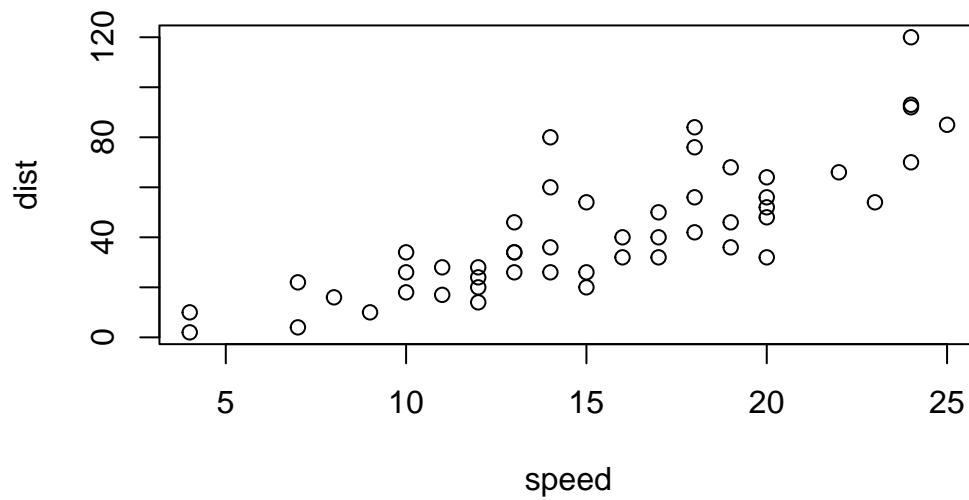
Our first plot

R has base graphics

```
head(cars)
```

```
  speed dist
1     4    2
2     4   10
3     7    4
4     7   22
5     8   16
6     9   10
```

```
plot(cars)
```



How would I plot this with `ggplot2`? NO!

We need to install and load the `ggplot2` package first. To install any package in R we use the `install.packages()` function.

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

Before I can use this package I need to load it with a `library()` call.

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



Every ggplot needs at least 3 layers:

- **Data** (i.e. 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!)

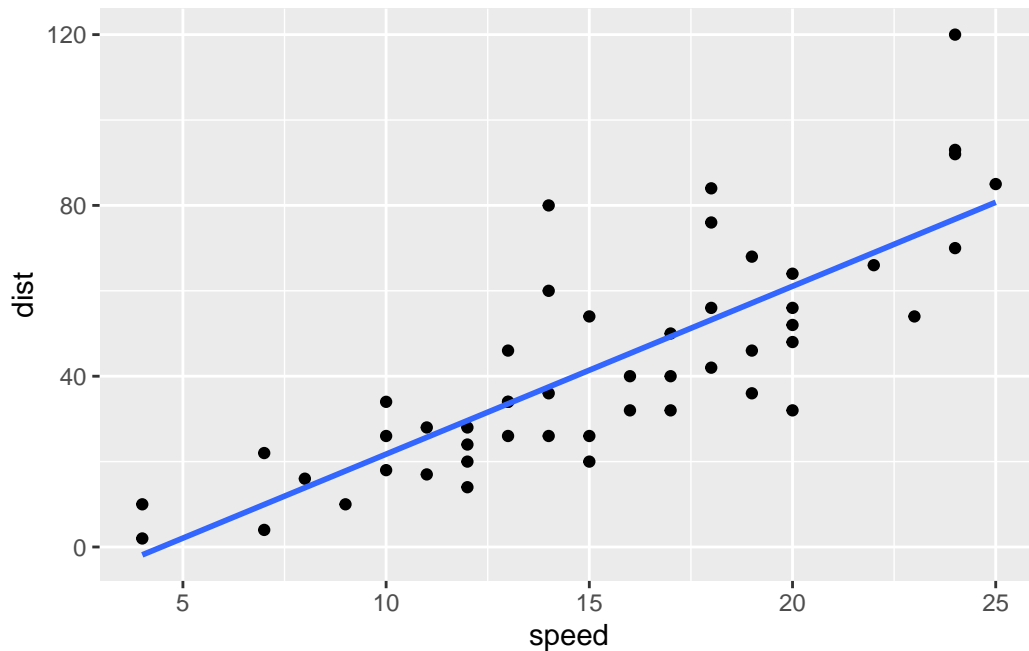
```
ggplot(data=cars) +  
  aes(x=speed, y=dist) +  
  geom_point()
```



Add another geom

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

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



A more interesting plot

First read 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

Q1. How many genes are in this dataset

```
nrow(genes)
```

```
[1] 5196
```

There are 5196 genes in this data set.

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

