

# Class 5: Data Visualization with ggplot

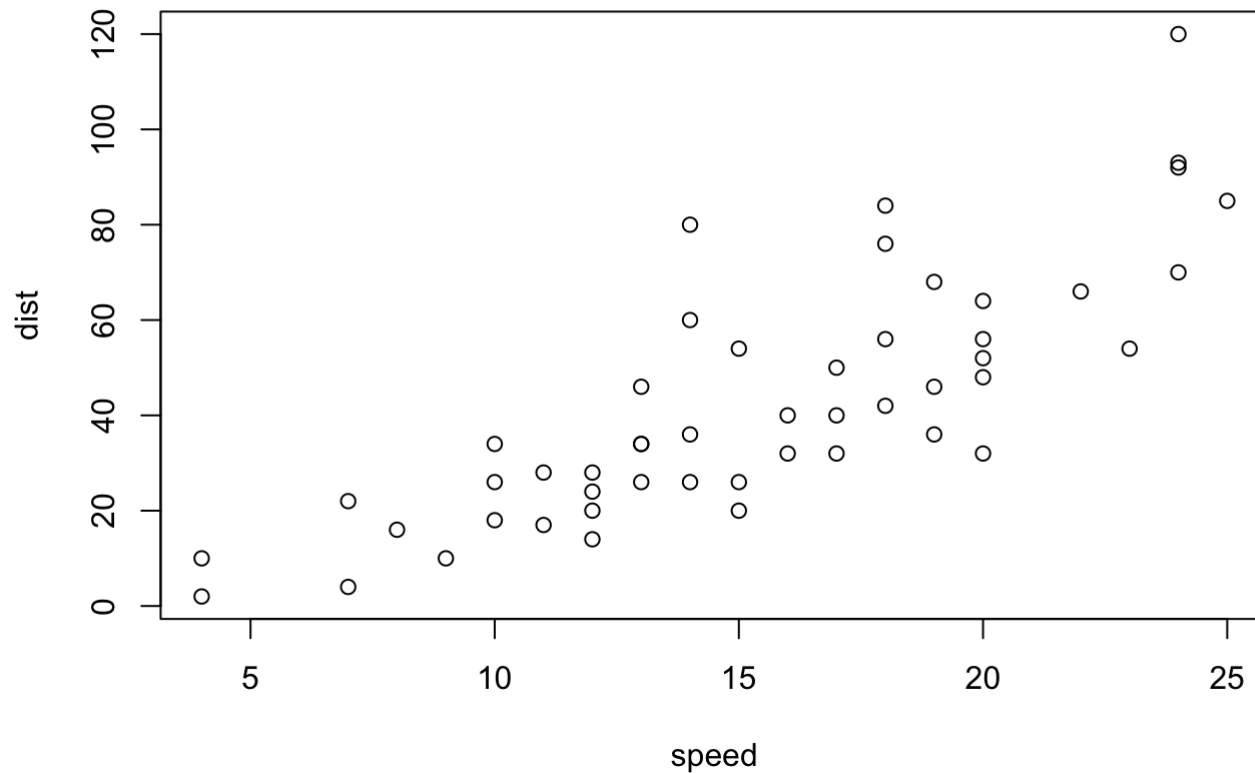
AUTHOR

Nichelle Camden

## Our first plot

R has base graphics

```
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, and put `ggplot` in quotes because it is a string.

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

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

```
library(ggplot2)
```

Registered S3 methods overwritten by 'tibble':

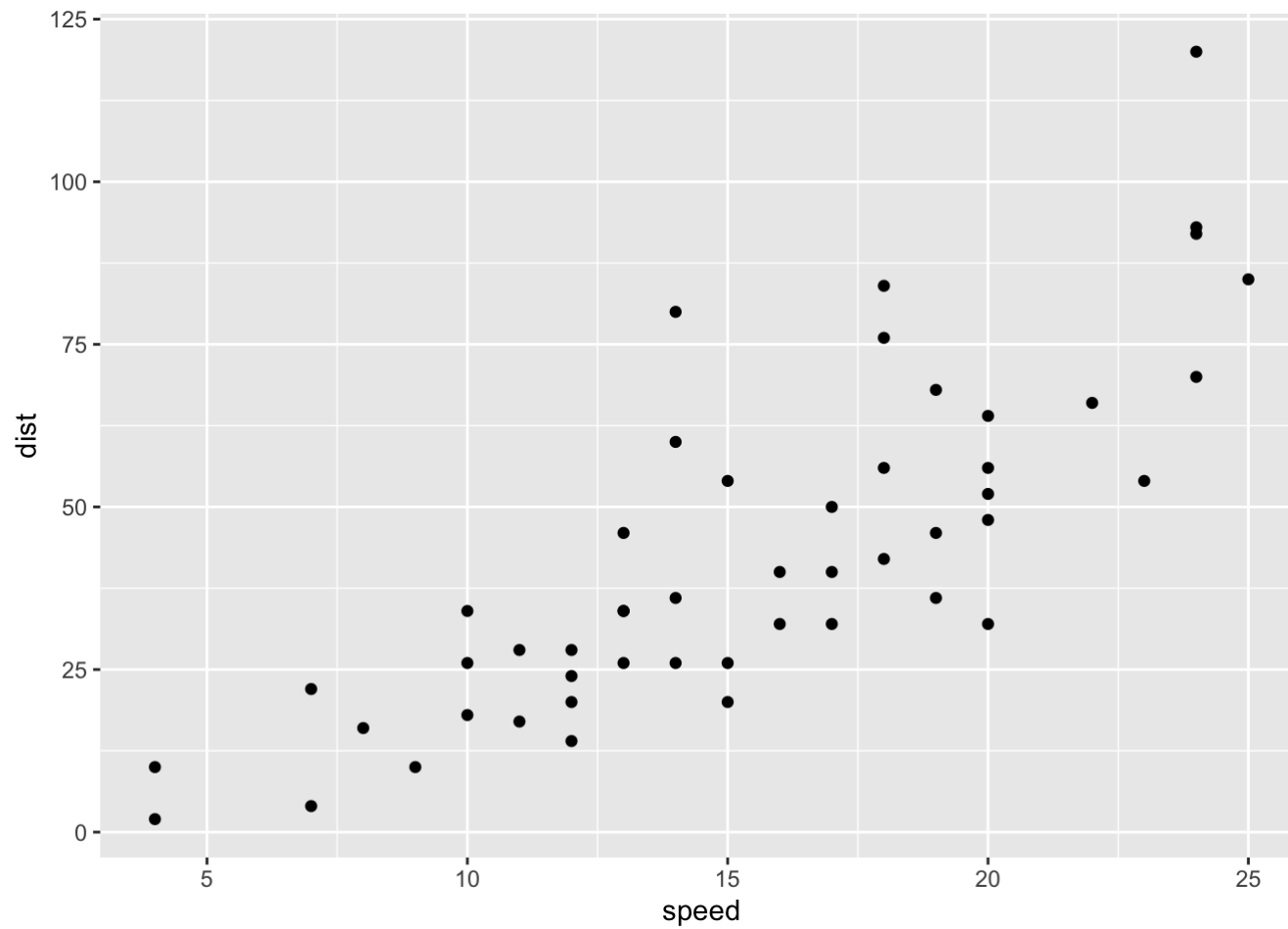
method	from
format.tbl	pillar
print.tbl	pillar

```
#ggplot(cars) <- won't work yet, needs more layers
```

Every `ggplot` needs at least 3 layers:

- **Data** (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)

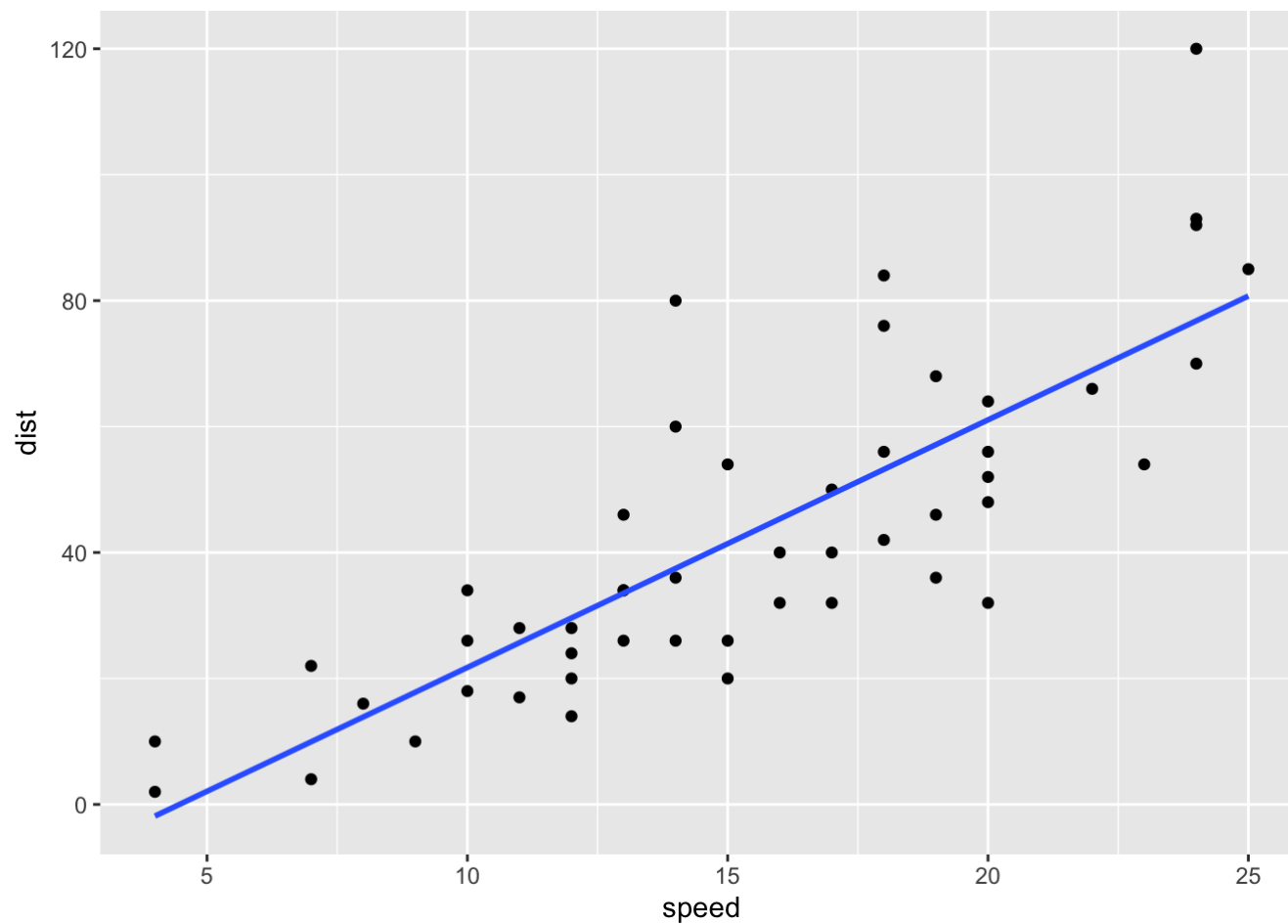
```
ggplot(data=cars) +  
  aes(x=speed, y=dist) + geom_point()
```



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

```
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
```

```
nrow(genes)
```

```
[1] 5196
```

```
colnames(genes)
```

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

```
ncol(genes)
```

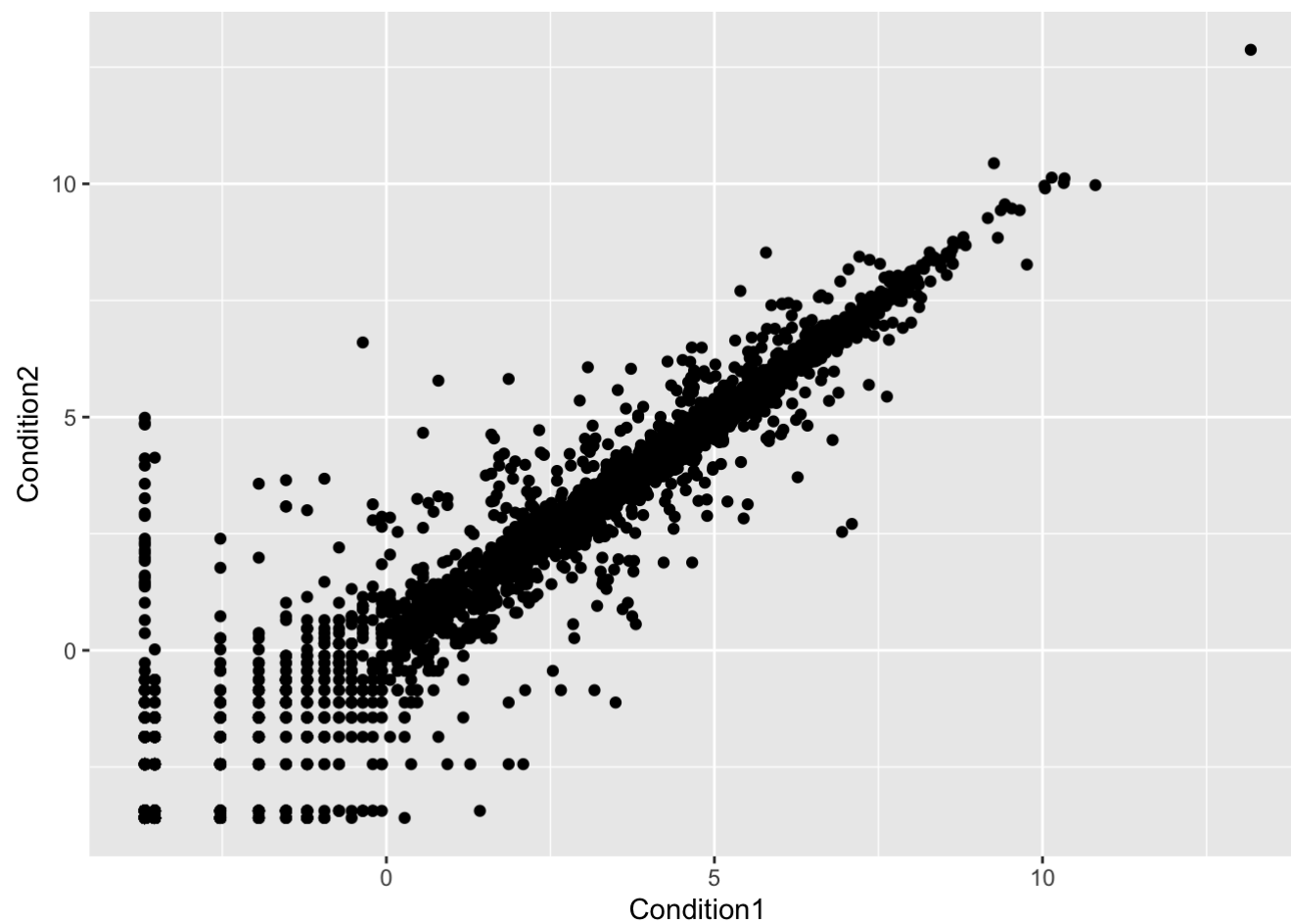
```
[1] 4
```

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

down	unchanging	up
72	4997	127

There are 5196 genes in this data set

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



```
p <- ggplot(genes) +  
  aes(x = Condition1, y = Condition2, col = State) +  
  geom_point()  
  
p + scale_colour_manual( values=c("blue","gray","red") ) +  
  labs(title = "Gene Expression Changes Upon Drug Treatment", x= "Control (no drug)", y= "Drug Treatment" )
```

