

Class 05: Data Visualization with ggplot2

Eric Jordahl

Table of contents

Our first plot	1
A More Interesting Plot	3

Our first plot

R has base graphics

```
plot(cars)
```



How would I plot this with `ggplot2`?

Every `ggplot` needs at least 3 layers:

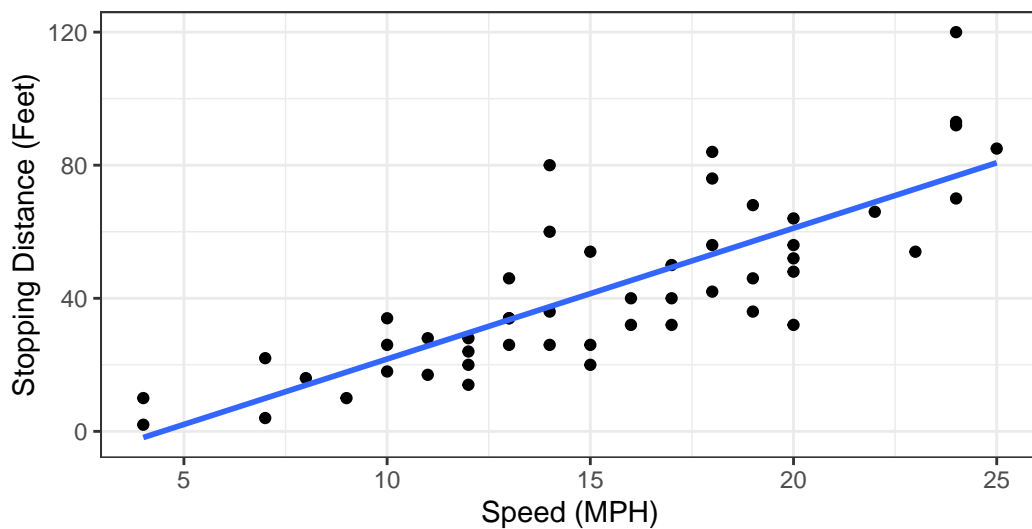
- **Data** (i.e. the `data.frame` we have),
- **Aes** (the aesthetic mapping of our data to what we want to plot)
- **Geoms** (How we want to plot this stuff!)

```
library(ggplot2)
ggplot(data=cars) + aes(x=speed, y=dist) + geom_point() + labs(title= "Car speed vs Stopping distance")
```

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

Car speed vs Stopping Distance

Eric Jordahl



Dataset: 'cars'

A More Interesting Plot

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

Q1. How many genes are in this data set?

There are 5196 genes in this data set.

Q2. How many columns are in this data set?

There are 4 columns in this data set with names Gene, Condition1, Condition2, State

Q3. How many upregulated genes are there?

```
up_genes <- table(genes$State=="up")
up_genes[2]
```

TRUE

127

The number of upregulated genes is 127

Q4. What fraction of total genes is up-regulated in this data set?

The fraction of total genes upregulated is 2.4441878

```
library(ggplot2)
p <- ggplot(data=genes) +
  aes(x=Condition1, y=Condition2, col=State) +
  geom_point(alpha=0.3) +
  labs(title="Gene Expression Changes Upon Drug Treatment", x="Control (No Drug)", y="Drug Treatment")
p + scale_colour_manual( values=c("skyblue", "gray", "orchid") )
```

