

Class 05 Data Viz with ggplot

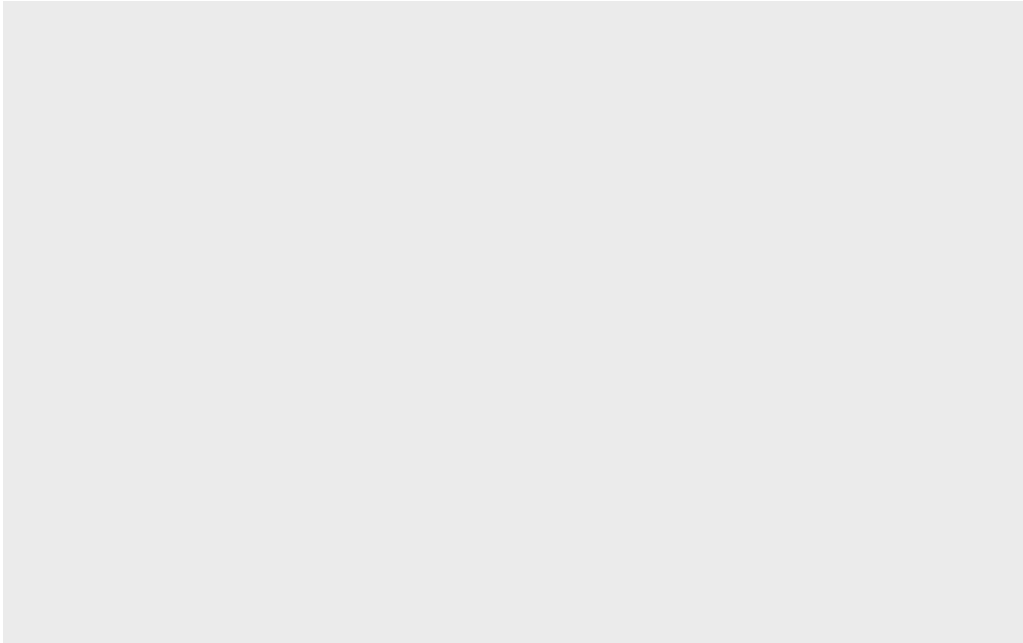
Ava Limtiaco (PID:A18007672)

Table of contents

Our first ggplot	3
we need data + ae + geoms	3
Add a line geom with geom_line()	5
Add a trend line close to the data	6
Gene expression plot	8
Custom color plot	9
Using different geoms	9

#Week 4 Data visualization lab #install.packages("ggplot2") #Any time I want to use this
#package I need to load it

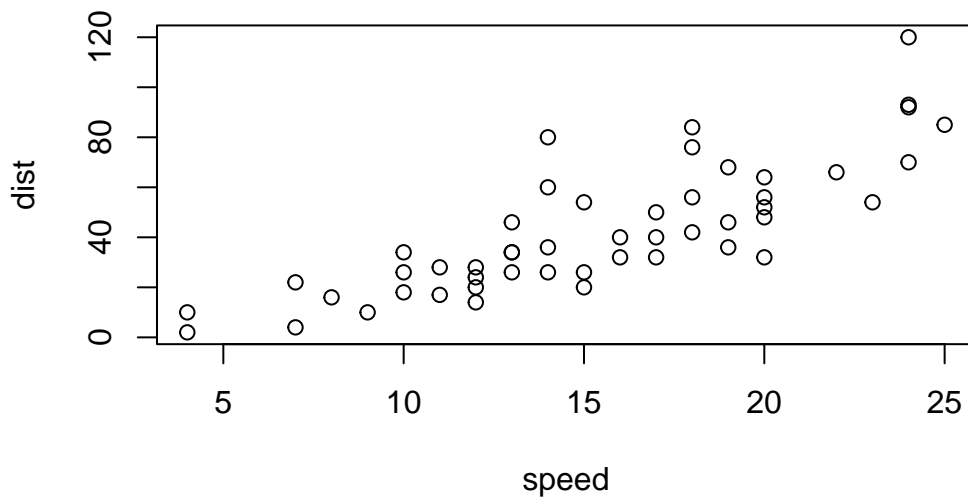
```
library(ggplot2)  
ggplot(cars)
```



```
View(cars)
```

#A quick base R plot - this is not a ggplot

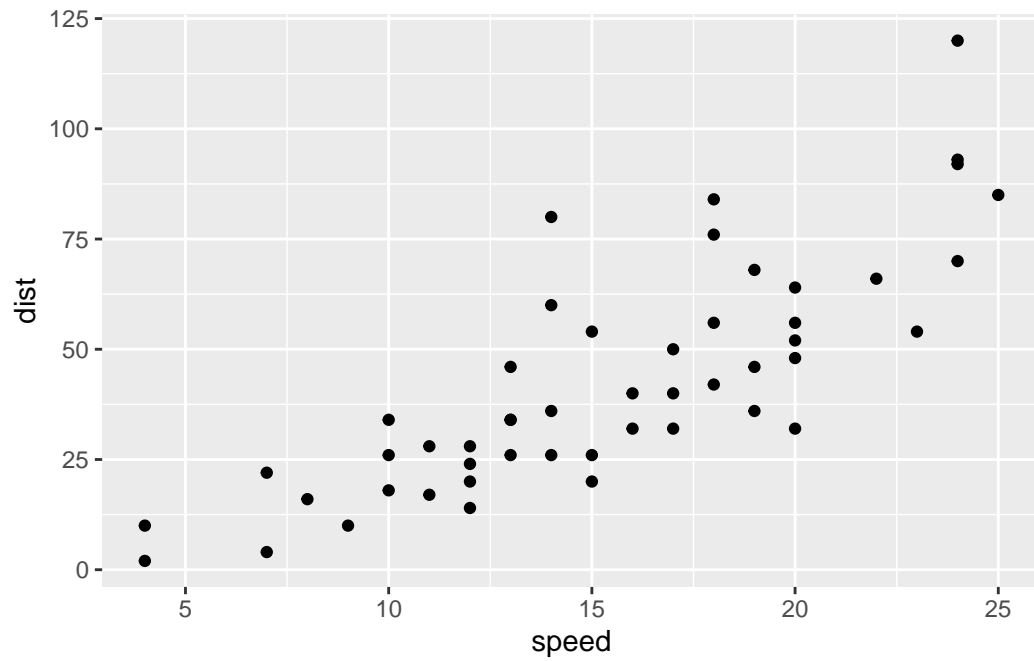
```
plot(cars)
```



Our first ggplot

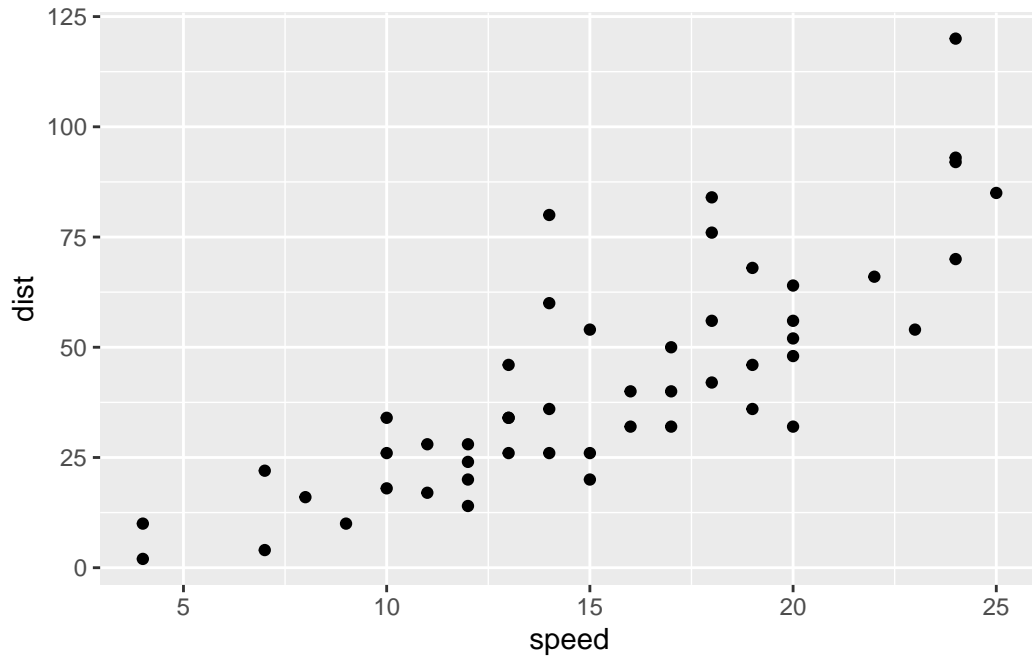
we need data + ae + geoms

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



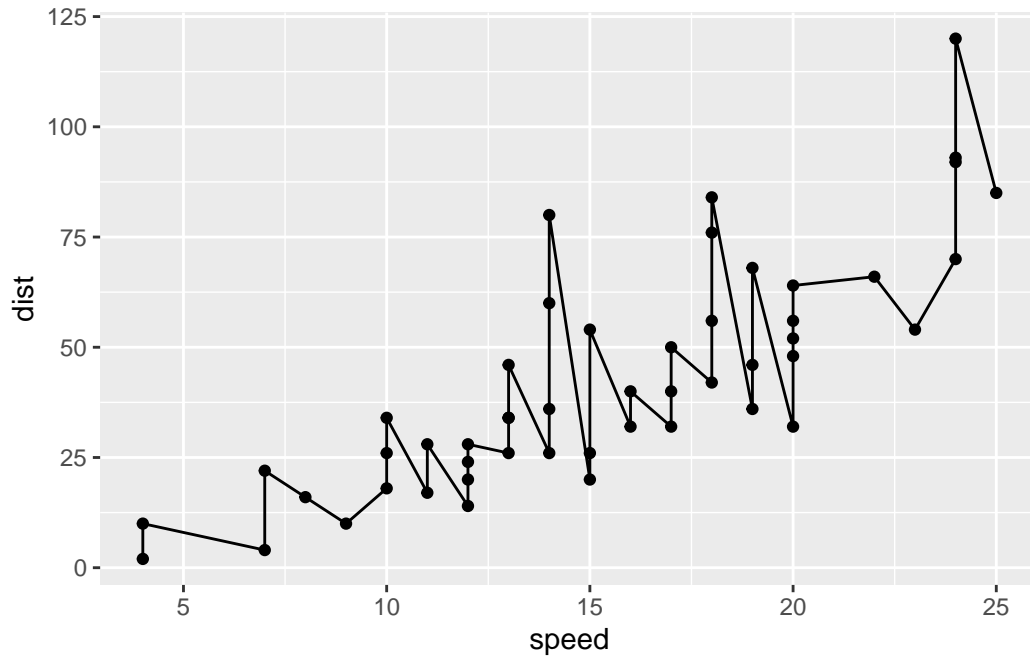
```
p <- ggplot(data=cars) +  
  aes (x=speed, y=dist)+  
  geom_point()
```

p



Add a line geom with `geom_line()`

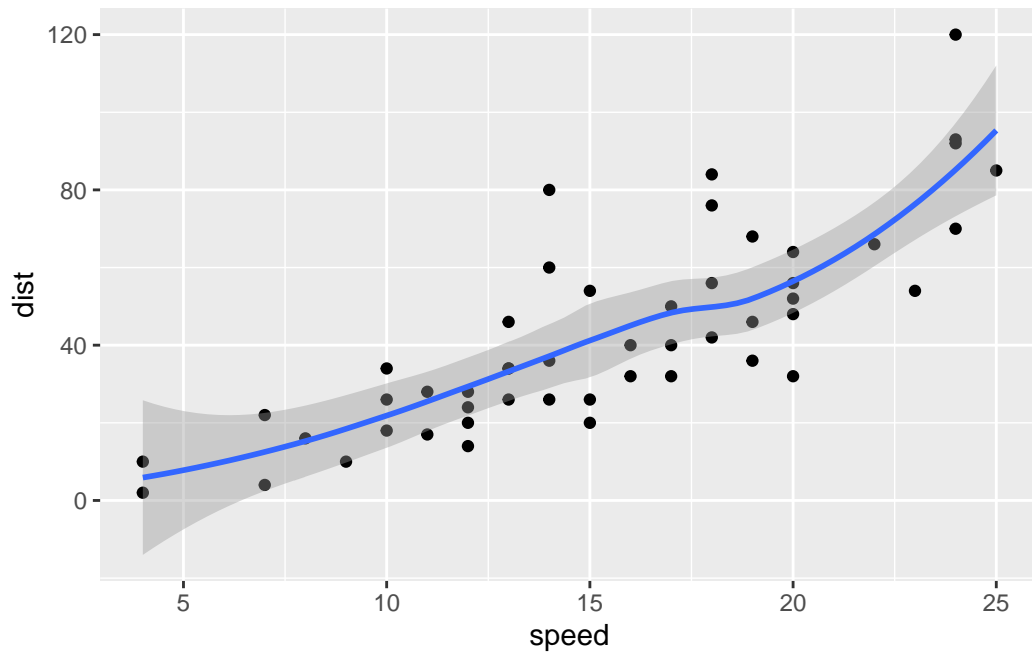
```
p + geom_line()
```



Add a trend line close to the data

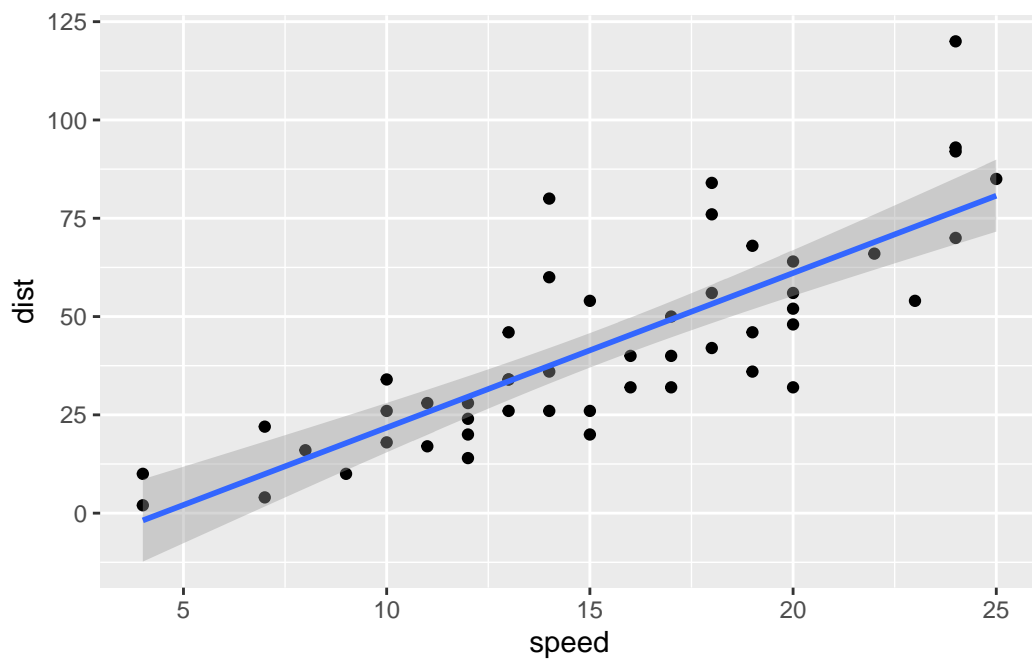
```
p + geom_smooth()
```

```
`geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



```
p + geom_smooth(method="lm")
```

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



Gene expression plot

Read input data into R

```
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 wee dataset?

```
nrow(genes)
```

```
[1] 5196
```

Q. How many columns are there?

```
ncol(genes)
```

```
[1] 4
```

Q. What are the column names?

```
colnames(genes)
```

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

Q. How many “up” and “down” regulated genes are there?

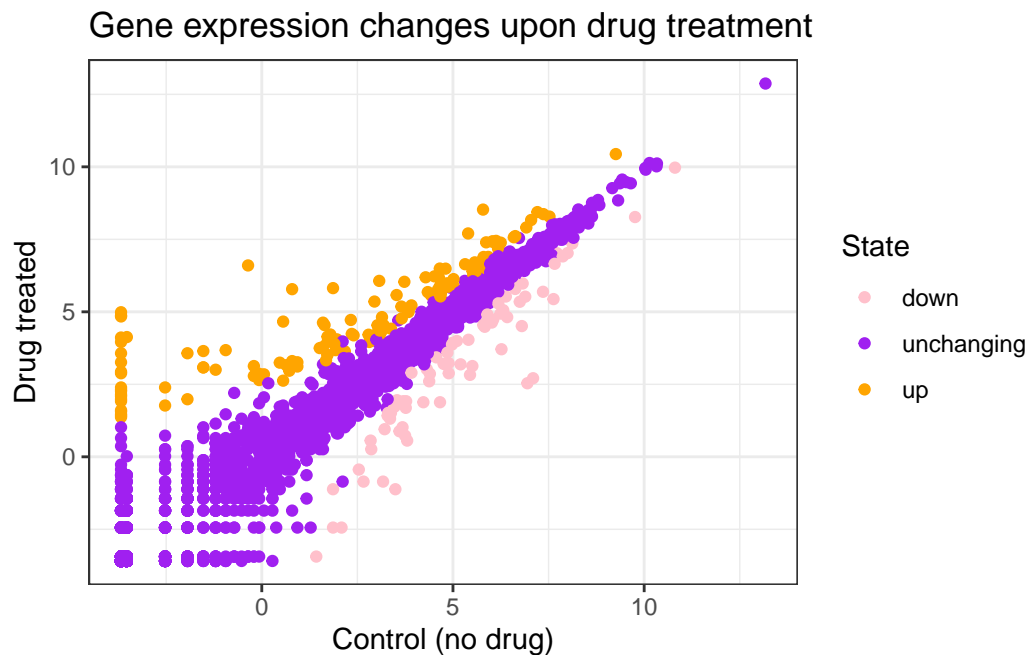
```
table(genes$State)
```

down	unchanging	up
72	4997	127

Custom color plot

Q. Make a first plot of this data

```
ggplot(genes) +  
  aes(x=Condition1, y=Condition2, col=State) + scale_color_manual( values=c("pink", "purple"  
  geom_point() +  
  labs(title="Gene expression changes upon drug treatment", x="Control (no drug)", y="Drug t  
  theme_bw()
```



Using different geoms

Let's plot some aspects of the in-built 'mtcars' dataset.

```
head(mtcars)
```

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1

Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

Q. Scatter plot of 'mpg' vs 'disp'

```
p1 <- ggplot(mtcars) +
  aes(x=mpg, y=disp) +
  geom_point()
```

Q. Boxplot of 'gear' vs 'disp'

```
p2 <- ggplot(mtcars) +
  aes(gear, disp, group=gear) +
  geom_boxplot()
```

Q. Barplot of 'carb'

```
p3 <- ggplot(mtcars) +
  aes(carb) +
  geom_bar()
```

Q. Smooth of 'disp' vs 'qsec'

```
p4 <-ggplot(mtcars) +
  aes(disp, qsec) +
  geom_smooth()
```

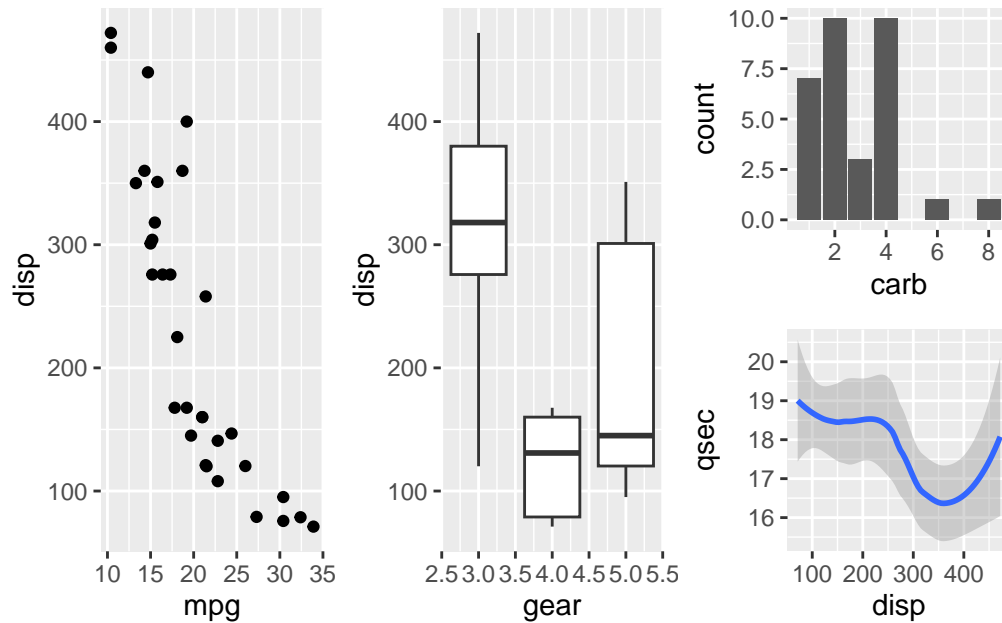
I want to combine all these plots into one figure with multiple pannels.

We can use the “patchwork” pacakge to do this.

```
library(patchwork)
```

```
( p1 | p2 | p3 / p4)
```

`geom_smooth()` using method = 'loess' and formula = 'y ~ x'



```
ggsave(filename="myplot.png",width=10, height=10)
```

`geom_smooth()` using method = 'loess' and formula = 'y ~ x'

```
# File location online
url <- "https://raw.githubusercontent.com/jennybc/gapminder/master/inst/extdata/gapminder.tsv"

gapminder <- read.delim(url)
```

A wee peak

```
head(gapminder)
```

	country	continent	year	lifeExp	pop	gdpPercap
1	Afghanistan	Asia	1952	28.801	8425333	779.4453
2	Afghanistan	Asia	1957	30.332	9240934	820.8530
3	Afghanistan	Asia	1962	31.997	10267083	853.1007
4	Afghanistan	Asia	1967	34.020	11537966	836.1971
5	Afghanistan	Asia	1972	36.088	13079460	739.9811
6	Afghanistan	Asia	1977	38.438	14880372	786.1134

Q. How many countries are in this dataset?

```
length(table(gapminder$country))
```

[1] 142

Q. Plot gdPercap vs lifeExpectancy color by continent

```
ggplot(gapminder) +  
  aes(x=gdPercap, y=lifeExp, color=continent) +  
  geom_point(alpha=0.3) +  
  facet_wrap(~continent) +  
  theme_bw()
```

