

Class05: Data Vis with ggplot

Neha (PID:A17567541)

Graphics systems in R

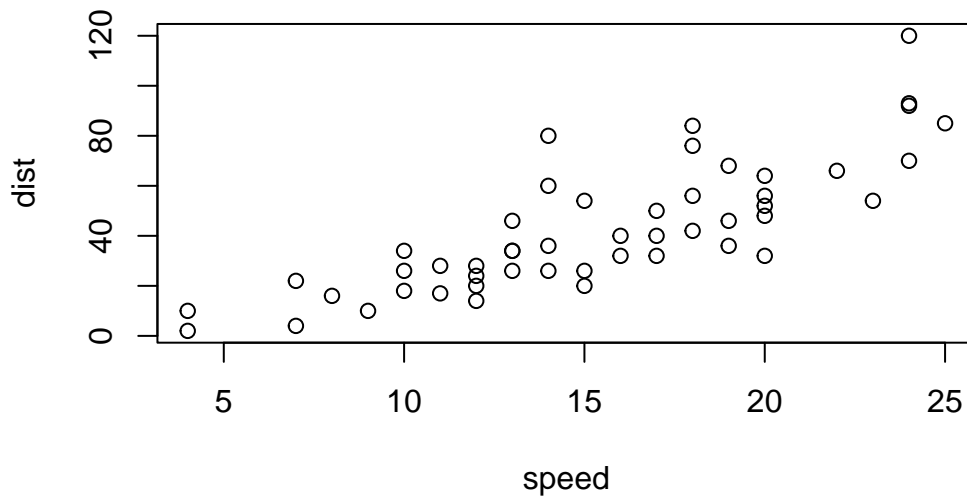
There are many graphic systems in R for making plots and figures

We have already played a little with “**baseR**” graphics and the `plot()` function

Today we will start learning about a popular graphics package called `ggplot2()`

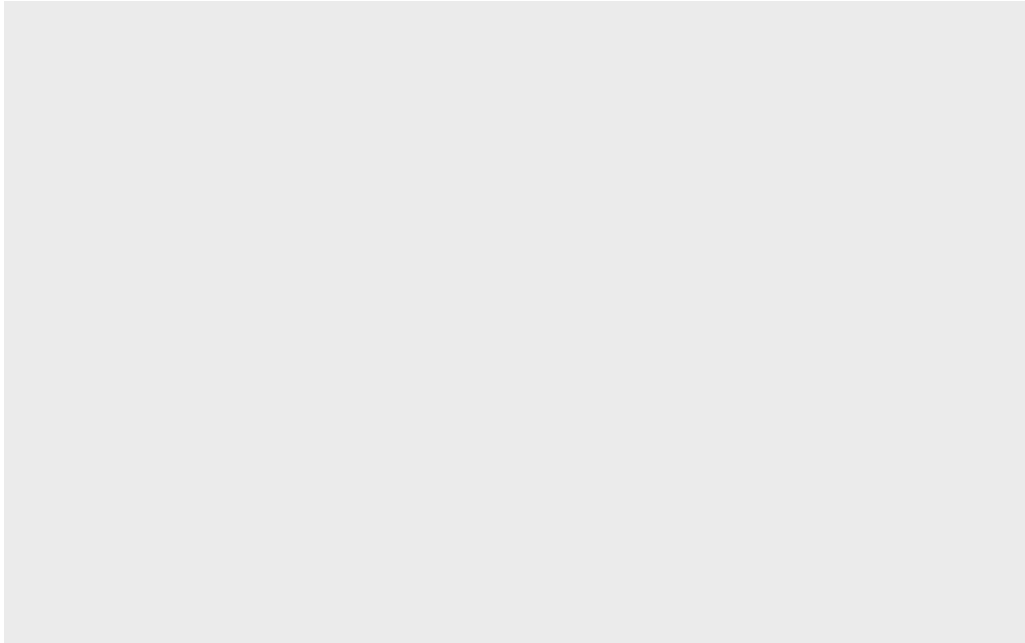
This is an add on package. We will need to install it. I install it like I install any package with the `install.packages()` function.

```
plot(cars)
```



Before I can use the functions from a package, I have to load up the package from my “library”. We use `library(ggplot2)` command to load it up.

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



Every ggplot is made up of at least 3 things: -data (the numbers etc. that will go into your plot) -aes (how the columns of data map to the plot aesthetics) -geoms (how the plot actually looks, points, bars, lines, etc.)

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



For simple plots ggplot is more verbose- it takes more code than base R plot.

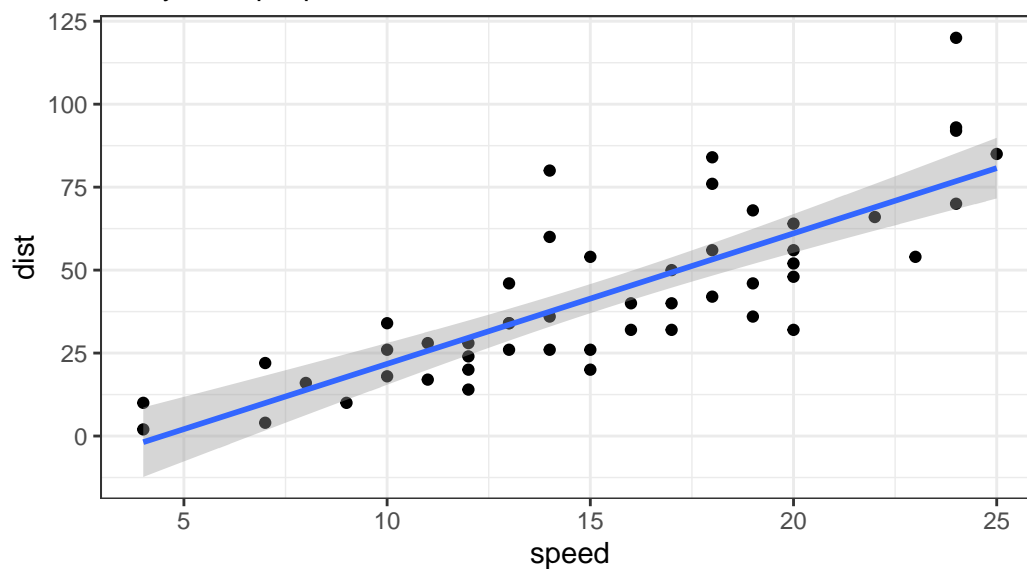
Add some more layers to our ggplot. The + sign needs to be at the end of each line

```
ggplot(cars) +  
  aes (x=speed, y=dist)+  
  geom_point() +  
  geom_smooth(method="lm")+  
  labs(title="Stopping distance of old cars", subtitle= "A silly example plot")+  
  theme_bw()
```

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

Stopping distance of old cars

A silly example 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)
```

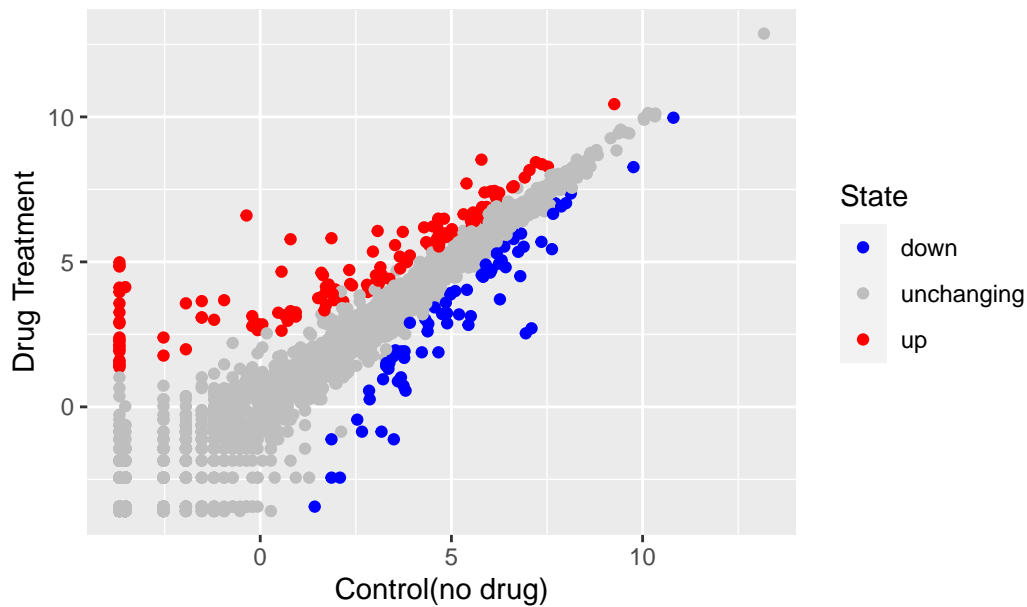
down	unchanging	up
72	4997	127

```
round(table(genes$State)/nrow(genes)*100,2)
```

down	unchanging	up
1.39	96.17	2.44

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

Gene Expression Changes Upon Drug Treatment



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

gapminder <- read.delim(url)

library(dplyr)
```

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

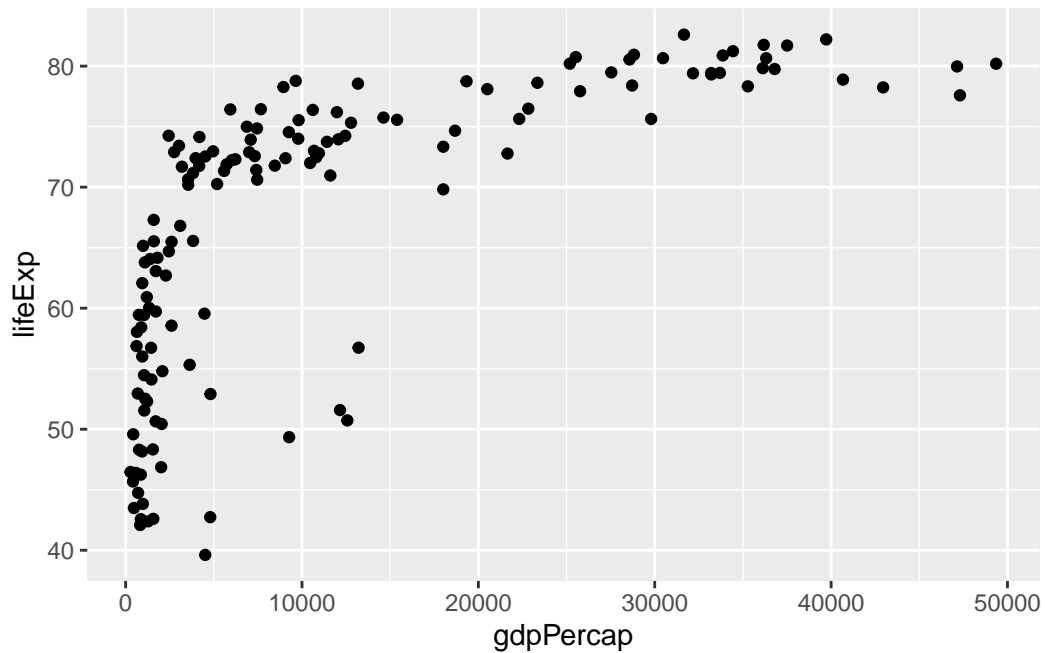
filter, lag

The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

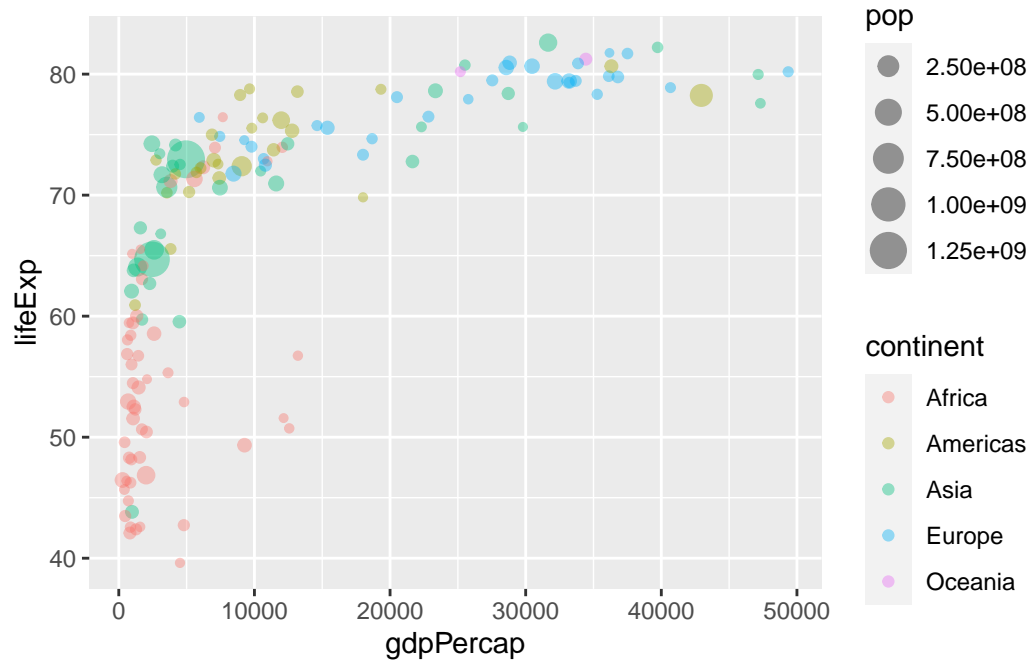
```
gapminder_2007 <- gapminder %>% filter(year==2007)
```

```
ggplot(gapminder_2007)+
  aes(x=gdpPercap,y=lifeExp)+
  geom_point()
```

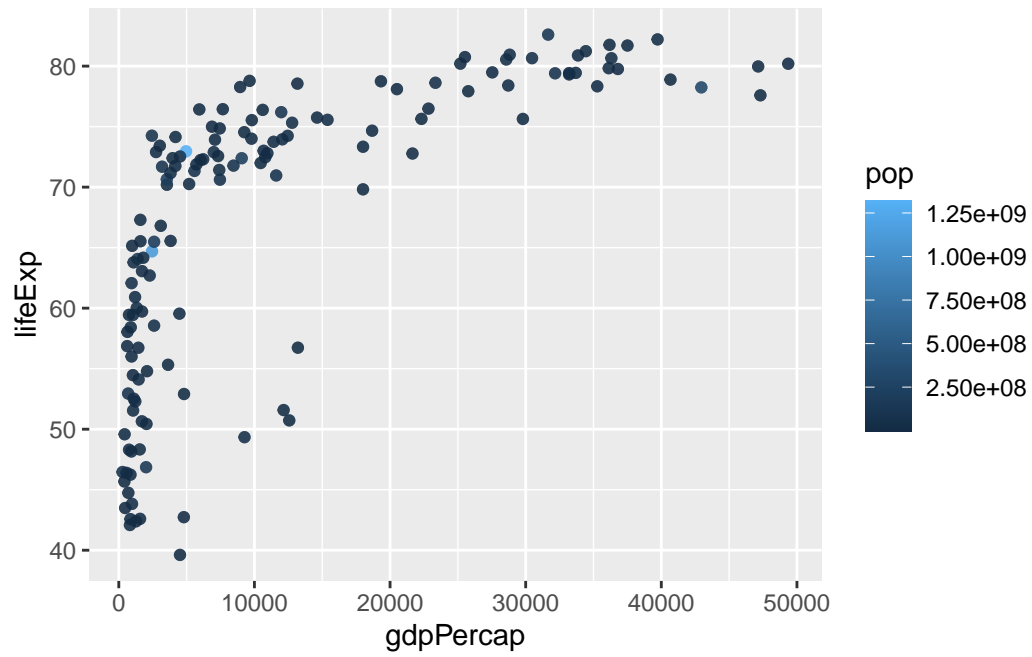


```
p<-ggplot(gapminder_2007)+
  aes(x=gdpPercap,y=lifeExp)+
  geom_point(alpha=0.4)
```

```
p +
  aes(color=continent, size=pop)
```



```
p+
  aes(color=pop) +
  geom_point(alpha=0.8)
```

```
p+aes(size=pop)
```

