

# class05: data vis with ggplot

Sneha Patel

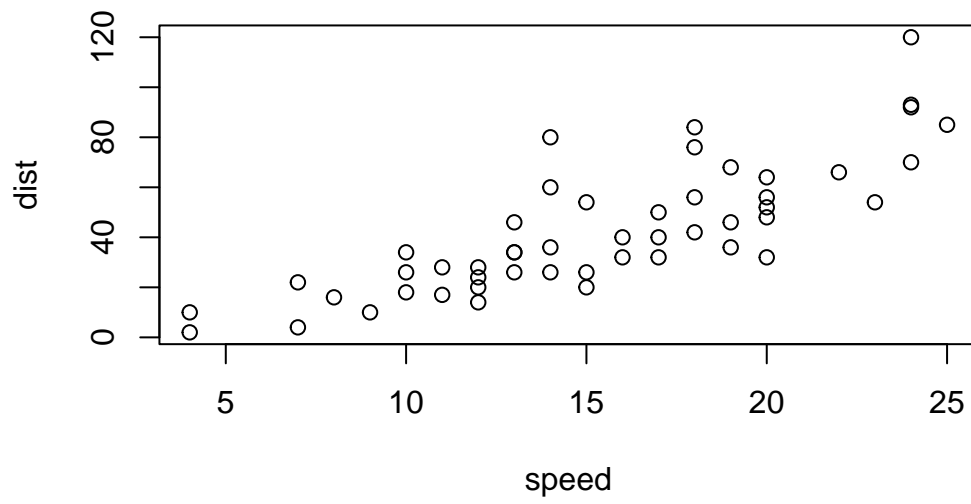
## Table of contents

Our first plot	1
A very cool plot	4

## Our first plot

R base graphics

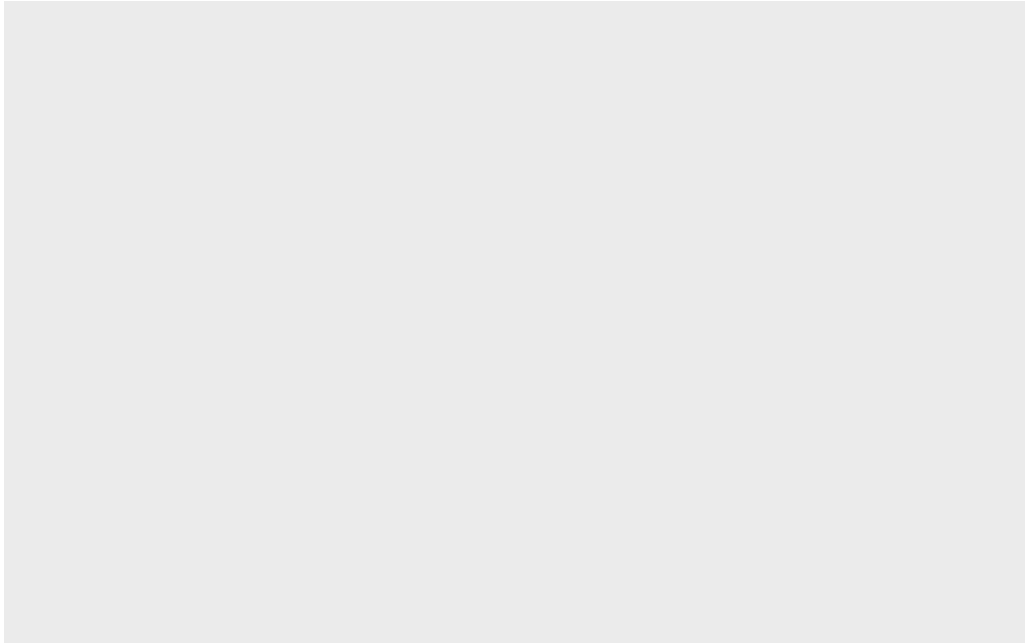
```
plot(cars)
```



must install and load ggplot2 library before you can use it `install.packages("ggplot2")`

```
library(ggplot2)
```

```
ggplot(cars)
```



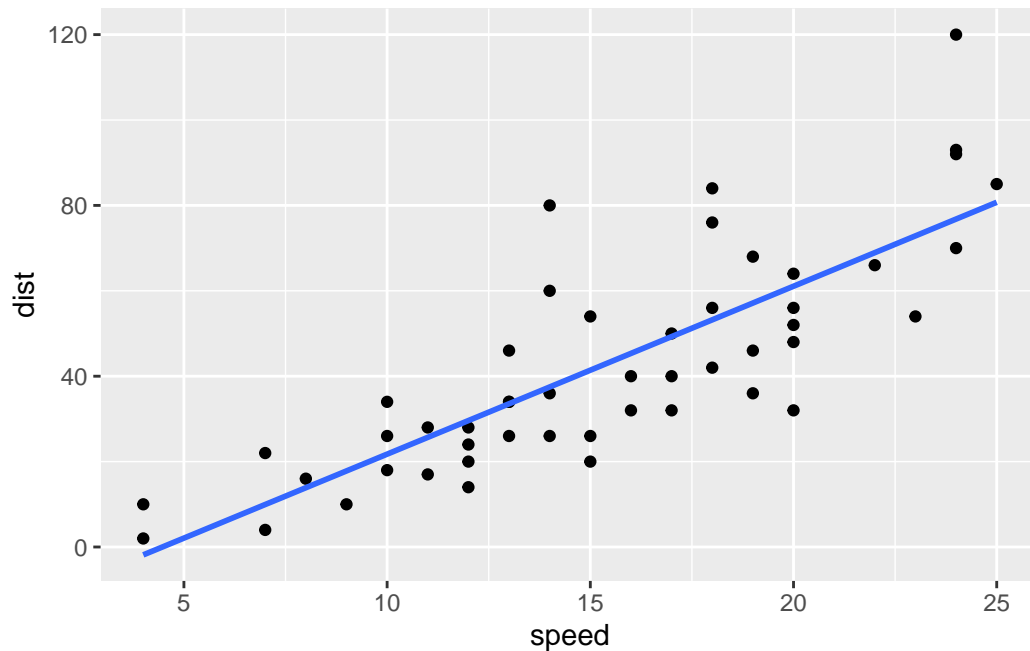
^ wont work

every ggplot needs at least 3 layers

- **data** (ie 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() +  
  geom_smooth(method=lm, se=FALSE)
```

`geom\_smooth()` using formula 'y ~ x'

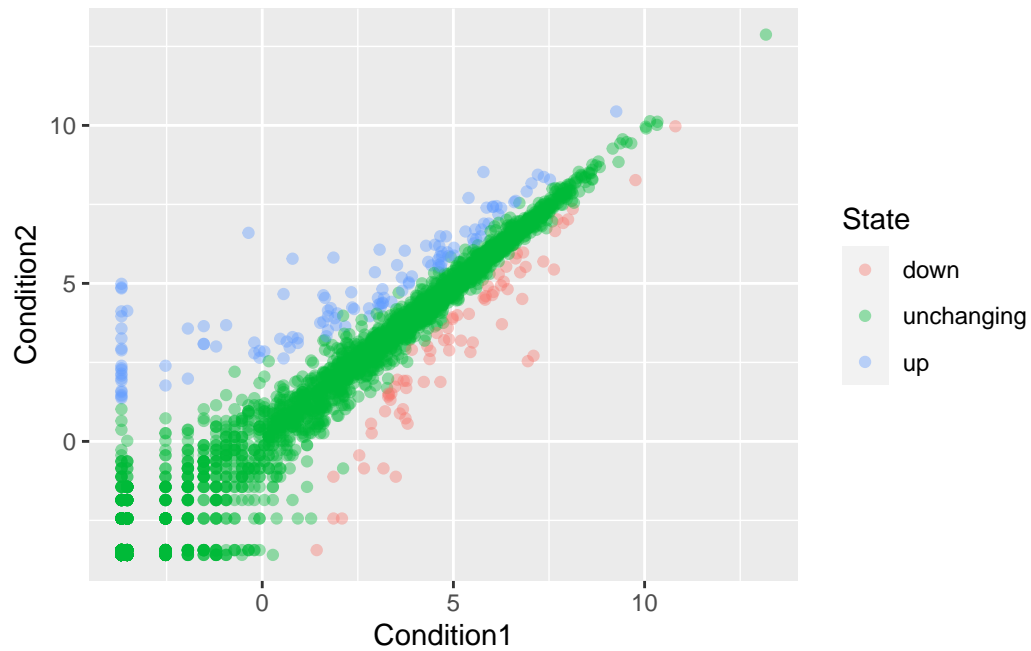


## A very cool plot

```
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

```
p <- ggplot(data=genes) + aes(x=Condition1, y=Condition2, col=State) + geom_point(alpha=0.5)
p
```



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

```
TRUE
127
```

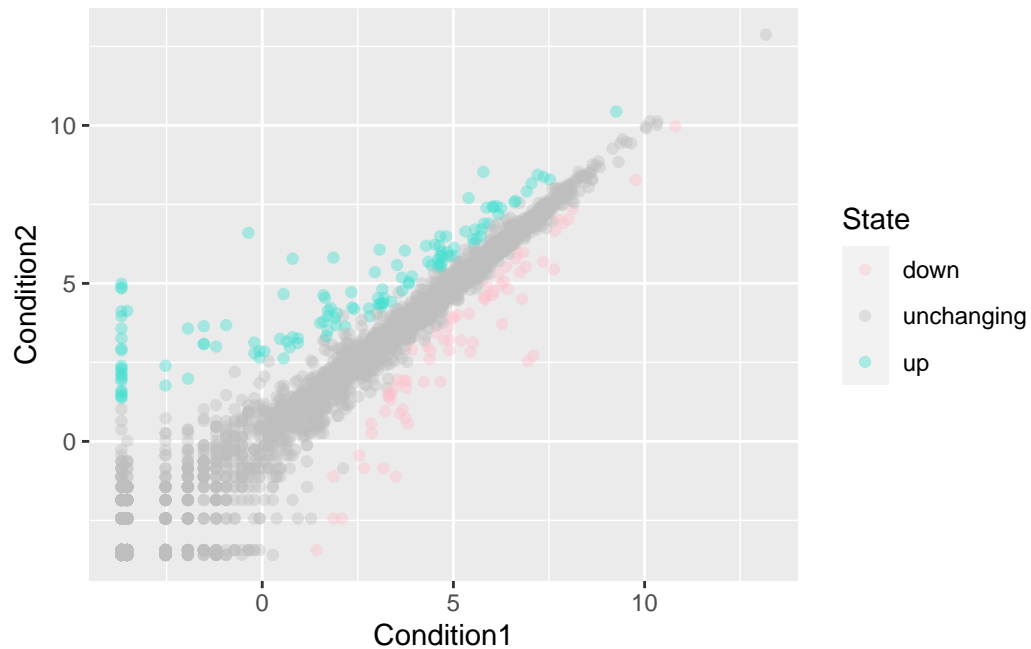
There are 5196 genes in this data set.

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

Here are how many genes are upregulated: 127

Fraction of total genes upregulated: 2.4441878

```
p + scale_colour_manual( values=c("pink","gray","turquoise") )
```



```
p + labs(title="Gene response to Drugs", x = "drug", y = "control")
```

```
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':

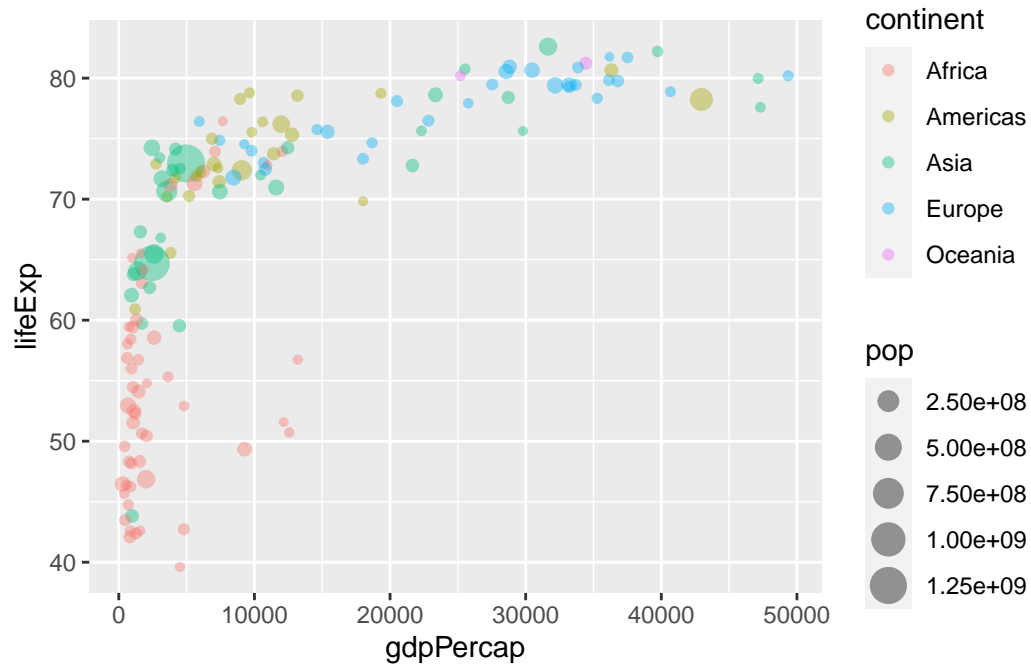
filter, lag

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

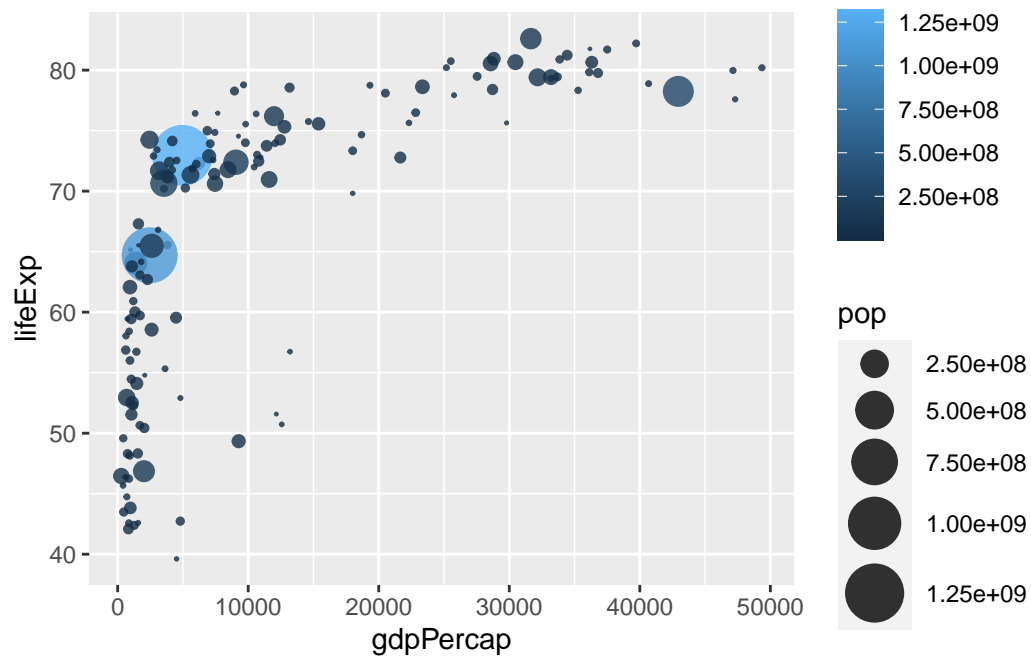
intersect, setdiff, setequal, union

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

```
ggplot(gapminder_2007) +
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +
  geom_point(alpha=0.4)
```



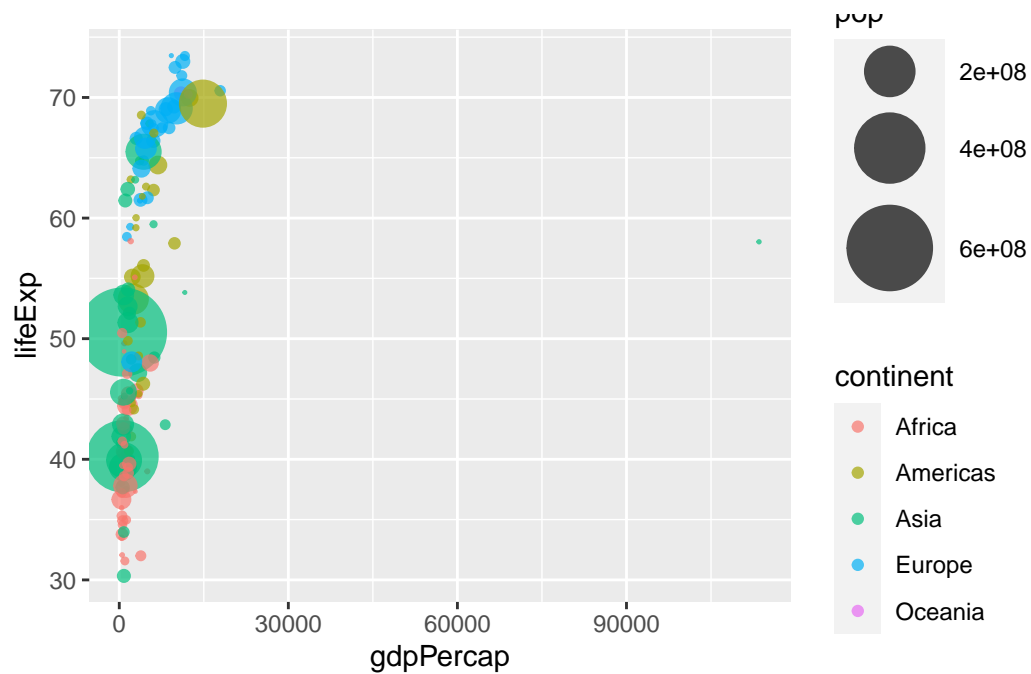
```
ggplot(gapminder_2007) +
  aes(x = gdpPercap, y = lifeExp, color = pop, size = pop) +
  geom_point(alpha=0.8) + scale_size_area(max_size=10)
```



```
gapminder_1957 <- gapminder %>% filter(year==1957)

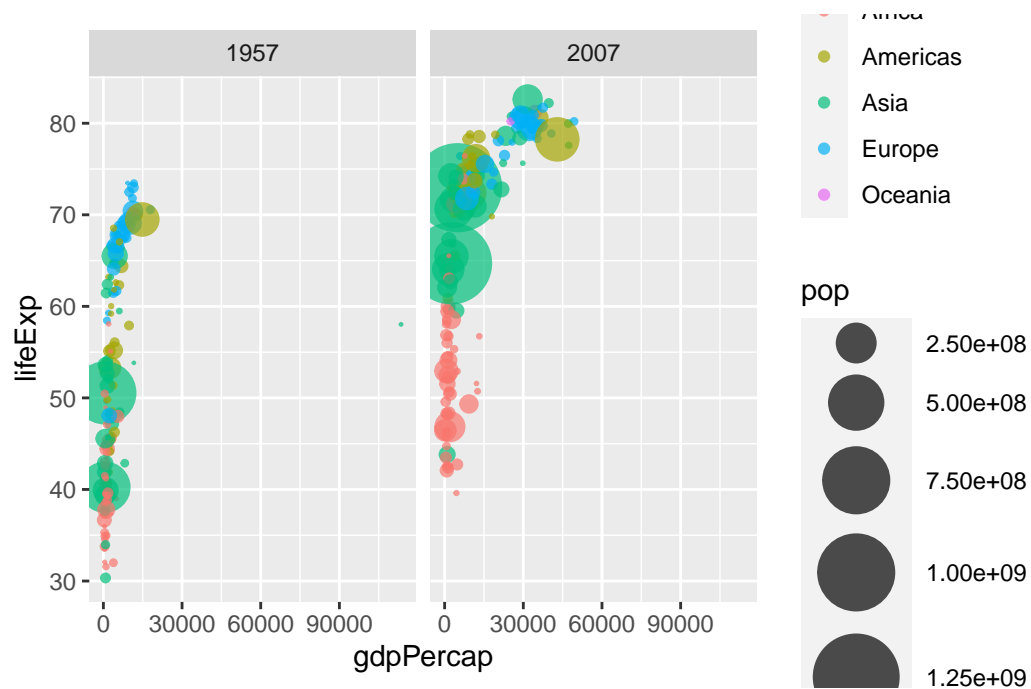
ggplot(gapminder_1957) +
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +
  geom_point(alpha=0.7) + scale_size_area(max_size=15)
```





```
gapminder_1957_2007 <- gapminder %>% filter(year==1957 | year==2007)

ggplot(gapminder_1957_2007) +
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +
  geom_point(alpha=0.7) + scale_size_area(max_size=15) + (facet_wrap(~year))
```



```
gapminder_top5 <- gapminder %>%
  filter(year==2007) %>%
  arrange(desc(pop)) %>%
  top_n(5, pop)

ggplot(gapminder_top5) +
  aes(x = reorder(country,-lifeExp), y = lifeExp, fill=gdpPerCap) + labs(title="Life expect
```

