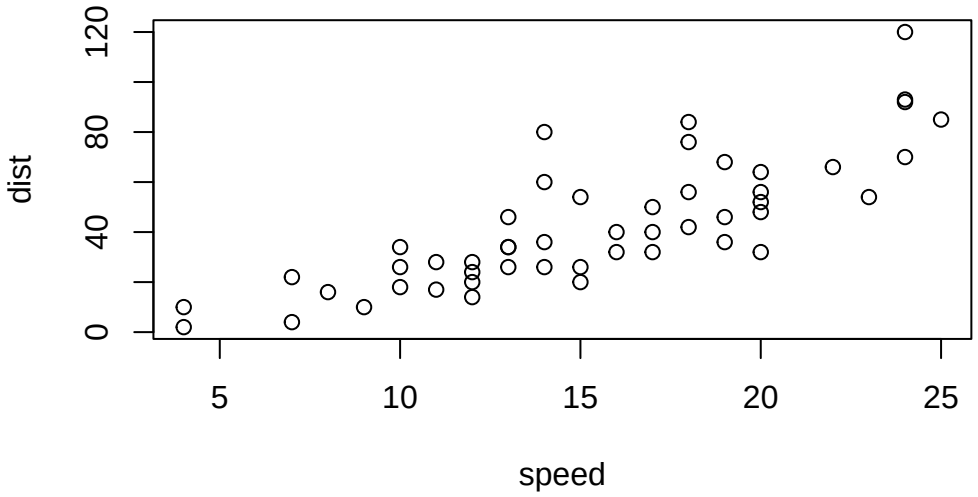
Class 5: Data Viz with ggplot

Iulia

2024-01-24

Graphics systems in R

```
plot(cars)
```



How can we make this is in `ggplot`

To install any packageL use `install.packages()` function

To use it we need to load up the package from the library. `library(ggplot2)`

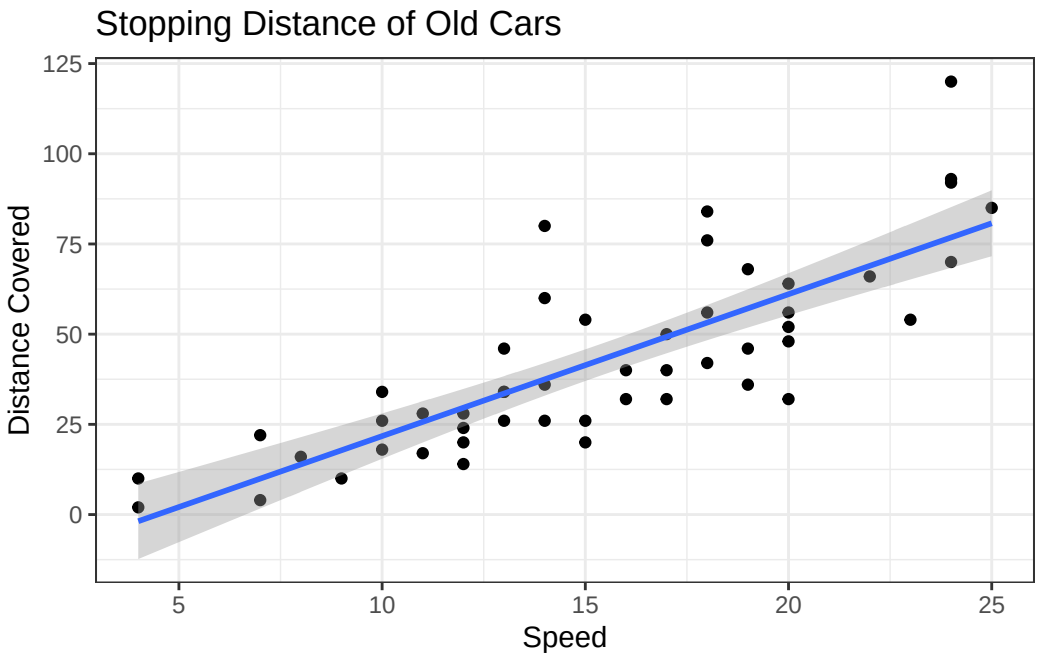
```
library(ggplot2)
ggplot(cars)
```


Every ggplot has at least 3 things -**data** (data.frame) -**aesthetics**, aes, -**geoms** (type of plot, line, points)

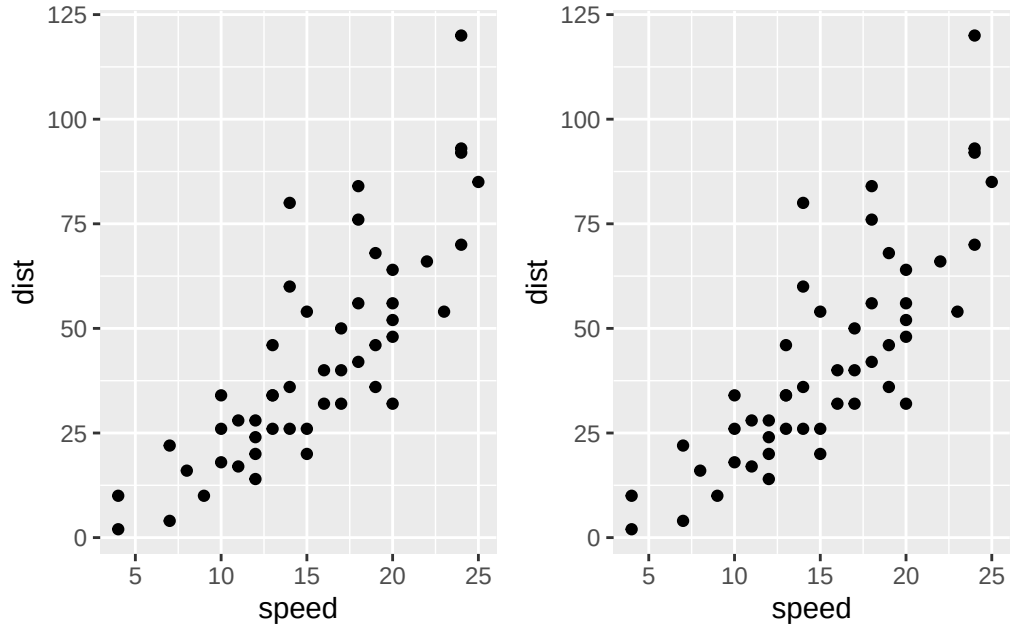
```
pl <- ggplot(cars) +
  aes(speed, dist) +
  geom_point()
```

```
ggplot(cars) +
  aes(speed, dist) +
  geom_point() +
  geom_smooth(method= "lm") +
  labs(title= "Stopping Distance of Old Cars", x = "Speed", y= "Distance Covered") +
  theme_bw()
```

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




```
library(patchwork)
(pl|pl)
```



url <- “https://bioboot.github.io/bimm143_S20/class-material/up_down_expression.txt” genes <- read.delim(url) head(genes)

```
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
```

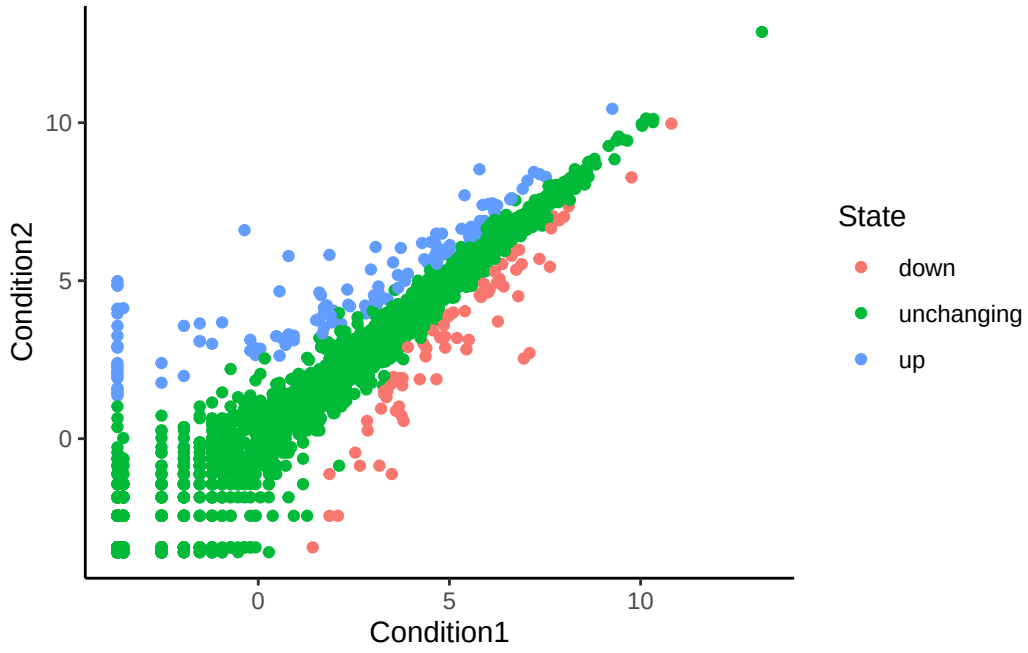
```
ncol(genes)
```

```
[1] 4
```

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

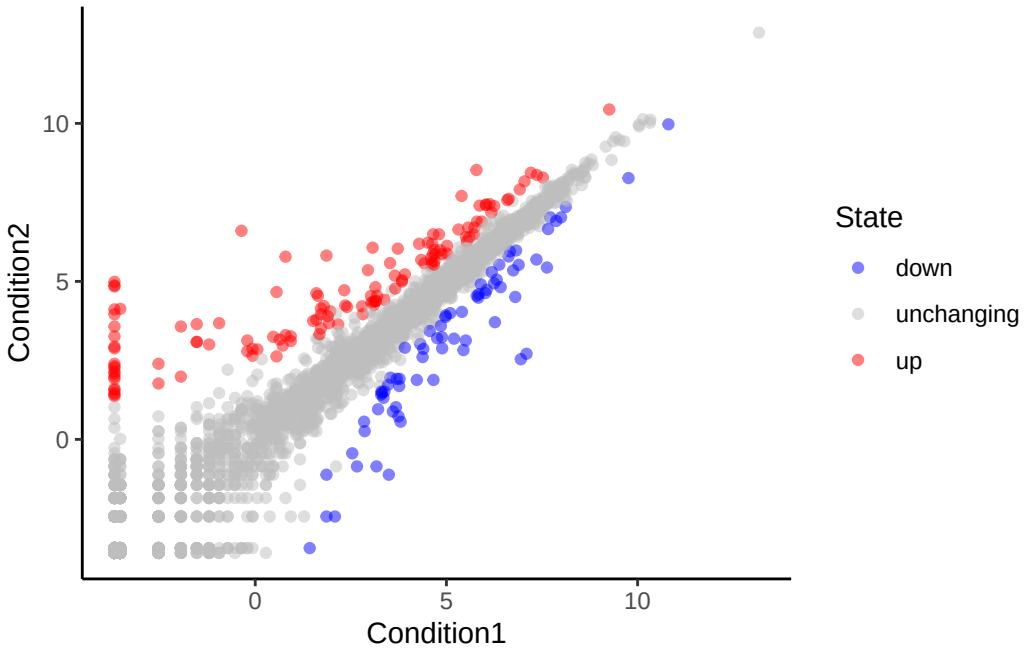
```
    down unchanging         up
    1.39      96.17       2.44
```

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

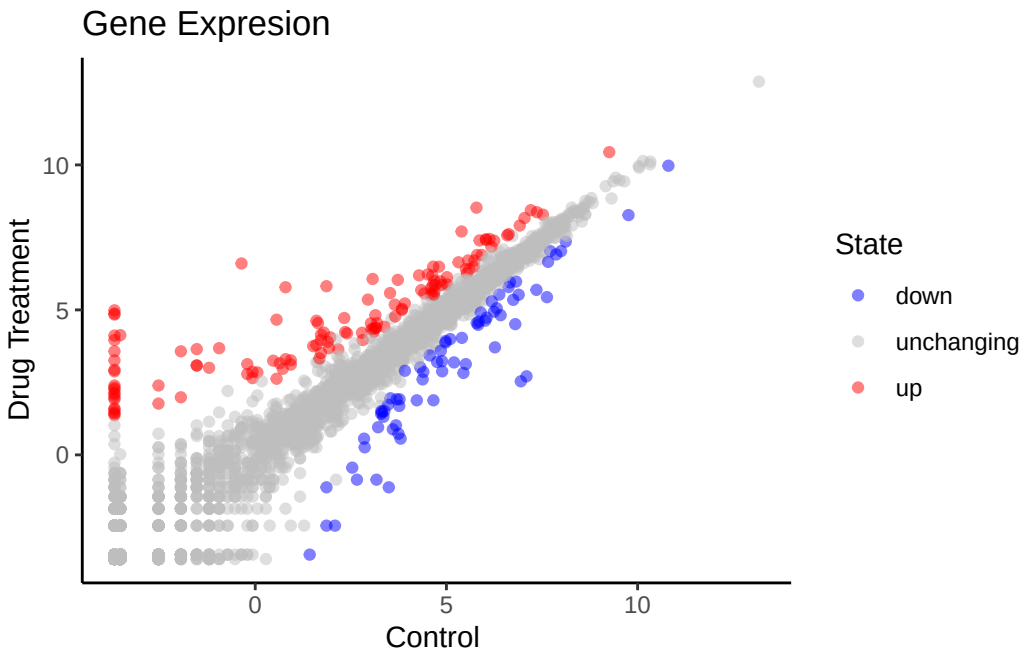



```
p <- ggplot(genes) +
  aes(x = Condition1, y = Condition2, col = State, name = Gene) +
  geom_point(alpha = 0.5) +
  theme_classic()
```

```
p + scale_colour_manual( values=c("blue","gray","red") )
```




```
p + scale_colour_manual(values=c("blue","gray","red")) +
    labs(title="Gene Expresion",
         x="Control ", y="Drug Treatment")
```




```
library(gapminder)
url1 <- "https://raw.githubusercontent.com/jennybc/gapminder/master/inst/extdata/gapminder
```

```
gapminder <- read.delim(url1)
```

```
#library(plotly)
#ggplotly(p)
```

```
library(dplyr)
```

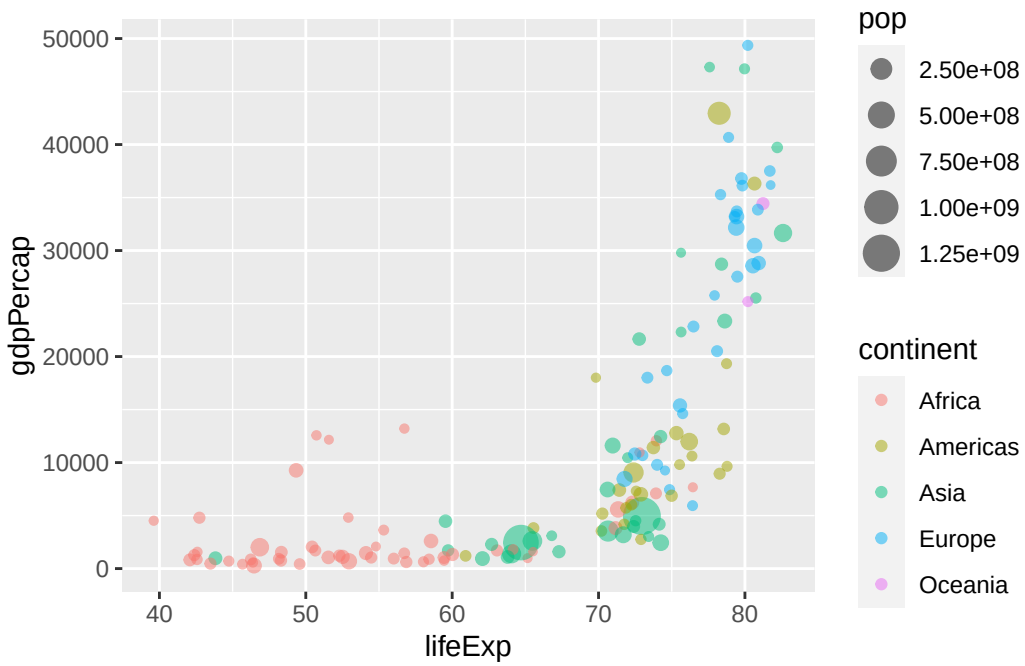
```
Attaching package: 'dplyr'

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)
```

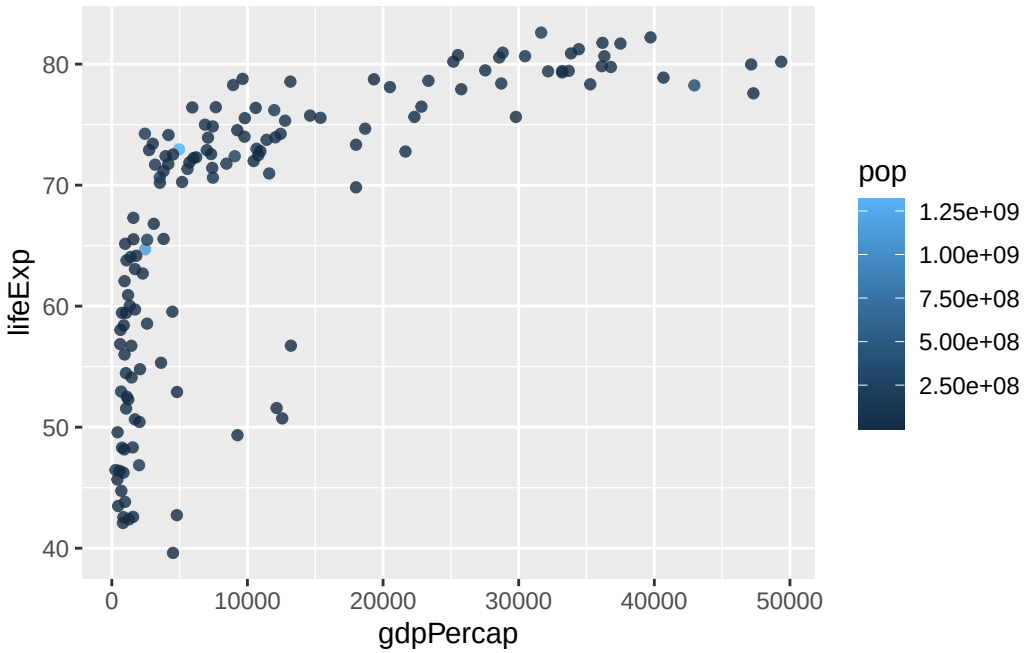
```
head(gapminder_2007)
```

```
      country continent year lifeExp      pop  gdpPercap
1 Afghanistan      Asia 2007  43.828 31889923   974.5803
2     Albania    Europe 2007  76.423  3600523  5937.0295
3     Algeria    Africa 2007  72.301 33333216  6223.3675
4      Angola    Africa 2007  42.731 12420476  4797.2313
5   Argentina  Americas 2007  75.320 40301927 12779.3796
6   Australia   Oceania 2007  81.235 20434176 34435.3674
```

```
ggplot(gapminder_2007) +
  aes(x = lifeExp, y = gdpPercap, color=continent, size=pop ) +
  geom_point(alpha = 0.5)
```




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


lifeExp
80
70
60
50
40
0
10000
20000
30000
40000
50000
gdpPercap
pop
1.25e+09
1.00e+09
7.50e+08
5.00e+08
2.50e+08