Class05:Data Visualization

Jacklyn (PID: A16410001)

Graphics Systems in R

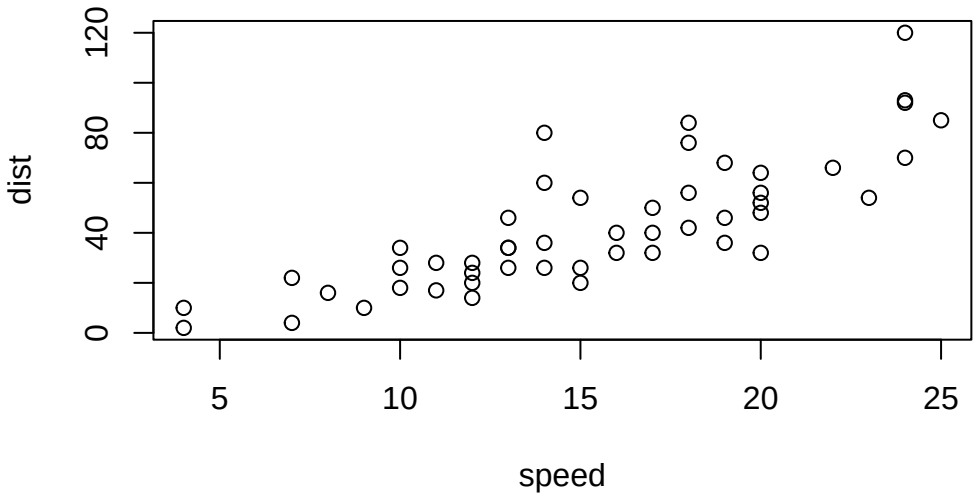
There are many graphics systems in R for making plots and figures.

We have already played a little with **"base R"** graphics and the 'plot()' function.

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

This is an add on package - i.e. we need to install it. I install it (like I install any package) with the 'install.packages()' function.

```
plot(cars)
```

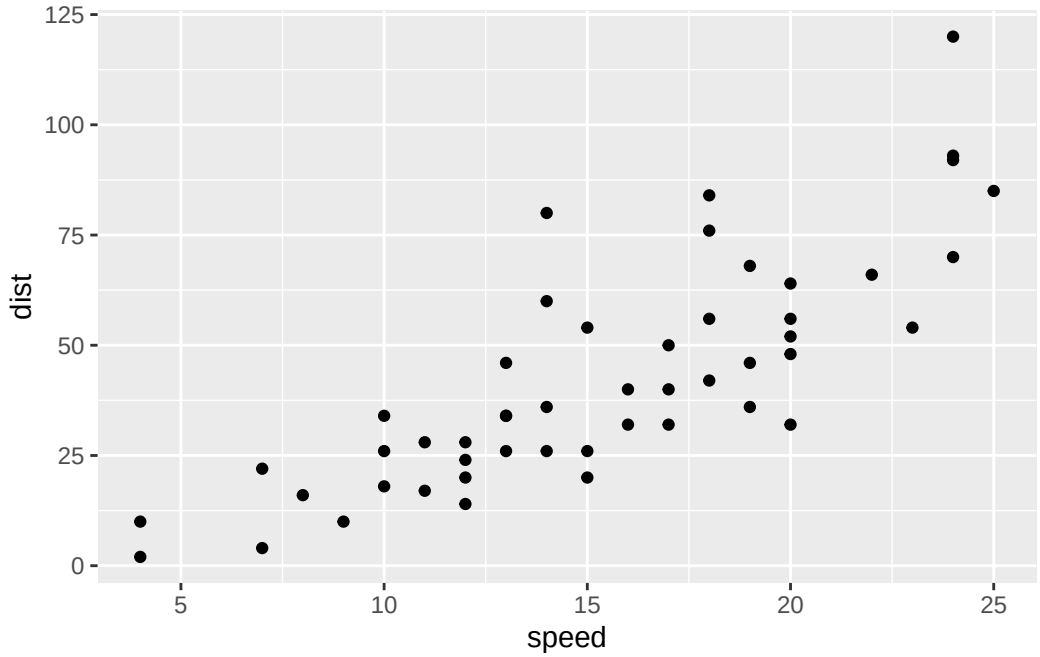


Before I can use the functions from a package I have to load up the package from my "library". We use the 'library(ggplot2)' command to load it up

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


Every ggplot is made up of at least 3 things: - data (the number etc. that will fo 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()
```

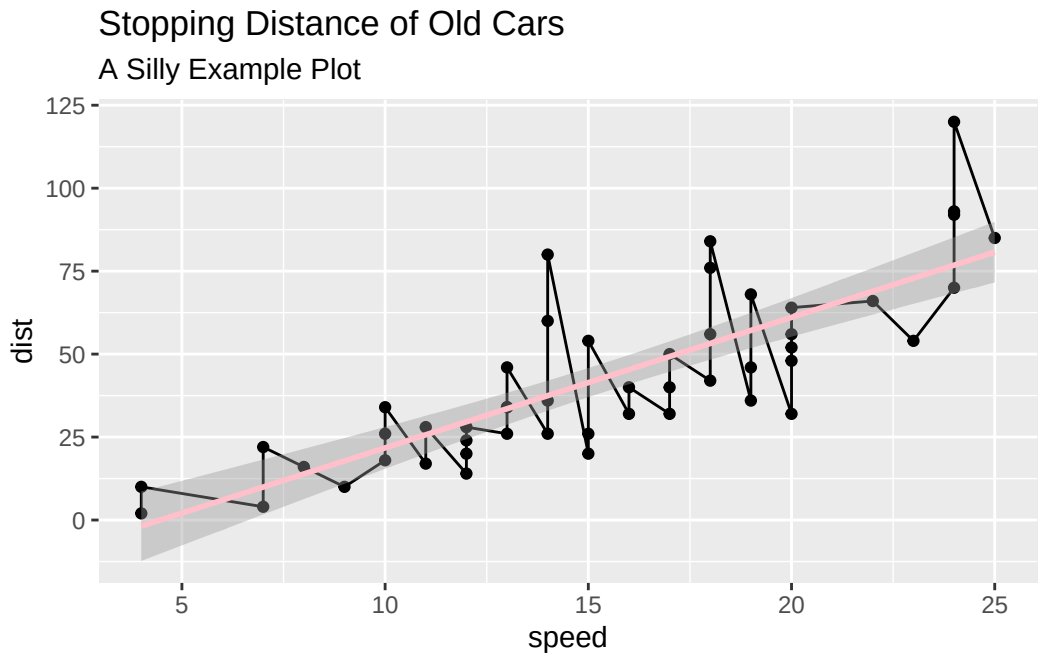


For simple plots ggplot is more verbose - it takes more code - than base R plot.

Add some more layers to you ggplot:

```
ggplot(cars) +
  aes(x=speed, y=dist) +
  geom_point() +
  geom_line() +
  geom_smooth(method="lm" , color="pink") +
  labs(title="Stopping Distance of Old Cars", subtitle = "A Silly Example Plot")
```

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

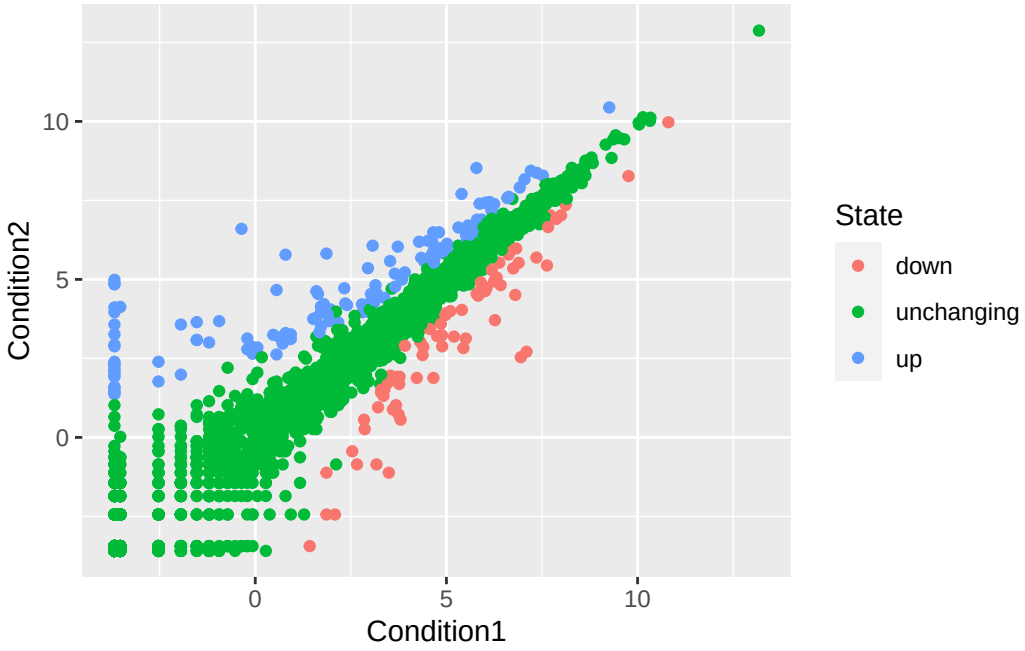


Gene Expression Changes Upon Drug Treatment

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

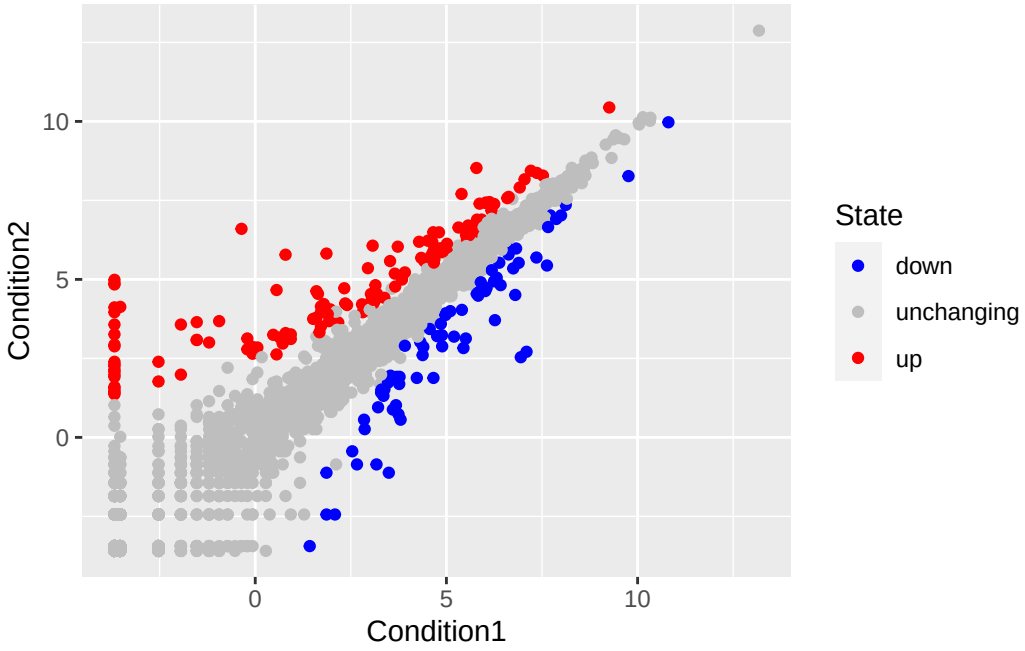
```
        Gene Condition1 Condition2      State
1      A4GNT -3.6808610 -3.4401355 unchanging
2       AAAS  4.5479580  4.3864126 unchanging
3      AASDH  3.7190695  3.4787276 unchanging
4       AATF  5.0784720  5.0151916 unchanging
5       AATK  0.4711421  0.5598642 unchanging
6 AB015752.4 -3.6808610 -3.5921390 unchanging
```

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

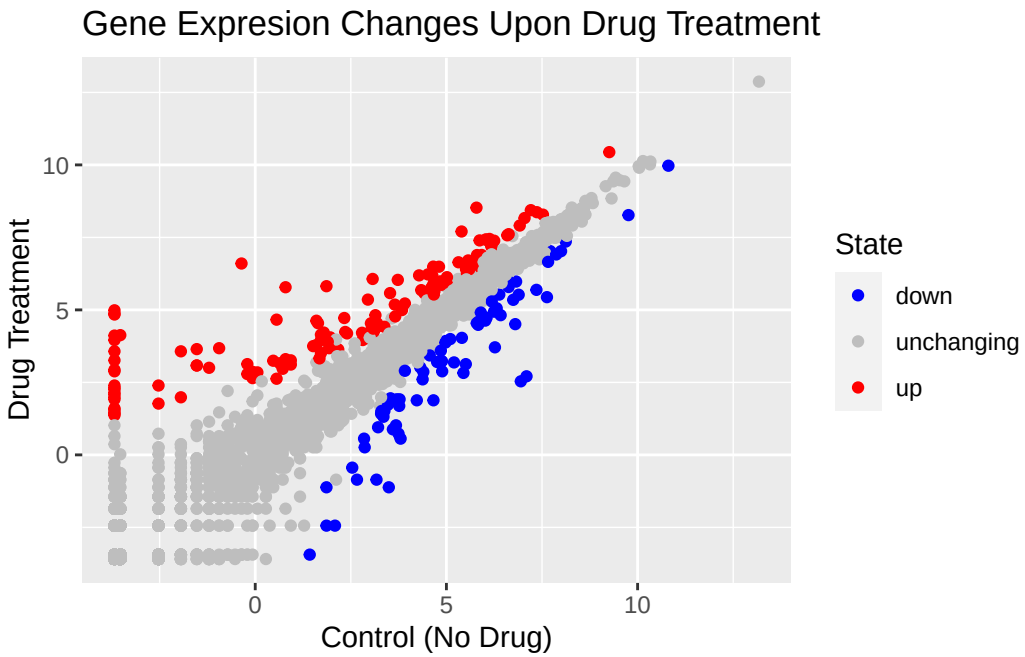



```
p <- ggplot(genes)+
  aes(x=Condition1, y=Condition2, col=State)+
  geom_point()
```

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




```
p + scale_color_manual (values=c("blue", "gray", "red"))+
  labs(title="Gene Expresion Changes Upon Drug Treatment",
       x="Control (No Drug)",
       y="Drug Treatment")
```



Gapminder

gapminder_2007

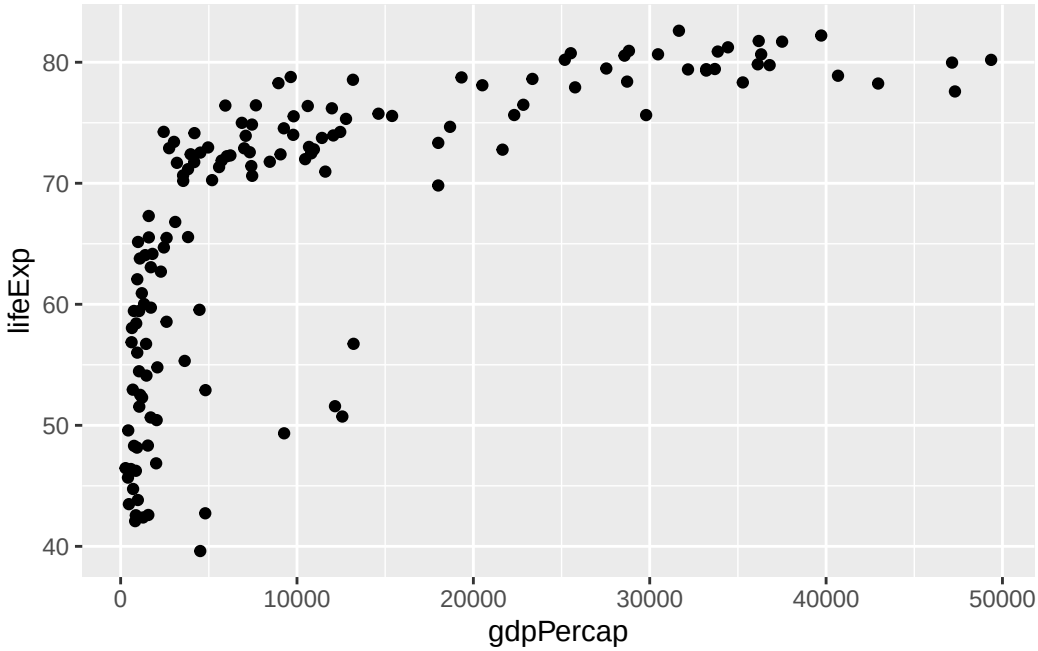
```
library (gapminder)
library(dplyr)
```

```
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)
```

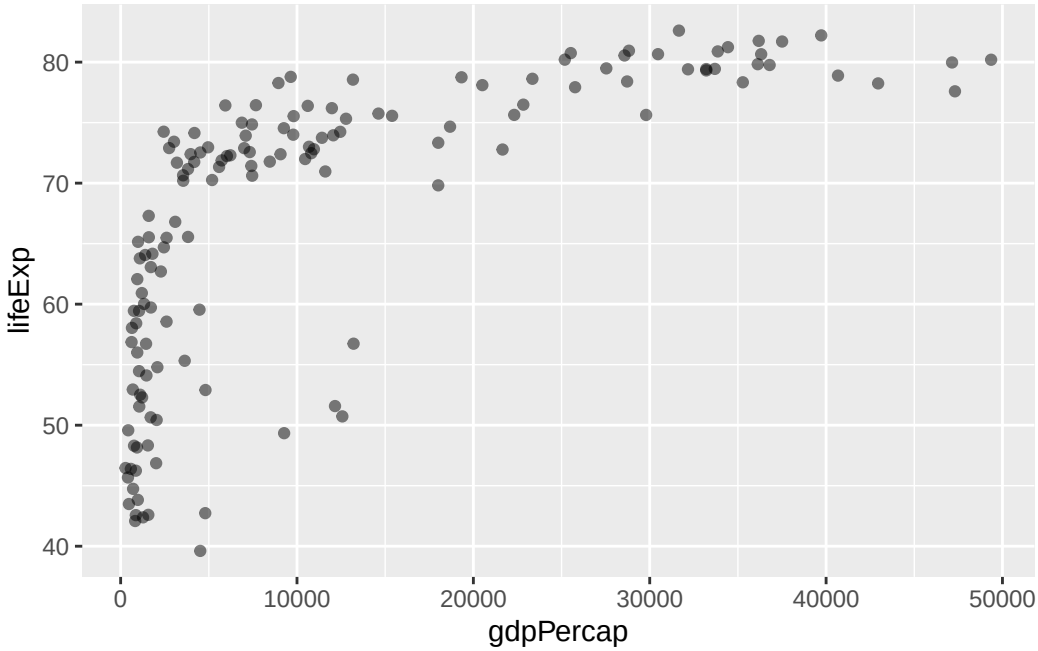
plotting

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



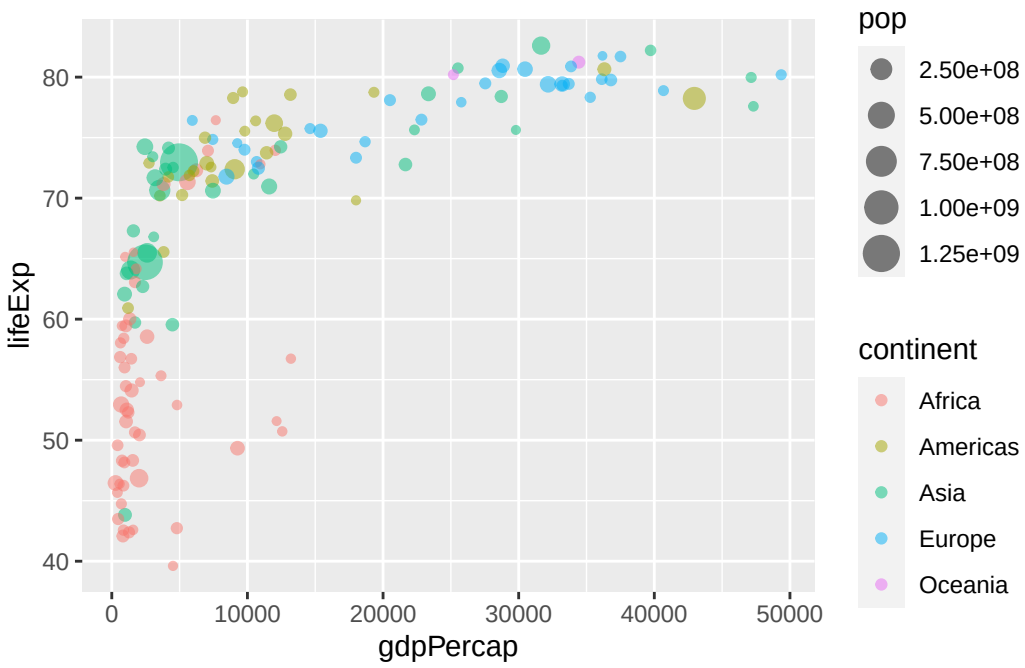
add "alpha=" to 'geom_point()' to make points more transparent

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



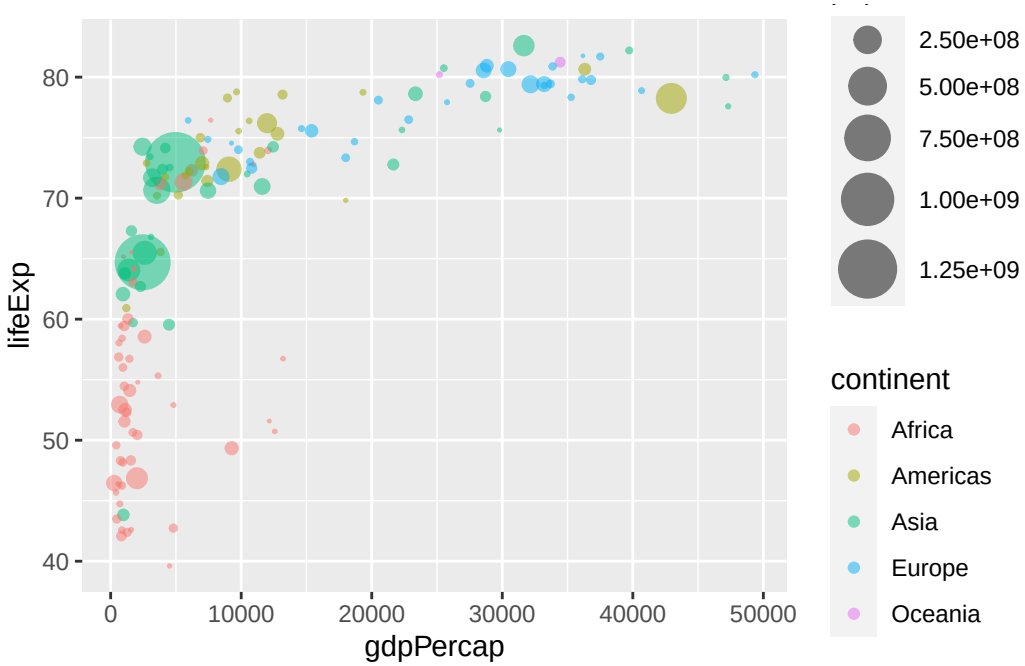
Adding Colors with 'aes()' function

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



Scaling Information

```
ggplot(gapminder_2007)+
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop)+
  geom_point(alpha=0.5)+
  scale_size_area(max_size=10)
```

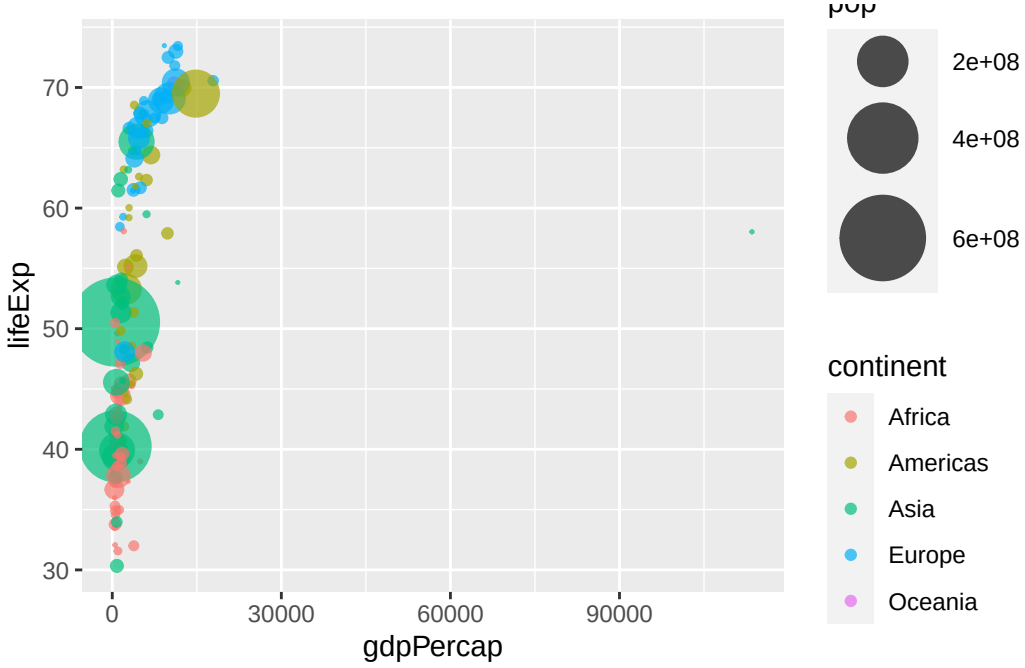


gapminder_1957

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

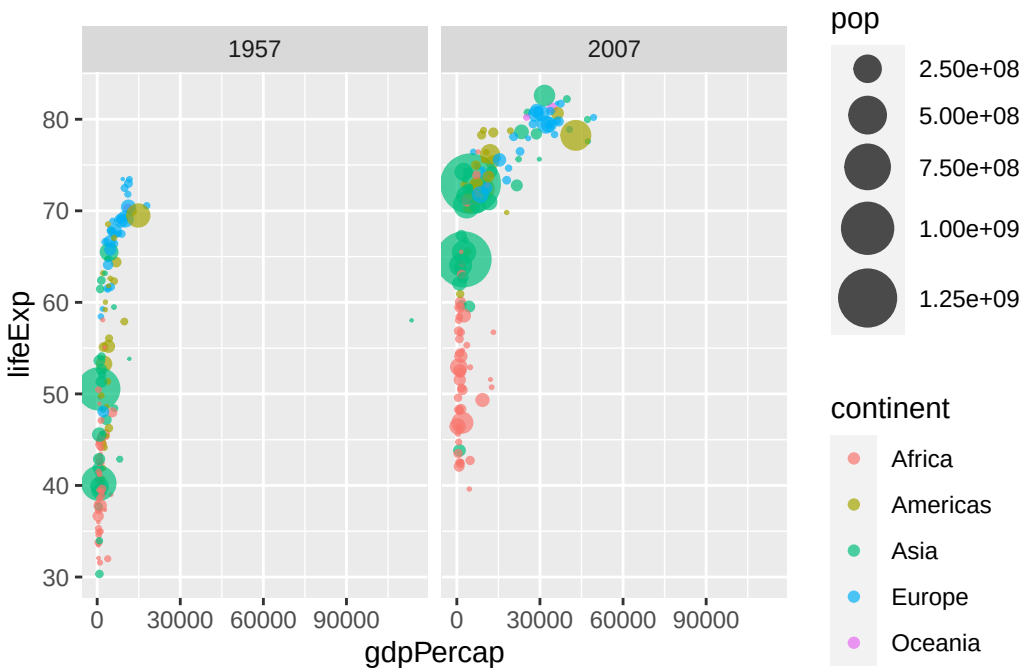
plotting

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



Plotting both 2007 and 1957

```
gapminder_1957 <- gapminder %>% filter(year==1957 | year==2007)

ggplot(gapminder_1957)+
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop)+
  geom_point(alpha=0.7)+
  scale_size_area(max_size=10)+
  facet_wrap(~year)
```


1957
2007
lifeExp
80
70
60
50
40
30
0
30000
60000
90000
gdpPercap
pop
2.50e+08
5.00e+08
7.50e+08
1.00e+09
1.25e+09
continent
Africa
Americas
Asia
Europe
Oceania