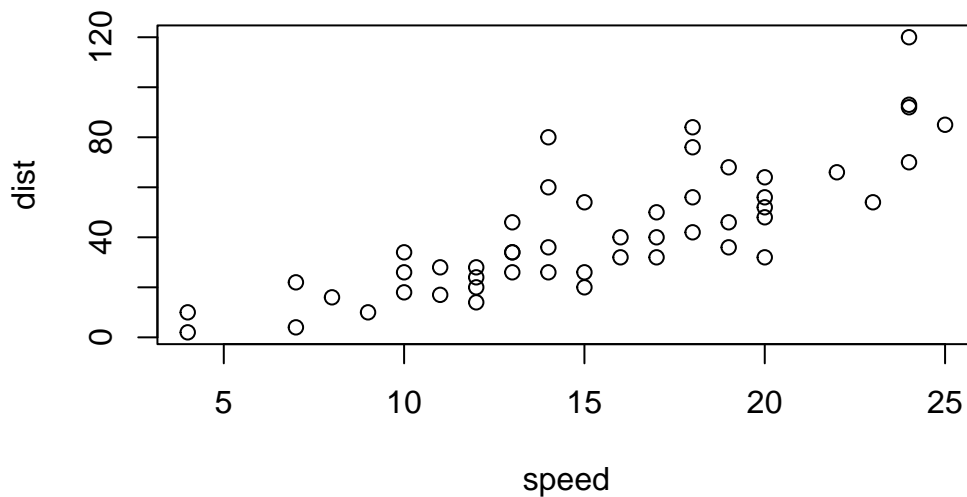


Class 5: Data visualization with ggplot

Sanket Garg (A59026686)

Using GGPLOT

```
plot(cars)
```

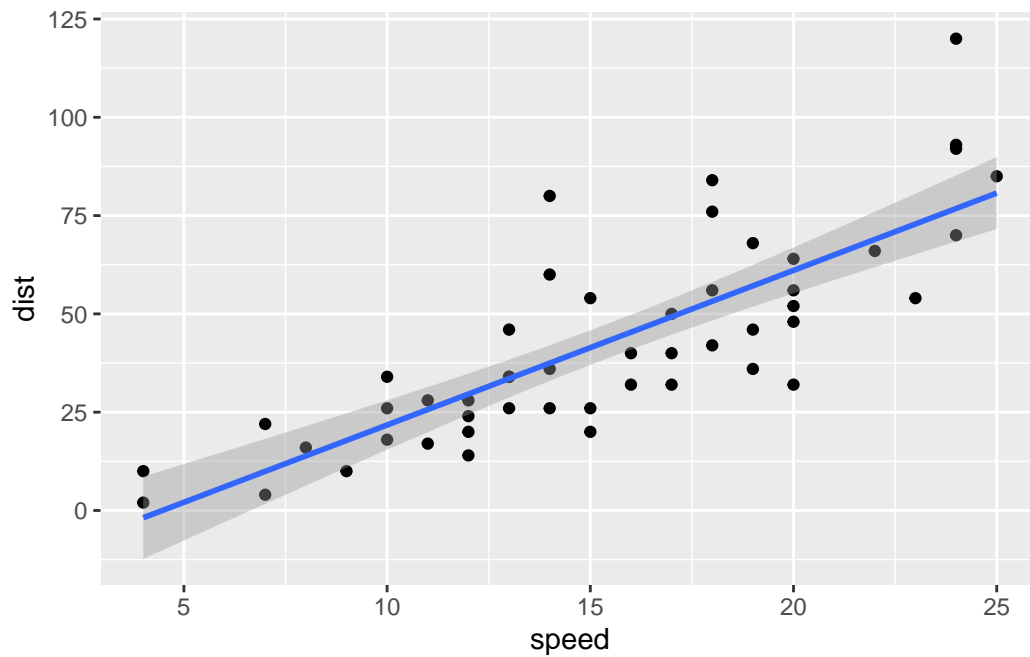


to use ggplot2, we first need to install it on our computers. To do this, we will use the function 'install.packages()'.

to use ggplot, I need to spell out at least 3 things: - data (the stuff I want to plot as a data.frame) - aesthetics (aes() values - how the data map to the plot) - geoms (how I want things drawn)

```
library(ggplot2)
ggplot(cars) +
  aes(x = speed, y = dist) +
  geom_point() +
  geom_smooth(method='lm')
```

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

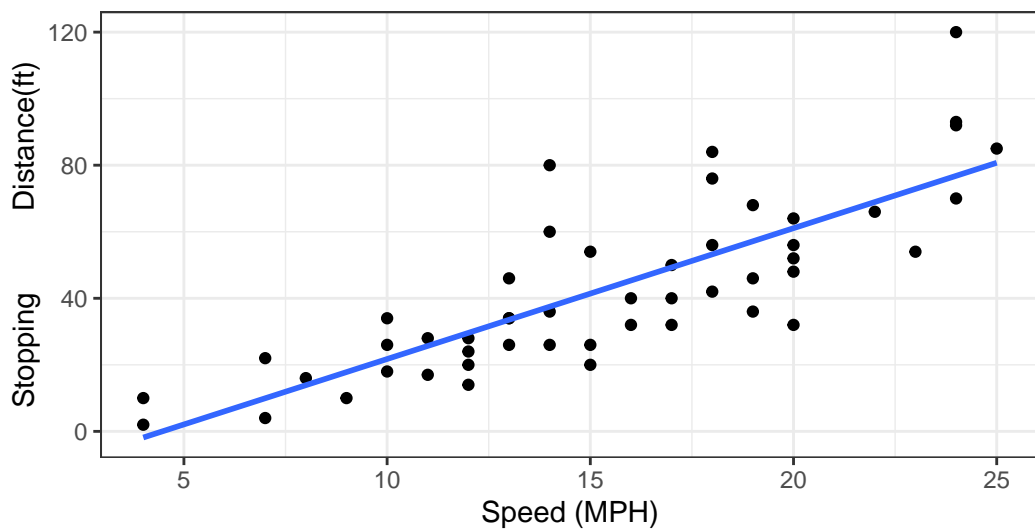


```
ggplot(cars) +
  aes(x = speed, y = dist) +
  geom_point() +
  labs(title="Speed and Stopping Distances of cars", x="Speed (MPH)", y="Stopping Distance") +
  geom_smooth(method = 'lm', se=FALSE) +
  theme_bw()
```

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

Speed and Stopping Distances of cars

Your informative subtitle text here



Dataset: 'cars'

How many genes are upregulated?

```
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 |

```
ncol(genes)
```

```
[1] 4
```

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

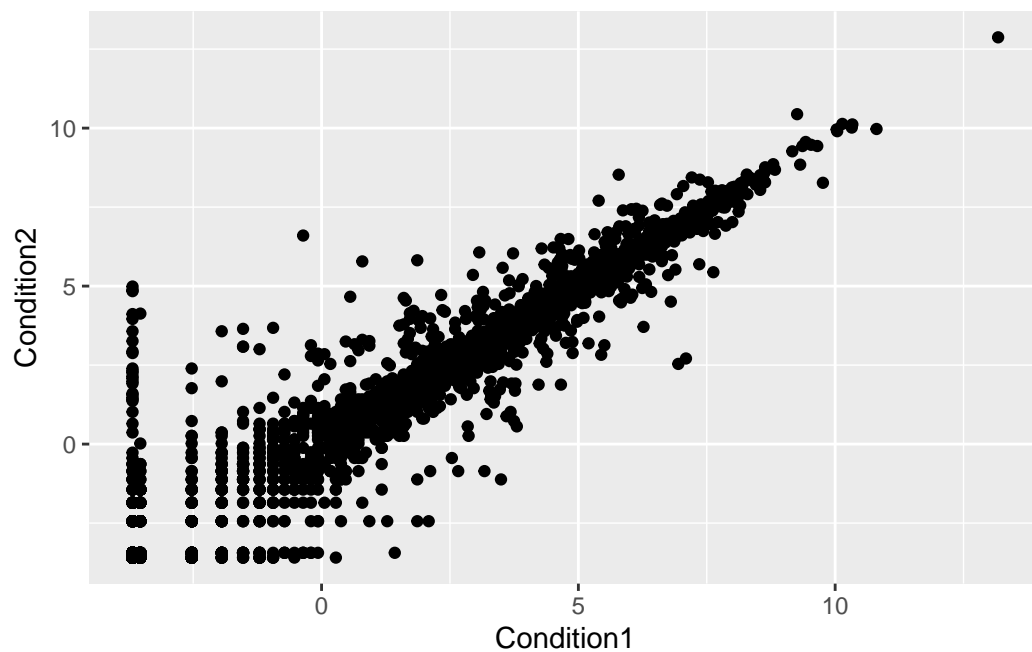
| down | unchanging | up |
|------|------------|-----|
| 72 | 4997 | 127 |

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

| down | unchanging | up |
|------|------------|------|
| 1.39 | 96.17 | 2.44 |

Making a scatter plot for the Genes dataset

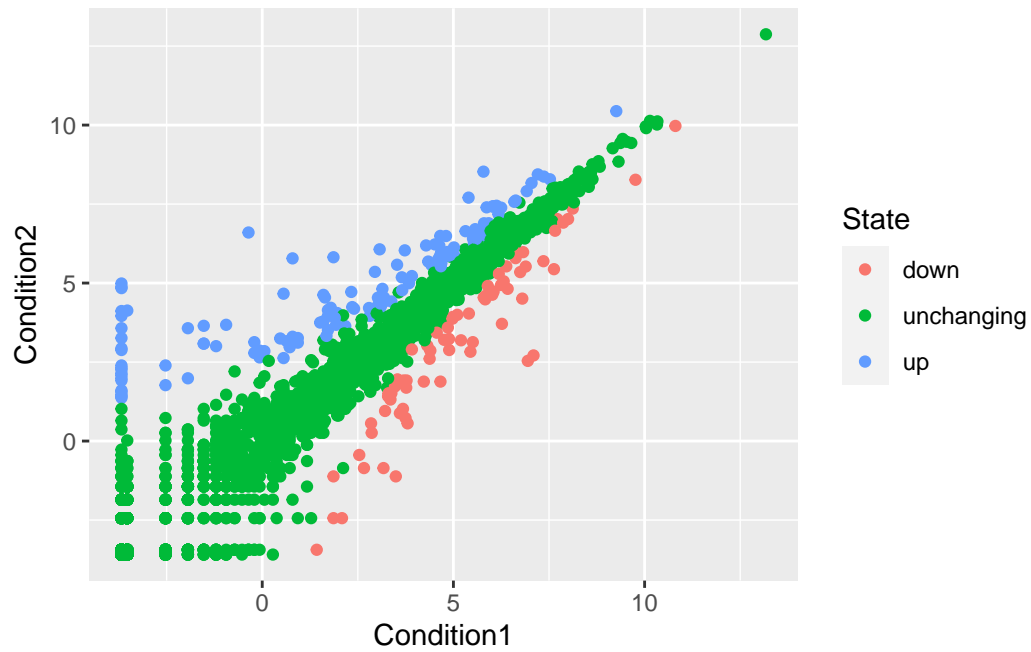
```
ggplot(genes) +  
aes(x=Condition1, y=Condition2) +  
geom_point()
```



Changing the color of the datapoints based on their State:

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

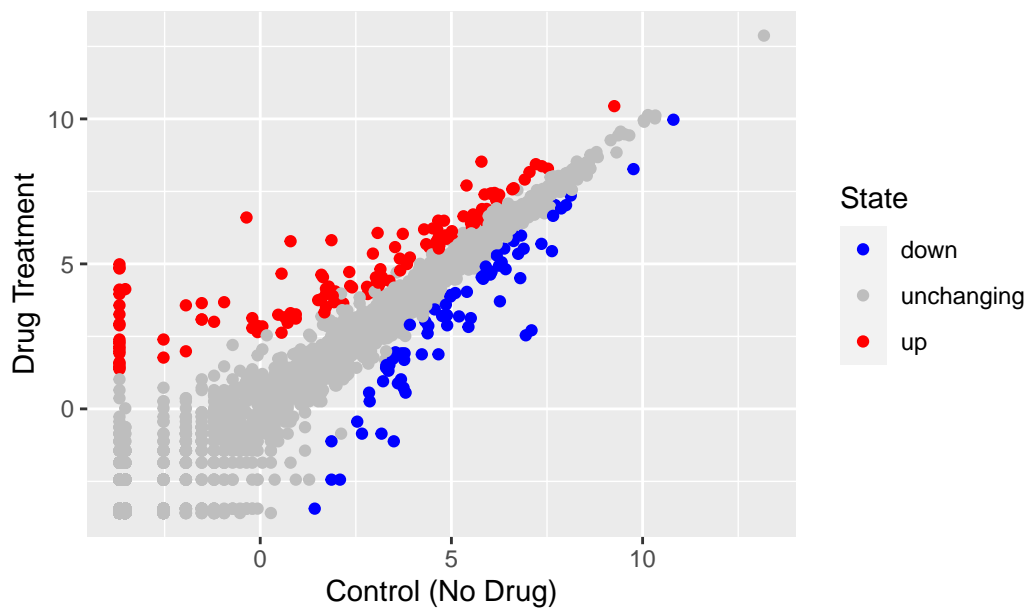
```
geom_point()
p
```



Adding title and changing the x and y labels:

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

Gene Expression Changes Upon Drug Treatment



Downloading the gapminder dataset:

```
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

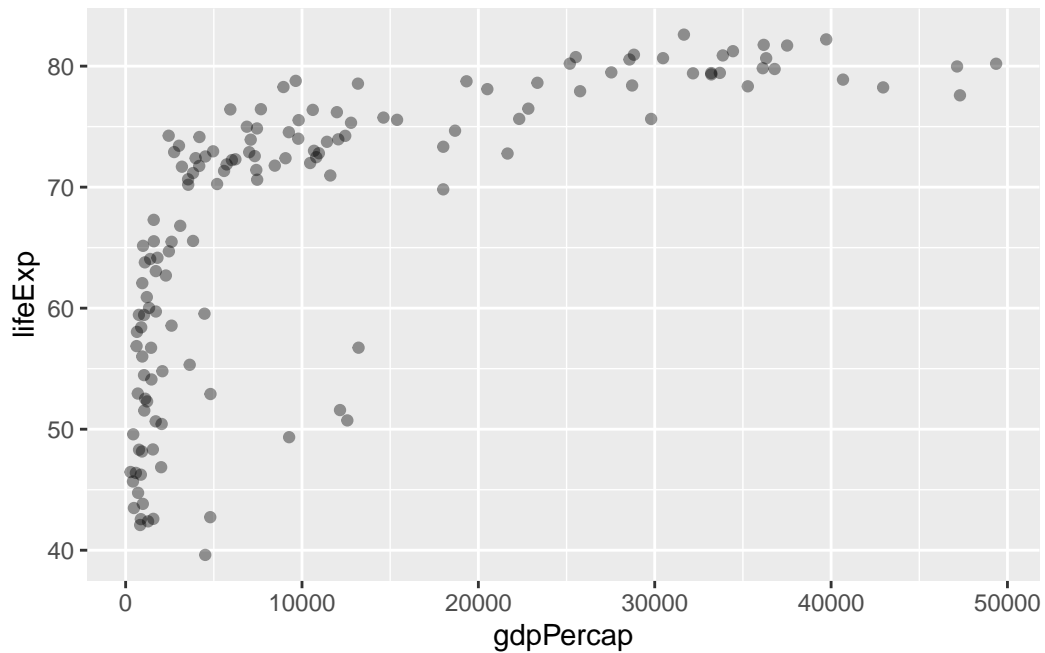
The following objects are masked from 'package:base':

intersect, setdiff, setequal, union

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

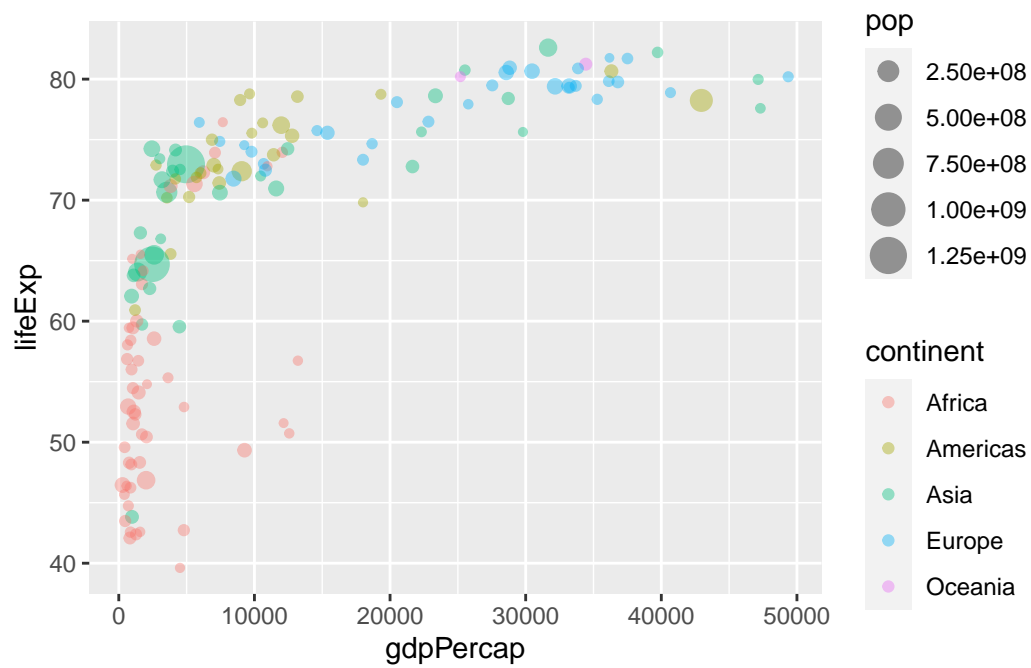
Plotting the gapminder dataset:

```
ggplot(gapminder_2007) +
  aes(x=gdpPerCap,y=lifeExp) +
  geom_point(alpha=0.4)
```



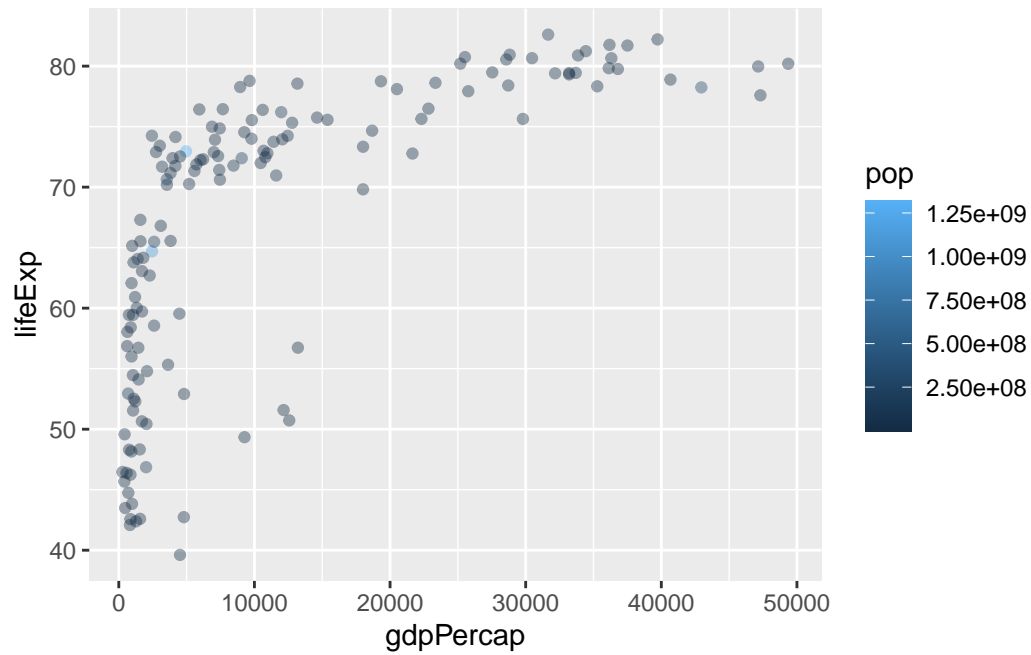
Controlling the aesthetics using the population and the continent aspects of the table:

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



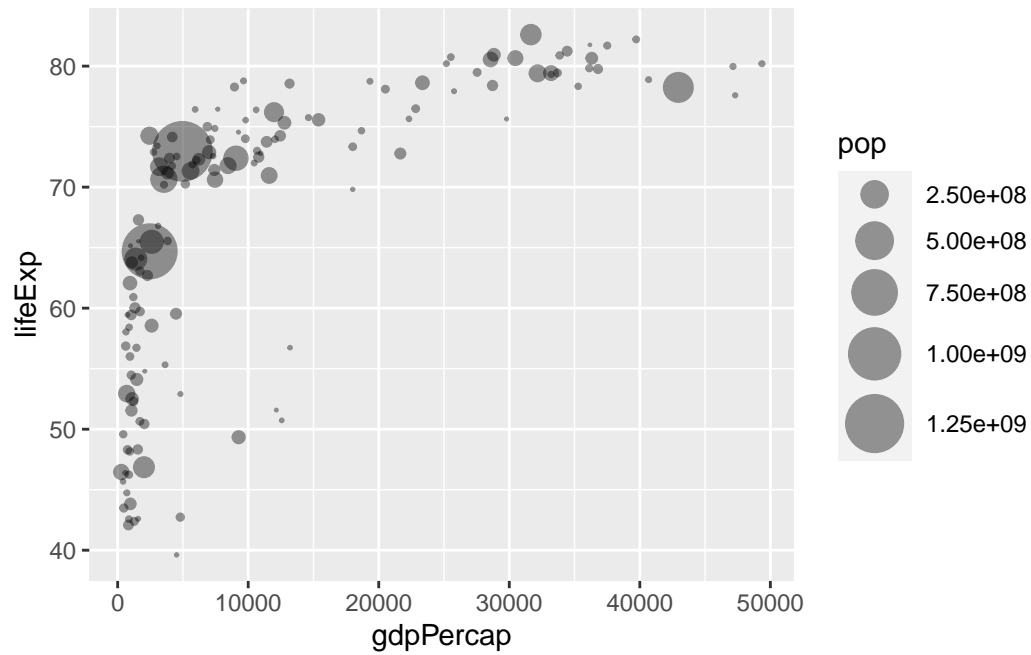
Plotting the scatter plot by using population aspect for color

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

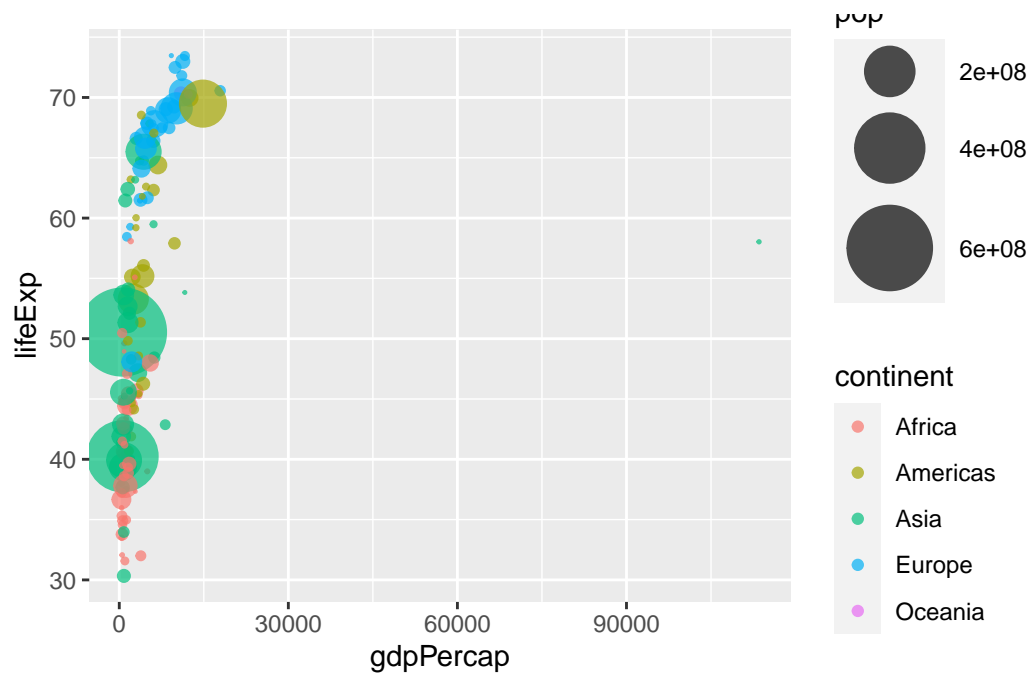
Controlling the size of the data points:

```
ggplot(gapminder_2007) +  
  aes(x=gdpPercap,y=lifeExp, size=pop) +  
  geom_point(alpha=0.4) +  
  scale_size_area(max_size = 10)
```



Plotting for 1957

```
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)
```



Comparing 1957 and 2007:

```
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)
```

