Class5: Data Visualization

Brian Wells (PID: A69026838)

Using GGPLOT

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

Before I use any package functions, I have to laod them up with a `library()` call, like so:

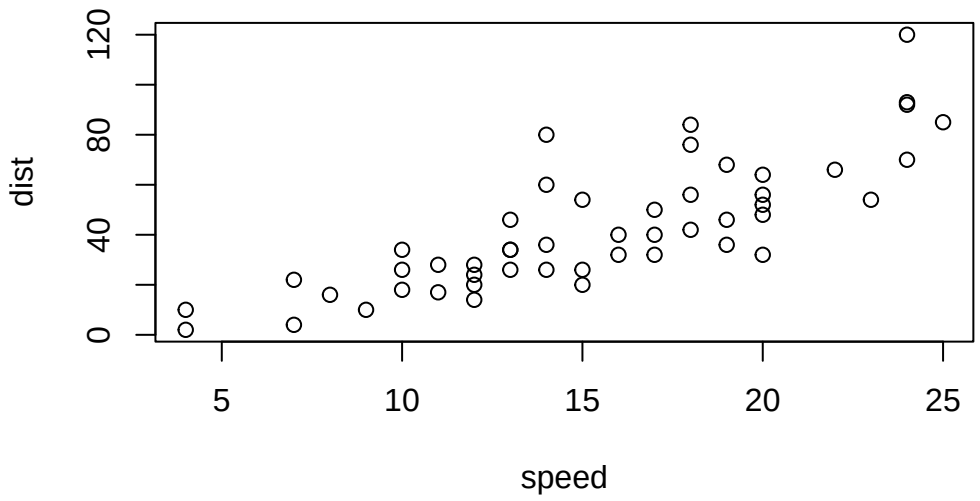
```
# install.packages("ggplot2")
library(ggplot2)
ggplot(cars)
```

```
head(cars)
```

```
  speed dist
1     4    2
2     4   10
3     7    4
4     7   22
5     8   16
6     9   10
```

There's always the "base R" graphics system, i.e. `plot()`

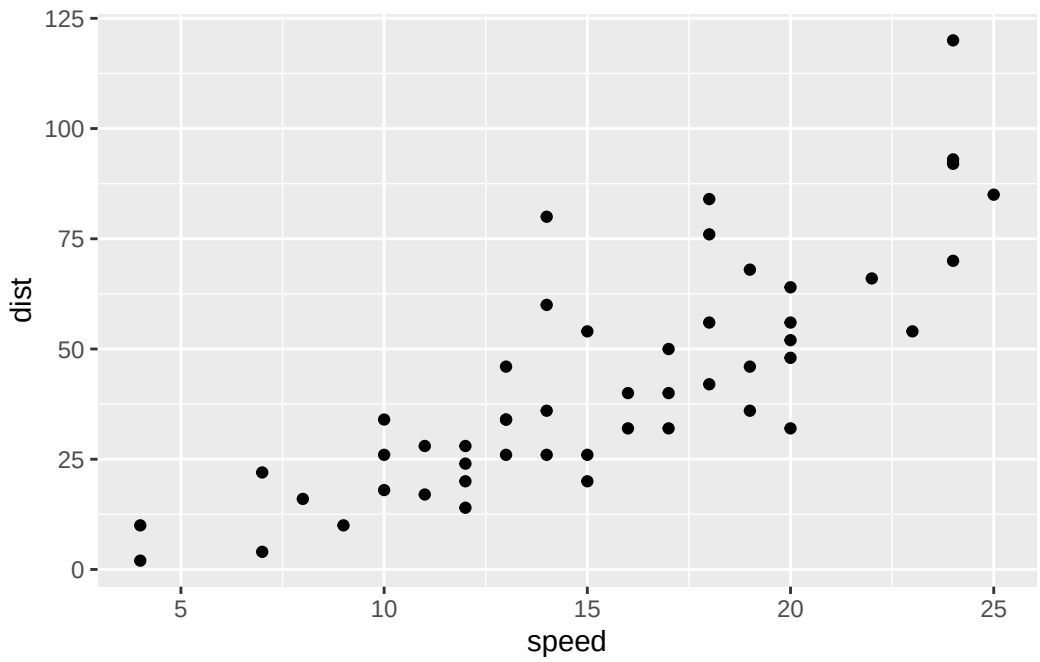
```
plot(cars)
```



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)

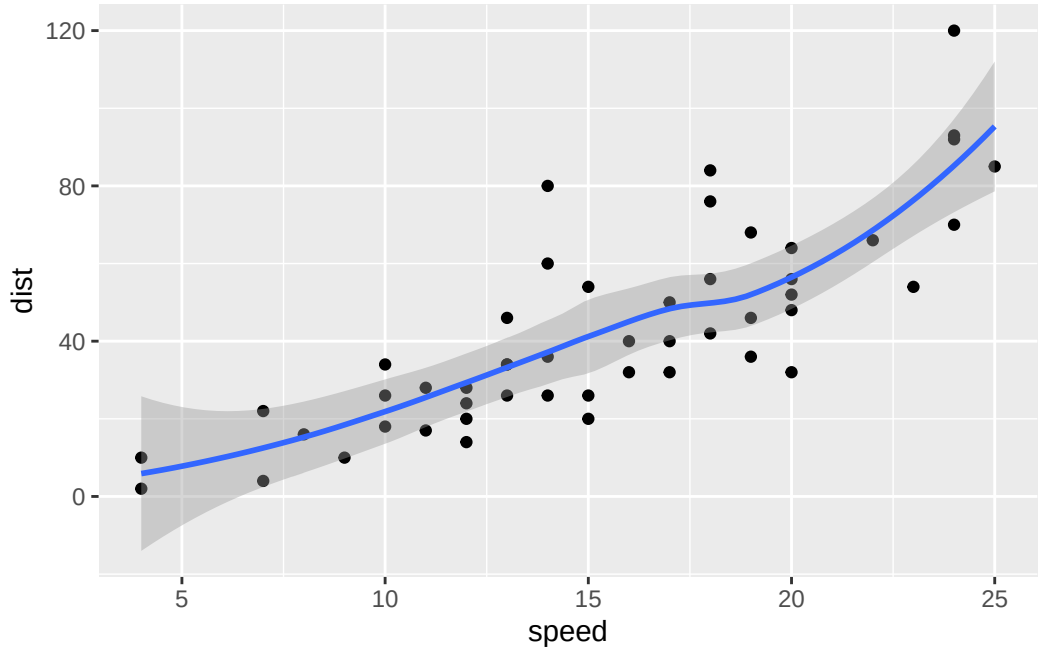
```
ggplot(cars) +
  aes(x=speed, y=dist) +
```

```
  geom_point()
```



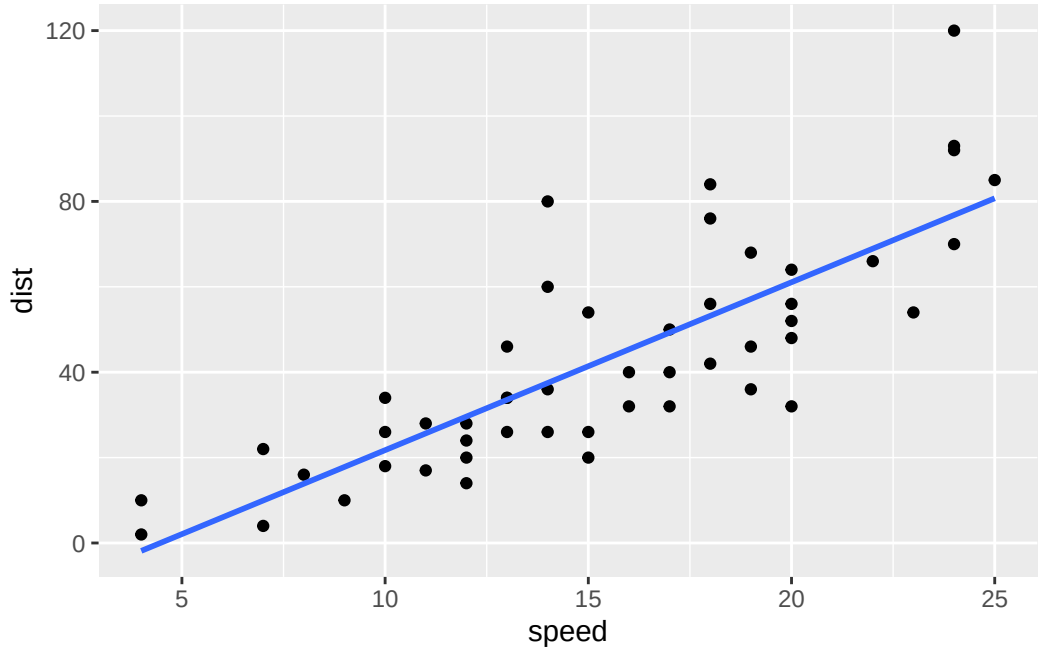

```
ggplot(cars) +
  aes(x=speed, y=dist) +
  geom_point() +
  geom_smooth()
```

```
`geom_smooth()` using method = 'loess' and formula = 'y ~ x'
```



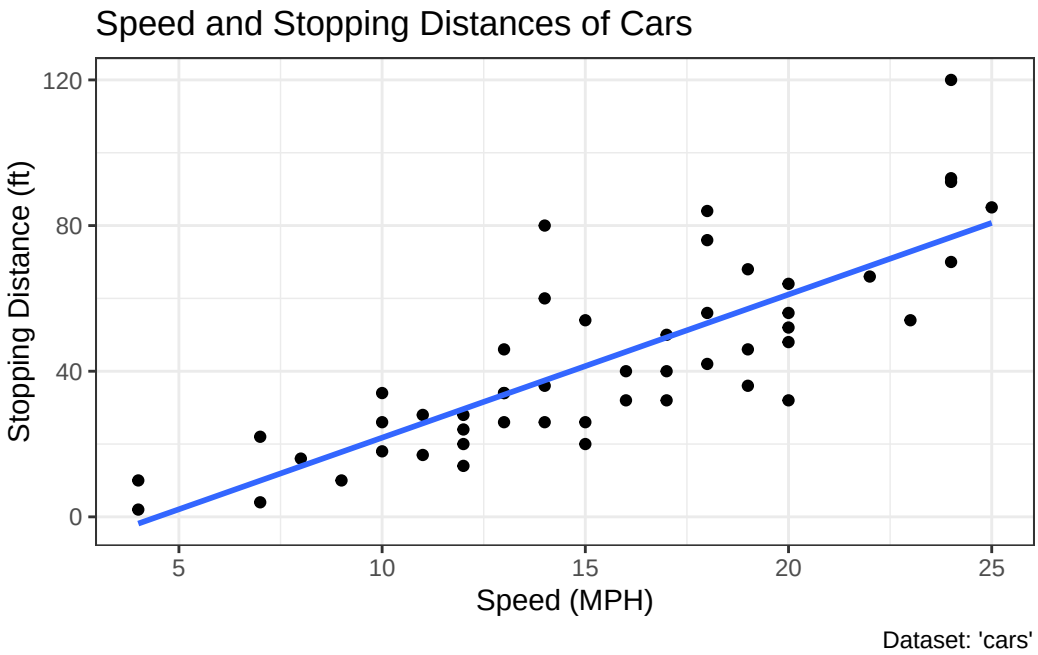

```
ggplot(cars) +
  aes(x=speed, y=dist) +
  geom_point() +
  geom_smooth(method="lm", se=FALSE)
```

```
`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 (ft)",
       caption="Dataset: 'cars'") +
  geom_smooth(method="lm", se=FALSE)+
  theme_bw()
```

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




```
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
```

to get how many genes there are (how many rows)

```
nrow(genes)
```

```
[1] 5196
```

```
colnames(genes)
```

```
[1] "Gene"       "Condition1" "Condition2" "State"
```

```
ncol(genes)
```

```
[1] 4
```

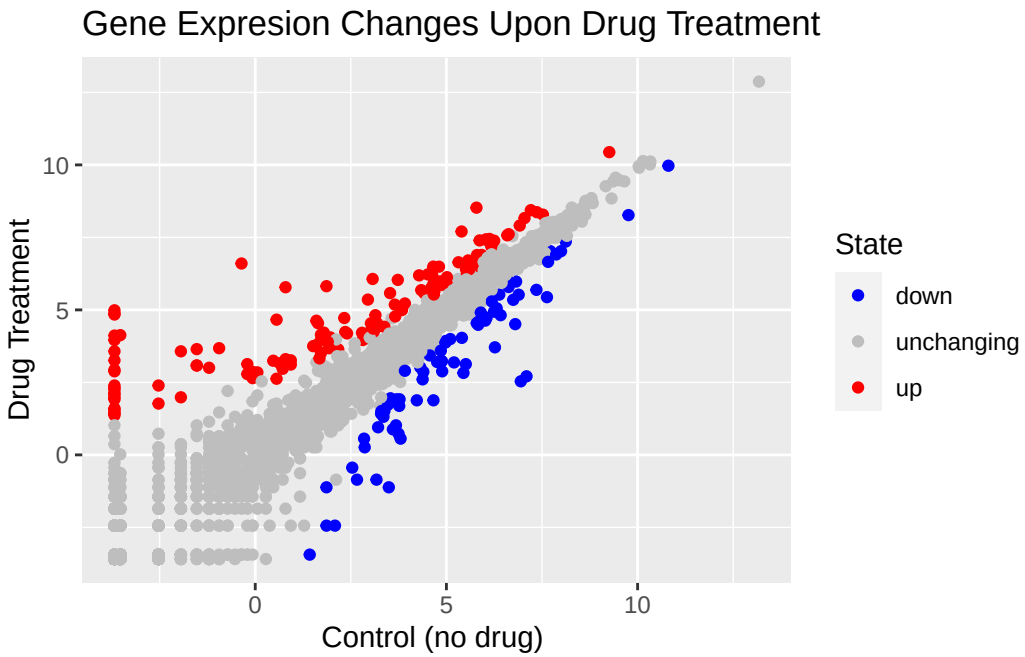
```
table(genes$State)
```

```
    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","grey","red"))+
  labs(title="Gene Expresion Changes Upon Drug Treatment",
       x="Control (no drug) ",
       y="Drug Treatment")
```

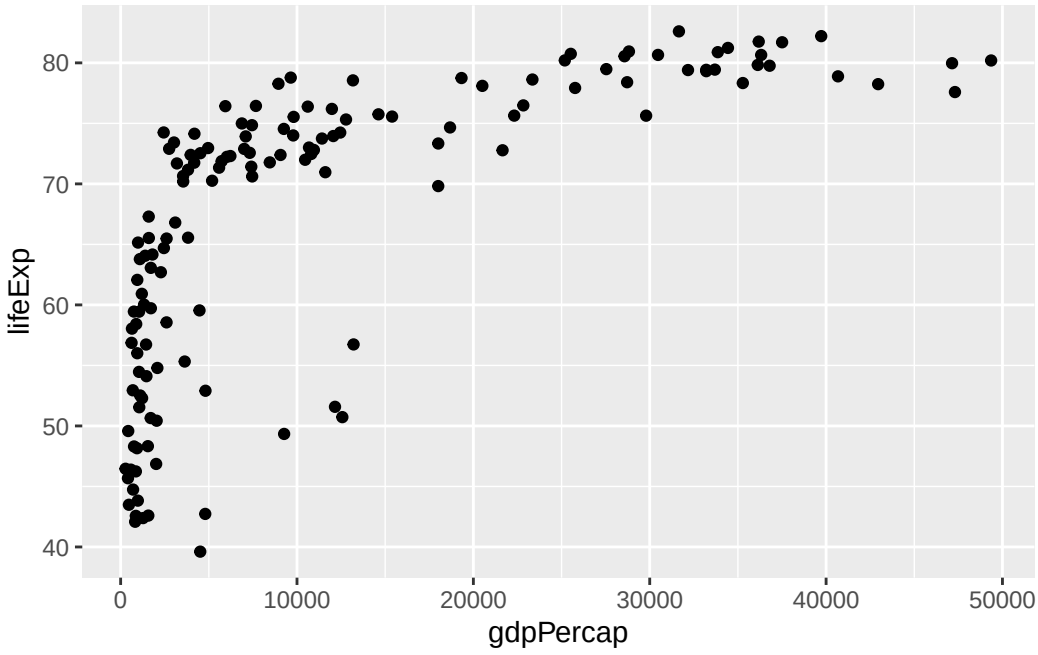



```
# File location online
url <- "https://raw.githubusercontent.com/jennybc/gapminder/master/inst/extdata/gapminder.
gapminder <- read.delim(url)
#install.packages("dplyr")  ## un-comment to install if needed
library(dplyr)
```

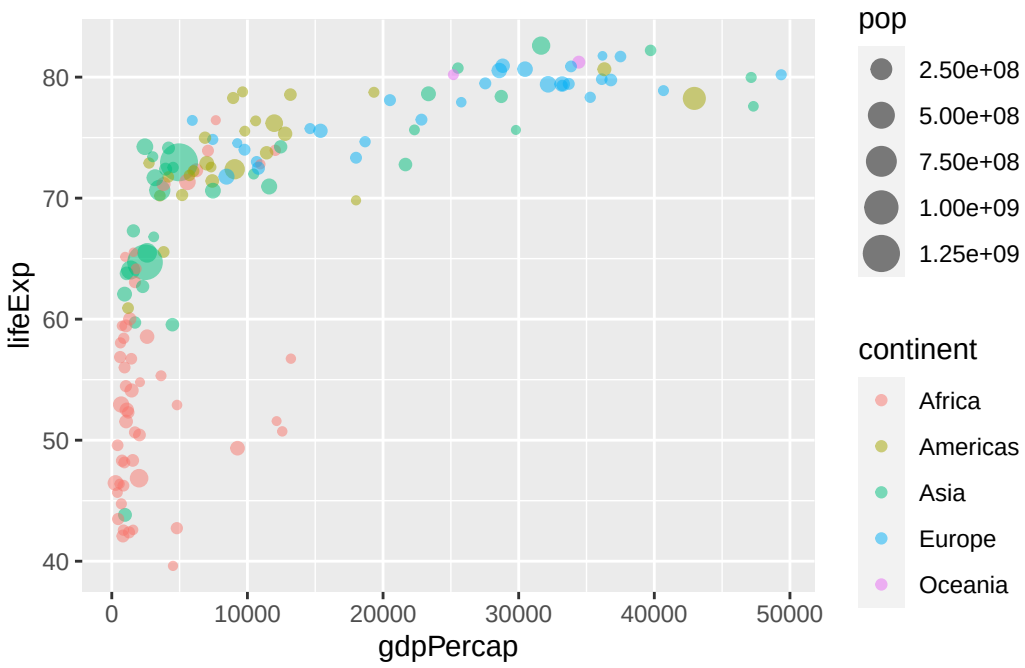
```
Attaching package: 'dplyr'

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) +
```

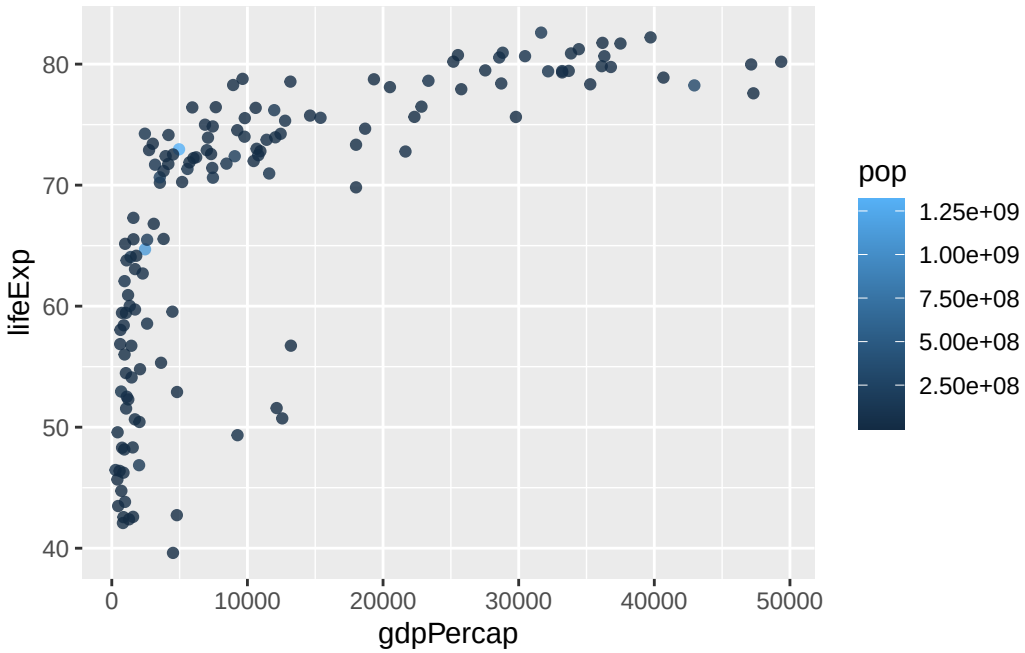
```
  geom_point()
```



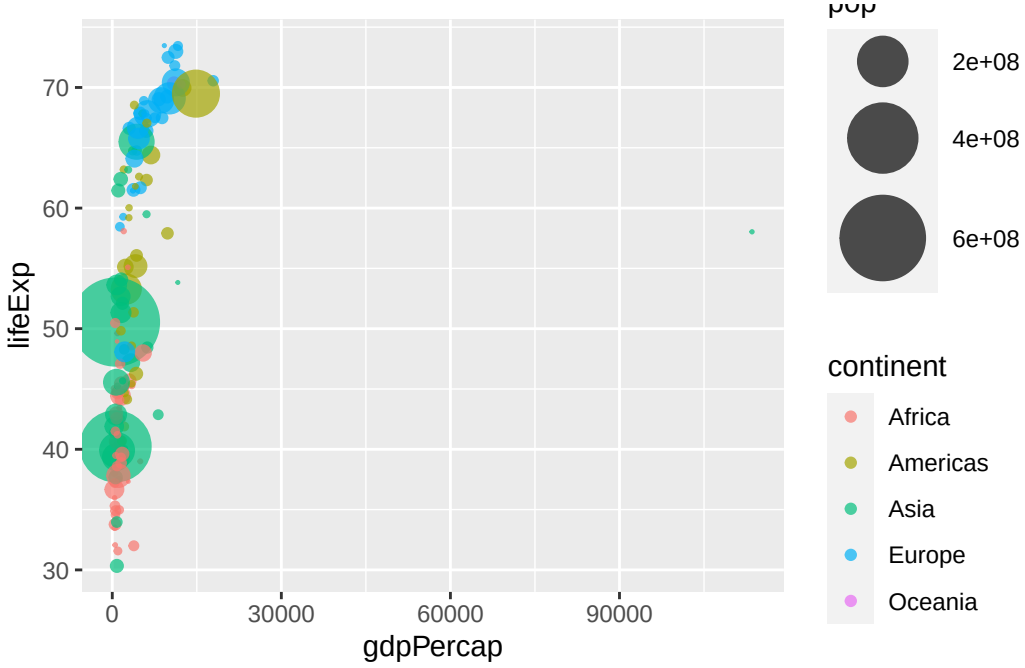

```
ggplot(gapminder_2007)+
  aes(x=gdpPercap, y=lifeExp, color=continent, size=pop) +
  geom_point (alpha=0.5)
```



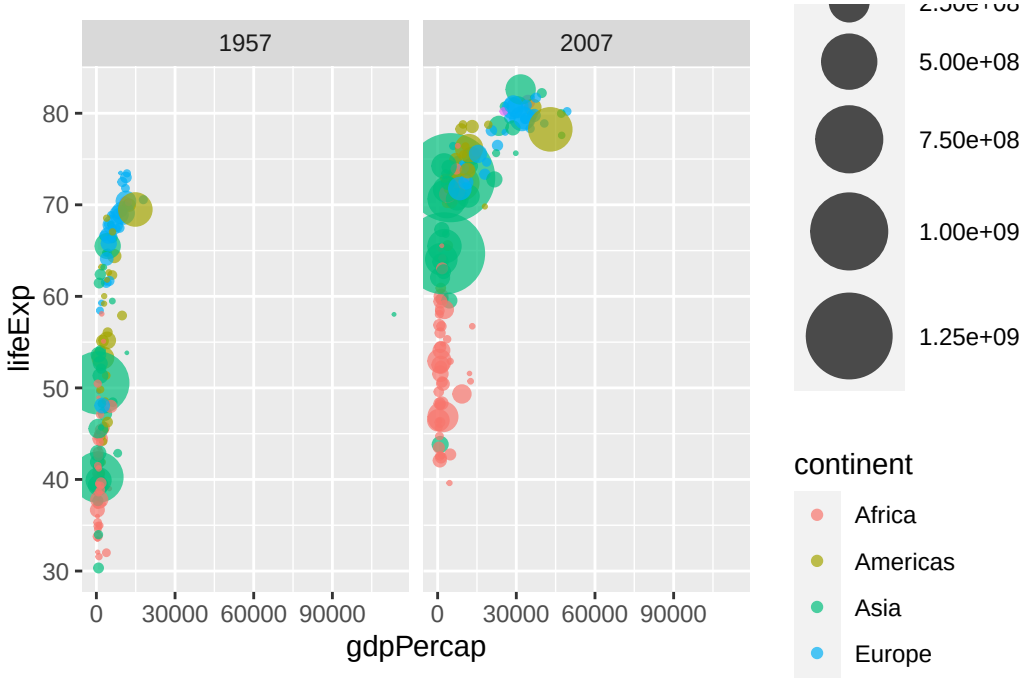

```
gapminder_2007 <- gapminder %>% filter(year==2007)
ggplot(gapminder_2007)+
  aes(x=gdpPercap, y=lifeExp, color=pop) +
  geom_point (alpha=0.8)
```




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

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


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