

Datacamp_Introduction_to_Tidyverse_Data visualization

dizhen

2019/4/3

Visualizing with ggplot2

```
library(ggplot2)
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
```

```
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.2.1 --

## v tibble  2.1.3      v purrr   0.3.2
## v tidyr   0.8.3      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.4.0

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
gapminder <- read.table(file = 'data/gapminder.tsv', sep = '\t', header = TRUE)
gapminder_2007 <- gapminder %>%
  filter(year == 2007)
```

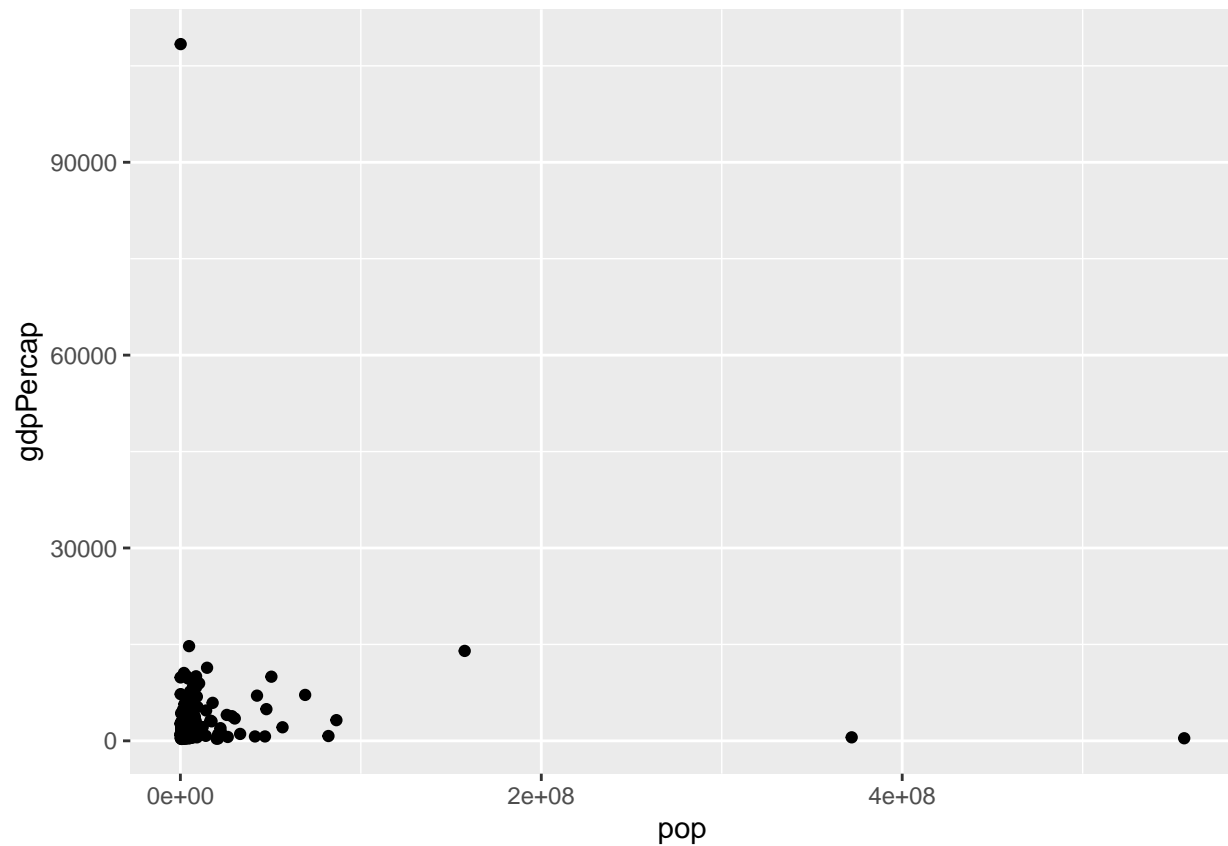
Practice

```
# Create gapminder_1952
gapminder_1952 <- gapminder %>% filter(year == 1952)
```

Comparing population and GDP per capita

```
gapminder_1952 <- gapminder %>%
  filter(year == 1952)

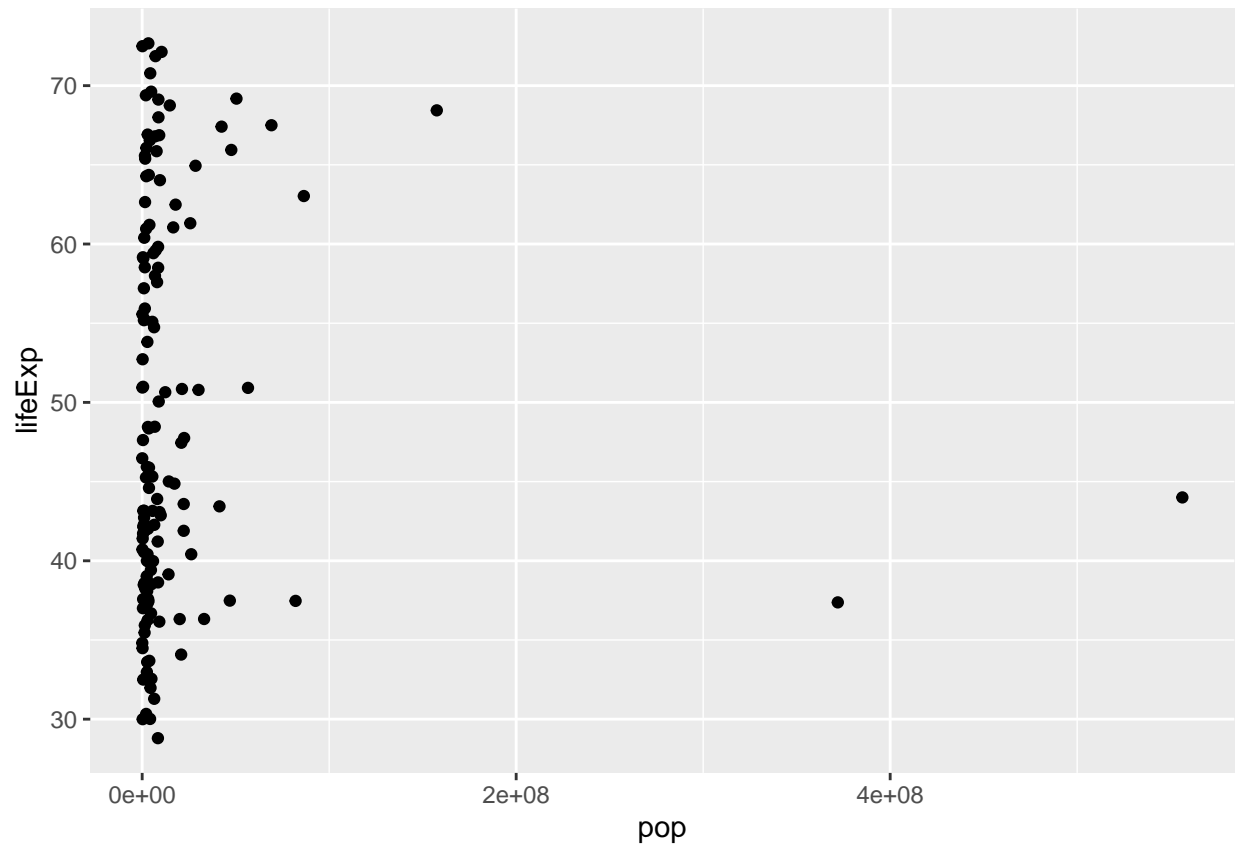
# Change to put pop on the x-axis and gdpPercap on the y-axis
ggplot(gapminder_1952, aes(x = pop, y = gdpPercap)) +
  geom_point()
```



Comparing population and life expectancy

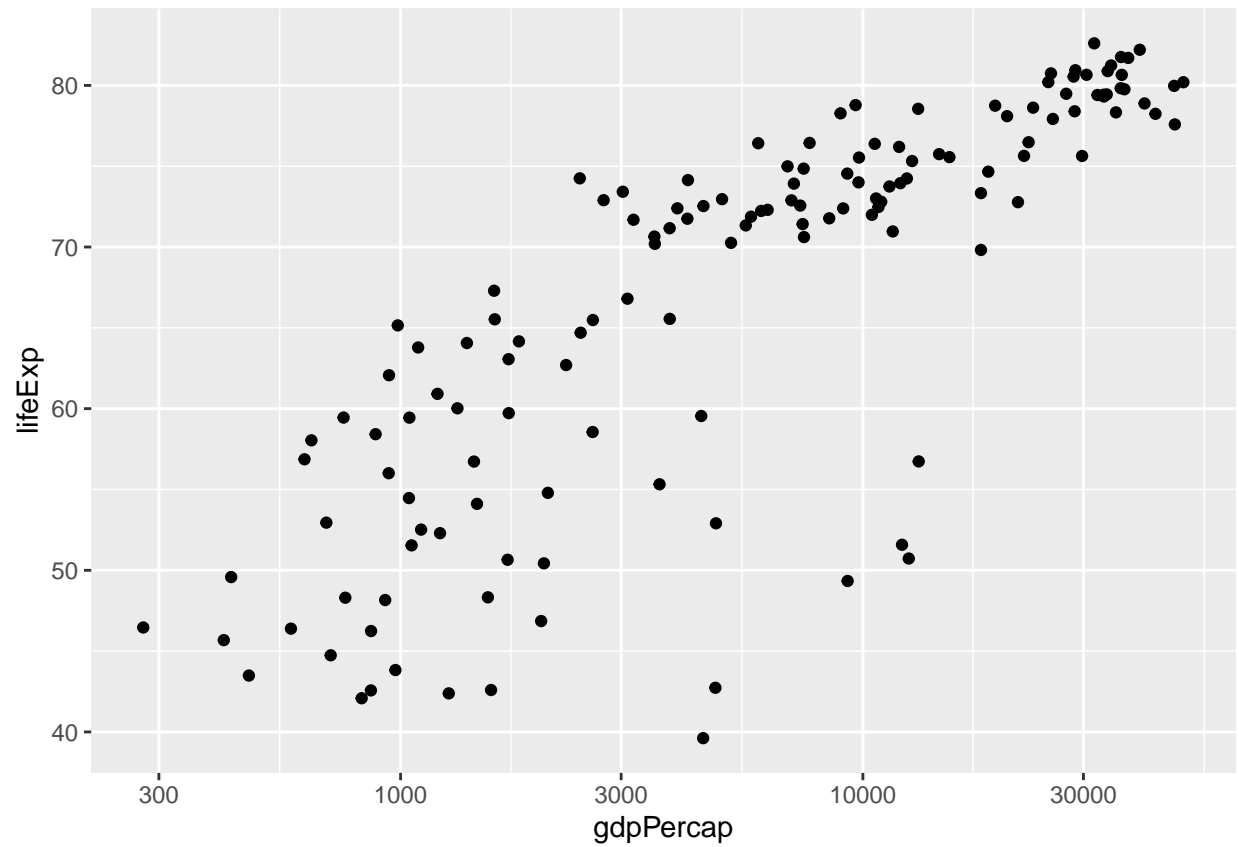
```
gapminder_1952 <- gapminder %>%
  filter(year == 1952)

# Create a scatter plot with pop on the x-axis and lifeExp on the y-axis
ggplot(gapminder_1952, aes(x = pop, y = lifeExp)) + geom_point()
```



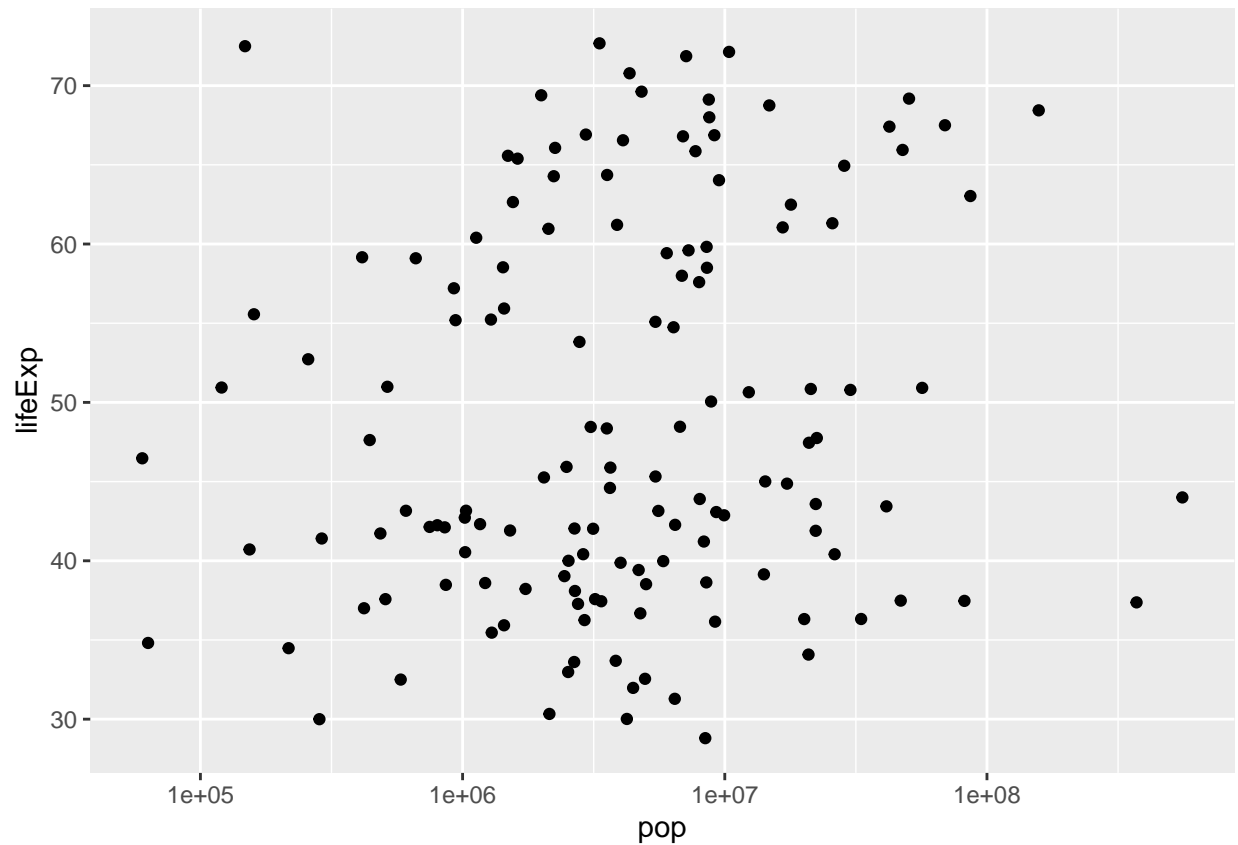
Log scales

```
# ggplot(gapminder_2007, aes(x = gdpPerCap, y = lifeExp)) +  
#   geom_point()  
  
ggplot(gapminder_2007, aes(x = gdpPerCap, y = lifeExp)) +  
  geom_point() +  
  scale_x_log10()
```

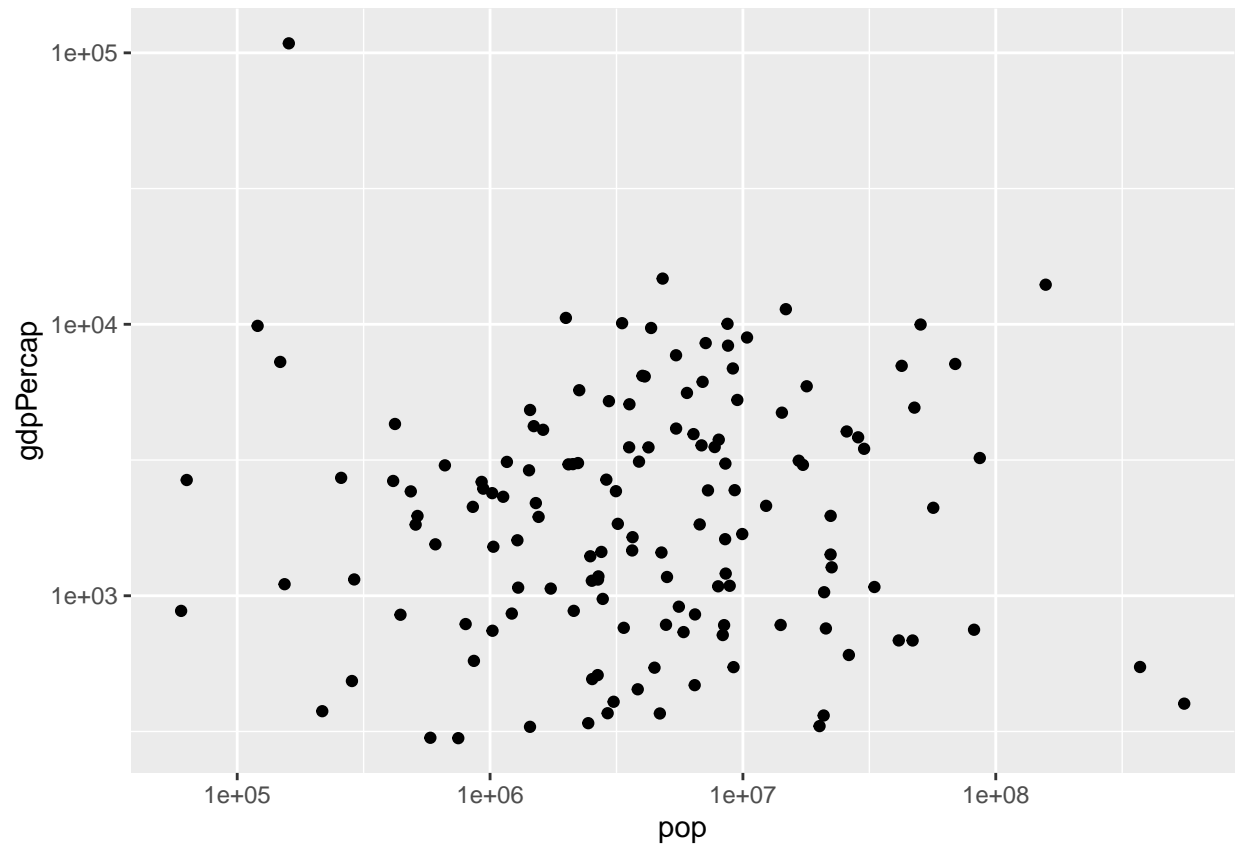


Practice

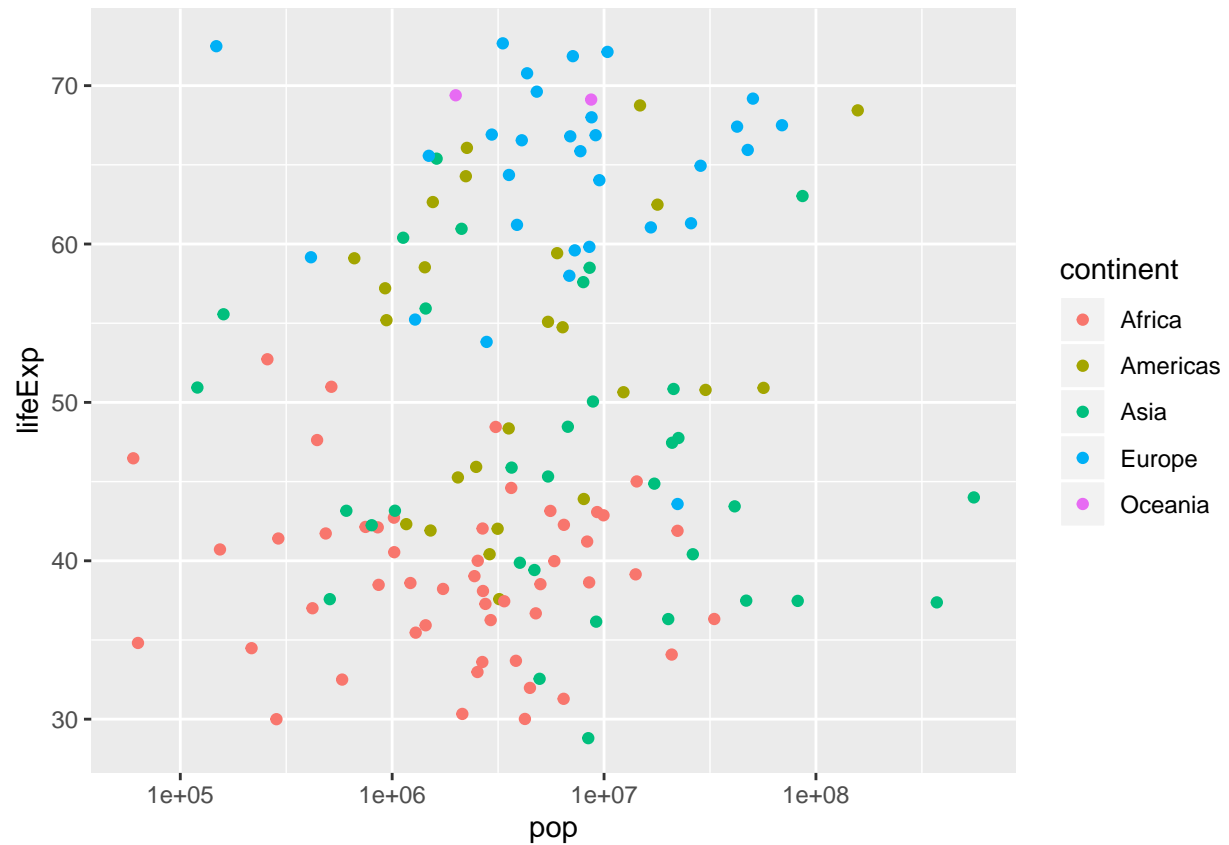
```
gapminder_1952 <- gapminder %>%  
  filter(year == 1952)  
  
# Change this plot to put the x-axis on a log scale  
ggplot(gapminder_1952, aes(x = pop, y = lifeExp)) +  
  geom_point() + scale_x_log10()
```



```
gapminder_1952 <- gapminder %>%  
  filter(year == 1952)  
  
# Scatter plot comparing pop and gdpPercap, with both axes on a log scale  
ggplot(gapminder_1952, aes(x = pop, y = gdpPercap)) + geom_point() + scale_y_log10() + scale_x_log10()
```

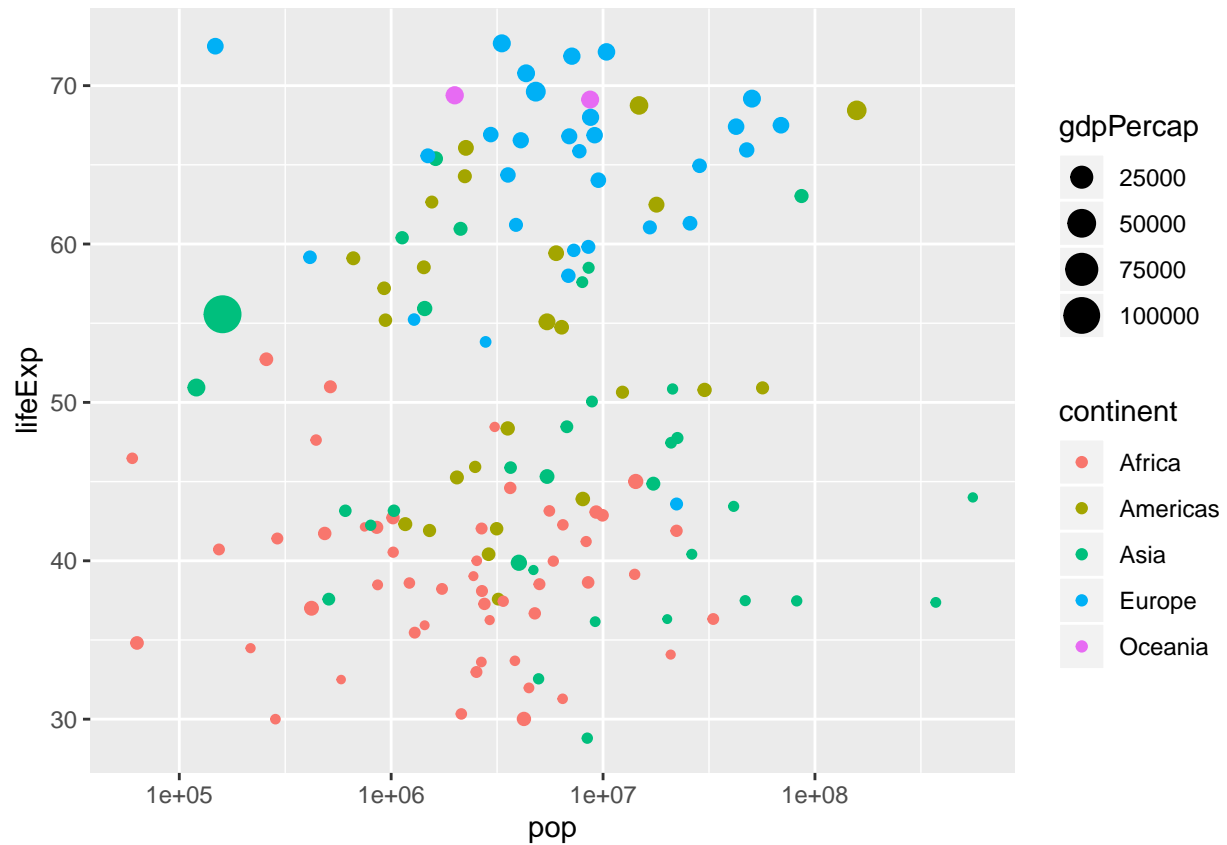


```
gapminder_1952 <- gapminder %>%  
  filter(year == 1952)  
  
# Scatter plot comparing pop and lifeExp, with color representing continent  
ggplot(gapminder_1952, aes(x = pop, y = lifeExp, color = continent)) + geom_point() + scale_x_log10()
```



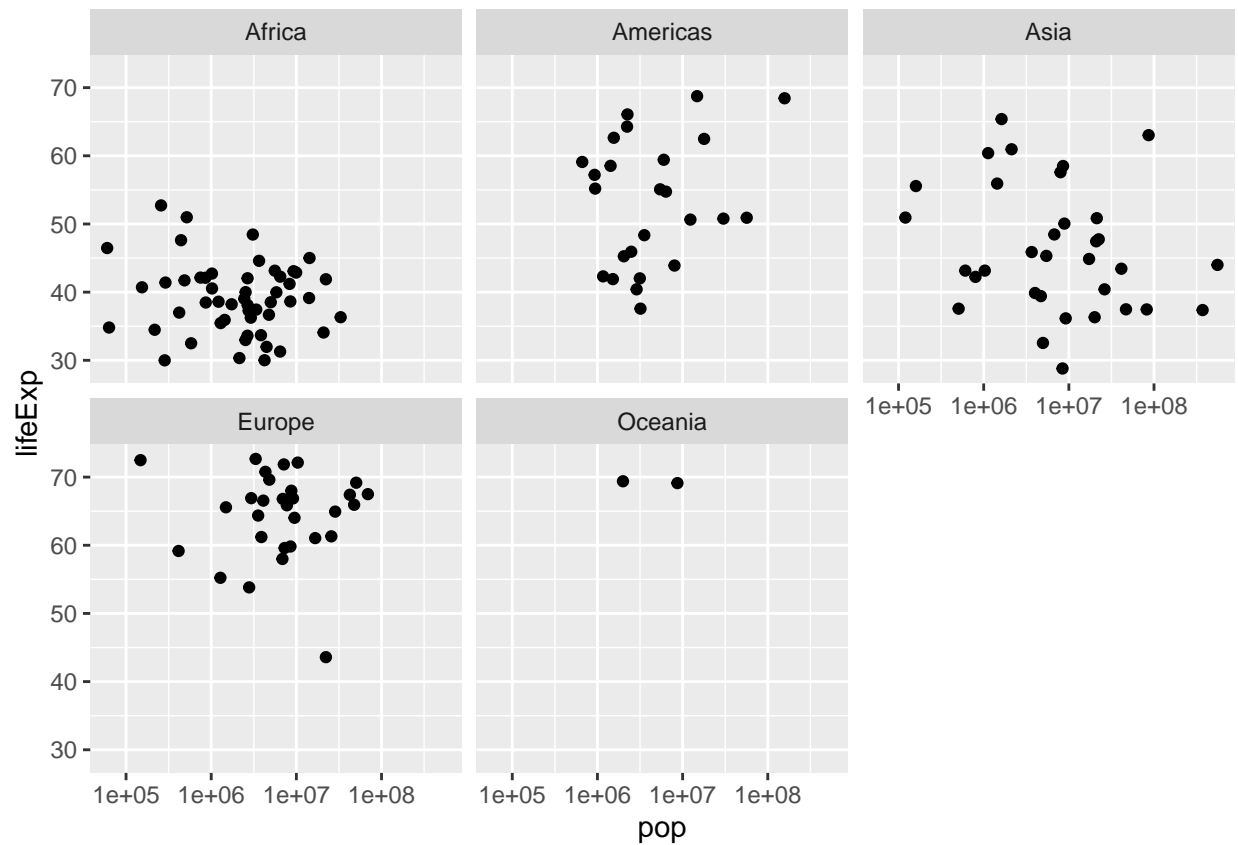
```
gapminder_1952 <- gapminder %>%
  filter(year == 1952)

# Add the size aesthetic to represent a country's gdpPerCap
ggplot(gapminder_1952, aes(x = pop, y = lifeExp, color = continent, size = gdpPerCap)) +
  geom_point() +
  scale_x_log10()
```



```
gapminder_1952 <- gapminder %>%
  filter(year == 1952)

# Scatter plot comparing pop and lifeExp, faceted by continent
ggplot(gapminder_1952, aes(x = pop, y = lifeExp)) + geom_point() + scale_x_log10() + facet_wrap(~continent)
```

*# Scatter plot comparing gdpPerCap and lifeExp, with color representing continent
and size representing population, faceted by year*

```
ggplot(gapminder, aes(x = gdpPerCap, y = lifeExp, color = continent, size = pop)) + geom_point() + scale_x.
```

