

## class05.R

vince

2022-02-01

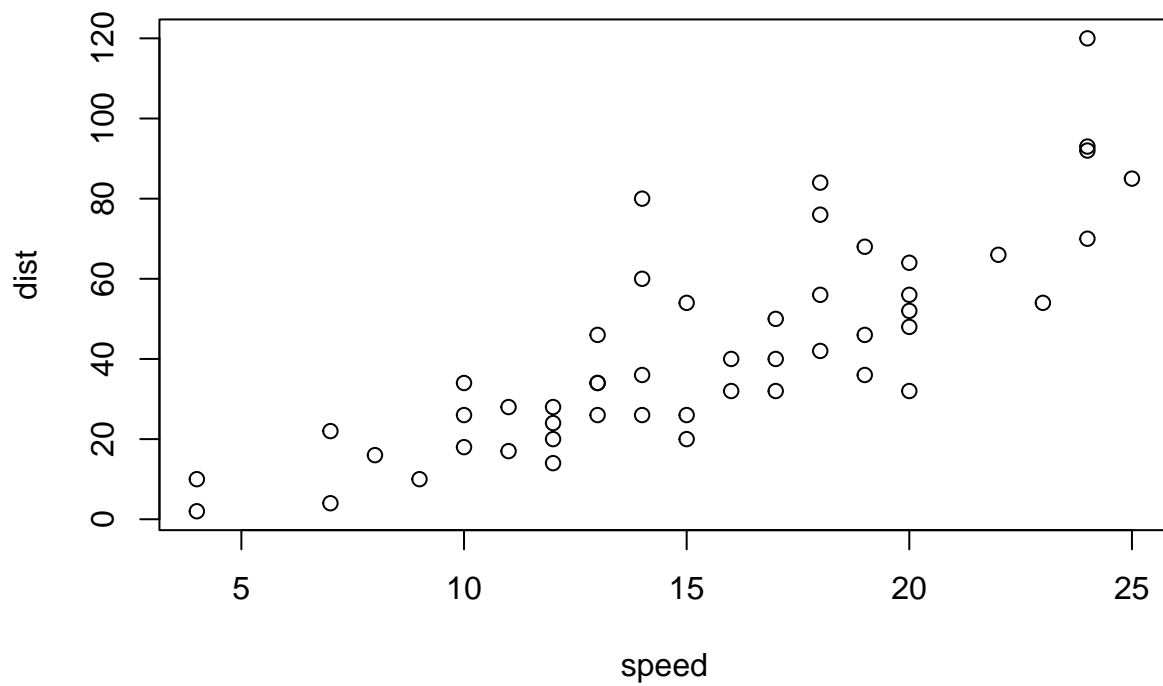
```
# Class 05 Data Visualization
```

```
# This is the "base" R plot  
plot(cars)
```

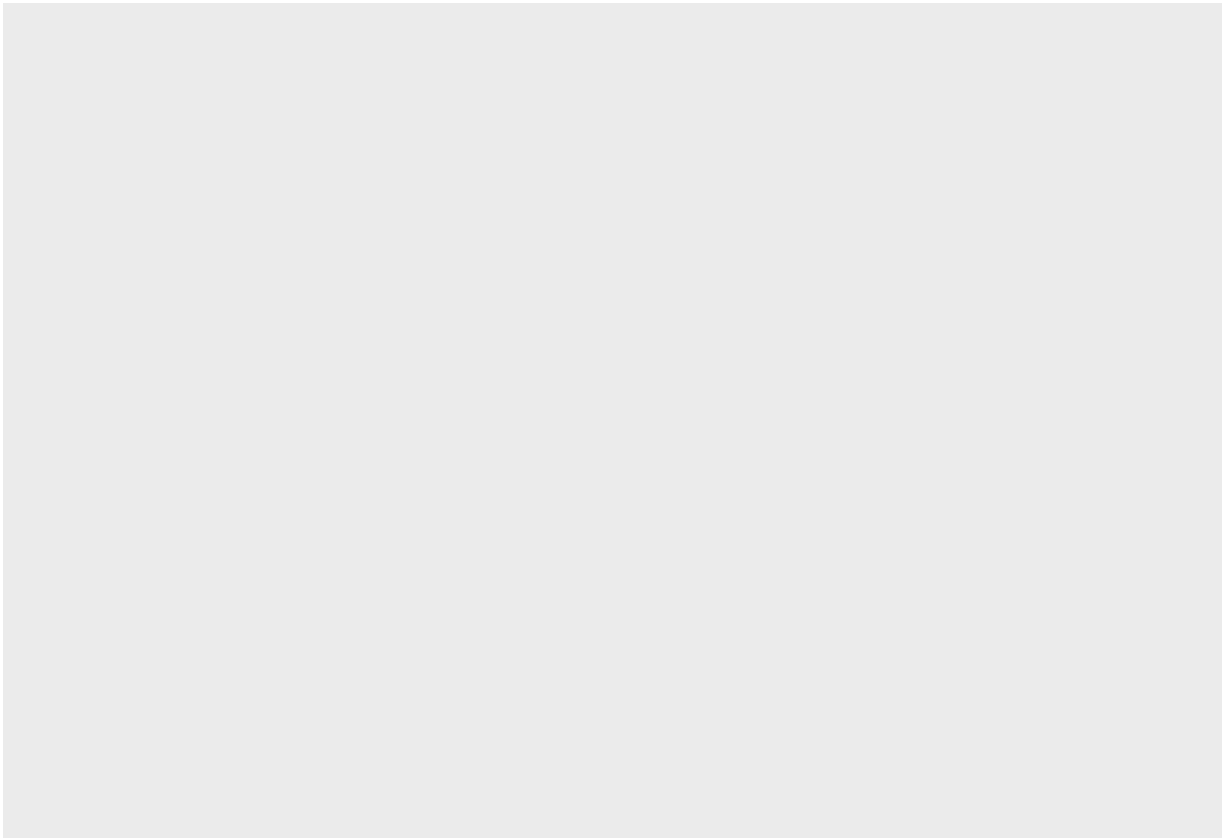
```
# Get a new plotting package/library called ggplot2  
# install.packages("ggplot2")
```

```
# Call/load the package into the R brain  
library(ggplot2)
```

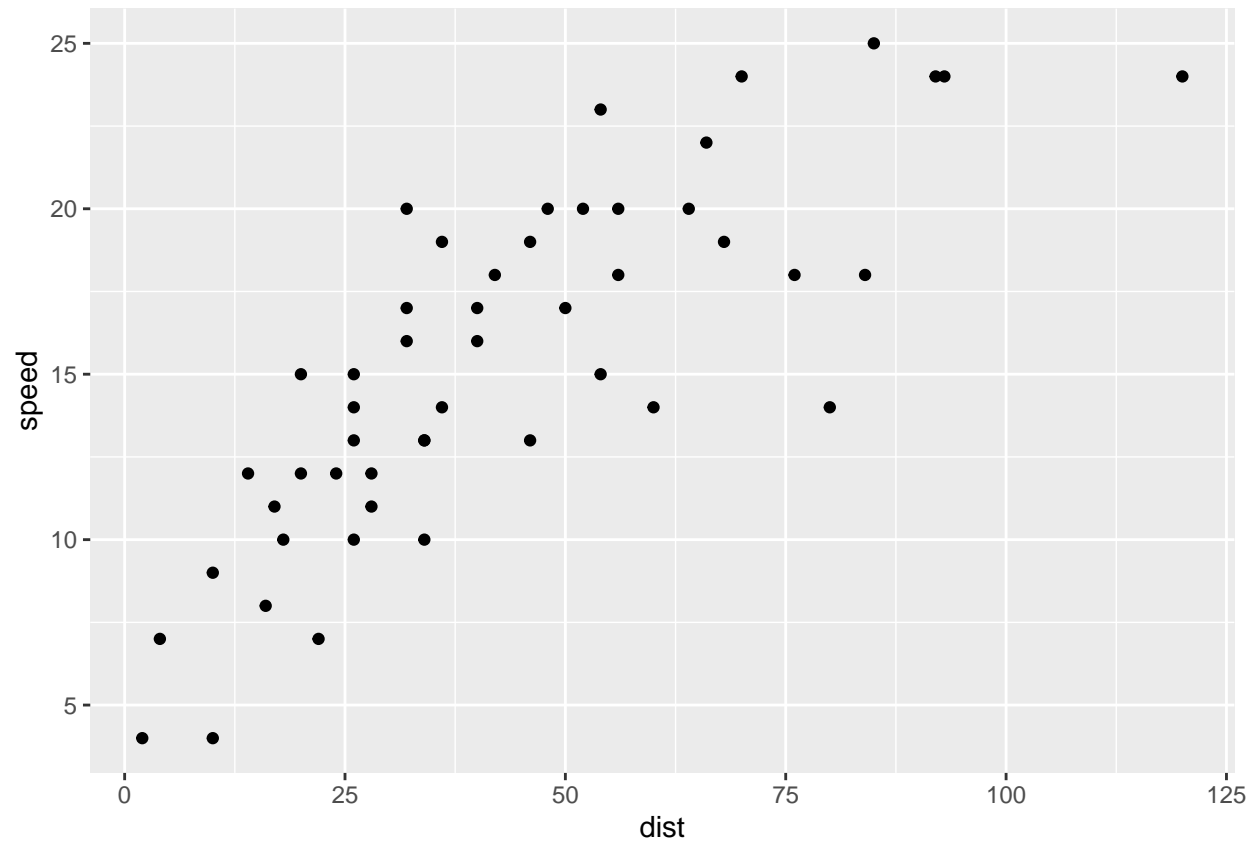
```
## Warning in register(): Can't find generic 'scale_type' in package ggplot2 to  
## register S3 method.
```



```
# Sets up the plot  
ggplot(cars)
```

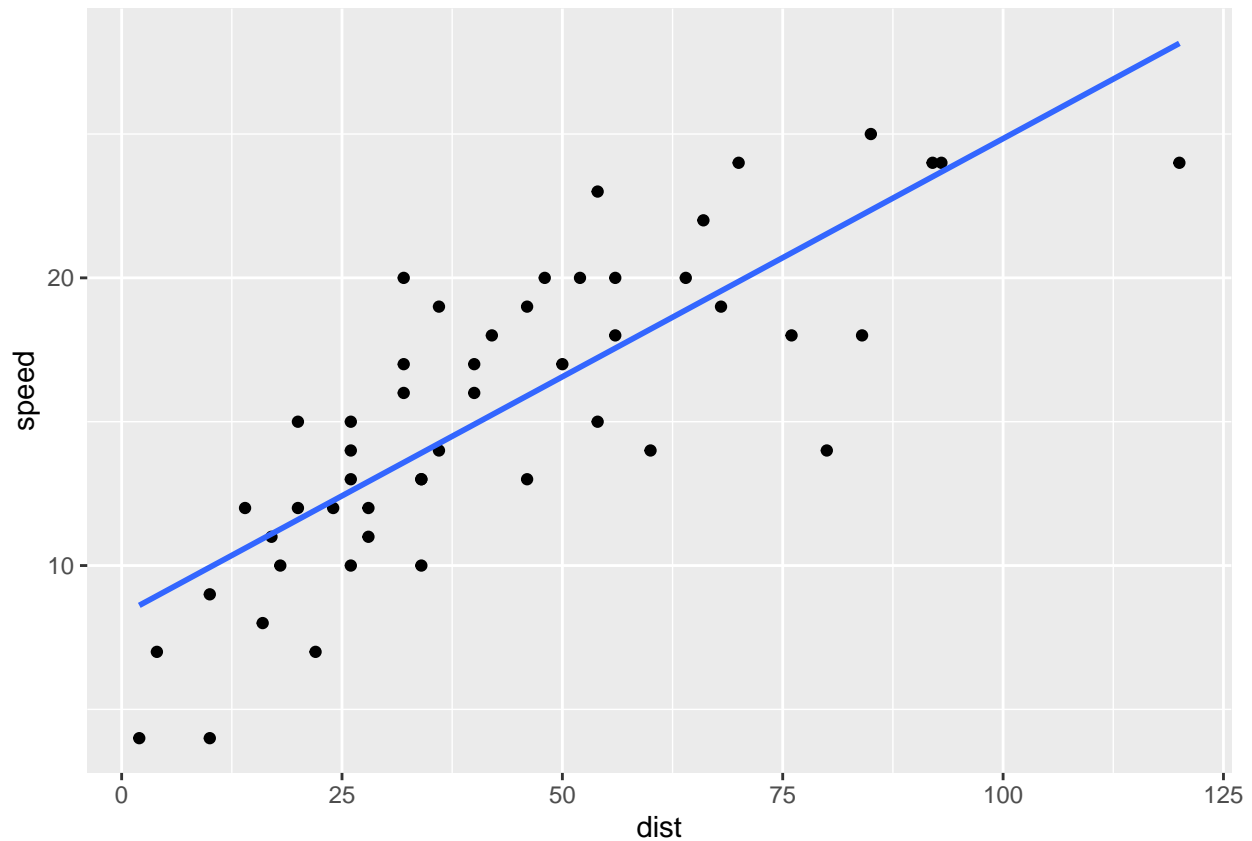


```
ggplot(data=cars) + aes(x=dist, y=speed) + geom_point()
```



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

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

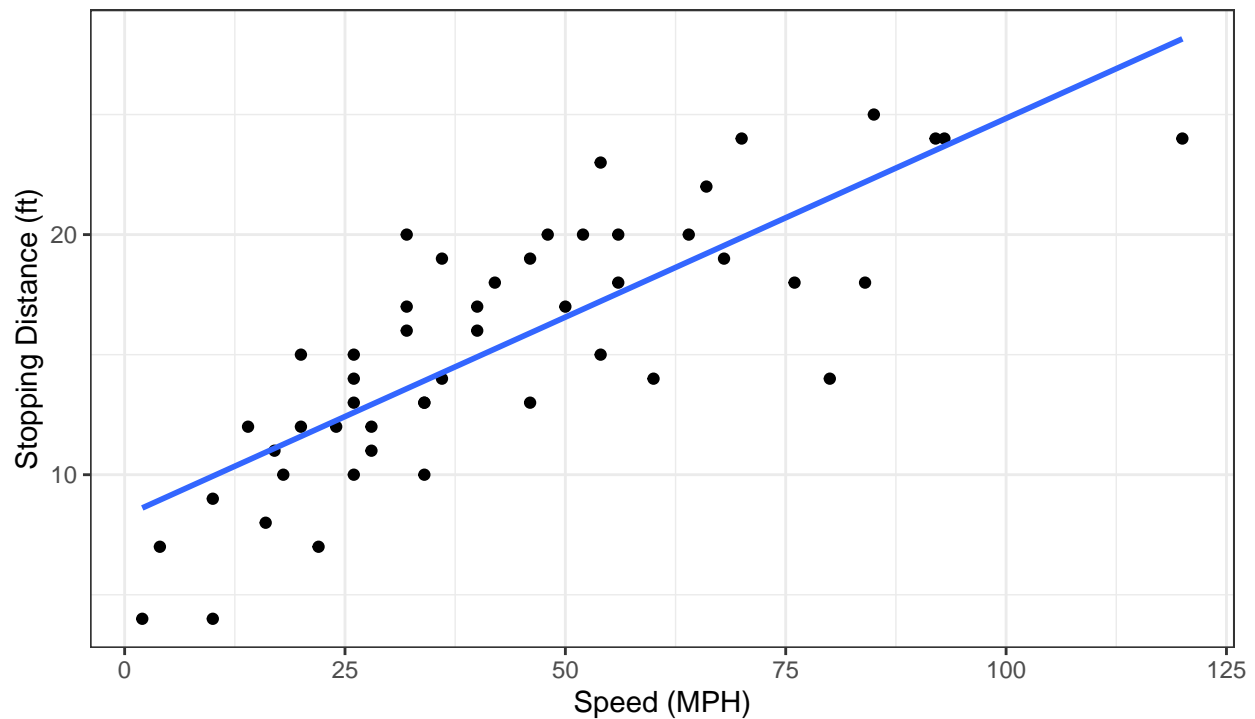


```
# Add more layers
ggplot(data=cars) +
  aes(x=dist, y=speed) +
  geom_point() +
  geom_smooth(method="lm", se=FALSE) +
  labs(title="Speed and Stopping Distances of Cars",
       x="Speed (MPH)",
       y="Stopping Distance (ft)",
       subtitle="For a given speed, how far must a car go to stop?",
       caption="Dataset: 'cars'",
       ) +
  theme_bw()
```

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

## Speed and Stopping Distances of Cars

For a given speed, how far must a car go to stop?



Dataset: 'cars'

```
# RNASeq experiment data set
# Read the data in R
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)
```

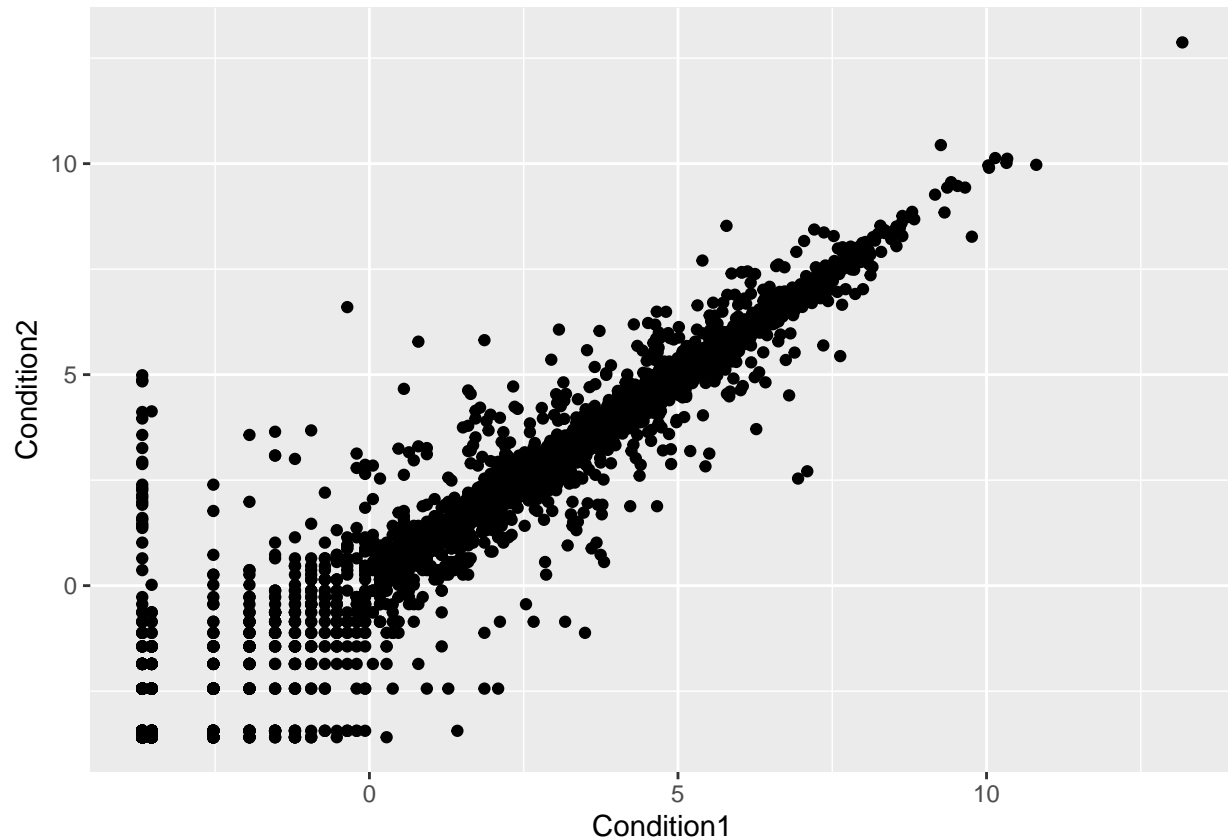
```
## [1] "Gene"      "Condition1" "Condition2" "State"
```

```
ncol(genes)
```

```
## [1] 4
```

```
# Plot the result
```

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



```
# Q. How many genes are up-regulated?
```

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

```
##  
##      down  unchanged      up  
##       72      4997     127
```

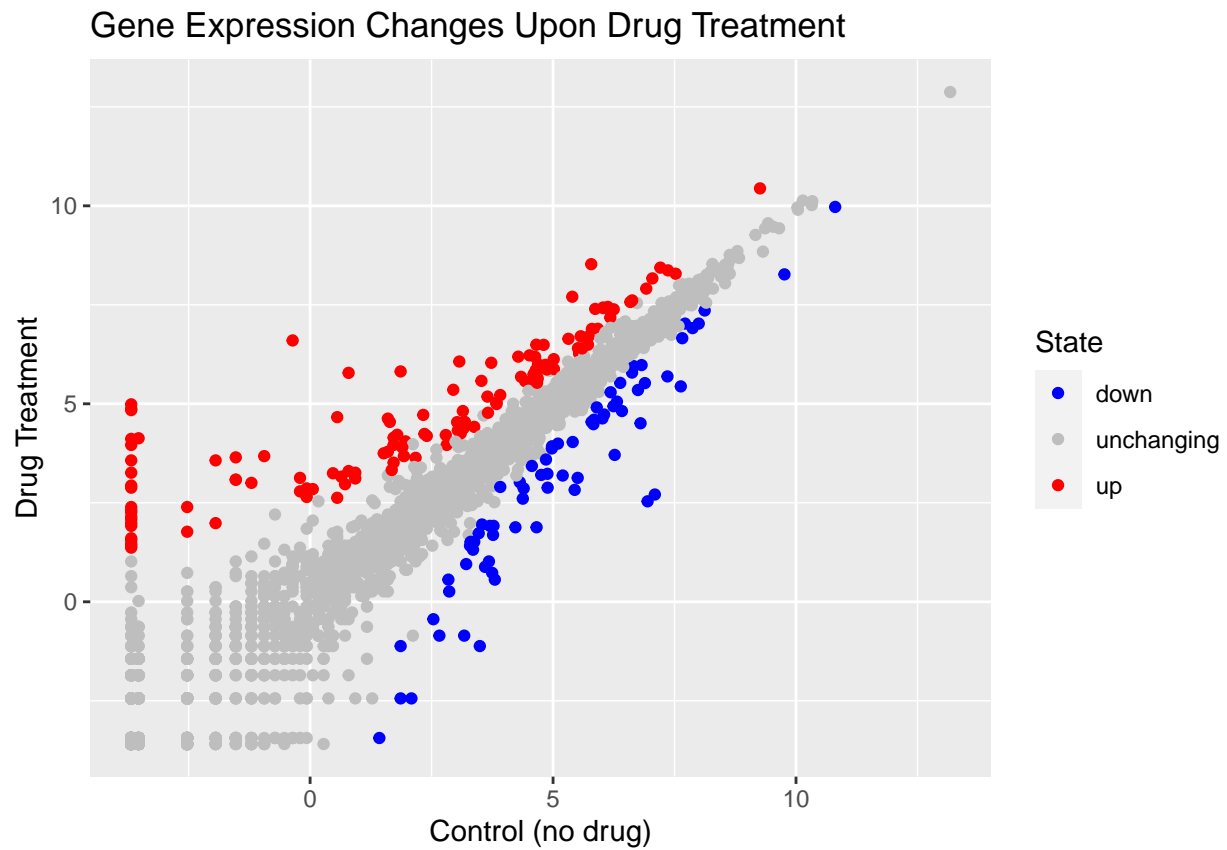
```
# Add some color to denote up/down values
```

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

```
#Add more layers
```

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

```
labs(title="Gene Expression Changes Upon Drug Treatment",
     x="Control (no drug)",
     y="Drug Treatment",
     )
```



```
# install.packages("gapminder")
library(gapminder)
```

```
# install.packages("dplyr")
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

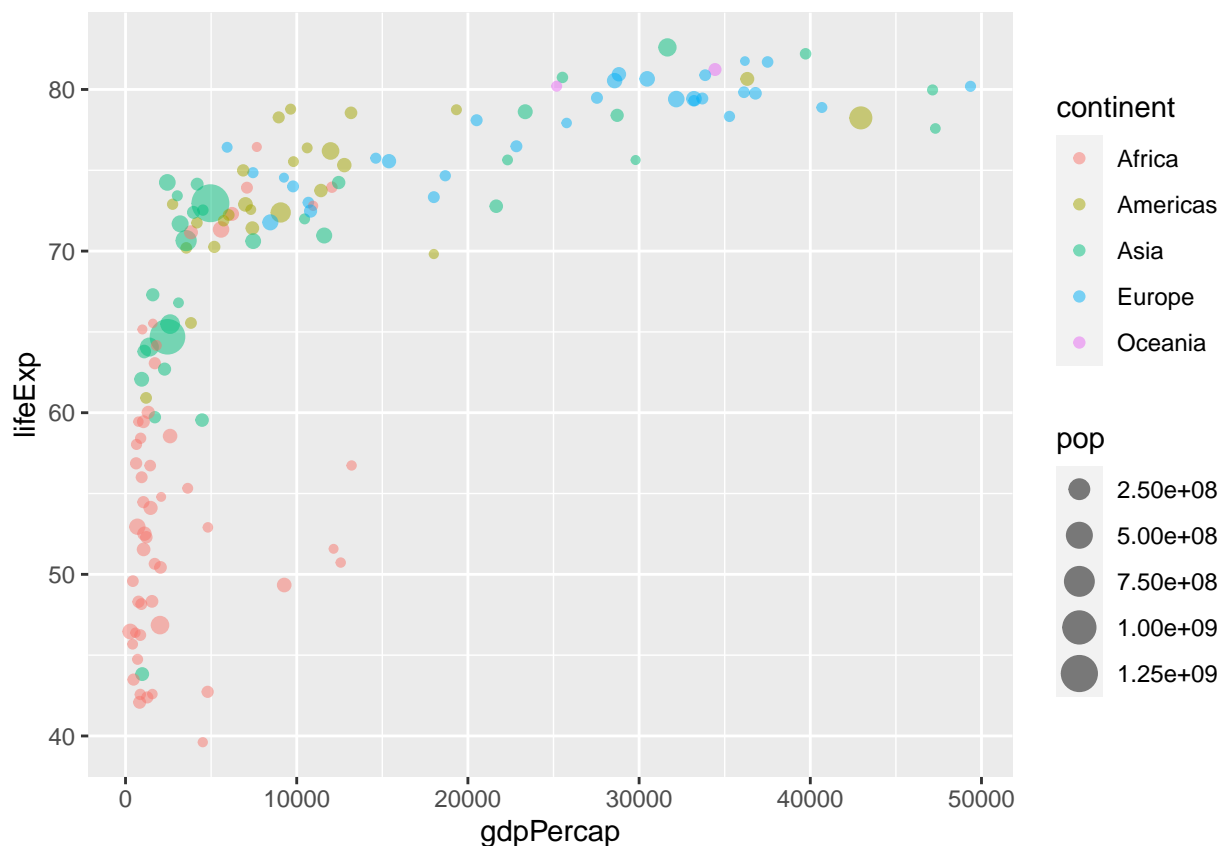
## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

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

```
# Basic scatter plot of gapminder_2007
```

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



```
# Basic bar chart of gapminder_top5
```

```
gapminder_top5 <- gapminder %>%  
  filter(year==2007) %>%  
  arrange(desc(pop)) %>%  
  top_n(5, pop)  
  
ggplot(gapminder_top5) +  
  aes(x=reorder(country, -pop), y=pop, fill=gdpPercap) +  
  geom_col()
```



