Ex1 R - Linear Regression One Variable

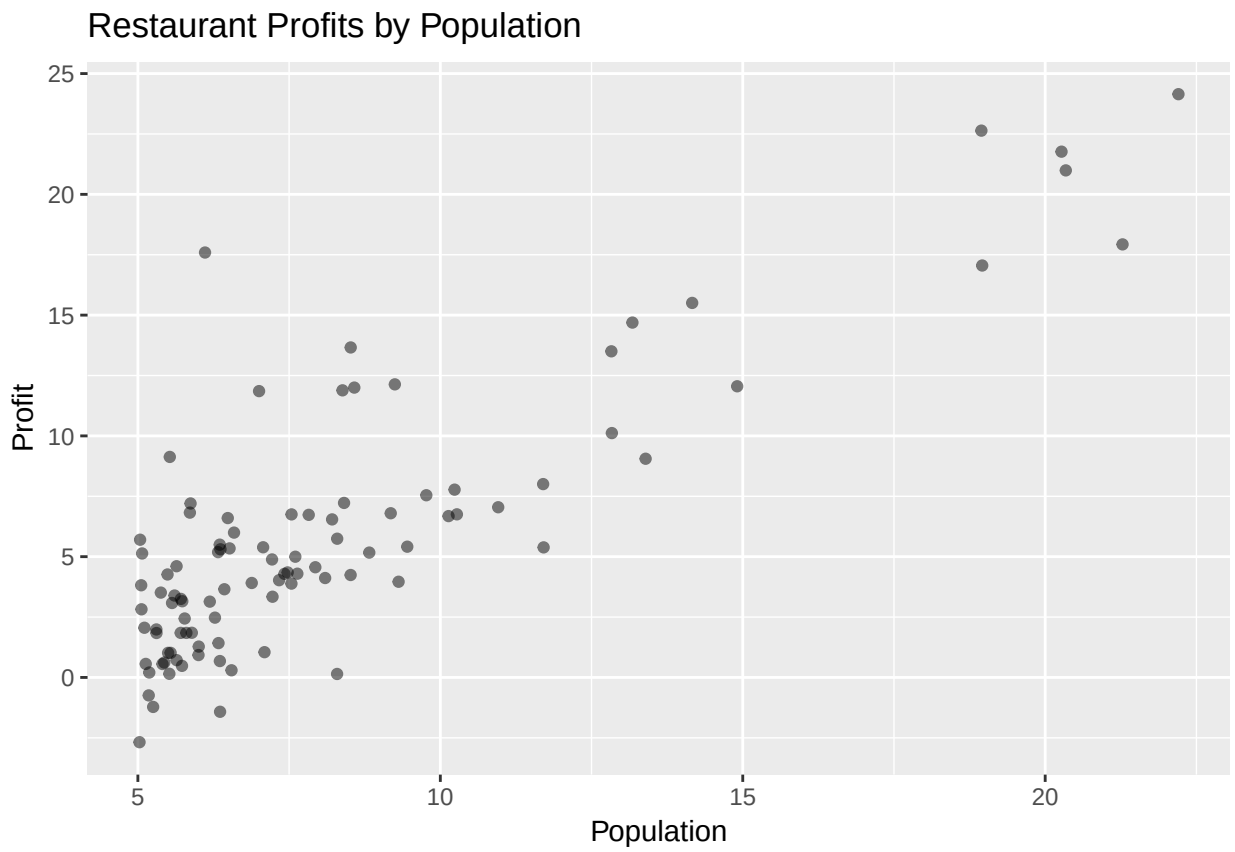
```
suppressPackageStartupMessages({
  library(readr)
  library(ggplot2)
  library(dplyr)
})
```

```
## Warning: package 'readr' was built under R version 3.3.3
```

```
dfRest <- read_csv("~/Projects/Machine Learning/machine-learning-ex1/ex1/ex1data1.txt",
                   col_names = FALSE,
                   col_types = 'dd')

colnames(dfRest) <- c('Population', 'Profit')
```

```
ggplot(dfRest) +
  geom_point(aes(x=Population, y=Profit), alpha=0.5) +
  ggtitle("Restaurant Profits by Population")
```




```
summary(lm(Profit ~ Population, data=dfRest))
```

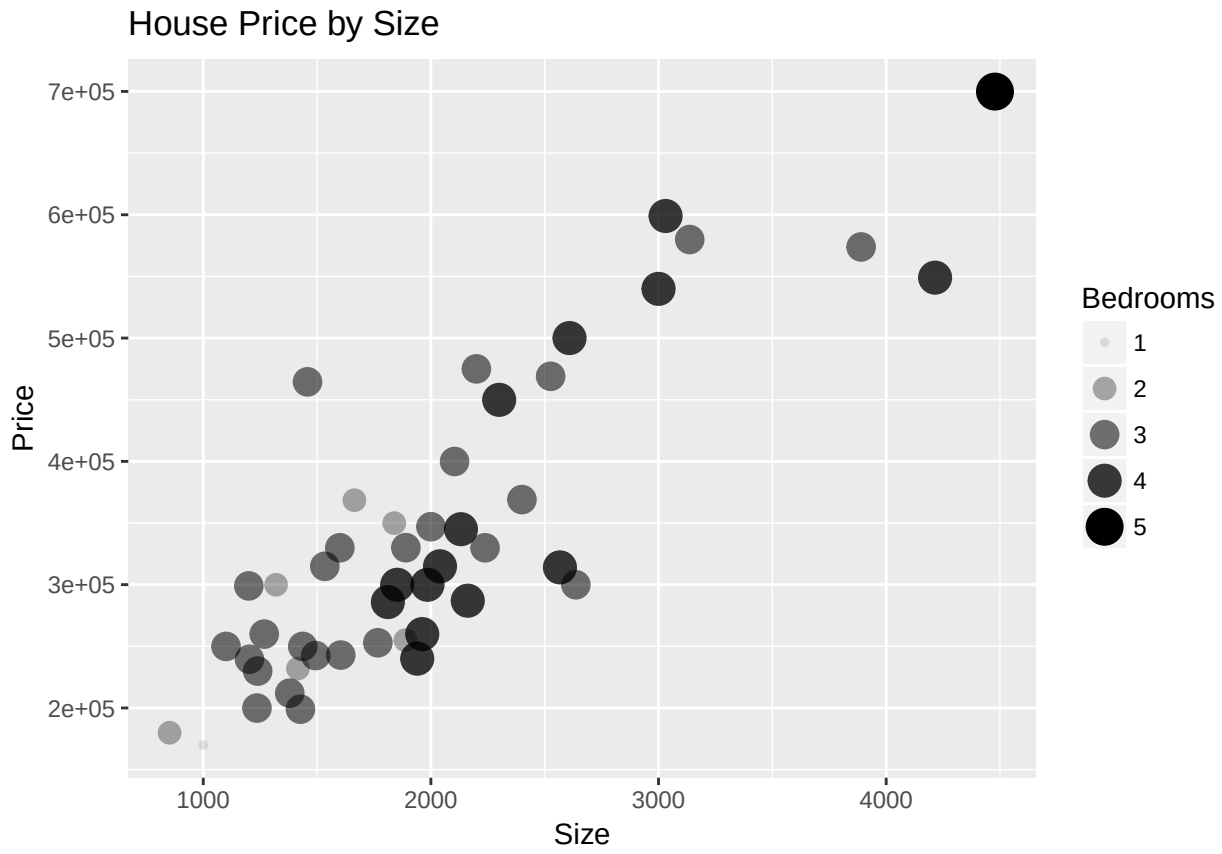
```
##
## Call:
## lm(formula = Profit ~ Population, data = dfRest)
##
```

```
## Residuals:
##     Min      1Q  Median      3Q     Max
## -5.8540 -1.9686 -0.5407  1.5360 14.1982
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.89578    0.71948  -5.415 4.61e-07 ***
## Population   1.19303    0.07974  14.961  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.024 on 95 degrees of freedom
## Multiple R-squared:  0.702,  Adjusted R-squared:  0.6989
## F-statistic: 223.8 on 1 and 95 DF,  p-value: < 2.2e-16
```

```
dfHouse <- read_csv("~/Projects/Machine Learning/machine-learning-ex1/ex1/ex1data2.txt",
                col_names = FALSE,
                col_types = 'ddd')

colnames(dfHouse) <- c('Size', 'Bedrooms', 'Price')
```

```
ggplot(dfHouse) +
  geom_point(aes(x=Size, y=Price,
                 alpha=Bedrooms, size=Bedrooms)) +
  ggtitle("House Price by Size")
```



```r
summary(lm(Price ~ Size + Bedrooms, data=dfHouse))
```

```
##
## Call:
## lm(formula = Price ~ Size + Bedrooms, data = dfHouse)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -130582  -43636  -10829   43698  198147
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  89597.9    41767.4   2.145   0.0375 *
## Size           139.2       14.8   9.409 4.22e-12 ***
## Bedrooms     -8738.0    15450.7  -0.566   0.5746
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 66070 on 44 degrees of freedom
## Multiple R-squared:  0.7329, Adjusted R-squared:  0.7208
## F-statistic: 60.38 on 2 and 44 DF,  p-value: 2.428e-13
```