SimpleRegression

```
dataset = read.csv(file='Salary_Data.csv')
dataset
```

```
##    YearsExperience Salary
## 1              1.1  39343
## 2              1.3  46205
## 3              1.5  37731
## 4              2.0  43525
## 5              2.2  39891
## 6              2.9  56642
## 7              3.0  60150
## 8              3.2  54445
## 9              3.2  64445
## 10             3.7  57189
## 11             3.9  63218
## 12             4.0  55794
## 13             4.0  56957
## 14             4.1  57081
## 15             4.5  61111
## 16             4.9  67938
## 17             5.1  66029
## 18             5.3  83088
## 19             5.9  81363
## 20             6.0  93940
## 21             6.8  91738
## 22             7.1  98273
## 23             7.9 101302
## 24             8.2 113812
## 25             8.7 109431
## 26             9.0 105582
## 27             9.5 116969
## 28             9.6 112635
## 29            10.3 122391
## 30            10.5 121872
```

```
# Splitting the dataset into the Training set and Test set
# install.packages('caTools')
library(caTools)
```

```
## Warning: package 'caTools' was built under R version 3.6.1
```

```
set.seed(123)
split = sample.split(dataset$Salary, SplitRatio = 2/3)
training_set = subset(dataset, split == TRUE)
test_set = subset(dataset, split == FALSE)
```

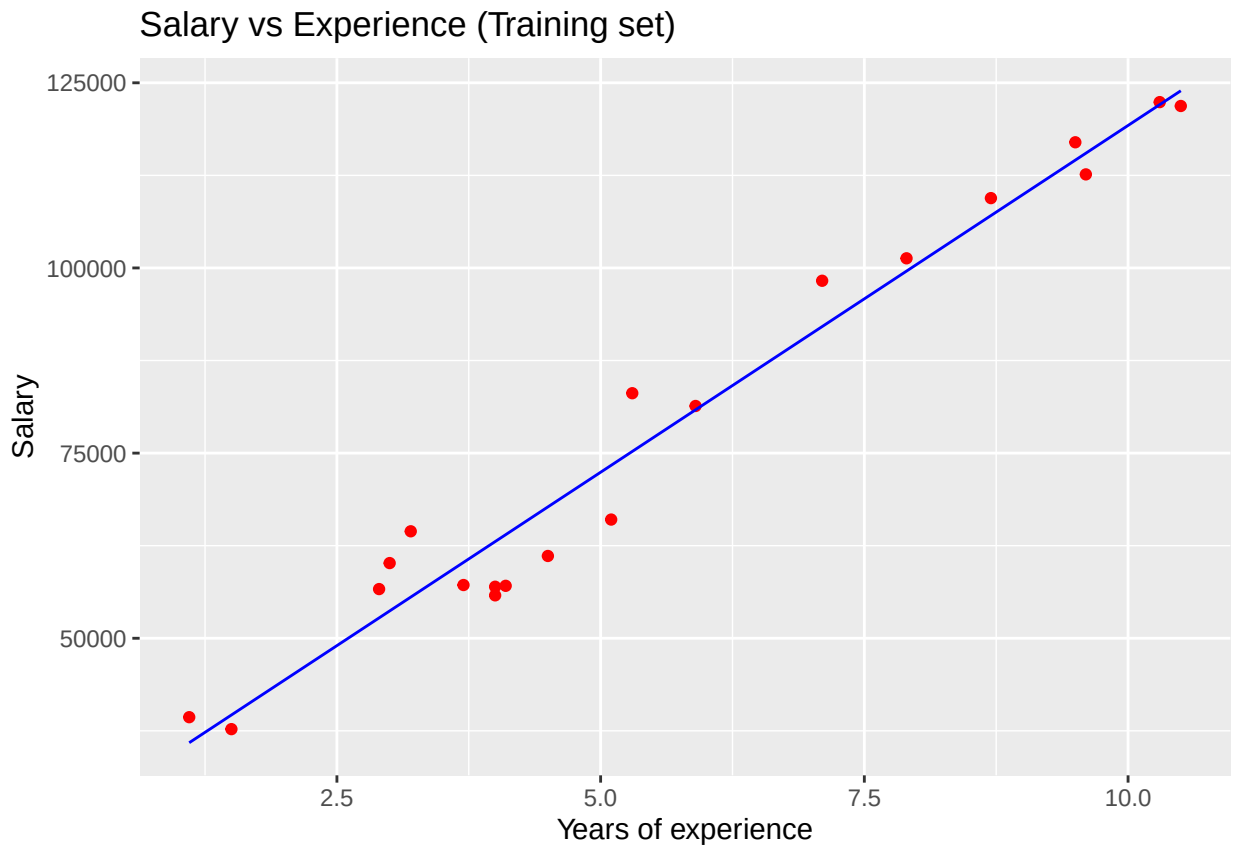
```
# Fitting Simple Linear Regression to the Training set
regressor = lm(formula = Salary ~ YearsExperience,
               data = training_set)
```

```
# Predicting the Test set results
y_pred = predict(regressor, newdata = test_set)
```

```
# Visualising the Training set results
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.6.1
```

```
ggplot() +
  geom_point(aes(x = training_set$YearsExperience, y = training_set$Salary),
             colour = 'red') +
  geom_line(aes(x = training_set$YearsExperience, y = predict(regressor, newdata = training_set)),
            colour = 'blue') +
  ggtitle('Salary vs Experience (Training set)') +
  xlab('Years of experience') +
  ylab('Salary')
```




```
# Visualising the Test set results
library(ggplot2)
ggplot() +
  geom_point(aes(x = test_set$YearsExperience, y = test_set$Salary),
             colour = 'red') +
  geom_line(aes(x = training_set$YearsExperience, y = predict(regressor, newdata = training_set)),
            colour = 'blue') +
```

```
  ggtitle('Salary vs Experience (Test set)') +
  xlab('Years of experience') +
  ylab('Salary')
```