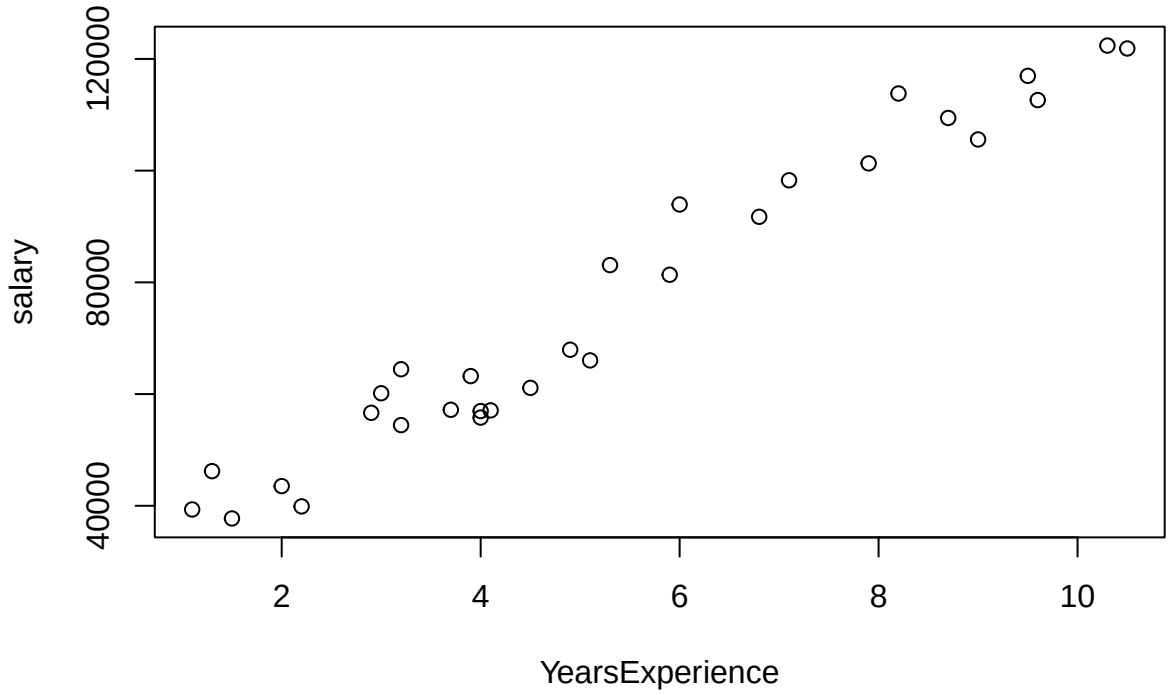
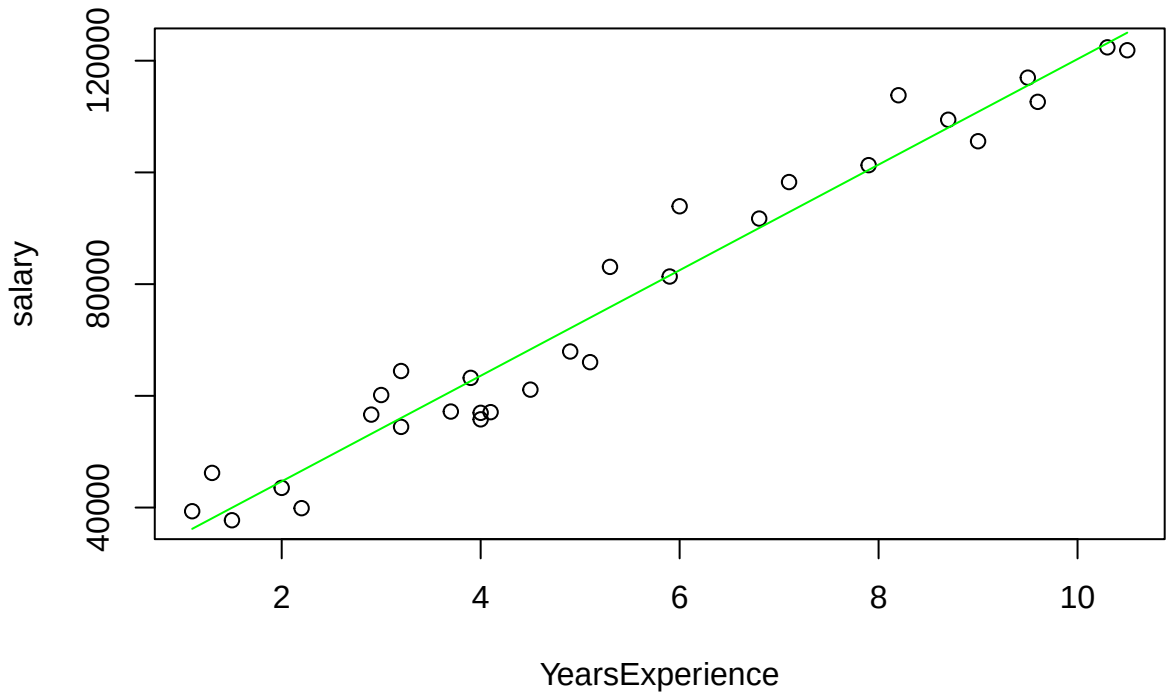
simple linear regression football_salaries

```
data<-read.csv("Salary_Data.csv")
plot(data$YearsExperience,data$Salary,xlab = "YearsExperience",ylab="salary")
```




```
modele=lm(Salary ~YearsExperience,data = data)
coeff=modele$coefficients
plot(data$YearsExperience,data$Salary,xlab = "YearsExperience",ylab="salary")
lines(data$YearsExperience,data$YearsExperience *coeff[2]+coeff[1],col="green")
```




```
summary(modele)
```

```
##
## Call:
## lm(formula = Salary ~ YearsExperience, data = data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -7958.0 -4088.5  -459.9  3372.6 11448.0
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      25792.2     2273.1   11.35 5.51e-12 ***
## YearsExperience   9450.0      378.8   24.95  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5788 on 28 degrees of freedom
## Multiple R-squared:  0.957,  Adjusted R-squared:  0.9554
## F-statistic: 622.5 on 1 and 28 DF,  p-value: < 2.2e-16
```

```
print(" On rejette l'hypothèse que l'un des coeff est nul ")
```

```
## [1] " On rejette l'hypothèse que l'un des coeff est nul "
```

```
x = seq(0,20,0.01)
intC = predict(modele, newdata=list(YearsExperience=x), interval="confidence")
intP = predict(modele, newdata=list(YearsExperience=x), interval="prediction")
plot(data$YearsExperience,data$Salary,xlab = "YearsExperience",ylab="salary")
abline(coeff[1], coeff[2], col="blue")
lines(x, intC[, 2], col="red")
lines(x, intC[, 3], col="red")
lines(x, intP[, 2], col="green")
```

```
lines(x, intP[, 3], col="green")
legend(2, 14, legend=c("Régression", "Intervalle de confiance", "Intervalle de prédiction")
       ,col=c("blue", "red", "green"), lty=1)
```