#Load data

data1<-read.csv("Salary\_Data.csv")

# Visualization

install.packages("lattice")

library(lattice)

colnames(data1)

dotplot(data1$Salary, main="salary",col="dodgerblue4")

dotplot(data1$YearsExperience, main="Years Experience", col="dodgerblue4")

boxplot(data1$Salary,col="dodgerblue4")

boxplot(data1$YearsExperience,col="dodgerblue4")

#Regression equation

#Syntax model<-lm(y~x,data=data set name)

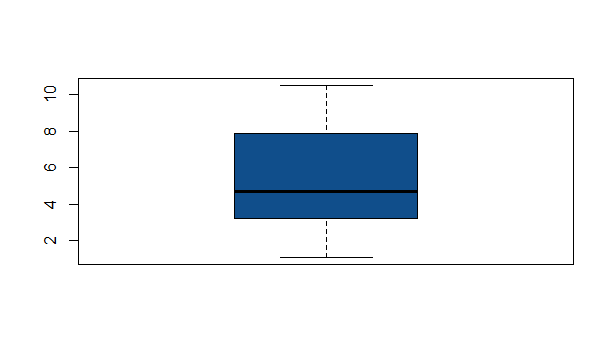
#column names

colnames(data1)

#Model building

model<- lm(Salary~YearsExperience,data =data1)

summary(model)



Right skewed…

Call:

lm(formula = Salary ~ YearsExperience, data = data1)

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

Conclusion:

Years of Experience p value is < 0.5 and R^2 value is 0.957. the model is found to be good.