1.giaNha <- read.csv("C:\\Users\\Thanh Vo\\Desktop\\BaiTapLon\\gia\_nha.csv", header = TRUE)

2a. new\_DF <- giaNha[, c(6, 9, 11, 14, 16, 23)]

b. new\_DF <- na.omit(new\_DF)

3

a.

new\_DF$price <- log(new\_DF$price)

new\_DF$sqft\_living15 <- log(new\_DF$sqft\_living15)

new\_DF$sqft\_above <- log(new\_DF$sqft\_above)

new\_DF$sqft\_living <- log(new\_DF$sqft\_living)

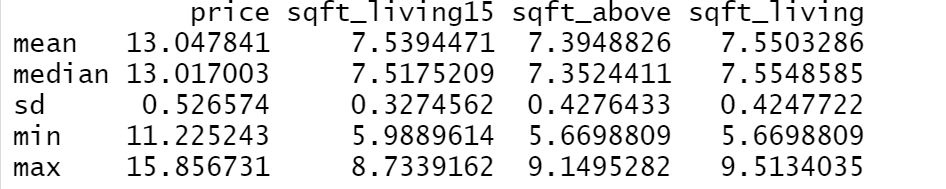
b.

new\_DF$price <- log(new\_DF$price)

new\_DF$sqft\_living15 <- log(new\_DF$sqft\_living15)

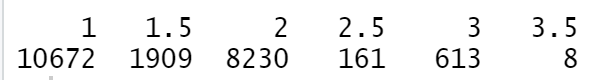
new\_DF$sqft\_above <- log(new\_DF$sqft\_above)

new\_DF$sqft\_living <- log(new\_DF$sqft\_living)

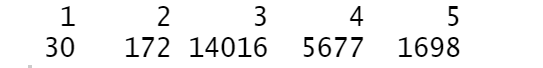


c.

table(new\_DF$floors)

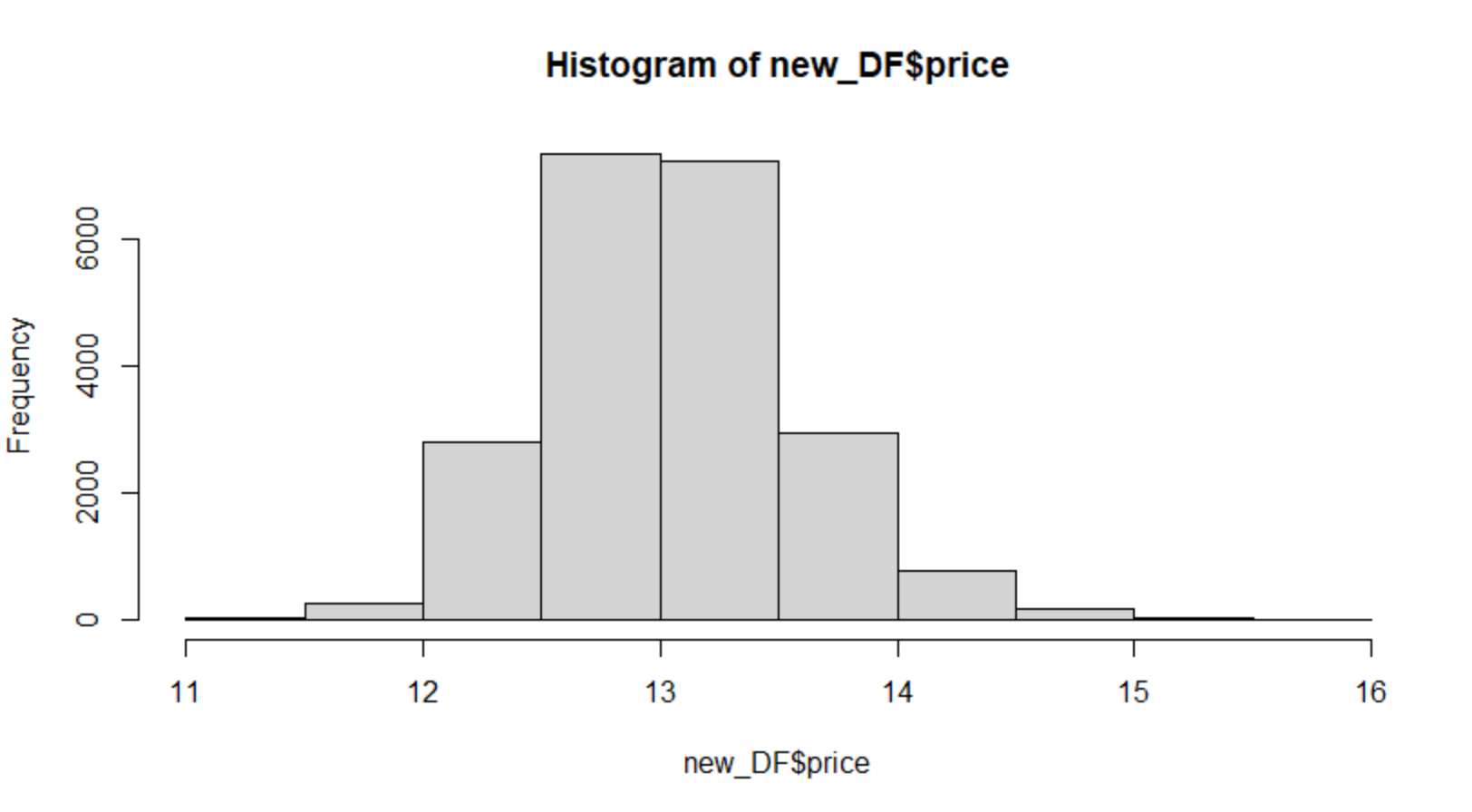


table(new\_DF$condition)



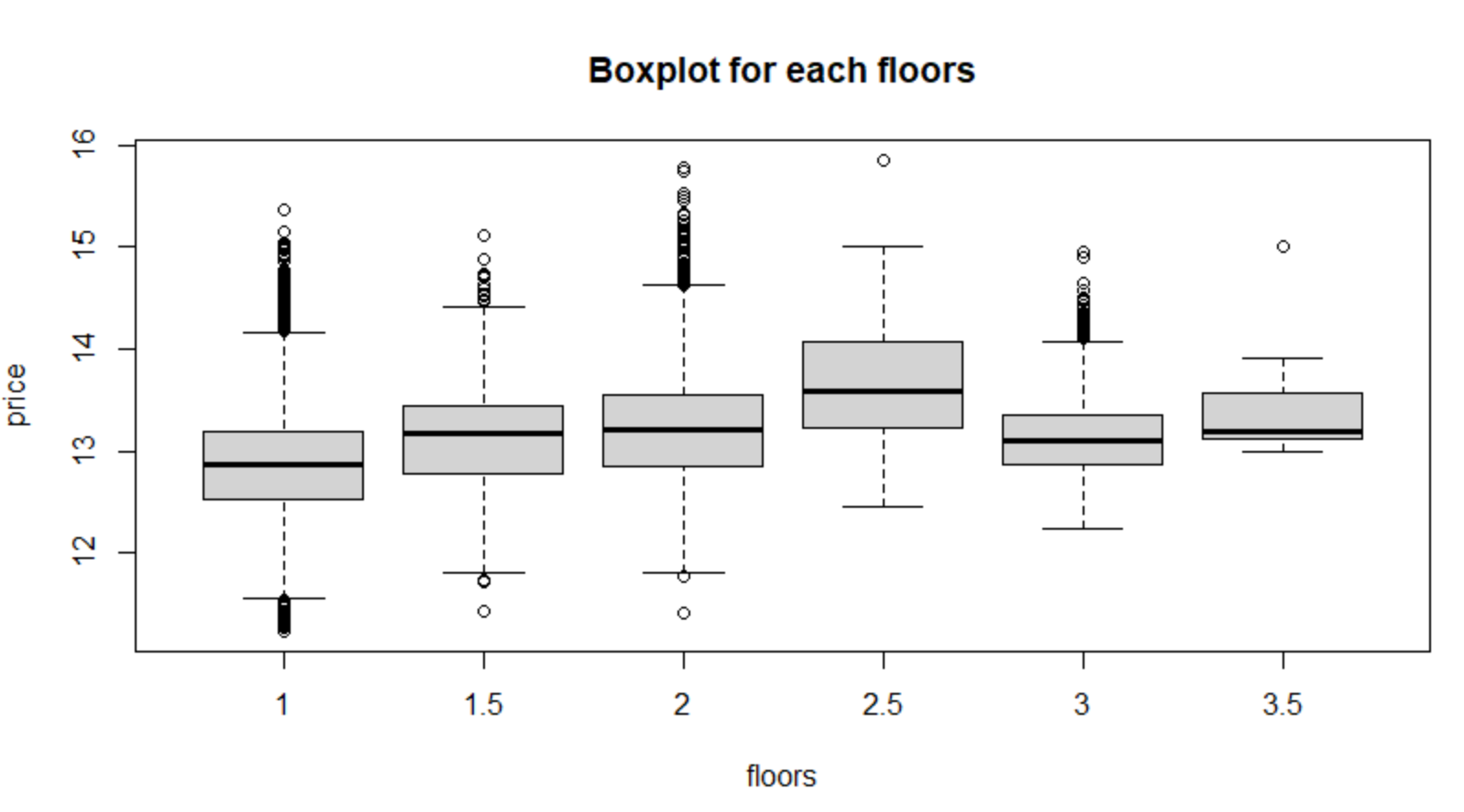
d.

hist(new\_DF$price)

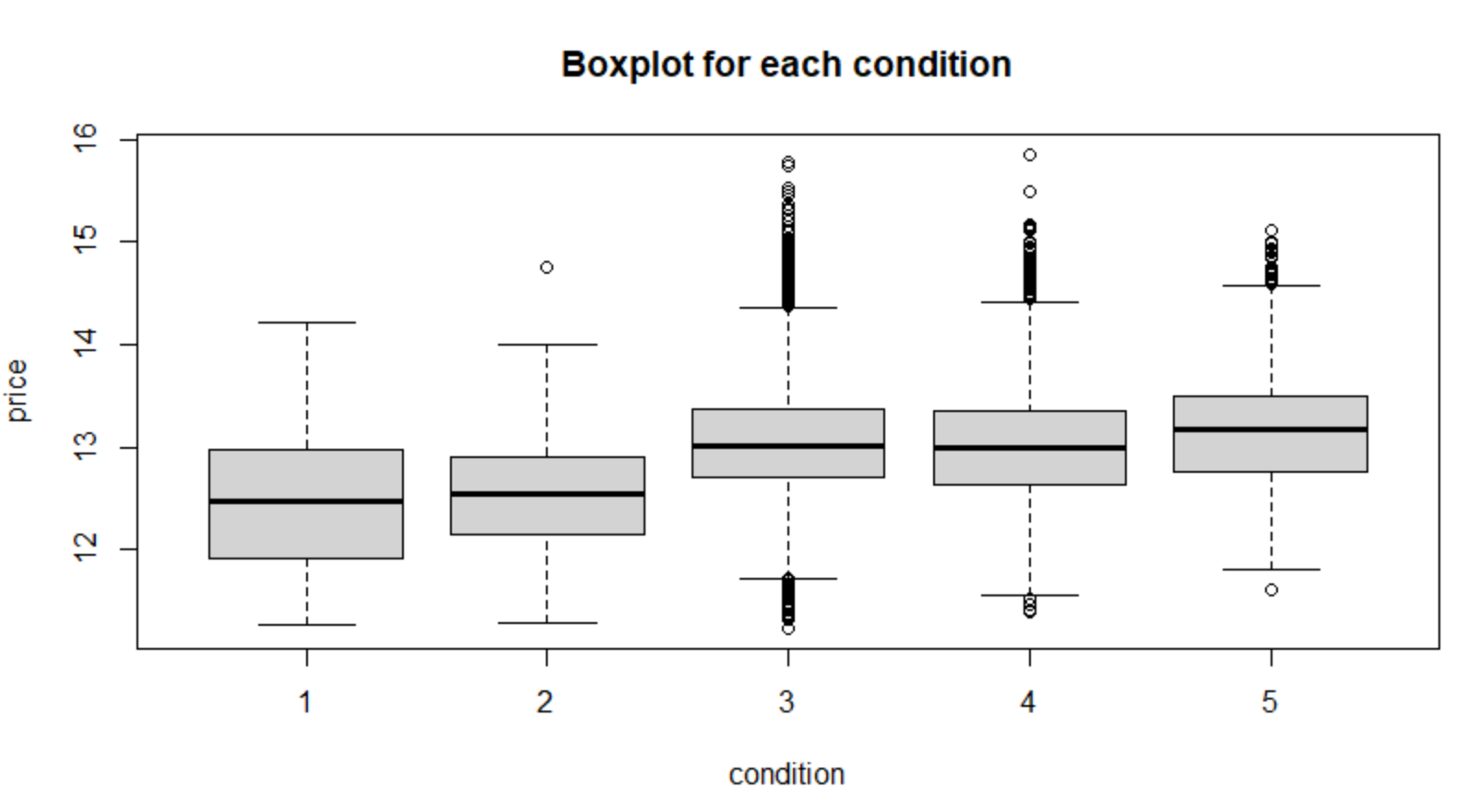


e.

boxplot(price~floors, data = new\_DF, main="Boxplot for each floors")

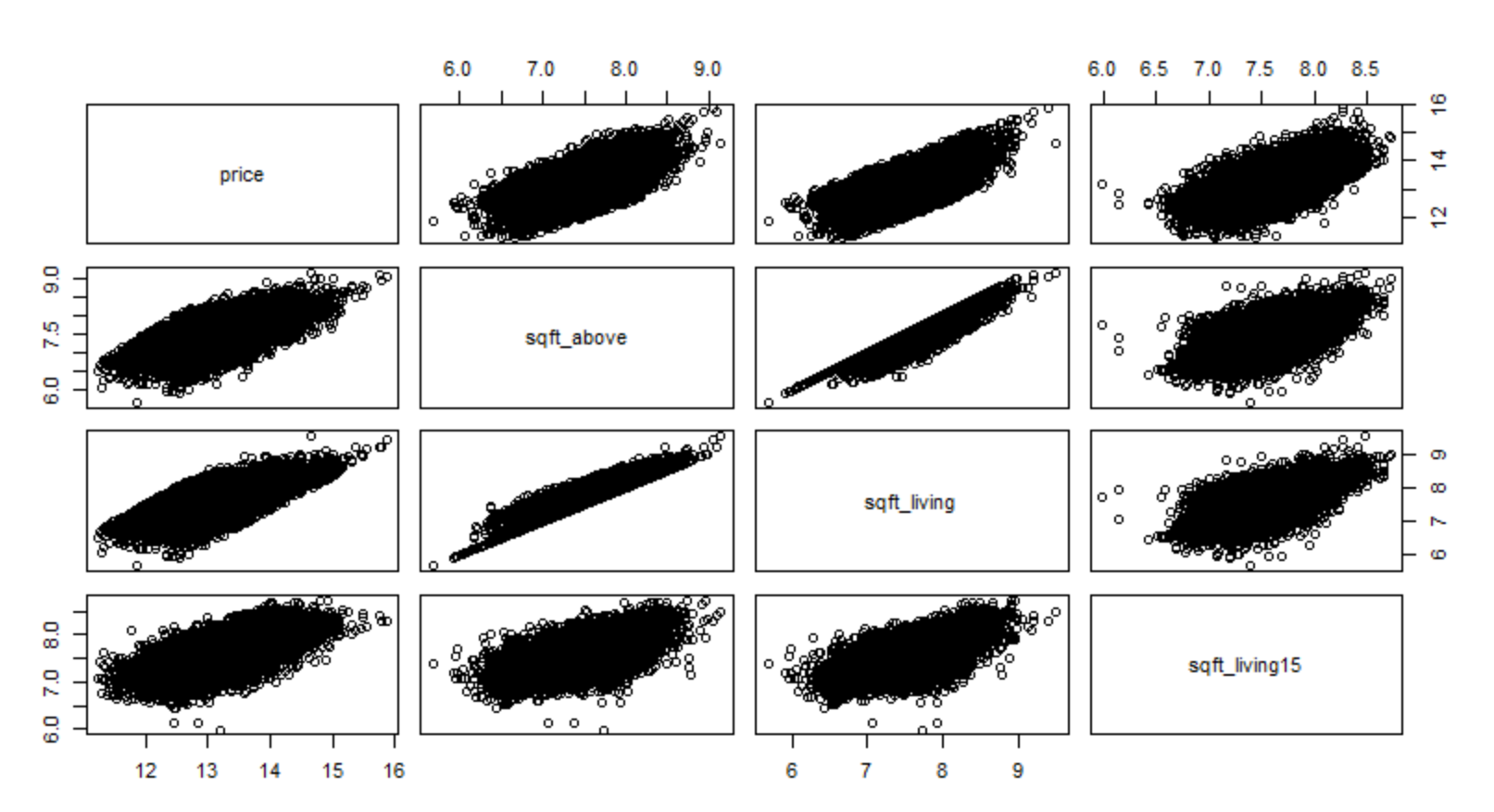


boxplot(price~condition, data = new\_DF, main="Boxplot for each condtion")



f.

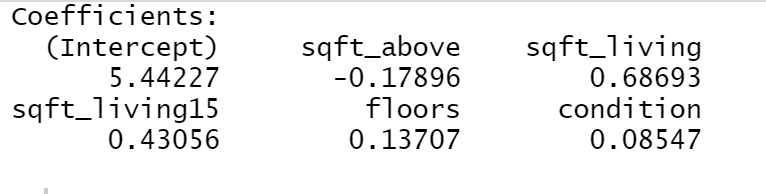
pairs(price~sqft\_above+sqft\_living+sqft\_living15, data = new\_DF)



4a.

mulfit <- lm(price~sqft\_above+sqft\_living+sqft\_living15+floors+condition , data = new\_DF)

mulfit



Mô hình:

price = 5.44227 – 0.17896(sqft\_above) + 0.68693(sqft\_living) + 0.43056(sqft\_living15) + 0.13707(floors) + 0.08547(sqft\_living)