**ASSIGNMENT No.3**

**NAME** : Bhavik Ransubhe

**CLASS** : TE(B) COMP

**ROLL NO**: 39055

**PROBLEM STATEMENT:**

Download Abalone dataset. (URL: http://archive.ics.uci.edu/ml/datasets/Abalone)

Data set has total 8 Number of Attributes.

Sex nominal M, F, and I (infant)

Length continuous mm Longest shell measurement

Diameter continuous mm perpendicular to length

Height continuous mm with meat in shell

Whole weight continuous grams whole abalone

Shucked weight continuous grams weight of meat

Viscera weight continuous grams gut weight (after bleeding)

Shell weight continuous grams after being dried

Rings (age/class of abalone)

Load the data from data file and split it into training and test datasets. Summarize the

properties in the training dataset. The number of rings is the value to predict: either as a

continuous value or as a classification problem.

Predict the age of abalone from physical measurements using linear regression or predict

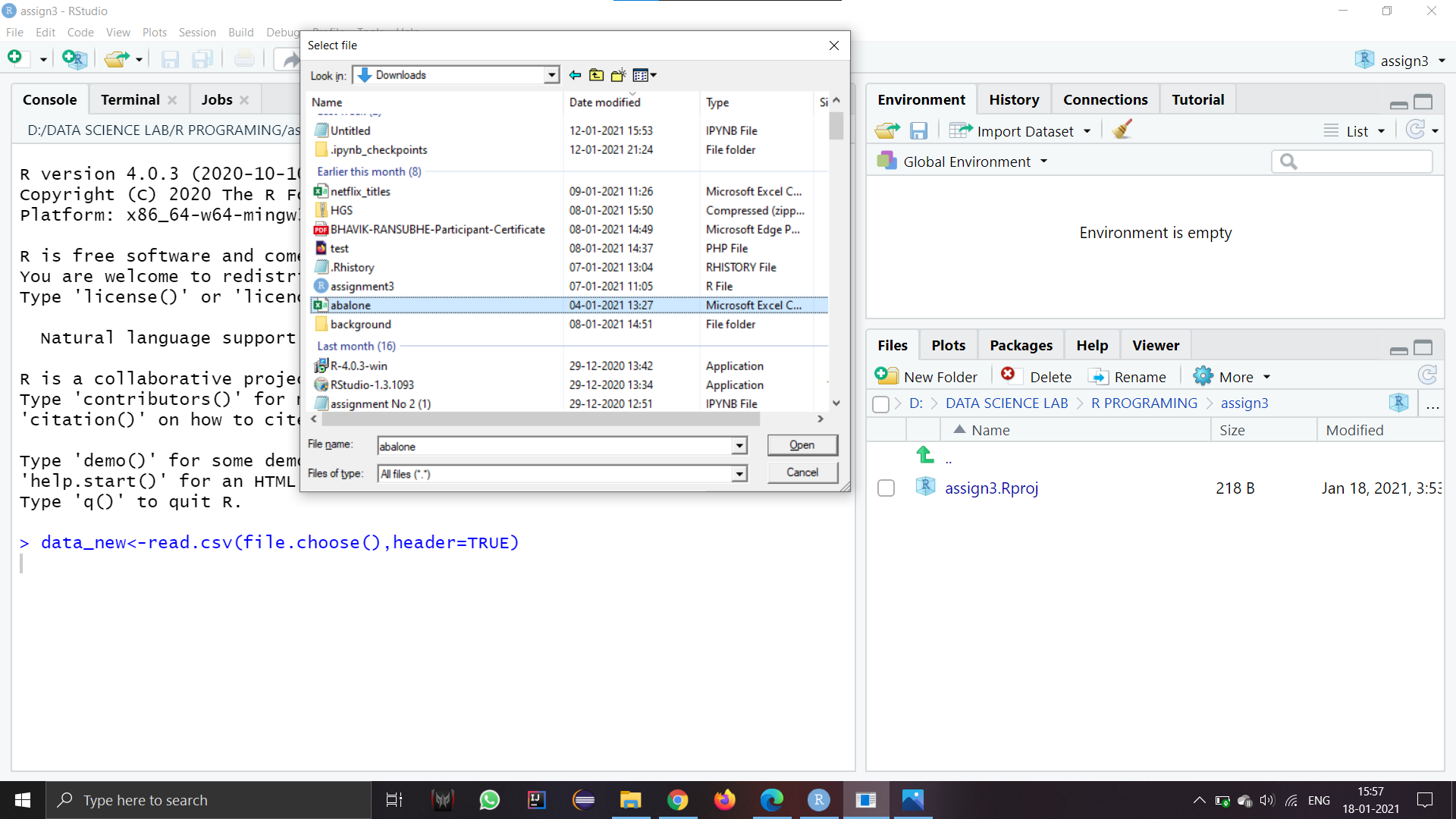
ring class as classification problem

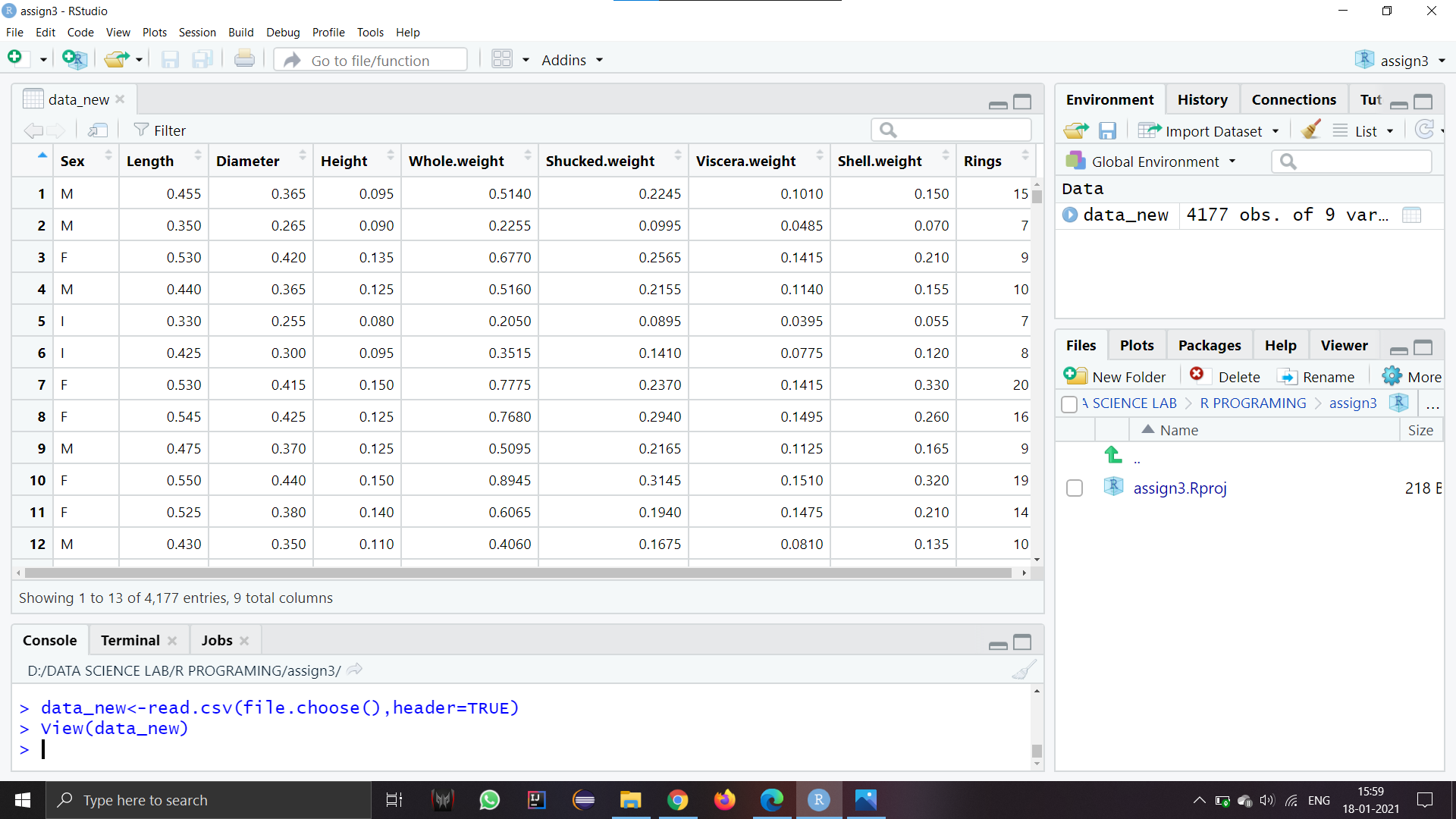
**IMPLEMENTATION /EXECUTATION:**

**1) Read the data using csv file:**

>data\_new<-read.csv(file.choose(),header=TRUE)

>View(data\_new)

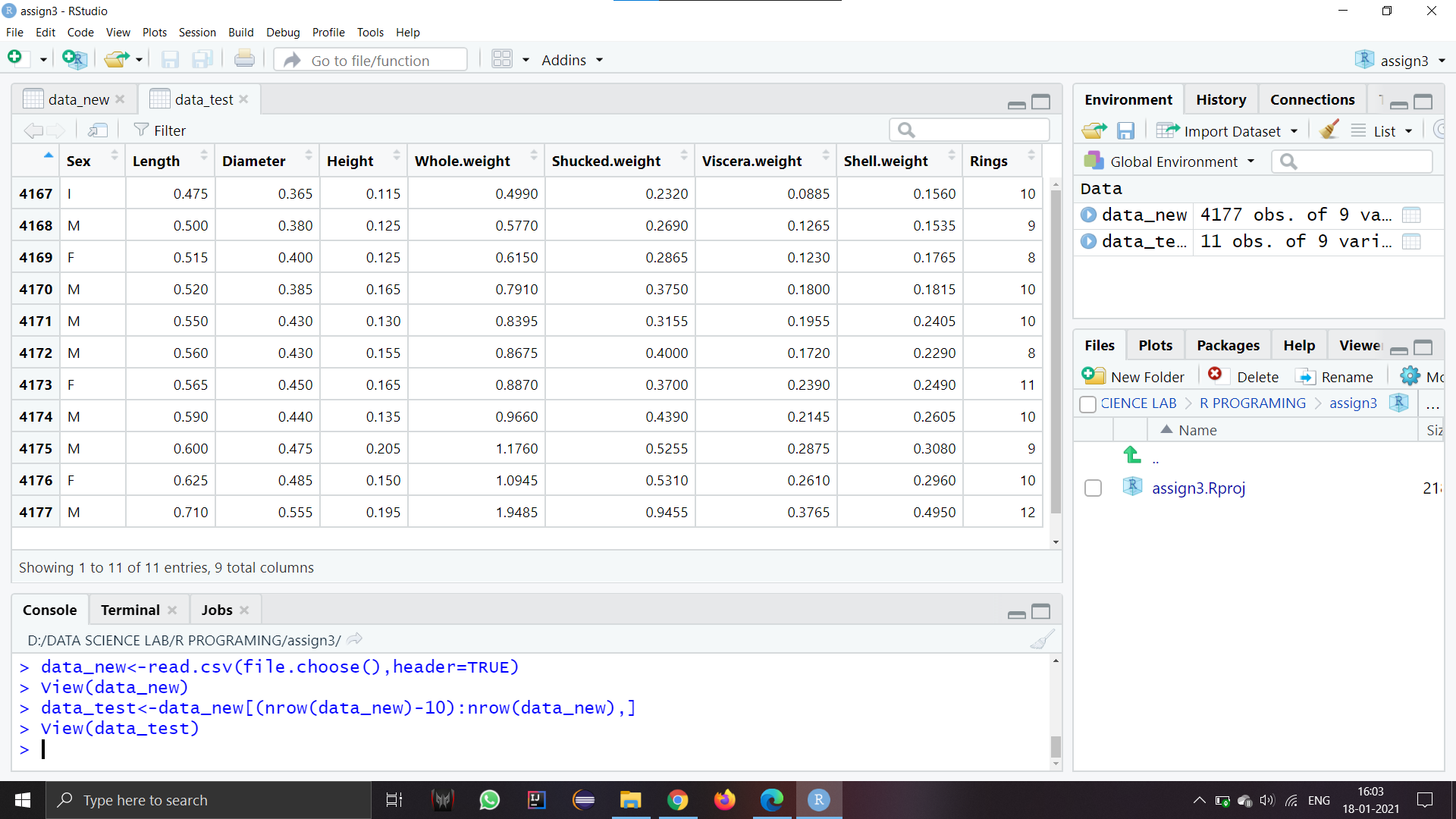




**2)Divide the dataset into train and test :**

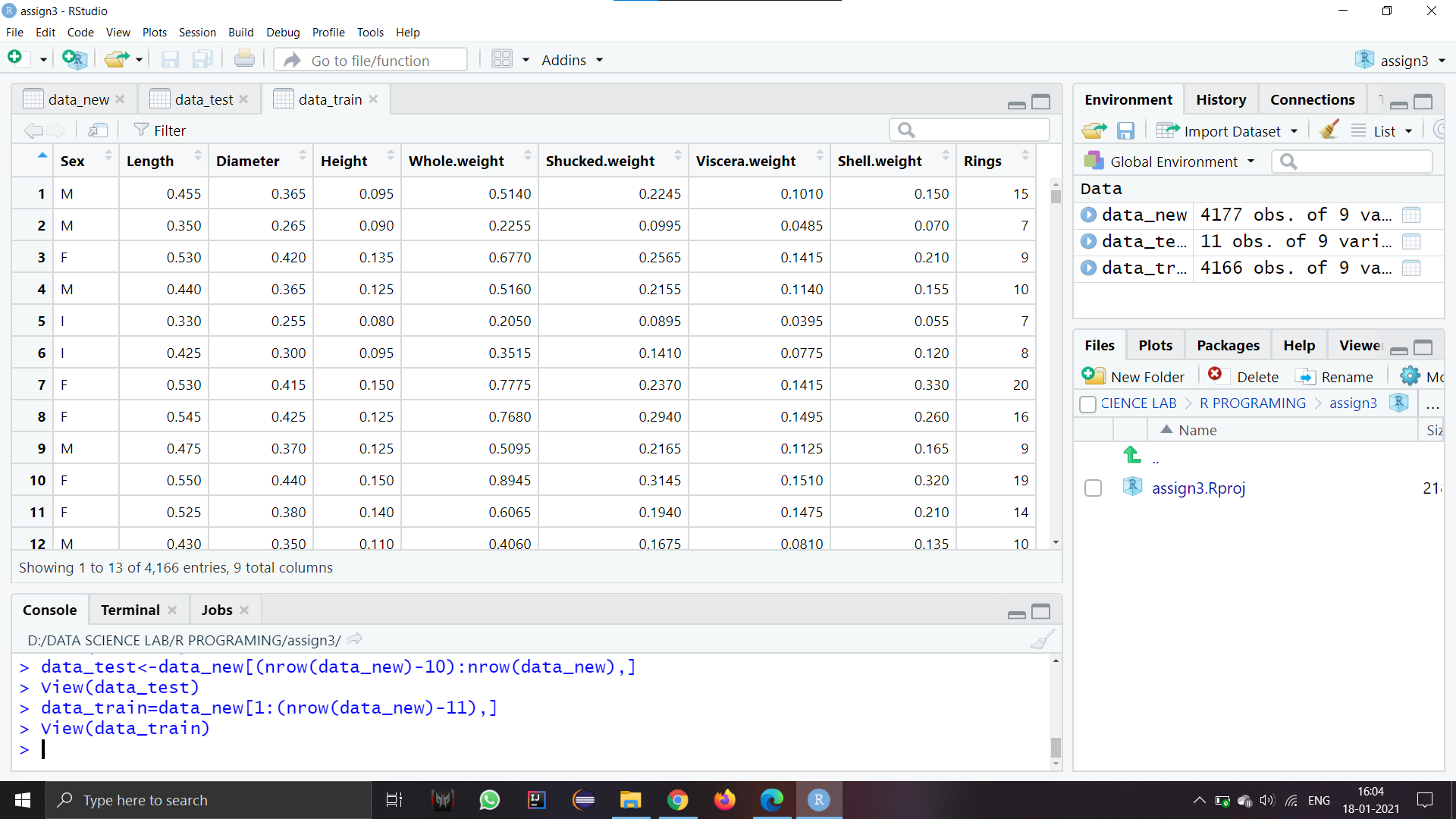
> data\_test<-data\_new[(nrow(data\_new)-10):nrow(data\_new),]

> View(data\_test)



> data\_train=data\_new[1:(nrow(data\_new)-11),]

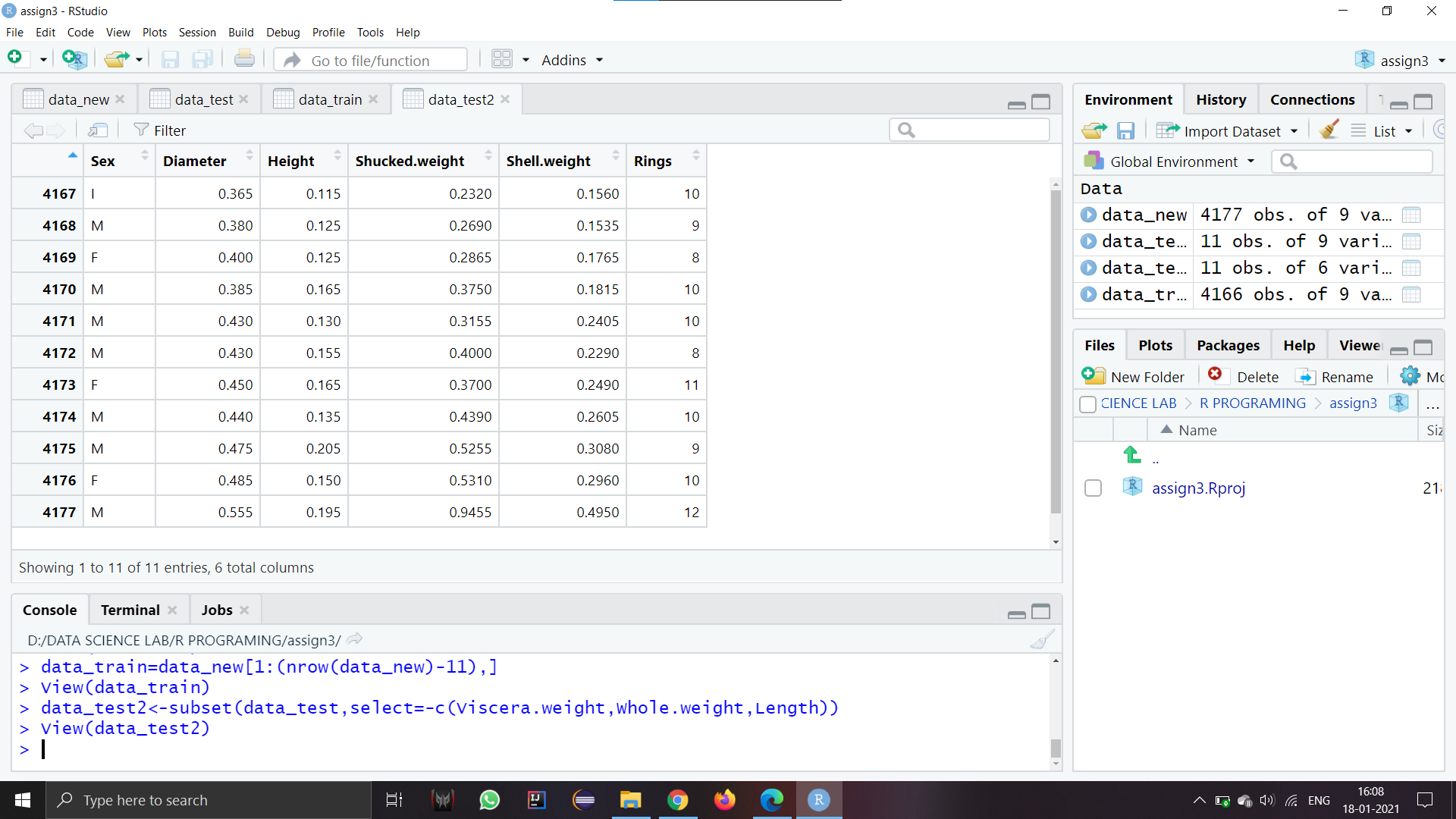
> View(data\_train)



**3)Remove attributes from dataset:**

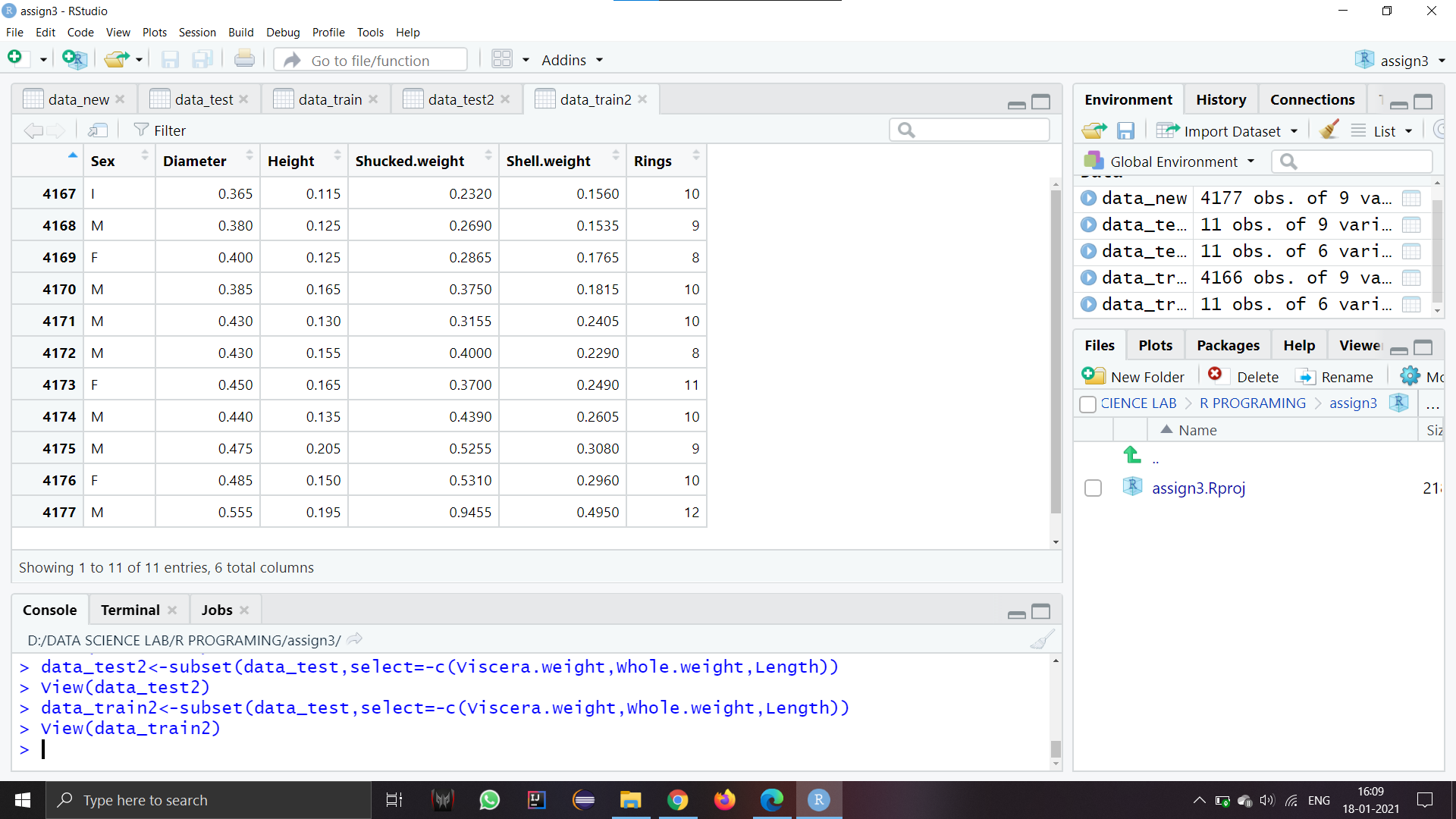
> data\_test2<-subset(data\_test,select=-c(Viscera.weight,Whole.weight,Length))

> View(data\_test2)



> data\_train2<-subset(data\_test,select=-c(Viscera.weight,Whole.weight,Length))

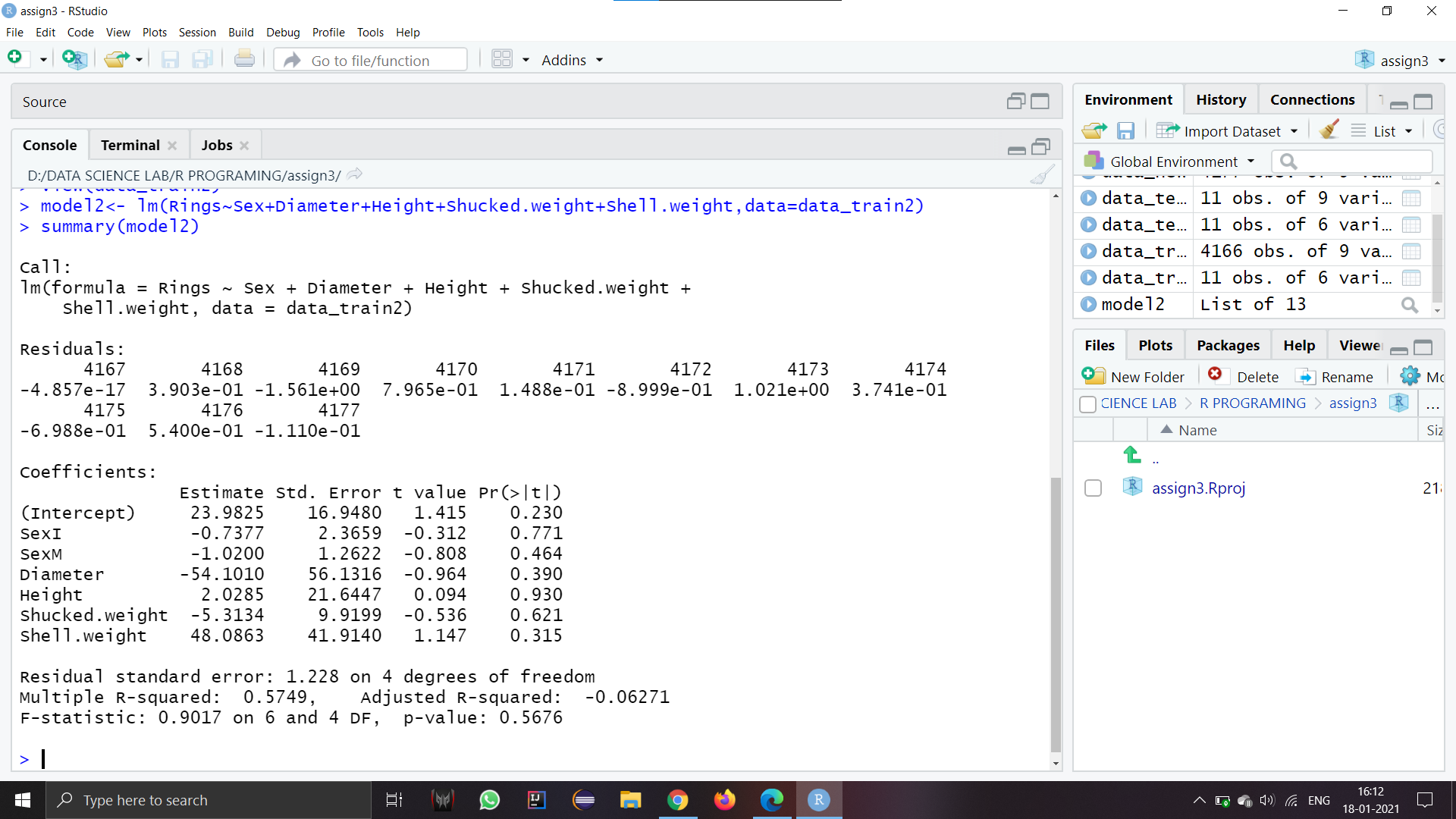
> View(data\_train2)



**4)Training the model for linear regression:**

> model2<- lm(Rings~Sex+Diameter+Height+Shucked.weight+Shell.weight,data=data\_train2)

> summary(model2)



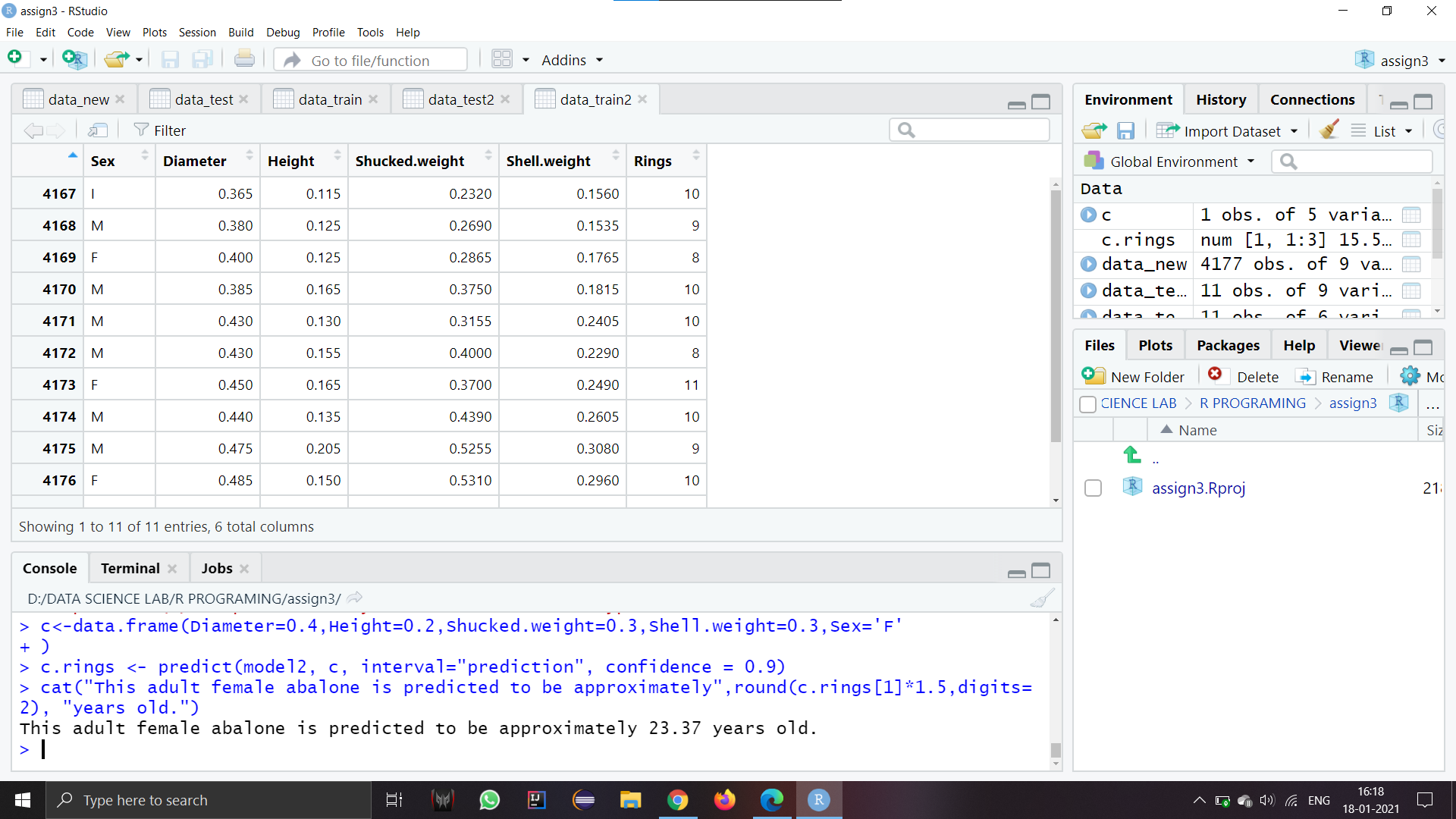
**5)Test for new data:**

c<-data.frame(Diameter=0.4,Height=0.2,Shucked.weight=0.3,Shell.weight=0.3,Sex='F')

**6)Predict using the model:**

> c.rings <- predict(model2, c, interval="prediction", confidence = 0.9)

> cat("This adult female abalone is predicted to be approximately", round(c.rings[1]\*1.5,digits=2), "years old.")



\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*