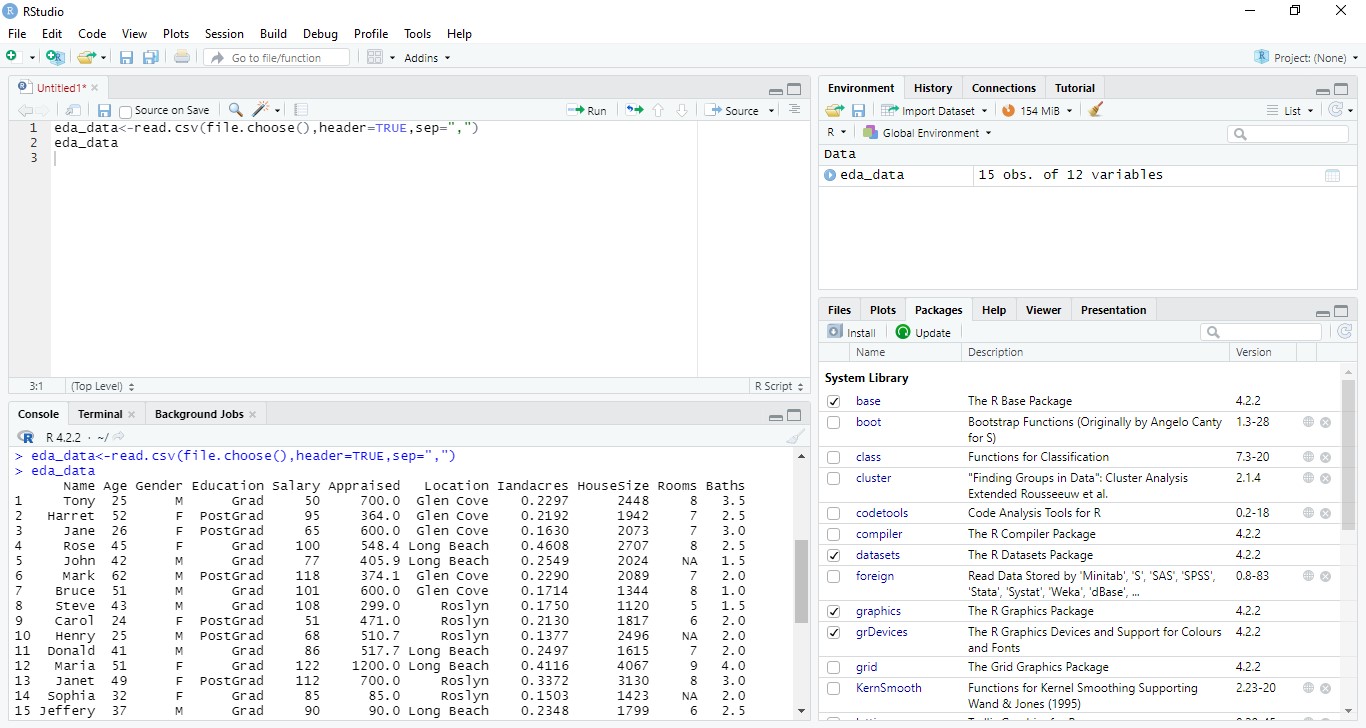
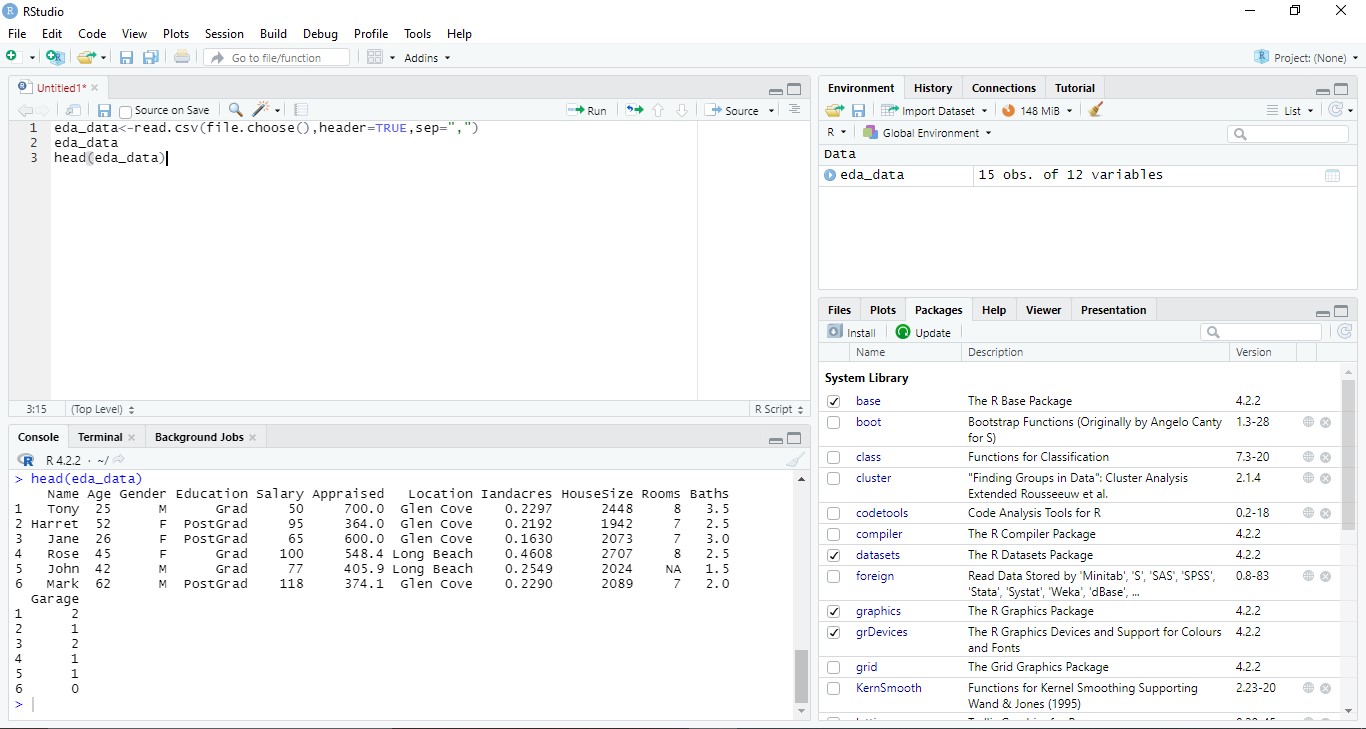
**Practical No. 01**

**AIM:** Exploratory Data Analysis(EDA) and Data Visualization

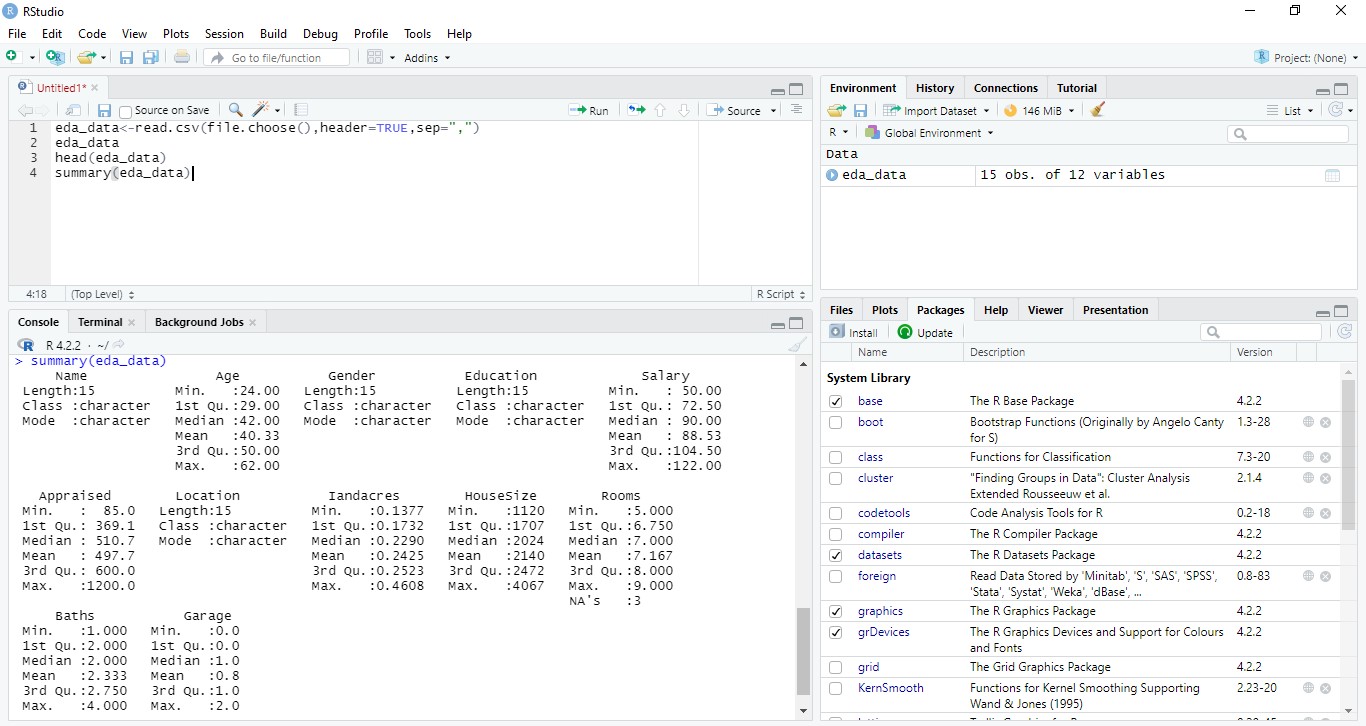
1. eda\_data<-read.csv(file.choose(),header=TRUE,sep=",") eda\_data



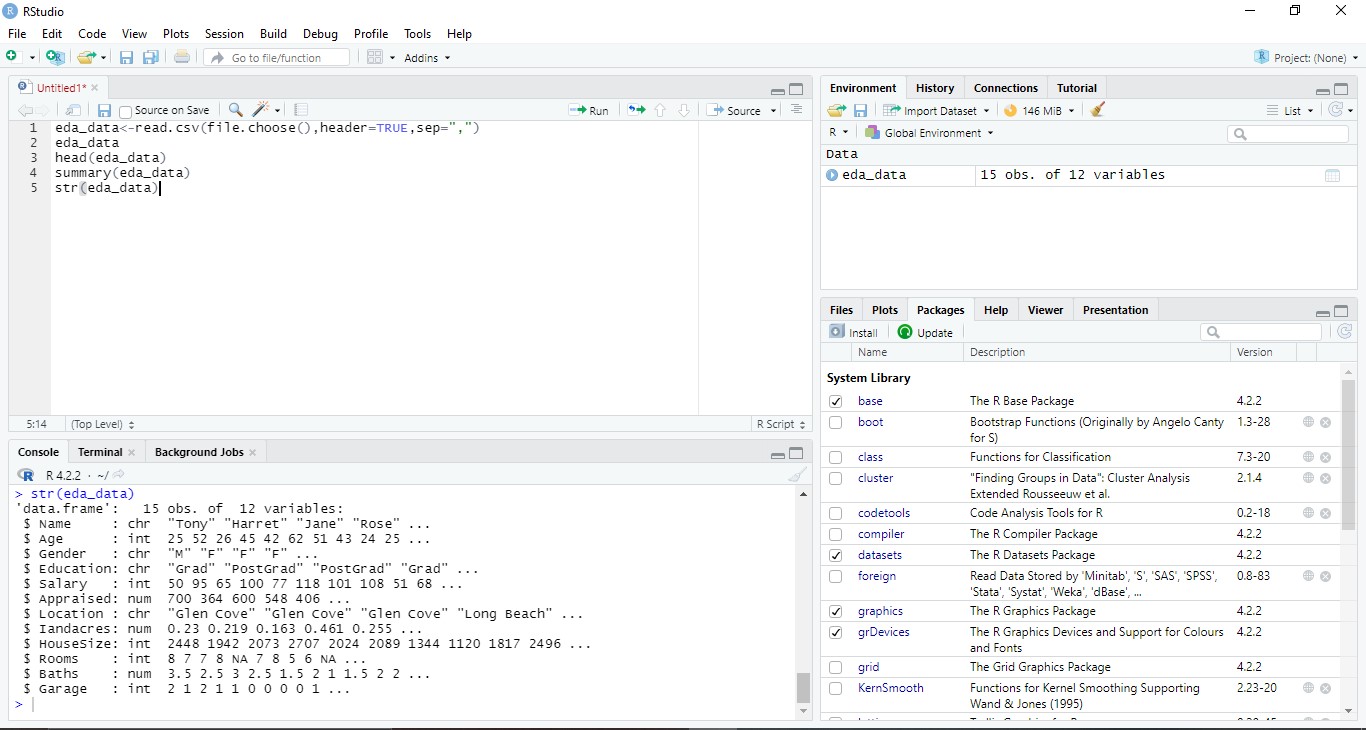
1. head(eda\_data)



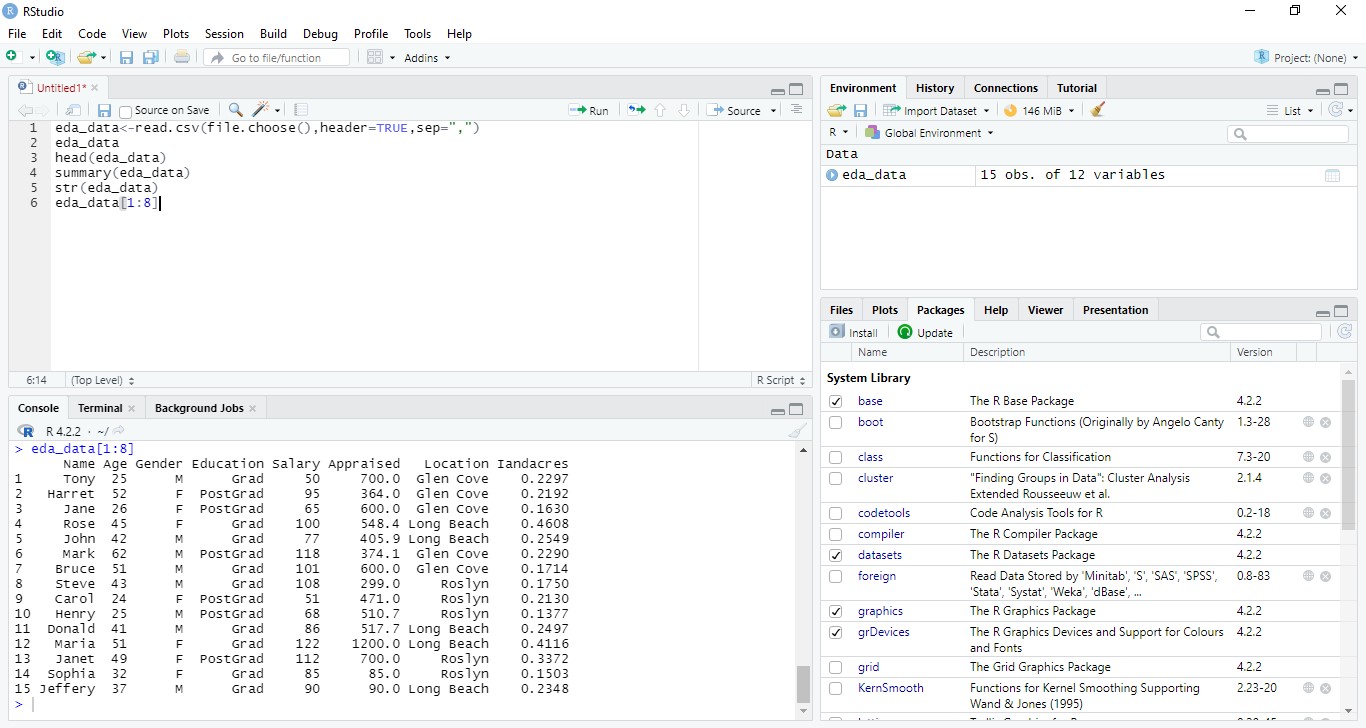
3)summary(eda\_data)



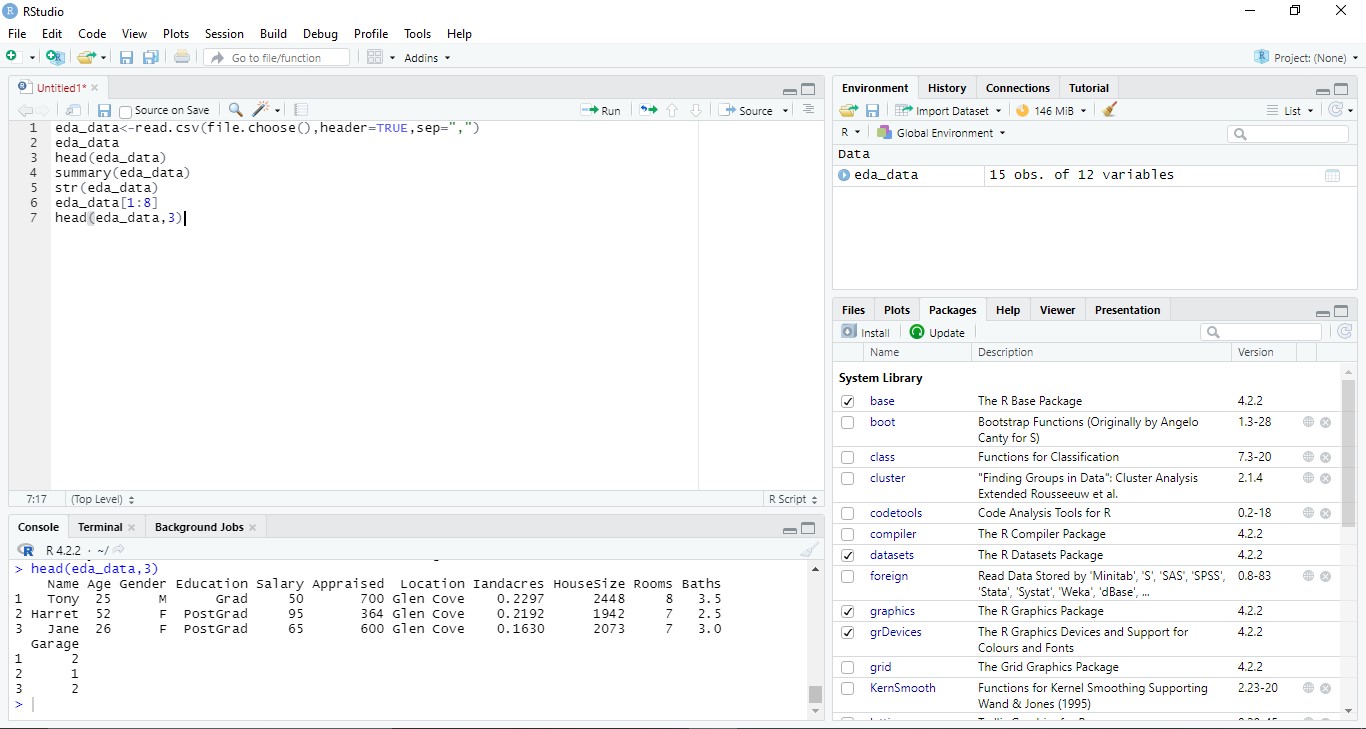
1. str(eda\_data)



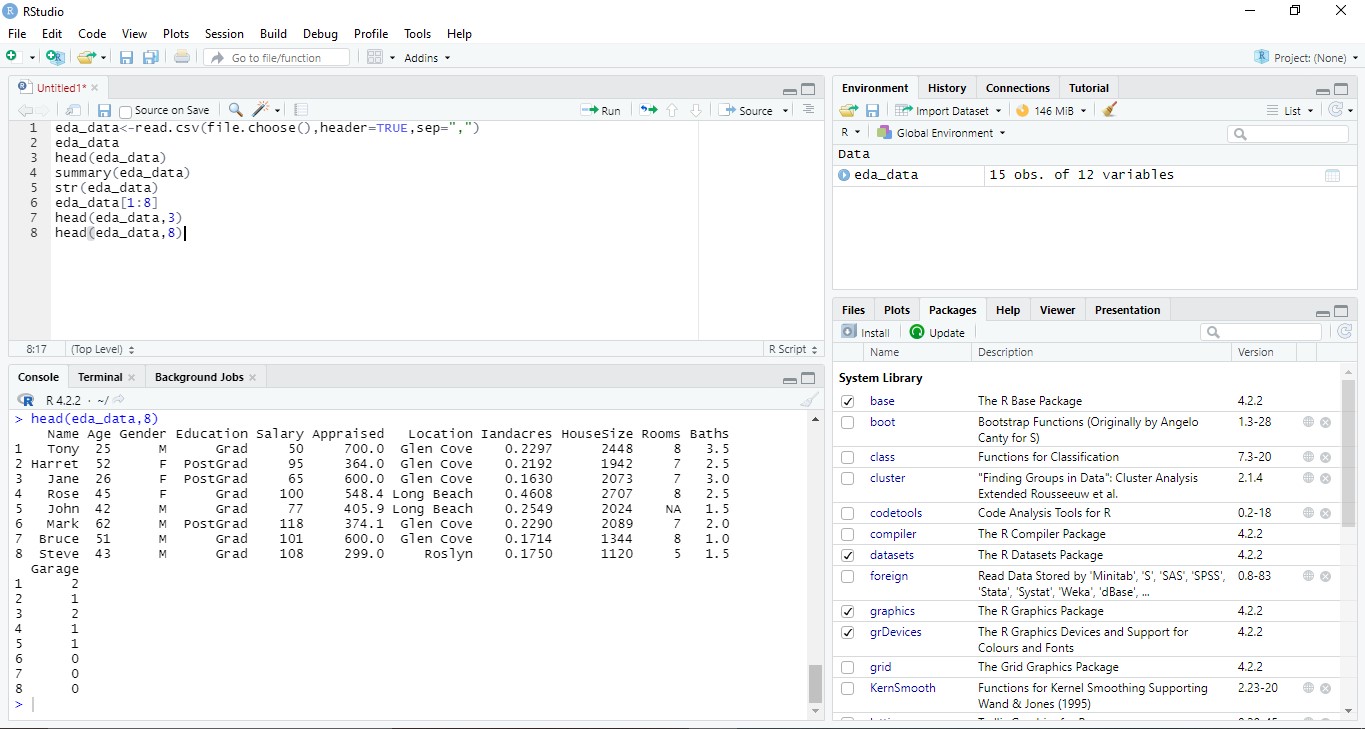
1. eda\_data[1:8]



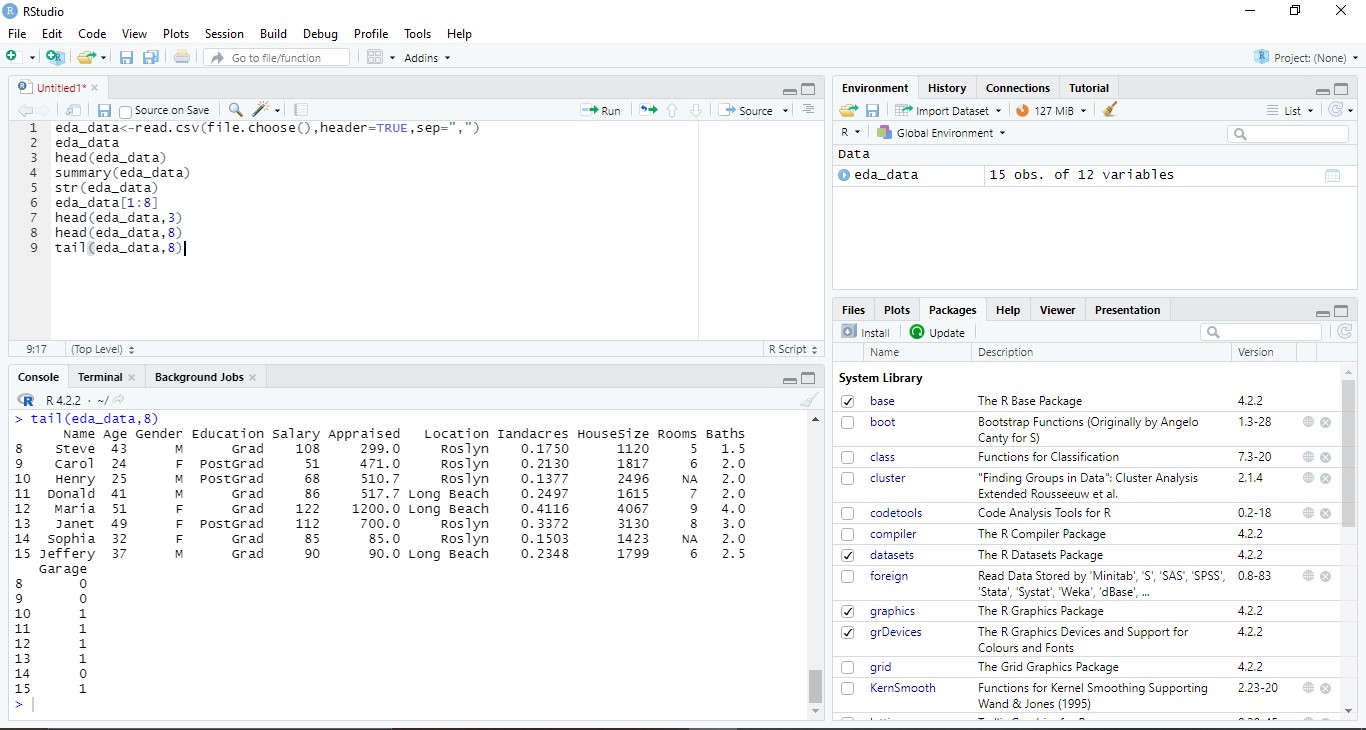
1. head(eda\_data,3)



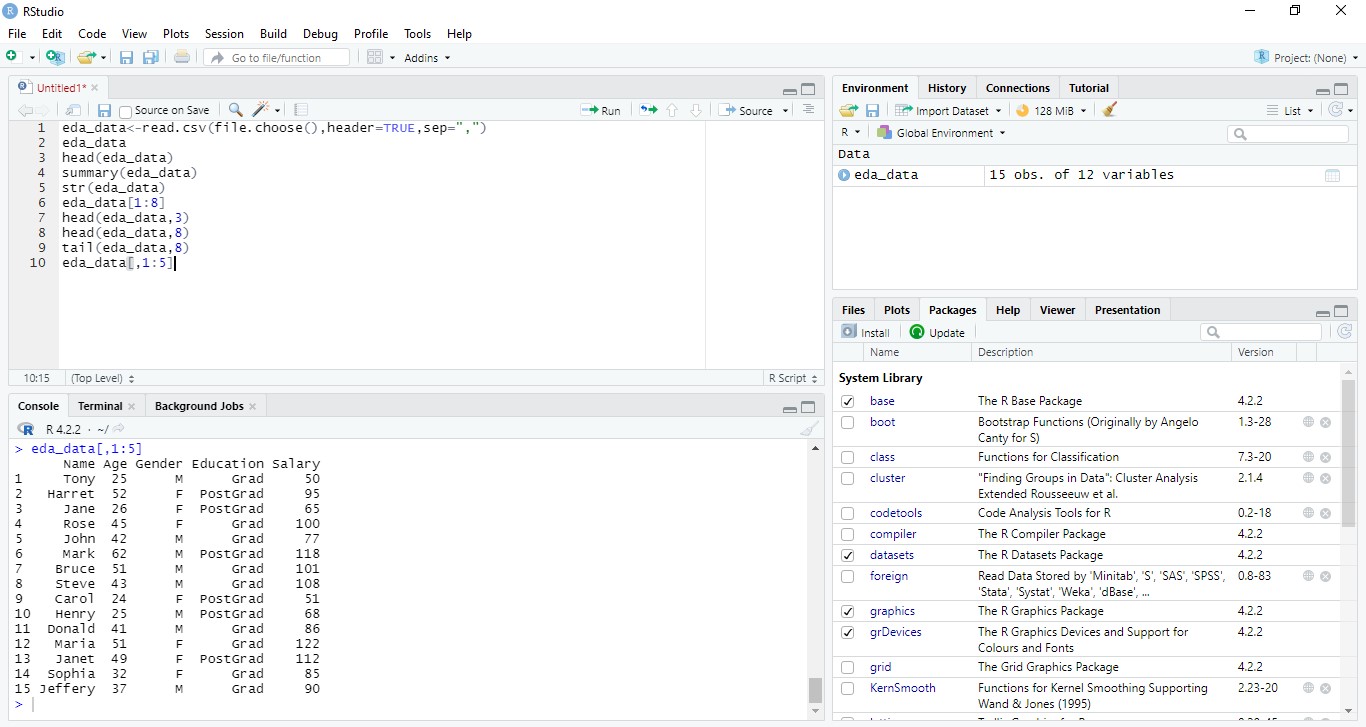
1. head(eda\_data,8)



1. tail(eda\_data,8)

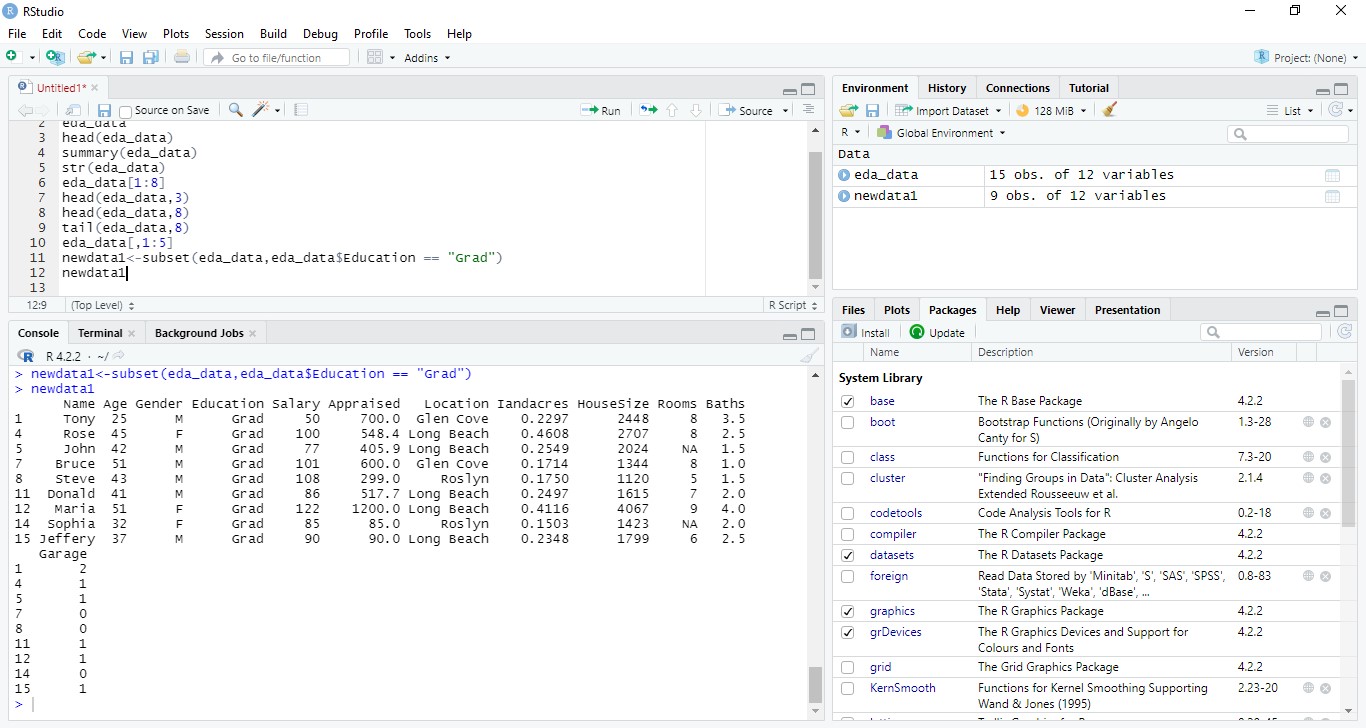


1. eda\_data[,1:5]



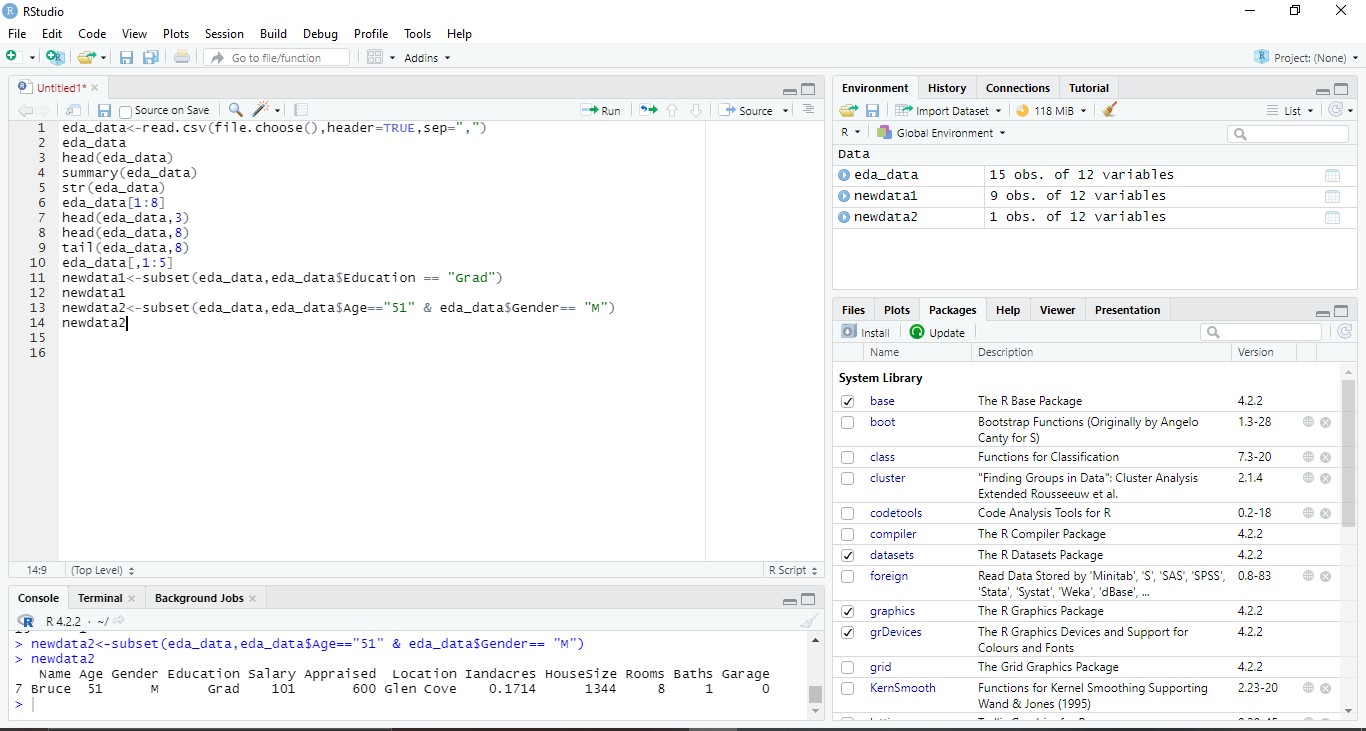
1. newdata1<-subset(eda\_data,eda\_data$Education == "Grad")

newdata1



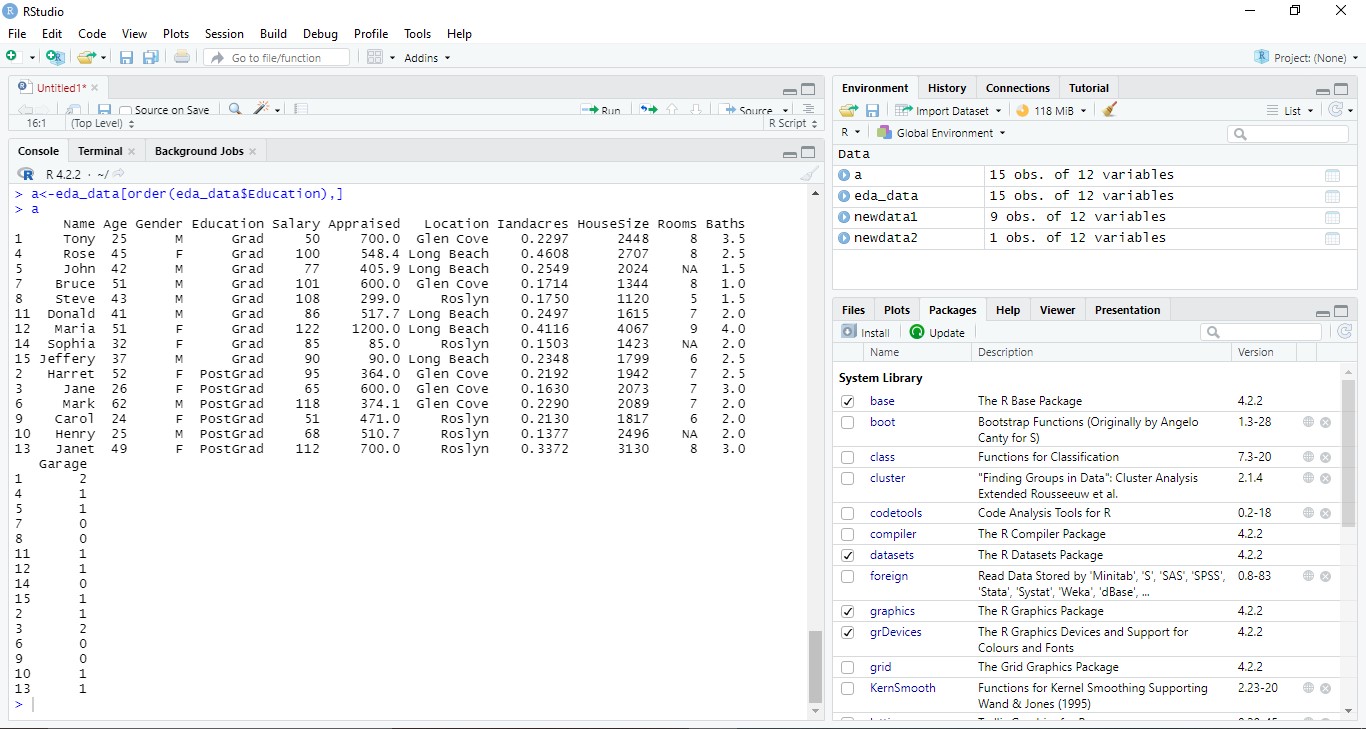
1. newdata2<-subset(eda\_data,eda\_data$Age=="51" & eda\_data$Gender== "M")

newdata2



1. a<-eda\_data[order(eda\_data$Education),]

a



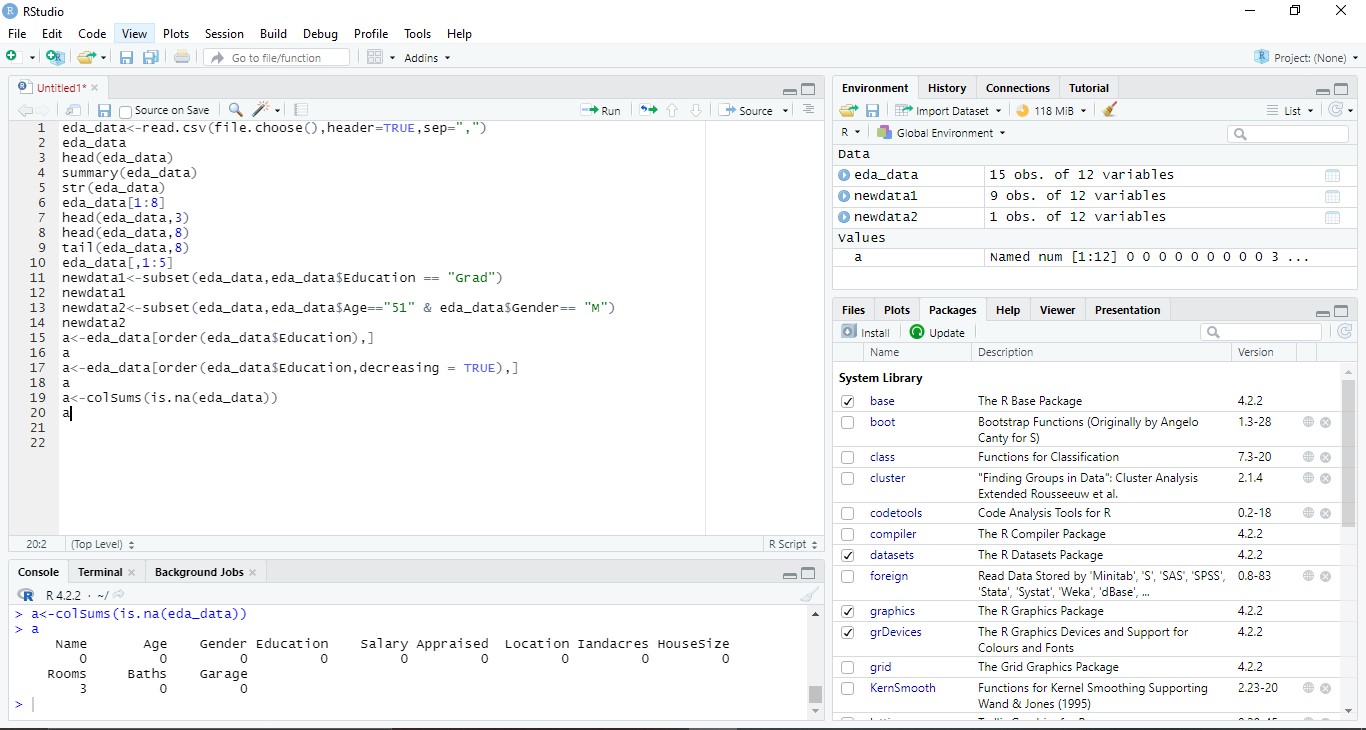
1. a<-eda\_data[order(eda\_data$Education,decreasing = TRUE),]

a



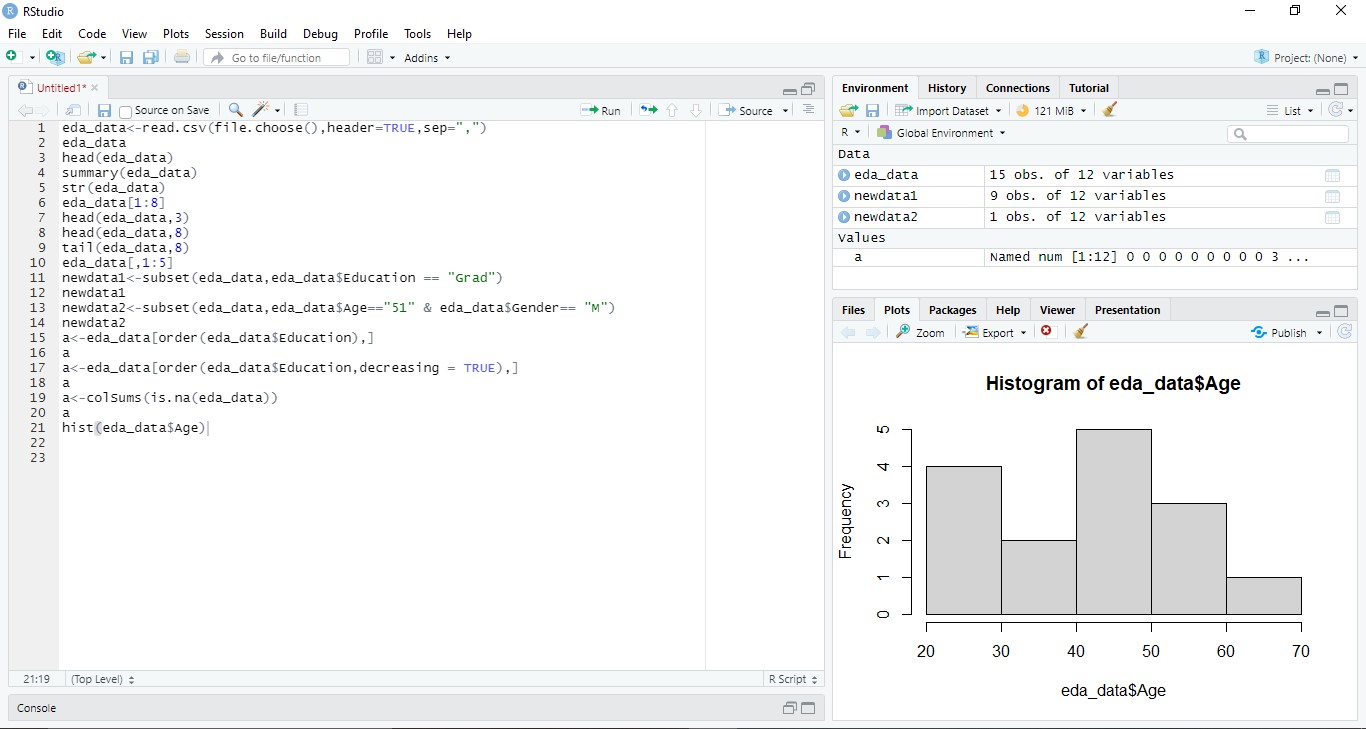
1. a<-colSums(is.na(eda\_data))

a

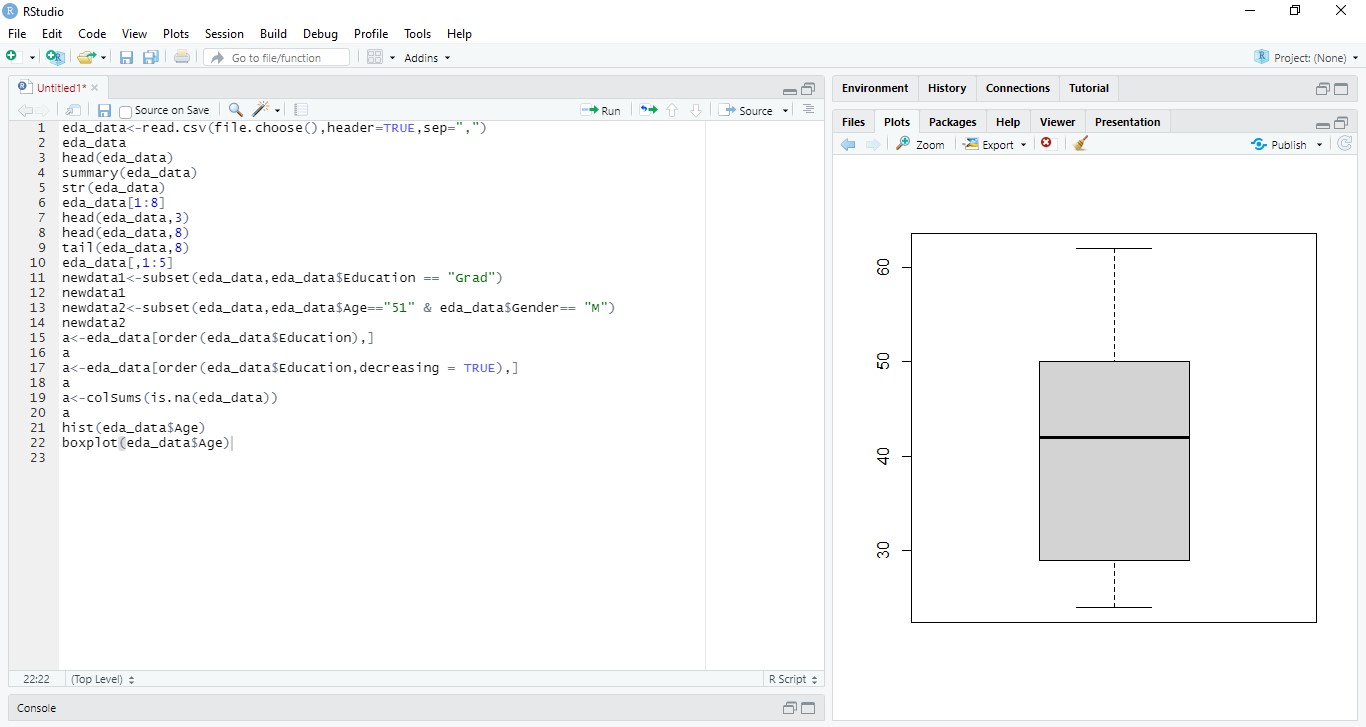


1. histogram

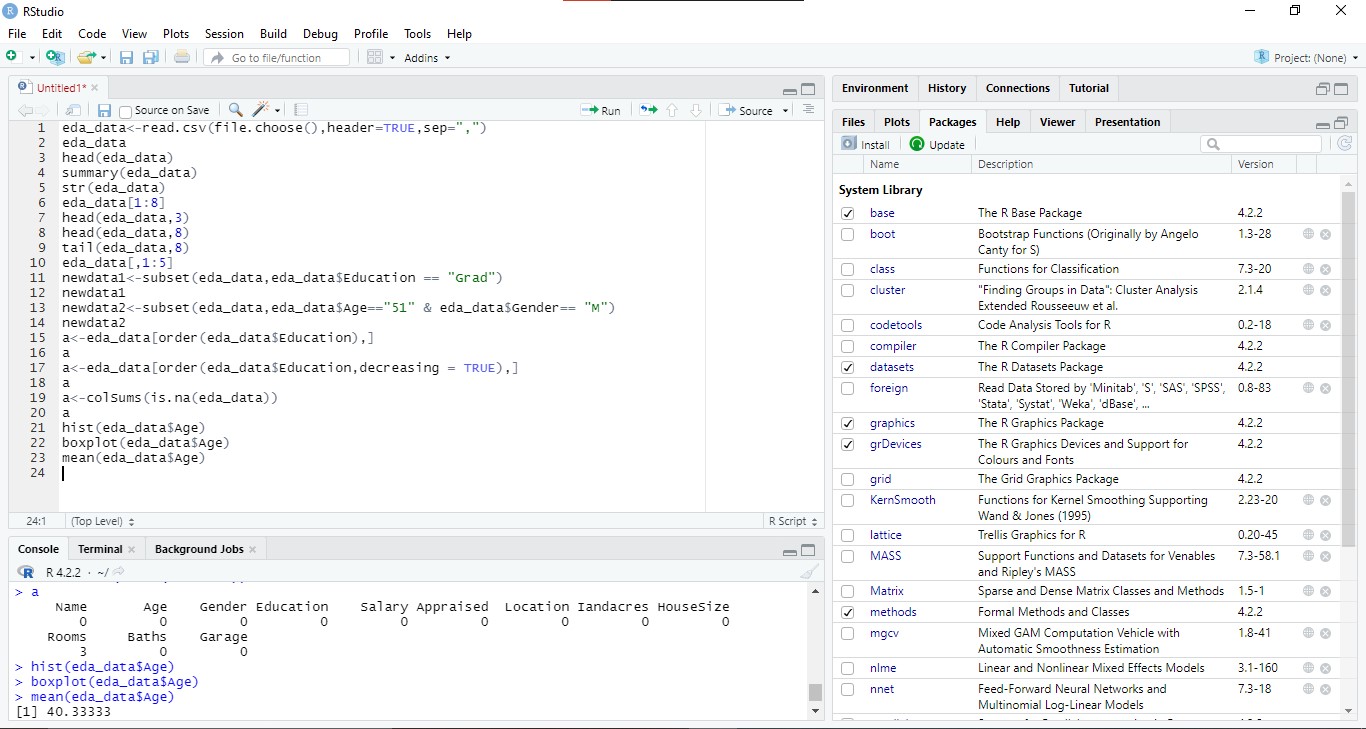
hist(eda\_data$Age)



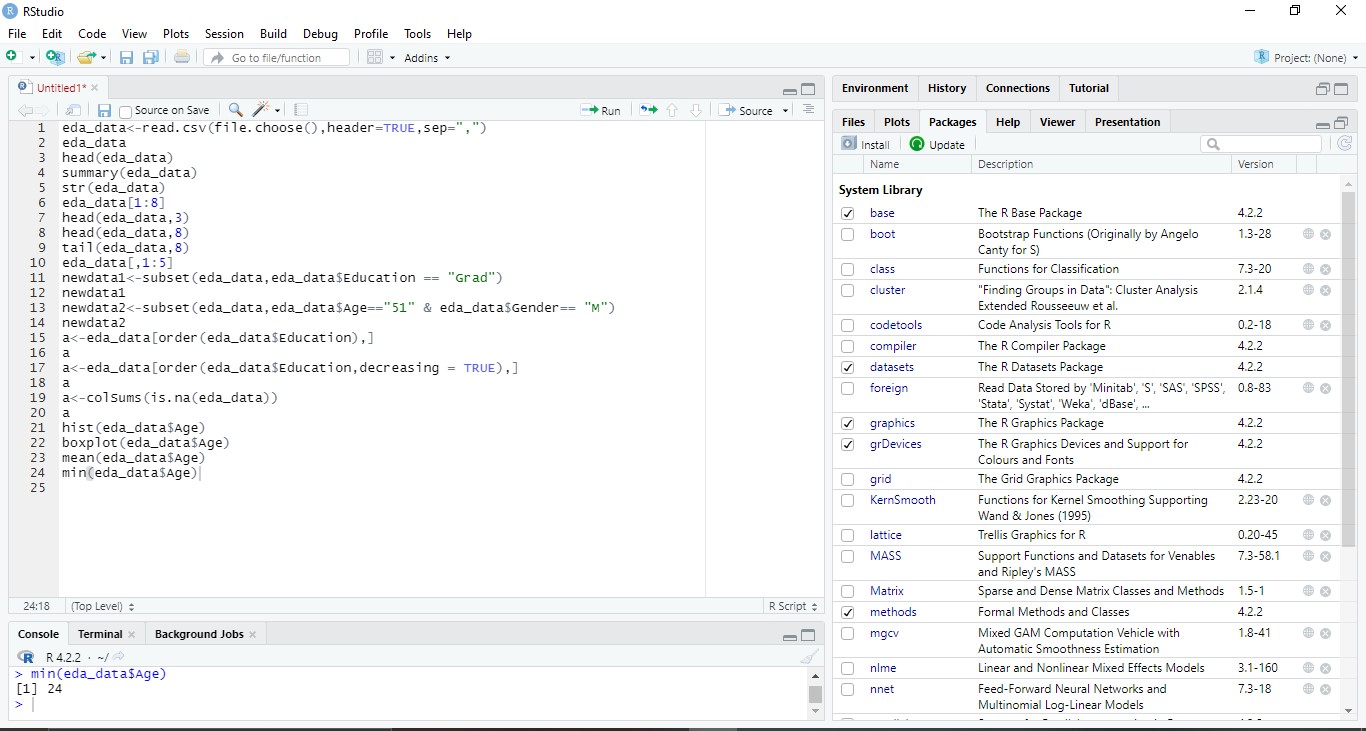
1. boxplot(eda\_data$Age)



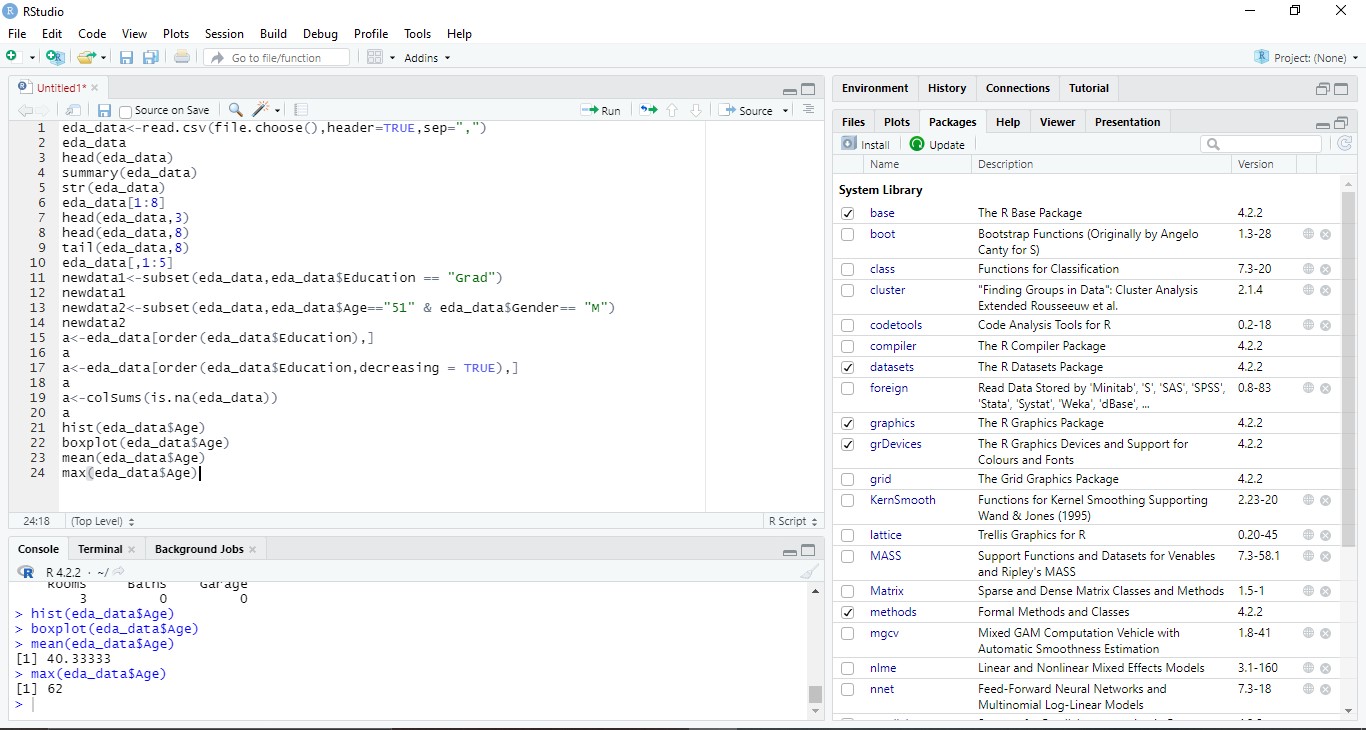
1. mean(eda\_data$Age)



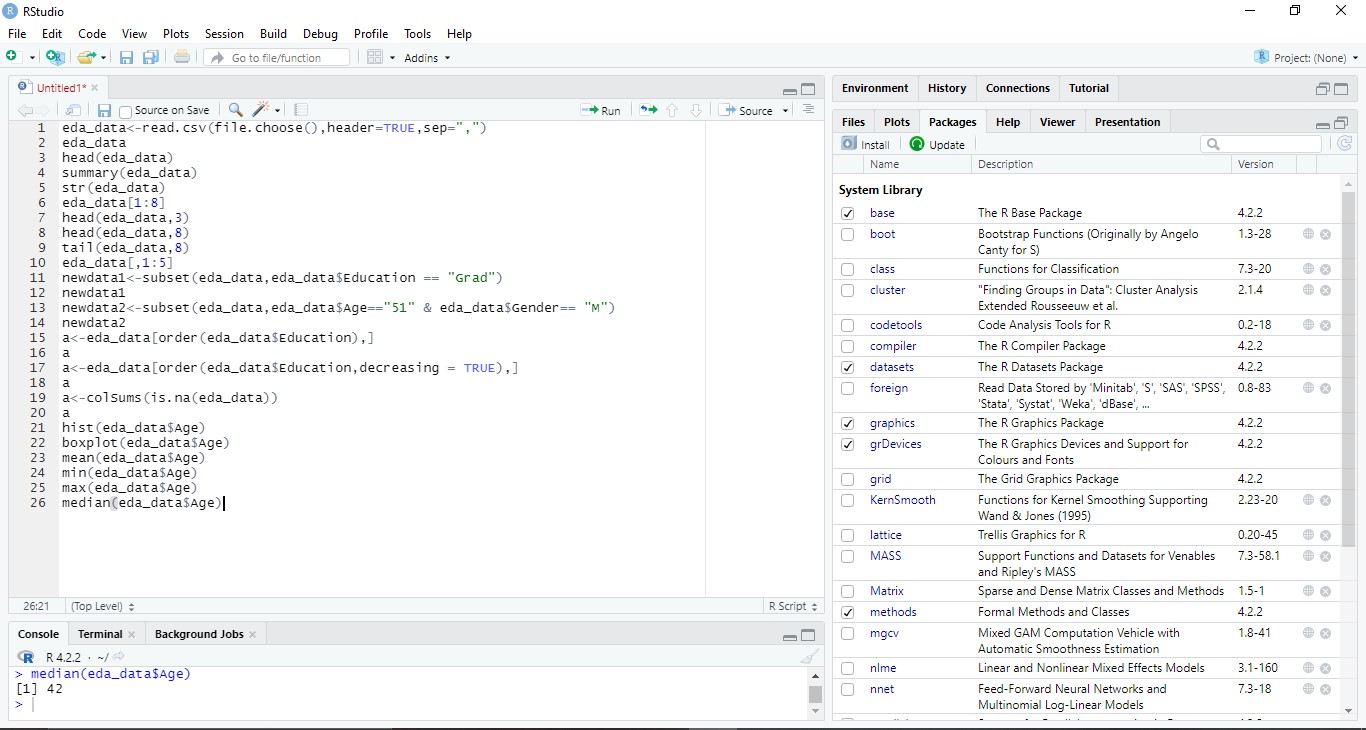
18)min(eda\_data$Age)



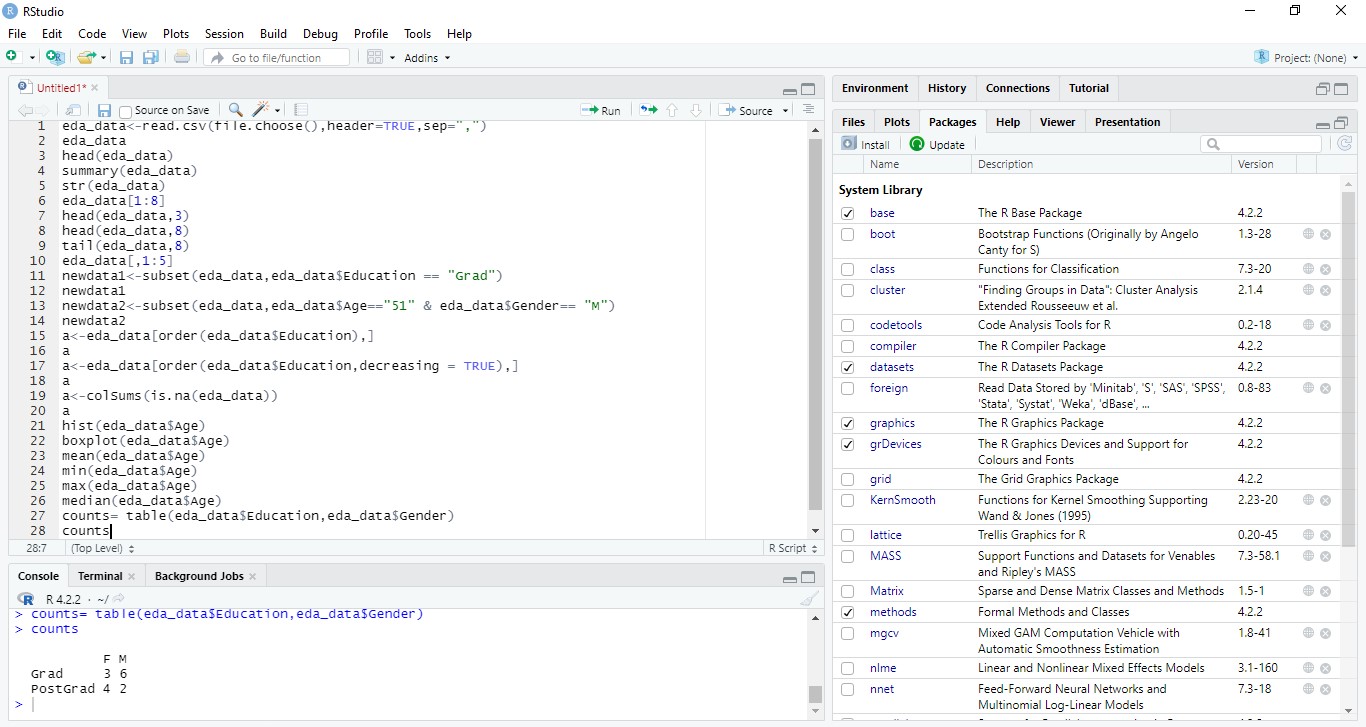
1. max(eda\_data$Age)



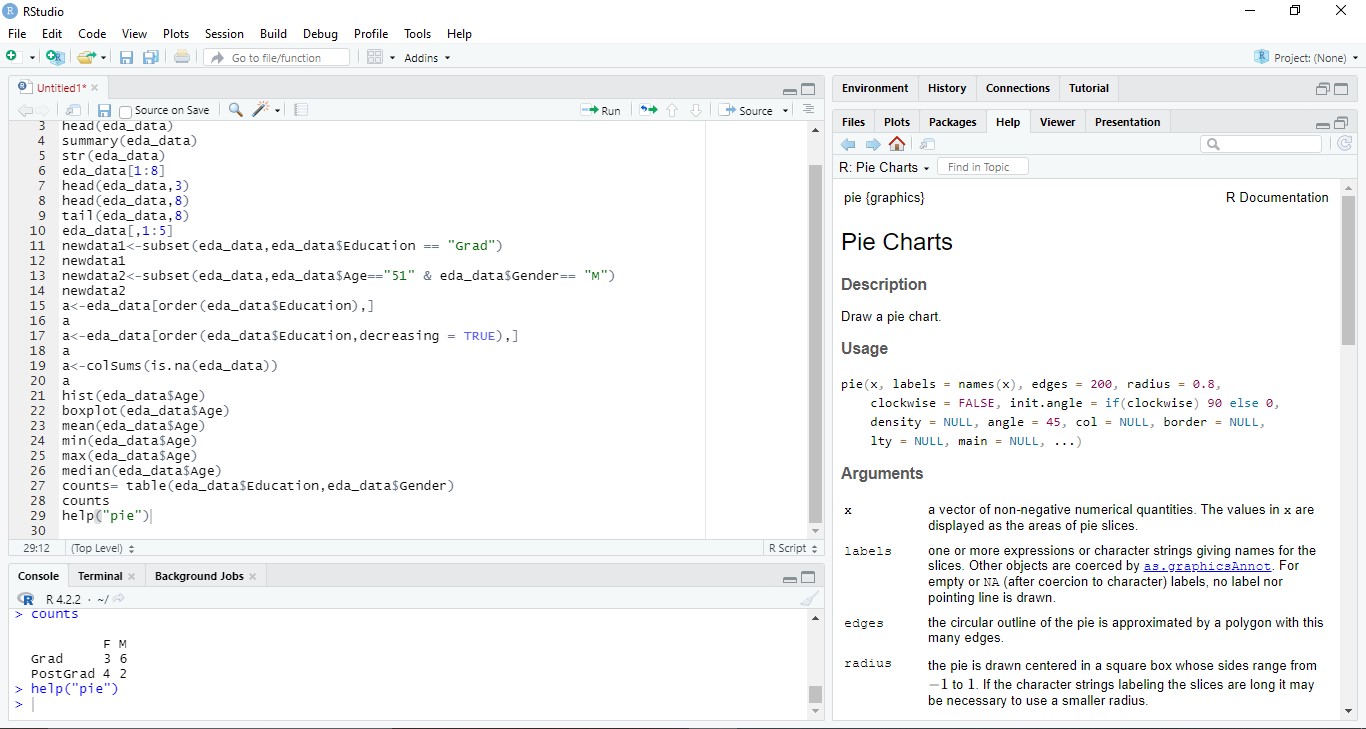
1. median(eda\_data$Age)



1. barplot counts= table(eda\_data$Education,eda\_data$Gender) counts

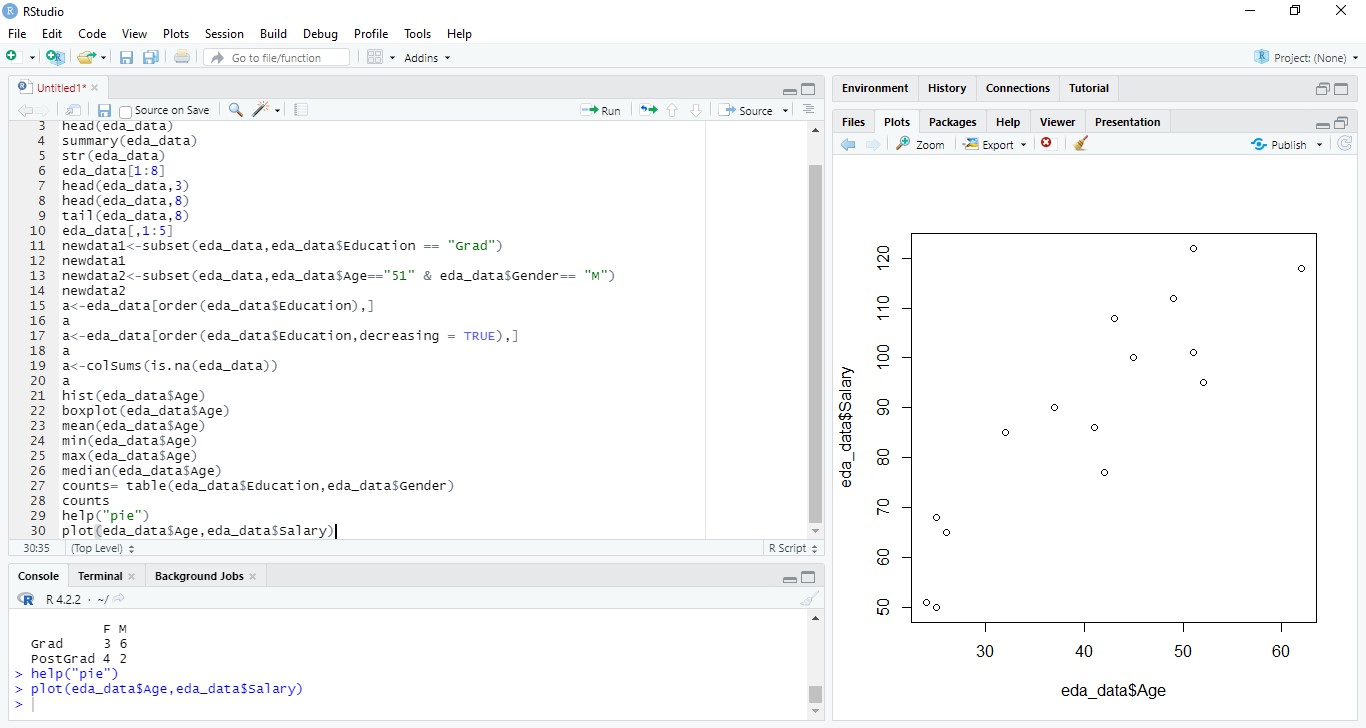


1. help("pie")#to know more about pie use this help



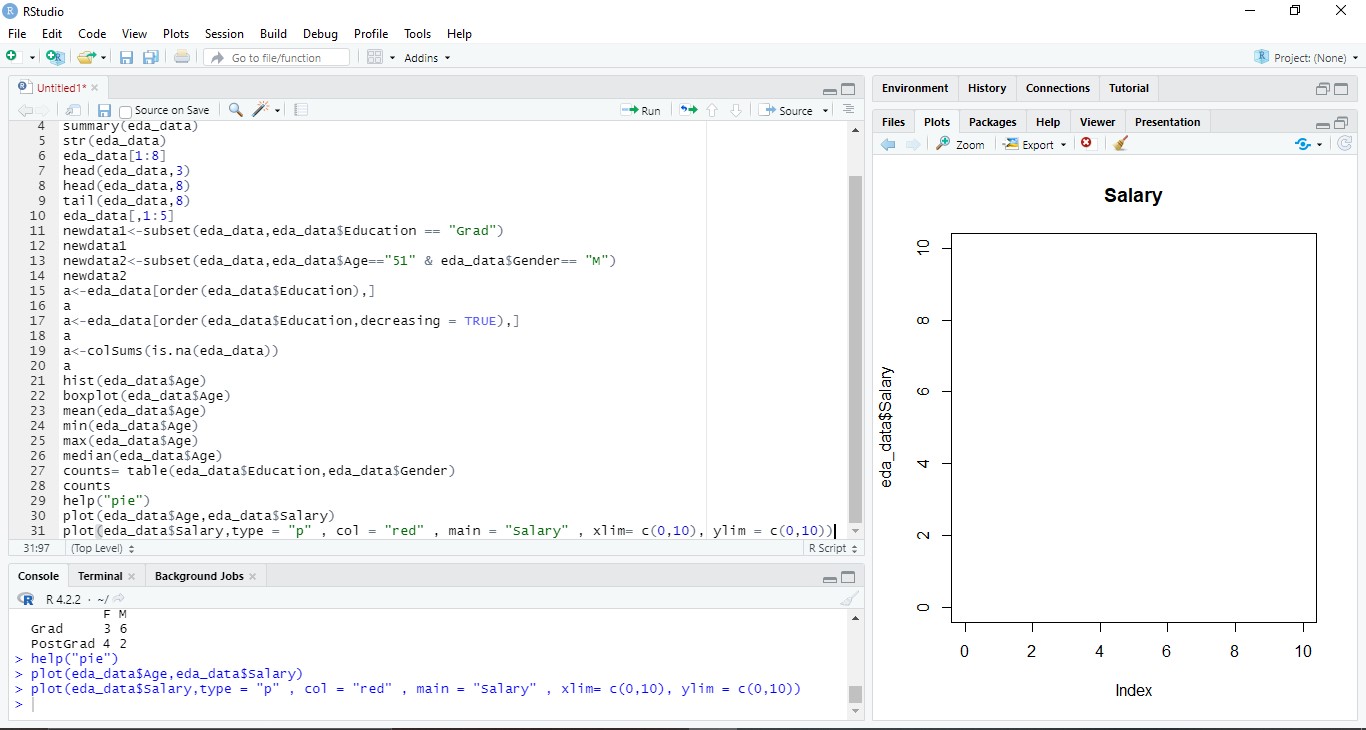
1. scatterplot

plot(eda\_data$Age,eda\_data$Salary)

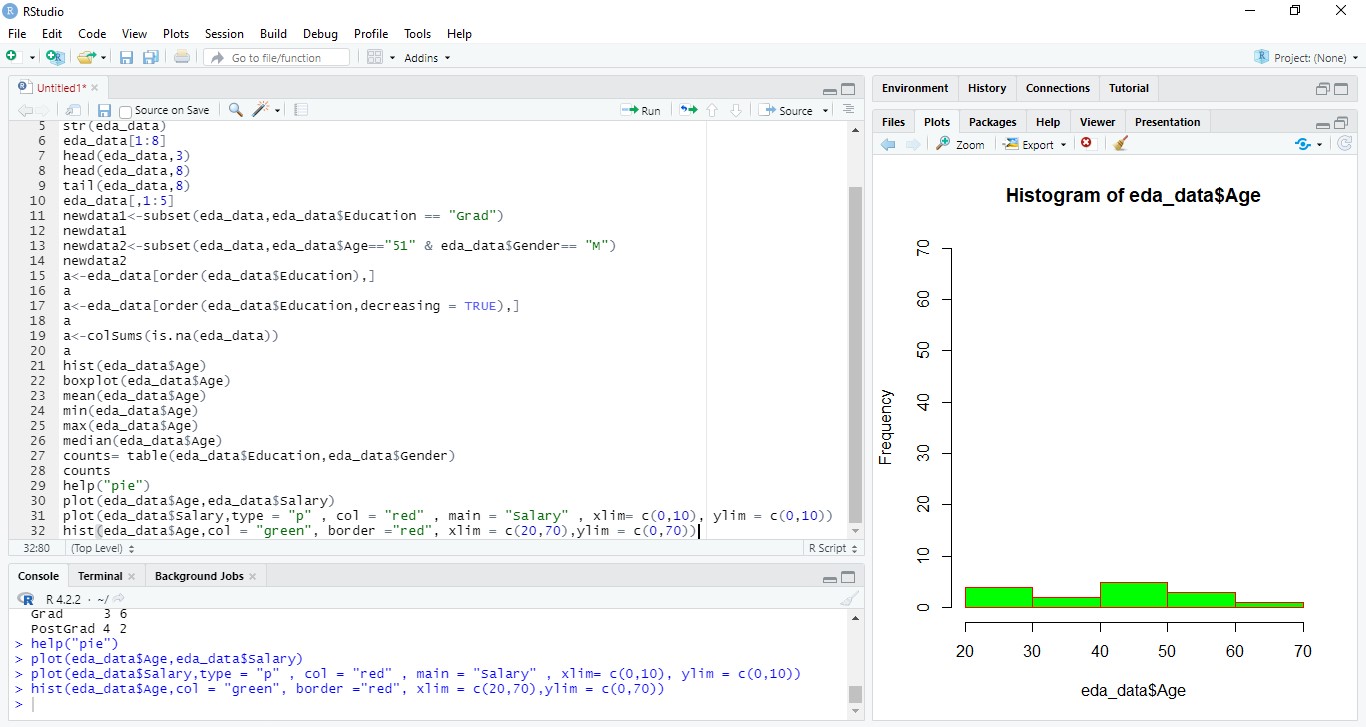


1. line-plot

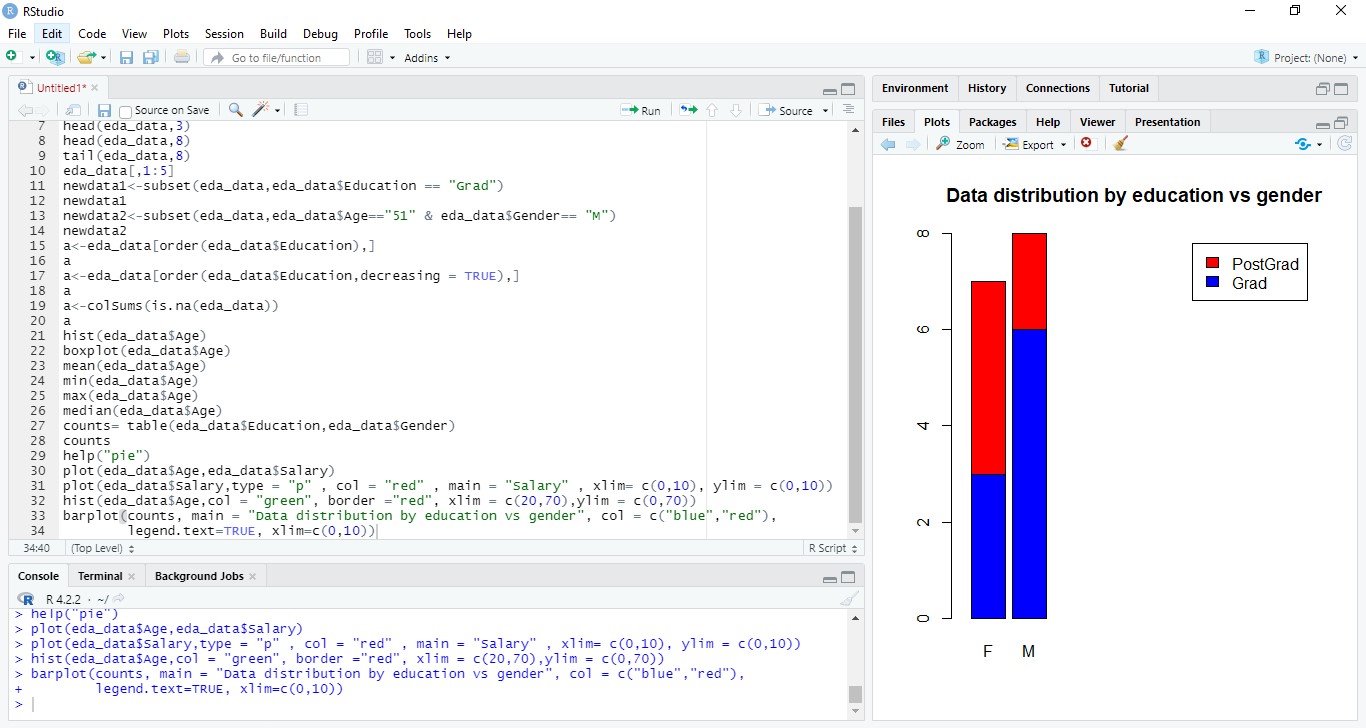
plot(eda\_data$Salary,type = "p" , col = "red" , main = "Salary" , xlim= c(0,10), ylim = c(0,10))



25)hist(eda\_data$Age,col = "green", border ="red", xlim = c(20,70),ylim = c(0,70))



26)barplot(counts, main = "Data distribution by education vs gender", col = c("blue","red"), legend.text=TRUE, xlim=c(0,10))



27)pie-chart x=table(eda\_data$Education) per<-round(x/sum(x)\*100,1)

pie(x, labels = per, main = "City pie chart", col = c("blue","red"))

