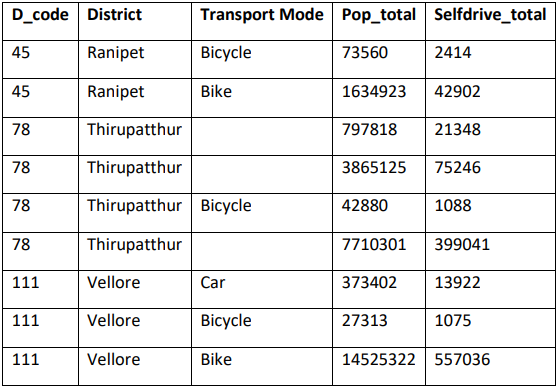
**Digital Assignment 1**

***Consider the following transportation dataset of three districts. It describes district code, District name, Transport mode, Total population and people who drove alone.***



**Write R code for the below questions**

1. **Create data frame for the above data**

**Code:**

trans.data <- data.frame(

D\_code = c(45,45,78,78,78,78,111,111,111),

District = c("Ranipet","Ranipet","Thirupatthur","Thirupatthur","Thirupatthur","Thirupatthur","Vellore","Vellore","Vellore"),

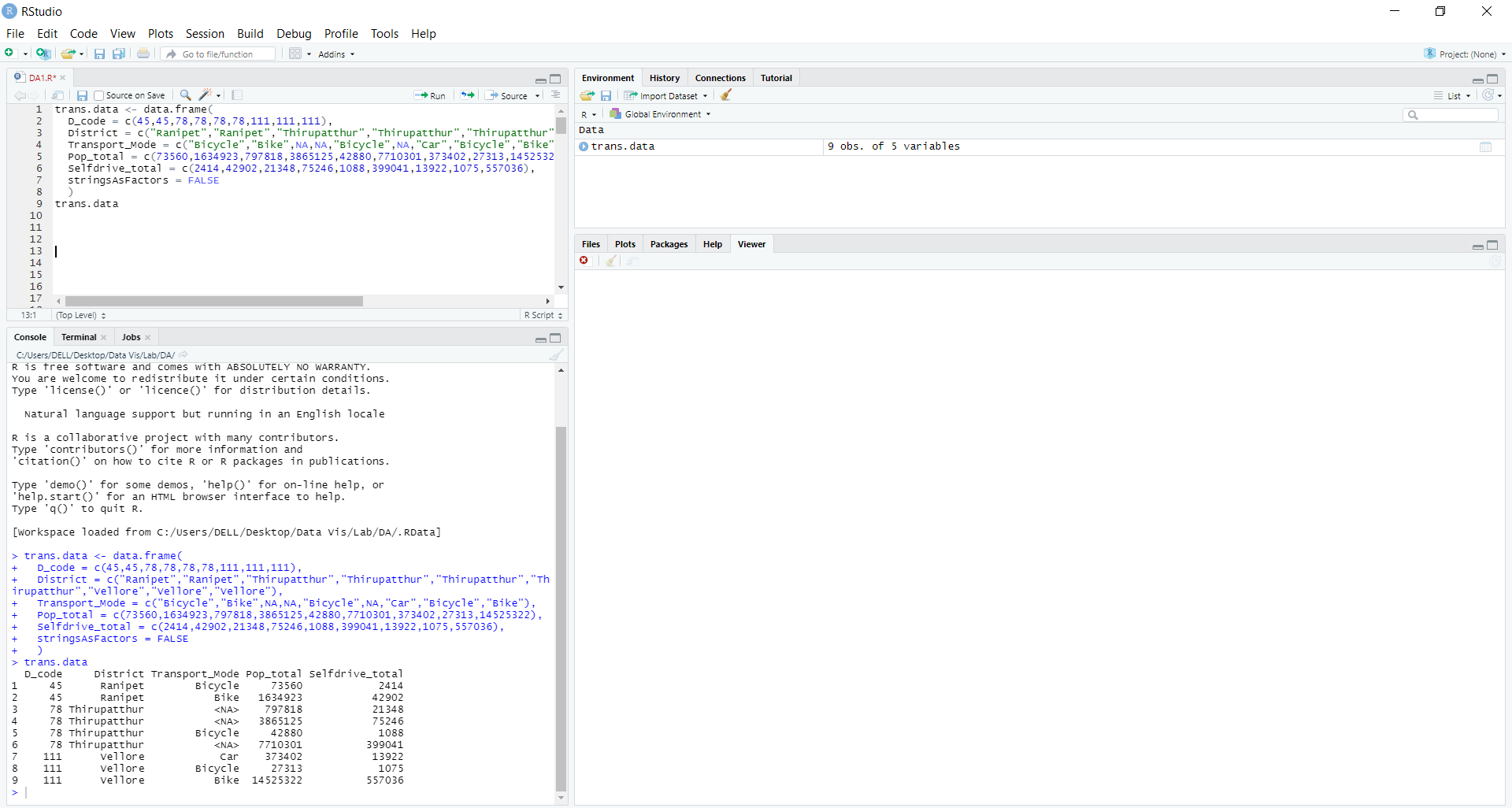
Transport\_Mode = c("Bicycle","Bike",NA,NA,"Bicycle",NA,"Car","Bicycle","Bike"),

Pop\_total = c(73560,1634923,797818,3865125,42880,7710301,373402,27313,14525322),

Selfdrive\_total = c(2414,42902,21348,75246,1088,399041,13922,1075,557036),

stringsAsFactors = FALSE

)

**Output:** 

1. **How many observations of ‘district’ are missing from the data frame**

**Code:**

i<-1

x=c()

for (n in 1:length(trans.data$Selfdrive\_total)){

if (is.na(trans.data[n, 'Transport\_Mode'])) {

x[i]<-trans.data[n, 'District']

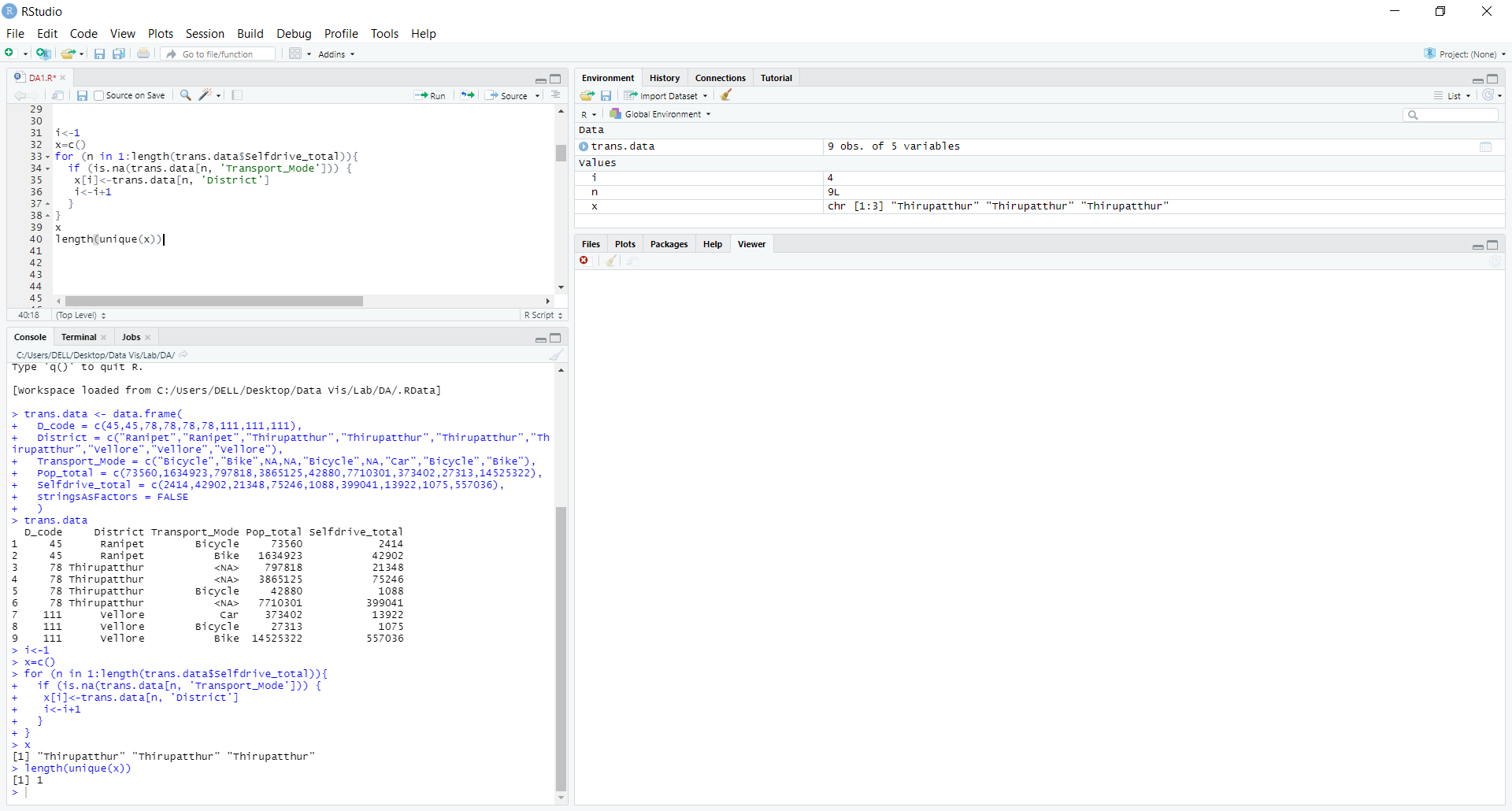
i<-i+1

}

}

x

length(unique(x))

**Output:** 

1. **Count the number of self-drive in each district.**

**Code:**

z=unique(trans.data$District)

y<-trans.data$D\_code

y<-unique(y)

sdt=c(0, 0, 0)

for (p in 1:length(y)){

for (q in 1:length(trans.data$Selfdrive\_total)){

if (trans.data$D\_code[q]==y[p]) {

sdt[p]=sdt[p]+trans.data$Selfdrive\_total[q]

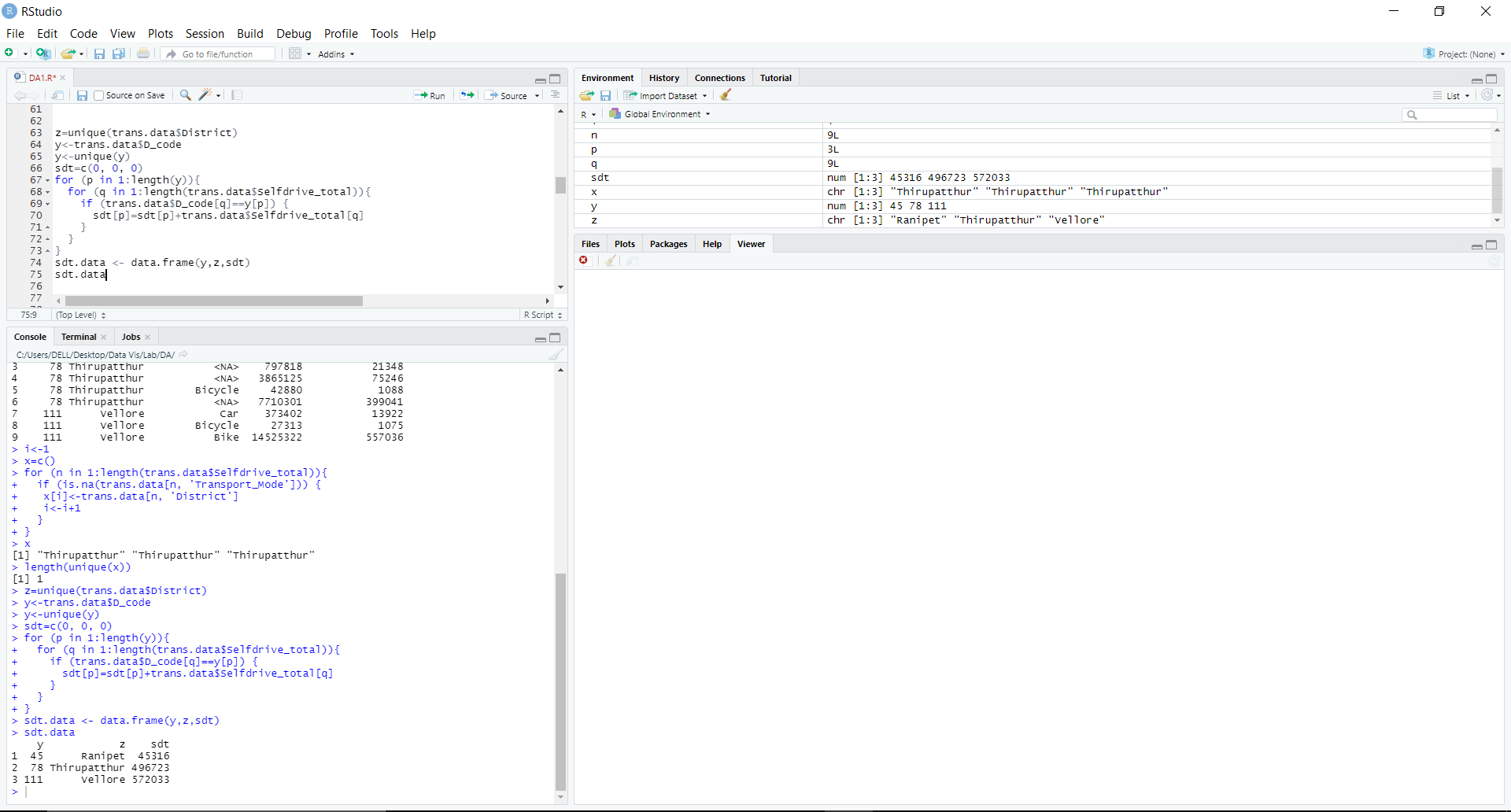
}

}

}

sdt.data <- data.frame(y,z,sdt)

sdt.data

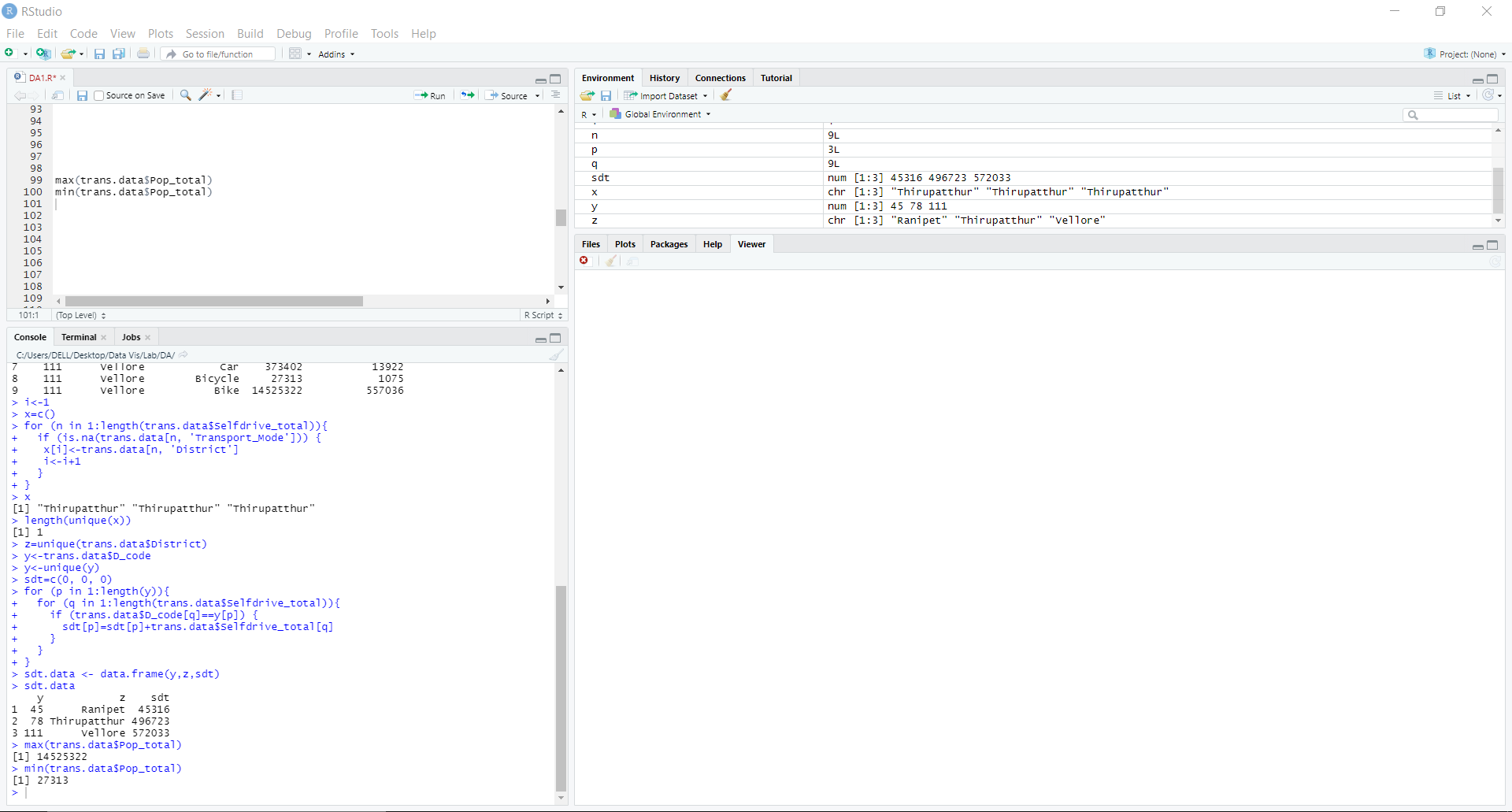
**Output:** 

1. **Print max and min of pop\_total.**

**Code:**

max(trans.data$Pop\_total)

min(trans.data$Pop\_total)

**Output:** 

1. **Derive new information/print "percentage of people who drove alone in all three districts" and also rank districts based on the % of people who used bicycle.**

**Code:**

popt=c(0, 0, 0)

for (p in 1:length(y)){

for (q in 1:length(trans.data$Selfdrive\_total)){

if (trans.data$D\_code[q]==y[p]) {

popt[p]=popt[p]+trans.data$Pop\_total[q]

}

}

}

sdt.data <- data.frame(y,z,sdt,popt)

sdt.data

sdtper=(sdt.data$sdt/sdt.data$popt)\*100

sdt.data <- data.frame(y,z,sdt,popt,sdtper)

sdt.data

Bicy=c()

d=1

for (h in 1:9){

if(!is.na(trans.data[h, 'Transport\_Mode']) & trans.data$Transport\_Mode[h]=="Bicycle"){

Bicy[d]=trans.data$Pop\_total[h]

d<-d+1

}

}

sdt.data <- data.frame(y,z,sdt,popt,sdtper,Bicy)

sdt.data

Bicyper=(sdt.data$Bicy/sdt.data$popt)\*100

sdt.data <- data.frame(y,z,sdt,popt,sdtper,Bicy,Bicyper)

sdt.data

**Output:** 