**Data Visualization Lab Assignment 3**

|  |  |  |  |
| --- | --- | --- | --- |
| Loan Sanctioned | Age | Income | Expenditure |
| 1000000 | 50 | 50000 | 40000 |
| 1200000 | 45 | 30000 | 25000 |
| 300000 | 20 | 10000 | 8000 |
| 3000000 | 35 | 60000 | 55000 |
| 4500000 | 30 | 50000 | 30000 |
| 3500000 | 40 | 60000 | 45000 |
| 5000000 | 45 | 80000 | 50000 |
|  |  |  |  |

Name: Om Ashish Mishra

Registration Number: 16BCE0789

Slot: C2

**Q1)**

|  |  |  |  |
| --- | --- | --- | --- |
| Loan Sanctioned | Age | Income | Expenditure |
| 1000000 | 50 | 50000 | 40000 |
| 1200000 | 45 | 30000 | 25000 |
| 300000 | 20 | 10000 | 8000 |
| 3000000 | 35 | 60000 | 55000 |
| 4500000 | 30 | 50000 | 30000 |
| 3500000 | 40 | 60000 | 45000 |
| 5000000 | 45 | 80000 | 50000 |
|  |  |  |  |

**For the data given in table fit a regression model for the attribute Loan Sanctioned. Use visualization techniques to identify suitable attributes (Age,Income,Expenditure) to construct the model.**

**After construction, simulate the model and obtain the predicted values. Show the residual plot for observed and simulated values for the attribute loan sanctioned.**

**CODE:**

loan = c(1000000, 1200000, 300000, 3000000, 4500000, 3500000, 5000000)

age = c(50, 45, 20, 35, 30, 40, 45)

income = c(50000, 30000, 10000, 60000, 50000, 60000, 80000)

expenditure = c(40000, 25000, 8000, 55000, 30000, 45000, 50000)

#TODO: MAKE COMPARITIVE GRAPHS BETWEEN ABOVE

df = data.frame(loan, age, income, expenditure)

df

model = lm(formula = loan ~ age + income + expenditure, data=df)

model

plot(model)

Y <- function(x, y, z) {

t = 1555398.52 - 71082.69\*x + 138.5\*y - 81.38\*z

return(t)

}

predicted = c()

original = loan

for(i in 1:length(loan)) {

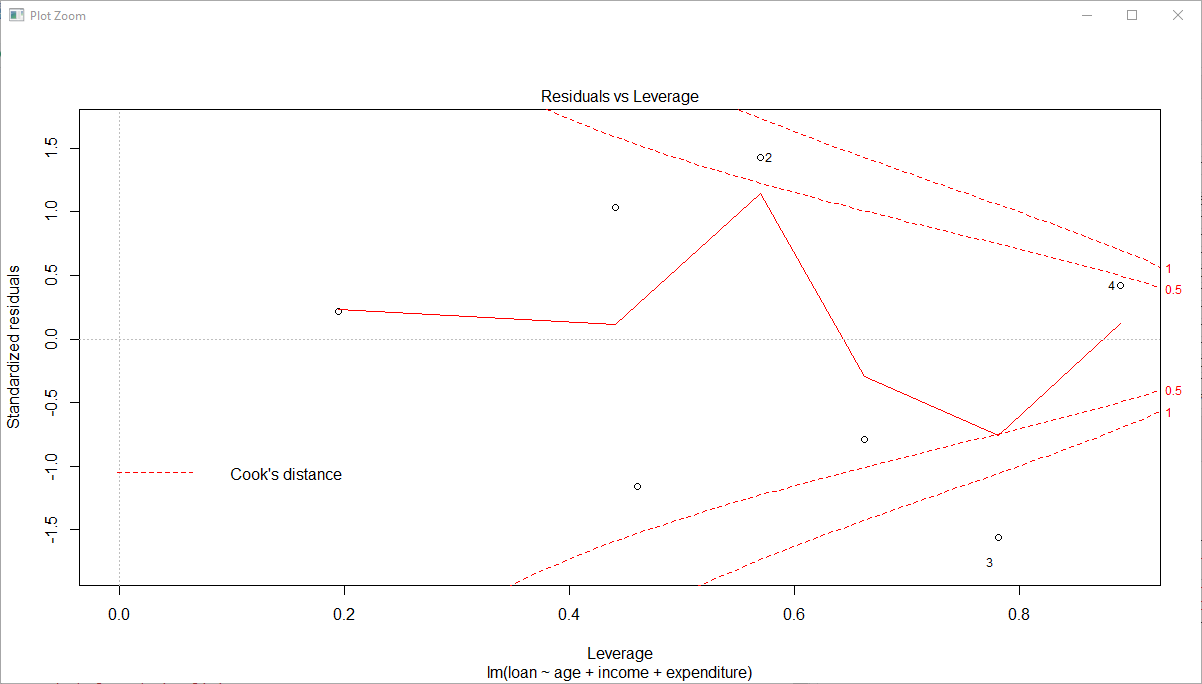
predicted[i] = Y(age[i], income[i], expenditure[i])

}

df2 = data.frame(predicted, original)

df2

**OUTPUT:**



Q2) **Generate tree data with 30 nodes with maximum of 5 children to each node and store it in a file. Visualize the stored data using radial spline, matrix and tree map.**

**CODE:**

Main <- Node$new("Tree")

a <- Main$AddChild("a")

b <- Main$AddChild("b")

c <- Main$AddChild("c")

d <- Main$AddChild("d")

e <- Main$AddChild("e")

f <- a$AddChild("f")

g <- a$AddChild("g")

h <- a$AddChild("h")

i <- a$AddChild("i")

j <- a$AddChild("j")

k <- b$AddChild("k")

l <- b$AddChild("l")

m <- b$AddChild("m")

n <- b$AddChild("n")

o <- b$AddChild("o")

p <- c$AddChild("p")

q <- c$AddChild("q")

h <- c$AddChild("r")

s <- c$AddChild("s")

t <- c$AddChild("t")

u <- d$AddChild("u")

v <- d$AddChild("v")

w <- d$AddChild("w")

x <- d$AddChild("x")

y <- d$AddChild("y")

z <- e$AddChild("z")

aa <- e$AddChild("aa")

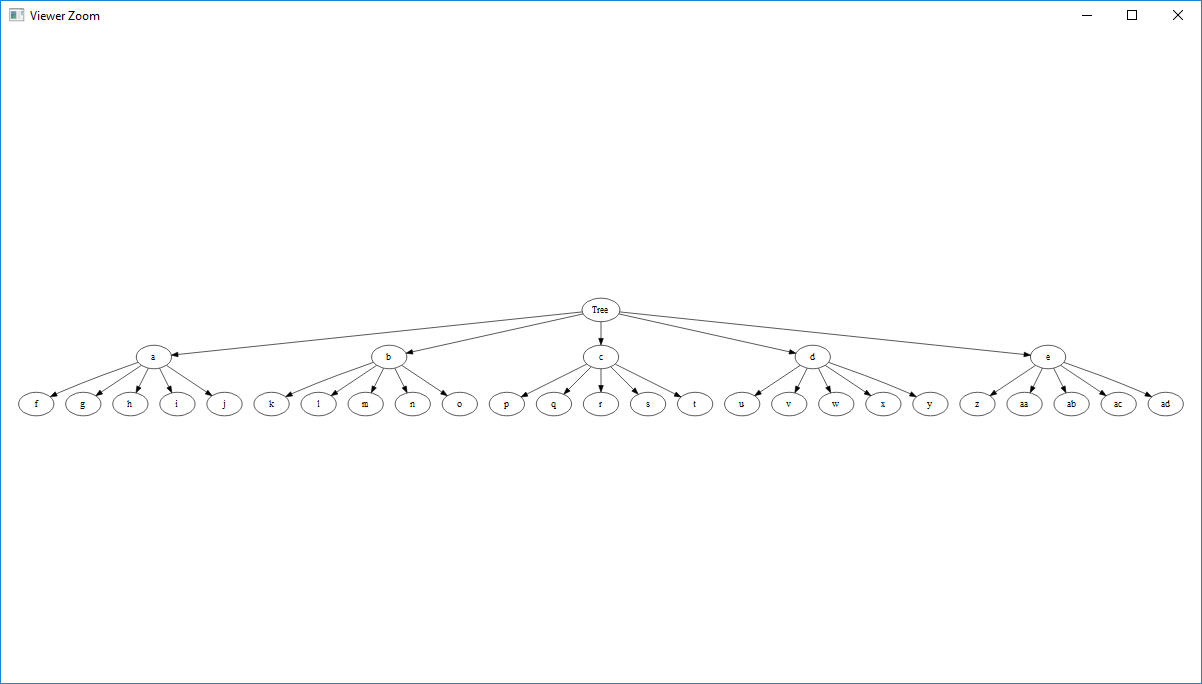
ab <- e$AddChild("ab")

ac <- e$AddChild("ac")

ad <- e$AddChild("ad")

plot(Main)

**OUTPUT:**



Q3)

**Use any geo-visual package in R and show India with states. Use color coding to differentiate southern and northern states. Also provide interactive feature for selecting a specific state**

**CODE:**

library(ggplot2)

library(ggmap)

library(maps)

library(mapdata)

usa <- map\_data("usa")

ggplot() + geom\_polygon(data = usa, aes(x=long, y = lat, group = group)) +

coord\_fixed(1.3)

states <- map\_data("state")

ggplot(data = states) +

geom\_polygon(aes(x = long, y = lat, fill = region, group = group), color = "white") +

coord\_fixed(1.3) # do this to leave off the color legend

**OUTPUT:**

