 **Problem Statement**

library(RcmdrPlugin.IPSUR)

data(RcmdrTestDrive)

Perform the below operations:

1. Compute the measures of central tendency for salary and reduction which variable has highest center?

#Solution1

#first find the measures of central tendency for salary and reduction

#for salary

library(RcmdrPlugin.IPSUR)

x<- c(mean(RcmdrTestDrive$salary),median(RcmdrTestDrive$salary))

x

#for reduction

y<- c(median(RcmdrTestDrive$reduction),mean(RcmdrTestDrive$reduction))

y

#now since we are looking for variable which has highest center

#we can check for this by plotting histogram or

#by checking kurtosis which describes the amount of peakedness of a distribution.

library(psych)

kurtosi(RcmdrTestDrive$salary)

kurtosi(RcmdrTestDrive$reduction)

#thus we can see variable reduction has more kurtosis thus more peaked hence more highest center

1. Which measure of center is more appropriate for before and after?

#If the distribution is fairly symmetric then the mean and median

#should be approximately the same

#by boxplot we can check for median where it lies

boxplot(RcmdrTestDrive$before,horizontal = T,col = "red",xlab="before",ylab="Boxplot")

#normal distributed

boxplot(RcmdrTestDrive$after,horizontal = T,col = "red",xlab="after",ylab="Boxplot")

#left skewed as the data is assymetrical distributed

#if we check the skewness of variables

skew (RcmdrTestDrive$before)

skew (RcmdrTestDrive$after)

#after more negative so data more on right side as compare to before variable

#thus, the median would likely be a good choice and it is more appropriate