SESSION 8: Exploratory Data Analytics

Assignment 2

1. Use the package -RcmdrPlugin.IPSUR.

data(RcmdrTestDrive)

And perform the below operations:

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

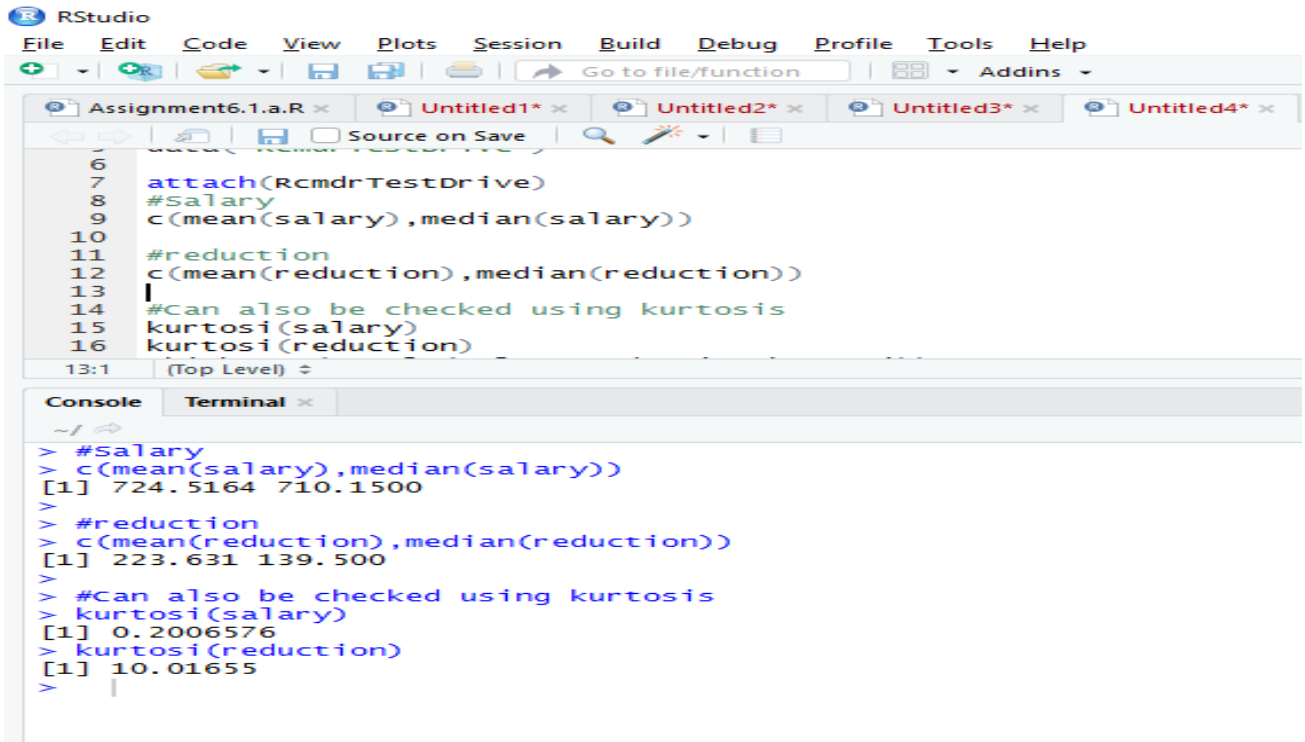
Answer:

```
library(RcmdrPlugin.IPSUR)
library(psych)
data("RcmdrTestDrive")
attach(RcmdrTestDrive)

#Salary
c(mean(salary),median(salary))
#reduction
c(mean(reduction),median(reduction))

#To check which has the highest center
#Can be done using kurtosis
kurtosi(salary)
kurtosi(reduction)
#Since reduction has more value, it is having the highest peak
```

Output:



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

Answer:

```
#if distribution is symmetric, then mean and median should be same,
#Boxplot can be used to check where the median lies in the distibution
windows()
par(mfrow=c(1,2))
boxplot(before,horizontal = T,main="Before")
boxplot(after,horizontal = T,main="Before")

#Skewness can be used to check the assymetric distribution
skew(before)
skew(after)
#after is more -ve means more distibution on the right side.
```

Output: