

Assignment - 2

Session 8 – Exploratory Data Analytics

5.Problem Statement

```
library(RcmdrPlugin.IPSUR)
```

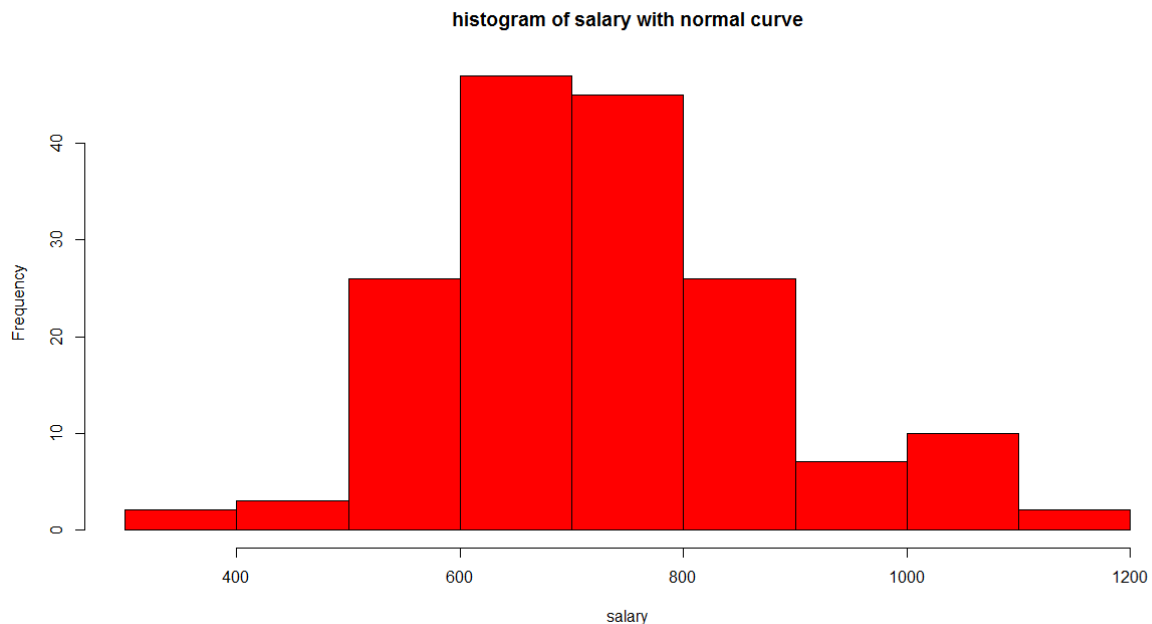
```
data(RcmdrTestDrive)
```

Perform the below operations:

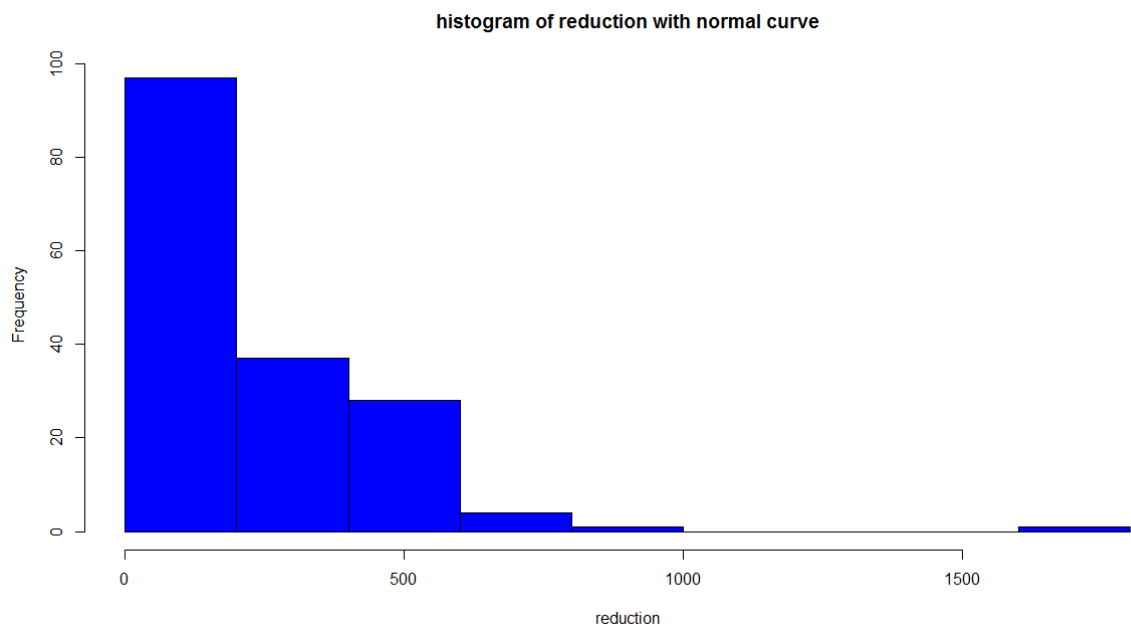
1. Compute the measures of central tendency for salary and reduction which variable has highest center?

Ans:

```
library(RcmdrPlugin.IPSUR)
> x<- c(mean(RcmdrTestDrive$salary),median(RcmdrTestDrive$salary))
> x
[1] 724.5164 710.1500
> y<- c(median(RcmdrTestDrive$reduction),mean(RcmdrTestDrive$reduction))
> y
[1] 139.500 223.631
> library(psych)
> kurtosi(RcmdrTestDrive$salary)
[1] 0.2006576
> kurtosi(RcmdrTestDrive$reduction)
[1] 10.01655
> x<-RcmdrTestDrive$salary
> h<-hist(x,breaks=10,col="red",xlab="salary",main="histogram of salary wi
th normal curve")
```



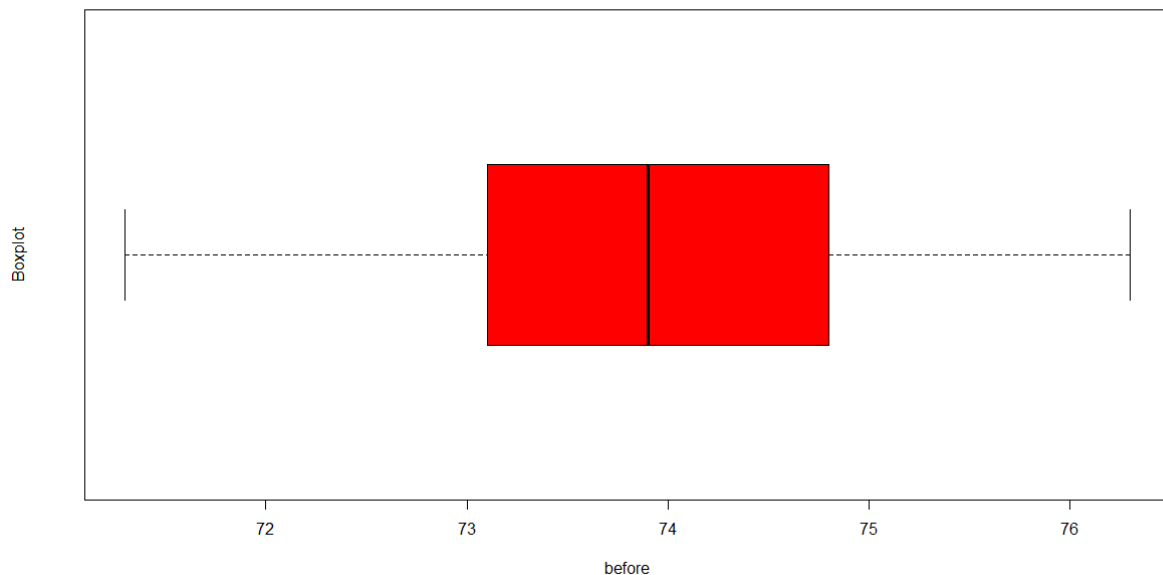
```
> y<-RcmdrTestDrive$reduction
> h<-hist(y,breaks=10,col="blue",xlab="reduction",main="histogram of reduc
tion with normal curve")
```



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

Ans:

```
> #by boxplot we can check for median where it lies
> boxplot(RcmdrTestDrive$before, horizontal=T, col="red", xlab="before", ylab="Boxplot")
```



```
> boxplot(RcmdrTestDrive$after, horizontal=T, col="red", xlab="after", ylab="Boxplot")
> #if we check the skewness of variables
> skew (RcmdrTestDrive$before)
```

```
[1] -0.03510369  
> skew (RcmdrTestDrive$after)  
[1] -1.164056
```

