ACADGILD

SESSION 8: Exploratory Data Analytics

Assignment 2

Table of Contents

1. Introduction

This assignment will help you understand the concepts learnt in the session.

2. Objective

This assignment will test your skills on Variables & Distributions in R.

3. Prerequisites

Not applicable.

4. Associated Data Files

Not applicable.

5. Problem Statement

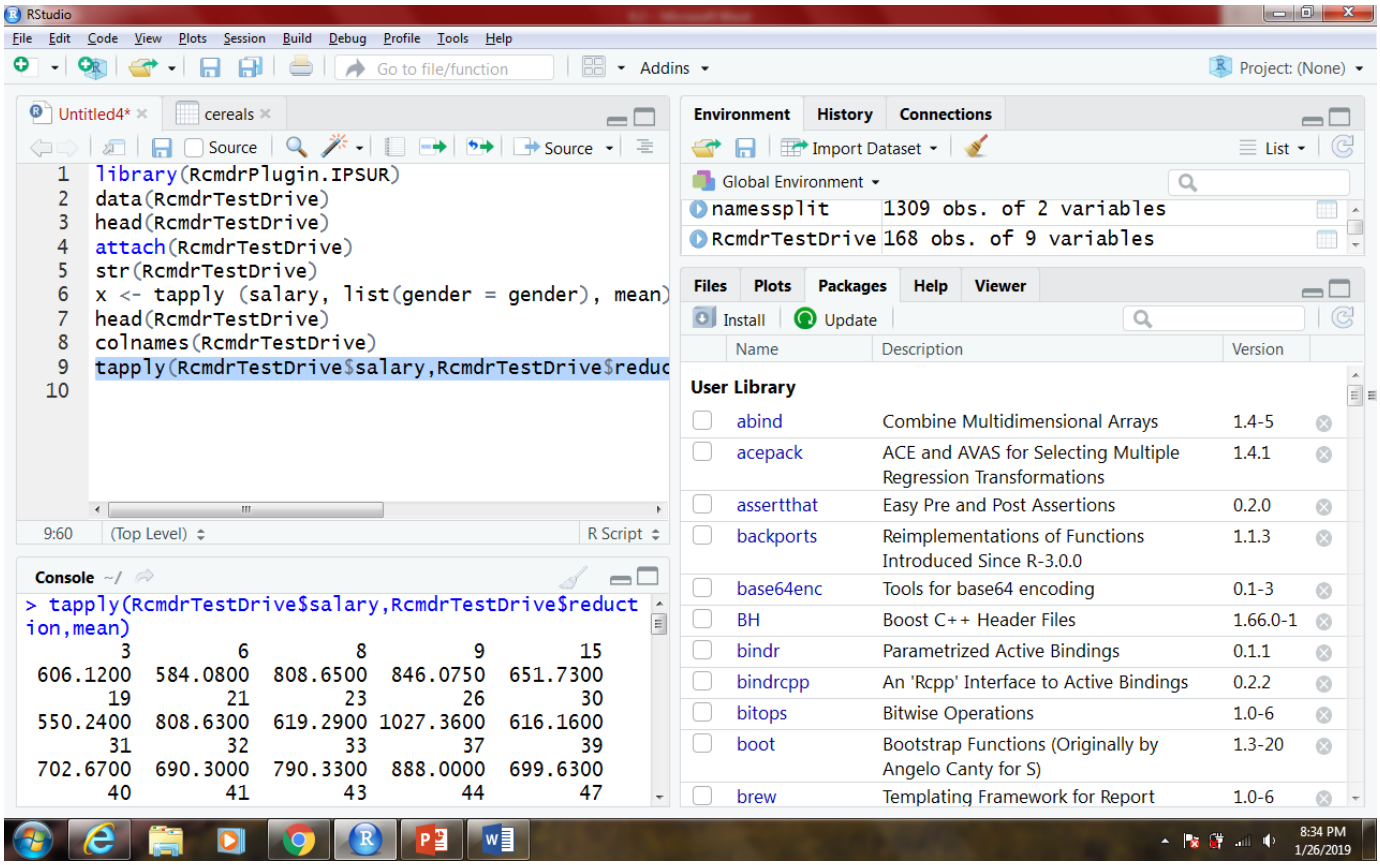
```
library(RcmdrPlugin.IPSUR)
data(RcmdrTestDrive)
```

Perform the below operations: -

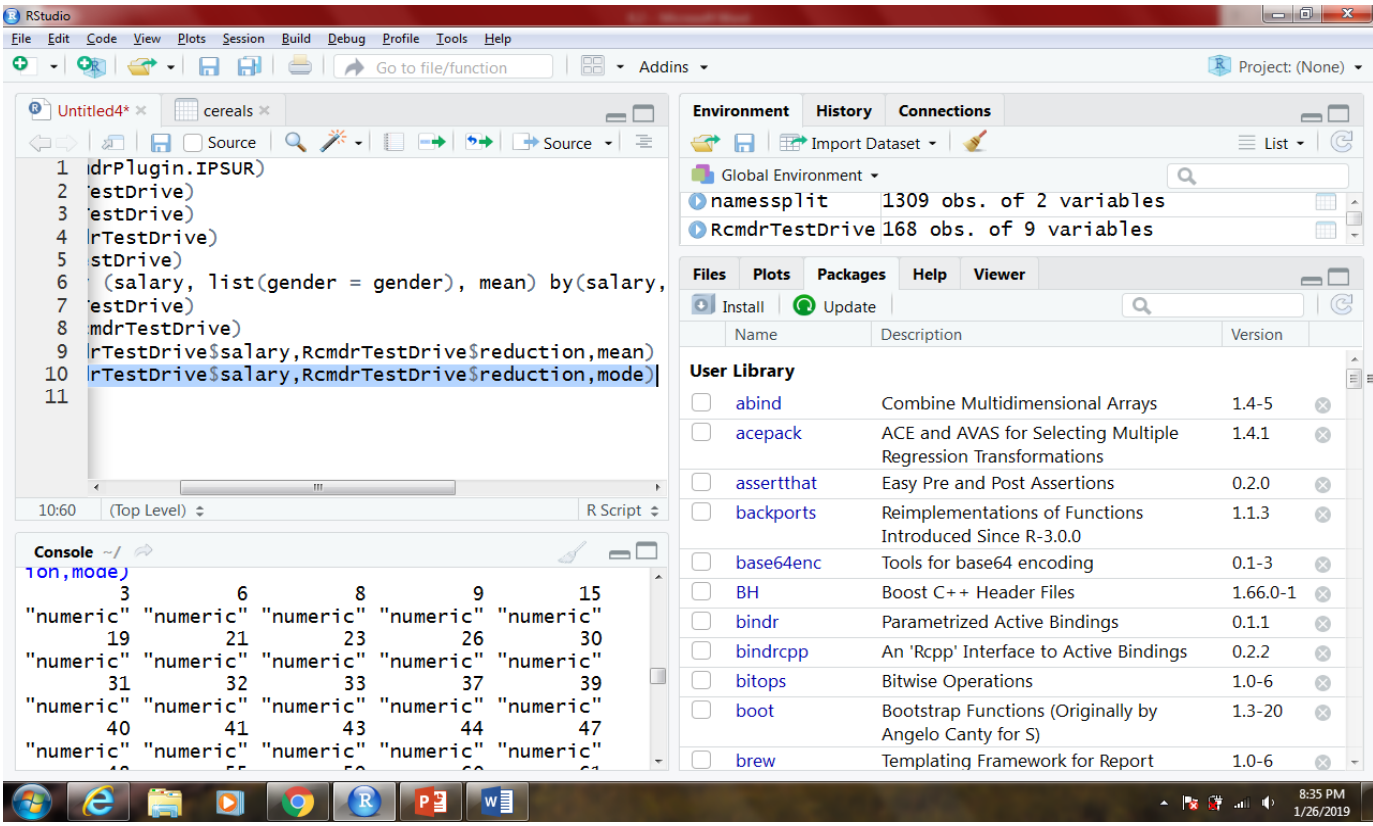
a. Compute the measures of central tendency for salary and reduction which variable has highest center?

Measures of Central Tendency are Mean, Mode and Median

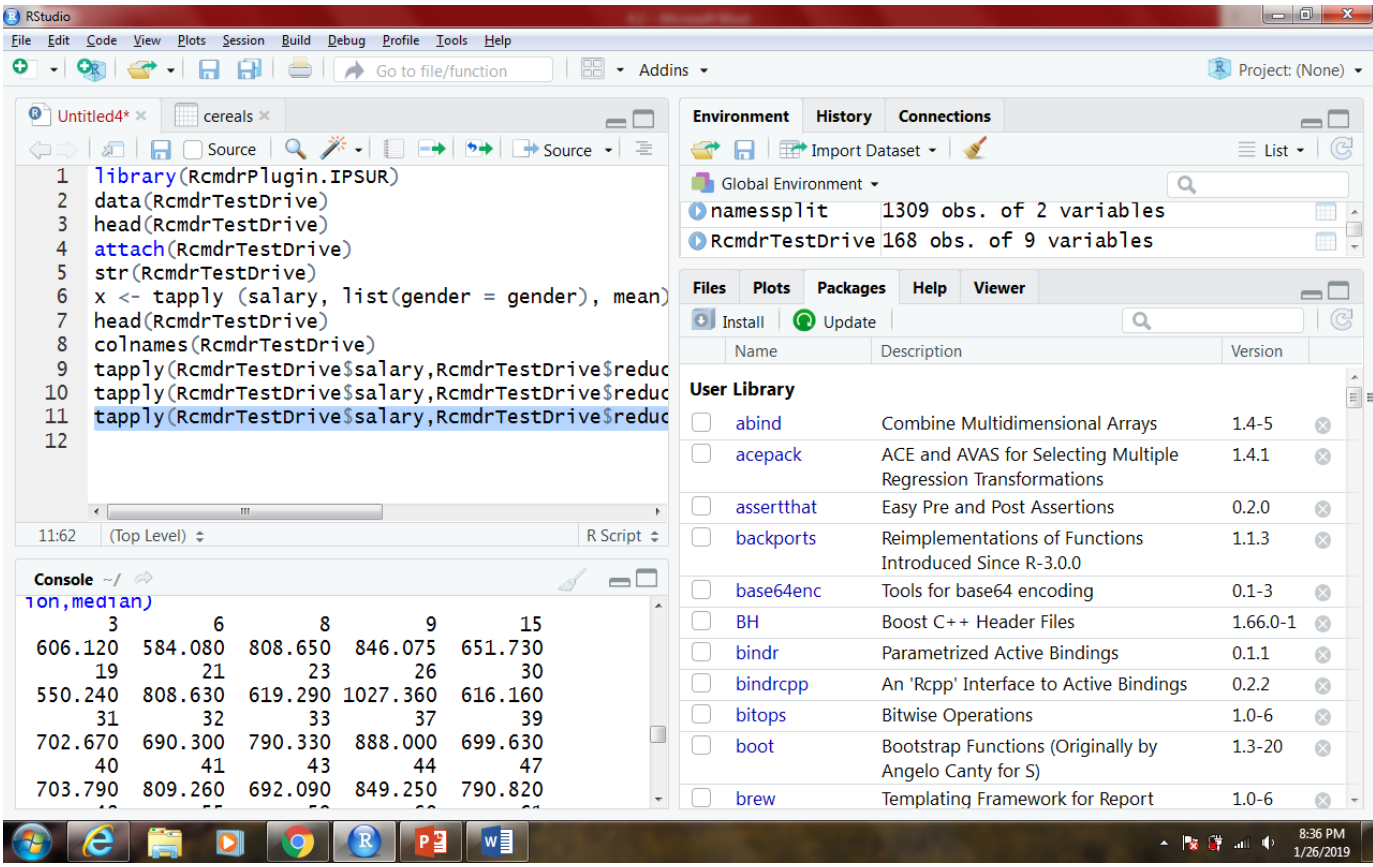
```
tapply(RcmdrTestDrive$salary,RcmdrTestDrive$reduction,mean)
```



tapply(RcmdrTestDrive$salary,RcmdrTestDrive$reduction,mode)

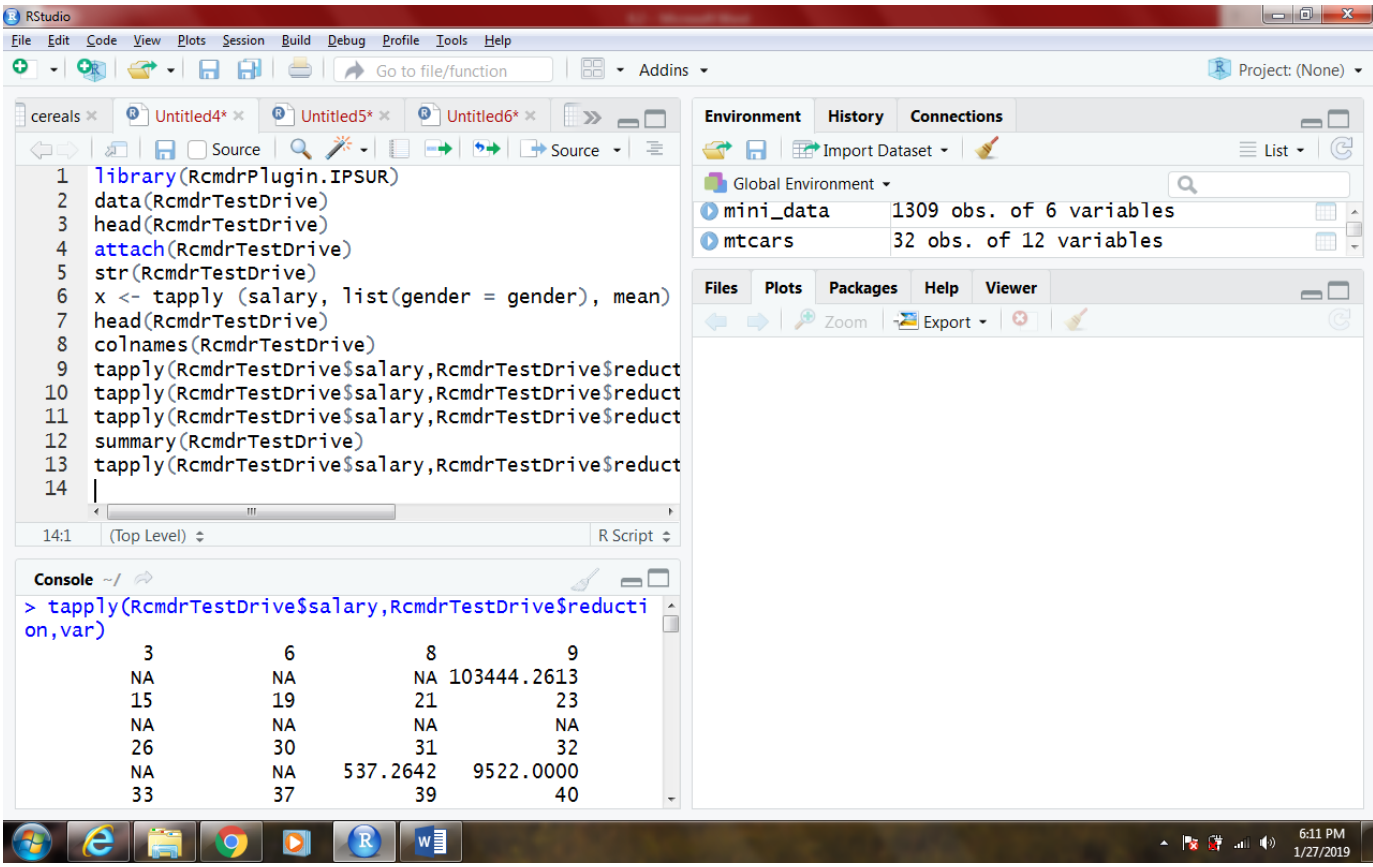


tapply(RcmdrTestDrive$salary,RcmdrTestDrive$reduction,median)



Variance:

tapply(RcmdrTestDrive$salary,RcmdrTestDrive$reduction,var)



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

As data is continuous , mean is the appropriate measure of central tendency.

6. Expected Format

1. R file should be submitted where applicable.
2. R file should be in PDF or in .r format
3. Proper screenshots of the outputs should be submitted as well
4. The r codes, if submitted in any other format, will be subjected to deduction in marks

Note: Your solution will not be entertained if it is any other format, e.g., .zip, .doc, .rtf etc.

7. Approximate Time to Complete Task

30 mins.