**MAT2001 – Statistics for Engineers - ELA (R Code Studio), Winter Semester 2020-2021**

**Lab Assessment - II**

**By: Jonathan Rufus Samuel (20BCT0332) Date: 17.6.2021**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Q1) Write R code to compute mean, median, mode and all measures of dispersions for the following frequency distribution:**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Expenditure** | **0-100** | **100-200** | **200-300** | **300-400** | **400-500** | **500-600** | **600-700** |
| **No. of families** | **10** | **6** | **9** | **16** | **4** | **24** | **27** |

**Also calculate (i) Mean Deviation about mean (ii) Mean Deviation about median (iii) Mean Deviation about mode.**

A: Code is as follows:

**1) Mean, Median & Mode**

#mean

h = 100

x = seq(50,650,h)

f = c(10,6,9,16,4,24,27)

N = sum(f)

Mean = sum(x\*f)/N

Mean

#median

cf = cumsum(f)

mc = min(which(cf>=N/2))

mcf=f[mc]

c=cf[mc-1]

l=x[mc]-h/2

Median=l+(h/mcf)\*((N/2-c))

Median

#mode

mo=which(f==max(f))

mo

f1=f[mo]

f1

f0=f[mo-1]

f0

f2=f[mo+1]

f2

l=x[mo]-h/2

l

#Mode=l+((f1-f0)/(2\*f1-f0-f2))\*h, but f2 does not exist. Therefore, f2 = 0

Mode=l+((f1-f0)/(2\*f1-f0-0))\*h

Mode

**Output (via Command Window):**

> #mean

> h = 100

> x = seq(50,650,h)

> f = c(10,6,9,16,4,24,27)

> N = sum(f)

> Mean = sum(x\*f)/N

> Mean

[1] 435.4167

> #median

> cf = cumsum(f)

> mc = min(which(cf>=N/2))

> mcf=f[mc]

> c=cf[mc-1]

> l=x[mc]-h/2

> Median=l+(h/mcf)\*((N/2-c))

> Median

[1] 512.5

> #mode

> mo=which(f==max(f))

> mo

[1] 7

> f1=f[mo]

> f1

[1] 27

> f0=f[mo-1]

> f0

[1] 24

> f2=f[mo+1]

> f2

[1] NA

> l=x[mo]-h/2

> l

[1] 600

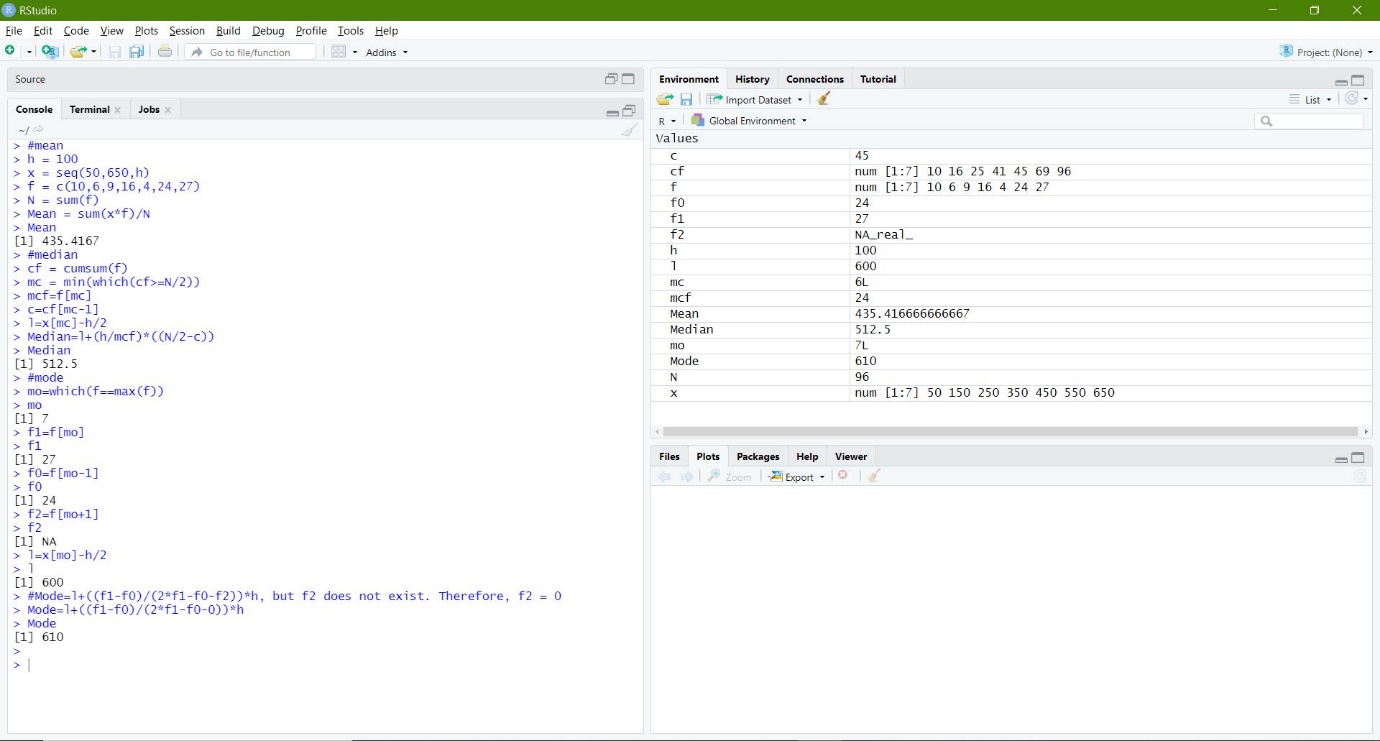
> #Mode=l+((f1-f0)/(2\*f1-f0-f2))\*h, but f2 does not exist. Therefore, f2 = 0

> Mode=l+((f1-f0)/(2\*f1-f0-0))\*h

> Mode

[1] 610

**Implementation on R Studio Code (via Command Window):**



**2) Mean Deviation about Mean (M1), about Median (M2) and Mode (M3)**

#mean

h = 100

x = seq(50,650,h)

f = c(10,6,9,16,4,24,27)

N = sum(f)

Mean = sum(x\*f)/N

Mean

#median

cf = cumsum(f)

mc = min(which(cf>=N/2))

mcf=f[mc]

c=cf[mc-1]

l=x[mc]-h/2

Median=l+(h/mcf)\*((N/2-c))

Median

#mode

mo=which(f==max(f))

mo

f1=f[mo]

f1

f0=f[mo-1]

f0

f2=f[mo+1]

f2

l=x[mo]-h/2

l

#Mode=l+((f1-f0)/(2\*f1-f0-f2))\*h, but f2 does not exist. Therefore, f2 = 0

Mode=l+((f1-f0)/(2\*f1-f0-0))\*h

Mode

#MD about mean = ∑f|x−Mean|/N

MD1 = sum(f\*abs(x - Mean))/sum(f)

MD1

#MD about mean = ∑f|x−Mean|/N

MD2 = sum(f\*abs(x - Median))/sum(f)

MD2

#MD about mean = ∑f|x−Mean|/N

MD3 = sum(f\*abs(x - Mode))/sum(f)

MD3

**Output (via Command Window):**

> #mean

> h = 100

> x = seq(50,650,h)

> f = c(10,6,9,16,4,24,27)

> N = sum(f)

> Mean = sum(x\*f)/N

> Mean

[1] 435.4167

> #median

> cf = cumsum(f)

> mc = min(which(cf>=N/2))

> mcf=f[mc]

> c=cf[mc-1]

> l=x[mc]-h/2

> Median=l+(h/mcf)\*((N/2-c))

> Median

[1] 512.5

> #mode

> mo=which(f==max(f))

> mo

[1] 7

> f1=f[mo]

> f1

[1] 27

> f0=f[mo-1]

> f0

[1] 24

> f2=f[mo+1]

> f2

[1] NA

> l=x[mo]-h/2

> l

[1] 600

> #Mode=l+((f1-f0)/(2\*f1-f0-f2))\*h, but f2 does not exist. Therefore, f2 = 0

> Mode=l+((f1-f0)/(2\*f1-f0-0))\*h

> Mode

[1] 610

> #MD about mean = ∑f|x−Mean|/N

> MD1 = sum(f\*abs(x - Mean))/sum(f)

> MD1

[1] 179.2101

> #MD about mean = ∑f|x−Mean|/N

> MD2 = sum(f\*abs(x - Median))/sum(f)

> MD2

[1] 173.1771

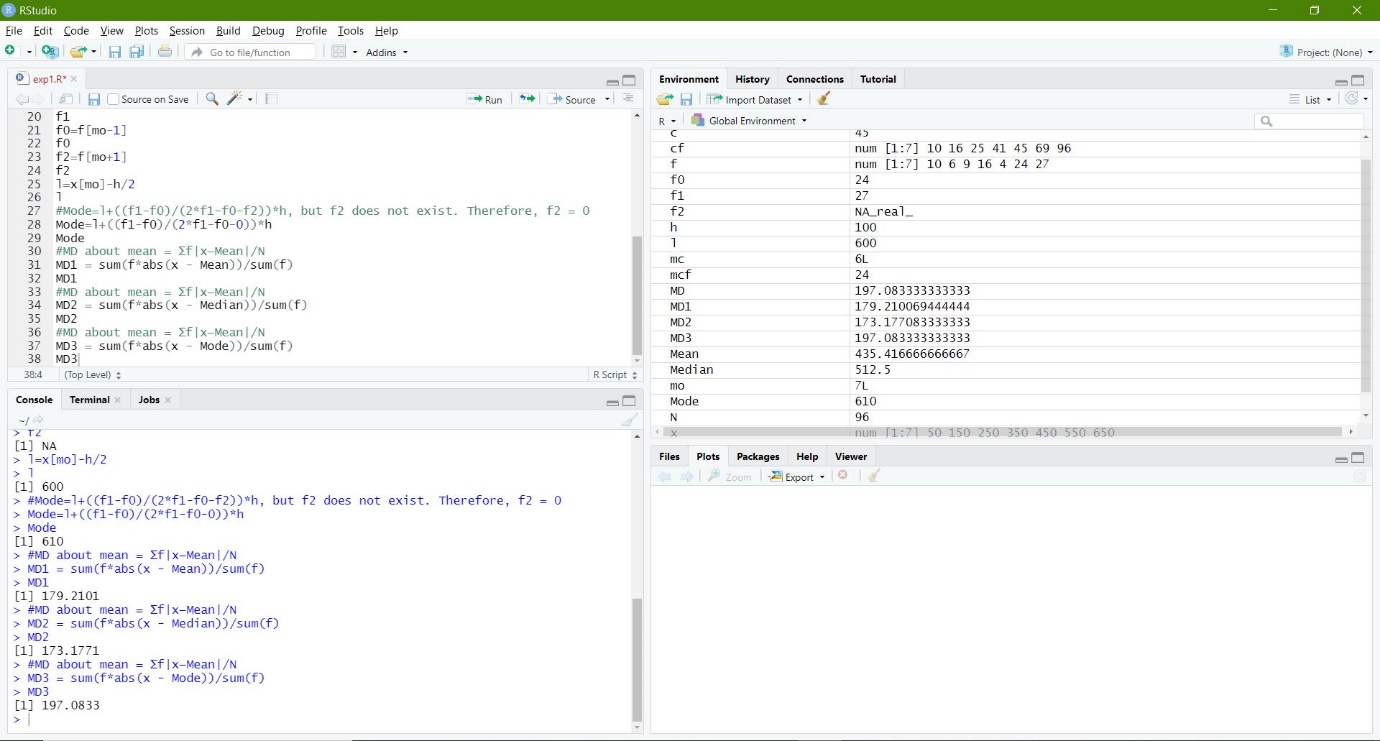
> #MD about mean = ∑f|x−Mean|/N

> MD3 = sum(f\*abs(x - Mode))/sum(f)

> MD3

[1] 197.0833

**Implementation on R Studio Code (via Command Window):**



\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_