

STATISTICS LAB

TASK 1

NAME: SPARSH ARYA
REGISTRATION NUMBER: 17BEC0656
SLOT: E2
TEACHER: SUJATHA V

Measures of central tendency

Problem1: Twenty students, graduates and undergraduates, were enrolled in a statistics course. Their ages were

18,19,19,19,19,20,20,20,20,20,21,21,21,21,22,23,24,27,30,36.

- a) Find Mean and Median of all students**
- b) Find median age of all students under 25 years.**
- c) Find modal age of all students**

R code-:

```
RGui (64-bit) - [R Console]
File Edit View Misc Packages Windows Help
[Icons: File Explorer, RStudio, R Console, R Packages, R Windows, R Help, R Stop, R Print]

R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> x=c(18,19,19,19,19,20,20,20,20,20,21,21,21,21,22,23,24,27,30,36)
> mean(x)
[1] 22
> median(x)
[1] 20.5
> y=x[x<25]
> md=median(y)
> md
[1] 20
> xr=table(x)
> mode=which(xr==max(xr))
> mode
20
3
.
```

Output:-

Mean=22

Median=20

**Mode= 20
3**

Problem 2 :

A survey of 25 faculty members is taken in a college to study their vocational mobility. They were asked the question “In addition to your present position ,at how many educational institutes have served on the faculty?. Following is the frequency distribution of their responses.

<i>X</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>3</i>
<i>f</i>	<i>8</i>	<i>11</i>	<i>5</i>	<i>1</i>

Find mean and median of the distribution.

R code:-

```
R Console
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> x=c(0,1,2,3)
> f=c(8,11,5,1)
> y=rep(x,f)
> mean=(sum(y))/(length(y))
> mean
[1] 0.96
> median(y)
[1] 1
> |
<
```

Output-:

Mean=0.96

Median=1

Problem 3 : Compute mean ,median and mode of for the following frequency Distribution:

<i>Height in Cm</i>	<i>145-150</i>	<i>150-155</i>	<i>155-160</i>	<i>160-165</i>	<i>165-170</i>	<i>170-175</i>	<i>175-180</i>	<i>180-185</i>
<i>No. of Adult men</i>	4	6	28	58	64	30	5	5

R code-:

```
R Console
Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[Previously saved workspace restored]

> mid=seq(147.5,182.5,5)
> mid
[1] 147.5 152.5 157.5 162.5 167.5 172.5 177.5 182.5
> f=c(4,6,28,58,64,30,5,5)
> fr.distr=data.frame(mid,f)
> fr.distr
  mid f
1 147.5 4
2 152.5 6
3 157.5 28
4 162.5 58
5 167.5 64
6 172.5 30
7 177.5 5
8 182.5 5
> mean=(sum(mid*f))/sum(f)
> mean
[1] 165.175
> midx=seq(147.5,182.5,5)
```

```
R Console

> midx=seq(147.5,182.5,5)
> frequency=c(4,6,28,58,64,30,5,5)
> fr.dist<-data.frame(midx,frequency)
> fr.dist
  midx frequency
1 147.5         4
2 152.5         6
3 157.5        28
4 162.5        58
5 167.5        64
6 172.5        30
7 177.5         5
8 182.5         5
> cl=cumsum(frequency)
> cl
[1] 4 10 38 96 160 190 195 200
> n=sum(frequency)
> n
[1] 200
> ml=min(which(cl>=n/2))
> ml
[1] 5
> h=5
> h
[1] 5
```

```
R Console

> h
[1] 5
> f=frequency[ml]
> f
[1] 64
> c=cl[ml-1]
> c
[1] 96
> l=mid[ml]-h/2
> l
[1] 165
> median=l+((n/2)-c)/f*h
> median
[1] 165.3125
> m=which(frequency==max(frequency))
> m
[1] 5
> fm=frequency[m]
> fm
[1] 64
> fl=frequency[m-1]
> fl
[1] 58
> f2=frequency[m+1]
> f2
```

```
R Console
> l=mid[m1]-h/2
> l
[1] 165
> median=l+((n/2)-c)/f)*h
> median
[1] 165.3125
> m=which(frequency==max(frequency))
> m
[1] 5
> fm=frequency[m]
> fm
[1] 64
> fl=frequency[m-1]
> fl
[1] 58
> f2=frequency[m+1]
> f2
[1] 30
> l=midx[m]-h/2
> l
[1] 165
> mode=l+((fm-fl)/(2*fm-fl-f2))*h
> mode
[1] 165.75
> |
```

Output:-

Mean=165.175

Median=165.3125

Mode= 165.75

Measure of dispersion

Problem 4: An entomologist studying morphological variation in species of mosquito recorded the following data on body length:

R code-:

```
R Console

> x=c(1.2,1.4,1.3,1.6,1.0,1.5,1.7,1.1,1.2,1.3)
> x
[1] 1.2 1.4 1.3 1.6 1.0 1.5 1.7 1.1 1.2 1.3
> summary(x)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 1.000  1.200   1.300   1.330   1.475   1.700
> range=1.7-1.0
> range
[1] 0.7
> var(x)
[1] 0.049
> sd=sqrt(var(x))
> sd
[1] 0.2213594
> cqd=(1.475-1.2)/(1.475+1.2)
> cqd
[1] 0.1028037
> y=(x-mean(x))
> y
[1] -0.13  0.07 -0.03  0.27 -0.33  0.17  0.37 -0.23 -0.13 -0.03
> y=abs(y)
> y
[1] 0.13 0.07 0.03 0.27 0.33 0.17 0.37 0.23 0.13 0.03
> mdl=sum(y)/length(y)
```

```
R Console

[1] 0.7
> var(x)
[1] 0.049
> sd=sqrt(var(x))
> sd
[1] 0.2213594
> cqd=(1.475-1.2)/(1.475+1.2)
> cqd
[1] 0.1028037
> y=(x-mean(x))
> y
[1] -0.13  0.07 -0.03  0.27 -0.33  0.17  0.37 -0.23 -0.13 -0.03
> y=abs(y)
> y
[1] 0.13 0.07 0.03 0.27 0.33 0.17 0.37 0.23 0.13 0.03
> mdl=sum(y)/length(y)
> mdl
[1] 0.176
> z=abs(x-median(x))
> z
[1] 0.1 0.1 0.0 0.3 0.3 0.2 0.4 0.2 0.1 0.0
> md2=sum(z)/length(z)
> md2
[1] 0.17
> |
```


Output-:

Variance =0.049

Standard deviation =0.221

Mean deviation = 0.176

Median deviation=0.17