

# EXPERIMENT-1

## CENTRAL TENDENCY AND DATA DISPERSION MEASURES USING R-TOOL

Name: S.G.DEVSACHIN

Reg.No: 192111088

Subject: CSA1672, Data warehouse and data mining

OUTPUT:

```
R Console
> sample <- read.csv("sample.csv", head=TRUE, sep=",")
> sample
  Zip.code Total.population Median.age Total.Males Total.Females Total.Households
1 91371 1 27 0 1 1
2 90001 57110 32 28000 25000 12098
3 90002 66266 41 21000 23000 12345
4 90003 62180 26 32000 23000 25672
5 90004 37681 27 13000 17000 10021
6 90005 58185 32 19000 18213 13486
7 90006 40920 37 38000 26444 16293
8 90007 32327 43 29000 10000 17303
9 90008 3800 22 1874 25002 13092
10 90009 103892 28 19456 31111 19363
11 90010 31103 30 26011 27211 20734
12 90011 11772 24 1367 19992 4167
13 90012 7005 25 10000 17322 2014
Average.households
1 1
2 4
3 4
4 2
5 2
6 3
7 4
8 2
9 1
10 4
11 2
12 3
13 1
> mean(sample$Total.Males)
[1] 18362.15
> median(sample$Total.Males)
[1] 19456
> mode(sample$Total.Males)
[1] "numeric"
> IQR=(sample$Total.Males)
> IQR(sample$Total.Males)
[1] 18000
> quantile(sample$Total.Males,0.25)
[25%
```

```
R Console
9          1
10         4
11         2
12         3
13         1
> mean(sample$Total.Males)
[1] 19362.15
> median(sample$Total.Males)
[1] 19456
> mode(sample$Total.Males)
[1] "numeric"
> IQR=(sample$Total.Males)
> IQR(sample$Total.Males)
[1] 18000
> quantile(sample$Total.Males,0.25)
25%
10000
> quantile(sample$Total.Males,0.75)
75%
28000
> range(sample$Total.Males)
[1] 0 38000
> mean(range(sample$Total.Males))
[1] 19000
> Lf<-quantile(sample$Total.Males,0.25)-1.5(IQR(sample$Total.Males)
+ )
Error: attempt to apply non-function
> Lf<-quantile(sample$Total.Males,0.25)-1.5(IQR(sample$Total.Males))
Error: attempt to apply non-function
> Lf<-quantile(sample$Total.Males,0.25)-1.5*(IQR(sample$Total.Males))
> print(Lf)
25%
-17000
> uf<-quantile(sample$Total.Males,0.25)+1.5*(IQR(sample$Total.Males))
> print(uf)
75%
37000
> outlier_values<-boxplot.stats(sample$Total.Males)$out
> print(outlier_values)
integer(0)
> |
```