#to read a csv file from excel

> data=read.csv(file.choose(),header=T)

> data

X State Literacy Male Female X..Change

1 - India 74.04 82.14 65.46 8.66

2 1 Kerala 94.00 96.11 92.07 3.14

3 2 Lakshadweep 91.85 95.56 87.95 5.19

4 3 Mizoram 91.33 93.35 89.27 2.53

5 4 Goa 88.70 92.65 84.66 6.69

6 5 Tripura 87.22 91.53 82.73 14.03

7 6 Daman and Diu 87.10 91.54 79.55 8.92

8 7 Andaman and Nicobar Islands 86.63 90.27 82.43 5.33

9 8 Delhi 86.21 90.94 80.76 4.54

10 9 Chandigarh 86.05 89.99 81.19 4.11

11 10 Puducherry 85.85 91.26 80.67 4.61

12 11 Himachal Pradesh 82.80 89.53 75.93 6.32

13 12 Maharashtra 82.34 88.38 75.87 5.46

14 13 Sikkim 81.42 86.55 75.61 12.61

15 14 Tamil Nadu 80.09 86.77 73.44 6.64

16 15 Nagaland 79.55 82.75 76.11 12.96

17 16 Uttarakhand 78.82 87.40 70.01 7.20

18 17 Gujarat 78.03 85.75 69.68 8.89

19 18 Manipur 76.94 83.58 70.26 10.33

20 19 West Bengal 76.26 81.69 70.54 7.62

21 20 Dadra and Nagar Haveli 76.24 85.17 64.32 18.61

22 21 Punjab 75.84 80.44 70.73 6.19

23 22 Haryana 75.55 84.06 65.94 7.64

24 23 Karnataka 75.36 82.47 68.08 8.72

25 24 Meghalaya 74.43 75.95 72.89 11.87

26 25 Orissa 72.87 81.59 64.01 9.79

27 26 Assam 72.19 77.85 66.27 8.94

28 27 Chhattisgarh 70.28 80.27 60.24 5.62

29 28 Madhya Pradesh 69.32 78.73 59.24 5.58

30 29 Uttar Pradesh 67.68 77.28 57.18 11.41

31 30 Jammu and Kashmir 67.16 76.75 56.43 11.64

32 31 Andhra Pradesh 67.02 74.88 59.15 6.55

33 32 Jharkhand 66.41 76.84 55.42 12.85

34 33 Rajasthan 66.11 79.19 52.12 5.70

35 34 Arunachal Pradesh 65.38 72.55 57.70 11.04

36 35 Bihar 61.80 71.20 51.50 14.80

> names(data)

[1] "X" "State" "Literacy" "Male" "Female" "X..Change"

> #mean

> mean = mean(data$Male)

> cat("the mean of male is",mean)

the mean of male is 84.24889

> #median

> median = median(data$Male)

> cat("the median of male is",median)

the median of male is 83.82

> #standard deviation

> standard\_deviation=sd(data$Male)

> cat("the standard deviation value of male is",standard\_deviation)

the standard deviation value of male is 6.624113

> #variance

> variance = var(data$Male)

> cat("the variance of male is",variance)

the variance of male is 43.87887

> #skewness

> numerator = 3\*(mean-median)

> value = numberator/standard\_deviation

> value = numerator/standard\_deviation

> cat("the skewness value is",value)

the skewness value is 0.1942398

> #mean deviation

> mean = mean(data$Male)

> column = data$Male - mean

> mean\_dev = mean(column)

> cat("the mean deviation of Male is",mean\_dev)

the mean deviation of Male is 3.552714e-15

> #geometric mean:

> geo\_mean <- function(data){}

> geo\_mean <- function(data){

+ log\_data <- log(data)

+ gm <- exp(mean(log\_data[is.finite(log\_data)]))

+ return (gm)

+ }

> geometric\_mean = geo\_mean(data$Male)

> cat("the geometric mean value of male is",geometric\_mean)

the geometric mean value of male is 83.99379

> #range

> range\_value = range(data$Male)

> cat("the range value of male is",range\_value)

the range value of male is 71.2 96.11

> #nth percentile

> Male = data$Male

> percentile = quantile(Male, c(.32, .57, .98))

> cat("the percentile value of male is",percentile)

the percentile value of male is 80.67 85.721 95.725

> #first and second quartile

> first = quantile(data$Male,0.25)

> second = quantile(data$Male,0.5)

> cat("the first and second quartile values are",first,second)

the first and second quartile values are 79.075 83.82

> #quartile deviation

> first = quantile(data$Male,0.25)

> third = quantile(data$Male,0.75)

> quartile\_deviation= (third-first)/2

> cat("the quartile deviation value of matches is",quartile\_deviation)

the quartile deviation value of matches is 5.4925

> #any two deciles

> first = quantile(data$Male,0.1)

> second = quantile(data$Male,0.2)

> cat("the two deciles value of male is",first,second)

the two deciles value of male is 76.35 77.85

> #sum of column values

> sum = sum(data$Male)

> cat("the sum of values of male column is",sum)

the sum of values of male column is 3032.96

> #minimum of coulmn values

> minimum = min(data$Male)

> cat("the minimum values of male column is",minimum)

the minimum values of male column is 71.2

> #maximum of column values

> maximum = max(data$Male)

> cat("the maximum value of male column is",maximum)

the maximum value of male column is 96.11

> #harmonic mean

> col = data$Male

> sum = 0

> for(i in col){

+ val = (1/i)

+ sum = sum + val

+ }

> numerator = length(data$Male)

> harmonic\_mean = (numerator/sum)

> cat("the harmonic mean is",harmonic\_mean)

the harmonic mean is 83.73708

|  |
| --- |
| > #to find binomial distribution  > #n = 324,p=0.5  > dbinom(20,324,0.5) #p(x=20)  [1] 1.07426e-66  > #we want p(20<=x<=25)  > dbinom(2-:25,30,0.5)  > dbinom(20:25,30,0.5)  [1] 0.0279816007 0.0133245718 0.0054509612 0.0018959865 0.0005529961 0.0001327191  > sum(dbinom(20:25,30,0.5))  [1] 0.04933884  > #cdf  > pbinom(25,30,0.5)  [1] 0.9999703 |
|  |
| |  | | --- | |  | |