Group 2 (DA02)Theory Assignment(TA03)

1)
Standard deviation - 0
Mean - c
Mode - c

2)
(a) new mean=((54*15)+70)/16=55
(b) new mean=((54*15)-38)/14=55.14
(c) new mean=((54*15)+(56*3))/18=54.33

3)
(a)median=(3+5)/2=4
mode=5
(b)median=75
mode=70

4) (a)GM=74.67
(b)GM=34249.04

5)
X=Mean(BMI)=24.5
Y=Mean(BP)=143
Z=Mean(BMI*BP)=3560.33
Covariance=Z-X*Y=56.83

6) L 1 = 20, N = 3194,sigma(freq.) = 950, f req median = 1500, width = 30, median= 32.94 years.

7)
(a) mean of the data is: $\bar{x}$ = 1/n sigma
x i = 809/27 = 30 .
The median : 25.

(b) data set has two values that occur with the same highest frequency and is, therefore, bimodal.

The modes of the data are 25 and 35.

(c) The midrange of the data is: (70 + 13)/2 = 41.5

(d) The first quartile of the data is: 20.
The third quartile of the data is: 35.

(e) The five number summary of a distribution consists of the minimum value, first quartile, median value, third quartile, and maximum value. It provides a good summary of the shape of the distribution and for this data is: 13, 20, 25, 35, 70.

8)
• Count: The current count can be stored as a value, and when x number of new values are added,
we can easily update count with count + x. This is a distributive measure and is easily updated for
incremental additions.

• Standard deviation: If we store the sum of the squared existing values and the count of the existing
values, we can easily generate the new standard deviation using the formula provided in the book.
We simply need to calculate the squared sum of the new numbers, add that to the existing squared
sum, update the count of the numbers, and plug that into the calculation to obtain the new standard
deviation. All of this is done without looking at the whole data set and is thus easy to compute.

• Median: To accurately calculate the median, we have to look at every value in the dataset. When we
add a new value or values, we have to sort the new set and then find the median based on that new
sorted set. This is much harder and thus makes the incremental addition of new values difficult.

9) Only mean is distributive and algebraic while median and mode are Hollistic.

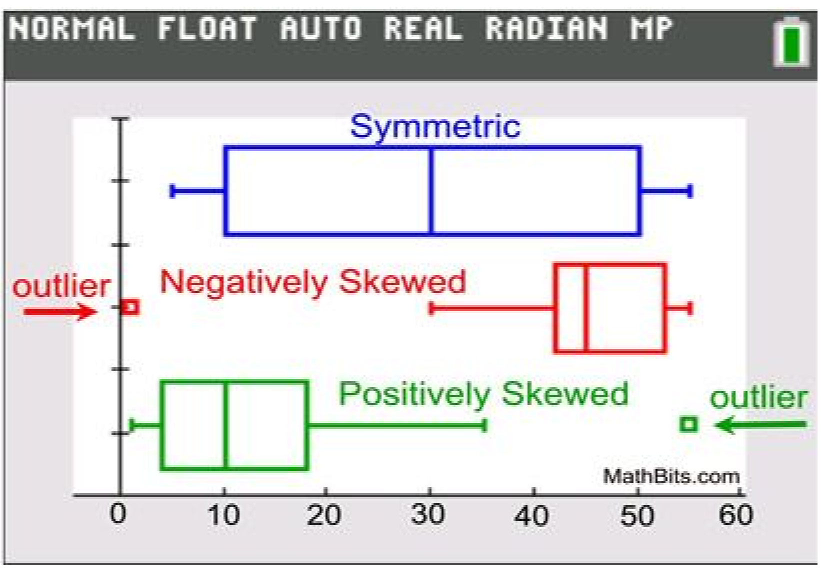
10)
AM is used when the sample elements are uniformly(linearly) distributed.

GM is an appropriate measure when values change exponentially and in case of skewed distribution that can be made symmetrical by a log transformation. GM is more commonly used in microbiological and serological research
HM is appropriate in situations where the reciprocals of values are more useful. HM is used when we want to determine the average sample size of a number of groups, each of which has a different sample size.
11)

let the line segment be ax+by+c with from x1 to x2 (x2>x1)

mean: m/2(x2-x1)+c
median: (m/2 (x2-x1)(x2+x1) +c(x2-x1)) / (2(c+m/2(x2-x1))
mode: x with maximum y
m>0 y2, if m <0 y1, if m=0 any x

12)
A symmetric distribution is one where the left and right hand sides of the distribution are roughly equally balanced around the mean. The mean is approximately equal to the median
A distribution that is skewed left
1. the mean is typically less than the median
2. the tail of the distribution is longer on the left hand side than on the right hand side
3. the median is closer to the third quartile than to the first quartile.
13)



14)
The standard deviation calculated with a divisor of n-1 is a standard deviation calculated from the sample as an estimate of the standard deviation of the population from which the sample was drawn. Because the observed values fall, on average, closer to the sample mean than to the population mean, the standard deviation which is calculated using deviations from the

sample mean underestimates the desired standard deviation of the population. Using n-1 instead of n as the divisor corrects for that by making the result a little bit bigger.
Note that the correction has a larger proportional effect when n is small than when it is large, which is what we want because when n is larger the sample mean is likely to be a good estimator of the population mean.

15)
Yes,standard deviation of sample can be zero if value of all members in sample is same.
Mean will be equal to the value of any member of such sample.
Example:- sample of values (2,2,2,2,2,2,2,2,2,2)

16)
Standard deviation will achieve its maximum value when half members of sample will have minimum possible value and other half members will have maximum possible value.
Example- sample of values (50,50,50,50,100,100,100,100)

17)
A.course =c2
B.course =c3
C.course =c2

SECTION 2:
1. A
2. D
3. C
4. A
5. A
6. C
7. C
8. D
9 B
10. B
11. C
12. C
13. B
14. B
15. C
16. A
17. B