[*https://dsft.code-data-ai.com/stats-1/*](https://dsft.code-data-ai.com/stats-1/)

*1. Given, there are 5 numbers in the data set: (8, 12, 16, 24, 4).*

*What will be the sum of deviations of individual data points from their mean ?*

I would first calculate the mean:

mean = (8 + 12 + 16 + 24 + 4) / 5 = 12.8

Then I would subtract the mean from each individual score to find the individual deviations individually.

Sum of deviations = (12.8 - 8) + (12.8 - 12) + (12.8 - 16) + (12.8 - 24) + (12.8 - 4)

4,8 + 0,8 +( - 3,2) + (-11,2) + 8,8 = 0

**Sum of deviations = 0**

*2. If some* ***outliers*** *are introduced to the dataset, what will happen to the* ***Standard Deviation*** *?*

*A) Standard Deviation is robust to outliers B) Standard Deviation will increase with the introduction of outliers. C) Standard Deviation will decrease with the introduction of outliers. D) Can not be determined.*

outliers = a data point that differs significantly from other observations.

***B) Standard Deviation will increase with the introduction of outliers.***

*3. Suppose the below* ***positively skewed distribution*** *has a* ***median*** *of* ***30****, which of the following statement is true?*

*A) Mean is greater than 30*

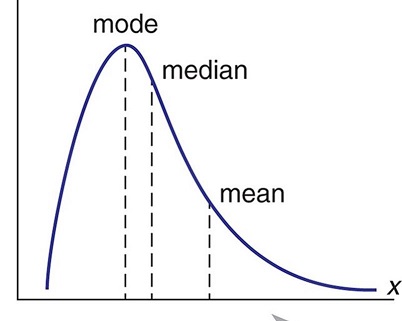
*B) Mean is less than 30*

*C) Mode is greater than 30*

*D) Mode is less than 30*

*E) Both A and D*

*F) Both B and C*



**Option E - Both A and D**

A positiv skewed distribituion as below may be represented like this:

Mean > Median > Mode

That means:

Mean > 30

Median = 30

Mode < 30

*4. Which value can be the possible value for the median of the below distribution?*

**B) 26**

10 to 15 = 36

15 to 20 = 54

20 to 25 = 69

25 to 30 = 82

30 to 35 = 55

35 to 40 = 43

40 to 45 = 25

45 to 50 = 22

50 to 55 = 17

403 / 2 = 202

til 20 = 36 + 54 = 90

til 25 = 90 + 69 = 159

til 30 = 159 + 82 = 241

**Median is between 25 and 30**

*5. What is the shape of the distribution ?*

**The graphic shows a positive skewed distribution**

*6. What would you consider to be the most appropriate measure of the center for this data?*

**Median**

*7. If Y axis represents the number of individuals and X axis – salary of the individual in thousands. How many individuals have salary less than 10 thousands ?*

*I would join the number of individuals in the interval from 0 to 5 and the intervale 5 to 10.*

*0 to 5 = 11*

*5 to 10 = 24*

*salaries 0 to 10 = 11 + 24 =* ***35***

*8. We have a set of positive numbers. If a single value of the set is altered what must change ?*

***A) Mean*** *B) Median C) Mode D) All of these*

*9. The chart shows hourly consultancy rate of 10 people.*

*Calculate the standard deviation of the salaries of the 10 employees.*

*First the mean:*

*mean = (25 + 25 + 25 + 40 + 40 + 35 + 50 + 50 + 50 + 50) / 10 = 39*

*Variance = ((39-25)\*\*2) + ((39-25)\*\*2) + ((39-25)\*\*2) + ((39-40)\*\*2) + ((39-40)\*\*2) + ((39-35)\*\*2) + ((39-50)\*\*2) + ((39-50)\*\*2) + ((39-50)\*\*2) + ((39-50)\*\*2) / (10-1) =*

*(196 + 196 + 196 + 1 + 1 + 16 + 121 + 121 + 121 + 121) / (10-1) = 121,11*

***standard deviation = Square root of Variance = 11***

*10. Which of the following random variables is discrete?*

A **discrete** distribution is one in which the data can only take on certain values, for **example** integers.

*A) the length of time a battery lasts* **>>> Continuous**

*B) the number of pens purchased by a student in a year* **>>> Discrete**

*C) the percentage of cows in a cattle firm that have been vaccinated*

**>>> Continuous**

*D) the distance between a pair of towns* **>>> Continuous**

*11. Which of the below normal distributions will have the greatest spread?*

***mu= mean***

***sigma = standard deviation***

*A) mu=5, sigma =1.5*

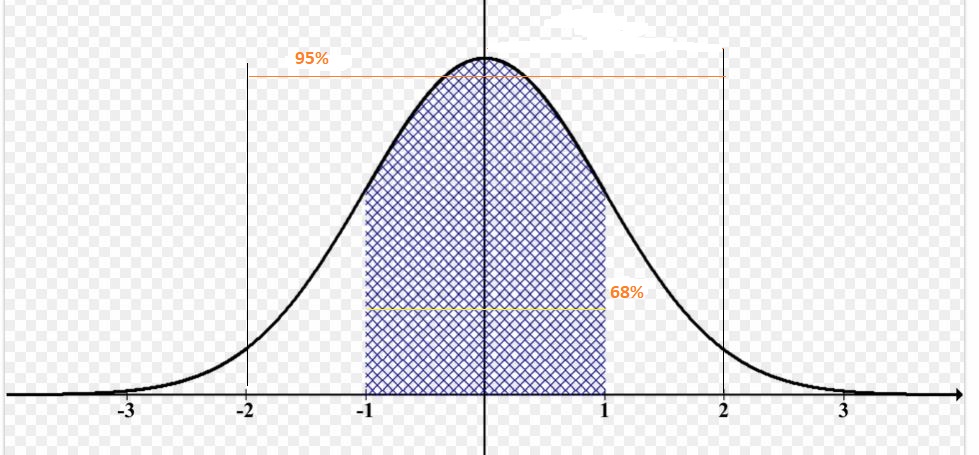
*B) mu=10, sigma =1.0*

*C) mu=5, sigma =1.65* ***>>> this one because it has the biggest standard deviation and the value on the mean does not affect the spread.***

*D) mu=8, sigma =1.2*

*E) mu=10, sigma =1.6*

*12. For a normal distribution with mu=10 & sigma =1.4, about 2.5% of the values lie above what value? (Assume that the number is above the mean value)*

**

*sigma 2 to 3 = 10 + 2 sigma (1.4 + 1.4) =* ***12.8***