

**NAME: Muhammad Azam Khan**

**Subject: BASIC STATISTICS**

**Submitted to: RAZA AHMED KHAN**

**ID: 13130**

**ASSIGNMENT**

**SUMMER-2020**

**Program: B.B.A / MMC**

**Course Title: BASIC STATISTICS / STATISTICS-I**

**Dated: 26th September, 2020**

**Instructor: Raza Ahmed Khan**

****

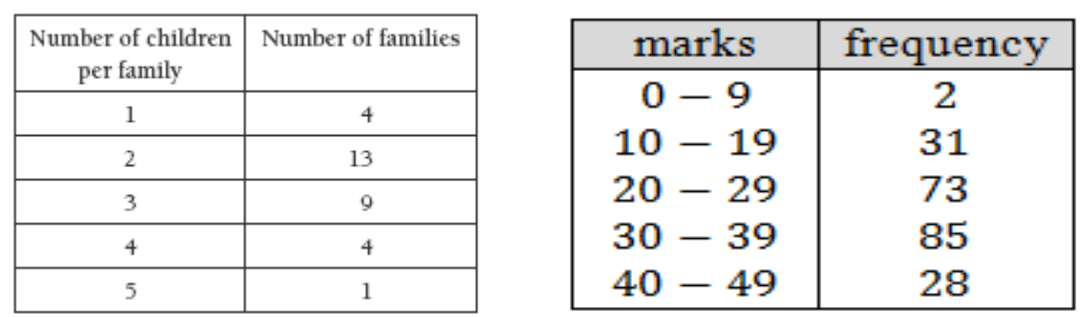
**Important Instructions:**

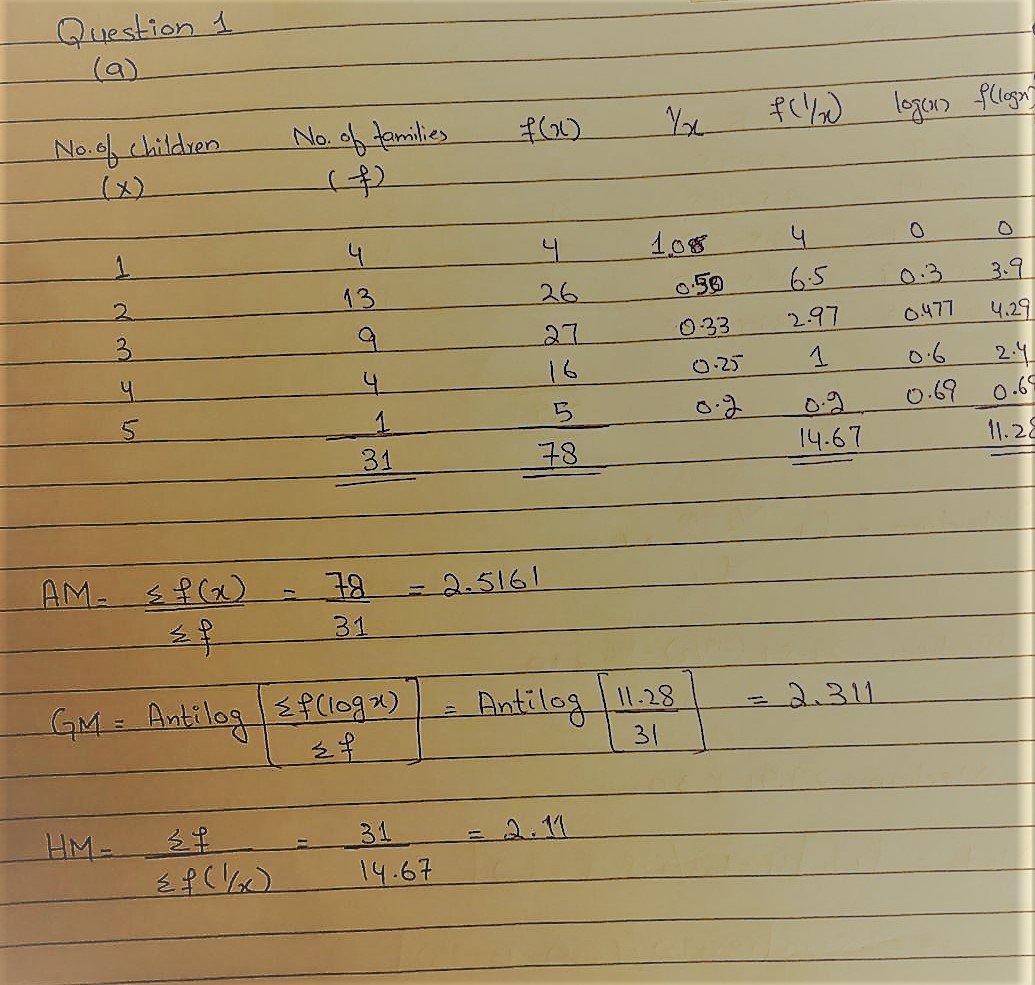
* **Assignment should be submitted within 4 – Hours.**
* **Submitted Document’s format should be in word, pdf or in jpg.**
* **No Assignment will be accepted after due date mentioned above.**
* **Attempt All Questions**

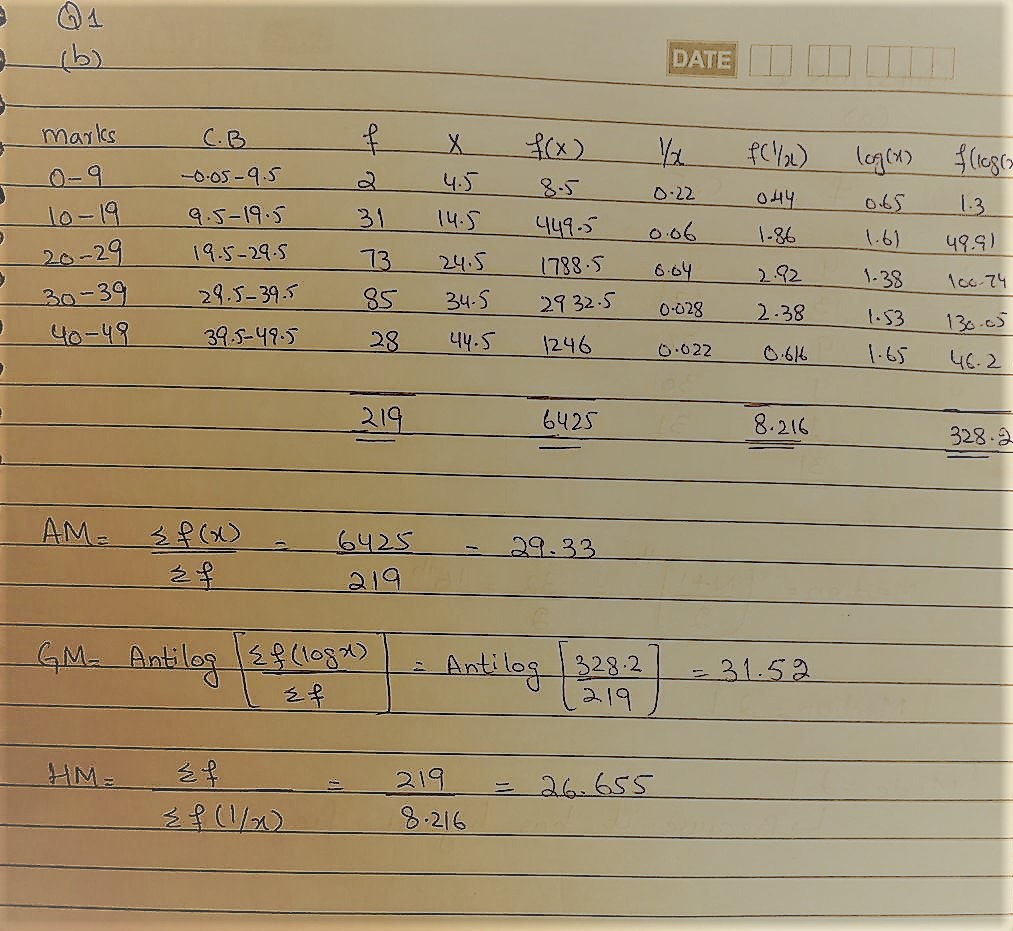
**Question No: 01** **12-MARKS**

**Find averages (A.M, G.M, H.M) of the following table (s) also justify their logical relationships for the following both of sections.**

**a.** **b.**

****

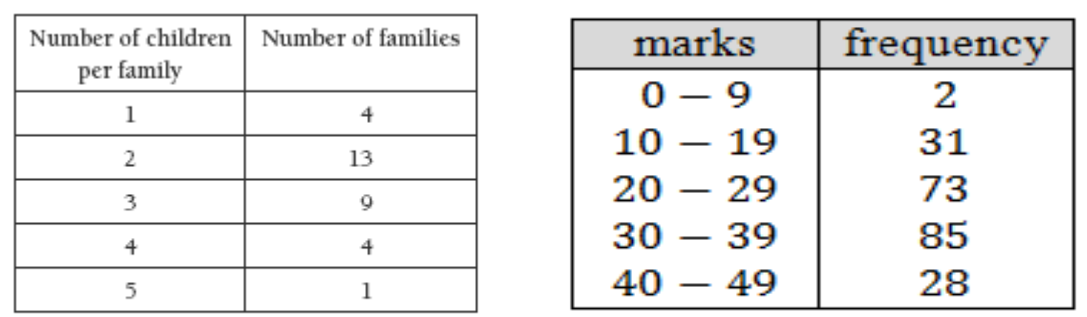
****

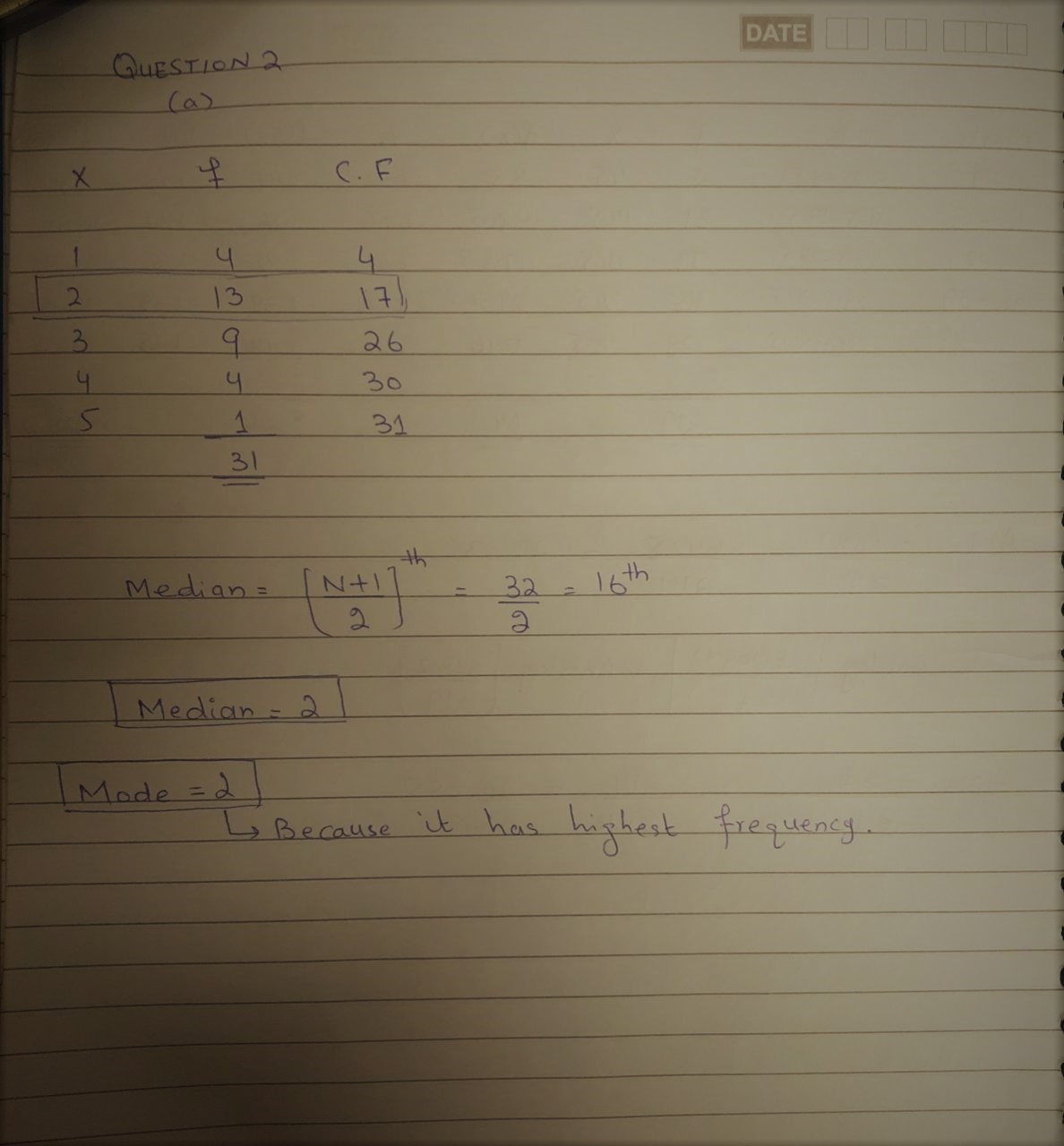
****

**Question No: 02** **14-MARKS**

**Find Median & Mode of the following tables also illustrate their results.**

**a.** **b.**

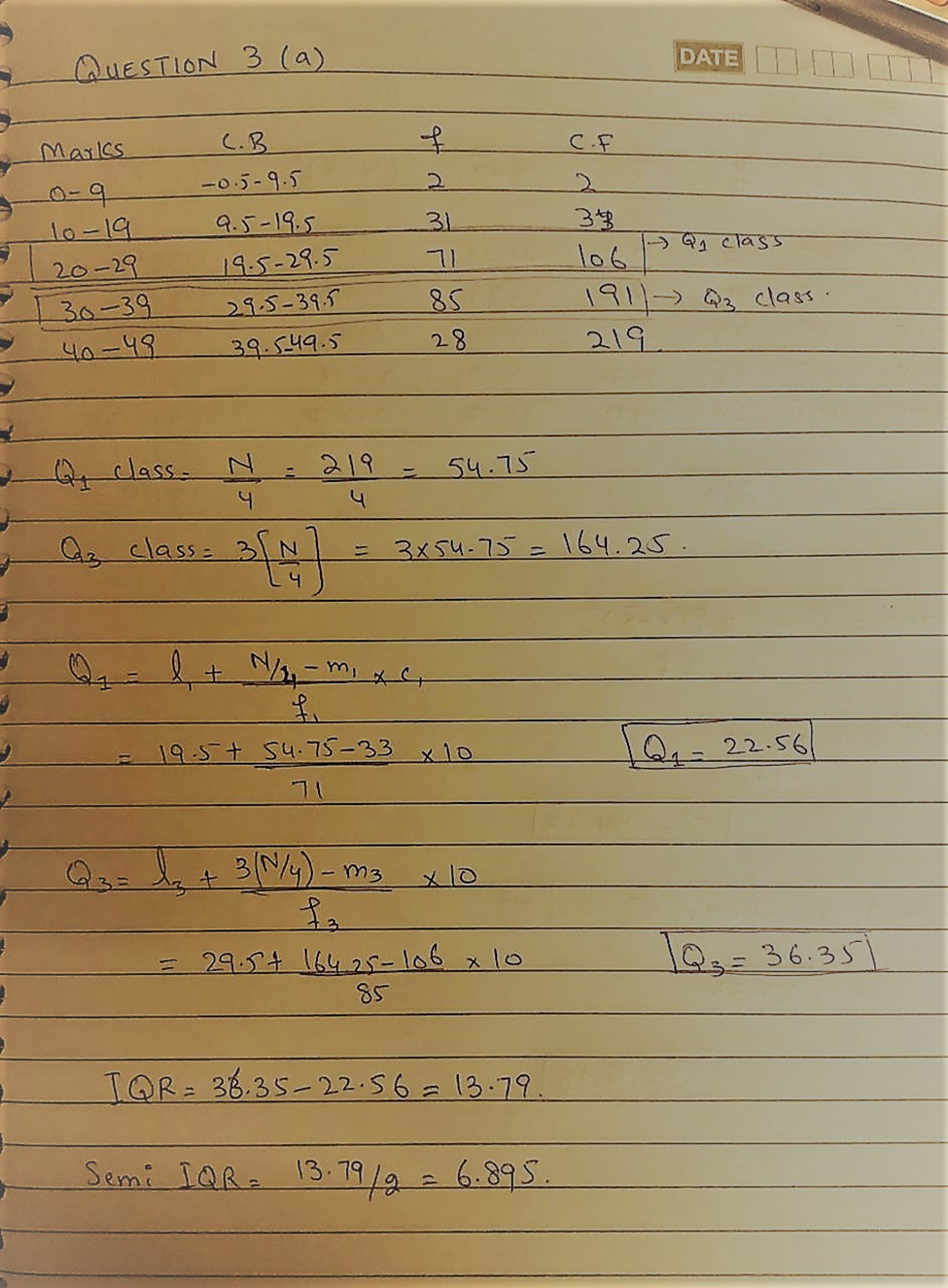
****

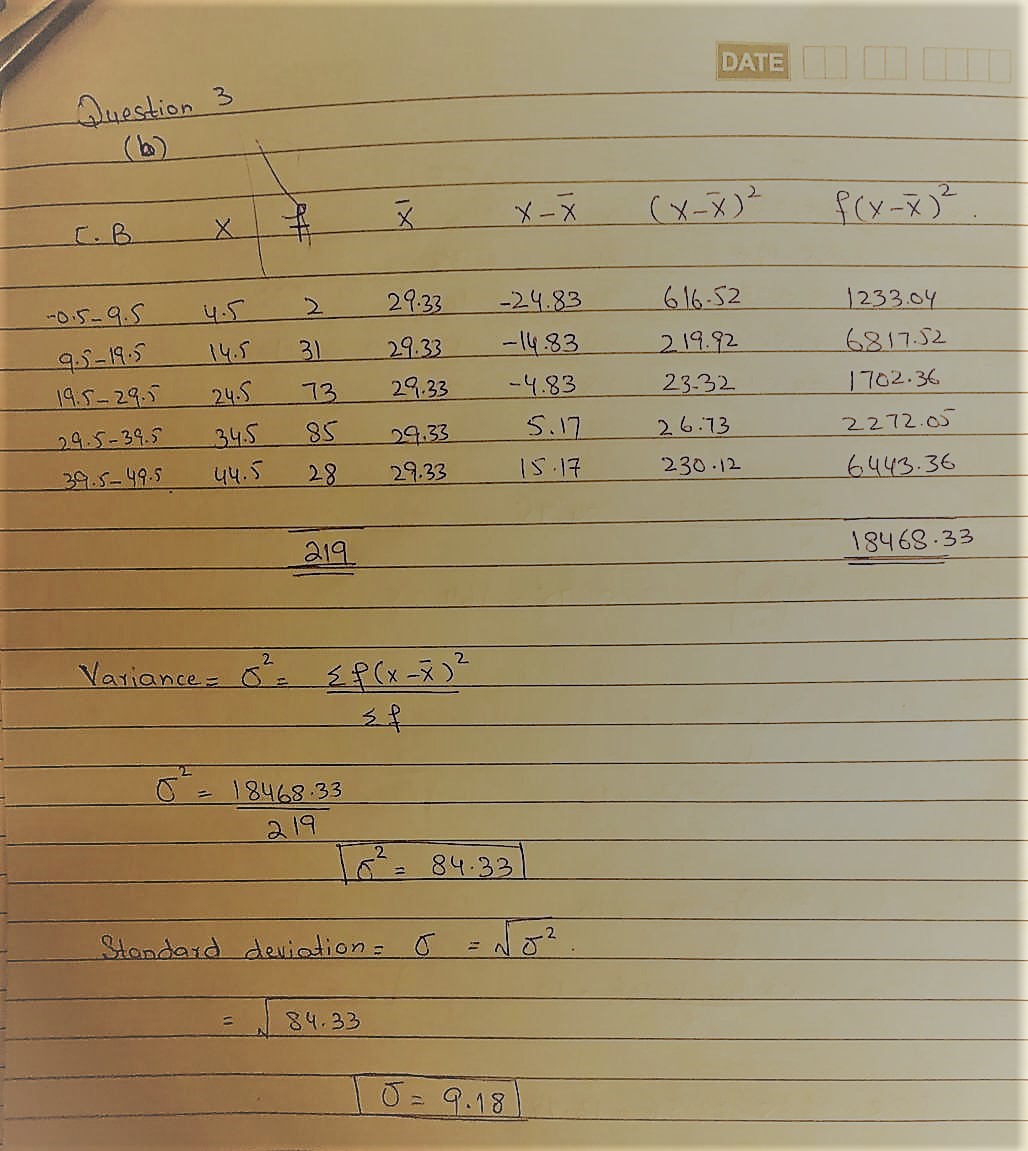
****

****

**Question No: 03** **12-MARKS**

1. Find Semi Quartile Range & Semi Inter Quartile Range of Q2(b)
2. Find Variance and Co-efficient of variance of Q2(b)



****

**Question No: 04** **12-MARKS**

Differentiate the followings with valid examples

* Range & Quartile Range & Semi Inter Quartile Range
* Variance & Standard Deviation &Coefficient of Variation

**Range:**

The range of a set of observation is the difference between two extreme values, i.e. the difference between the maximum and minimum values. Therefore, it indicates the limits within all observations fall.

**In the form of an equation: Range = Highest value – Lowest value.**

**OR**

The range is defined as the difference between the largest score in the set of data and the smallest score in the set of data, XL – XS.

Range is used when you have ordinal data or you are presenting your results to people with little or no knowledge of statistics. It is rarely used in scientific work as it is fairly insensitive it depends on only two scores in the set of data, XL and XS.

**Example:**

What is the range of the following data: 4 8 1 6 6 2 9 3 6 9

The largest score (XL) is 9; the smallest score (XS) is 1; the range is XL - XS = 9 - 1 = 8

Two very different sets of data can have the same range: 1 1 1 1 9 vs 1 3 5 7 9

**Range for Ungrouped data:**

Let us consider a set of observations 𝑥1, 𝑥2, 𝑥3……………, 𝑥 𝑛 and 𝑋 𝐻 is Maximum and 𝑋 𝐿 is Minimum.

Then Range = 𝑋 𝐻 − 𝑋 𝐿.

Find out the range of the set of observations, -7, -2, -4, 0, 8.

**Solution:**

Here, maximum value, 𝑋 𝐻 = 8 and minimum value, 𝑋 𝐿 = −7 Range = 𝑋 𝐻 − 𝑋 𝐿 = 8 − −7 = 8 + 7 = 15

**Range for Grouped data:**

In this case, the range is the difference between the upper boundary of the highest class and the lower boundary of the lowest class.

Then Range = 𝑋 𝑈 − 𝑋 𝐿

Where, 𝑋 𝑈= the upper boundary of the highest class. 𝑋 𝐿= the lowest boundary of the highest class.

**Example:** Determine the range from the following frequency distribution.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Salary (TK.) | 1700-1800 | 1800-1900 | 1900-2000 | 2000-2100 | 2100-2200 |
| No. of workers | 420 | 460 | 500 | 300 | 200 |

**Solution:**

From the given frequency distribution,

We have,

The upper boundary of the highest class, 𝑋 𝑈 = 𝑇𝐾. 2200

And the lowest boundary of the highest class, 𝑋 𝐿 = 𝑇𝐾. 1700

Then Range = 𝑋 𝑈 − 𝑋 𝐿 = 𝑇𝐾. 2200 − 𝑇𝐾. 1700 = 𝑇𝐾. 500

**Quartile Range:**

The quartile range is a measure of where the “middle fifty” is in a data set. Where a range is a measure of where the beginning and end are in a set, an interquartile range is a measure of where the bulk of the values lie. That’s why it’s preferred over many other measures of spread when reporting things like school performance or SAT scores.

The interquartile range formula is the first quartile subtracted from the third quartile:

IQR = Q3 – Q1.

**Example:**

Find the IQR for the following data set: 3, 5, 7, 8, 9, 11, 15, 16, 20, 21.

* Step 1: **Put the numbers in order**.  
  3, 5, 7, 8, 9, 11, 15, 16, 20, 21.
* Step 2: **Make a mark in the center of the data**:  
  3, 5, 7, 8, 9, **|** 11, 15, 16, 20, 21.
* Step 3: **Place parentheses around the numbers above and below the mark you made in Step 2–it makes Q1 and Q3 easier to spot.**  
  (3, 5, 7, 8, 9), **|** (11, 15, 16, 20, 21).
* Step 4: **Find Q1 and Q3**  
  Q1 is the median (the middle) of the lower half of the data, and Q3 is the median (the middle) of the upper half of the data.  
  (3, 5, **7**, 8, 9), **|** (11, 15, **16,** 20, 21). Q1 = 7 and Q3 = 16.
* Step 5: **Subtract Q1 from Q3**.  
  16 – 7 = 9.

9 IS IQR.

**Semi Inter Quartile Range:**

The semi-interquartile range is defined as the measures of dispersion. Semi interquartile range also is defined as half of the interquartile range. It is computed as one half of the difference between the 75th percentile (Q3) and the 25th percentile (Q1). The semi-interquartile range is one-half of the difference between the first and third quartiles.

The Formula for Semi Interquartile Range is

Semi Interquartile Range = (Q3– Q1) / 2

**Example:**

Find the Quartile Deviation for the following set of data:

{490, 540, 590, 600, 620, 650, 680, 770, 830, 840, 890, 900}

**Step 1:** Find the first quartile, Q1.

This is the median of the lower half of the set {490, 540, 590, 600, 620, 650}.

Q1 = (590 + 600) / 2 = 595.

**Step 2:** Find the third quartile, Q3.

This is the median of the upper half of the set {680, 770, 830, 840, 890, 900}.

Q3 = (830 + 840) / 2 = 835.

**Step 3:** Subtract Step 1 from Step 2.

835 – 595 = 240.

Step 4: Divide by 2.

240 / 2 = 120

The quartile deviation/ Semi Interquartile Range for this set of data is 12.

**Variance:**

Variance (σ2) in statistics is a measurement of the spread between numbers in a data set. That is, it measures how far each number in the set is from the mean and therefore from every other number in the set.

**Example:**

Find the variance for the following set of data representing trees in California (heights in feet): 3, 21, 98, 203, 17, 9

Step 1: Add up the numbers in your given data set.

3 + 21 + 98 + 203 + 17 + 9 = 351

Step 2: Square your answer:

351 × 351 = 123,201

…and divide by the number of items. We have 6 items in our example so:

123,201 / 6 = 20,533.5

Set this number aside for a moment.

Step 3: Take your set of original numbers from Step 1, and square them individually this time:

3 × 3 + 21 × 21 + 98 × 98 + 203 × 203 + 17 × 17 + 9 × 9

Add those numbers (the squares) together:

9 + 441 + 9604 + 41209 + 289 + 81 = 51,633

Step 4: Subtract the amount in Step 2 from the amount in Step 3.

51,633 – 20,533.5 = 31,099.5

Set this number aside for a moment.

Step 5: Subtract 1 from the number of items in your data set\*. For our example:

6 – 1 = 5

Step 6: Divide the number in Step 4 by the number in Step 5. This gives you the variance:

31,099.5 / 5 = 6,219.9

**Standard Deviation:**

Standard deviation is the positive square root of the mean-square deviations of the observations from their arithmetic mean.

It measures the dispersion (or spread) of figures around the mean. A large number for the standard deviation means there is a wide spread of values around the mean, whereas a small number for the standard deviation implies that the values are grouped close together around the mean.

The formula:

σ = √ {∑ (x -) ẍ 2 / n}

**Example:**

Find the standard deviation for the following results:  
{12, 15, 17, 20, 30, 31, 43, 44, 54}

Step 1: Add the numbers up:  
12 + 15 + 17 + 20 + 30 + 31 + 43 + 44 + 54 = 266.

Step 2: Square the answer from Step 1:  
266 x 266 = 70756

Step 3: Divide your answer from Step 2 by the number of items (n) in your set. In this example, we have 9 items, so:  
70756 / 9 = 7861.777777777777 (dividing by n)

Set this number aside for a moment. You’ll need it in a later step.

Step 4: Square the original numbers {12, 15, 17, 20, 30, 31, 43, 44, 54} one at a time, then add them up:

(12 x 12) + (15 x 15) + (17 x 17) + (20 x 20) + (30 x 30) + (31 x 31) + (43 x 43) + (44 x 44) + (54 x 54) = 9620

Step 5: Subtract Step 4 from Step 3.

9620 – 7861.777777777777 = 1758.2222222222226

Step 6: Subtract 1 from n. We have 9 items, so n = 9:

9 – 1 = 8

Step 7: Divide Step 5 by Step 6 to get the [**variance**](https://calculushowto.com/standard-deviation-definition/#definitionvar):  
1758.2222222222226 / 8 = 219.77777777777783

Step 8: Take the square root of Step 7:  
√(219.77777777777783) = 14.824903971958058  
The standard deviation is 14.825.

**Coefficient of Variation:**

Coefficient of Variation Formula Coefficient of Variation is expressed as the ratio of standard deviation and mean. It is often abbreviated as CV. Coefficient of variation is the measure of variability of the data. When the value of coefficient of variation is higher, it means that the data has high variability and less stability. When the value of coefficient of variation is lower, it means the data has less variability and high stability. The formula for coefficient of variation is given below: Coefficient of Variation = Standard Deviation / Mean.

**Example:**

Find the coefficient of variation of the following sample set of numbers.

{1, 5, 6, 8, 10, 40, 65, 88}.

Given sample set: {1, 5, 6, 8, 10, 40, 65, 88}.

Sample mean = (1 + 5 + 6 + 8 + 10 + 40 + 65 + 88)/8 = 223/8 = 27.875

∑ni=1(xi−x¯)2=(1−27.875)2+(5−27.875)2+(6−27.875)2+(8−27.875)2+(10−27.875)2+(40−27.875)2+(65−27.875)2+(88−27.875)2=7578.875

Variance:

∑ni=1(xi−x¯)2n−1=7578.8757=1082.696

Standard deviation:

σ=∑ni=1(xi−x¯)2n−1−−−−−−−−−−√=1082.696−−−−−−−√=32.904

Coefficient of variation = 32.901/27.875=1.180