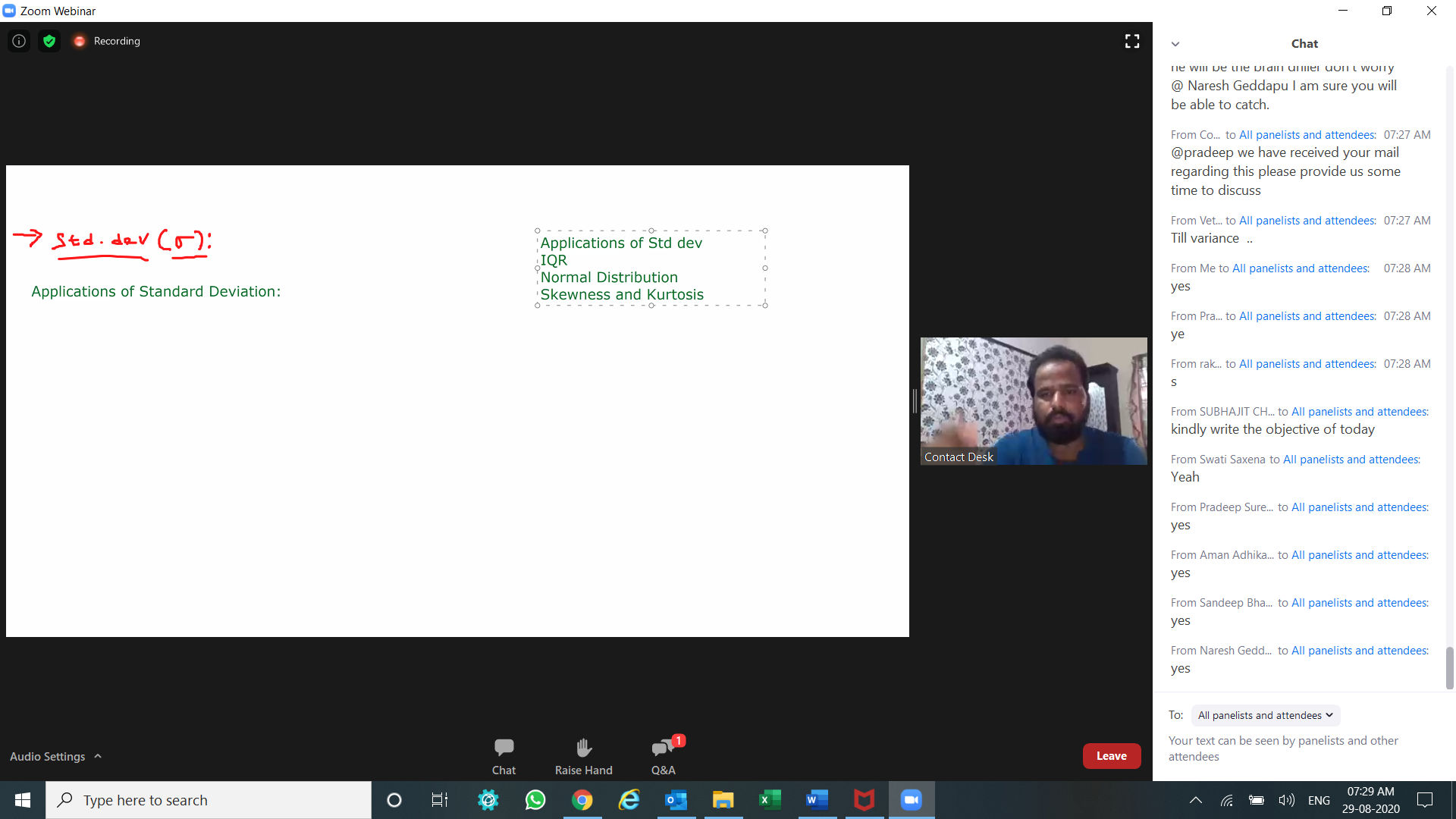
guys pls change the chat to "All panelist and attendees" from "All panelist"

Because even we might have same doubt or ques.

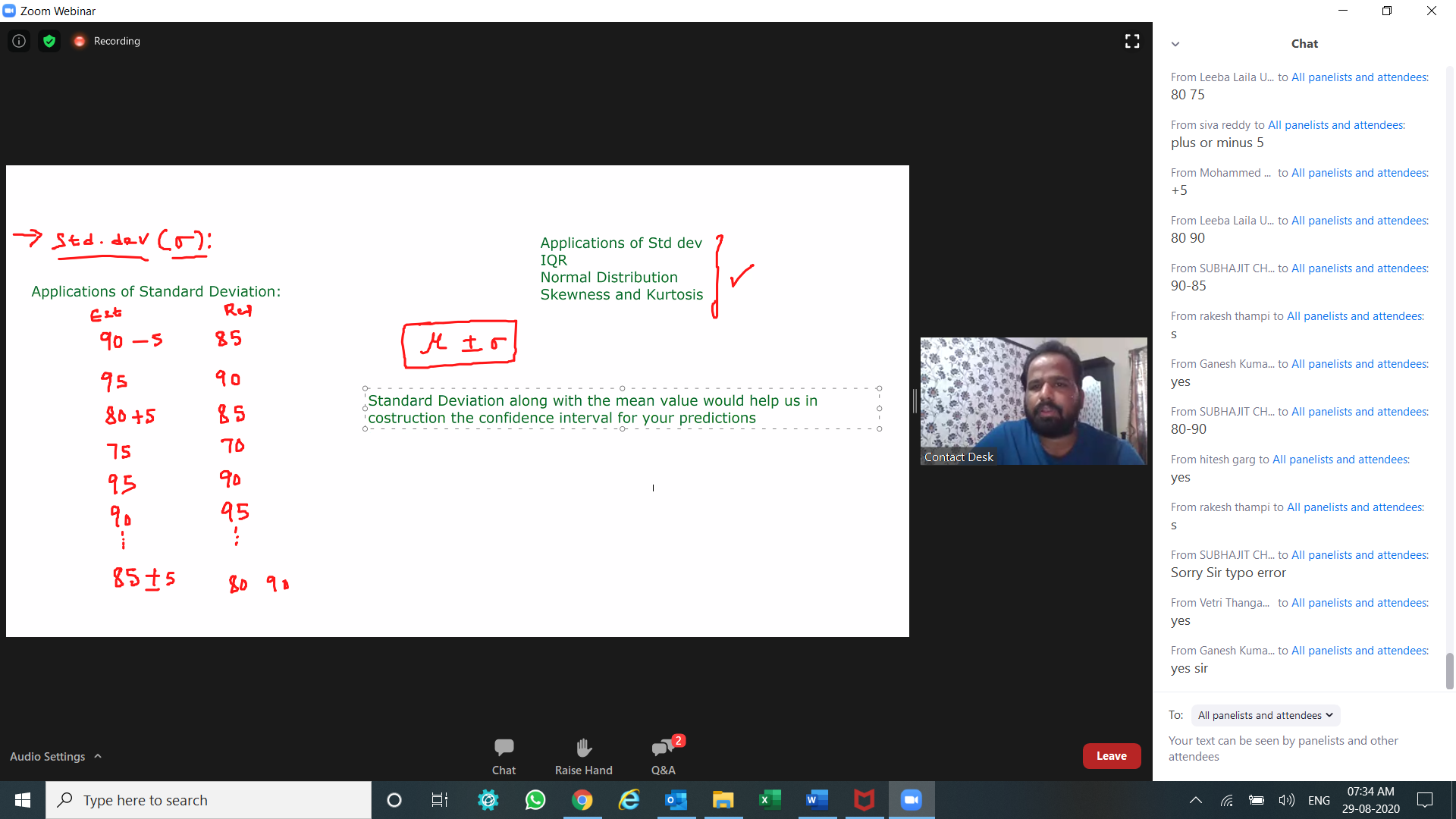
@Jasmeet Singh

<https://medium.com/@manavpal8/central-tendency-9806a1fc3e6>

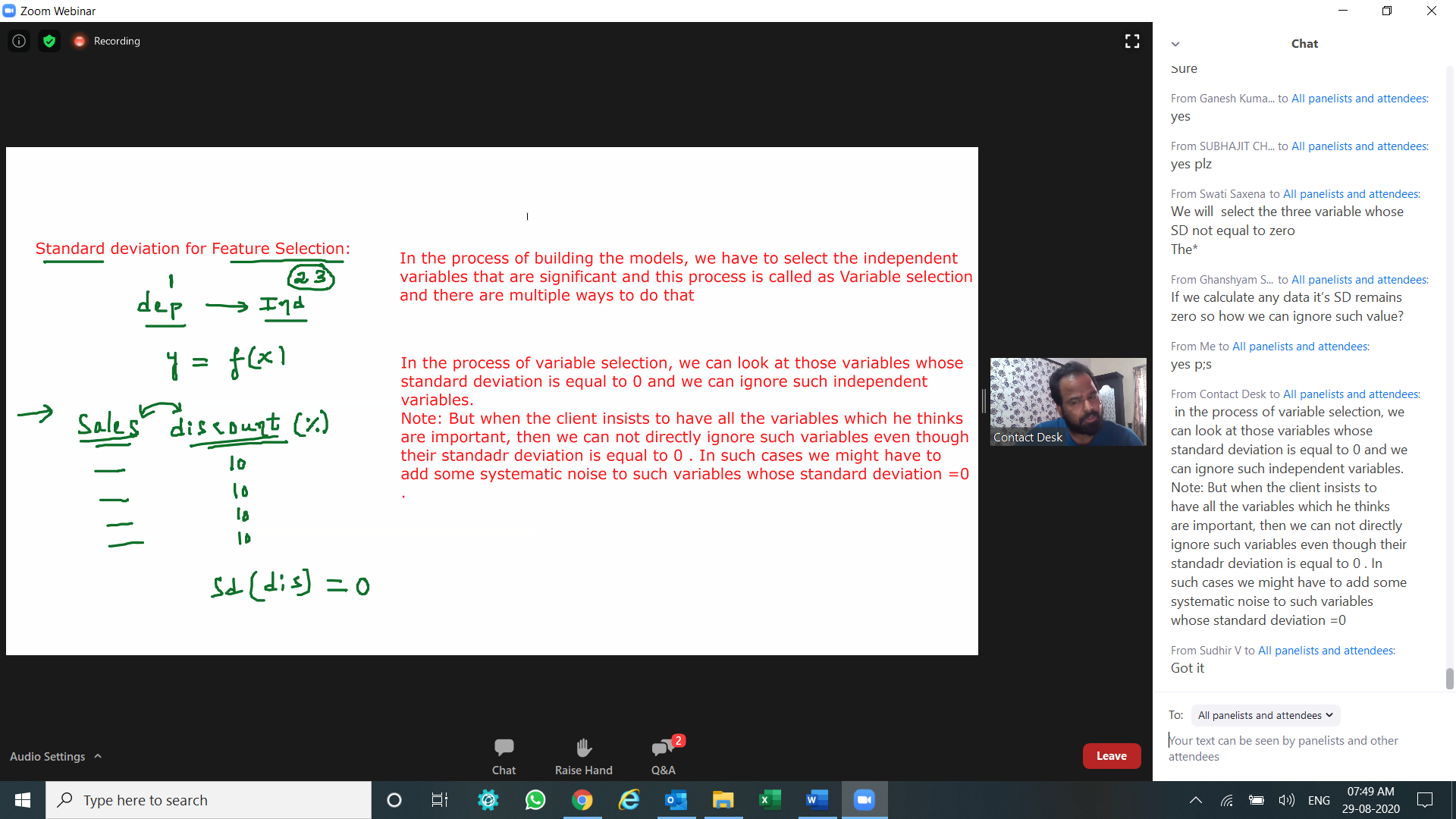
<https://medium.com/@chaitali.mehta310/explanation-on-iqr-inter-quater-range-2aa0a4c86e92>

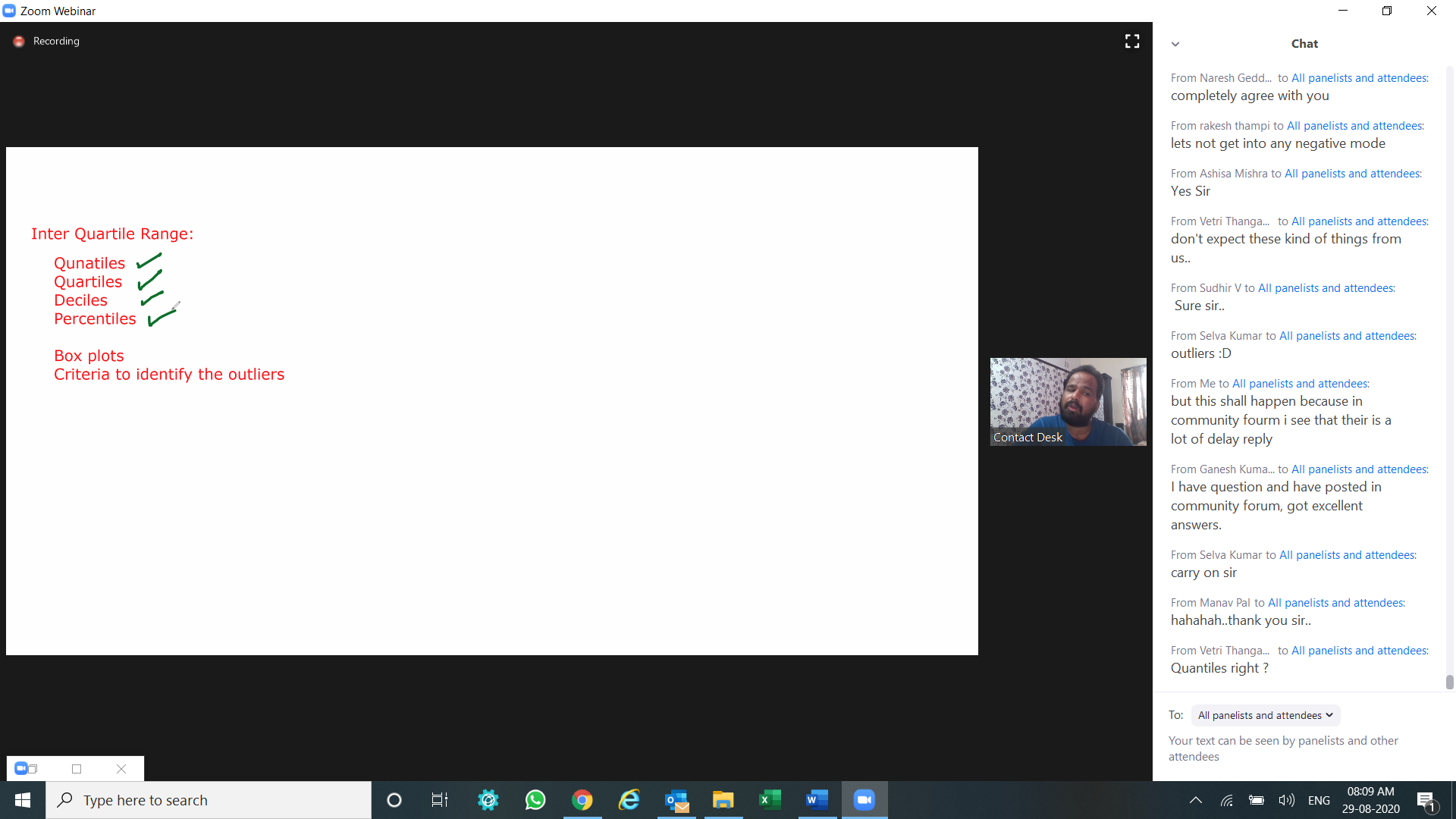


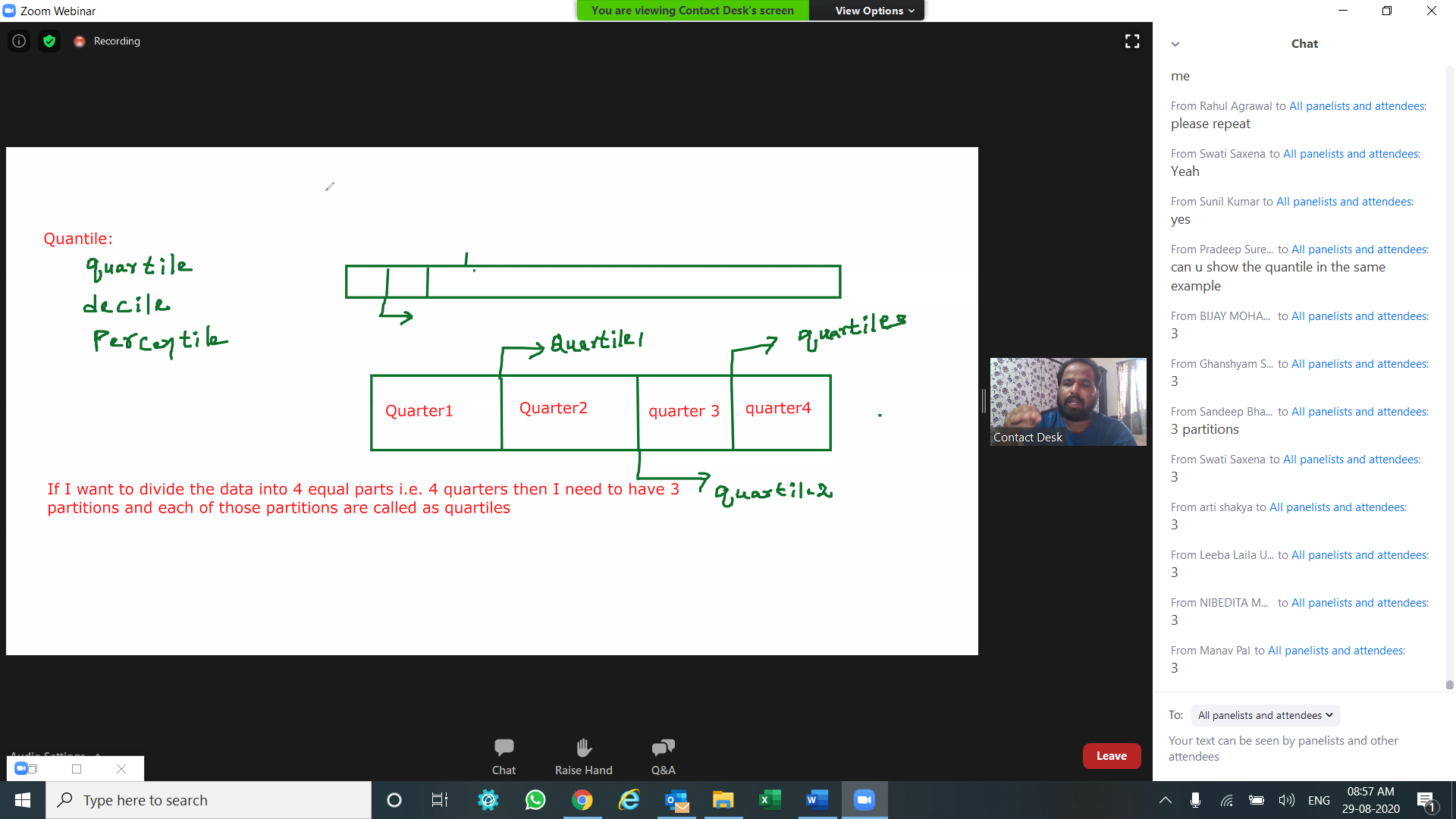
just to confirm, confidence interval here is the range of values we expect to achieve at certain level of confidence?



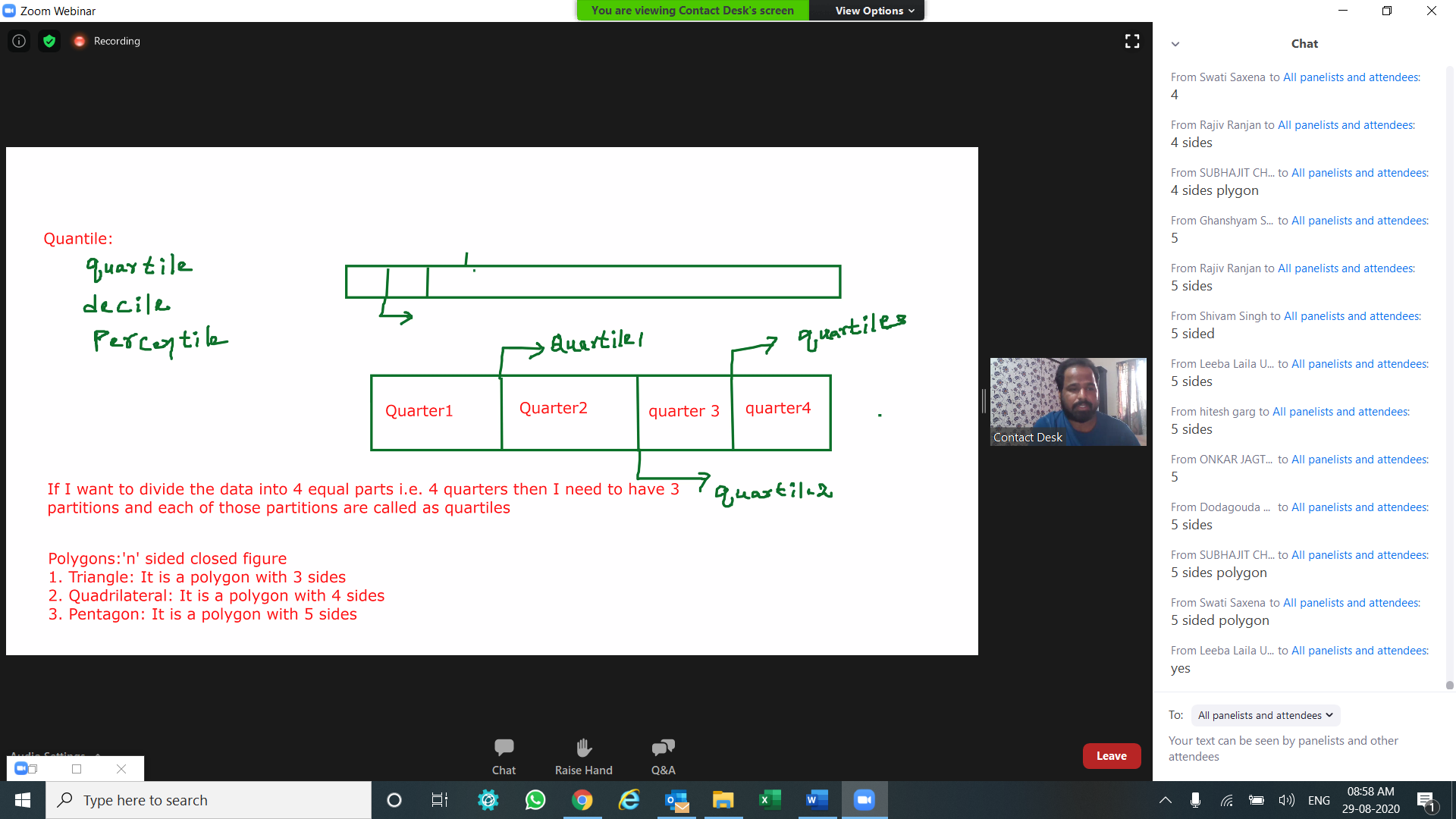
in the process of variable selection, we can look at those variables whose standard deviation is equal to 0 and we can ignore such independent variables.Note: But when the client insists to have all the variables which he thinks are important, then we can not directly ignore such variables even though their standard deviation is equal to 0 . In such cases we might have to add some systematic noise to such variables whose standard deviation =0



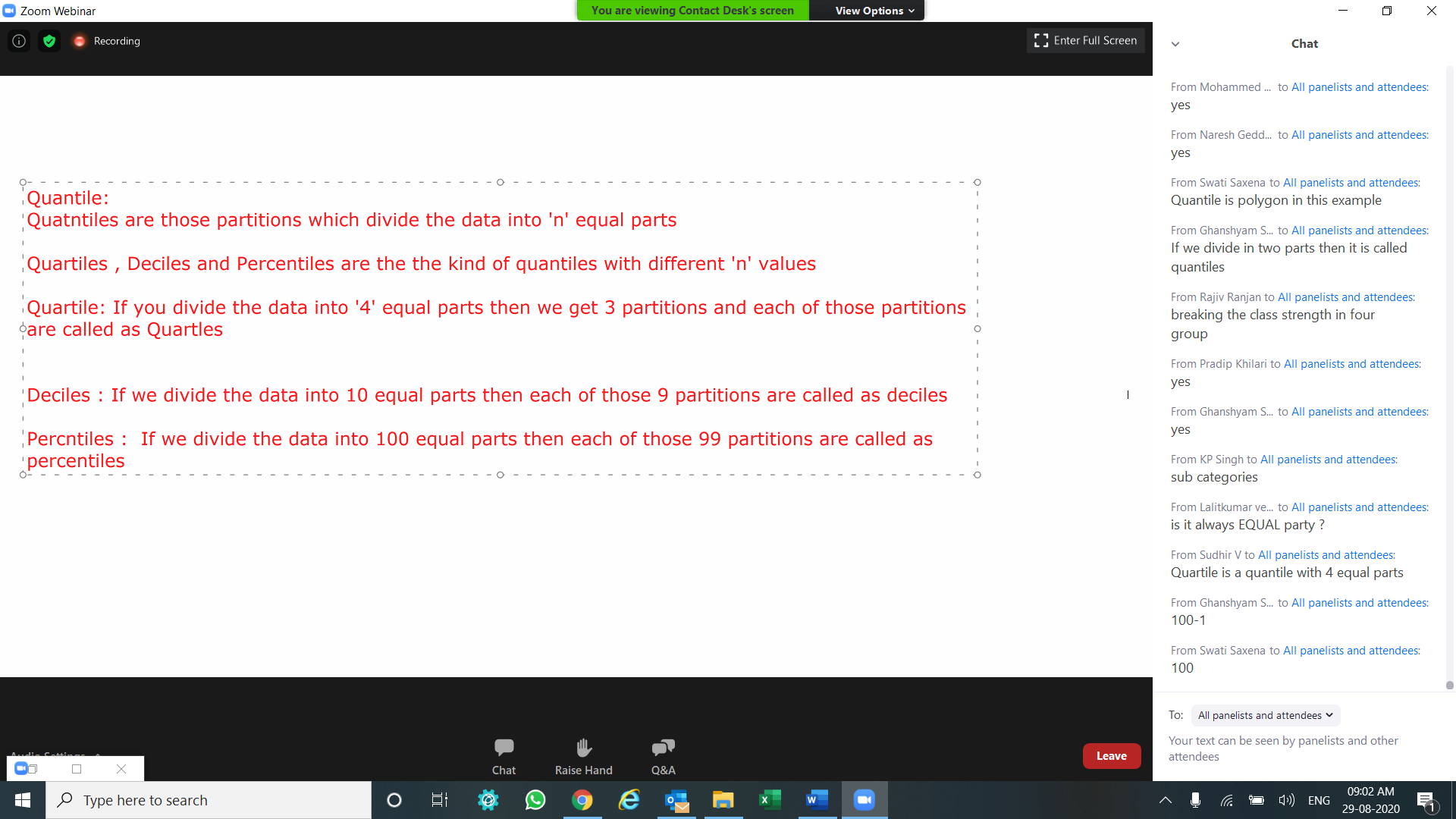


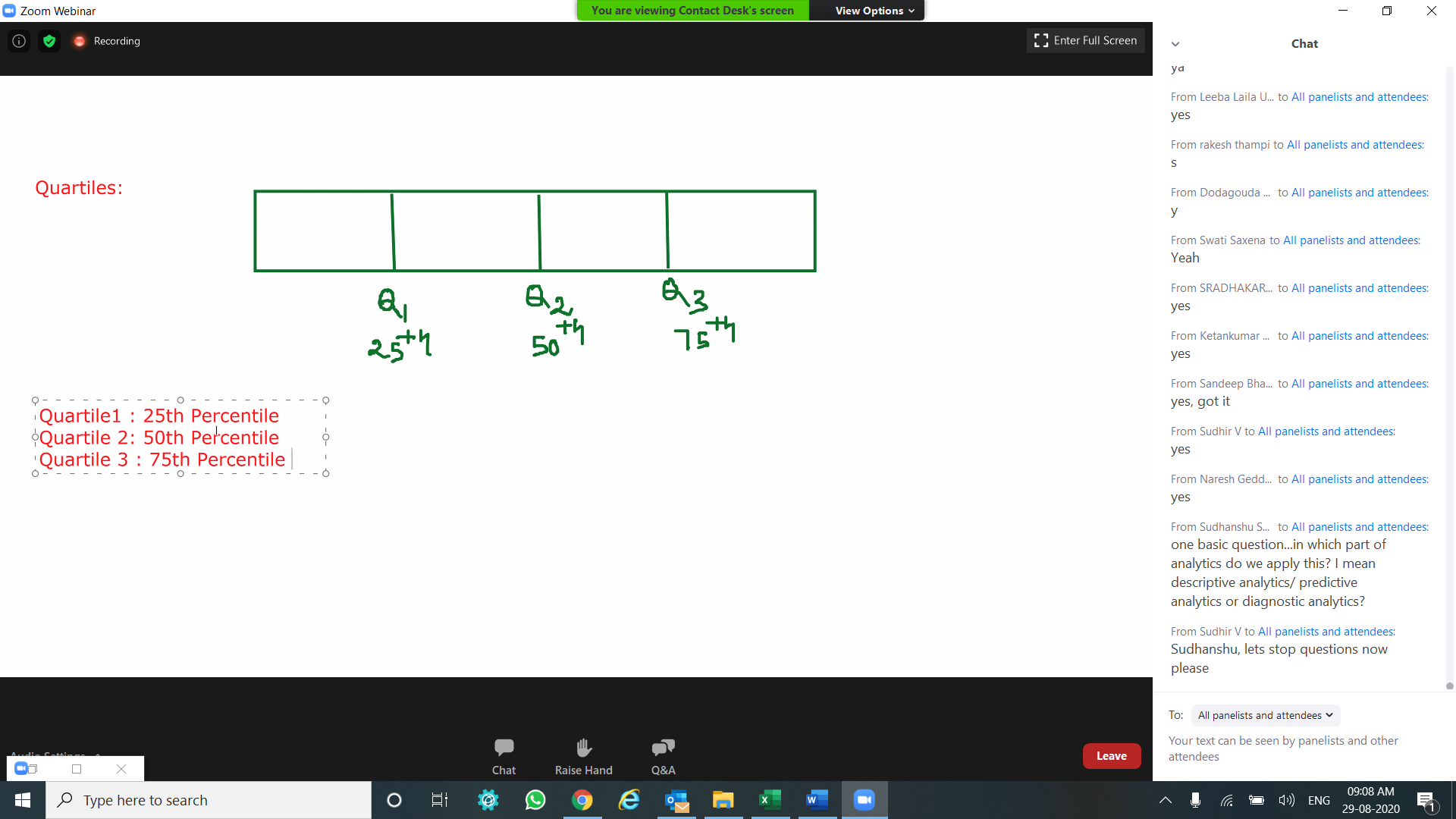


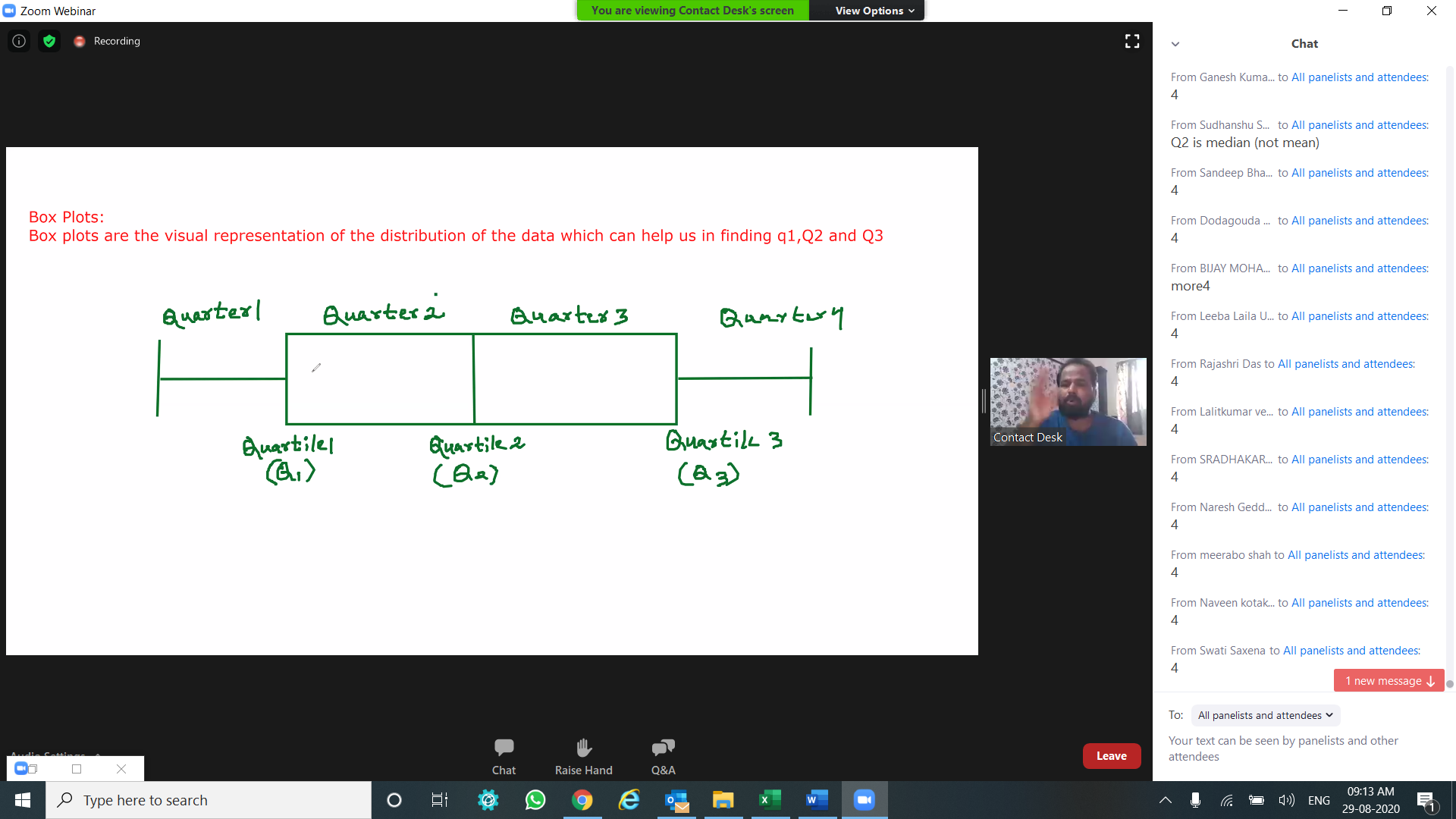
A polygon is any 2-dimensional shape formed with straight lines. Triangles, quadrilaterals, pentagons, and hexagons are all examples of polygons



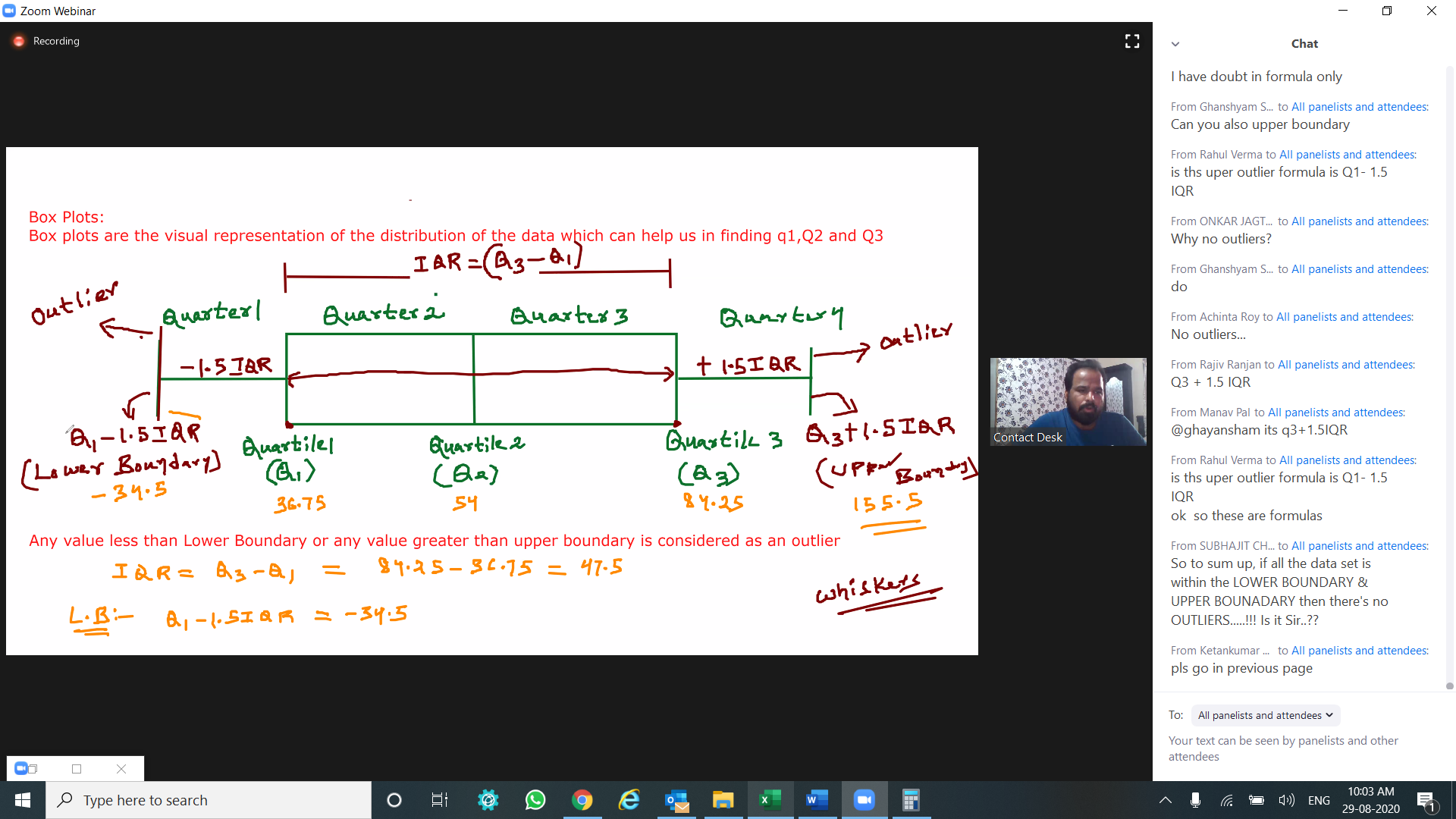
in this example, Quartile are a type of quantile which divide data into 4 parts.





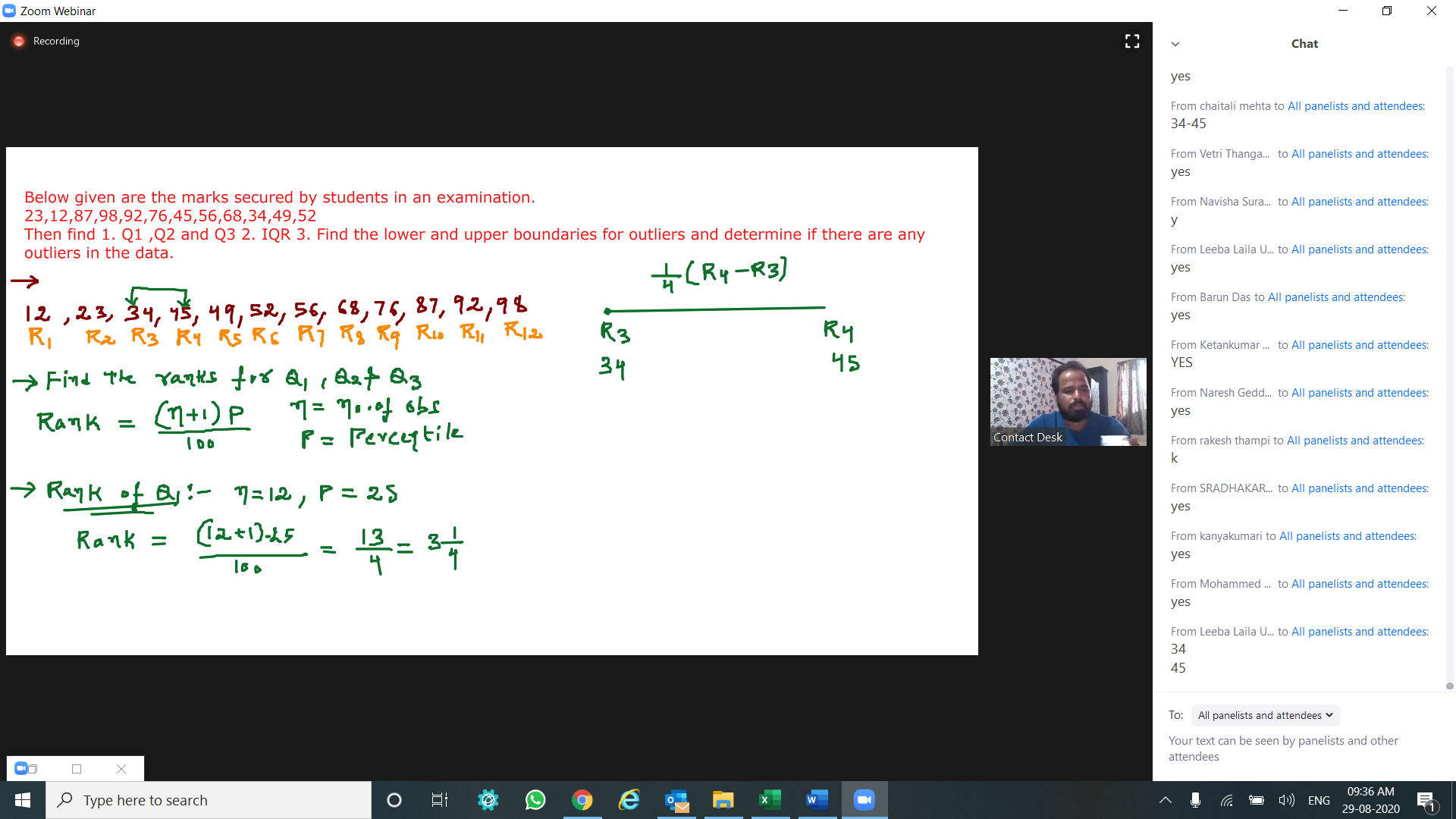


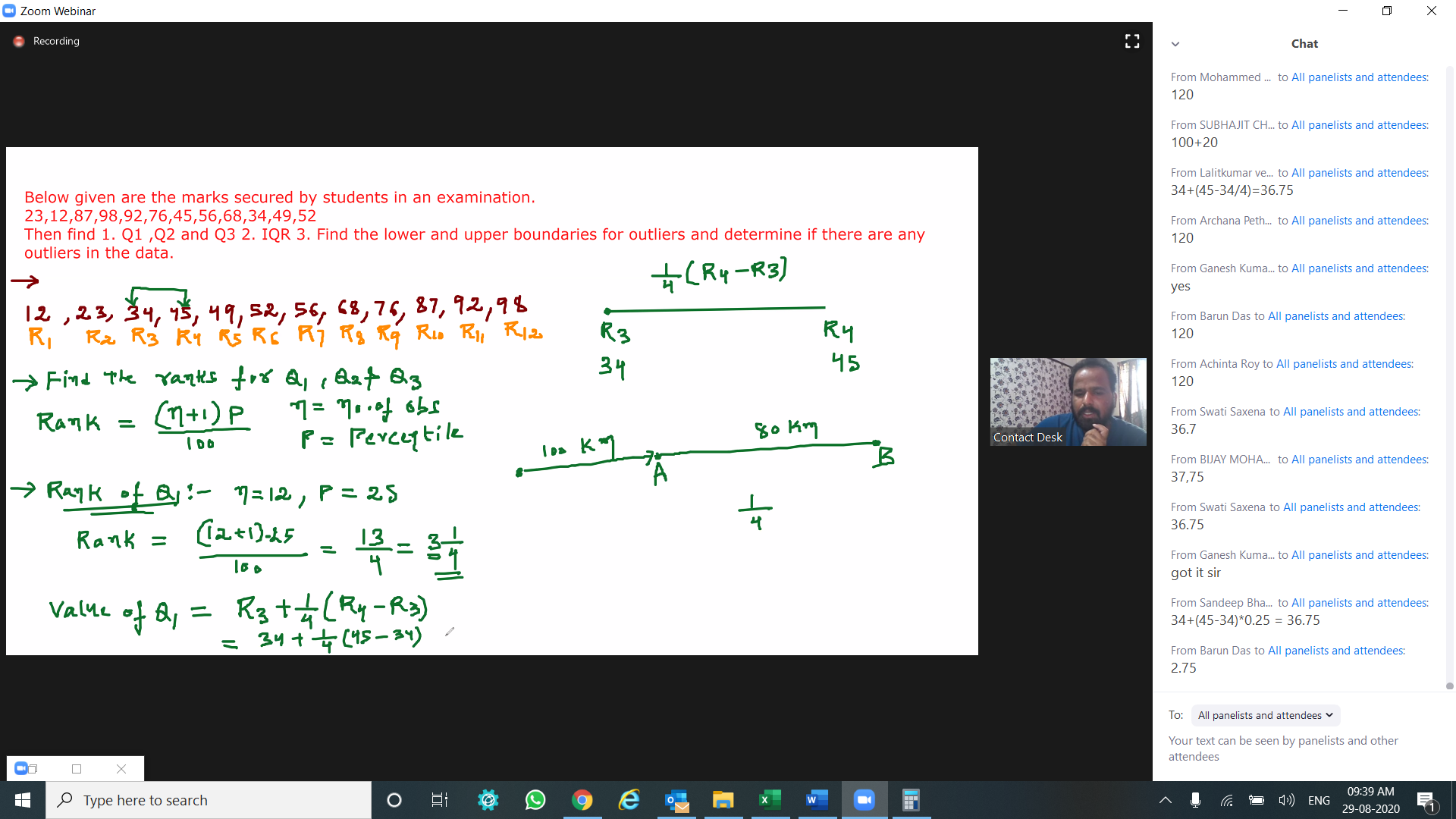
interquartile range (IQR)

whiskers - whiskers are the boundaries | whiskers does not show 0 percentile or 100 percentile since we have outliers (if present in the sample). 

So to sum up, if all the data set is within the LOWER BOUNDARY & UPPER BOUNADARY then there's no OUTLIERS.

e.g.,





IQR=Q3-Q1 = 84.25-36.75 = 47.5

Lower boundary = IQR-1.5 Q1 =47.5 - 1.5\*36.75 = -7.625

Upper boundary = IQR+1.5 Q3 =84.25 + 1.5\*84.75 = 174.625

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| R1 | R2 | R3 | R4 | R5 | R6 | R7 | R8 | R9 | R10 |
| 32 | 545 | 980 | 1256 | 9898 | 12567 | 12568 | 32984 | 33000 | 99000 |

| n Value = 10 |

Percentile 1 (P1) = 10+1\*0.25 = 2.75 (this lies in b/w R2 & R3) |  
Percentile 2 (P2) = 10+1\*0.50 = 5.5 (this lies in b/w R5 & R6) |

Percentile 3 (P3) = 10+1\*0.75 = 8.25 (this lies in b/w R8 & R9) |

Quartile 1 (Q1) = R2+(R3-R2)\*0.25 = 653.75 |

Quartile 2 (Q2) = R5+(R6-R5)\*0.50 = 11232.5 |

Quartile 3 (Q3) = R8+(R9-R8)\*0.75 = 32996 |

IQR = Q3 - Q1 = 32342.25 |

Lower boundary = Q1-1.5\*IQR = -47859.625 |

Upper boundary = Q3+1.5\*IQR = 81509.375 |

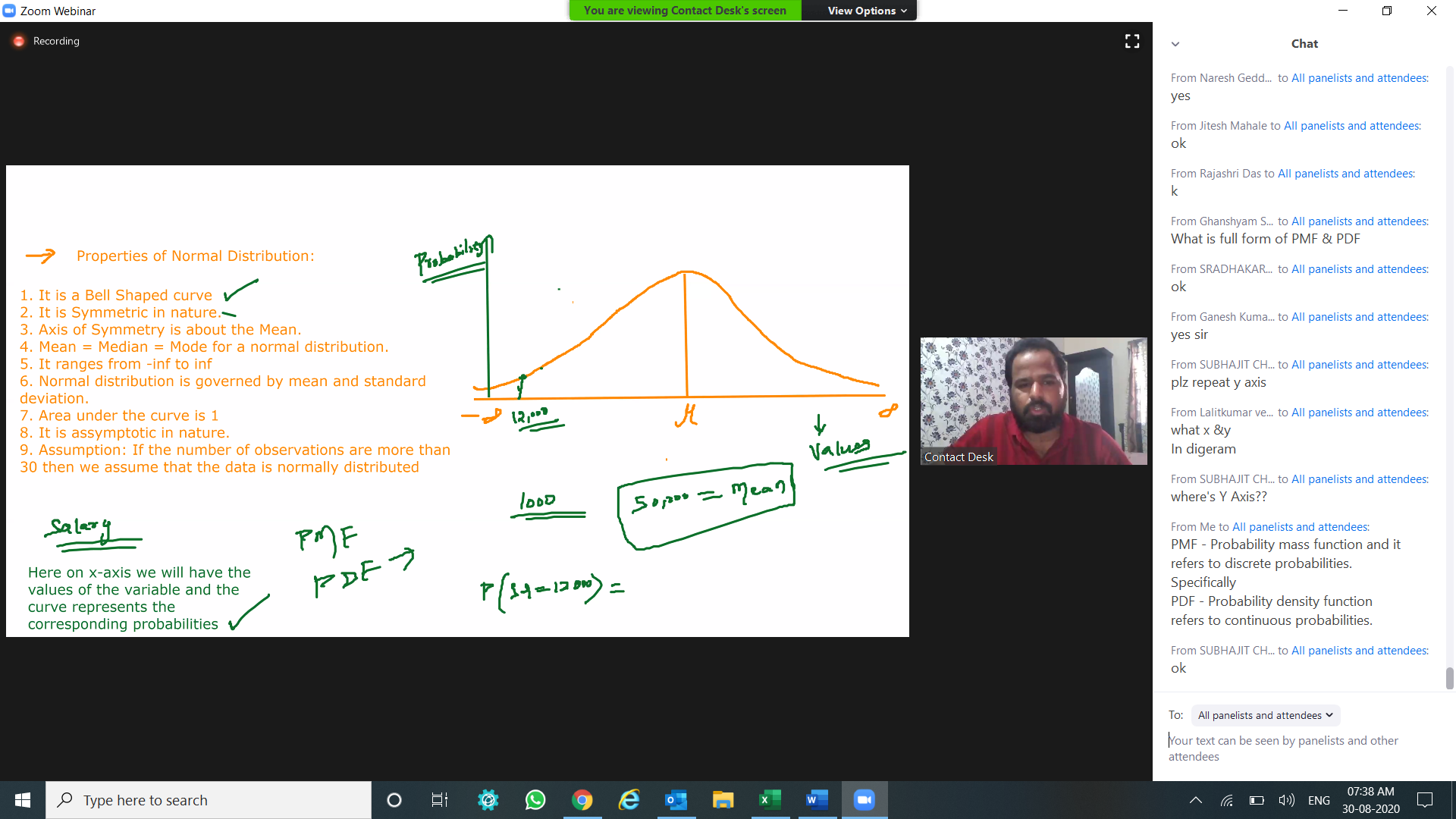
And the

Outliers = 99000

1. It is a Bell-Shaped curve 2. It is Symmetric in nature.3. Axis of Symmetry is about the Mean.4. Mean = Median = Mode for a normal distribution.5. It ranges from -inf to inf 6. Normal distribution is governed by mean and standard deviation.7. Area under the curve is 1

8. It is asymptotic in nature.

9. Assumption: If the number of observations/variable is more than 30 then we assume that the data is normally distributed.



PMF - Probability mass function and it refers to discrete probabilities. Specifically

PDF - Probability density function refers to continuous probabilities.

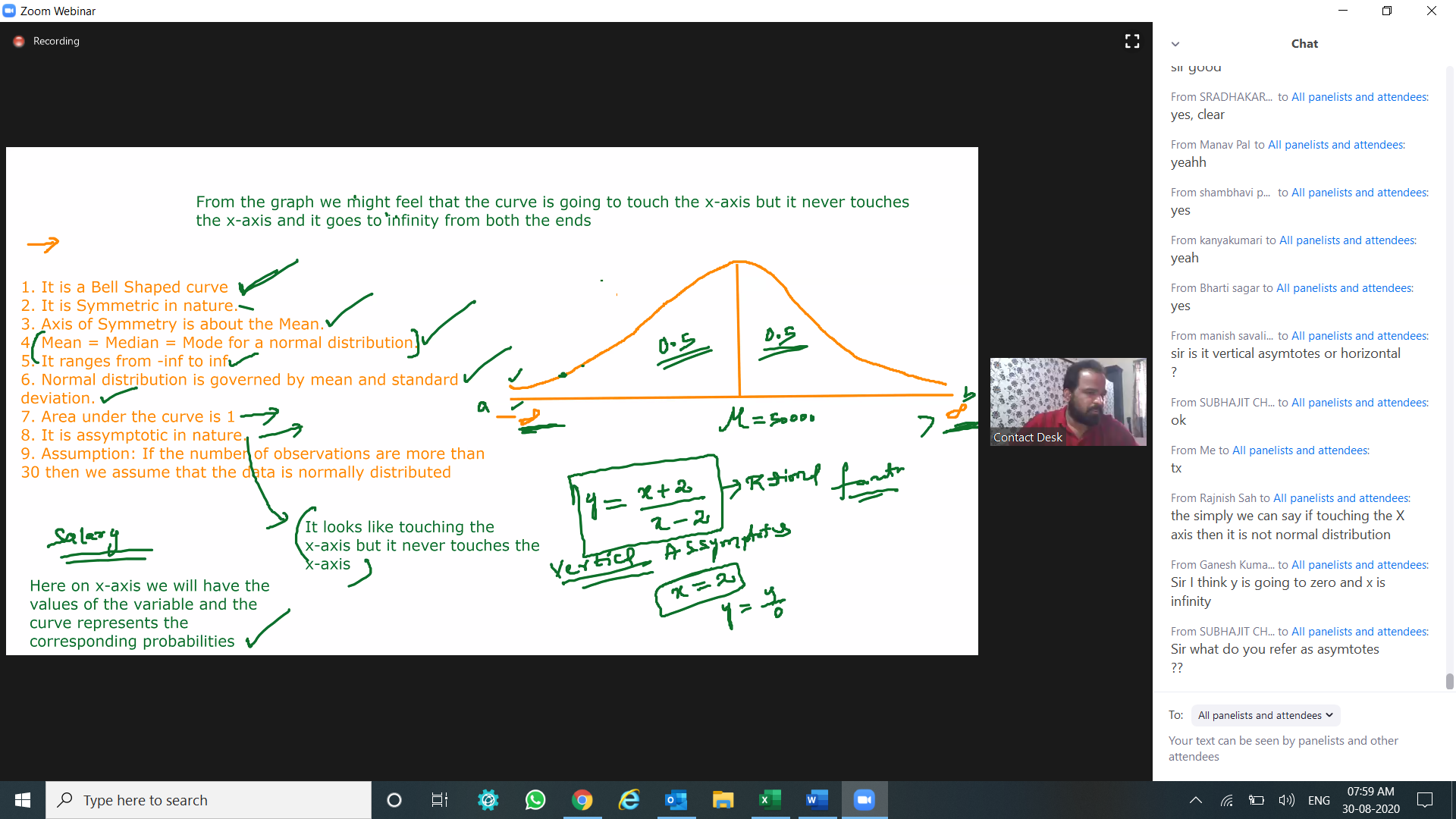
symmetrical about diameter and its image == mirror image in the above image bell curve

area under curve represents sum of all probabilities, hence area = 1

Asymptotic:

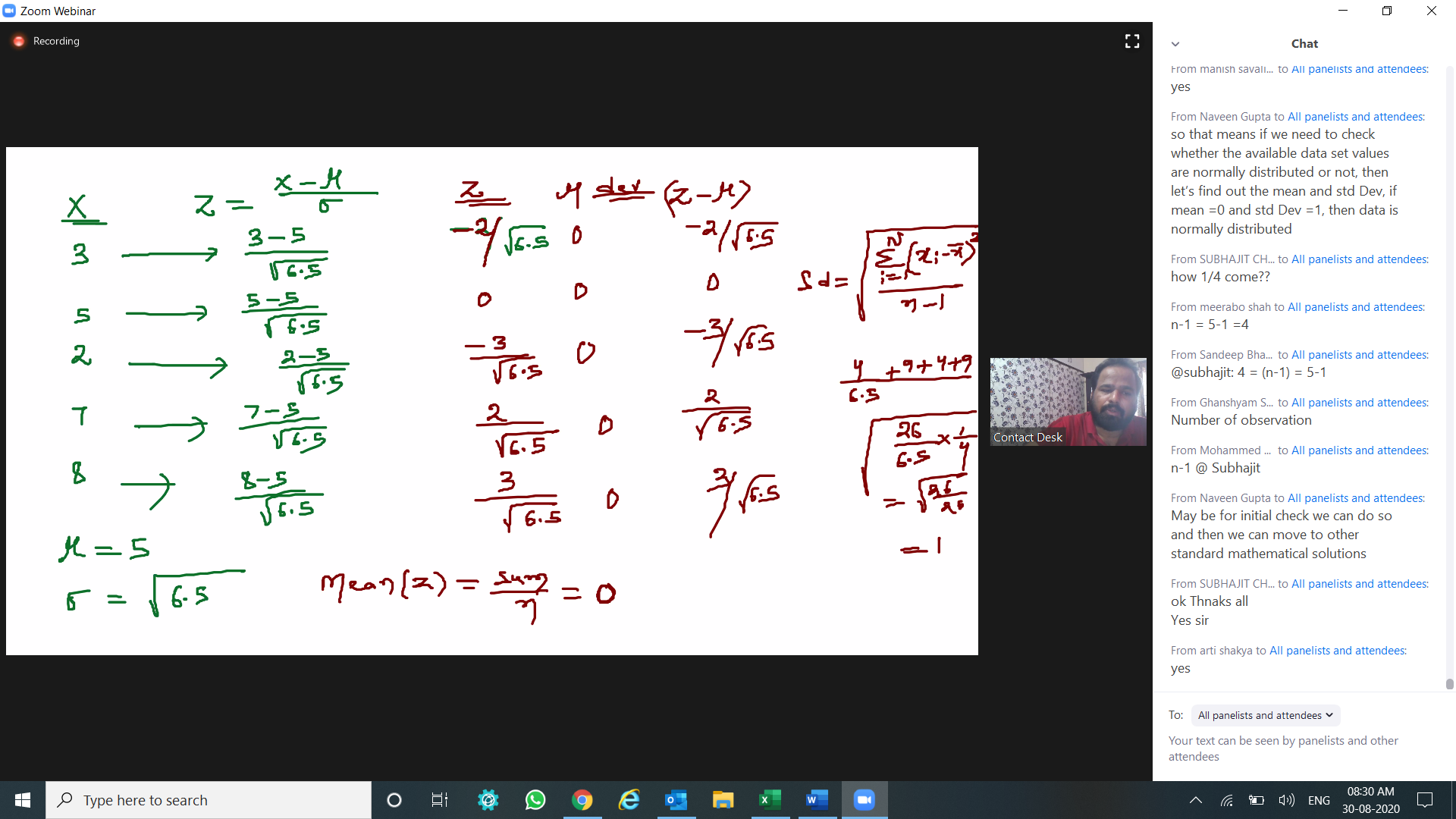
as per my understanding back to your salary eg., in the image - no empl will have "0" as his/her salary so the x-axis will not touch

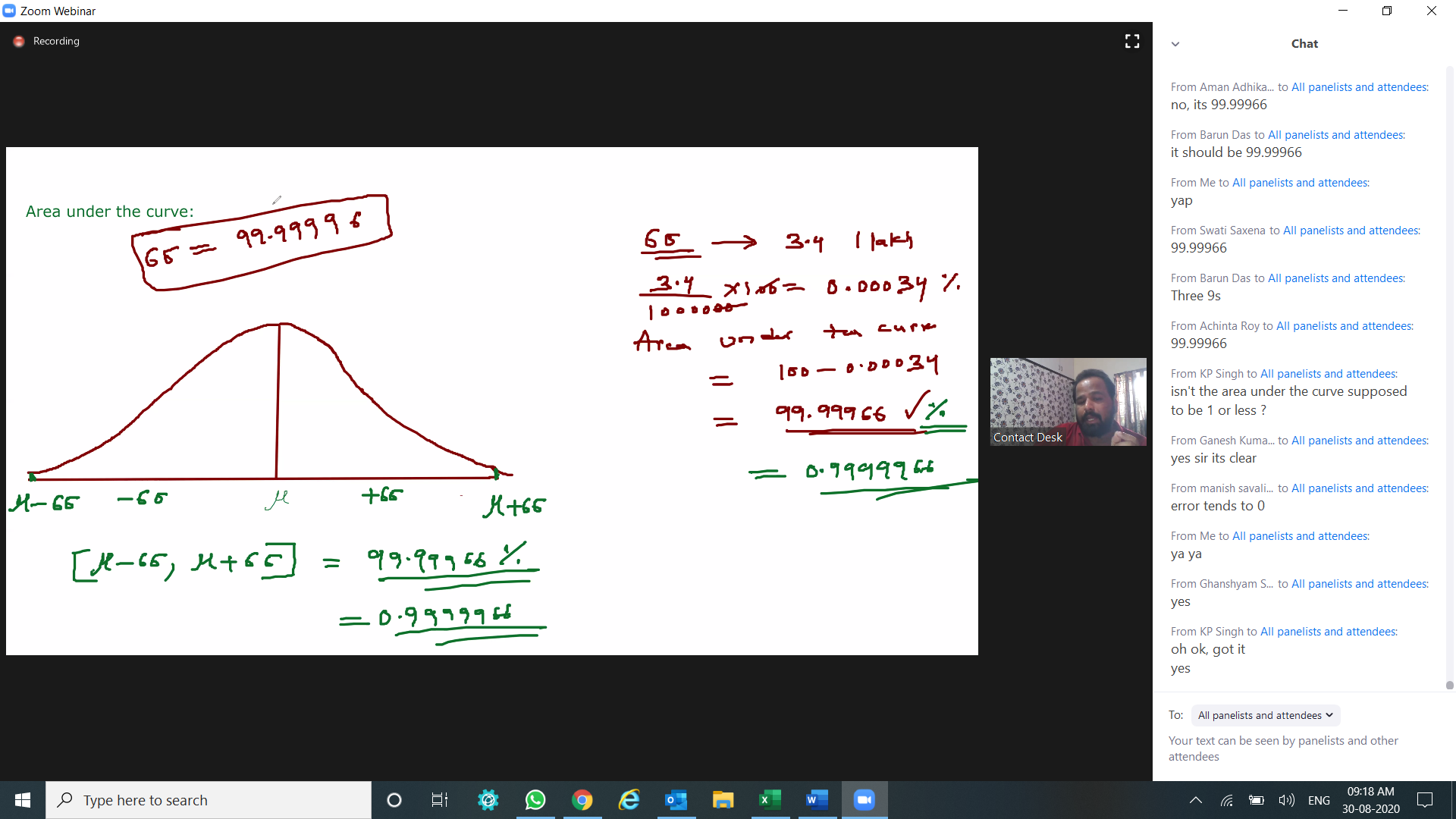
It is asymptotic is in bell in nature.

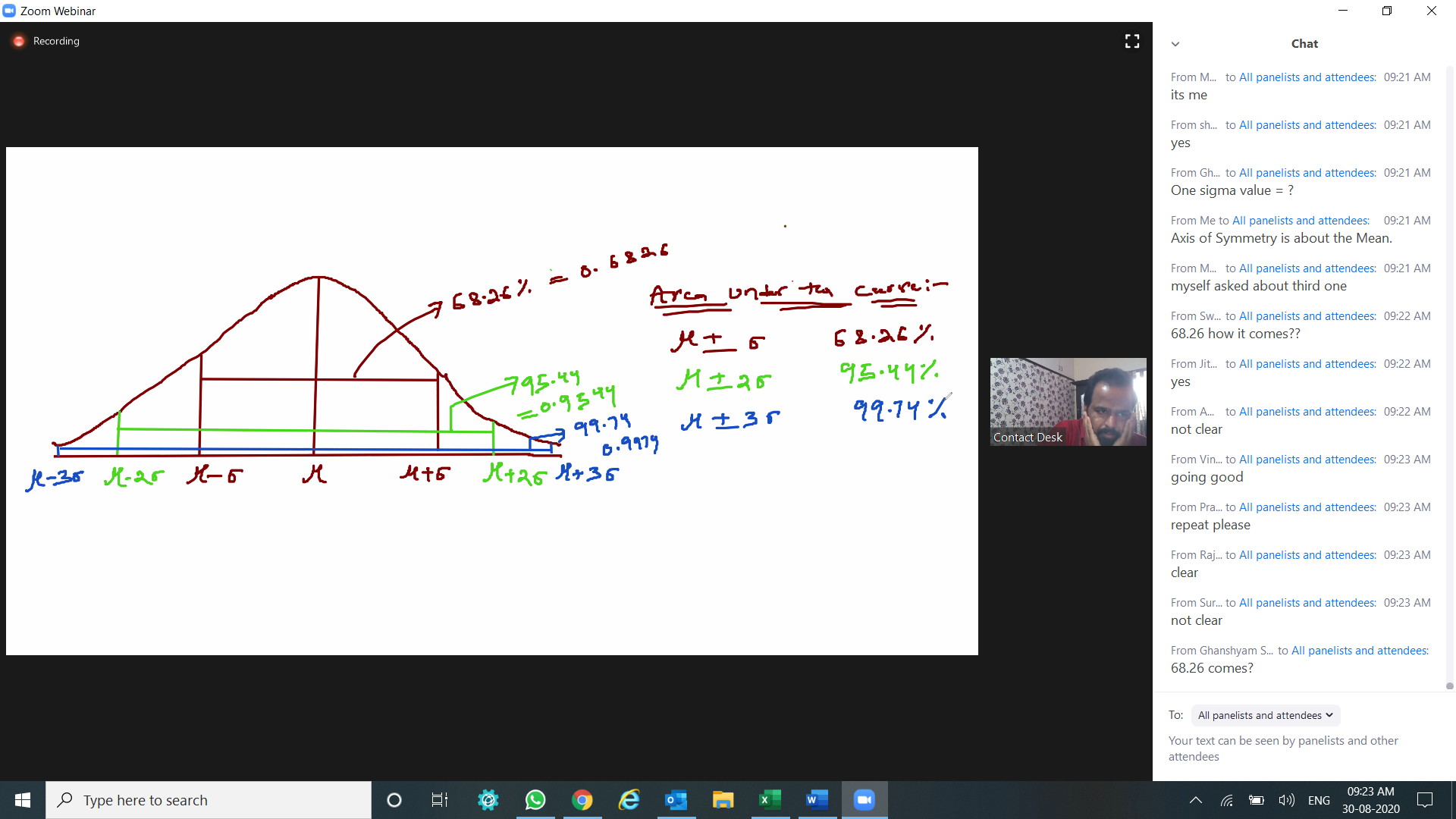


To know more about asymptote

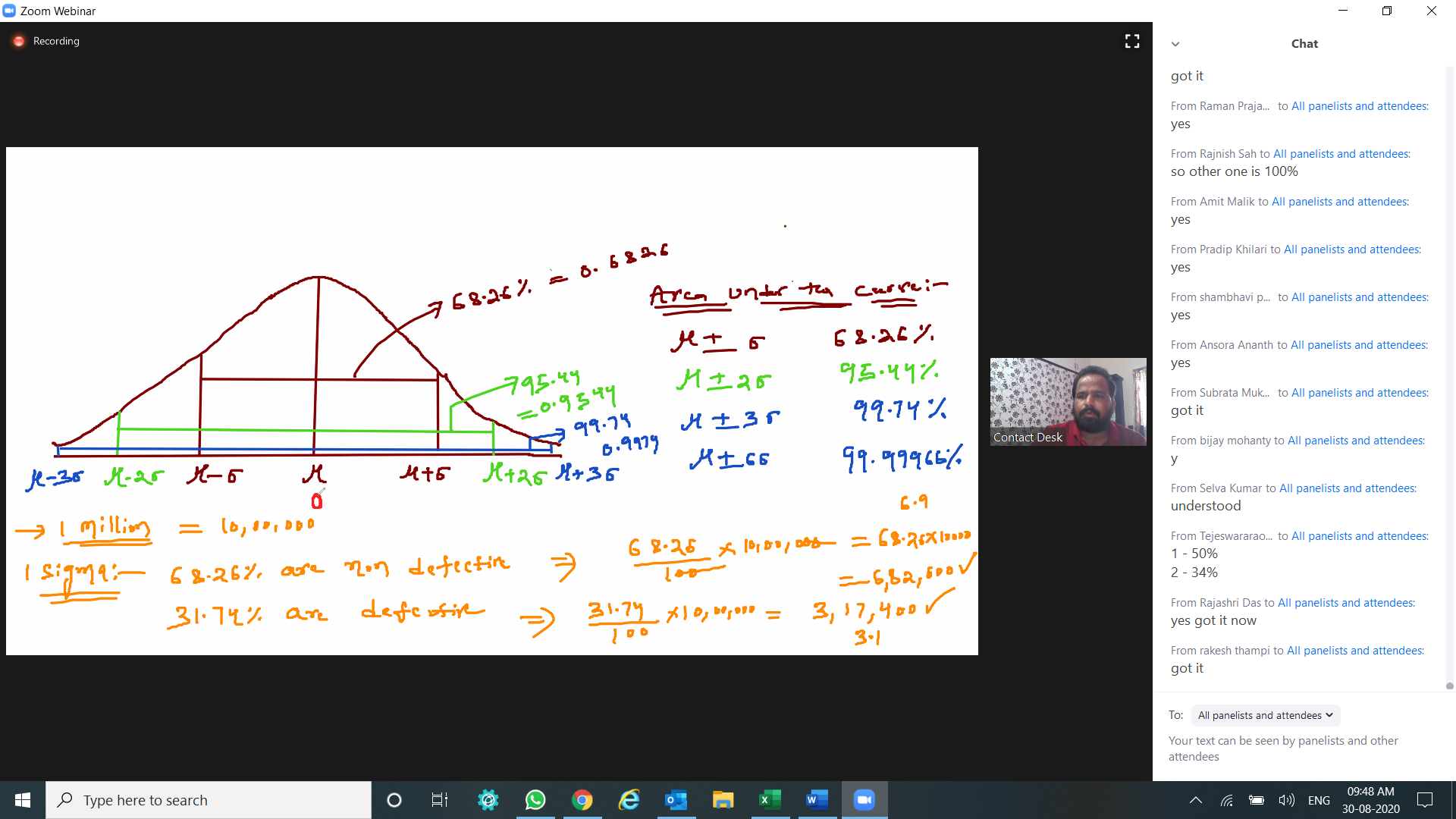
mathsisfun.com/algebra/asymptote.html





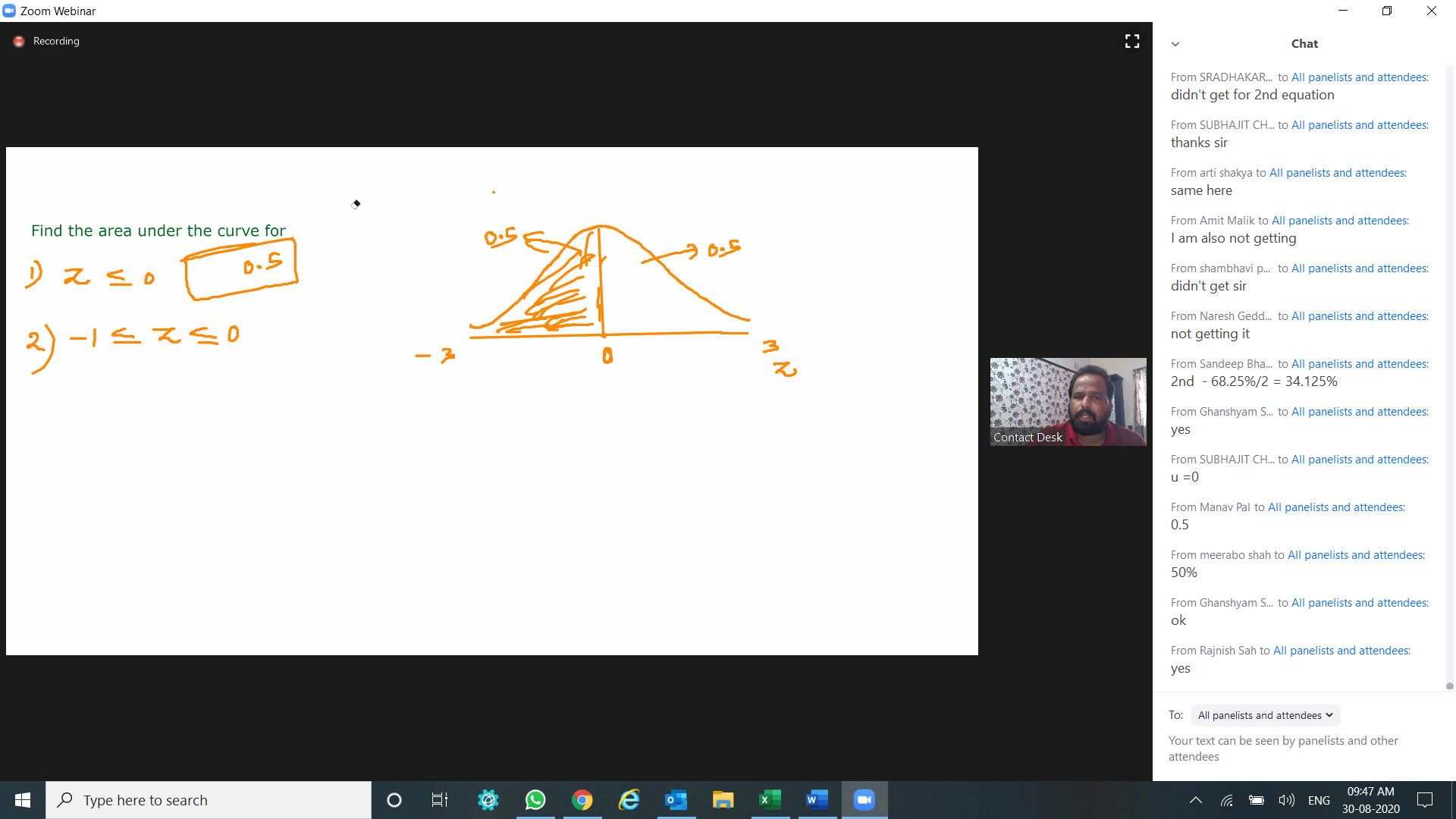


By the above examples, I could see that six sigma analysis would be bring few employee termination from the company due to this pandemic. Company has done this to avoid future loss via 6 sigma.



Find the area under the curve for

Z < = 0



-1 < = Z < = 0

