Data Science with R Programming

1.Stastics

2.Regression

3.Visulaization

4.Machine Learning

5.Deep learning

Getting user from user

Vector

Matrix transverse

Slicing

Dataframes

Filling Nan

Splicing

Factors

String as factors

Factors are the data objects which are used to categorize the data and store it as levels

Random function

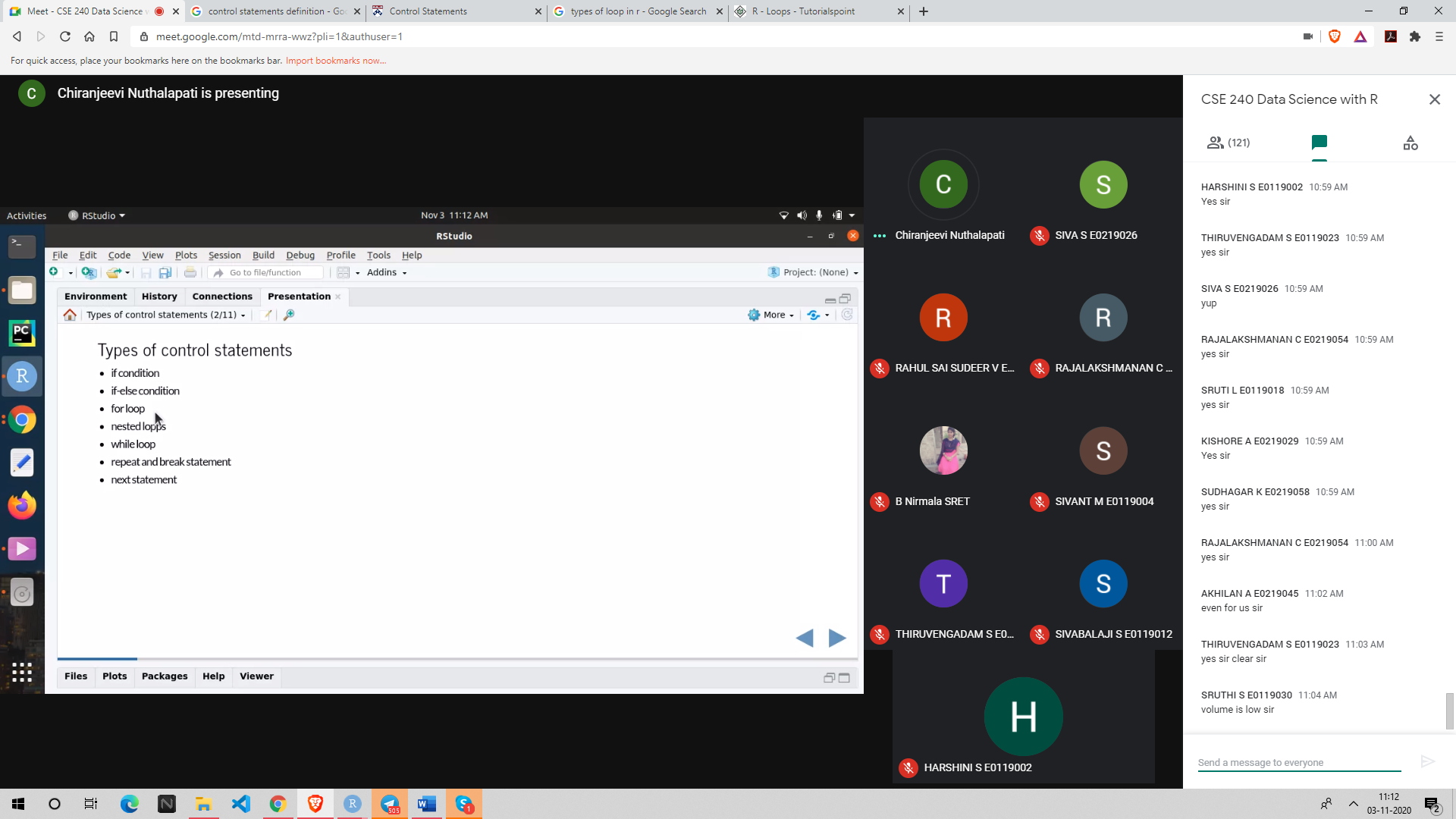
Sample

Runif

Decision statements

TRUE

FALSE



if

If else

For

Nested loops

Repeat loop

Break statement

Functions

Variables globally assigned

Data visualiztaion:

Charts,grap,map

Graphical represnetion of information and data

Understanding large amount of data

Ggplot has 6 layers ggplot – (grammar of graphics)

Data

Aestatics

Geometry

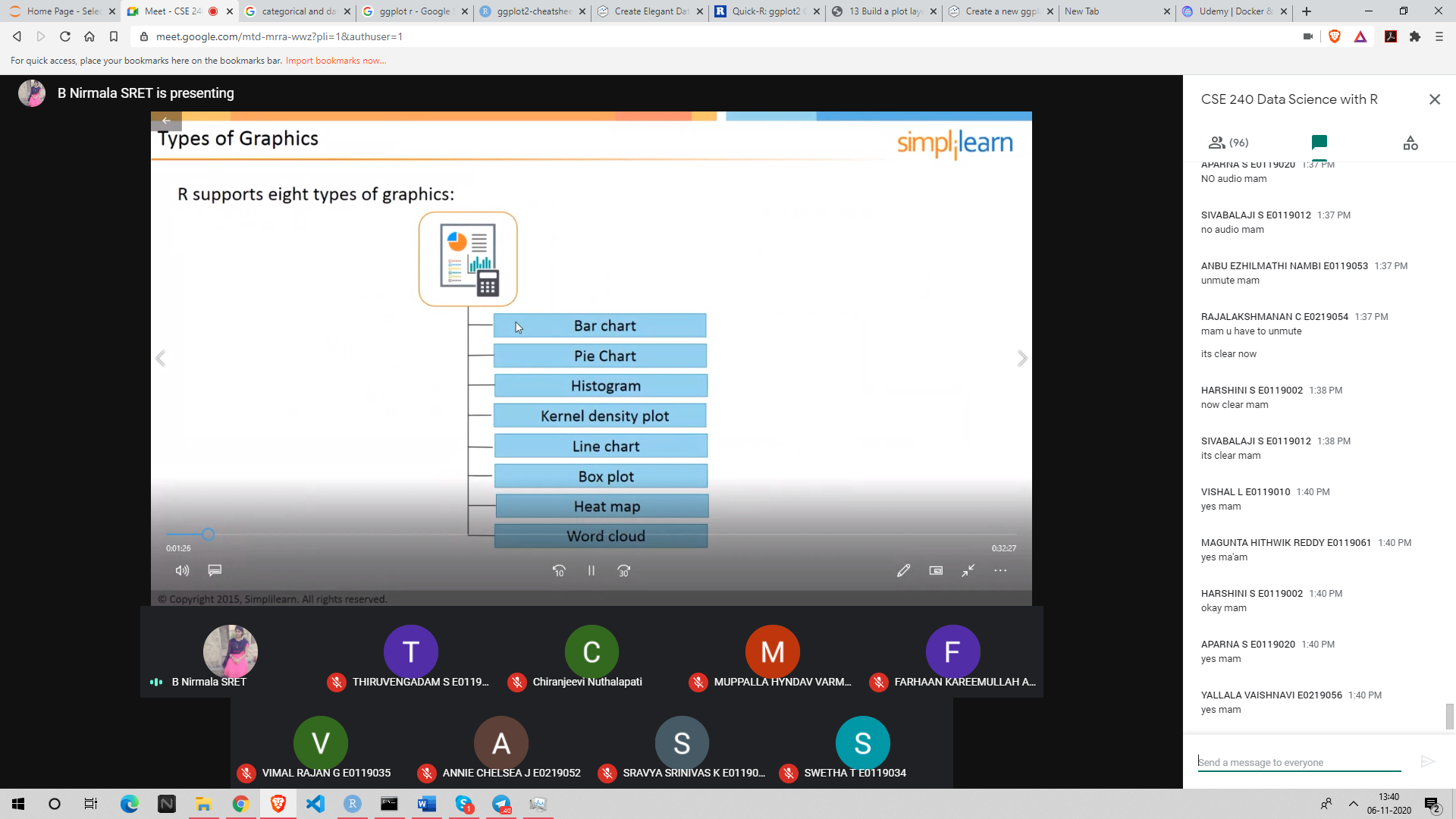
Facets

Statistics

Coordinates

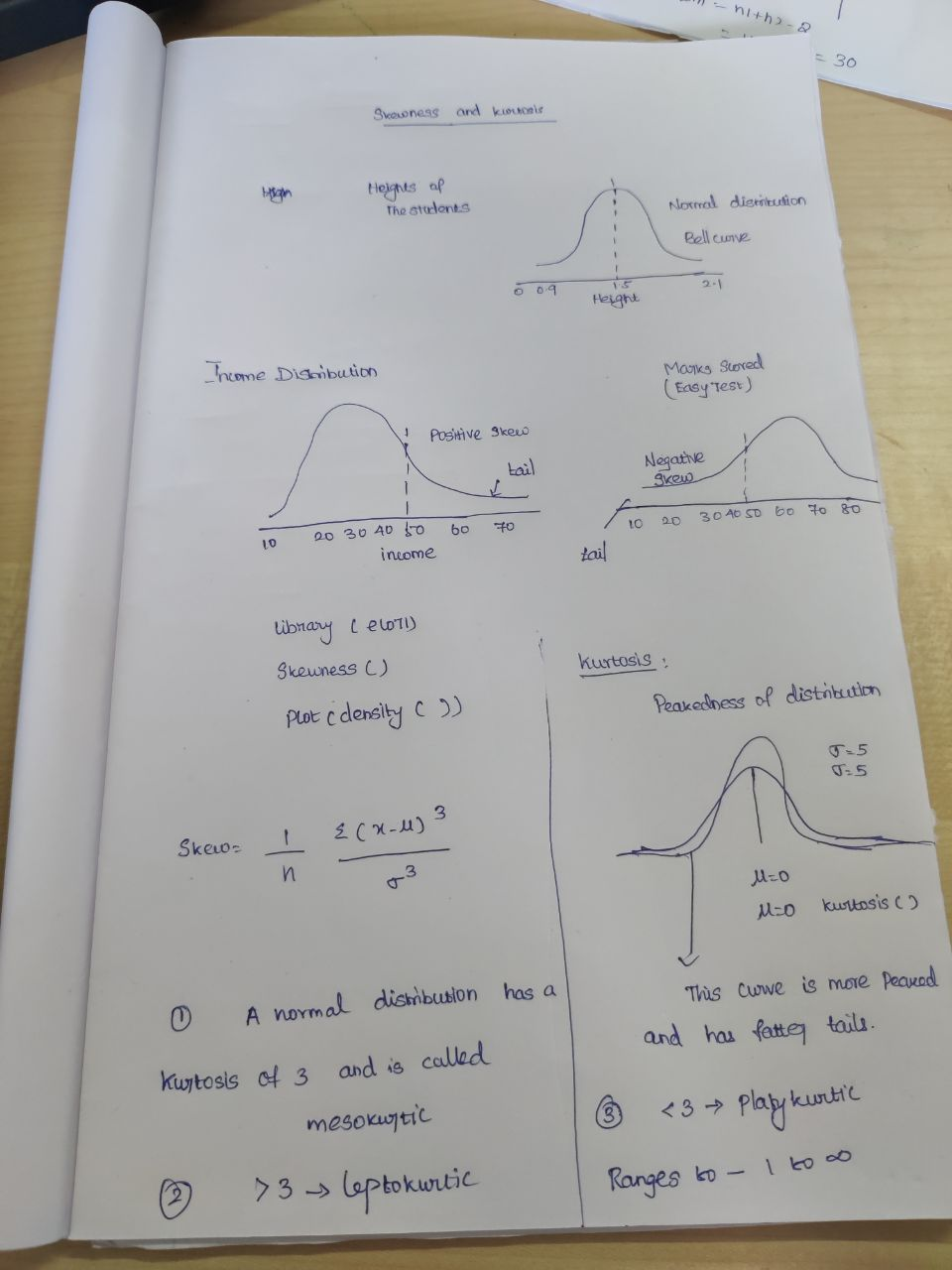
Themes

Plots except ggplot in r

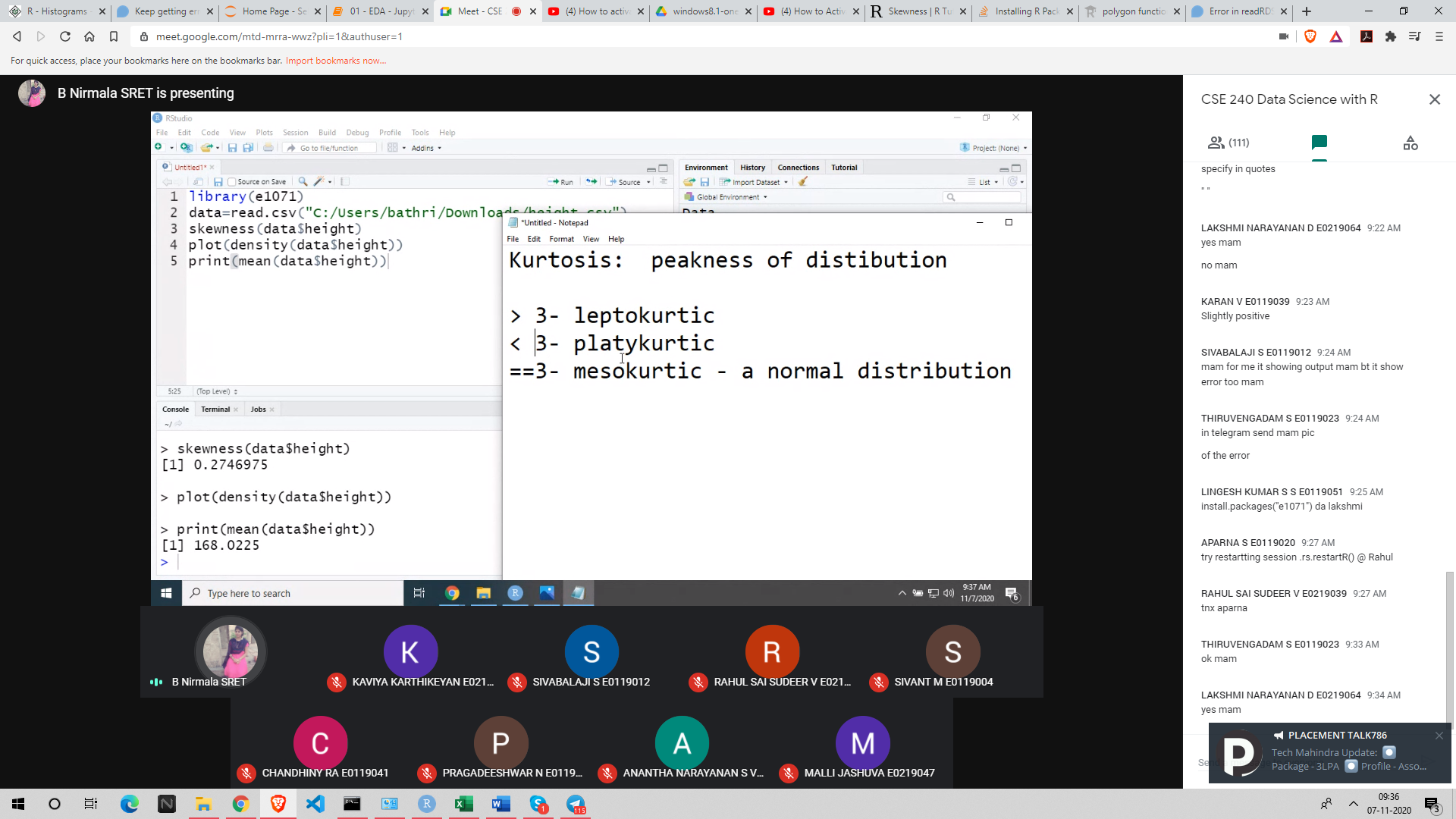


Module 4

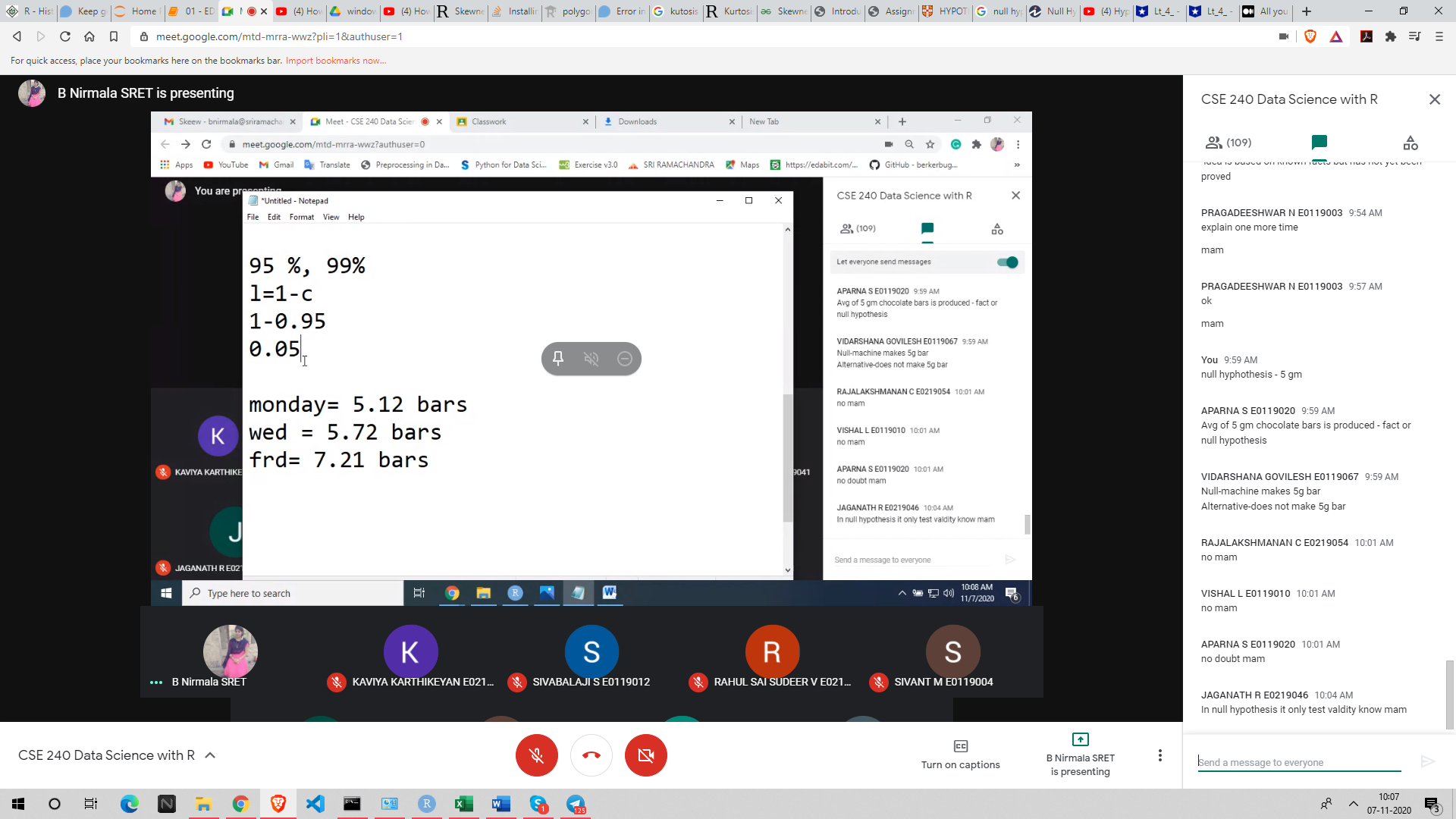
Data Summarization



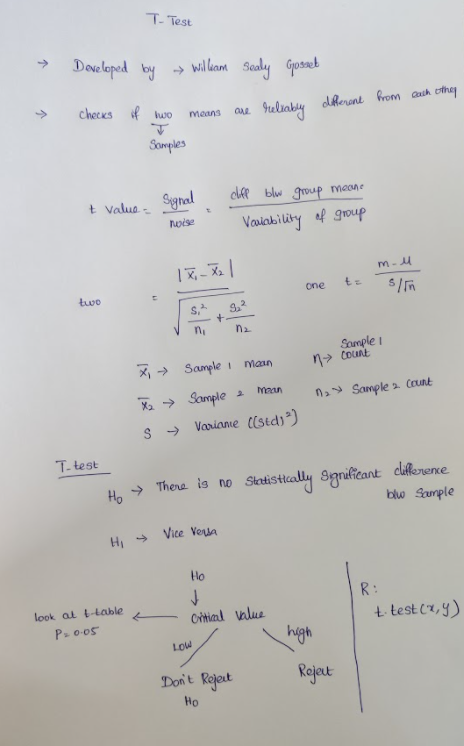
Skewness & kurtosis



Kurtosis levels



When to use t-test



1.when data is independent

2.normally distributed

3.similar amunt of variance

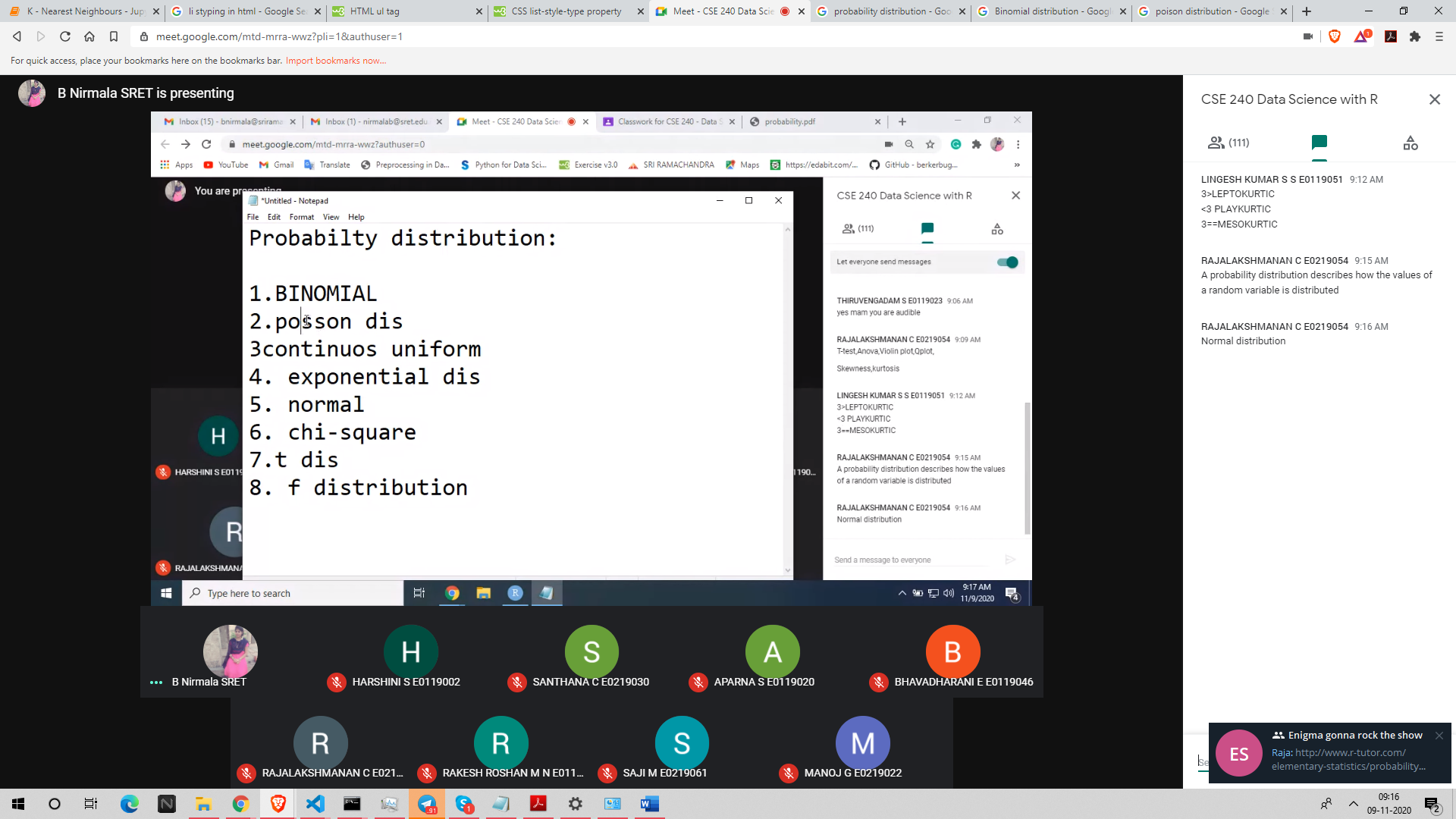
Data summarization:

Mean

Median

Mode

Standard deviation



<http://www.r-tutor.com/elementary-statistics/probability-distributions>

Suppose ther e12 mcq in english class each question has 5 possible answers and only one crct finf a probabilty of having 4 or less crct answer if a student attempt the answer randomly