Descriptive Statistic — Summarization of data without adding or subtracting anything of Measures of Central tendency at a specific instance. 1) Measures of Central tendency 2 Measures of dispersion 3 Measures of symmetricity D Measures of Central tendency

Central

1,2,3,4,5 what is one value around which all the data is revolving?

—3 \* CT reforesents the center paint ef a dataset. (1) Mean 2) median EDA and feature Engineering 3 Mode

Population - 
$$\{1, 2, 3, 4, 5\}$$

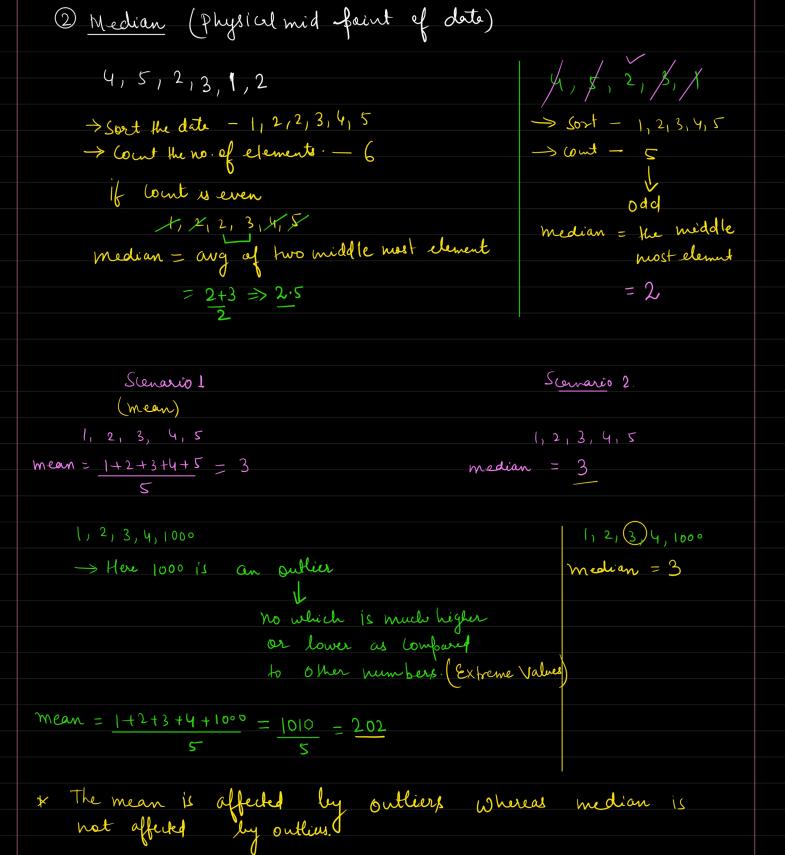
$$M = \sum_{i=1}^{N} \chi_{i}^{2}$$

$$\chi_{i}^{2} = \sum_{i=1}^{N} \chi_{i}^{2}$$

$$\chi_{i}^{2} = \sum_{i=1}^{N} \chi_{i}^{2}$$

$$\leq \Rightarrow$$
 Summetion  $\Rightarrow 1+2+3+4+5 = 15 = 3$ 

Summing up all the observations and dividing by no of observation.



1 1 1

3 mode Frequency maximum				
3 modre — Frequency maximum				
\\ \begin{align*} \begin{align*} 2, \\ 3, \\ 1, \\ 1, \\ 4, \\ 4, \\ 3, \\ 4, \\ 2 \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \				
mode = 4				
	_			
Use Case	s of Central	l tendency		
	0	(		
Λ α.			Cal /	
Age	e Jender	Weight	Salary (k)	
25	M	86		
26	\ \ \ \	70	50	
7 24	rs M	30	60	
2-3	<u> </u>	<del></del>	70	
25	F	60	45	
- Ago is a firmary variable				
-> Age is Continuous variable -> impute the missing   mull value with mean				
> inpute the missing full value with mean				
25+26+23+25 = 24.75				
> gender > Calegorical data				
Tighest frequency				
> M,M,M, F >> Mode - M				
* Median is not affected by outlier.				
date/Age				
Continuous data Calegorical				
Outlier				
$\langle A \rangle = \langle A $				
is present owner replace impute missing present values with mode				
median				
mean				