

Department of Computer Science and Engineering
Data Science

Semester: 1 Subject Statistics-for AIDS Academic Year: 2033-2024	
Estimates of Locatation: - Variable with measured data	
- Mean night have thousands of distinc	L-
* Median values of basic step in emploring	
* Robust your data is getting a "typical val	ue
+ Outlier . for each variable cue an estimale	-
- of where the most of the date is	
- located (i.e) its contral fendancy).
The central tendency can be found by the following	: -
The most basic estimate of location is the mean, or	
arrage value. The mean is the sure of all values	
divided by the number of values. The formula	
to compute the mean for a set of nealure x, x2 x	in i
$Mean = \overline{x} = S \times \overline{x}$	
Consider the following set of numbers: 13 5 12 3. The mean is (3+5+1+2) = 11/4 = 2.75	
The mean is (3+5+1+2) = 11/4 = 2.75	
Exemple:	
Median:	
The median is the middle number on a sortled -	
list of the data	
The middle or the central number become the	
niedian.	

Subject Incharge But Sarala Magage No. 5



Department of Computer Science and Engineering Data Science

Semester: D Subject Statistics for ATDS Academic Year: 20 93 20 94
* To calculate the median in an even numbers of
data set, select the middle two numbers and
find the average of Et.
niddle number is considered as madian. Example:
middle number is considered as median.
Example:-
Consider that we have a data set [1,2,5,4,5]
The mean for this dalasel = 1+2+5+4+5 = 15-3
S The state of the
If suppose a additional data point 60 is added
to the existing dalaset. The date set becomes at like
this: [1,2,3,4,5,50]. Now the mean for
the dalacet is 1+2+3+4+5+50 65 13
in a complete different formal of distribution
in a complete different formal of distribution
and this is harmful to do any statistics.
The predictions worst be accurate. In this can's
we calculate the median. The median value for
the above example is: 3+4 = 3.5. Nova the value
Tot is closed to the mean 3.
Compared to the mean, which uses all the observations,
the median depends only on the values in the certie of
Subject Incharge : Brof Sarala Manpage No. b Department of CSE-Data Science APSIT



Department of Computer Science and Engineering
Data Science

Semester: Subject Statistics for ATDS Academic Year: 2023 2024
the forted data. This night seem to be a disadvanto
but there are many inclances in which the median is
a better motric for location.
In the above example the problem caused was due
to the value 50. This value is known as out-lies
200
the other values in adalaset.
the other values in adalacet.
* Being an outlier in itself does not make a dala
value invalid. Still outlier are the result of
dala egrogs.
* when outliers are the result of bad data, the mean
will result in a poor estimate of location, while
the median will still be valid.
* In any case, outlier should be identified and are
esually worthy of further investigation.
TOURIDIE -
Tet consider a example to demonstrate the concept of Estimates of location.
of Estimates of location.
Consider there is a botch of petudente who have
scored the following marks in one of the subject.
\$20,10,40,303. The school has introduced a system
that if the average score of the batch is less
than 35 , then extra class has to be arranged for

Subject Incharge Prof Sarala Marypage No. 7

Department of CSE-Data Science | APSIT



Semester: Subject: Bloch Stice for AIDS Academic Year: 2023-2024
the students, so that the nesults can be Emproved!
Her taking the mean, we can come to the conclusion
whether special class is needed for this class or
Mean of . {20, 10, 40, 30 }
Mean = 20+10+40+30 100
Mean = 25
The conclusion draw from the mean value is that they very much need the special class to be conducted immediately.
Just imagine a new student is added to the batch
and his score is 90, Now the new dataset is
120,10,40,30,903.
Mean = 20+10+40+30+90 = 190
January State of State of the S
Mean = 38
Because of just one student the Mean >35, whenis
the conclusion results that the batch does not
nequire special class. Because of this the
require special class. Because of this the students will suffer. The conclusion is not
accurate:





Semester: V Subject: Statistics for ATDS Academic Year: 20.33. 20.94
The solution for this problem is median. The outlier in
this race is 90.
Median:
Sort the dataset = \$10,20, 30,40,903
The middle value is [30]
Nedian= 30
30 is near to the mean value. The conclusion is that
batch needs special class. Because of median the
Conclusion is correct
The disadvantage faced by the mean is somehow
solved by the meshian.
If supporte another student is added to the batch
with the score of 100. Let's check the median.
Median {10,20,30,40,90,100}
= 50140
Median = 35
Mode:-
Mode is the most frequent number - that is, the number
that occurs the highest number of times.
Example:
Consider the delaset of 10, 20, 10, 30, 40, 10, 90 4
20→1 90→1
30 →1 Mode=3



A.P. SHAH INSTITUTE OF TECHNOLOGY

Department of Computer Science and Engineering

Date Science



r , D !! /	Statistics for Al		emic Year: 202	
Cakulate the center	of tendercy	for the	below	giren
data:	. 0	Ma 15.	lowel	
Formulas to calcul	ale Mean,	Meelias	arces	2330
Mode for Grouped.	dalā	A C	CONTRACTOR OF THE PARTY OF THE	
		Design College	100000	
Mean = X = :	Ef x		11/11	
	Ef	ALL PROPERTY OF		
1	Est.		A DESCRIPTION	-
Median = L -	- 1/2 -cf	1	Telegraph I	
	4		The same	200
The state of the s		1	CRC16HOD	Bear Inc.
Mode = L+	/ fi-to)h	- intra-	
	2/1-fo-fo	1	100	
1	(
			HAN THE	
			:	State
	A STANSON OF THE PERSON OF THE	THE PERSON	SINE	1
	MULTINESS OF THE PARTY OF	Sand I	22127	the contract
Seem the ends		AND THE REAL PROPERTY.	TENTO.	15年1
	MALEUN COMPANY	ASSESSED IN	mollins)	HEREN
	1 1 7 7		£ 12-2 mod 4	
	A contract		S toronto	
	The second			



PARTICIPATE CHARTAGE TRUST

A.P. SHAH INSTITUTE OF TECHNOLOGY



Semester: 0.	Subject : s	glatitics fo	ATDS A	cademic Year: 2023- 2024
Brownlas	Wann Day For	mused do	la .	the second state of
(alulate	the mean.	for the s	given follo	roing dala
Find average	ge number of	defective	bulb Po	each boxes.
No at date	dire No. of	Class	f.x	The state of the s
Bulbs.		mark (x)	(Total	L:
0-2	3	(1) AL	1 = 3 ain	Statedire
2-4	4	3	12	Build
4-6	5	5	25	
6-8	3	4	21	LE BURELLE
8-10	1	9	9	
Total	£4)-16		5-fix=7	0
Mean	= X = Ef. >			The Marian
	Zf.		Marian St.	XM - 1 (b / 2)
	X = 70 (Total No of	defective	bulbe)
A STATE OF THE PARTY OF THE PAR		Total 1		
STREET, SQU	= 4-38	\$ 4.		The state of the state of
Average 1	lumber of de	Pective bu	167 -	The state of the s
		(Mean)	9=14	I. Carlot
	7-7-9-5-1			
		THE PARTY NAMED IN	124	a saute
A THE PARTY	A CONTRACTOR			THE PARTY OF
-			THE RES	
	the state of the s		100	
-	THE PERSON NAMED IN			



B CHAH INSTITUTE OF TECHNOLO



Semester: 1 Example - Ma Find the mea	Frequency	Cumulative	.,	
	(4)	Frequerry	2	Total No of
(0-10)	5 lodish	5	Lis	2.
10-20	10	15		+ 20 .
20-30	12	27 ->	બ .	_
(30-40	15	42)	Median cla	u).
(40-50)	18 Eliebalt	60		
Total	≤fi=60		Deliver to	The state of
Here 1 Find N/a =	Total No o	of students	= 60 = 30	
Here 1	V = 60 .	A sludente	- 60 - 70	1
Find N/a =	Total No. 0			
Find N/a =	Total No o	e frequenc	y crosse	1 30. The
Check when the	Total No o	e frequenc	y crosse in clas	1 30. The
Check when the cumulative =	Cumulative forequency C	e frequence rouses zo	y crosse in class	1 30. The 1 Preterval
Check when the	Cumulative forequency C	e frequence rouses zo	y crosse in class 1 → Lower 1 N/2 → 30	1 30. The 1 Preterval 1 mit →30
Check when the cumulative = 30-40. Median =	Total No. of 2. Cumulative frequency C	e frequence rouses zo	y crosse in clay 1→ Lower 1 N/2 → 30 cf → cf o	1 30. The 1 Preteral 1 mit → 30
Check when the cumulative = 30-40. Median =	Cumulative frequency C	e frequence rouses zo cf h	y crosse in clay 1 → Lower 1 N/2 → 30 cf → cf o	1 30. The 1 30. The 1 Preteral 1 → 30 1 paerious
Check when the cumulative = = = = = = = = = = = = = = = = = = =	Cumulative frequency C	e frequence rouses zo cf h	y crosse in class 1 → Lower 1 N/2 → 30 cf → cf o cta → 27 f → freq	s Preteral imit → 30 t paerious
Check when the cumulative =	Cumulative frequency C	e frequence rouses zo cf h h=10.	y crosse in clay 1 → Lower 1 N/2 → 30 cf → cf o cta → 27 f → frequence	1 30. The 1 30. The 1 Preteral 1 → 30 1 paerious



PARTIE AND TO CHARTENIA TRUETS

A.P. SHAH INSTITUTE OF TECHNOLOGY



Semester:	4		Subject :	Statis	treefer	AIDS	Academic Y	ear: 2023-202.4
Exami	le:-	Mede	for	Grou	ped d	ala	10000	
- Consid	er th	e befor	o giv	en d	مام ه	and co	lalate	the Made
Mark	4	Freque				- 44		ALTERNATION OF THE PARTY OF THE
0-10		2	0					
10-20		15		→ f	14,4			
20-30		6	-	-> Ma	dal clas	1 (HI)		
30-40		5	_	→-f2		641		Topic Williams
40-50		2						
								- A
M	ode:	-L+/	f1	fo)	L			BECKET SE
		(-	211	0-12				Was lat
-0-	> val	ue of m	odal	lais.	→6			Market
Co -	Pre	quency	6	prevle	our cle	aus -	-5-	
£2 -	Pre	quency	of	nent	clas	1 - 5	THE	1.19
		es wid				1100	USE COL	TO AT LOSS
		ves Linu						
	1000000							
Ma	J.	20+[6.	-5	7			
PLOC	26 =			5-5	HIC)——		
	-		- 0	3 - 3				
		90+1	- 1		1 .	9		5
	-	20+	12		×10	= 20	+ 1 x	10
			14-	5-5			~	
-		-						
Mod	Le =	25					-	Bulling
		-						



PARTINWASATW CHARITABLE TRUST

A.P. SHAH INSTITUTE OF TECHNOLOGY



Paleulett the n	Trequency (fi)	cf.	
0-20	5	ب	A Property Land
20-40	8	13	The state of the
(40-60	15	28	Hedian d
60-80	16	44	
80-100	6	50	
Hotel	£1:=50		100 4
= 40+	[25 - 13] x 20 15] x 20 15] x 20 15] x 20		
= 40+16 Median =3	= 56.	TALL!	-15 YA
	NAME OF TAXABLE PARTY.	1	



A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Date Science



class Internal	Trequency fi	
8-4	.1	
4.5	#	
5-6	28	
6-7	78	
7-8	84> Noc	lal class (fx)
8-9	45	
9-10	28	A STATE OF THE STA
10-11	7	21 - 41
11-12_	2_	Daniel de la company
fi -> 84., fo	- 18 - 12 - 45 - h + (84 - 78	-1 , L - = ·
Mode = I		
	(2×84) -78 - 45)	
	1 (6 168-78-45) - 74	K2 4515