PADENO.: OATE: / 1

Na	me		Sal	rshi	Si	no	7h
$\mathcal{B}_{\mathcal{C}}$	utch		- 77	02		. (
F	ou	No	. 11	19	061	2 3	3.
1	1	1		× *	11.1	١	*
	\		- 1	λ;	-4:	4	1.

Pata Mining and Data Wavehouse.

Discuss the privacy vissues in data mining with examples.

Ans - Poivacy Issues in data Mining-

1. Minimal Protection Setup-

Most of the time data is protected by security measures like anti-visuses, usernames, passwords etc. which doesn't protect the data in long term.

2. Access Contololer

Millimum

-Access contoiol verifies the identity of the person trying to access data. I Single layer access contoiol is not a very secured option.

	Sakshi (1906123)
	JUNS NI (1906123)
	DATE: / /
4:-	2. Non-verified Data Updation -
	Many times data is collected the roughly and updated without the verification of source.
	Thoroughly and updated without the
21	Obrification of source.
4.	Secretarity - Anchident Frequetion
	Security - joschitect Evaluation _
	To save money and time, a lot of
	onganizations Skip the process of audil- of
	the security whitech, which makes
	Casicy to hack the data collection.
- Control	A Comment of the Comm
<u>5</u> .	filtering and Validating External Sources -
10/1	Whenever an unauthorised device in
. 4	able to connect to the 3 ecuvily system.
	il- gives an entory point for vunera-
	bilities.
	ex - People bring office work home and access the official data via their
	and access the official data via their
	personal devices, which can create
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	a loophole.
· (*)	
<u> </u>	The state of the s
	The state of the s

PAGENO.: DAYE: / /

ADEM.	
2)	for a certain dataset, the values of The attoribute Age is given as
	The attoribute Age is given as
The state of the s	Jollous.
	Y
AND THE REAL PROPERTY.	25, 30, 15, 16, 33, 35, 70, 52, 13, 25, 33, 25,
With the second	40, 36,35,19,25,16,20,13,22,45,21,35,20.
a)	Find the mean, median, mode and mid-stange
	of the data.
	10 10 15 1/ 1/ 10 20 20 21 22 25 25
Soln.	13, 13, 15, 16, 16, 19, 20, 20, 21, 22, 25, 25,
	(25) 25, 30, 33, 33, 35, 35, 35, 36, 40, 45, 52, 70.
	25
N. C.	25
	$\Rightarrow \pi = 1 [2x13 + 15 + 2x16 + 19 + 2x20 +$
	25
N. Sanara	21+22+4×25+30+2×33+3×35
A Marie	
	+ 36+40+45+52+70.
	$-1 \times 1 = 1 \times 719$
	25
A STATE OF THE STA	= 28.76.
N. Maria	
1 1 1 2 1 2 1 SICH	

PAGINO:

	31 100	DATE: / /
		\cdot ; $n \rightarrow odd$.
	· Med	$lian = \pi \left[\frac{n+1}{2}\right] = \pi \left[\frac{25+1}{2}\right] = \pi \left[\frac{25+1}{2}\right]$
	iles de la des	
Land Company	, ,	Median = 25
T.	10 1 11 N. S. S.	
	deage	no. of students /people.
N.	13	2
	15-1	¥ 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
	16	2
	19	
STERNAL PRO	2011	2 1 2
	21 41.1	Γ
	22	
	522	· 4 ->: Highes-wo of people are of age 25.
	30	
	33	21 M L 05
	35	31 Mode = 25.
\$ 1 /2 A 1 -	36	
	in How	
	45	The day of the second s
	32	
	IO	
(6)		
1	77id3	Hange = min + max.
		2
	Carlotte Agent Company	
	31	midrange = 13+70 = 83
		2 2
		· midrange = 41.5
NA STATE	APPLICATION OF THE PROPERTY OF	

Bran

PAGE NO.

b) Calculate the varionce and Standard deviation.

Solu- n= 25.

En = 719.

n = 28,76

variance $(\sigma^2) = 1 \sum_{n=1}^{\infty} (n; -\pi)^2$

 $= \frac{1}{24} \left[\frac{(13 - 28.76)^2 \times 2 + (15 - 28.76)^2}{24} \right]^2$

 $+(16-28.76)^{2}\times2+(19-28.76)^{2}+(20-28.76)^{2}\times2$

 $+(21-28.76)^2+(22-28.76)^2+(25-28.76)^2\times4$

 $+(30-28.76)^2+(33-28.76)^2$ $\times 2+(35-28.76)^2$ $\times 3$

 $+(36-28-76)^2+(40-28.76)^2+(45-28.76)^2$

+(52-78.76)2+(70-28.76)2

 $\Rightarrow 6^{-2} = 1$ 2 x 248.3776 + 189.3376

+ 162.8176 *2 + 95.2576 + 2x 76.7376

PAGENO.: DATEL: / /

	+ 60.2176 + 45.6976 + 4×14.1376
	+ 1.5376 + 2x 17,9776 + 3x 38,9376
AL IVAL	+ 52.4176 + 126.3376 + 263.7376
J. Inc.	+ 540.0976 + 1,700,7376
3/4/	
	$\Rightarrow 6^2 = 1 \times 4260.56$
	$76^2 = 177.5233.$
	(a) (c) (b) (c) (c) (c) (d) (c) (c) (c) (c) (c) (c) (c) (c) (c) (c
	Standard deviation (+)= 5-2.
	7 6 - 177.5233
	- - 12.09
	The state of the s
20.00- 60.070 miles 2.477 miles	
Jan.	1. P. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1.
halling at the	CH CHANNE IN

	Sakshi (1906123)
	PAGE NO.:
- 01	Caire Ha Para Managhar Quinan
-9	Give the five-number summary of the data.
	of rock day.
SALIN	
72-6	13, 13, 15, 16, 16, 19, 20, 20, 21, 22,
-	25, 25, 21, 25, 30, 33, 35, 35, 35, 36,
	40,45,52,70.
100	
67 67 67	Minimum = 13.
	D., a 4 1:10 1 10.1.
	Quatitile I = (N+1) +4 term
	= (6.5) th term
	i.e. between 19 al 20. [6th - 7th]
	= 19+20 = 19.5
	$\frac{1}{2}$
	Median = 25
	54
	Quartile 3 = (N+1)3. the term
	4
	= 10 TU
	i.e. between 25 cost (101)
	i.e. between 35 apst [1944 - 20ta]
	- 3T215
3-	= 37+35 = 35

Maximum = 70

PAOTINO: DATE: / /

de Create a boxplot of the data.	
4 5 CONTE OF STORY	
Soln - Minimum = 13.	
Q1 = 19.5	
02 (MecUan) = 25.	
Q3 = 35.	
Maximum = 70	e .
	70
13 19 20 25 35	70
FE SAL ALIVE TO THE SECOND STATE OF THE SECOND SECO	,
	data
e) Duram a histogram to represent the	awq.
301n-	
Harry Constitution of the second of the seco	
10	is some
q -	n dan A M
8	
	90,5 (78)
4	
3 2 5 5 5	1
	0
9 - 9	
	1 99
0 5, 10 15 20 25 30, 35 40 45 50 55 60	65 70
Age of people.	1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1

Sakshi	1906	123)

PARTITION POL	n.c		
DAYE:	1	,	

3.	Evecate a dissimilarily matrix for	114
	the given data	, 0
9		

 						1 1		
	Item	Colows	Porice	Size.	.5	1.0		
 					1. p. 141	,	lke	~
	1	Blue	500	Small	-17	da	P-M	-
	2	Govern	100	Louige			7	
	3	Red	300	Small			_ F ·	
	4	Green	600	Medium	1			
		and the same of th	000	, ccum				

$$50 \text{ m} - P \rightarrow n0. \text{ of attributes} = 03$$

$$d(2,1) = 3-0 = 1$$

$$d(2,1) = 3-0 = 1.$$

$$d(3,1) = 3-01 = 0.67$$

$$d(3,2) = 3-0 = 1$$

$$(1(4,1) = 3-0 = 1$$

$$d(4,2) = 3-1 = 0.67$$

$$d(4,3) = 3-0 = 1$$

$$\begin{bmatrix}
d_{(2,1)} & d_{(2,2)} \\
d_{(3,1)} & d_{(3,2)} & cl_{(3,3)} \\
d_{(4,1)} & cl_{(4,2)} & cl_{(4,3)} & cl_{(4,4)}
\end{bmatrix}$$

Sakshi (1906123) 0 0 0 0.67 0 0.67