

Data Science & Artificial Intelligence



Warehousing

Part- 03

By- Vishvadeep Gothi sir





Topic : Datawarehousing



#Q. A student wrote 2 quizzes. In the first quiz, he scored 80 and in other, he scored 75. The mean and standard deviation of first quiz are 70 and 15 respectively, while the mean and standard deviation of the second quiz are 54 and 12, respectively. Which of the following is true?

- ☒ **A** Z score of second quiz is *greater* than that of first quiz,
- ☐ **B** Z score of first quiz is *greater* than that of second quiz,
- ☐ **C** Z score of both quizzes are equal
- ☐ **D** None

$$\frac{x_i - \mu}{\sigma}$$

Quiz 1:-

$$x_i = 80$$

$$\mu = 70$$

$$\sigma = 15$$

$$Z = \frac{80 - 70}{15}$$

$$= \frac{10}{15}$$

$$= 0.67$$

Quiz 2:-

$$x_i = 75$$

$$\mu = 54$$

$$\sigma = 12$$

$$Z = \frac{75 - 54}{12}$$

$$= \frac{21}{12}$$

$$= 1.75$$



Topic : Datawarehousing



#Q. Following Group Of Data: 26, 29, 62, 39, 43, 48, 23, 31, 40, 30.
The min-max normalized value of data 39 is _____?

↓
range \Rightarrow 0 to 1

$$x' = \frac{x - \min}{\max - \min} = \frac{39 - 23}{62 - 23} = 0.41$$



Topic : Datawarehousing



$$\text{Ans} = 0.85$$

#Q. The weight of chocolate bars from a particular chocolate factory has a mean of 80 grams with a standard deviation of 2 grams. What is the z-score corresponding to a weight of 81.7 grams?

$$\begin{aligned}\mu &= 80 \text{ grams} \\ \sigma &= 2 \text{ grams} \\ x &= 81.7 \text{ grams}\end{aligned}$$

$$\begin{aligned}Z &= \frac{81.7 - 80}{2} \\ &= \frac{1.7}{2} \\ &= 0.85\end{aligned}$$



Topic : Datawarehousing



#Q. Following Group Of Data: 124, 543, 247, 391, 443, 119, 296.
The min-max normalized value of data 247 in range [1,5] is _____?

newmin
newmax



Topic : Datawarehousing



$$Value' = \frac{Value_i - min}{max - min} (NewMax - NewMin) + NewMin$$

$$= \left(\frac{247 - 119}{543 - 119} \right) * (5 - 1) + 1$$

$$= 2.2075$$

$$\approx 2.21$$



Topic : Datawarehousing



$$\text{Ans} = -2.7$$

#Q. Books in the library are found to have an average length of 350 pages with a standard deviation of 100 pages. What is the z-score corresponding to a book of length 80 pages?

$$x_i = 80$$

$$\mu = 350$$

$$\sigma = 100$$

$$Z = \frac{80 - 350}{100}$$

$$= -2.7$$



Topic : Datawarehousing



Ans = 10

#Q. Huffman tree is constructed for the following data: {A, B, C, D, E}, with frequencies {0.25, 0.12, 0.19, 0.14, 0.20} respectively. BADE is encoded in _____ bits?

A : 0.25

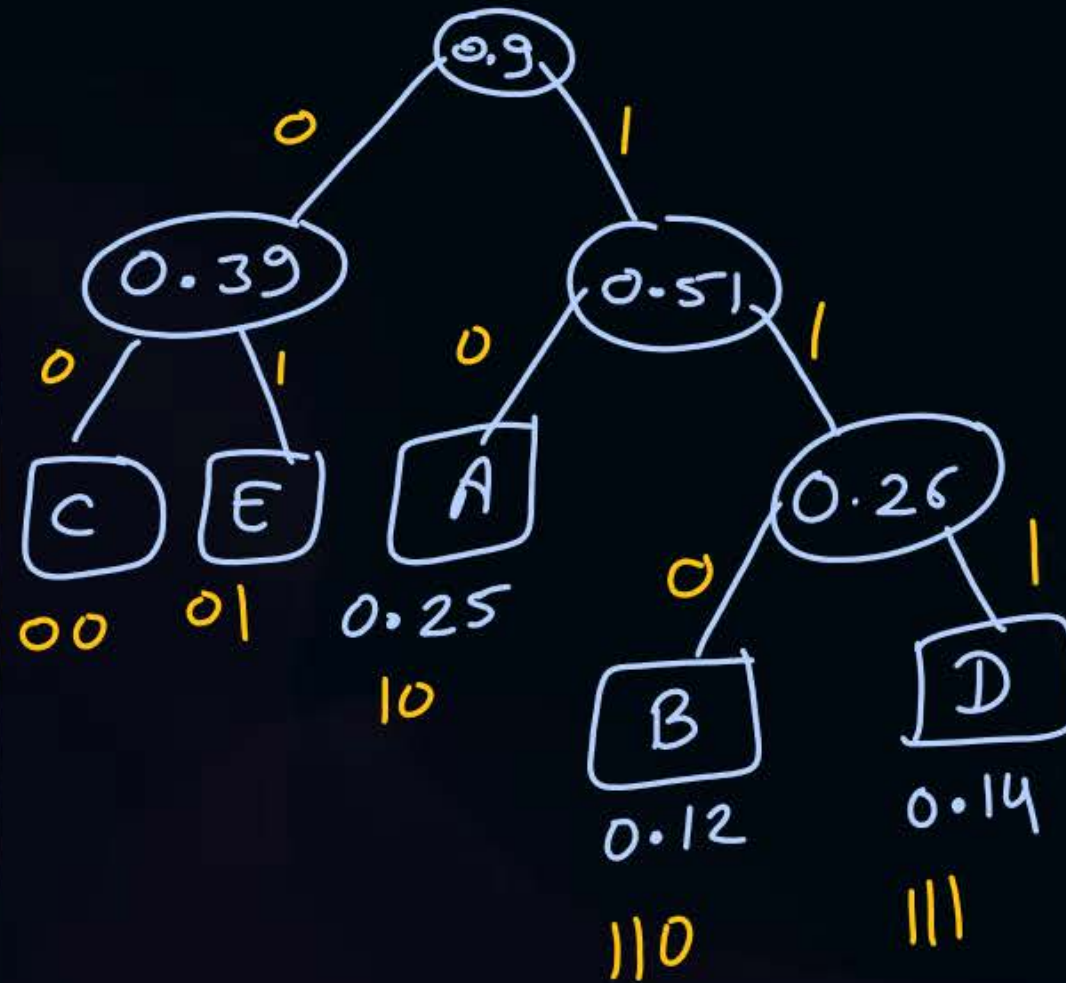
B : 0.12

C : 0.19

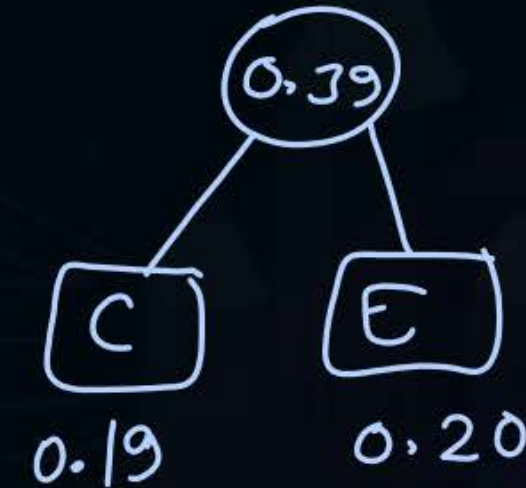
D : 0.14

E : 0.20

→ 0.26



110 10 111 01
10 bits





Topic : Datawarehousing



#Q. Following Group Of Data: 26, 29, 62, 39, 43, 48, 23, 31, 40, 30.
The z-score value of data 43 is _____?

$$\mu = 37.1$$

	$(x - \mu)^2$
26	123.21
29	65.61
62	620.01
39	3.61
43	34.81
48	118.81
23	198.81
31	37.21
40	8.41
30	50.41
	<hr/> 1260.9

$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$$
$$= 11.23$$

$$Z = \frac{43 - 37.1}{11.23} = 0.52 \text{ to } 0.53$$



Topic : Datawarehousing



Ans = 36 bits

#Q. Following message with just 2 unique characters A and M.

AMAAMAAMAMAMAAMMAMAM

Total number of bits needed to encode the above message in Lempel-Ziv-Welch Algorithm?

A → 0
M → 1

Position	A	M	AA	MA	AM	AMA	MAM	MM	AMAM
Sequence	1	2	3	4	5	6	7	8	9
Numerical representation	ϕA	ϕM	1A	2A	1M	5A	4A	2M	6M
Code	0000	0001	0010	0100	0011	1010	1000	0101	1101

$4 * 9 = 36$ bits

Ques In prev. questⁿ no. of bits saved by compression?

Solⁿ no. of characters = 20

without compression size of message = $20 * 8$ bits
= 160 bits

no. of bits saved = $160 - 36$
= 124 bits Ans.



Topic : Datawarehousing

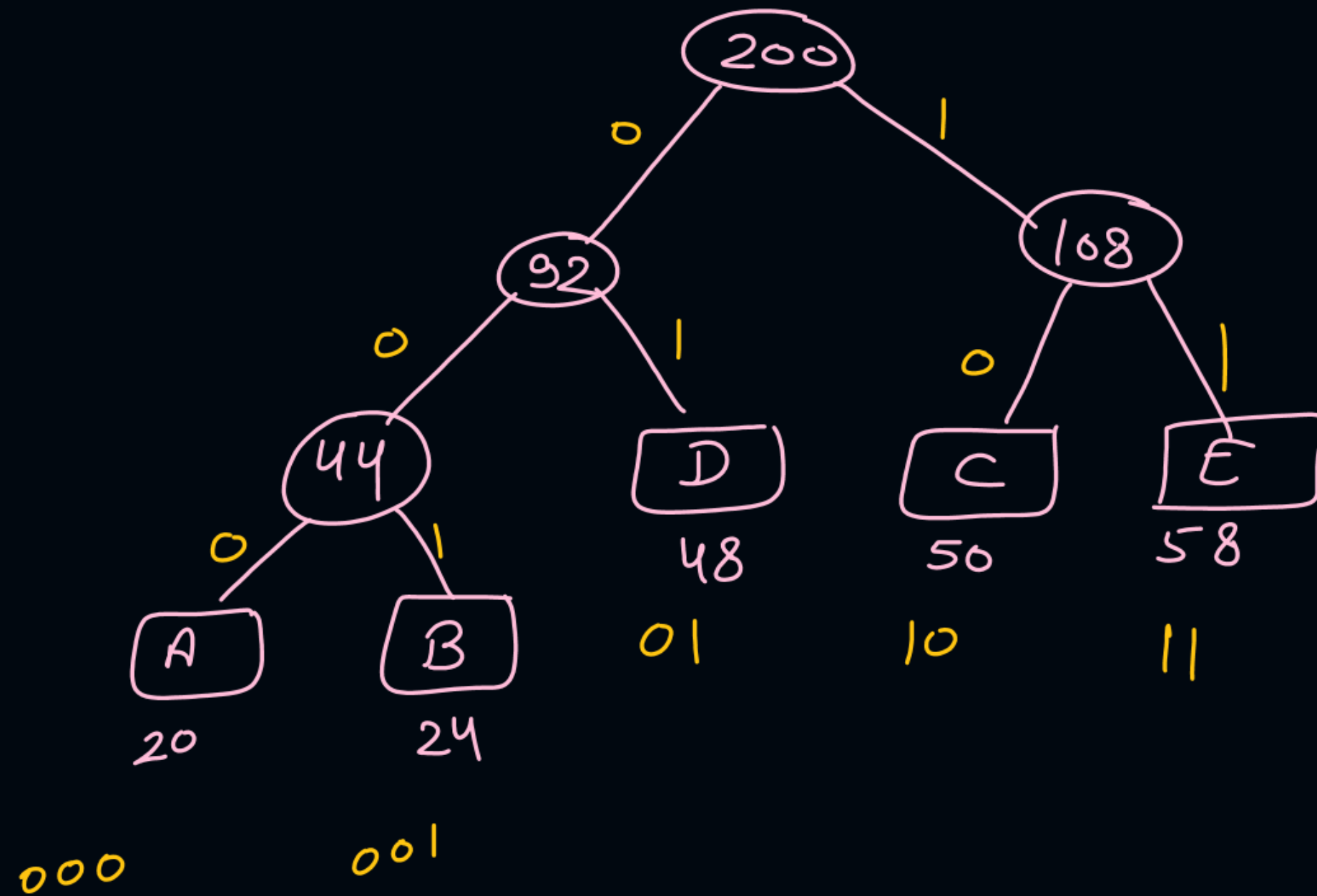


Ans = 444

#Q. Assume a message having 200 characters total with

- 98 {
- A: 20 times 000 $\Rightarrow 20 * 3 = 60$ bits
 - B: 24 times 001 $\Rightarrow 24 * 3 = 72$ bits
 - C: 50 times 10 $\Rightarrow 50 * 2 = 100$ bits
 - D: 48 times 01 $\Rightarrow 48 * 2 = 96$ bits
 - E: 58 times 11 $\Rightarrow 58 * 2 = 116$ bits
- Total = 444 bits

The size of encoded message is ____bits?



ques) In prev. questⁿ no. of bits saved by compression?

Solⁿ

$$\begin{array}{rcl} 200 * 8 & = & 1600 \text{ bits w/o compression} \\ & & 444 \text{ bits with } \underline{\hspace{1cm}} 11 \hspace{1cm} \\ \hline \text{saved} & = & \boxed{1156 \text{ bits}} \text{ Ans.} \end{array}$$



Happy Learning

THANK - YOU