

## University of Mumbai

Program: Computer Engineering

Curriculum Scheme: Rev2019 Examination: TE Semester V

Course Code: CSC504 and Course Name: Data Warehousing and Mining

Time: 2 hour 30 minutes

Max. Marks: 80

---

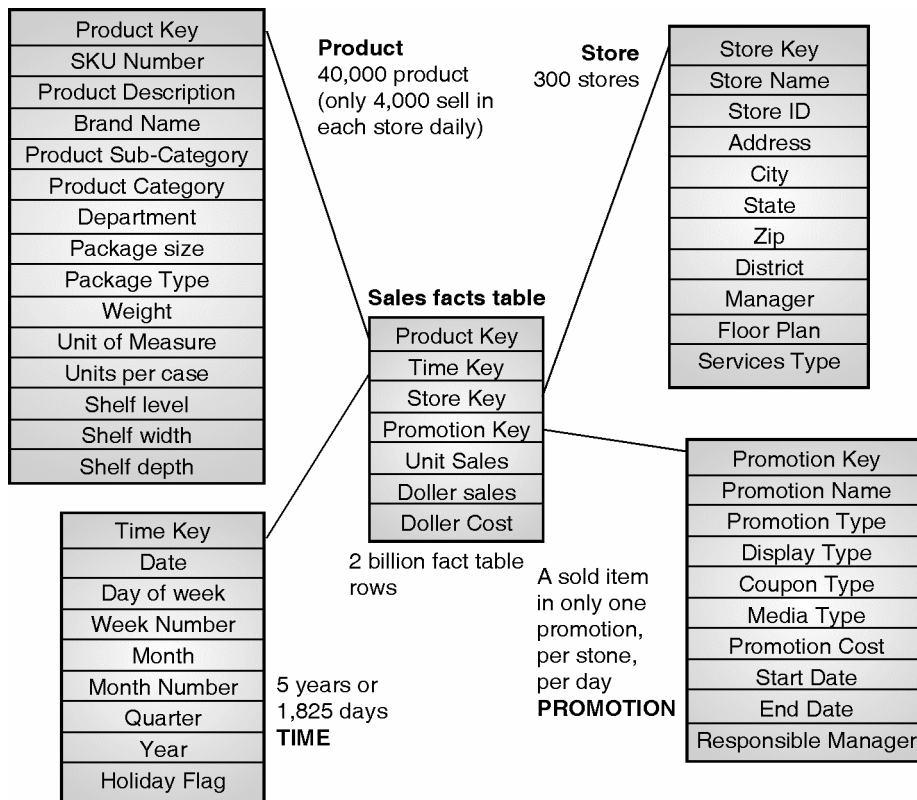
**Q1. Choose the correct option for following questions. All the Questions are compulsory and carry equal marks**

Question Number	Correct Option (Enter either 'A' or 'B' or 'C' or 'D')
Q1.	C
Q2.	C
Q3.	B
Q4	D
Q5	A

Q6	B
Q7	B
Q8.	D
Q9.	B
Q10.	C

**Q2**

A.



(b) Time period = 5 years · 365 days = 1825

There are 300 stores,

Each stores daily sale = 4000

Promotion = 1

Maximum number of fact table records:  $1825 \cdot 300 \cdot 4000 \cdot 1 = 2 \text{ billion}$

B.  
cency Matrix

	0				
	1	0			
	1.41	2.24	0		
	1.41	1	2	0	
	1.58	2.12	0.71	1.58	0

4 : In the last step there are only two clusters to be combined they are, (A,B,D) and (C,E).

Step 1 : Closest clusters are merged where the distance is the smallest measured by looking at the maximum distance between any two point.

Since C, E is minimum we can combine clusters C, E.

	A	B	(C, E)	D
A	0			
B	1	0		
(C,E)	1.5 8	2.2 4	0	
D	1.4 1	1	2	0

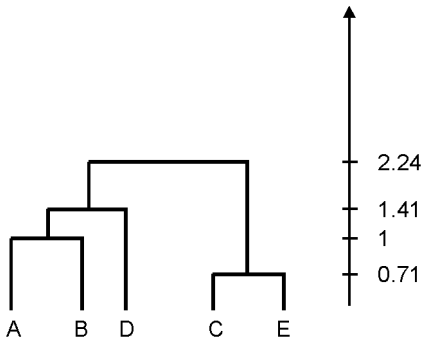
Step 2 : Now A and B is having minimum closest measure value therefore we merge these two clusters.

	(A, B)	(C, E)	D
(A,B)	0		
(C, E)	2.24	0	
D	1.41	2	0

Step 3 : Cluster (A,B) and D can be merged together as they are having minimum distance value.

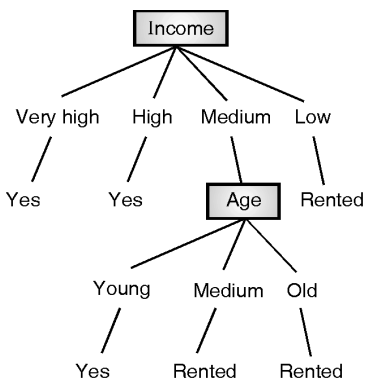
	(A, B, D)	(C, E)
(A,B,D)	0	
(C, E)	2.24	0

4A, 4B & 4C answers present in Notes given on classroom, refer same.

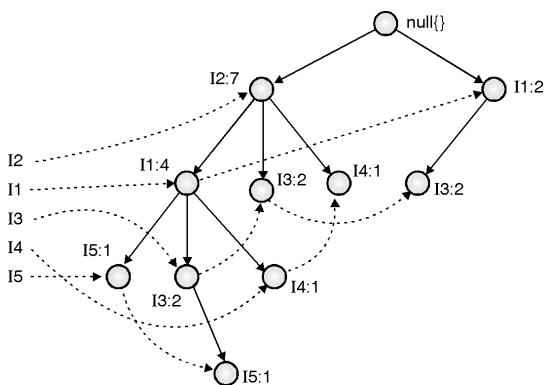


Final Dendrogram.

C.)



3A.)



3B.) Same as 2B, just change the distance formula and use min distance formula for Single link.

3B.) Solved in Class during lectures.