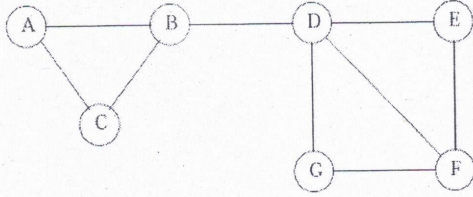


Option B:	1 hash functions in 1pass
Option C:	2 hash functions in 1pass
Option D:	2 hash functions in 2 different passes
8	Two k-cliques are adjacent when they share
Option A:	$2*k$ nodes
Option B:	$k+1$ nodes
Option C:	$k-1$ nodes
Option D:	k nodes
9	<p>Assume that a text file contains the following text. In a map-reduce logic of finding frequency of occurrence of each word in this file, what is the output of map function?</p> <p>This is a exam Yes it is exam</p>
Option A:	(This,1), (is, 2), (a, 1), (exam, 2), (Yes, 1), (it, 1)
Option B:	(This,1), (is, 1), (a, 1), (exam, 1), (Yes, 1), (it, 1), (is, 1), (exam,1)
Option C:	(This,1), (is, 1), (a, 1), (exam,1)
Option D:	(This,1), (is, 1), (a, 1), (exam, 2), (Yes, 1), (it, 1), (is, 1)
10	<p>In a map-reduce logic of finding Matrix-Vector Multiplication, what is the output of the map function?</p> $\begin{matrix} 3 & 4 & 1 \\ 5 & 6 & * & 2 \\ 7 & 8 & \end{matrix}$
Option A:	(1,11), (2,17), (3,23)
Option B:	(1,1,3), (1,2,4), (2,1,5), (2,2,6), (3,1,7), (3,2,8)
Option C:	(1,1,3), (1,2,4), (2,1,5), (2,2,6), (3,1,7), (3,2,8), (1, 1), (2, 2)
Option D:	(1, 3), (1, 8), (2, 5), (2, 12), (3, 7), (3, 16)

Q 2.	<p>Attempt ANY TWO QUESTIONS out of THREE</p> <p>Each question is for 10 marks</p>
A	Recall all NoSQL design patterns with examples. Justify CAP with suitable examples.
B	Explain with example Collaborative based filtering in a recommendation system.
C	<p>Apply Matrix - Matrix Multiplication using MapReduce model and solve the following example</p> $\begin{matrix} 1 & 2 & 3 & 4 \\ 1 & 2 & * & 3 & 4 \end{matrix}$
Q 3.	<p>Attempt ANY TWO QUESTIONS out of THREE</p> <p>Each question is for 10 marks</p>
A	<p>Apply PCY algorithm to find frequent itemset for the given dataset with minimum support 50% with hash function $h(ij) = i*j \% 8$</p> <p>T1-- 1,2,4,5</p> <p>T2-- 2,4,5</p>

		T3-- 1,2,4 T4-- 1,2,5
	B	Figure is an example of a social-network graph. Use the Girvan-Newman approach to find the between-ness of every edge. 
	C	Discuss all phases of the CURE algorithm for clustering with suitable example.
Q 4.		Attempt any FOUR Questions out of SIX Each question is for 5 marks
	A	What are the five Vs of Big Data? Explain.
	B	Recall Hadoop architecture with diagrams and give its advantages.
	C	Discuss any 5 different relational algebra operations with examples.
	D	Mention problems of PageRank along with its solution.
	E	State Bloom filter and explain with the help of an example.
	F	Explain KNN with proper example.