Feedback — Week 1 Quiz

Help Center

Thank you. Your submission for this quiz was received.

You submitted this quiz on **Wed 11 Feb 2015 2:17 AM PST**. You got a score of **9.80** out of **10.00**.

Question 1

Which of the following statements is true? Select all that apply.

Your Answer		Score	Explanation
☐ Data entry is part of the data mining phase.	✔	0.20	This is false since none of these tasks involve data analysis.
☑ The output of the pattern evaluation phase can be used in the data mining to improve the quality of patterns discovered.	✔	0.20	
☑ Clustering and classification are examples of data mining tasks.	✔	0.20	
☐ In the knowledge discovery process, pattern evaluation takes place before the data preprocessing phase.	✔	0.20	This is false since pattern mining takes place after data preprocessing.
☑ Data collection and data cleaning are examples of data mining tasks.	✖	0.00	This is false since none of these tasks involve data analysis.
Total		0.80 / 1.00	

Question Explanation

The correct answers are: "Clustering and classification are examples of data mining tasks." and "The output of the pattern evaluation phase can be used in the data mining to improve the quality of patterns discovered."

Question 2

Tid	Items bought
10	Beer, Nuts, Diaper
20	Beer, Coffee, Diaper, Nuts
30	Beer, Diaper, Eggs
40	Beer, Nuts, Eggs, Milk
50	Nuts, Coffee, Diaper, Eggs, Milk

Table 1: Transactions from a database.

Given the transactions in Table 1, *minsup s* = 50%, and *minconf c* = 50%, which of the following is an association rule? Select all that apply.

Your Answer		Score	Explanation
☑ Nuts ⇒ Diaper	✔	0.20	
☐ Coffee ⇒ Milk	✔	0.20	
☑ Beer ⇒ Nuts	✔	0.20	
☐ Nuts ⇒ Eggs	✔	0.20	
☐ Diaper ⇒ Eggs	✔	0.20	
Total		1.00 / 1.00	

Question Explanation

The correct answers are: "Beer ⇒ Nuts" and "Nuts ⇒ Diaper"

Beer ⇒ *Nuts* has support = $^3/_5$ and confidence $^3/_4$
Nuts ⇒ *Diaper* has support = $^3/_5$ and confidence $^3/_4$
Coffee ⇒ *Milk* has support = $^1/_5$ and confidence $^1/_2$
Nuts ⇒ *Eggs* has support = $^2/_5$ and confidence $^1/_2$
Diaper ⇒ *Eggs* has support = $^2/_5$ and confidence $^1/_2$

Question 3

Tid	Items bought
10	Beer, Nuts, Diaper
20	Beer, Coffee, Diaper, Nuts
30	Beer, Diaper, Eggs
40	Beer, Nuts, Eggs, Milk
50	Nuts, Coffee, Diaper, Eggs, Milk

Table 1: Transactions from a database.

Given the transactions in Table 1, what is the confidence and relative support of the association rule {Diaper} ⇒ {Coffee, Nuts}?

Your Answer		Score	Explanation
◉ support s = 0.4, confidence c = 0.5	✔	1.00	
○ support s = 0.8, confidence c = 0.5			
○ support s = 0.4, confidence c = 1			
○ support s = 0.8, confidence c = 1			
○ None of the above			
Total		1.00 / 1.00	

Question Explanation

The correct answer is: "support s = 0.4, confidence c = 0.5"

Question 4

Consider the database containing the transaction T_1 : {a_1, a_2, a_3, a_4, a_5}, T_2 : {a_2, a_3, a_4, a_5, a_6}. Let *minsup* = 1. Which of the following is a closed frequent pattern? Select all that apply.

Your Answer		Score	Explanation
☐ {a_1, a_2, a_3, a_4, a_5, a_6}	✔	0.20	
☐ {a_2, a_5}	✔	0.20	

☑ $\{a_2, a_3, a_4, a_5\}$	✔	0.20
☑ $\{a_1, a_2, a_3, a_4, a_5\}$	✔	0.20
☑ $\{a_2, a_3, a_4, a_5, a_6\}$	✔	0.20
Total		1.00 / 1.00

Question Explanation

The correct answers are: "$\{a_2, a_3, a_4, a_5\}$", "$\{a_1, a_2, a_3, a_4, a_5\}$", and "$\{a_2, a_3, a_4, a_5, a_6\}$".

"$\{a_2, a_5\}$" and "$\{a_1, a_2, a_3, a_4, a_5, a_6\}$" are not closed frequent patterns because "$\{a_2, a_3, a_4, a_5\}$" is a super patter of "$\{a_2, a_5\}$" with the same support and "$\{a_1, a_2, a_3, a_4, a_5, a_6\}$" has support 0, hence not a frequent pattern given *minsup* = 1.

Question 5

Consider the database containing the transactions $T_1 : \{a_1, a_2, a_3, a_4, a_5\}$, $T_2 : \{a_2, a_3, a_4, a_5, a_6\}$.

Let *minsup* = 1. Which of the following is a max frequent pattern? Select all that apply.

Your Answer		Score	Explanation
☑ $\{a_1, a_2, a_3, a_4, a_5\}$	✔	0.20	
☑ $\{a_2, a_3, a_4, a_5, a_6,\}$	✔	0.20	
☐ $\{a_2, a_5,\}$	✔	0.20	
☐ $\{a_2, a_3, a_4, a_5\}$	✔	0.20	
☐ $\{a_1, a_2, a_3, a_4, a_5, a_6\}$	✔	0.20	
Total		1.00 / 1.00	

Question Explanation

The correct answers are: "$\{a_1, a_2, a_3, a_4, a_5\}$" and "$\{a_2, a_3, a_4, a_5, a_6,\}$".

"$\{a_1, a_2, a_3, a_4, a_5\}$" and "$\{a_2, a_3, a_4, a_5, a_6,\}$" are super patterns of "$\{a_2, a_5,\}$" and "$\{a_2, a_3, a_4, a_5\}$", hence "$\{a_1, a_2, a_3, a_4, a_5\}$" and "$\{a_2, a_3, a_4, a_5\}$" cannot be max frequent patterns." $\{a_1, a_2, a_3, a_4, a_5, a_6\}$" has support 0, hence it is not a frequent pattern.

Question 6

Rank the following sets by their cardinality for a given database: {all frequent patterns}, {closed

frequent patterns}, {max frequent patterns}

Your Answer		Score	Explanation
○ Ranking in impossible without further information.			
○ {all frequent patterns} ≥ {max frequent patterns}, {all frequent patterns} ≥ {closed frequent patterns}, but the order of {max frequent patterns} and {closed frequent patterns} cannot be determined without furthur information.			
○ {all frequent patterns} ≥ {max frequent patterns} = {closed frequent patterns}, i.e. the set of max frequent patterns and the set of closed frequent are identical.			
◉ {all frequent patterns} ≥ {closed frequent patterns} ≥ {max frequent patterns}	✔	1.00	
○ {all frequent patterns} ≥ {max frequent patterns} ≥ {closed frequent patterns}			
Total		1.00 / 1.00	

Question Explanation

The correct answer is "{all frequent patterns} ≥ {closed frequent patterns} ≥ {max frequent patterns}"

By definition, closed frequent patterns are a subset of all frequent patterns, and max frequent patterns are a subset of closed frequent patterns. We can remove all patterns that have a super pattern in the closed frequent pattern set to obtain the max frequent patterns, hence max frequent patterns are a subset of closed frequent patterns.

Question 7

If we know the support of itemset {*a*, *b*} is 10, which of the following numbers are the possible supports of itemset {*a*, *b*, *c*}?

Your Answer		Score	Explanation
☑ 10	✔	0.33	
☑ 9	✔	0.33	
☐ 11	✔	0.33	
Total		1.00 / 1.00	

Question Explanation

The correct answers are: "9" and "10".

The support of {*a*, *b*, *c*} should be no more than 10, which is the support of the {*a*, *b*}.

Question 8

If we know the support of itemset {*a*} is 50, and the support of itemset {*a*, *b*, *c*} is 10, which of the following numbers are the possible supports of itemset {*a*, *w*}?

Your Answer		Score	Explanation
☑ 50	✔	0.20	
☑ 5	✔	0.20	
☑ 30	✔	0.20	
☑ 10	✔	0.20	
☐ 100	✔	0.20	
Total		1.00 / 1.00	

Question Explanation

The correct answers are: "50", "30", "10", and "5".

The support of {*a*, *w*} should be no more than the support of {*a*}. However, there is no relation between the supports of the {*a*, *b*, *c*} and {*a*, *w*}.

Question 9

Considering Apriori Algorithm, assume we have 5 items (A to E) in total. In the 1-st scan, we find out **all** frequent items A, B, C, and E. How many size-2 (i.e. containing 2 items, e.g. A, B) itemsets should be considered in 2-nd scan, i.e. are potential to be size-2 frequent itemsets?

Your Answer		Score	Explanation
○ 25			
◉ 6	✔	1.00	
○ 4			

10		
Total	1.00 / 1.00	

Question Explanation

The correct answer is "6".

6 in total, i.e. AB, AC, AE, BC, BE, CE. They are all possible combinations of the 4 frequent items.

Question 10

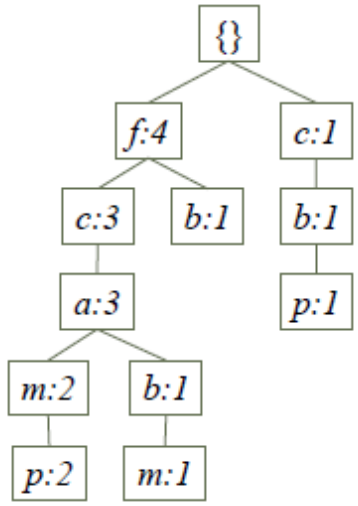

Figure 1: FP Tree

Given the FP-tree as shown in Figure 1, how many transactions do we have in total?

Your Answer		Score	Explanation
4			
1			
2			
5	✔	1.00	
3			
Total		1.00 / 1.00	

Question Explanation

The correct answer is: "5".

You can sum up the supports of all leaves or sum up the supports of the root's children.