	Lehrveranstaltung	Databases and Information Systems 2020		
	Aufgabenzettel	7		
	STiNE-Gruppe 14	Simon Weidmann, Aram Yesildeniz		
	Ausgabe	7. Juli 2020	Abgabe	15. Juli 2020

Report

1 How the algorithm and software works

See in the lower part of Apriori.java.

Before the algorithm, we parse all the transactions into a list.

First, the frequent 1-itemsets are being fetched from the transactions. Then they are filtered using MIN_SUP.


Then we loop for the next steps: We generate candidates given the algorithm from the extra slides, slide 7, including pruning.

Then we measure how well they are supported. Then we kick out all those under the support threshold.

We loop until one step has an empty result.

2 List how many frequent itemsets you obtained containing one, two, three, ... items.

1	379
2	11
3	1

	Lehrveranstaltung	Databases and Information Systems 2020		
	Aufgabenzettel	7		
	STiNE-Gruppe 14	Simon Weidmann, Aram Yesildeniz		
	Ausgabe	7. Juli 2020	Abgabe	15. Juli 2020

3 Provide the found frequent itemsets containing two or more items together with their computed support value.

Itemset:

(Articles: [39, 704] | Support: 1.39%)
 (Articles: [39, 825] | Support: 1.51%)
 (Articles: [217, 283] | Support: 1.04%)
 (Articles: [217, 346] | Support: 1.49%)
 (Articles: [227, 390] | Support: 1.01%)
 (Articles: [368, 682] | Support: 1.28%)
 (Articles: [368, 692] | Support: 1.14%)
 (Articles: [368, 829] | Support: 1.27%)
 (Articles: [390, 722] | Support: 1.03%)
 (Articles: [704, 825] | Support: 1.35%)
 (Articles: [789, 829] | Support: 1.19%)

Itemset:

(Articles: [39, 704, 825] | Support: 1.28%)