Minh Tran – INF 553 – HW 2

The frequent itemsets (containing a set of products purchased by the same customer at a particular day) are the following two:

First itemset: ('4710011401128', '4710011401135', '4710011405133', '4710011406123')

Second itemset: ('4710011401128', '4710011401135', '4710011405133', '4710011409056')

If we cross-reference these items to the product category, the following information are extracted:

• First itemset:

Т	RANSACTION_DT	CUSTOME	AGE_GRO	PIN_CODE	PRODUCT	PRODUCT_ID	AMOUNT	ASSET	SALES_PRI
	11/1/2000	217361	45-49	115	120103	'4710011401128'	2	46	56
	11/1/2000	1717723	40-44	115	120103	'4710011401135'	1	23	29
	11/1/2000	1873139	30-34	221	120103	'4710011405133'	1	23	29
	11/1/2000	1849332	45-49	Others	120103	'4710011406123')	2	46	55

• Second itemset:

11/1/2000	217361	45-49	115	120103	4710011401128	2	46	56
11/1/2000	1717723	40-44	115	120103	4710011401135	1	23	29
11/1/2000	1849332	45-49	Others	120103	4710011405133	1	23	28
11/1/2000	1849332	45-49	Others	120103	4710011409056	1	23	28

Notice that all 8 items in 2 itemset belong to the same product_subclass (120103), which means that they are likely serving a similar purpose (i.e. bathroom items or baby-related) and possibly located in the same section of the superstore. This will facilitate the shopping process for customers.