Data-Warehouse-Technologien
Exercise sheet 11

Assignment 1: What is Business Intelligence? Describe the Business Intelligence Process.

Assignment 2: Given is a shopping basket, with the following 10 transactions:

TID	Items
1	Anhaltinisch Flüssig, Berliner Bräu, Erfurter Bock, Riesling
2	Anhaltinisch Flüssig, Erfurter Bock, MT, Riesling
3	Ilmenauer Pils, MT
4	Anhaltinisch Flüssig, Dornfelder, Kölnische Weiße, Riesling
5	Berliner Bräu, Dornfelder, Kölnische Weiße, MT, Riesling
6	Ilmenauer Pils, Kölnische Weiße, MT
7	Anhaltinisch Flüssig, Erfurter Bock
8	Dornfelder, Erfurter Bock, Riesling
9	Berliner Bräu, Dornfelder, MT
10	Anhaltinisch Flüssig, Erfurter Bock

Determine possible rules using the Apriori-algorithm. Hereby, only consider items with a support $\geq 30\%$.

Assignment 3: Construct a decision tree determining the characteristics of the column "Too expensive" with the facts of the given table. Assume that year has only three possible values and that the prices can be classified in the following categories: cheap(0-9 €), medium (10-29 €), expensive (30-50 €). By equal selectivity of two attributes, use the order of the attributes within the table.

TID	sort of wine	year	price	Too expensive?
1	Riesling	2010	7 €	Ja
2	Merlot	2010	5 €	Nein
3	Riesling	1980	30 €	Nein
4	Pinot Noir	1980	42 €	Ja
5	Müller Thurgau	2002	24 €	Nein
6	Merlot	2002	17 €	Ja
7	Pinot Noir	2010	8 €	Nein
8	Merlot	2002	9 €	Nein
9	Pinot Noir	2002	33 €	Ja
10	Müller Thurgau	1980	28 €	Nein

Given the wines of the following table, determine the decision of your decision tree.

TID	sort of wine	year	price	Too expensive?
11	Müller Thurgau	2010	45 €	?
12	Merlot	2002	8 €	?
13	Riesling	1980	44 €	?
14	Pinot Noir	2010	22 €	?

Is the result of your decision tree always comprehensible? How can the results of decision trees be improved?

 Good Luck!