

Business Analytics minor



UC Leuven
Limburg
MOVING MINDS

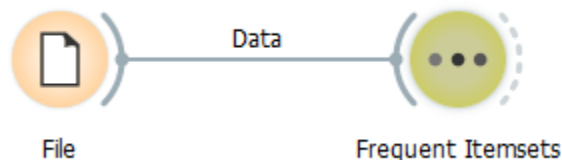
Week 7: Exercises



Exercise 9.1: Titanic

1. Get the itemsets with minimal support of 60%:

- Install the Associate Add-On to Orange
 - Options > Add-Ons > Associate
- File
 - Titanic
- Frequent Itemsets
 - Minimal support: 60%
 - Click Find Itemsets



The screenshot shows the 'Frequent Itemsets' widget interface. The left panel contains settings for finding itemsets, and the right panel displays the results in a table.

Find itemsets settings:

- Minimal support: 60% (slider)
- Max. number of itemsets: 10000 (slider)
- ☐ Find Itemsets (button)

Filter itemsets settings:

- Contains:
- Min. items: 1, Max. items: 999 (spinners)
- ☒ Apply these filters in search
- ☒ Send Selection Automatically (button)

Itemsets table:

Itemsets	Support	%
survived=no	1490	67.7
▼ age=adult	2092	95.05
survived=no	1438	65.33
▼ sex=male	1667	75.74
survived=no	1329	60.38
▼ sex=male	1731	78.65
survived=no	1364	61.97



Exercise 9.1: Titanic

1. Get the itemsets with minimal support of 60%:

▪ Frequent 1-itemsets?

- Survived = no
- Sex = male
- Age = adult

▪ Frequent 2-itemsets?

- Adult and Did-Not-Survive
- Adult and Male
- Male and Did-Not-Survive

▪ Frequent 3-itemset?

- Adult and Male and Did-Not-Survive
- Monotonicity: all subsets of a frequent set must be frequent

▪ How many people did not survive?

- 1490

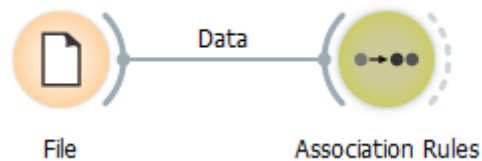
Itemsets	Support	%
survived=no	1490	67.7
▼ age=adult	2092	95.05
survived=no	1438	65.33
▼ sex=male	1667	75.74
survived=no	1329	60.38
▼ sex=male	1731	78.65
survived=no	1364	61.97



Exercise 9.1: Titanic

2. Find the association rules with minimal support 40% and minimal confidence 20%.

- File
 - Titanic
- Association Rules
 - Minimal support: 40%
 - Minimal Confidence: 20%
 - Click Find Rules



Association Rules

Info

Number of rules: 14
Filtered rules: 14
Selected rules: 0
Selected examples: 0

Find association rules

Minimal support: 40%
Minimal confidence: 20%
Max. number of rules: 10000

☐ Induce classification (itemset → class) rules

Filter rules

Antecedent

Contains:
Min. items: 1 Max. items: 999

Consequent

Contains:
Min. items: 1 Max. items: 999

☒ Apply these filters in search

☒ Send Selection Automatically

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent	Consequent
0.757	0.963	0.786	1.209	1.013	0.010	sex=male	→ age=adult
0.757	0.797	0.950	0.827	1.013	0.010	age=adult	→ sex=male
0.653	0.687	0.950	0.712	1.015	0.010	age=adult	→ survived=no
0.653	0.965	0.677	1.404	1.015	0.010	survived=no	→ age=adult
0.620	0.788	0.786	0.861	1.164	0.087	sex=male	→ survived=no
0.620	0.915	0.677	1.162	1.164	0.087	survived=no	→ sex=male
0.604	0.797	0.757	0.894	1.178	0.091	age=adult, sex=male	→ survived=no
0.604	0.768	0.786	0.831	1.175	0.090	sex=male	→ age=adult, survived=no
0.604	0.635	0.950	0.652	1.025	0.015	age=adult	→ sex=male, survived=no
0.604	0.974	0.620	1.534	1.025	0.015	sex=male, survived=no	→ age=adult
0.604	0.924	0.6352772466539197			0.090	age=adult, survived=no	→ sex=male
0.604	0.892	0.677	1.119	1.178	0.091	survived=no	→ age=adult, sex=male
0.402	0.423	0.950	0.423	1.052	0.020	age=adult	→ status=crew
0.402	1.000	0.402	2.364	1.052	0.020	status=crew	→ age=adult



Exercise 9.1: Titanic

2. Find the association rules with minimal support 40% and minimal confidence 20%.

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent		Consequent
0.757	0.963	0.786	1.209	1.013	0.010	sex= male	→	age= adult
0.757	0.797	0.950	0.827	1.013	0.010	age= adult	→	sex= male
0.653	0.687	0.950	0.712	1.015	0.010	age= adult	→	survived= no
0.653	0.965	0.677	1.404	1.015	0.010	survived= no	→	age= adult
0.620	0.788	0.786	0.861	1.164	0.087	sex= male	→	survived= no
0.620	0.915	0.677	1.162	1.164	0.087	survived= no	→	sex= male
0.604	0.797	0.757	0.894	1.178	0.091	age= adult, sex= male	→	survived= no
0.604	0.768	0.786	0.831	1.175	0.090	sex= male	→	age= adult, survived= no
0.604	0.635	0.950	0.652	1.025	0.015	age= adult	→	sex= male, survived= no
0.604	0.974	0.620	1.534	1.025	0.015	sex= male, survived= no	→	age= adult
0.604	0.924	0.653	1.204	1.175	0.090	age= adult, survived= no	→	sex= male
0.604	0.892	0.677	1.119	1.178	0.091	survived= no	→	age= adult, sex= male
0.402	0.423	0.950	0.423	1.052	0.020	age= adult	→	status= crew
0.402	1.000	0.402	2.364	1.052	0.020	status= crew	→	age= adult



Exercise 9.1: Titanic

3. In your opinion, which association rules are the strongest?
- Consider the following measures:

$$\text{Support} = \frac{\text{count}(A \text{ and } C)}{\text{total}}$$

$$\text{Confidence} = \frac{\text{count}(A \text{ and } C)}{\text{count}(A)}$$

$$\text{Coverage} = \text{support}(A)$$

$$\text{Strength} = \frac{\text{support}(C)}{\text{support}(A)}$$

$$\text{Lift} = \frac{\text{support}(A \rightarrow C)}{\text{support}(A) * \text{support}(C)}$$

$$\text{Leverage} = \text{support}(A \rightarrow C) - \text{support}(A) * \text{support}(C)$$



Exercise 9.1: Titanic

4. Calculate the conditional probability of a male not surviving. *(If a passenger is male, what are the chances of not surviving?)*
- Conditional probability: $P(A|B) = P(A \cap B) / P(B)$
 $P(\text{DNS} \mid \text{Male}) = P(\text{DNS} \cap \text{Male}) / P(\text{Male})$
 $= 0.6197 / 0.7865 = 0.788$
 - Compare to the confidence of Rule: Male \rightarrow Does not survive
0.788



Exercise 9.1: Titanic

5. Look at the rules more in depth.

- Which two rules would you choose?
- How do these findings compare to the exercise using Naïve Bayes Classification?
- Any conclusions?



Exercise 9.1: Titanic

The following options are no longer available in Orange:

- Association Rules Filter (question 2)
- Association Rules Explorer (question 5)



Want more practice?

- Check out this multiple choice quiz from the University of British Columbia:
 - <https://www.ugrad.cs.ubc.ca/~cs100/2016W2/exercises/apriori.html>



UC Leuven
Limburg

MOVING MINDS

Any questions?