Part 2: Exercise

The sinking of the RMS Titanic is one of the most infamous shipwrecks in history. On April 15, 1912, during her maiden voyage, the Titanic sank after colliding with an iceberg, killing 1502 out of 2224 passengers and crew. This sensational tragedy shocked the international community and led to better safety regulations for ships.

One of the reasons that the shipwreck led to such loss of life was that there were not enough lifeboats for the passengers and crew. Although there was some element of luck involved in surviving the sinking, some groups of people were more likely to survive than others, such as women, children, and the upper-class.

In this exercise, we ask you to complete the analysis of what sorts of people were likely to survive. In particular, we ask you to apply the Association Rule mining to predict which passengers survived from the tragedy.

titanic_preprocessed.csv data set can be downloaded from the Blackboard. We have pre-processed the data for you so this is in the format required to use the ARutils.prepare_data() function. The columns are:

- **Age_Adult**
- **Age_Child**
- **Sex_Female**
- **Sex_Male**
- **Crew_Member**
- **First_class**
- **Second_class**
- **Third_class**
- **Survived**

Each variable takes the value 'Yes' or 'No'; the data takes the value 'Yes' when the attribute applies to the individual in the data.

For this exercise, you should apply Association Rule mining and then use a lambda function similar to the one used in step 24 to filter on those rules that have 'Survived' on the RHS of the rule.

source : https://www.kaggle.com/c/titanic