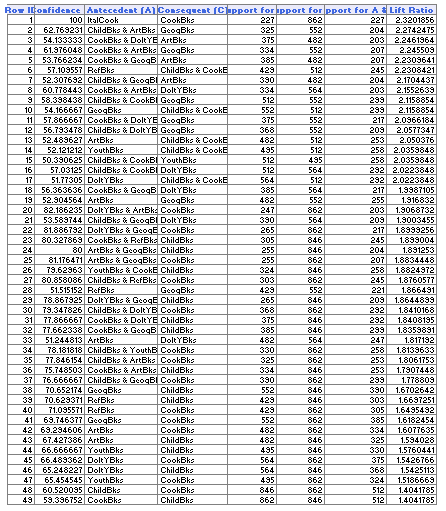
Assignment #2 Item Set/Assoc. Rules by Joshua Troup

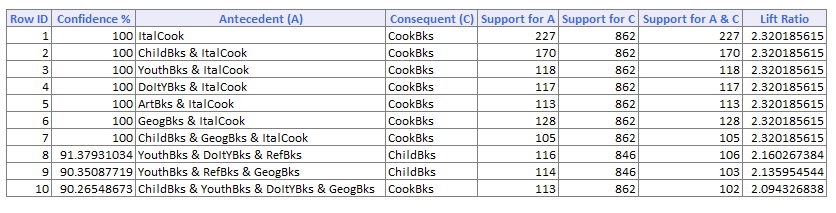
1. Think of the objectives here which is to find groups of customers. Is AR the right approach? If yes, why. If not, what algorithmic approach will you suggest?

**Association rules is the correct approach to take in this situation. The rules are meant to help discover relationships between unrelated data using if/then statements. (Example: customer buys a motorcycle, he/she is 90% likely to buy motorcycle accessories or apparel.)**

1. Find some association rules for this dataset (using XLMiner default settings of 10 % of dataset as “Minimum support” and minimum confidence at 50%.) - 8 points



1. a) How many rules are found? **49**
2. b) Explain the first 2 rules as a data analytics practitioner. **Rule 1 states if ItalCook is purchased, there is a 100% confidence level that CookBks will be purchased with a 2.43 Lift Ratio which is very likely. Rule 2 states if ChildBks & ArtBks are purchased, then there is a 62% confidence level that GeogBks will be purchased with a 2.27 Lift Ratio which is very likely.**
3. c) What are the measures traditionally used to describe how good an association rule is? **Support, Confidence, and Lift**
4. d) For second rule, calculate the confidence and lift from the output values using the formula ( you learnt in data mining course) (2) **Confidence is 62%. Lift is 2.27 Lift is greater than 1 so items in Rule 2 are likely to be bought and is considered to be a good rule.**
5. Change the minimum support to 100 transactions and confidence thresholds to 90% to find additional decision rules. List the top 10 rules as a Table



1. Change the “Minimum support” to 60 transactions and the minimum confidence to 20%. How many rules do you have now? Do you have any insights with this final run?

**1343 rules. Rule 1 states if CookBks & ItalArt are purchased, there is a 91% confidence level ArtBks & ItalArt will be purchased with a 16.18 Lift Ratio. Rule 2 is very comparable in results as Rule 1 with the Antecdent A and Consequent C switched. All 1343 rules have a Lift Ratio greater than one so these rules are considered to be good rules based on a 60 minimum support and 20% confidence level.**