**Association rule** learning is a **rule**-based machine learning method for discovering interesting relations between variables in large databases. It is intended to identify strong **rules** discovered in databases using some measures of interestingness. <https://en.wikipedia.org/wiki/Association_rule_learning>

With a large number of people trooping into stores and supermarkets daily to sear

Association Rules is one of the very important concepts of machine learning being used in market basket analysis. Market Basket Analysis is the study of customer transaction databases to determine dependencies between purchases of different items.

**Association rule** learning is a **rule**-based machine learning method for discovering interesting relations between variables in large databases. It identifies frequent if-then associations called association rules which consists of an antecedent (if) and a consequent (then). For example,

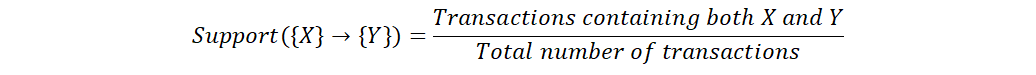
*“If tea and milk, then sugar” (“If tea and milk are purchased, then sugar would also be bought by the customer”)*

– Antecedent: Red and White

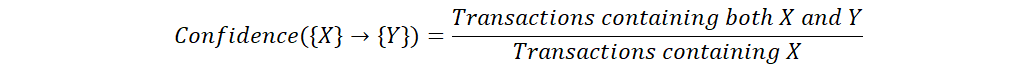
– Consequent: Green.

There are three common ways to measure association:

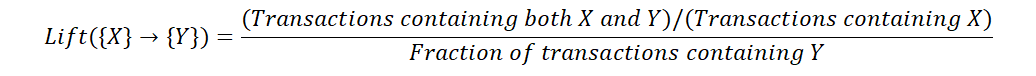
**Support** is an indication of how frequently the items appear in the data. Mathematically, support is the fraction of the total number of transactions in which the itemset occurs.



**Confidence** indicates the number of times the if-then statements are found true. Technically, confidence is the conditional probability of occurrence of consequent given the antecedent.



Lift can be used to compare confidence with expected confidence. This says how likely item Y is purchased when item X is purchased, while controlling for how popular item Y is. Mathematically,



Looking back at the multitude of concepts that have been introduced to me in the Datascence nigeria boot camp, there is a lot to write and share. I choose to start with Association Rules because of two reasons. First, this was one of the concepts which I enjoyed learning the most and second, there are a limited resources available online to get a good grasp.

before basket my dataframe was in the form of single transactions in multiple rows.  
i wanted to convert it in the form where every item represents a column and single transactions in one row, so that if multiple items were bought in single transaction they will be represented by 1.  
so this single statement does multiple things, it group the data by transaction and items and fill the values with the quantitites sum, and if the item was not sold on that transaction it will be filled by NaN so fillna(0) will replace those NaN with a 0.