

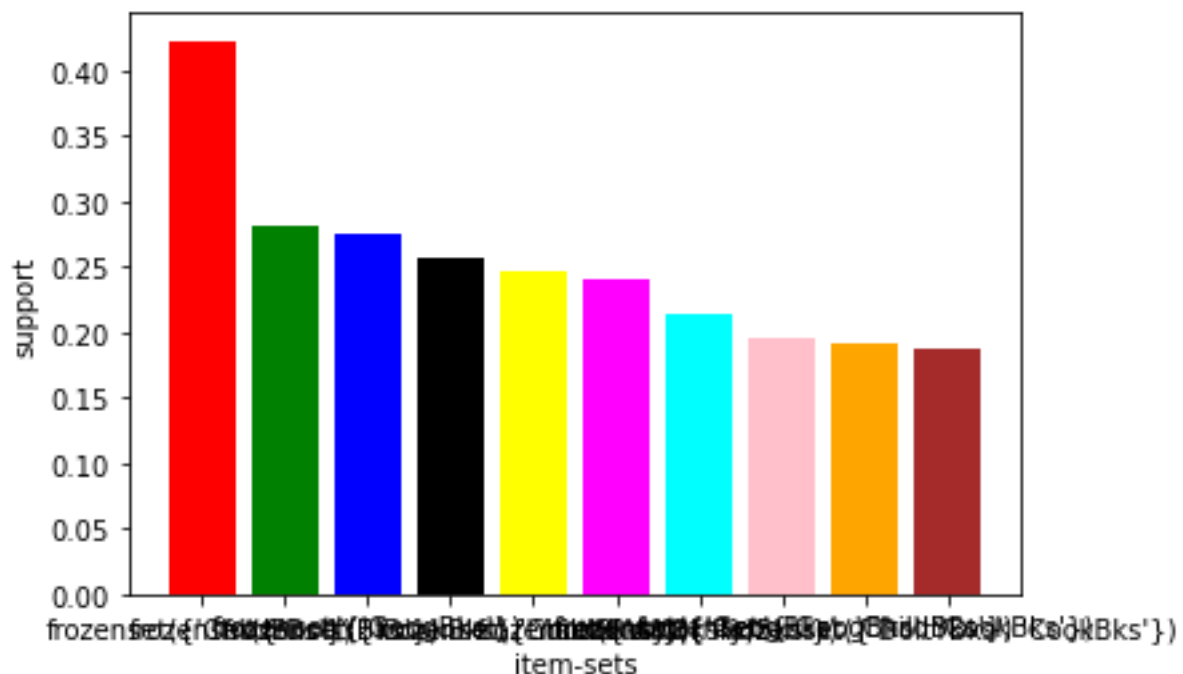
## Association Rule on Books Dataset

There are 2000 transaction in the dataset

First, the rules were generated with a support of 5%, i.e. the item in consideration should be present in at least 5% of the transactions

Top 10 frequent itemset:

| Index | support | itemsets                           |
|-------|---------|------------------------------------|
| 2     | 0.431   | frozenset({'CookBks'})             |
| 0     | 0.423   | frozenset({'ChildBks'})            |
| 3     | 0.282   | frozenset({'DoItYBks'})            |
| 6     | 0.276   | frozenset({'GeogBks'})             |
| 10    | 0.256   | frozenset({'CookBks', 'ChildBks'}) |
| 1     | 0.2475  | frozenset({'YouthBks'})            |
| 5     | 0.241   | frozenset({'ArtBks'})              |
| 4     | 0.2145  | frozenset({'RefBks'})              |
| 14    | 0.195   | frozenset({'GeogBks', 'ChildBks'}) |
| 25    | 0.1925  | frozenset({'GeogBks', 'CookBks'})  |



Frequent itemsets are the ones which occur at least a minimum number of times in the transactions. Technically, these are the itemsets for which support value (fraction of transactions containing the itemset) is above a minimum threshold (min\_support = 0.05).

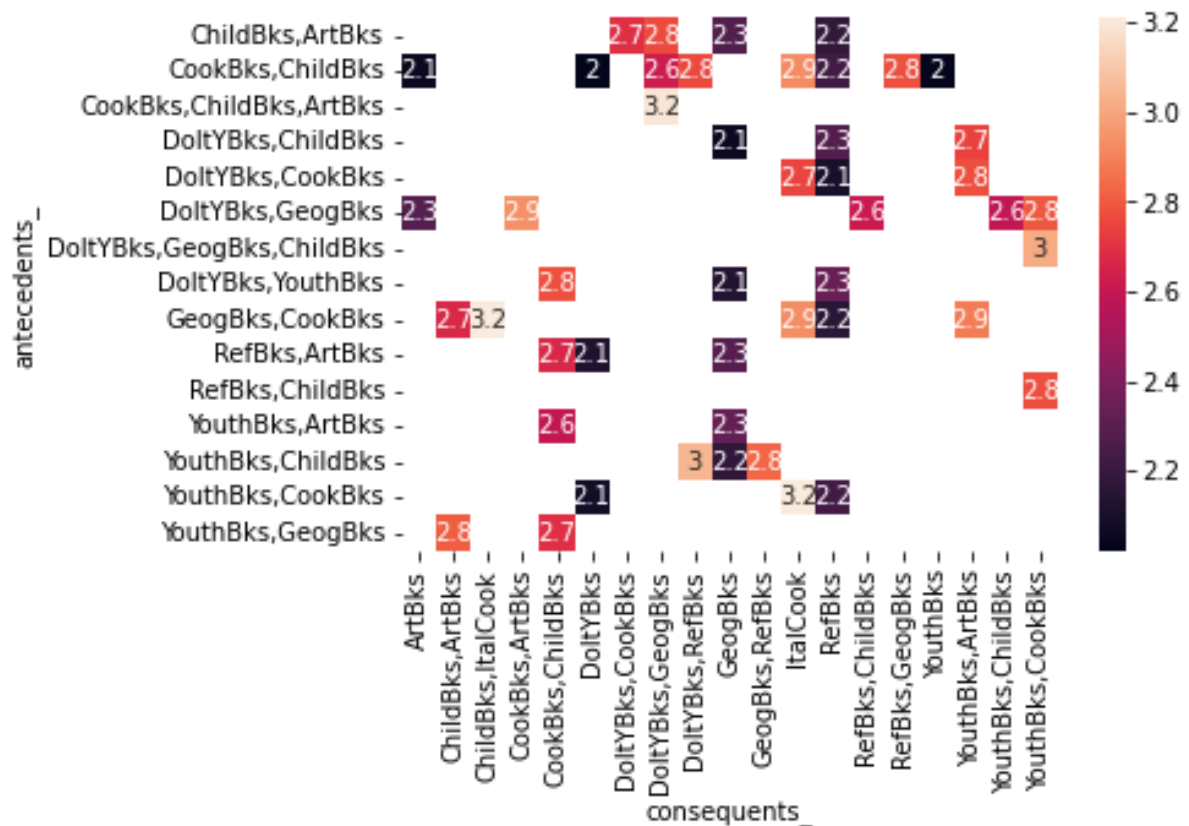
- 43 percent of 100 purchases are "COOKBOOK"
- 42 percent of 100 purchases are "CHILD BOOK"
- 28 percent of 100 purchases are "DO IT YOURSELF BOOK"
- 27 percent of 100 purchases are "GEOGRAPHY BOOK"

### Rules after applying lift ratio

| antecedents                         | consequents                         | edent su | quent su | support | confidence | lift     | leverage   | conviction |
|-------------------------------------|-------------------------------------|----------|----------|---------|------------|----------|------------|------------|
| frozenset({'CookBks', 'YouthBks'})  | frozenset({'ItalCook'})             | 0.162    | 0.1135   | 0.059   | 0.3641975  | 3.208789 | 0.040613   | 1.394301   |
| frozenset({'ItalCook'})             | frozenset({'CookBks', 'ArtBks'})    | 0.1135   | 0.167    | 0.0565  | 0.4977974  | 2.980822 | 0.0375455  | 1.658693   |
| frozenset({'GeogBks', 'CookBks'})   | frozenset({'ItalCook'})             | 0.1925   | 0.1135   | 0.064   | 0.3324675  | 2.929229 | 0.04215125 | 1.328025   |
| frozenset({'CookBks', 'ChildBks'})  | frozenset({'ItalCook'})             | 0.256    | 0.1135   | 0.085   | 0.3320312  | 2.925385 | 0.055944   | 1.327158   |
| frozenset({'DoItYBks', 'CookBks'})  | frozenset({'ItalCook'})             | 0.1875   | 0.1135   | 0.0585  | 0.312      | 2.748899 | 0.03721875 | 1.288517   |
| frozenset({'ItalCook'})             | frozenset({'GeogBks', 'ChildBks'})  | 0.1135   | 0.195    | 0.0525  | 0.4625551  | 2.372077 | 0.0303675  | 1.497828   |
| frozenset({'RefBks'})               | frozenset({'YouthBks', 'ChildBks'}) | 0.2145   | 0.165    | 0.083   | 0.3869464  | 2.34513  | 0.0476075  | 1.362034   |
| frozenset({'DoItYBks', 'YouthBks'}) | frozenset({'RefBks'})               | 0.1155   | 0.2145   | 0.058   | 0.5021645  | 2.341093 | 0.03322525 | 1.57783    |
| frozenset({'YouthBks', 'ArtBks'})   | frozenset({'GeogBks'})              | 0.101    | 0.276    | 0.065   | 0.6435644  | 2.331755 | 0.037124   | 2.031222   |
| frozenset({'CookBks'})              | frozenset({'ItalCook'})             | 0.431    | 0.1135   | 0.1135  | 0.2633411  | 2.320186 | 0.0645815  | 1.203406   |
| frozenset({'GeogBks'})              | frozenset({'ArtBks', 'RefBks'})     | 0.276    | 0.0895   | 0.0565  | 0.2047101  | 2.287264 | 0.031798   | 1.144866   |

From the above snapshot, we can say that any cart consisting 'Cookbook and Youthbook' would likely have 'Italian Cookbook' too with 36% confidence

## Data Visualization for Association Rules



From the above image we can see some possibilities of given antecedents what would the consequents be. For example, {Youth books, Geography books} → {Child books, Artbooks}.

This can help in many ways, one of which is demand planning / product placements in the bookstore.

Next, the rules were generated with a support of 0.8%, i.e. the item in consideration should be present in at least 0.8% of the transactions.

| antecedents                          | consequents                          | edent su | quent su | support | confidence | lift     | leverage   | conviction |
|--------------------------------------|--------------------------------------|----------|----------|---------|------------|----------|------------|------------|
| frozenset({'ItalArt', 'RefBks'})     | frozenset({'ItalAtlas'})             | 0.02     | 0.037    | 0.0165  | 0.825      | 22.2973  | 0.01576    | 5.502857   |
| frozenset({'ItalArt'})               | frozenset({'ItalAtlas', 'ArtBks'})   | 0.0485   | 0.018    | 0.0165  | 0.3402062  | 18.90034 | 0.015627   | 1.488344   |
| frozenset({'ItalCook', 'ArtBks'})    | frozenset({'ItalArt'})               | 0.0565   | 0.0485   | 0.0375  | 0.6637168  | 13.68488 | 0.03475975 | 2.829461   |
| frozenset({'ItalAtlas'})             | frozenset({'ItalCook', 'RefBks'})    | 0.037    | 0.0465   | 0.023   | 0.6216216  | 13.36821 | 0.0212795  | 2.519964   |
| frozenset({'ItalArt'})               | frozenset({'GeogBks', 'ItalAtlas'})  | 0.0485   | 0.0205   | 0.0115  | 0.2371134  | 11.56651 | 0.01050575 | 1.283939   |
| frozenset({'ItalCook', 'ItalAtlas'}) | frozenset({'ItalArt'})               | 0.023    | 0.0485   | 0.0125  | 0.5434783  | 11.20574 | 0.0113845  | 2.084238   |
| frozenset({'ItalCook', 'Florence'})  | frozenset({'ItalArt'})               | 0.0175   | 0.0485   | 0.0095  | 0.5428571  | 11.19293 | 0.00865125 | 2.081406   |
| frozenset({'ItalArt', 'ChildBks'})   | frozenset({'ItalAtlas'})             | 0.036    | 0.037    | 0.0145  | 0.4027778  | 10.88589 | 0.013168   | 1.612465   |
| frozenset({'ItalArt'})               | frozenset({'DoItYBks', 'ItalAtlas'}) | 0.0485   | 0.019    | 0.0095  | 0.1958763  | 10.30928 | 0.0085785  | 1.219962   |
| frozenset({'ItalAtlas', 'YouthBks'}) | frozenset({'ItalArt'})               | 0.0175   | 0.0485   | 0.0085  | 0.4857143  | 10.01473 | 0.00765125 | 1.850139   |
| frozenset({'CookBks', 'ItalAtlas'})  | frozenset({'ItalArt'})               | 0.0285   | 0.0485   | 0.013   | 0.4561404  | 9.404956 | 0.01161775 | 1.749532   |
| frozenset({'ItalArt'})               | frozenset({'ItalAtlas'})             | 0.0485   | 0.037    | 0.0165  | 0.3402062  | 9.194762 | 0.0147055  | 1.459547   |

Compared to the previous set of rules generated, with 0.8% support threshold, naturally more rules could be generated. The interesting observation here is the top contender has replaced {Cookbook, Youth book} → {Italian Cookbook} to {Italian Art, Ref books} → Italian Atlas, and this time with 82.5% confidence.