In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
from mlxtend.frequent_patterns import apriori, association_rules
from mlxtend.preprocessing import TransactionEncoder
```

In [2]:
```
book = pd.read_csv('book.csv')
book
```

Out[2]:

| | ChildBks | YouthBks | CookBks | DoItYBks | RefBks | ArtBks | GeogBks | ItalCook | ItalAtlas | ItalA |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | 0 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | |
| **1** | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| **2** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| **3** | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | |
| **4** | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | |
| **1995** | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 0 | |
| **1996** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| **1997** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |
| **1998** | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | |
| **1999** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | |

2000 rows × 11 columns

In [3]:
```
book.shape
```

Out[3]:
```
(2000, 11)
```

In [4]:
```
book.dtypes
```

Out[4]:
```
ChildBks     int64
YouthBks     int64
CookBks      int64
DoItYBks     int64
RefBks       int64
ArtBks       int64
GeogBks      int64
ItalCook     int64
ItalAtlas    int64
ItalArt      int64
Florence     int64
dtype: object
```

In [5]:
```
book.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2000 entries, 0 to 1999
Data columns (total 11 columns):
 #   Column     Non-Null Count  Dtype
---  ------     --------------  -----
 0   ChildBks   2000 non-null   int64
 1   YouthBks   2000 non-null   int64
 2   CookBks    2000 non-null   int64
 3   DoItYBks   2000 non-null   int64
 4   RefBks     2000 non-null   int64
 5   ArtBks     2000 non-null   int64
 6   GeogBks    2000 non-null   int64
 7   ItalCook   2000 non-null   int64
 8   ItalAtlas  2000 non-null   int64
 9   ItalArt    2000 non-null   int64
 10  Florence   2000 non-null   int64
dtypes: int64(11)
memory usage: 171.9 KB
```

In [7]:
```
book.isna().sum()
```

Out[7]:
```
ChildBks     0
YouthBks     0
CookBks      0
DoItYBks     0
RefBks       0
ArtBks       0
GeogBks      0
ItalCook     0
ItalAtlas    0
ItalArt      0
Florence     0
dtype: int64
```

## 1. Value of support

In [45]:
```
# with 5% support
freq_itemsets = apriori(book, min_support=0.05, use_colnames=True)
freq_itemsets
```

Out[45]:

| | support | itemsets |
|---|---|---|
| **0** | 0.4230 | (ChildBks) |
| **1** | 0.2475 | (YouthBks) |
| **2** | 0.4310 | (CookBks) |
| **3** | 0.2820 | (DoItYBks) |
| **4** | 0.2145 | (RefBks) |
| **...** | ... | ... |
| **95** | 0.0600 | (CookBks, DoItYBks, GeogBks, YouthBks) |
| **96** | 0.0560 | (ArtBks, CookBks, GeogBks, YouthBks) |
| **97** | 0.0650 | (ArtBks, CookBks, DoItYBks, GeogBks) |
| **98** | 0.0510 | (CookBks, ChildBks, DoItYBks, GeogBks, YouthBks) |
| **99** | 0.0535 | (ArtBks, ChildBks, CookBks, DoItYBks, GeogBks) |

100 rows × 2 columns

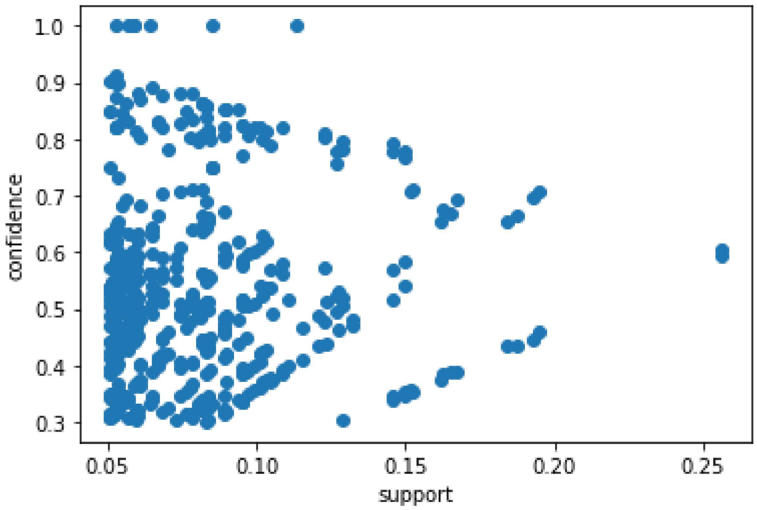
In [46]:
```
# 30% confidence
AR1 = association_rules(freq_itemsets, metric='confidence', min_threshold=0.3)
AR1
```

Out[46]:

| | antecedents | consequents | antecedent support | consequent support | support | confidence | lift | leverage |
|---|---|---|---|---|---|---|---|---|
| **0** | (ChildBks) | (YouthBks) | 0.4230 | 0.2475 | 0.1650 | 0.390071 | 1.576044 | 0.060308 |
| **1** | (YouthBks) | (ChildBks) | 0.2475 | 0.4230 | 0.1650 | 0.666667 | 1.576044 | 0.060308 |
| **2** | (ChildBks) | (CookBks) | 0.4230 | 0.4310 | 0.2560 | 0.605201 | 1.404179 | 0.073687 |
| **3** | (CookBks) | (ChildBks) | 0.4310 | 0.4230 | 0.2560 | 0.593968 | 1.404179 | 0.073687 |
| **4** | (ChildBks) | (DoItYBks) | 0.4230 | 0.2820 | 0.1840 | 0.434988 | 1.542511 | 0.064714 |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... |
| **473** | (ChildBks, ArtBks) | (CookBks, DoItYBks, GeogBks) | 0.1625 | 0.1085 | 0.0535 | 0.329231 | 3.034385 | 0.035869 |
| **474** | (ArtBks, CookBks) | (ChildBks, DoItYBks, GeogBks) | 0.1670 | 0.1045 | 0.0535 | 0.320359 | 3.065639 | 0.036048 |
| **475** | (ArtBks, DoItYBks) | (ChildBks, GeogBks, CookBks) | 0.1235 | 0.1495 | 0.0535 | 0.433198 | 2.897648 | 0.035037 |
| **476** | (ArtBks, GeogBks) | (ChildBks, DoItYBks, CookBks) | 0.1275 | 0.1460 | 0.0535 | 0.419608 | 2.874026 | 0.034885 |
| **477** | (GeogBks, DoItYBks) | (ChildBks, ArtBks, CookBks) | 0.1325 | 0.1265 | 0.0535 | 0.403774 | 3.191886 | 0.036739 |

478 rows × 9 columns

In [47]:
```
plt.scatter(AR1['support'], AR1['confidence'])
plt.xlabel('support')
plt.ylabel('confidence')
plt.show()
```

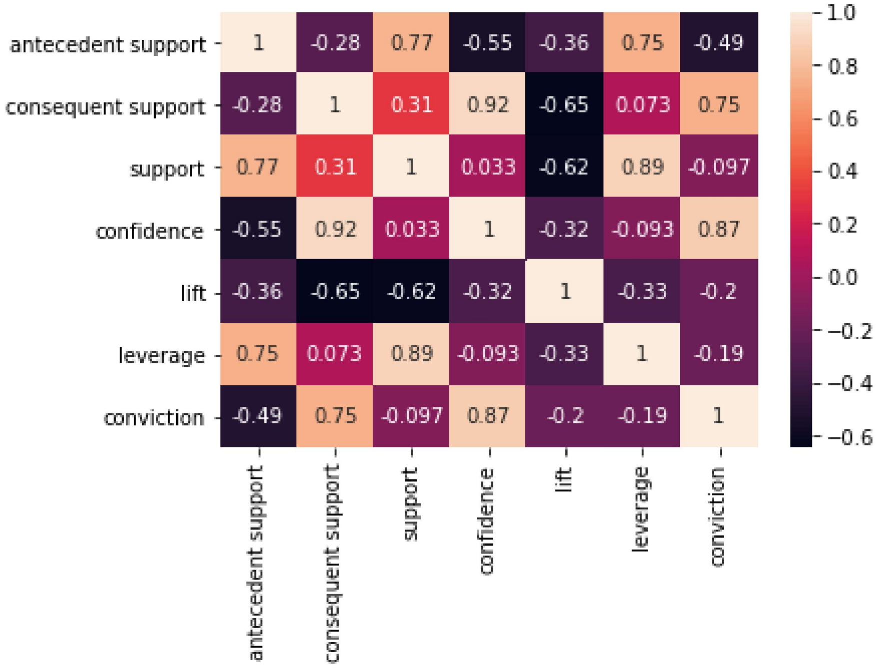



In [48]:
```
corr_AR1 = AR1.corr()
corr_AR1
```

Out[48]:

| | antecedent support | consequent support | support | confidence | lift | leverage | conviction |
|---|---|---|---|---|---|---|---|
| **antecedent support** | 1.000000 | -0.275429 | 0.768106 | -0.546549 | -0.355572 | 0.749441 | -0.490575 |
| **consequent support** | -0.275429 | 1.000000 | 0.309742 | 0.917072 | -0.645453 | 0.073186 | 0.753839 |
| **support** | 0.768106 | 0.309742 | 1.000000 | 0.032897 | -0.624631 | 0.892221 | -0.096799 |
| **confidence** | -0.546549 | 0.917072 | 0.032897 | 1.000000 | -0.321874 | -0.092935 | 0.866620 |
| **lift** | -0.355572 | -0.645453 | -0.624631 | -0.321874 | 1.000000 | -0.325328 | -0.201533 |
| **leverage** | 0.749441 | 0.073186 | 0.892221 | -0.092935 | -0.325328 | 1.000000 | -0.191227 |
| **conviction** | -0.490575 | 0.753839 | -0.096799 | 0.866620 | -0.201533 | -0.191227 | 1.000000 |

In [49]:
```
sns.heatmap(data=corr_AR1, annot=True)
plt.show()
```




In [ ]:

## 2. Value of support

In [50]:
```
# with 10% support
freq_itemsets_2 = apriori(book, min_support=0.10, use_colnames=True)
freq_itemsets_2
```

Out[50]:

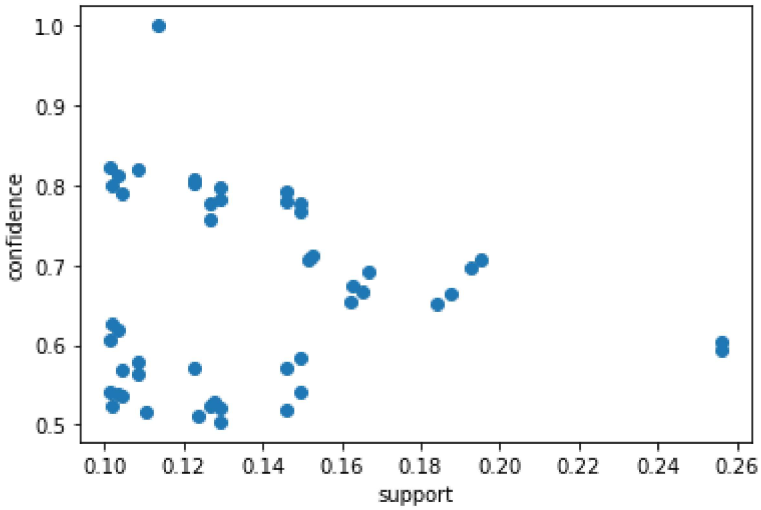
| | support | itemsets |
|---|---|---|
| 0 | 0.4230 | (ChildBks) |
| 1 | 0.2475 | (YouthBks) |
| 2 | 0.4310 | (CookBks) |
| 3 | 0.2820 | (DoItYBks) |
| 4 | 0.2145 | (RefBks) |
| 5 | 0.2410 | (ArtBks) |
| 6 | 0.2760 | (GeogBks) |
| 7 | 0.1135 | (ItalCook) |
| 8 | 0.1085 | (Florence) |
| 9 | 0.1650 | (ChildBks, YouthBks) |
| 10 | 0.2560 | (ChildBks, CookBks) |
| 11 | 0.1840 | (ChildBks, DoItYBks) |
| 12 | 0.1515 | (ChildBks, RefBks) |
| 13 | 0.1625 | (ArtBks, ChildBks) |
| 14 | 0.1950 | (ChildBks, GeogBks) |
| 15 | 0.1620 | (CookBks, YouthBks) |
| 16 | 0.1155 | (DoItYBks, YouthBks) |
| 17 | 0.1010 | (ArtBks, YouthBks) |
| 18 | 0.1205 | (GeogBks, YouthBks) |
| 19 | 0.1875 | (CookBks, DoItYBks) |
| 20 | 0.1525 | (CookBks, RefBks) |
| 21 | 0.1670 | (ArtBks, CookBks) |
| 22 | 0.1925 | (CookBks, GeogBks) |
| 23 | 0.1135 | (ItalCook, CookBks) |
| 24 | 0.1055 | (RefBks, DoItYBks) |
| 25 | 0.1235 | (ArtBks, DoItYBks) |
| 26 | 0.1325 | (GeogBks, DoItYBks) |
| 27 | 0.1105 | (GeogBks, RefBks) |
| 28 | 0.1275 | (ArtBks, GeogBks) |
| 29 | 0.1290 | (CookBks, ChildBks, YouthBks) |
| 30 | 0.1460 | (ChildBks, DoItYBks, CookBks) |
| 31 | 0.1225 | (ChildBks, RefBks, CookBks) |

| | support | itemsets |
|---|---|---|
| **32** | 0.1265 | (ArtBks, ChildBks, CookBks) |
| **33** | 0.1495 | (ChildBks, GeogBks, CookBks) |
| **34** | 0.1045 | (ChildBks, DoItYBks, GeogBks) |
| **35** | 0.1020 | (ArtBks, ChildBks, GeogBks) |
| **36** | 0.1015 | (ArtBks, CookBks, DoItYBks) |
| **37** | 0.1085 | (CookBks, DoItYBks, GeogBks) |
| **38** | 0.1035 | (ArtBks, CookBks, GeogBks) |

In [51]:
```python
# 50% confidence
AR2 = association_rules(freq_itemsets_2, metric='confidence', min_threshold=0.5)
AR2
```

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **8** | (DoItYBks) | (CookBks) | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] | [illegible] |
| **9** | (RefBks) | (CookBks) | 0.2145 | 0.4310 | 0.1525 | 0.710956 | 1.649549 | 0.060050 |
| **10** | (ArtBks) | (CookBks) | 0.2410 | 0.4310 | 0.1670 | 0.692946 | 1.607763 | 0.063129 |
| **11** | (GeogBks) | (CookBks) | 0.2760 | 0.4310 | 0.1925 | 0.697464 | 1.618245 | 0.073544 |
| **12** | (ItalCook) | (CookBks) | 0.1135 | 0.4310 | 0.1135 | 1.000000 | 2.320186 | 0.064582 |
| **13** | (ArtBks) | (DoItYBks) | 0.2410 | 0.2820 | 0.1235 | 0.512448 | 1.817192 | 0.055538 |
| **14** | (RefBks) | (GeogBks) | 0.2145 | 0.2760 | 0.1105 | 0.515152 | 1.866491 | 0.051298 |
| **15** | (ArtBks) | (GeogBks) | 0.2410 | 0.2760 | 0.1275 | 0.529046 | 1.916832 | 0.060984 |
| **16** | (ChildBks, CookBks) | (YouthBks) | 0.2560 | 0.2475 | 0.1290 | 0.503906 | 2.035985 | 0.065640 |
| **17** | (CookBks, YouthBks) | (ChildBks) | 0.1620 | 0.4230 | 0.1290 | 0.796296 | 1.882497 | 0.060474 |
| **18** | (ChildBks, YouthBks) | (CookBks) | 0.1650 | 0.4310 | 0.1290 | 0.781818 | 1.813963 | 0.057885 |

In [52]:
```
plt.scatter(AR2['support'], AR2['confidence'])
plt.xlabel('support')
plt.ylabel('confidence')
plt.show()
```

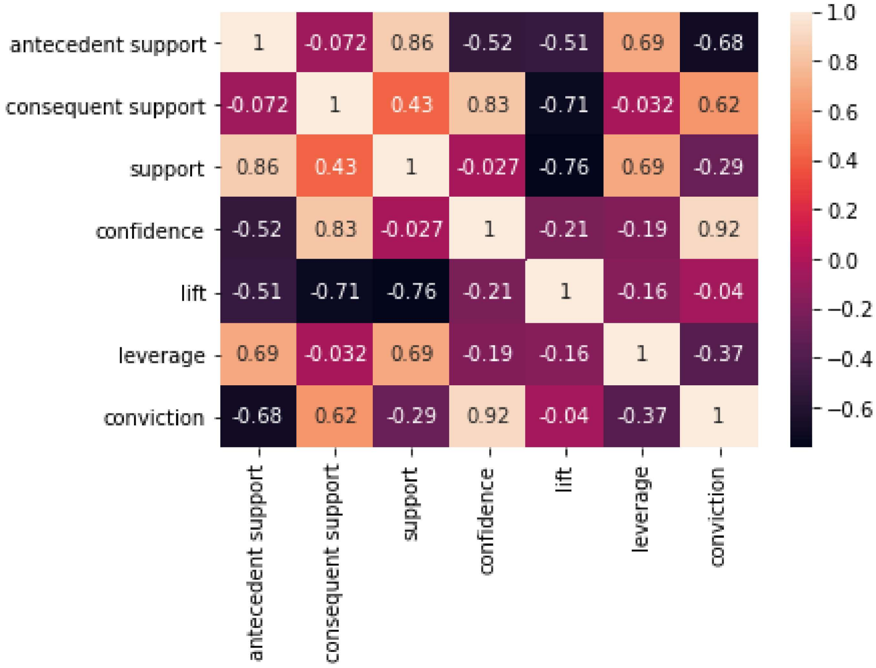



In [53]:
```
corr_AR2 = AR2.corr()
corr_AR2
```

Out[53]:

| | antecedent support | consequent support | support | confidence | lift | leverage | conviction |
|---|---|---|---|---|---|---|---|
| **antecedent support** | 1.000000 | -0.071635 | 0.858437 | -0.519173 | -0.512713 | 0.690605 | -0.683617 |
| **consequent support** | -0.071635 | 1.000000 | 0.426340 | 0.831627 | -0.711482 | -0.031692 | 0.621042 |
| **support** | 0.858437 | 0.426340 | 1.000000 | -0.027358 | -0.762058 | 0.692109 | -0.291754 |
| **confidence** | -0.519173 | 0.831627 | -0.027358 | 1.000000 | -0.208019 | -0.185236 | 0.917344 |
| **lift** | -0.512713 | -0.711482 | -0.762058 | -0.208019 | 1.000000 | -0.157036 | -0.040188 |
| **leverage** | 0.690605 | -0.031692 | 0.692109 | -0.185236 | -0.157036 | 1.000000 | -0.365942 |
| **conviction** | -0.683617 | 0.621042 | -0.291754 | 0.917344 | -0.040188 | -0.365942 | 1.000000 |

In [54]:
```
sns.heatmap(data=corr_AR2, annot=True)
plt.show()
```




In [ ]:

## 3. Value of support

In [55]:
```
# with 15% support
freq_itemsets_3 = apriori(book, min_support=0.15, use_colnames=True)
freq_itemsets_3
```

Out[55]:

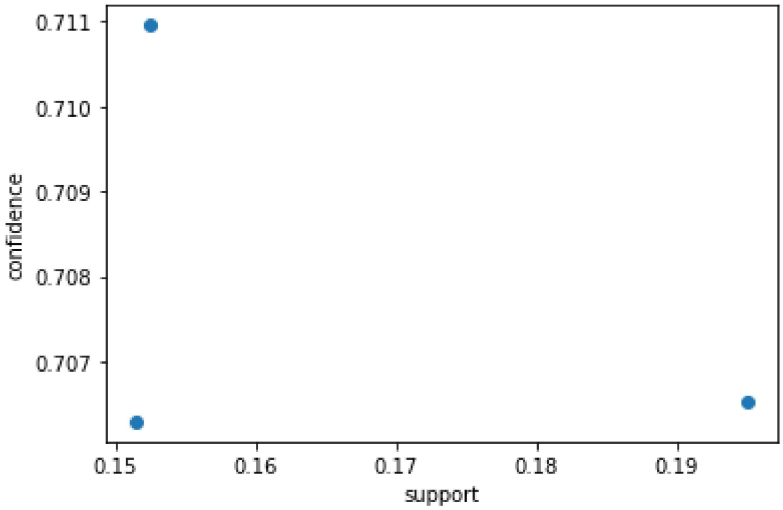
| | support | itemsets |
|---|---|---|
| 0 | 0.4230 | (ChildBks) |
| 1 | 0.2475 | (YouthBks) |
| 2 | 0.4310 | (CookBks) |
| 3 | 0.2820 | (DoItYBks) |
| 4 | 0.2145 | (RefBks) |
| 5 | 0.2410 | (ArtBks) |
| 6 | 0.2760 | (GeogBks) |
| 7 | 0.1650 | (ChildBks, YouthBks) |
| 8 | 0.2560 | (ChildBks, CookBks) |
| 9 | 0.1840 | (ChildBks, DoItYBks) |
| 10 | 0.1515 | (ChildBks, RefBks) |
| 11 | 0.1625 | (ArtBks, ChildBks) |
| 12 | 0.1950 | (ChildBks, GeogBks) |
| 13 | 0.1620 | (CookBks, YouthBks) |
| 14 | 0.1875 | (CookBks, DoItYBks) |
| 15 | 0.1525 | (CookBks, RefBks) |
| 16 | 0.1670 | (ArtBks, CookBks) |
| 17 | 0.1925 | (CookBks, GeogBks) |

In [56]:
```
# 70% confidence
AR3 = association_rules(freq_itemsets_3, metric='confidence', min_threshold=0.7)
AR3
```

Out[56]:

| | antecedents | consequents | antecedent support | consequent support | support | confidence | lift | leverage | c |
|---|---|---|---|---|---|---|---|---|---|
| 0 | (RefBks) | (ChildBks) | 0.2145 | 0.423 | 0.1515 | 0.706294 | 1.669725 | 0.060767 | |
| 1 | (GeogBks) | (ChildBks) | 0.2760 | 0.423 | 0.1950 | 0.706522 | 1.670264 | 0.078252 | |
| 2 | (RefBks) | (CookBks) | 0.2145 | 0.431 | 0.1525 | 0.710956 | 1.649549 | 0.060050 | |

In [57]:
```
plt.scatter(AR3['support'], AR3['confidence'])
plt.xlabel('support')
plt.ylabel('confidence')
plt.show()
```

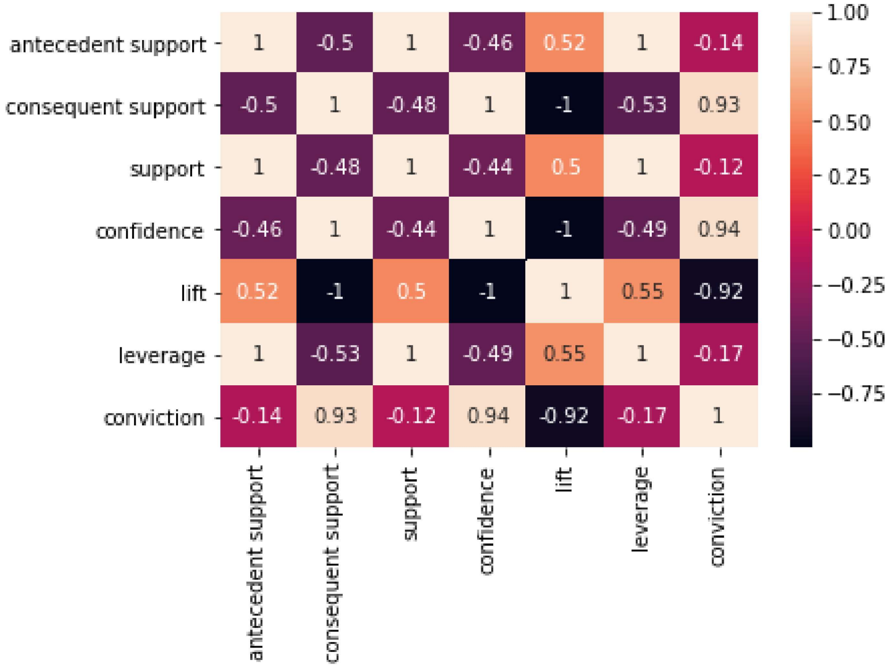



In [58]:
```
corr_AR3 = AR3.corr()
corr_AR3
```

Out[58]:

| | antecedent support | consequent support | support | confidence | lift | leverage | conviction |
|---|---|---|---|---|---|---|---|
| **antecedent support** | 1.000000 | -0.500000 | 0.999797 | -0.461960 | 0.519640 | 0.999397 | -0.136385 |
| **consequent support** | -0.500000 | 1.000000 | -0.482460 | 0.999059 | -0.999739 | -0.529775 | 0.926126 |
| **support** | 0.999797 | -0.482460 | 1.000000 | -0.444008 | 0.502331 | 0.998495 | -0.116410 |
| **confidence** | -0.461960 | 0.999059 | -0.444008 | 1.000000 | -0.997808 | -0.492483 | 0.941618 |
| **lift** | 0.519640 | -0.999739 | 0.502331 | -0.997808 | 1.000000 | 0.548999 | -0.917273 |
| **leverage** | 0.999397 | -0.529775 | 0.998495 | -0.492483 | 0.548999 | 1.000000 | -0.170708 |
| **conviction** | -0.136385 | 0.926126 | -0.116410 | 0.941618 | -0.917273 | -0.170708 | 1.000000 |

In [59]:
```
sns.heatmap(data=corr_AR3, annot=True)
plt.show()
```




In [ ]: