1. Basic Rule Creation

The 'databse' below has four transactions. What ssociation rules can be found in this set. if the minimum support (ie coverage) is 60% and minimum confidence (ie accuracy) is 80%?

Tans_id	ItemList
T1	{K, A, D, B}
T2	{D,A,C,E,B}
T3	{C, A, B, E }
T3	{B, A, D}

Show each step of your calculation in details (ie item sets containing one item, two items like this), as each step will be graded individually.

$$Support_A = \frac{Num\ A}{Num\ Transactions}$$

$$Support_A = \frac{Num\ of\ times\ A\ occured}{Total\ Num\ Transactions}$$

Item	Frequency	Support
A	4	4/4 = 100%
B	4	4/4 = 100%
C	2	2/4 = 50%
D	3	3/4 = 75%
E	2	2/4 = 50%
K	1	1/4 = 25%

Requireing minimum Support of 60% C, E, K Are eliminated

Possble Pairs, AB, BD, AD

Item Pairs	Frequency	Support
A, B	4	4/4 = 100%
B, D	3	3/4 = 75%
A, D	3	3/4 = 75%

No Pairs are eliminated

Possible Rules:

$$(A \implies B), (B \implies A), (B \implies D), (D \implies B), (A \implies D)\ and (D \implies A)$$

$$Confidence(A \implies B) = \frac{Support(A \cup B)}{Support(A)} = \frac{100}{100} = 1$$

$$Confifence(B \implies A) = \frac{Support(A \cup B)}{Support(B)} = \frac{100}{100} = 1$$

$$Confidence(B \implies D) = \frac{Support(B \cup D)}{Support(B)} = \frac{3 \times 4}{4 \times 4} = 0.75$$

$$Confidence(D \implies B) = \frac{Support(B \cup D)}{Support(D)} = \frac{3 \times 4}{4 \times 3} = 1$$

$$Condifidence(A \implies D) = \frac{Support(A \cup D)}{Support(A)} = \frac{3 \times 4}{4 \times 4} = 0.75$$

$$Confidence(D \implies A) = \frac{Support(A \cup D)}{Support(D)} = \frac{3 \times 4}{4 \times 3} = 1$$

Due to minimum confidence of 80%

$$(B \implies D)\ and\ (A \implies D)$$

Are eliminated

Association Rules:

$$(A \implies B), (B \implies A), (D \implies B)\ and\ (D \implies A)$$

2. Using the XML Decument bellow (library with books), define the following queries in XQuery:

a. Give the titles of all books sorted by price.

```
for $book in doc("library.xml")/bib/book
order by $book/price
return $book/title
```

b. How many books were written by Abiteboul?

```
count(doc("library.xml")/bib/book[author="Abiteboul"])
```

c. Give for each author, the number of books they have written.

```
for $author in doc("library.xml")/bib/book/author
return count(doc("library.xml")/bib/book[author=$author])
```

Privided XML Document (library.xml)

```
<?xml version="1.0"?>
<bib>
   <book year="1994">
       <title>TCP/IP</title>
       <author>Stevens</author>
       <publisher>Addison-Wesley</publisher>
       <price>65.95</price>
   </book>
   <book year="1994">
       <title>Principles of Databses</title>
       <author> Abiteboul</author>
       <publisher>Addison-Wesley</publisher>
       <price>35.89</price>
   </book>
   <book year="1994">
       <title>Advanced Programming in the Unix enviroment</title>
       <author>Stevens</author>
       <publisher>Addison-Wesley</publisher>
       <price>65.95</price>
   </book>
   <book year="2000">
       <title> Data on the Web </title>
       <author>Abiteboul</author>
       <author>Bunenman</author>
       <author>Suciu</author>
       <publisher>Morgan Kaufmann Publishers</publisher>
       <price>39.95</price>
   </book>
   <book year="1992">
       <title>The Economics of Technology and Content for Digital
TV</title>
       <editor>
           <affiliation>CITI</affiliation>
       </editor>
       <publisher>Kluwer Academix Publishers</publisher>
       <price>129.95</price>
   </book>
 </bib>
```

========== Originality Declaration ==========

Name: Hector Ramirez

Panther-ID: 5708475

Course: COP-4751

Assignment#: 3

Due: 11/12/2021

I hereby certify that this work is my own and none of it is the work of any other person.

Signature: Hector Ramirez

==========