

Lista 6

Breno Pires Santos

Matricula: 808238

Questão 01

Considere que em um determinado supermercado foram efetuadas as seguintes transações:

N°	Leite	Café	Cerveja	Pão	Manteiga	Arroz	Feijão
1	Não	Sim	Não	Sim	Sim	Não	Não
2	Sim	Não	Sim	Sim	Sim	Não	Não
3	Não	Sim	Não	Sim	Sim	Não	Não
4	Sim	Sim	Não	Sim	Sim	Não	Não
5	Não	Não	Sim	Não	Não	Não	Não
6	Não	Não	Não	Não	Sim	Não	Não
7	Não	Não	Não	Sim	Não	Não	Não
8	Não	Não	Não	Não	Não	Não	Sim
9	Não	Não	Não	Não	Não	Sim	Sim
10	Não	Não	Não	Não	Não	Sim	Não

Utilizando-se o algoritmo Apriori, um suporte mínimo aceitável de **0.3** e confiança de **0.8**, o número de **ItemSets 1, 2, 3 e de regras** a partir desta base de dados são:

Resposta

ItemSets 1: 5

ItemSets 2: 5

ItemSets 3: 1

Regras: 3

Questão 02

Considerando-se o código que está em **Módulos/Apriori.ipynb**, rode o código com a base acima e confira os resultados.

```
ItemSets 1: Arroz
ItemSets 2: Manteiga
ItemSets 3: Pão
ItemSets 4: Manteiga, Arroz
ItemSets 5: Pão, Arroz
ItemSets 6: Manteiga, Café
ItemSets 7: Pão, Café
ItemSets 8: Feijão, Manteiga
ItemSets 9: Feijão, Pão
ItemSets 10: Pão, Leite
ItemSets 11: Manteiga, Pão
ItemSets 12: Manteiga, Pão, Arroz
ItemSets 13: Pão, Manteiga, Café
```

Regras:

```
Arroz => Manteiga (confiança: 0.67, lift: 0.95)
Arroz => Pão (confiança: 0.83, lift: 1.04)
Pão => Arroz (confiança: 0.62, lift: 1.04)
Café => Manteiga (confiança: 1.00, lift: 1.43)
Café => Pão (confiança: 0.75, lift: 0.94)
Feijão => Manteiga (confiança: 0.75, lift: 1.07)
Feijão => Pão (confiança: 0.75, lift: 0.94)
Leite => Pão (confiança: 1.00, lift: 1.25)
Manteiga => Pão (confiança: 0.86, lift: 1.07)
Pão => Manteiga (confiança: 0.75, lift: 1.07)
Arroz => Manteiga, Pão (confiança: 0.67, lift: 1.11)
Manteiga, Arroz => Pão (confiança: 1.00, lift: 1.25)
Pão, Arroz => Manteiga (confiança: 0.80, lift: 1.14)
Manteiga, Pão => Arroz (confiança: 0.67, lift: 1.11)
Café => Manteiga, Pão (confiança: 0.75, lift: 1.25)
Manteiga, Café => Pão (confiança: 0.75, lift: 0.94)
Pão, Café => Manteiga (confiança: 1.00, lift: 1.43)
PS C:\Users\Breno\Documents\IA\Lista6> □
```

Questão 03

Considerando-se o código que está em `Módulos/Apriori.ipynb`, altere-o para que ele imprima os temsets gerados, com os respectivos suportes

```
ItemSet 1: Arroz | Suporte: 0.60
ItemSet 2: Manteiga | Suporte: 0.70
ItemSet 3: Pão | Suporte: 0.80
ItemSet 4: Arroz, Manteiga | Suporte: 0.40
ItemSet 5: Arroz, Pão | Suporte: 0.50
ItemSet 6: Café, Manteiga | Suporte: 0.40
ItemSet 7: Café, Pão | Suporte: 0.30
ItemSet 8: Feijão, Manteiga | Suporte: 0.30
ItemSet 9: Feijão, Pão | Suporte: 0.30
ItemSet 10: Leite, Pão | Suporte: 0.30
ItemSet 11: Pão, Manteiga | Suporte: 0.60
ItemSet 12: Pão, Arroz, Manteiga | Suporte: 0.40
ItemSet 13: Pão, Café, Manteiga | Suporte: 0.30
```

Questão 04

Considerando-se o código que está em **Módulos/Apriori.ipynb**, altere-o para que ele gere regras de associação quando não há presença do produto. Ou seja, gostaria de ver regras da seguinte forma:

Quem não leva álcool leva detergente;

```
quem leva arroz => quem não leva café (confiança: 0.67, lift: 1.11)
quem não leva café => quem leva arroz (confiança: 0.67, lift: 1.11)
quem leva arroz => quem não leva cerveja (confiança: 0.83, lift: 1.39)
quem não leva cerveja => quem leva arroz (confiança: 0.83, lift: 1.39)
quem leva arroz => quem não leva feijão (confiança: 0.67, lift: 1.11)
quem não leva feijão => quem leva arroz (confiança: 0.67, lift: 1.11)
quem leva arroz => quem não leva leite (confiança: 1.00, lift: 1.43)
quem não leva leite => quem leva arroz (confiança: 0.86, lift: 1.43)
quem leva arroz => quem leva manteiga (confiança: 0.67, lift: 0.95)
quem leva arroz => quem leva pão (confiança: 0.83, lift: 1.04)
quem leva pão => quem leva arroz (confiança: 0.62, lift: 1.04)
quem não leva arroz => quem leva cerveja (confiança: 0.75, lift: 1.87)
quem leva cerveja => quem não leva arroz (confiança: 0.75, lift: 1.87)
quem não leva arroz => quem leva leite (confiança: 0.75, lift: 2.50)
quem leva leite => quem não leva arroz (confiança: 1.00, lift: 2.50)
quem não leva arroz => quem leva manteiga (confiança: 0.75, lift: 1.07)
quem não leva arroz => quem leva pão (confiança: 0.75, lift: 0.94)
quem leva café => quem não leva cerveja (confiança: 0.75, lift: 1.25)
quem leva café => quem não leva feijão (confiança: 0.75, lift: 1.25)
quem leva café => quem não leva leite (confiança: 0.75, lift: 1.07)
quem leva café => quem leva manteiga (confiança: 1.00, lift: 1.43)
quem leva café => quem leva pão (confiança: 0.75, lift: 0.94)
quem leva cerveja => quem não leva café (confiança: 0.75, lift: 1.25)
```



```

quem leva pão => quem não leva café (confiança: 0.62, lift: 1.04)
quem não leva cerveja => quem não leva feijão (confiança: 0.67, lift: 1.11)
quem não leva feijão => quem não leva cerveja (confiança: 0.67, lift: 1.11)
quem não leva cerveja => quem não leva leite (confiança: 0.83, lift: 1.19)
quem não leva leite => quem não leva cerveja (confiança: 0.71, lift: 1.19)
quem não leva cerveja => quem leva manteiga (confiança: 0.83, lift: 1.19)
quem leva manteiga => quem não leva cerveja (confiança: 0.71, lift: 1.19)
quem leva pão => quem não leva cerveja (confiança: 1.00, lift: 1.25)
quem leva pão => quem não leva leite (confiança: 0.75, lift: 1.25)
quem leva feijão => quem não leva leite (confiança: 0.75, lift: 1.07)
quem leva feijão => quem leva manteiga (confiança: 0.75, lift: 1.07)
quem leva feijão => quem leva pão (confiança: 0.75, lift: 0.94)
quem não leva feijão => quem não leva leite (confiança: 0.67, lift: 0.95)
quem não leva feijão => quem leva manteiga (confiança: 0.67, lift: 0.95)
quem não leva feijão => quem leva pão (confiança: 0.83, lift: 1.04)
quem leva pão => quem não leva feijão (confiança: 0.62, lift: 1.04)
quem leva leite => quem leva pão (confiança: 1.00, lift: 1.25)
quem não leva leite => quem leva manteiga (confiança: 0.71, lift: 1.02)
quem leva manteiga => quem não leva leite (confiança: 0.71, lift: 1.02)
quem não leva leite => quem leva pão (confiança: 0.71, lift: 0.89)
quem leva pão => quem não leva leite (confiança: 0.62, lift: 0.89)
quem leva manteiga => quem leva pão (confiança: 0.86, lift: 1.07)
quem leva pão => quem leva manteiga (confiança: 0.75, lift: 1.07)
quem não leva café e quem leva arroz => quem não leva cerveja (confiança: 0.75, lift: 1.25)

```

```

quem não leva leite e quem leva arroz => quem não leva feijão (confiança: 0.67, lift: 1.11)
quem não leva leite e quem não leva feijão => quem leva arroz (confiança: 1.00, lift: 1.67)
quem não leva feijão e quem leva arroz => quem leva pão (confiança: 0.75, lift: 0.94)
quem leva pão e quem leva arroz => quem não leva feijão (confiança: 0.60, lift: 1.00)
quem não leva feijão e quem leva pão => quem leva arroz (confiança: 0.60, lift: 1.00)
quem leva arroz => quem leva manteiga e quem não leva leite (confiança: 0.67, lift: 1.33)
quem não leva leite e quem leva arroz => quem leva manteiga (confiança: 0.67, lift: 0.95)
quem leva manteiga e quem leva arroz => quem não leva leite (confiança: 1.00, lift: 1.43)
quem leva manteiga e quem não leva leite => quem leva arroz (confiança: 0.80, lift: 1.33)
quem leva arroz => quem não leva leite e quem leva pão (confiança: 0.83, lift: 1.67)
quem não leva leite => quem leva arroz e quem leva pão (confiança: 0.71, lift: 1.43)
quem leva pão => quem não leva leite e quem leva arroz (confiança: 0.62, lift: 1.04)
quem não leva leite e quem leva arroz => quem leva pão (confiança: 0.83, lift: 1.04)
quem leva pão e quem leva arroz => quem não leva leite (confiança: 1.00, lift: 1.43)
quem não leva leite e quem leva pão => quem leva arroz (confiança: 1.00, lift: 1.67)
quem leva arroz => quem leva manteiga e quem leva pão (confiança: 0.67, lift: 1.11)
quem leva manteiga e quem leva arroz => quem leva pão (confiança: 1.00, lift: 1.25)
quem leva pão e quem leva arroz => quem leva manteiga (confiança: 0.80, lift: 1.14)
quem leva manteiga e quem leva pão => quem leva arroz (confiança: 0.67, lift: 1.11)
quem não leva arroz => quem leva leite e quem leva pão (confiança: 0.75, lift: 2.50)
quem leva leite => quem não leva arroz e quem leva pão (confiança: 1.00, lift: 3.33)
quem leva leite e quem não leva arroz => quem leva pão (confiança: 1.00, lift: 1.25)
quem leva pão e quem não leva arroz => quem leva leite (confiança: 1.00, lift: 3.33)
quem leva leite e quem leva pão => quem não leva arroz (confiança: 1.00, lift: 2.50)

```

Questão 05

Investigue o funcionamento da biblioteca **mlxtend** para geração de regras de associação.

<https://github.com/rasbt/mlxtendLinks to an external site.>

https://github.com/rasbt/mlxtend/blob/master/mlxtend/frequent_patterns/apriori.py

A biblioteca **mlxtend** é usada em Python para aplicar o algoritmo Apriori e gerar regras de associação. Ela funciona a partir de um DataFrame onde cada linha representa uma transação, e cada coluna indica se o item está presente (True ou 1) ou não (False ou 0).

O processo tem duas etapas principais:

1. **apriori()**: encontra os conjuntos de itens (itemsets) mais frequentes nas transações, de acordo com um suporte mínimo.

2. **association_rules()**: gera as regras com base nesses itemsets, usando métricas como confiança e lift.

As funções retornam DataFrames com as combinações encontradas e suas respectivas estatísticas. Isso facilita bastante a análise, já que dá pra ver quais itens costumam aparecer juntos.

Por padrão, o mlxtend só trabalha com a **presença** dos itens. Para gerar regras do tipo “quem **não** leva pão tende a levar leite”, é necessário adaptar os dados, criando colunas como pão_sim e pão_não.

Questão 06

Faça uma resenha do artigo “**A comprehensive review of visualization methods for association rule**” que está no CANVAS.

O artigo apresenta uma revisão detalhada sobre métodos de visualização de regras de associação, que são usadas para descobrir relações entre itens em bases de dados transacionais, como no famoso exemplo do carrinho de supermercado.

A principal ideia do texto é mostrar que, embora existam muitos algoritmos para descobrir regras de associação (como o Apriori), os resultados costumam ser difíceis de interpretar — especialmente quando o número de regras é grande. Por isso, visualizar essas regras de forma clara é fundamental para que usuários possam entendê-las e tomar decisões baseadas nos dados.

Os autores organizam os métodos de visualização em duas categorias principais: os tradicionais, como gráficos de dispersão (scatter plots), gráficos de redes e matrizes; e os modernos ou inovadores, como mapas de metrô, diagramas de Sankey e visualizações moleculares. Cada método é avaliado com base em aspectos como interatividade, tipo de dado suportado (binário, categórico, numérico) e facilidade de uso.

Além disso, o artigo discute os desafios e limitações atuais da área, como o excesso de regras geradas, a dificuldade de representar múltiplas métricas ao mesmo tempo (como suporte, confiança e lift), e a necessidade de técnicas mais interativas. Também são apresentadas sugestões de pesquisa futura, como o uso de inteligência artificial explicável (XAI) e o desenvolvimento de ferramentas que combinem mineração e visualização em um só processo.