## 实验4：关联规则挖掘算法实现及应用

【实验目标】：

实现Apriori算法产生频繁项集和强规则，并在Groceries数据集上进行应用。

【实验任务】：

1. 阐述Apriori算法产生频繁项集和强规则的原理及过程。

2. 编程实现Apriori算法（包括频繁项集产生和强规则生成）。

3. 设置支持度阈值*min\_sup*=0.01，置信度阈值*min\_conf*=0.3，在Groceries数据集上进行关联规则挖掘，并对挖掘结果进行分析（可列出全部或部分结果，挑选部分结果进行分析，可从挖掘到的频繁项集和强规则的含义等进行分析）。注：若因电脑性能不足，在上述阈值设置下不能跑出结果，可适当更改支持度阈值设置。

4. 改变支持度阈值和置信度阈值，观察Apriori算法生成的频繁项集数量和强规则数量，以及算法生成频繁项集所用时间和产生强规则所用时间。

5. 实验总结与思考。

【实验要求】：

（1）独立完成实验，拒绝抄袭。撰写实验报告，需包含实验目的，实验原理及过程，实验结果展示等。

（2）实现Apriori算法时，可根据能力只实现频繁项集生成，或者频繁项集生成和强规则产生都实现，鼓励有余力的同学都实现。

（3）可根据自己能力借助合适的工具（如sklearn库）完成实验，但需清楚算法原理。鼓励自己动手编程实现算法，评分时，自己动手编程实现>借助工具。

（4）附代码（使用工具的附操作过程）。