

HOMEWORK #3 (10分)

一、数据立方体练习：(2分)

假定某一数据仓库包含4个维：date(日期)，spectator(观众)，location(地点)和game(节目)；2个度量count和charge，count是观众的人数，charge是观众在某日期某地点观看某节目的费用。观众分三类：学生、成年人和老人，每类观众有不同的收费标准。

- 1) 画出该数据仓库的星型模式图Star Schema（自己定义维表属性）。
- 2) 从基本方体[date, spectator, location, game]开始，为列出2018年学生观众在清华大学大礼堂的总付费，应当执行哪些OLAP操作，并说明原因。

二、频繁项集与关联规则：(4分)

数据库中有四条购物记录，如下表所示。设 $\min_sup=2$

- 1) 利用 Apriori 算法找出所有含物品 E 的频繁项集。
- 2) 设 confidence=75%，写出满足下面元规则的关联规则：buy(x) & buy(y) \rightarrow buy(z)。

TID	Items
T01	{K, A, D, B}
T02	{D, A, C, E, B}
T03	{C, A, B, E}
T04	{B, A, D}

三、约束的性质：(4分)

假定候选项集中的所有元素都是整数，判断下列对于 S 的约束是否满足 Antimonotone、Monotone、Succinct 三条性质（对于 Succinct 含混不清的可以不填）

Constraint	Antimonotone	Monotone	Succinct
$0 \notin S$			
S 中的正数数量大于 5			
S 中只包含 3 的倍数			
$\min(S) > 0 \ \& \ \max(S) < 10$			
S 的方差小于 1			

有问题请邮件联系助教 shijx12@163.com。