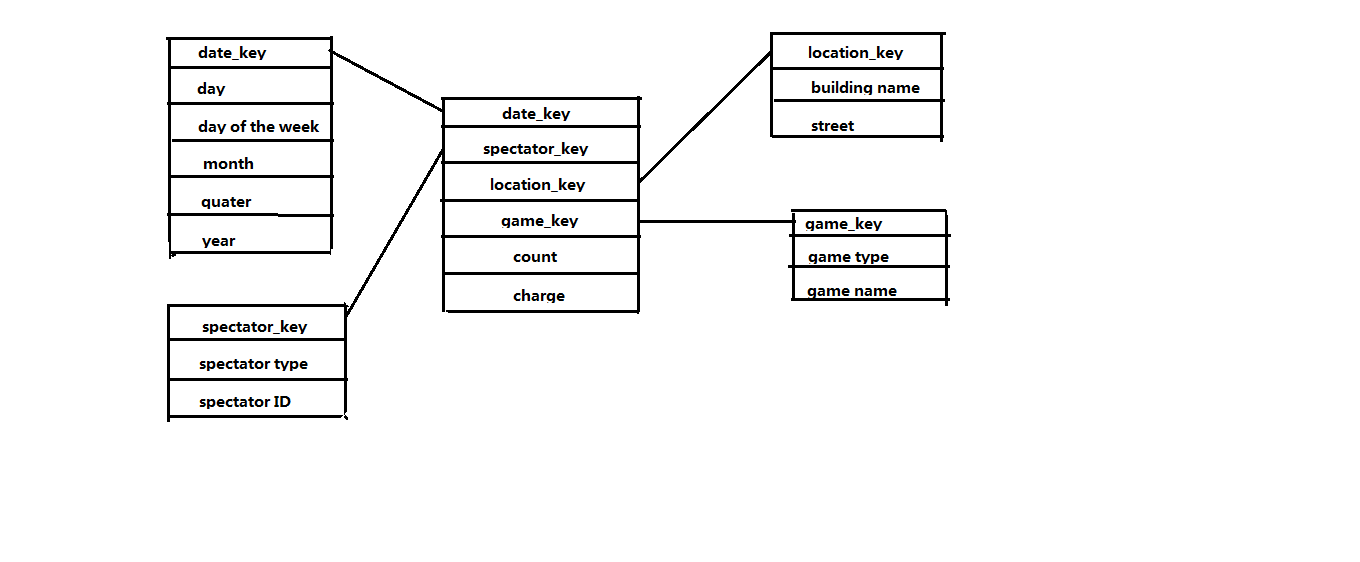
一、

1）星型模式图如下

 2）

假设数据立方体初始时time按照day聚集，location按照street聚集，spectator按照spectator type聚集，game按照game type聚集，显示的度量为Charge。

需要执行的OLAP操作如下：

date\_key上卷到year，因为列出的时间是2018年以年为单位；

location\_key 下钻到building name，因为列出的地点大礼堂精确到了建筑物；

切块（time = 2018）and （spectator type = student）and （building name = DaLITang），因为要按照给出的条件筛选出满足条件的数据；

由于要计算满足条件的数据的总付费，经过以上OLAP操作可以通过求和得到所需的分析结果。

二、

1)

为执行Apriori算法，先对项集中的项进行字典序排序，

结果如下：

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

1. 第一次迭代，候选1项集中，K只出现了一次，小于min\_sup，进行剪枝，候选一项集为{A，B，C，D，E}
2. 使用L1连接产生候选2项集，CD只出现了一次，小于min\_sup，进行剪枝，候选2项集为{AB，AC，AD，AE，BC，BD，BE，CE}
3. 使用L2连接产生候选3项集，ACD，BCD只出现了一次，小于min\_sup，进行剪枝，候选3项集为{ABC，ABD，ABE，ACE，BCE}
4. 使用L3连接产生候选4项集，ABCD，ABDE只出现了一次，小于min\_sup，进行剪枝，候选4项集为{ABCE}
5. 使用L3连接产生候选5项集为空集，算法结束。

综上可得，包含物品E的频繁项集有E，AE，BE，CE，ABE，ACE，BCE, ABCE ;

2)

P(ABC | AB) = 2 / 4 = 50%;

P(ABC | BC) = 2 / 2 = 100%; buy(B) & buy(C) -> buy(A)

P(ABC | AC) = 2 / 2 = 100%; buy(A) & buy(C) -> buy(B)

P(ABD | AB) = 3 / 4 = 75%；buy(A) & buy(B) -> buy(D)

P(ABD | AD) = 3 / 3 = 100%; buy(A) & buy(D) -> buy(B)

P(ABD | BD) = 3 / 3 = 100%; buy(B) & buy(D) -> buy(A)

P(ABE | AB) = 2 / 4 = 50%;

P(ABE | BE) = 2 / 2 = 100%; buy(B) & buy(E) -> buy(A)

P(ABE | AE) = 2 / 2 = 100%; buy(A) & buy(E) -> buy(B)

P(ACE | AC) = 2 / 2 = 100%; buy(A) & buy(C) -> buy(E)

P(ACE | AE) = 2 / 2 = 100%; buy(A) & buy(E) -> buy(C)

P(ACE | CE) = 2 / 2 = 100%; buy(C) & buy(E) -> buy(A)

P(BCE | BC) = 2 / 2 = 100%; buy(B) & buy(C) -> buy(E)

P(BCE | BE) = 2 / 2 = 100%; buy(B) & buy(E) -> buy(C)

P(BCE | CE) = 2 / 2 = 100%; buy(C) & buy(E) -> buy(B)

三、

|  |  |  |  |
| --- | --- | --- | --- |
| Constraint | Antimonotone | Monotone | Succinct |
| 0 ∉ S | Yes | No | Yes |
| S 中的正数数量大于 5 | No | Yes | Week |
| S 中只包含 3 的倍数 | Yes | No | Yes |
| min(S) > 0 & max(S) <10 | Yes | No | Yes |
| S 的方差小于 1 | No | No | No |