数据挖掘大作业二：关联规则挖掘

芦美艳 2120171042

1. 数据源

Building\_Permits.csv

2. 实验内容

1. 数据集处理，关联规则挖掘数据形式；

选择“**Permit Type（5），Current Status（14），Structural Notification（1），Voluntary Soft-Story Retrofit（1），Fire Only Permit（1），TIDF Compliance（2），Existing Construction Type（5），Proposed Construction Type（5），Site Permit（1）”十个属性，每个属性对应一到多个值，每个值为一个项，构成共33项，每个事务以list存储进行分析。**

e.g.

1. 频繁项集；

#频繁1项集

[frozenset({'Proposed Construction Type = 5.0'}), frozenset({'Permit Type = 8'}), frozenset({'Existing Construction Type = 5.0'}), frozenset({'Current Status = complete'})],

#频繁2项集

[frozenset({'Current Status = complete', 'Permit Type = 8'}), frozenset({'Permit Type = 8', 'Existing Construction Type = 5.0'}), frozenset({'Proposed Construction Type = 5.0', 'Existing Construction Type = 5.0'}), frozenset({'Permit Type = 8', 'Proposed Construction Type = 5.0'})],

#频繁3项集

[frozenset({'Permit Type = 8', 'Proposed Construction Type = 5.0', 'Existing Construction Type = 5.0'})],

[]

1. 关联规则，支持度(support)和置信度(conf)

frozenset({'Permit Type = 8'})-->frozenset({'Current Status = complete'}) support:0.5594669504318345 conf:0.6387156468735212  
frozenset({'Current Status = complete'})-->frozenset({'Permit Type = 8'}) support:0.5594669504318345 conf:0.9315477723409734  
frozenset({'Permit Type = 8'})-->frozenset({'Existing Construction Type = 5.0'}) support:0.6254732856541859 conf:0.7140718034199504  
frozenset({'Existing Construction Type = 5.0'})-->frozenset({'Permit Type = 8'}) support:0.6254732856541859 conf:0.8919188354653727  
frozenset({'Proposed Construction Type = 5.0'})-->frozenset({'Existing Construction Type = 5.0'}) support:0.6925561137370387 conf:0.9786679722333933  
frozenset({'Existing Construction Type = 5.0'})-->frozenset({'Proposed Construction Type = 5.0'}) support:0.6925561137370387 conf:0.9875782973092191  
frozenset({'Permit Type = 8'})-->frozenset({'Proposed Construction Type = 5.0'}) support:0.6324519290257121 conf:0.7220389741561368  
frozenset({'Proposed Construction Type = 5.0'})-->frozenset({'Permit Type = 8'}) support:0.6324519290257121 conf:0.8937332797118428  
frozenset({'Permit Type = 8'})-->frozenset({'Existing Construction Type = 5.0', 'Proposed Construction Type = 5.0'}) support:0.6248422381152714 conf:0.7133513677682741  
frozenset({'Proposed Construction Type = 5.0'})-->frozenset({'Existing Construction Type = 5.0', 'Permit Type = 8'}) support:0.6248422381152714 conf:0.8829798394852337  
frozenset({'Existing Construction Type = 5.0'})-->frozenset({'Proposed Construction Type = 5.0', 'Permit Type = 8'}) support:0.6248422381152714 conf:0.8910189677988531

1. 使用Lift对规则进行评价

frozenset({'Current Status = complete'})—> frozenset({'Permit Type = 8'}), lift：1.063501852156049

frozenset({'Permit Type = 8'}) --> frozenset({'Current Status = complete'}), lift: 1.063501852156049

frozenset({'Permit Type = 8'}) --> frozenset({'Existing Construction Type = 5.0'}), lift: 1.018259461999004

frozenset({'Existing Construction Type = 5.0'})--> frozenset({'Permit Type = 8'}), lift: 1.018259461999004

frozenset({'Proposed Construction Type = 5.0'})--> frozenset({'Existing Construction Type = 5.0'}), lift: 1.3955710309653

frozenset({'Existing Construction Type = 5.0'})--> frozenset({'Proposed Construction Type = 5.0'}), lift: 1.3955710309653

frozenset({'Proposed Construction Type = 5.0'})--> frozenset({'Permit Type = 8'}), lift: 1.0203309229310673

frozenset({'Permit Type = 8'})--> frozenset({'Proposed Construction Type = 5.0'}), lift: 1.0203309229310673

frozenset({'Proposed Construction Type = 5.0'})--> frozenset({'Existing Construction Type = 5.0', 'Permit Type = 8'}), lift: 1.4116987243695314

frozenset({'Permit Type = 8'})--> frozenset({'Existing Construction Type = 5.0', 'Proposed Construction Type = 5.0'}), lift: 1.0300268146057134

frozenset({'Existing Construction Type = 5.0'})--> frozenset({'Permit Type = 8', 'Proposed Construction Type = 5.0'}), lift: 1.4088327142451156