## 问题三

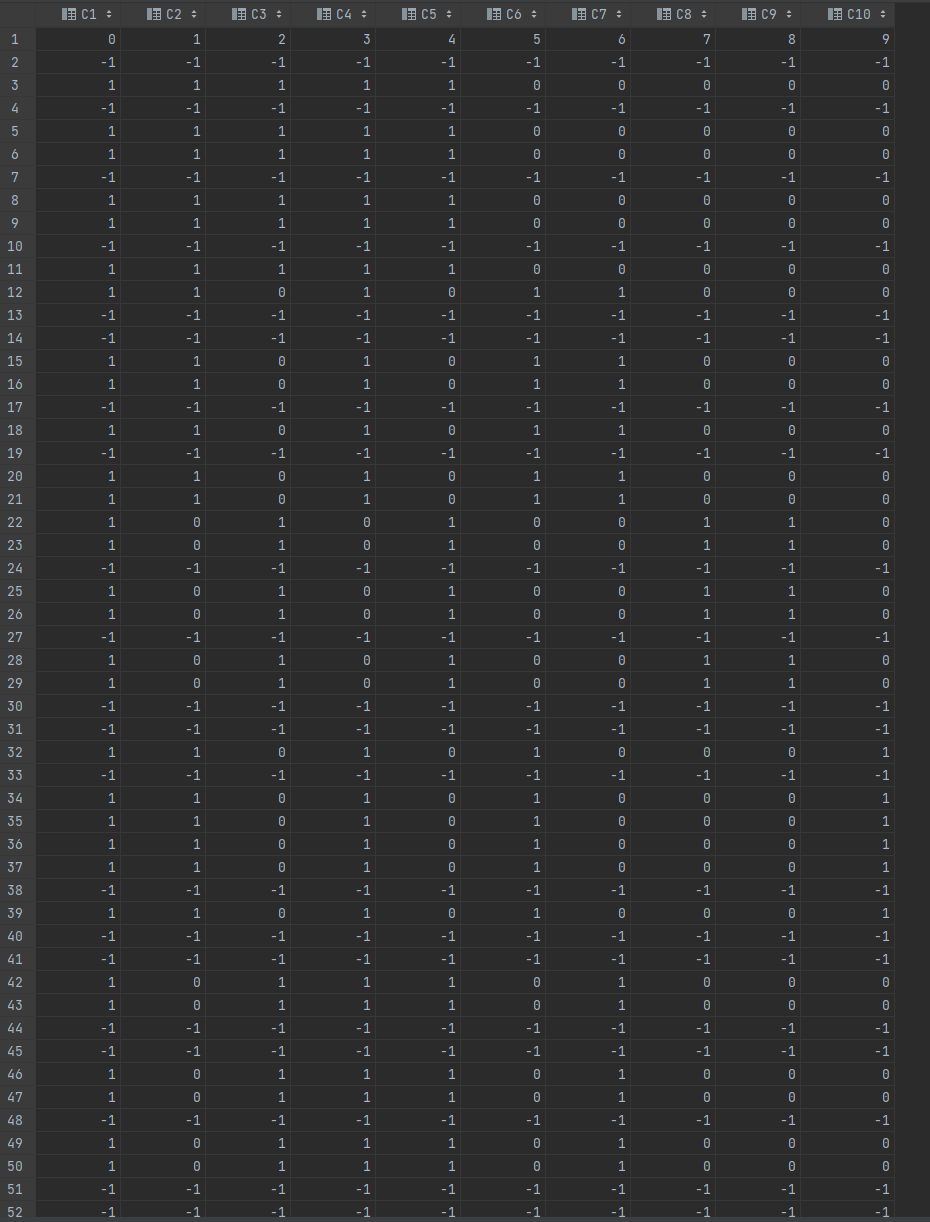
对于该问题，我们采用Z-score规约进行调分。如此我们便可以使用调分后的打分直接作平均值来评估小组的论文得分情况。

由于题目中存在四个打分步骤，而每个打分步骤结束以后都要进行调分，对调分后的结果进行排名，然后选出这一步骤中的淘汰小组和获奖小组。同时，由于一旦某个小组获得了奖项或者被淘汰，那么该小组的评委打分数据就不能再次被使用。这就造成了数据集过少和数据集状态问题。而对于缺失的数据集，直接使用Z-score规约会使得结果出现很大的误差，所以我们需要填充数据集。为此我们采用了掩蔽矩阵，领域预测和Z-score规约来设计我们的调分算法。

### 模型部分组件构成

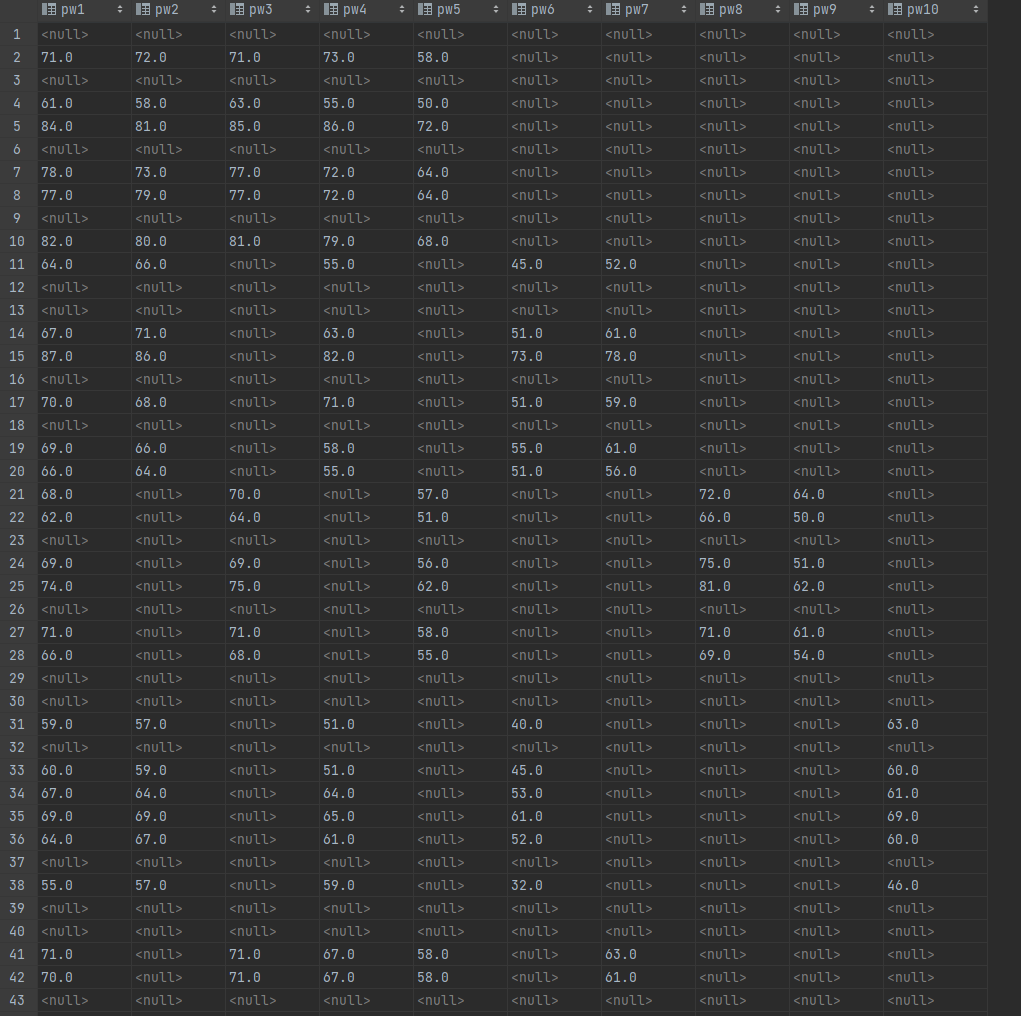
#### 掩蔽矩阵

为了解决数据集状态问题，即每个步骤中可用的数据都是变化的，我们设计了掩蔽矩阵来描述每个步骤中可以被使用的数据。以下为矩阵数值的公式和部分矩阵展示



步骤二的掩蔽矩阵

在根据给定的分组情况获取对用的掩蔽矩阵后，由于掩蔽矩阵可以与原数据对应，我们使用掩蔽矩阵对原数据进行处理，保留本轮需要使用的数据，将不需要的数据置为NaN，最终得到的处理完毕的数据如下所示



步骤二的数据

#### 领域预测

由于每个评委不可能对所有论文打分，所以出现了数据集的缺失。同时，如果每个评委对每篇论文都打分了那么这样的结果一定是最好的。为了使现有的缺失数据集达到后者的完备状态以进行更好地调分，我们采用了领域预测的方式，通过计算每个评委在每篇论文的打分上的相关性，预测该评委在其未打分的小组上的可能的打分，使得每个小组都被全部十位评委打分，从而使数据集达到如问题二的数据的完备形式。

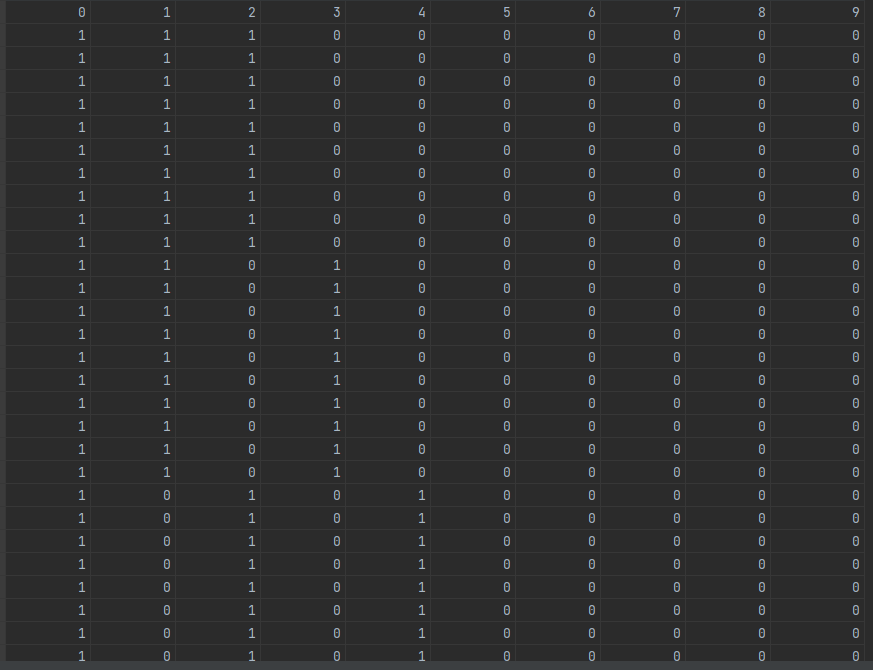
#### Z-score规约

在获取了完备数据集的情况下，我们便可以参照问题二的解法，对于数据集作Z-score规约，从而使得每个评委的打分情况全被被规约到一个统一的标准，使直接的数值平均能反映小组得分的真实情况。

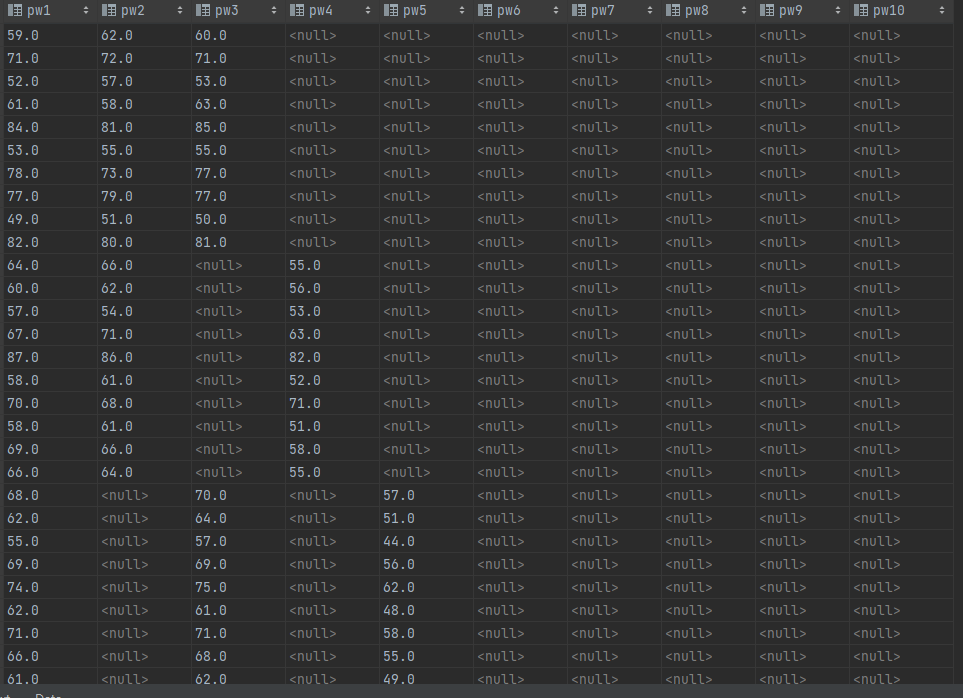
### 模型详细说明

#### 步骤一

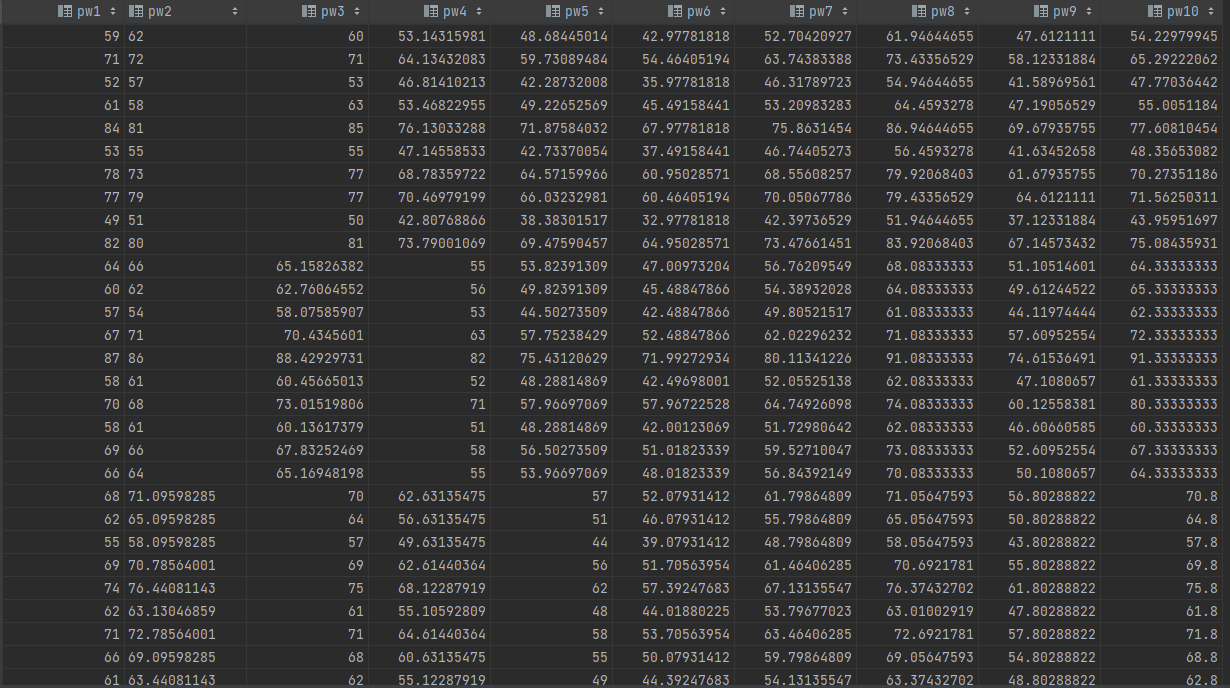
在步骤一中，各个评委的分组情况是给定的，因此我们需要根据分组情况，获取在这一轮中的掩蔽矩阵，步骤一的掩蔽矩阵中的部分数据如下图：



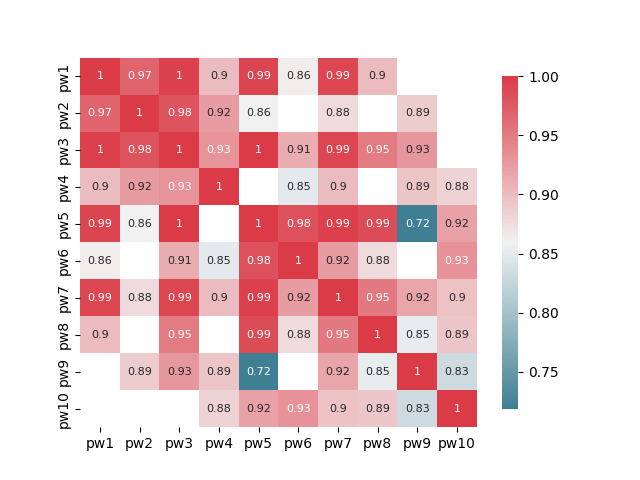
在获取了掩蔽矩阵后，我们根据该掩蔽矩阵对于原始的数据作掩蔽处理，获取在步骤一中可见的数据。步骤一的可见数据部分数据如下：



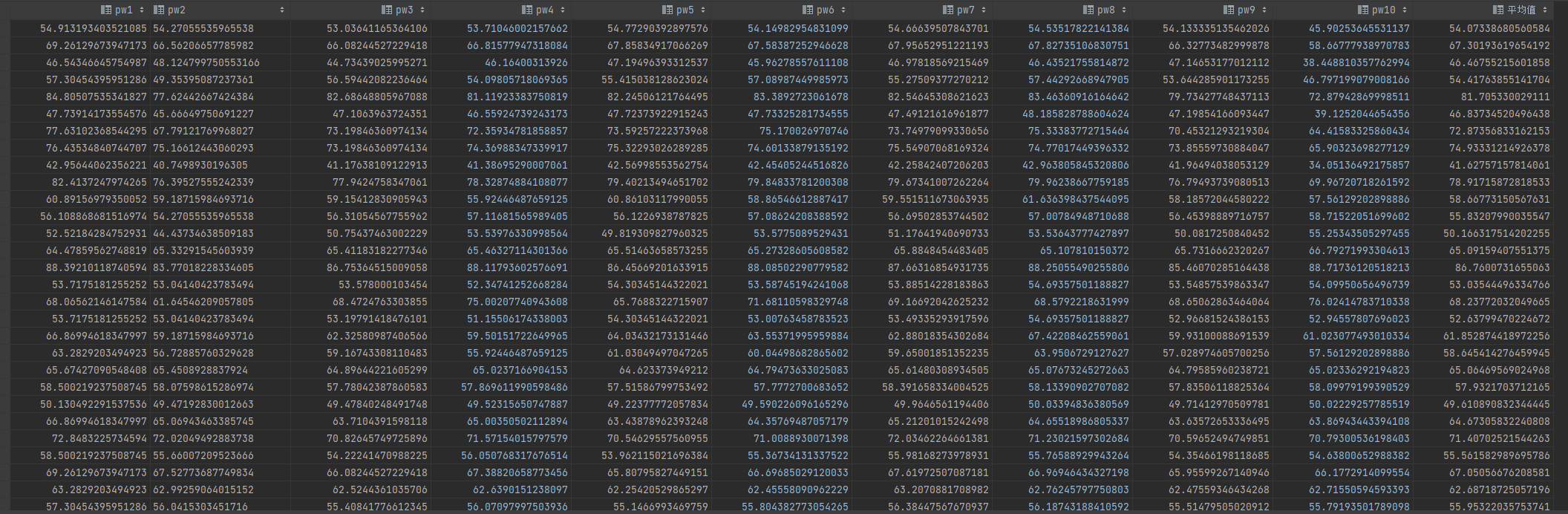
在获取了可见数据后，为了扩充数据集，我们对于该可见数据作领域预测，填补表中的数据空缺。填补的部分结果如下：



领域预测计算的相关性矩阵如下：



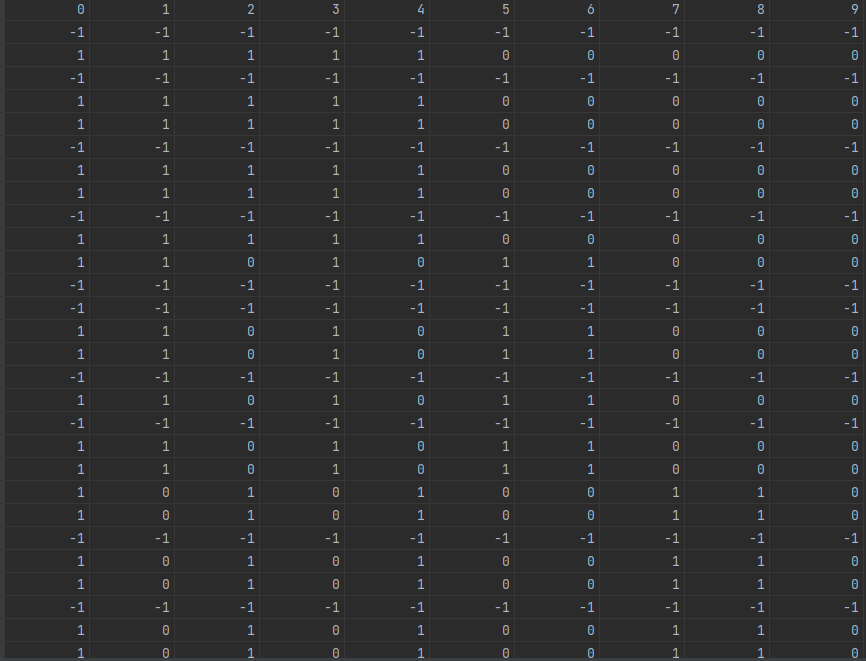
在获取到填补结果后，便可以对该结果作Z-score规约，然后对于Z-score规约后的结果，我们可以直接根据掩蔽矩阵，计算该小组在步骤一中的对应的三个评委的打分的平均值，最终部分结果如下：



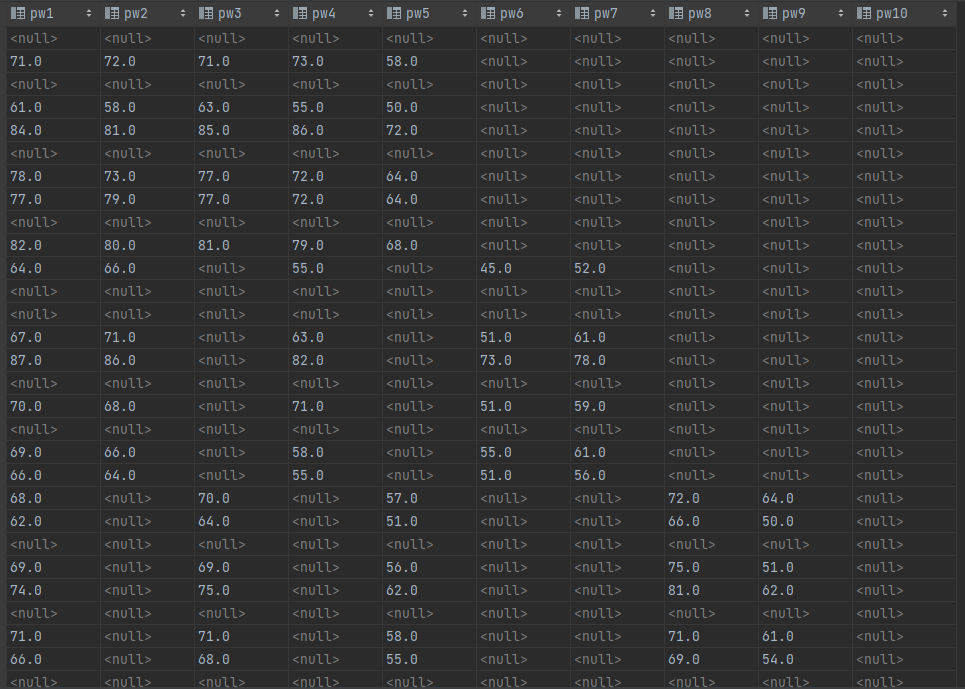
根据步骤一的排序方式，我们对每个10组大组进行排序，淘汰该大组中排名为后40%的小组，更新掩蔽矩阵，将已经被淘汰的小组对应的掩蔽矩阵中的值置为-1.

#### 步骤二

在步骤二中，各个评委的分组情况是给定的，因此我们需要根据分组情况，获取在这一轮中的掩蔽矩阵，步骤二的掩蔽矩阵中的部分数据如下图：

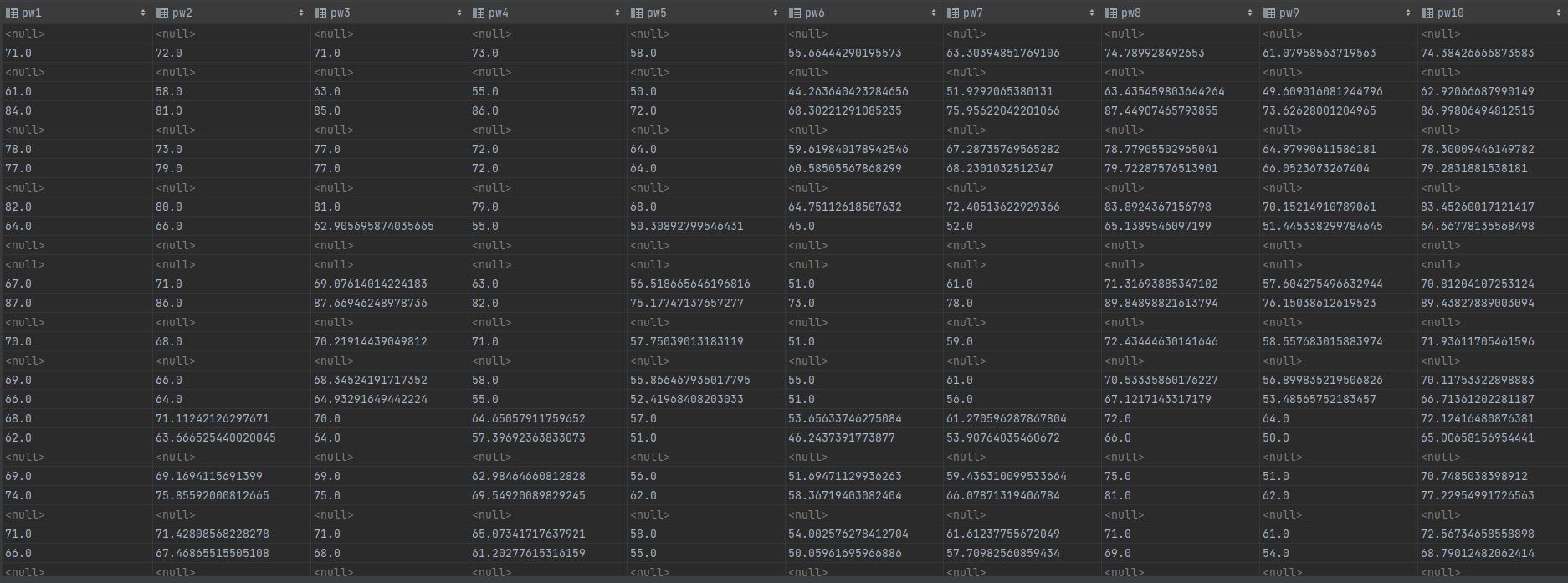


在获取了掩蔽矩阵后，我们根据该掩蔽矩阵对于原始的数据作掩蔽处理，获取在步骤二中可见的数据。步骤二的可见数据的部分数据如下：

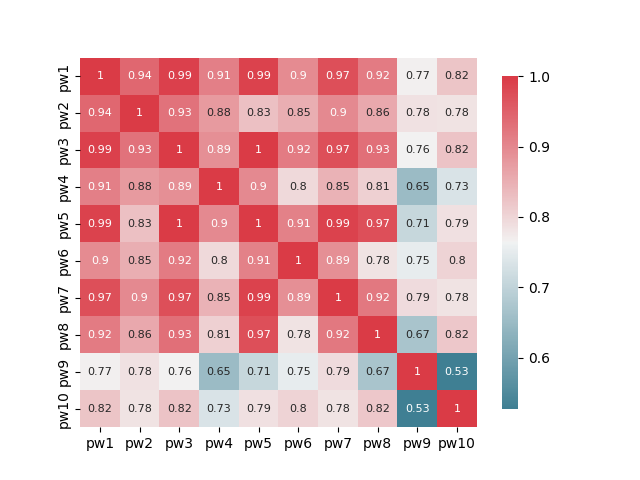


如图所示，第一组在步骤一中已被淘汰，因此在步骤二中，第一组的数据便不再可用。

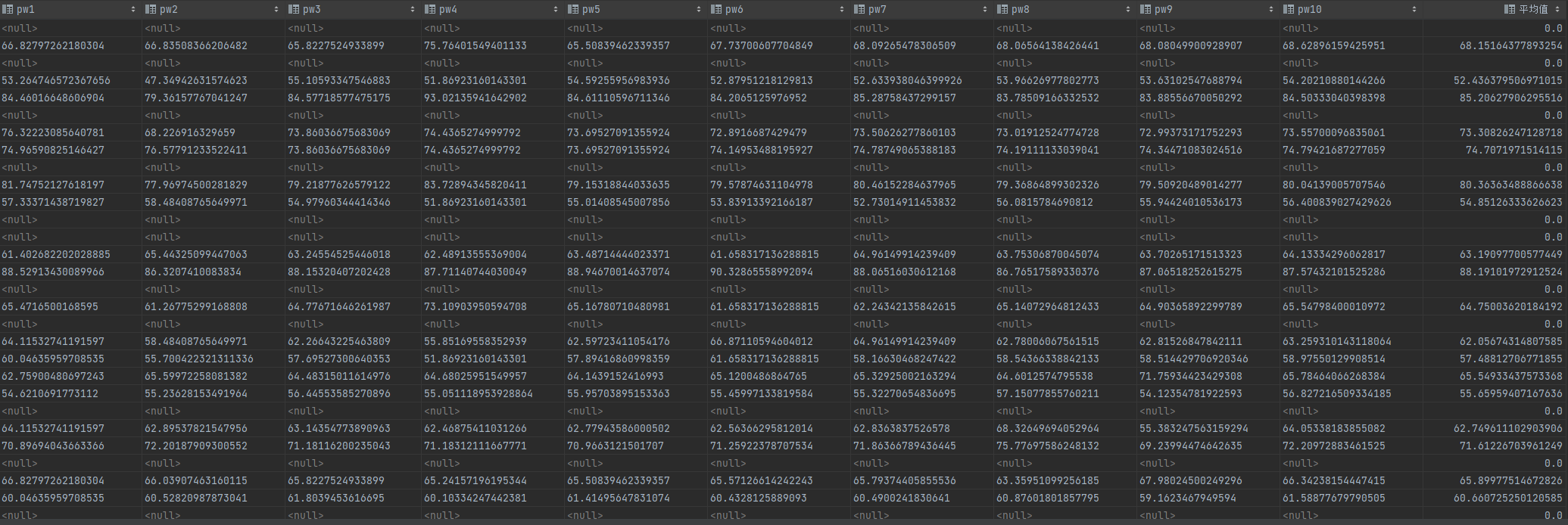
在获取了可见数据后，为了扩充数据集，我们对于该可见数据作领域预测，填补表中的数据空缺。填补的部分结果如下：



领域预测计算的相关性矩阵如下：



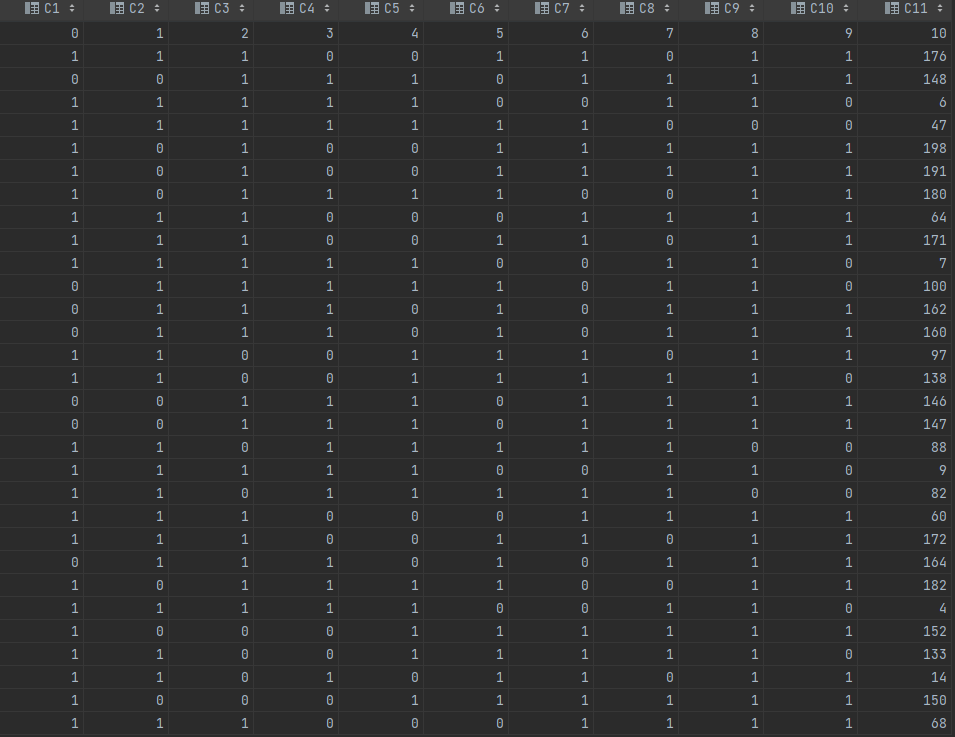
在获取到填补结果后，便可以对该结果作Z-score规约，然后对于Z-score规约后的结果，我们可以直接根据掩蔽矩阵，计算该小组在步骤二中的对应的五个评委的打分的平均值，最终部分结果如下：

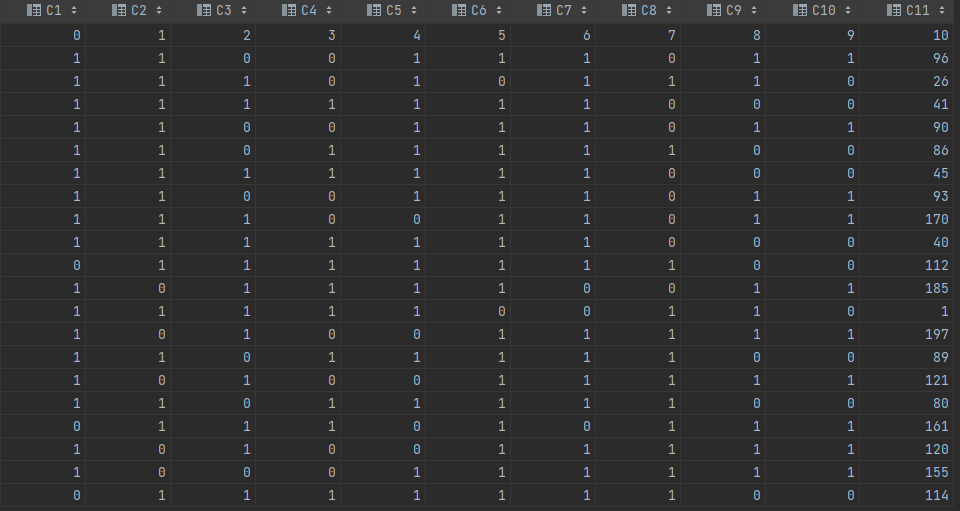


根据步骤二的排序方式，我们对每个10组大组的所有未被淘汰的论文进行排序，淘汰排名为后20名的小组，排名31到40的小组，获二等奖。排名61到100的小组，获三等奖。

#### 步骤三

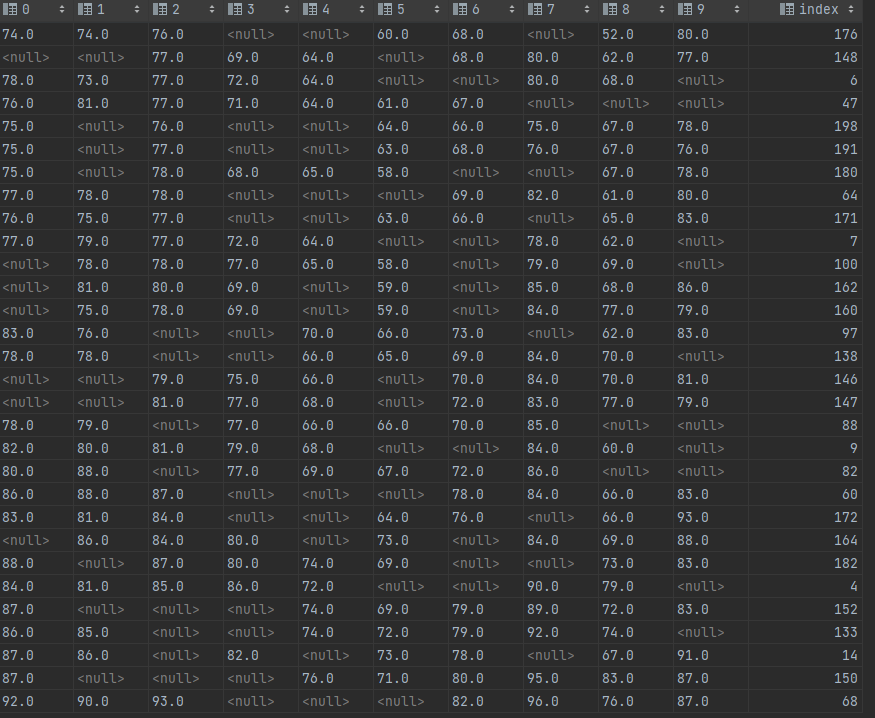
在步骤三中，我们需要对步骤二中排名1到30和41到60的论文，每次额外随机抽取两位评委进行评分，我们对两组论文分别随机选取评委，生成他们对应的掩蔽矩阵，如下图所示：

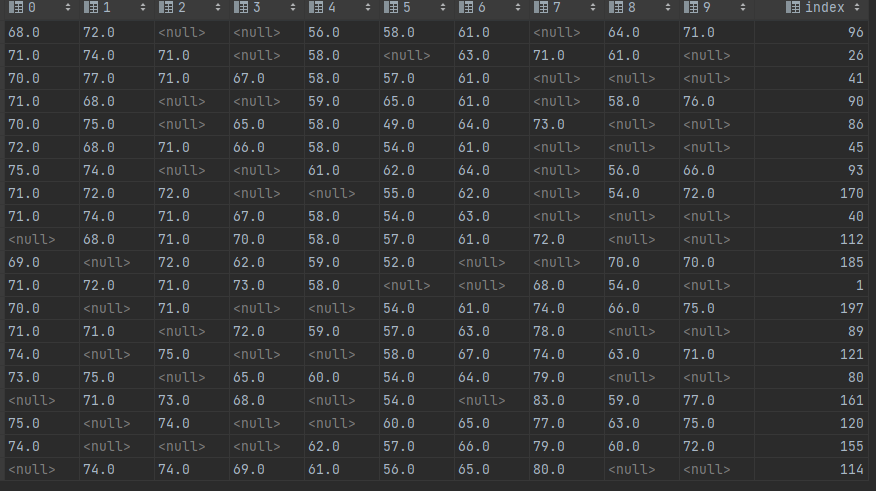




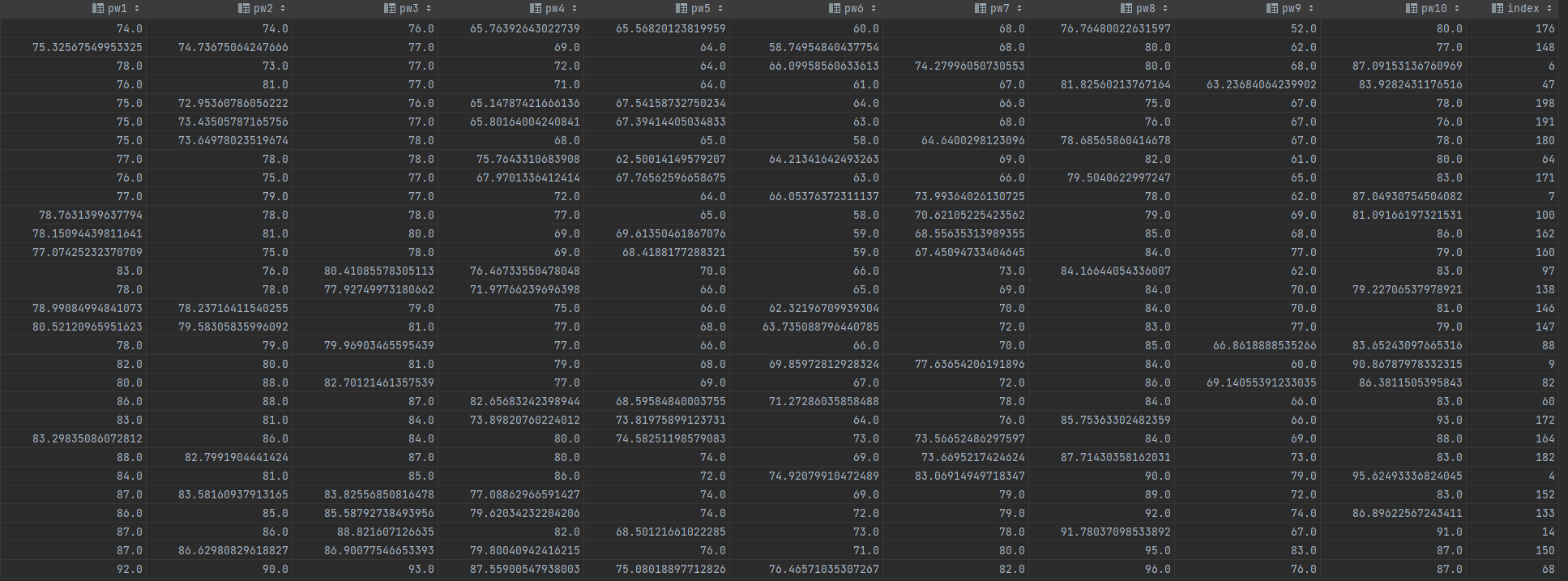
需要指出的是，由于先前的已经被淘汰的小组和已经获奖的小组的数据无法使用，于是我们缩小了掩蔽矩阵，添加了新的列来表征对应的小组号。

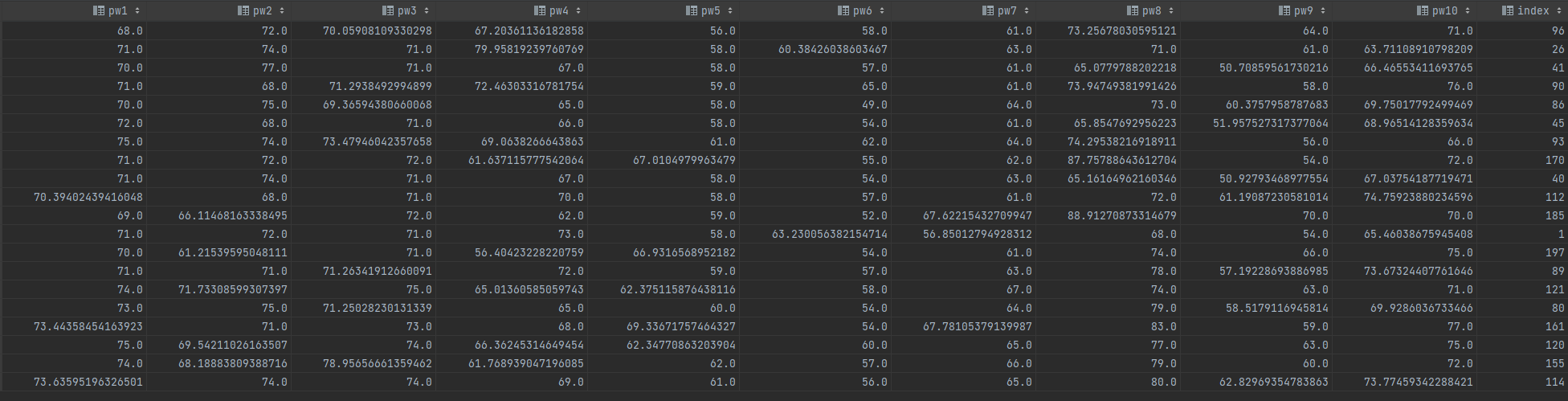
在获取了掩蔽矩阵后，我们根据该掩蔽矩阵对于原始的数据作掩蔽处理，获取在步骤三中可见的数据。步骤三的可见数据的部分数据如下：



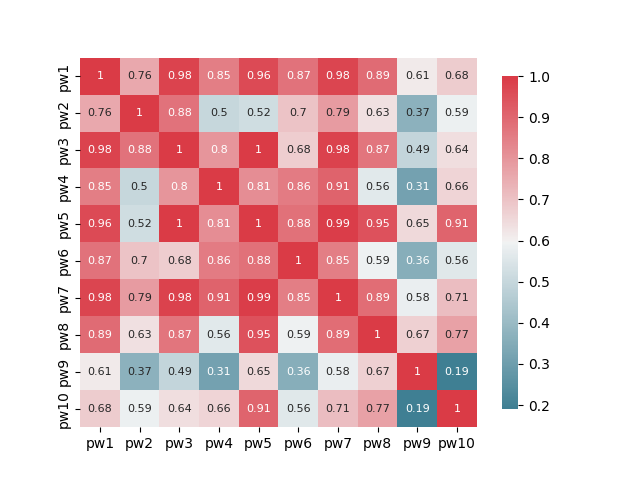


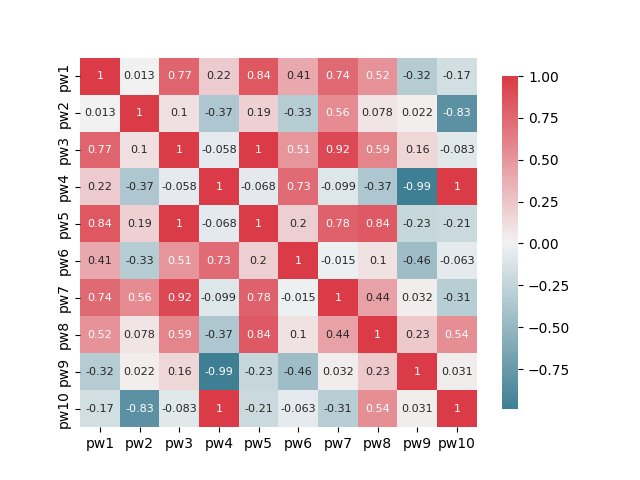
在获取了可见数据后，为了扩充数据集，我们对于该可见数据作领域预测，填补表中的数据空缺。填补的部分结果如下：



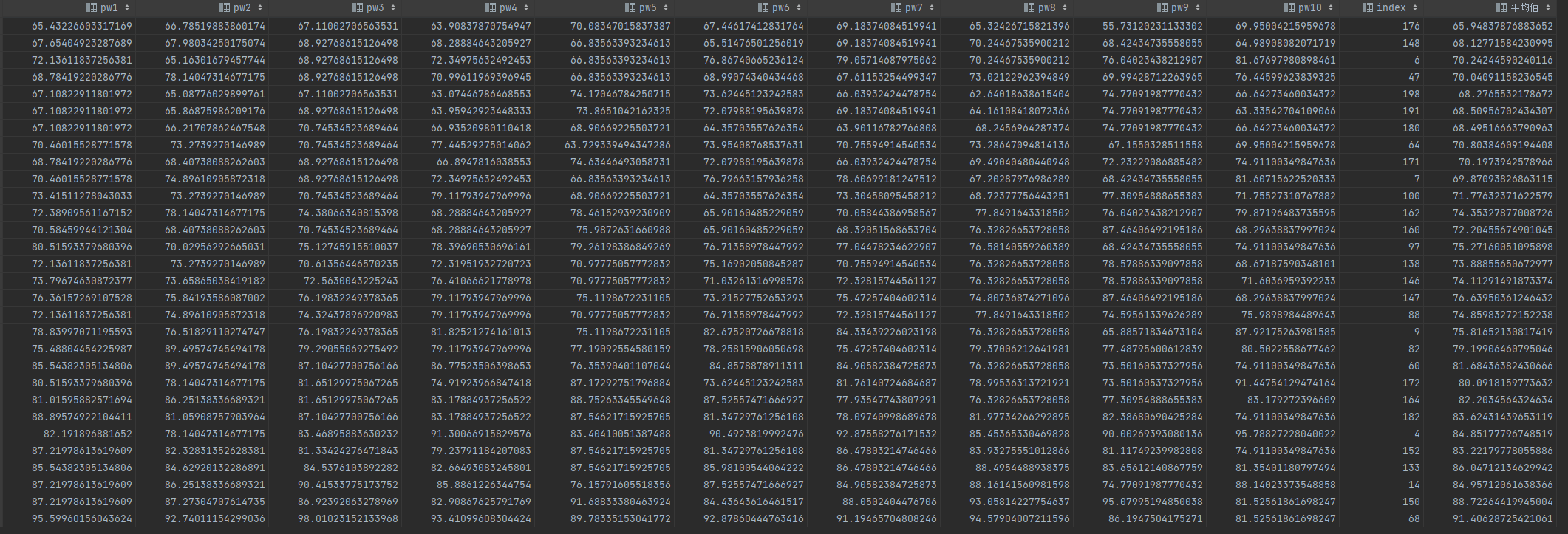


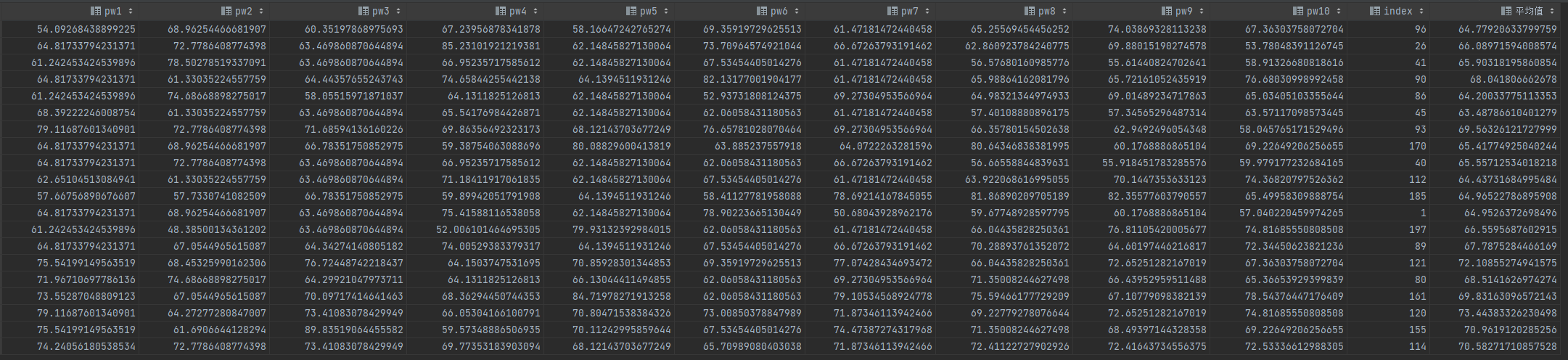
领域预测计算的相关性矩阵如下：





在获取到填补结果后，便可以对该结果作Z-score规约，然后对于Z-score规约后的结果，我们可以直接根据掩蔽矩阵，计算该小组在步骤三中的对应的七个评委的打分的平均值，最终部分结果如下：



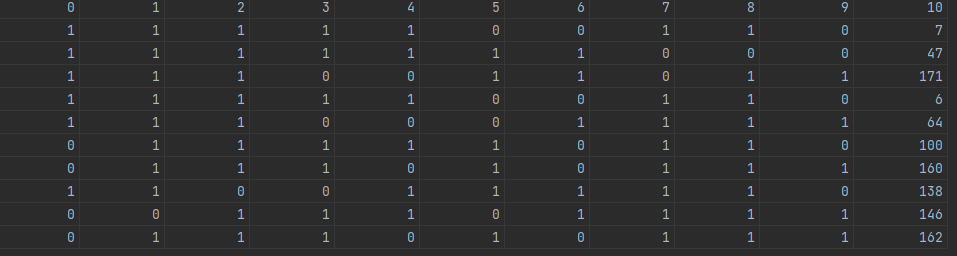


根据步骤三的排序方式，原排名41到60的20篇论文，7位评委给分平均，新排名41-50的获二等奖，新排名51-60的获三等奖；

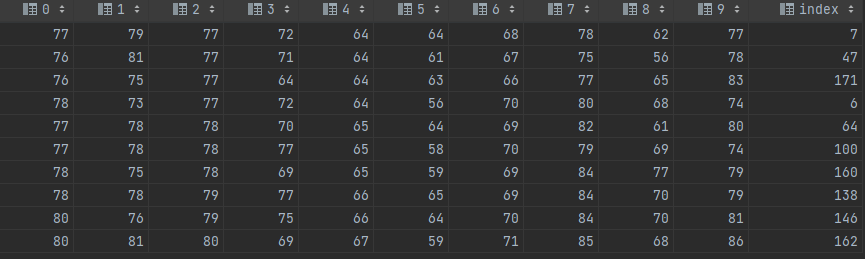
原排名1到30的30篇论文，7位评委给分平均，新排名26-30名获二等奖，新排名1-15名获一等奖。

#### 步骤四

在步骤四中，我们需要对步骤三中排名16到25的论文，让全部十位评委进行评分，我们生成对应的掩蔽矩阵，如下图所示：

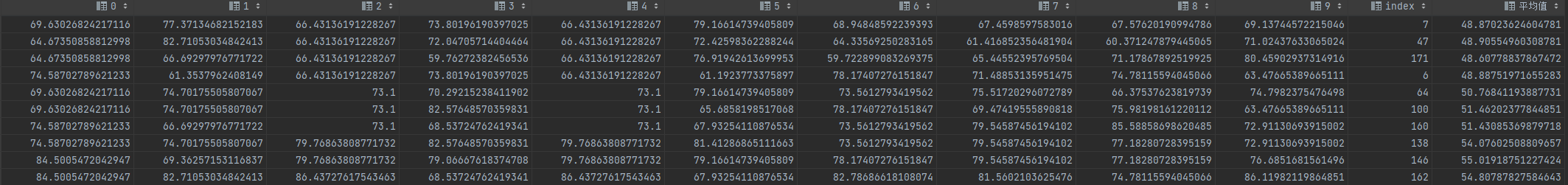


在获取了掩蔽矩阵后，我们根据该掩蔽矩阵对于原始的数据作掩蔽处理，获取在步骤四中可见的数据。步骤四的可见数据如下：



由于步骤四中全部十位评委都参与了打分，因此我们不需要使用领域预测填补数据。

在获取到填补结果后，便可以对该结果作Z-score规约，然后对于Z-score规约后的结果，我们可以直接根据掩蔽矩阵，计算该小组在步骤三中的对应的七个评委的打分的平均值，最终部分结果如下：



根据步骤四的排序方式，最终排名16-20的获一等奖，最终排名21-25的获二等奖。

### 评奖结果展示

第一步淘汰:
  
第1大组：
  
第9组：,得分：41.62757157814061
  
第3组：,得分：46.46755215601858
  
第6组：,得分：46.83734520496438
  
第1组：,得分：54.07338680560584
  
第2大组：
  
第13组：,得分：50.166317514202255
  
第18组：,得分：52.63799470224672
  
第16组：,得分：53.03544496334766
  
第12组：,得分：55.83207990035547
  
第3大组：
  
第23组：,得分：49.610890832344445
  
第26组：,得分：55.561582989695786
  
第29组：,得分：55.95322035753741
  
第30组：,得分：57.14197457737627
  
第4大组：
  
第39组：,得分：33.67957730702623
  
第37组：,得分：45.13321526030566
  
第40组：,得分：49.48776975396152
  
第32组：,得分：50.24582331941354
  
第5大组：
  
第47组：,得分：40.29612954344943
  
第43组：,得分：42.285224689203794
  
第44组：,得分：44.67736345794551
  
第50组：,得分：48.653614568461656
  
第6大组：
  
第53组：,得分：46.364672417773846
  
第57组：,得分：47.14357980980223
  
第56组：,得分：51.82391835484545
  
第51组：,得分：55.33346256730272
  
第7大组：
  
第67组：,得分：37.76898353529103
  
第63组：,得分：39.99722560293798
  
第66组：,得分：49.869159946684135
  
第64组：,得分：50.87550613831246
  
第8大组：
  
第78组：,得分：31.398463560754838
  
第80组：,得分：37.812845022042765
  
第72组：,得分：52.13017322194364
  
第74组：,得分：54.12969130695421
  
第9大组：
  
第85组：,得分：49.16192404340551
  
第82组：,得分：58.454126279952
  
第86组：,得分：60.79599442502964
  
第84组：,得分：61.2755719247744
  
第10大组：
  
第92组：,得分：49.91656177138068
  
第99组：,得分：55.06172563578984
  
第93组：,得分：59.00600433741463
  
第96组：,得分：61.44130056345612
  
第11大组：
  
第109组：,得分：52.80426415118512
  
第108组：,得分：53.97259856423148
  
第106组：,得分：57.16016236288374
  
第105组：,得分：57.53695546001668
  
第12大组：
  
第116组：,得分：44.69576453607479
  
第120组：,得分：52.2649122525901
  
第117组：,得分：52.66831398135178
  
第111组：,得分：55.05857075621725
  
第13大组：
  
第126组：,得分：40.770968172157474
  
第125组：,得分：56.13166982926713
  
第127组：,得分：58.5311522842757
  
第130组：,得分：58.92648663635611
  
第14大组：
  
第131组：,得分：37.51539544146138
  
第140组：,得分：53.919174032454855
  
第138组：,得分：57.85516380431371
  
第135组：,得分：58.60143347780309
  
第15大组：
  
第145组：,得分：49.427299740465145
  
第146组：,得分：50.60915253572554
  
第142组：,得分：51.36501083237072
  
第150组：,得分：54.51095222400619
  
第16大组：
  
第155组：,得分：36.03082398155854
  
第154组：,得分：48.88360395567121
  
第158组：,得分：49.67875744907388
  
第152组：,得分：52.81242375576321
  
第17大组：
  
第167组：,得分：34.049089147169944
  
第169组：,得分：39.07565947428069
  
第170组：,得分：43.35060102286068
  
第164组：,得分：46.02544394807003
  
第18大组：
  
第179组：,得分：51.49537313019601
  
第174组：,得分：52.69923058941942
  
第175组：,得分：64.48418525317908
  
第180组：,得分：66.08411212470793
  
第19大组：
  
第182组：,得分：30.614532915736746
  
第184组：,得分：51.66374619582359
  
第188组：,得分：53.12793938598188
  
第190组：,得分：57.85231737314734
  
第20大组：
  
第195组：,得分：46.58969111991704
  
第194组：,得分：46.96406395362399
  
第200组：,得分：54.29352742955749
  
第197组：,得分：54.67881400571556
  
第二步淘汰:
  
第38组：,得分：43.61374744112697
  
第31组：,得分：48.9388142945343
  
第33组：,得分：50.314913666880265
  
第193组：,得分：51.88974551725745
  
第4组：,得分：52.436379506971015
  
第143组：,得分：52.780130209241136
  
第77组：,得分：54.5519982632609
  
第60组：,得分：54.74446826697863
  
第11组：,得分：54.85126333626623
  
第73组：,得分：55.46705543057279
  
第22组：,得分：55.65959407167636
  
第45组：,得分：55.93246683912294
  
第191组：,得分：55.97418604632105
  
第52组：,得分：56.315461938448365
  
第144组：,得分：56.64885200722707
  
第160组：,得分：57.19066023675593
  
第20组：,得分：57.48812706771855
  
第112组：,得分：57.6032730225793
  
第36组：,得分：58.10635609046301
  
第185组：,得分：58.27281577838613
  
第二步获二等奖:
  
第118组：,得分：58.43681551454781
  
第59组：,得分：58.57581036283182
  
第141组：,得分：58.79474274121752
  
第107组：,得分：58.8251027381753
  
第55组：,得分：58.90850055225599
  
第133组：,得分：58.93877967501074
  
第34组：,得分：59.39388076248555
  
第114组：,得分：59.75620607445977
  
第110组：,得分：59.85040154458056
  
第58组：,得分：59.95232699304991
  
第71组：,得分：60.08399822141032
  
第79组：,得分：60.31727749646528
  
第28组：,得分：60.660725250120585
  
第137组：,得分：60.794596452037936
  
第157组：,得分：60.860741399545056
  
第159组：,得分：60.88088441247364
  
第136组：,得分：60.91702143616068
  
第187组：,得分：61.66185654285745
  
第95组：,得分：61.72453855337094
  
第19组：,得分：62.05674314807585
  
第124组：,得分：62.62173690785424
  
第123组：,得分：62.74006818467808
  
第24组：,得分：62.74961110290391
  
第103组：,得分：62.83057132176706
  
第14组：,得分：63.19097700577449
  
第104组：,得分：63.33278086941671
  
第119组：,得分：63.340011894208416
  
第88组：,得分：63.37173343873786
  
第75组：,得分：63.56895988357068
  
第176组：,得分：63.68087604799259
  
第54组：,得分：64.10933810731505
  
第68组：,得分：64.32443155074687
  
第70组：,得分：64.44763227815702
  
第62组：,得分：64.60312572938324
  
第17组：,得分：64.75003620184192
  
第128组：,得分：64.85158128589937
  
第189组：,得分：65.21395784118504
  
第21组：,得分：65.54933437573368
  
第49组：,得分：65.63790126559874
  
第35组：,得分：65.69244337641274
  
第二步获三等奖:
  
第76组：,得分：70.88876358460627
  
第168组：,得分：70.94210913179819
  
第100组：,得分：71.34028658163267
  
第178组：,得分：71.41491927586644
  
第25组：,得分：71.61226703961249
  
第132组：,得分：71.74036117133691
  
第129组：,得分：71.76405730756399
  
第196组：,得分：71.93407546343917
  
第166组：,得分：71.94227599338558
  
第102组：,得分：72.45888971091973
  
第三步的1-30名中，获二等奖为：
  
第177组：,得分：65.94837876883652
  
第149组：,得分：68.12771584230995
  
第199组：,得分：68.2765532178672
  
第181组：,得分：68.49516663790963
  
第192组：,得分：68.50956702434307
  
第三步的1-30名中，获一等奖为：
  
第89组：,得分：74.85983272152238
  
第98组：,得分：75.27160051095898
  
第10组：,得分：75.81652130817419
  
第148组：,得分：76.63950361246432
  
第83组：,得分：79.19906460795046
  
第173组：,得分：80.0918159773632
  
第61组：,得分：81.68436382430666
  
第165组：,得分：82.2034564324634
  
第153组：,得分：83.22179778055886
  
第183组：,得分：83.62431439653119
  
第5组：,得分：84.85177796748519
  
第15组：,得分：84.95712061638366
  
第134组：,得分：86.04712134629942
  
第151组：,得分：88.72264419945004
  
第69组：,得分：91.4062872542106
  
第四步中，获二等奖为：
  
第172组：,得分：48.60778837867472
  
第8组：,得分：48.87023624604781
  
第7组：,得分：48.88751971655283
  
第48组：,得分：48.90554960308781
  
第65组：,得分：50.76841193887731
  
第四步中，获一等奖为：
  
第161组：,得分：51.43085369879718
  
第101组：,得分：51.46202377844851
  
第139组：,得分：54.07602508809657
  
第163组：,得分：54.80787827584643
  
第147组：,得分：55.01918751227424