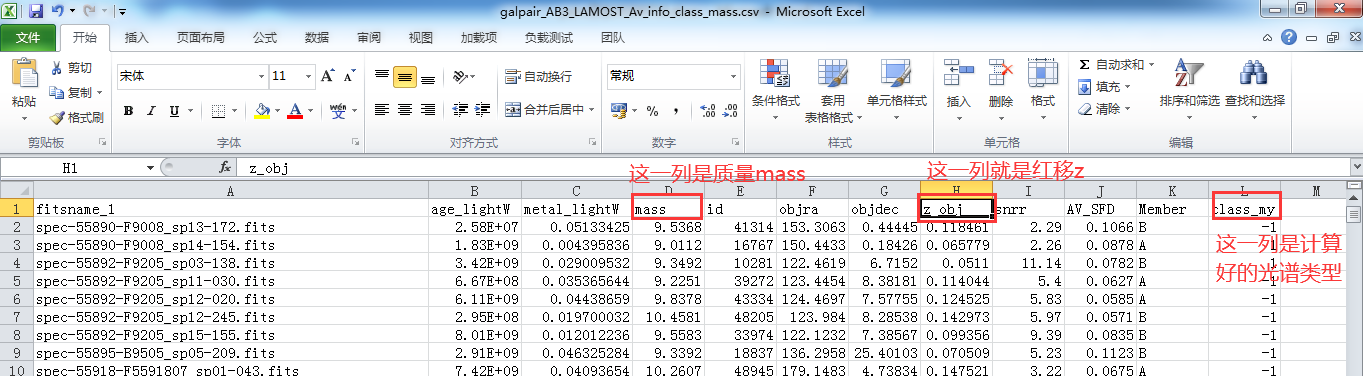
目标：给LAMOST星系样本找到control样本（1.control样本没有伴星系；2.每个成员与其control样本具有相同的红移(z)和质量分布。），每一个LAMOST星系只一个control样本即可。

一、给定一个数据表，共约2379个记录，其中红移列是z\_obj，质量列mass，星系光谱的分类列是class\_my.

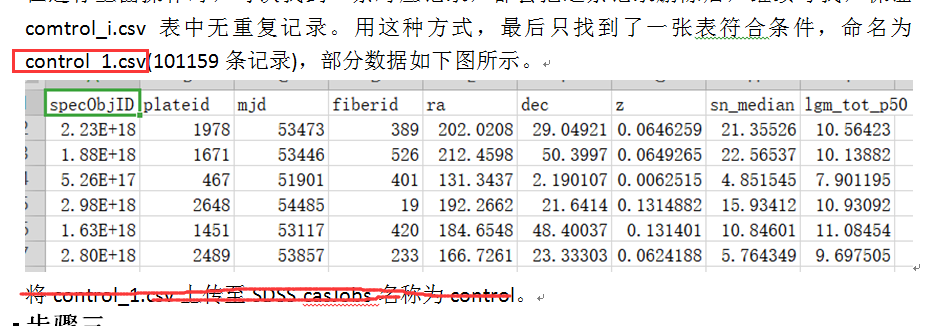


二、在星系对研究.docx文档中，从**步骤二**的第4步开始做（前面不用做了），但有的地方跟星系对研究.docx文档中的步骤不一样。

**步骤二**

4.从galpair\_AB3\_LAMOST\_Av\_info\_class\_mass.csv (后面称为G表)中取出每一行记录，在b1.csv(后面称为B表)中迭代搜索1个相同红移和质量的记录，形成control样本集。

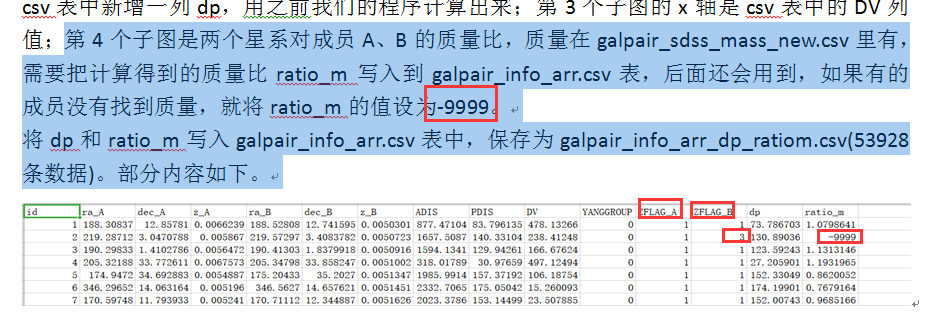
具体方法步骤和代码可参考文档里的说明，注意做到生成control\_1\_LAMOST.csv即可，不用做下图中的画红色删除线的步骤了。理论上这个表里的记录数应该和galpair\_AB3\_LAMOST\_Av\_info\_class\_mass.csv表中的记录数一样，即2379行记录。



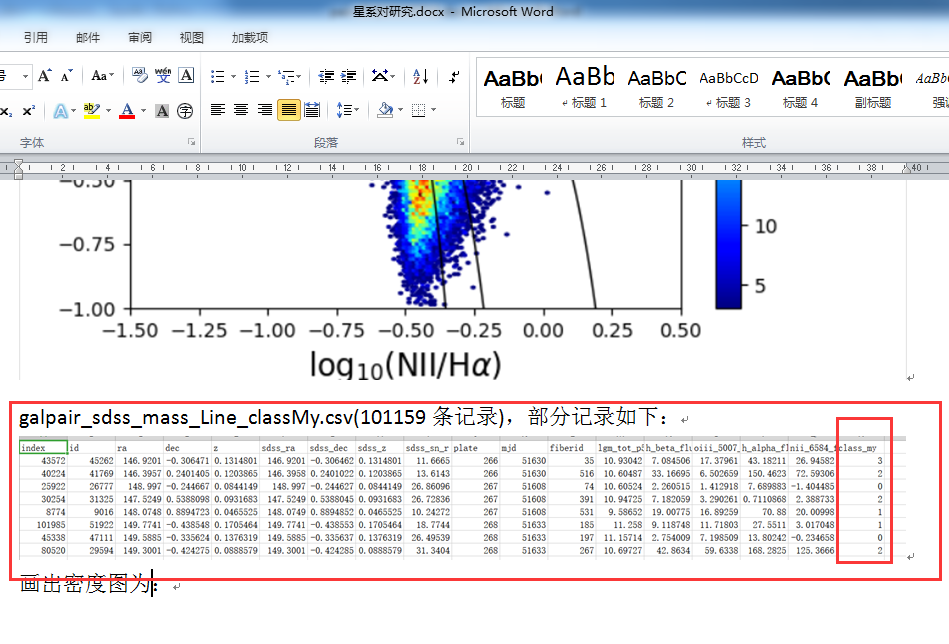
**步骤三**

1. 用星系对表galpair\_info\_arr.csv画如下图（详见星系对研究.docx文档中描述）

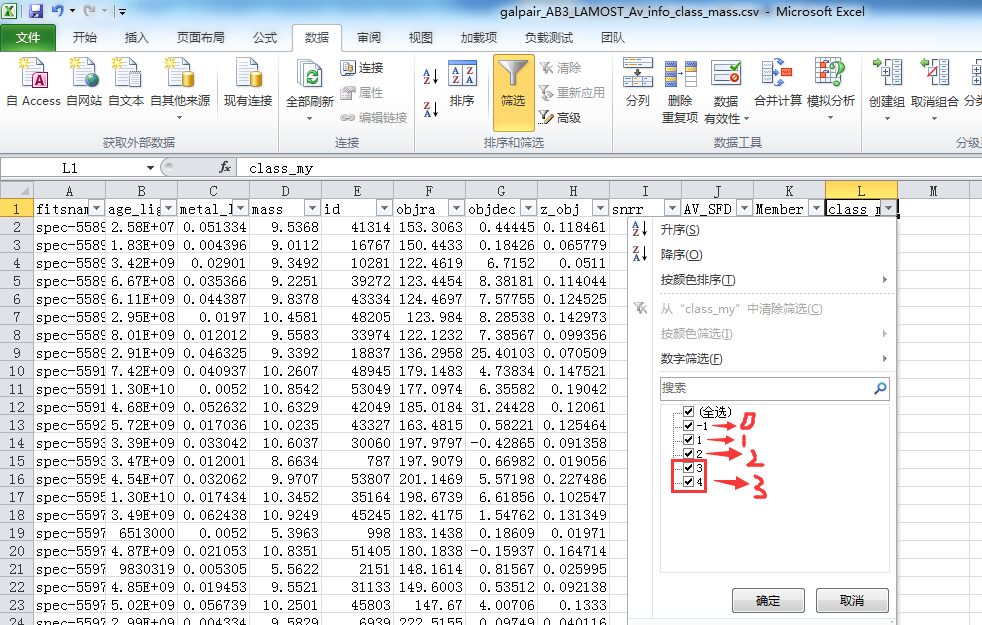
在这一步中的第4个子图的画法需要修改，下面描述中有时要把ratio\_m的值设为-9999，是因为有的成员没有找到质量，现在在galpair\_AB3\_LAMOST\_Av\_info\_class\_mass.csv这个表中有了2000多个星系有了质量了，可以补充之前李义做的工作，即在下面表中ZFLAG\_A或ZFLAG\_B这两列的值等于3时，就可以在galpair\_AB3\_LAMOST\_Av\_info\_class\_mass.csv表中找到相应星系的质量了，这样下面表中的两个成员A和B的质量比可以补充原来工作，但可能还有一些成员找不到质量，则将ratio\_m的值设为-9999。



2和3步骤是为了得到galpair\_sdss\_mass\_Line\_classMy.csv里的class\_my这一列，表示每个星系的类型（星系是分为-1，0，1，2，3五类），这两步对于本次任务不用再做了，因为galpair\_AB3\_LAMOST\_Av\_info\_class\_mass.csv这个表里有class\_my这一列，但是这个表中的有class\_my这一列有-1，1，2，3，4五个值，与上述表中的有class\_my的值不一致，需要这样修改来对应：



在galpair\_AB3\_LAMOST\_Av\_info\_class\_mass.csv表里有class\_my=-1时对应galpair\_sdss\_mass\_Line\_classMy.csv表里class\_my=0，这样可以把galpair\_AB3\_LAMOST\_Av\_info\_class\_mass.csv表里有class\_my=-1修改为class\_my=0，其它对应下图进行修改。



**步骤四**

前面步骤都是为这一步做准备的，本次任务前面步骤中对2379条LAMOST星系寻找对应的control样本、处理星系的类型，这2000多条光谱是对李义学前期工作的补充，所以在**步骤四**中需要把这2000多条光谱得到的control样本和星系类型等信息合并到前期李义的相应表里，然后再按照星系对研究.docx文档的描述执行**步骤四**。

根据查看 galpair\_sdss\_mass\_Line\_classMy.csv ， control\_Line\_classMy.csv 可以添加galpair\_index两个表联系起来