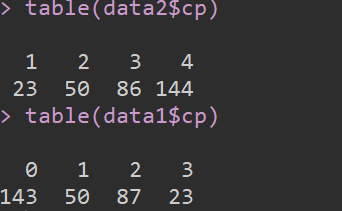
Variable Description

我们从kaggle得到的数据和原始uci的数据部分变量有较大的差异，但是它们使用的是同样的变量说明，这明显是不合理的。因此我们需要通过对比两个数据库，得到正确的变量解释。以下说明中data1为kaggle，data2为uci。

cp：chest pain type

原变量解释

-- Value 1: typical angina  
-- Value 2: atypical angina  
-- Value 3: non-anginal pain  
-- Value 4: asymptomatic



所以在kaggle数据集中

Cp:

0 asymptomatic 无症状

1 atypical angina 不典型心绞痛

2 non-anginal pain 非心绞痛疼痛

3 typical angina 典型心绞痛

Trestbps：静息血压

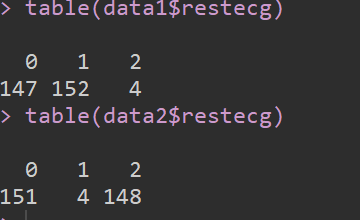
Chol：胆固醇

Fbs： 空腹血糖 >120mg 为1

Restecg: resting electrocardiographic results 静息心电图结果

原变量解释

19 restecg: resting electrocardiographic results  
-- Value 0: normal 正常  
-- Value 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV)  
-- Value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria 左心室肥大



所以在kaggle数据集中：

Restecg

0 left

1 normal

2 abnormaility

Thalach 最大心跳数

Exang 运动时是否有心绞痛 1为有

Oldpeak 运动相对于休息的ST depression

slope: the slope of the peak exercise ST segment

原变量解释：

 slope: the slope of the peak exercise ST segment  
-- Value 1: upsloping  
-- Value 2: flat  
-- Value 3: downsloping

|  |
| --- |
| > table(data1$slope)  0 1 2  21 140 142  > table(data2$slope)  1 2 3  142 140 21 |
|  |
| |  | | --- | | > | |
|  |

所以在kaggle数据集中

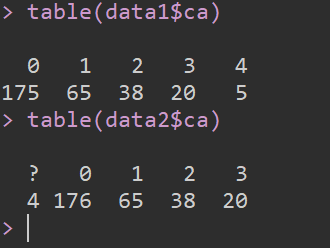
0: downsloping; 1: flat; 2: upsloping

注释：

运动耐力测试

在运动过程中，J点（S波和ST段交界处的拐点）变得压低，在峰值运动时压低最大。因此，运动期间正常的ST段会急剧向上倾斜，所以斜率为负为不正常

Ca：被透视荧光检查标注颜色的大血管的数量

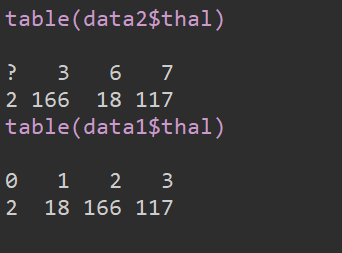


可以看到原数据集中缺失值数量为4，在kaggle中直接标记为属于标签4，这是不合理的，建议去除这四个观测。

thal: 钴示踪检测

原数据集中：

thal: 3 = normal; 6 = fixed defect; 7 = reversable defect



所以，在kaggle数据集中

1 = fixed defect; 2 = normal; 3 = reversable defect，0为缺失值 应该删除

当心肌部分在静止时摄取核示踪剂时检测到冠状动脉狭窄，但在心脏受压时则没有。这称为“可逆缺陷”。先前梗塞的疤痕心肌不会吸收示踪剂，被称为“固定缺陷”。

理论上应该3更严重?

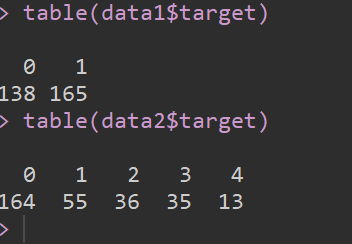
Target

原数据集中:

The "goal" field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4

在两个数据集中都没有找到field字段，根据观察应该被改为了target字段

根据上述说明，在nci数据集中，0为无心脏病



所以，在kaggle数据集中，

0为有心脏病，1为无心脏病。 0 无心脏病 1 有心脏病

165 138