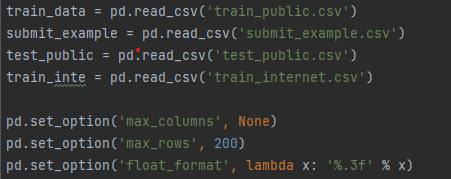
程序说明

对于这个程序，我们通过pandas对于训练数据进行读取，



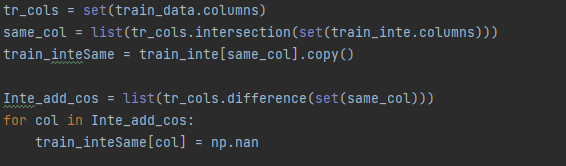
再设置训练代码块train\_model,展示代码块display\_importances以及数据处理单元，对数据进行分析处理。



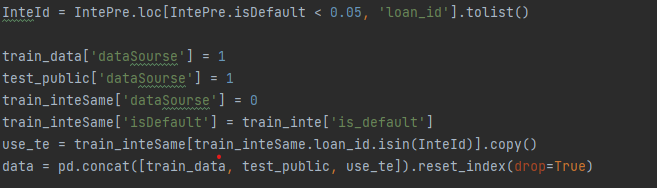
数据预处理，并对internet表进行统一处理，构建时间特征年和月,同时internet表也构建相同特征



获取train\_public表与internet表相同字段的internet数据，然后添加缺少的特征并赋值为空值



选择阈值0.05，从internet表中提取预测小于该概率的样本，并对不同来源的样本赋予来源值



最终得出结果

