二、科比投篮数据分析，可以使用pandas工具包

1. 查看数据集的样本数及特征数，特征的数据类型，以及是否存在缺失值
2. 统计科比每年的进球数与未进球数，利用条形图可视化，并将结果以文件名“科比1.csv”保存；指出哪年是科比进球最多的年
3. 统计1996-2000年、2001-2005年、2006-2010年、2011-2016年所进球数占总进球数的比例，并将结果以文件名“科比2.csv”保存
4. 利用AdaBoost，KNN，CART，Naïve Bayes，Random forest在训练集（有标签的数据集）上进行3折交叉验证，并基于AUC来比较不同模型的预测效果
5. 对于训练集中的负样本（即未进球），预测哪些负样本最有可能接近进球（即差点命中），并指出前10个