TMDB数据预测电影票房收入

1、数据导入

import pandas as pd

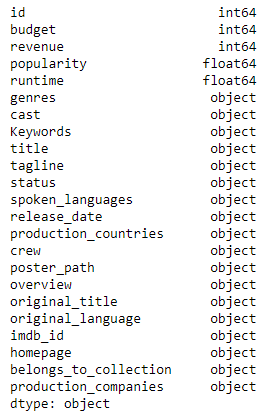
data = pd.read\_csv('input/train.csv')

test\_data = pd.read\_csv('input/test.csv')

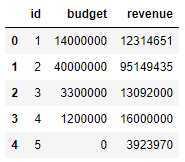
sample = pd.read\_csv('input/sample\_submission.csv')

2、数据类型

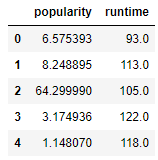
data.dtypes.sort\_values()



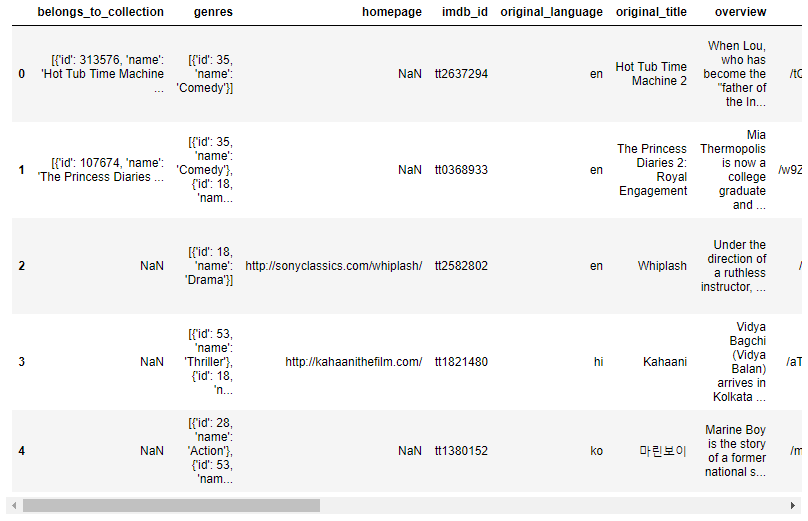
data.select\_dtypes(include='int64').head()



data.select\_dtypes(include='float64').head()



data.select\_dtypes(include='object').head()



3、缺失数据处理

4、数据探索

5、特征工程

6、训练数据准备

7、模型建立、训练、误差检验

8、提交结果