代码部分

import pandas as pd  
  
# 读取CSV文件  
data = pd.read\_csv('C:/Users/48844/Desktop/B题-企业财务数据分析与造假识别/数据/financial\_data.csv')  
  
# 计算相关性  
correlation = data.corr(method="pearson").abs()  
  
# 选择与利润总额相关性最高的5个指标  
top\_5\_correlated = correlation['LRZE'].nlargest(6)[0:]  
  
print(top\_5\_correlated)

pearson方法计算结果：

YYSR 0.782726

YWFY 0.772832

YYCB 0.737736

YYSJJFJ 0.565440

ZCJZSS 0.238524

spearman方法计算结果:

YYSR 0.609192

YYSJJFJ 0.608179

YWFY 0.570395

YYCB 0.510354

ZCBCL 0.475194

利润总额与各项指标之间的关系应更趋向于线性相关，但是pearson相关系数计算公式要求适用场景是呈正态分布的连续变量。显然题目中所给数据并不符合，而spearman相关系数计算公式并没有pearson相关系数计算公式要求严格，而且计算结果更为均匀合理，所以我们采用spearman相关系数计算公式来进行计算。

由于不清楚是正线性相关，负线性相关，所以我们在代码中加入abs函数进行就绝对值计算再进行比较其相关程度，最终得出五个相关程度最高的指标。