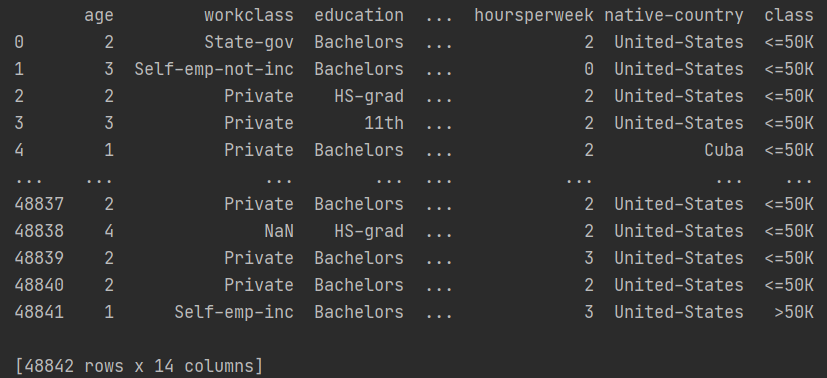
Adult.py

1. read\_csv\_1:

输入：adult.csv

输出：



1. num\_rows

输入：初始dataframe

输出：

1. column\_names:

输入：初始dataframe

输出：

['age', 'workclass', 'education', 'education-num', 'marital-status', 'occupation', 'relationship', 'race', 'sex', 'capitalgain', 'capitalloss', 'hoursperweek', 'native-country', 'class']

1. Missing\_values

输入：初始dataframe

输出：6465

1. Columns\_with\_missing\_values

输入：初始dataframe

输出：['workclass', 'occupation', 'native-country']

1. Bachelors\_masters\_percentage

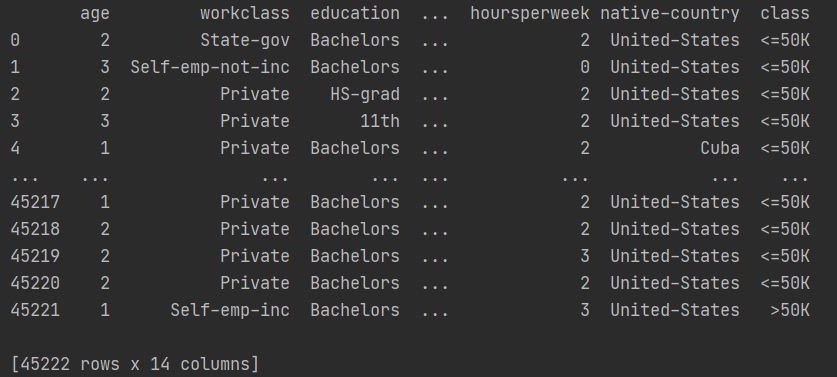
输入：初始dataframe

输出：21.871

1. data\_frame\_without\_missing\_values

输入：初始dataframe

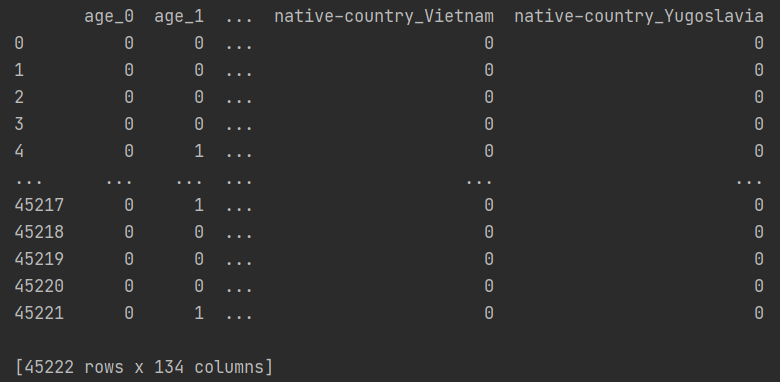
输出：



1. one\_hot\_encoding

输入：初始dataframe

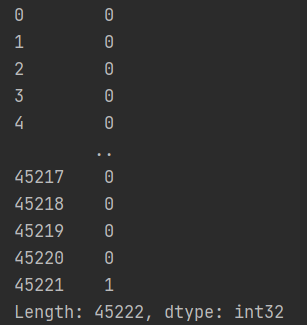
输出：



1. Label\_encoding

输入：初始dataframe

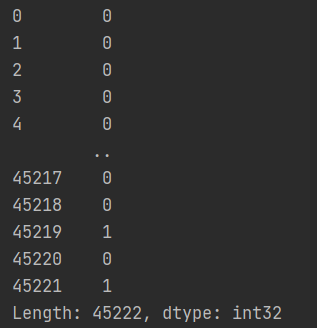
输出：



1. dt\_predict

输入：one hot和label的结果

输出：



1. dt\_error\_rate

输入：label和predict的结果

输出：0.08279156162929548