    pandas的DataFrame，有时需要处理一些字符串类型列，运用Series.str列内置方法很方便。

1.one hot 独热编码,get\_dummies

series=data['列名'].str.get\_dummies(sep=',')

实现DataFrame中列有多值，且想把这列one hot下

2.切分字符串，split()

series=data['列名'].str.split(',')

把DataFrame列中字符串以','分隔开，每个元素分开后存入一个列表里

series=data['列名'].str.split(',',expand=True)

参数expand，这个参数取True时，会把切割出来的内容当做一列，产生多列。

series=data['列名'].str.split(',',expand=True)[0]

可以只要第一列。

3.替换，replace()

series=data['列名'].str.replace(',','-')

用‘-’代替‘，’

4.是否包含表达式，contains()

series=data['列名'].str.contains('we')

返回的是布尔值series

5.查找所有符合正则表达式的字符findall()

series=data['列名'].str.findall("[a-z]")

以数组的形式返回

6.计算字符串的长度，len()

series=data['列名'].str.len()

7.去除前后的空白字符，strip()

series=data['列名'].str.strip()

rstrip() 去除后面的空白字符

lstrip() 去除前面的空白字符

8.isalnum() 是否全部是数字和字母组成

isalpha() 是否全部是字母

isdigit() 是否全部都是数字

isspace() 是否空格

islower() 是否全部小写

isupper() 是否全部大写

istitle() 是否只有首字母为大写，其他字母为小写

--------------------- 本文来自 月耀\_717 的CSDN 博客 ，全文地址请点击：https://blog.csdn.net/yueyao121107/article/details/79730934?utm\_source=copy