

## pandas 常用api

笔记本: 深度学习

创建时间: 2019/10/2 15:42

更新时间: 2019/10/2 16:25

作者: beyourselfwb@163.com

---

`import pandas as pd`

打开 `pd.read_csv('a.csv')` or `pd.read_csv('a.csv', names=['q1', 'q2', 'label', 'category'])`

保存 `pd.to_csv('b.csv', index=False, header=None)`

选取几列 `df = df[['q1', 'q2']]`

重命名 `df =`

`df.rename(columns={'q1_new': 'q1', 'q2_new': 'q1'})`

按照某列排序 `df =`

`df.sort_values(by='id')` 默认升序

两个表concat `df = pd.concat([df1,`

`df2])` 前提: 两个表是一样的

两个表根据id来merge `df = pd.merge(df1, df2, on='id')`

根据条件筛选 `df =`

`df.loc[(df['label']==0) & (df['category']=='apple')]`

根据切片取指定行列 `df = df.iloc[:10]`

随机打散 `df =`

`df.sample(frac=1).reset_index(drop=True)`

直接对列进行修改      `df['q1'] =`  
`df.q1.apply(lambda x: x.replace(',', ' '))`  
根据多列进行修改      `df['delta'] =`  
`df.apply(lambda x: x['len1']-x['len2'], axis=1)`

查看某列的类别分布

`df.category.value_counts()`

获取所有类别      `labels = [_label for`  
`_label, count in df.label.value_counts().items()]`

获取df的列名列表      `columns = [i for i in`  
`df.columns]`