

pandas数据加载

Text
CSV
Excel
HTML
MySQL
MongoDB

Text

```
In [4]: import pandas as pd
```

```
In [5]: pd.read_table('data/01.txt')
```

Out[5]:

	email
0	jack@csdn.com
1	lily@qq.com
2	marry@163.com

```
In [6]: pd.read_table('data/02.txt')
```

Out[6]:

	Name	Company	Mail
0	Jack	CSDN	jack@csdn.net
1	Lily	tencent	lily@qq.com
2	Marry	wangyi	marry@163.com

```
In [7]: pd.read_table('data/03.txt', sep=':')
```

Out[7]:

	jack	jack@csdn.com
0	lily	lily@qq.com
1	marry	marry@163.com

```
In [8]: pd.read_table('data/03.txt', sep=':', header=None)
```

Out[8]:

	0	1
0	jack	jack@csdn.com
1	lily	lily@qq.com
2	marry	marry@163.com

```
In [9]: pd.read_table('data/03.txt', sep=':', names=['name','email'])
```

Out[9]:

	name	email
0	jack	jack@csdn.com
1	lily	lily@qq.com
2	marry	marry@163.com

CSV

```
In [10]: pd.read_csv('data/04.csv')
```

Out[10]:

	Name	Phone	Addr
0	jack	15964522054	Beijing, China
1	lily	17934243244	Shanghai, China
2	marry	15354354567	Shenzhen, China

Excel

```
In [12]: pd.read_excel('data/05.xlsx')
```

Out[12]:

	Name	Phone	Addr
0	jack	15964522054	Beijing, China
1	lily	17934243244	Shanghai, China
2	marry	15354354567	Shenzhen, China

在使用上面这条语句的时候，需要先安装 lxrd 模块: `pip3 install xlrd`

HTML

```
In [14]: tables = pd.read_html('data/06.html', header=0)    # 其中的 header=0指用第一行做为head头
```

```
In [15]: tables[0]
```

```
Out[15]:
```

	ID	名称	IP	CPU	内存	硬盘
0	1	tencent	202.54.21.124	2	1	50
1	2	nfp	204.44.68.49	1	1	20
2	3	DigitalOcean	159.65.130.172	2	2	25

```
In [16]: len(tables)
```

```
Out[16]: 1
```

使用 read_html() 前需要安装 lxml 模块: `pip3 install lxml`

MySQL

使用之前需要安装 pymysql

```
In [17]: import pymysql
```

```
In [19]: conn = pymysql.connect(host='127.0.0.1', user='root', password='xxx  
xxx', db='auto', charset='utf8')  #这里charset不要写utf-8, 而是utf8
```

```
In [20]: sql = 'select * from monitor'  
data = pd.read_sql(sql, conn)  
conn.close()
```

```
In [21]: data
```

```
Out[21]:
```

	id	host_id	cpu_used	vir_mem_used	mem_used	addtime	time
0	6	3	10.2	29.7	29.7	2018-08-05 02:31:00	10:31
1	7	3	10.0	29.8	29.8	2018-08-05 02:31:40	10:31
2	8	3	2.0	29.7	29.7	2018-08-05 02:34:23	10:34
3	9	3	3.9	30.3	30.3	2018-08-05 02:37:06	10:37
4	10	3	2.0	28.6	28.6	2018-08-05 02:43:47	10:43
5	11	3	8.0	27.2	27.2	2018-08-05 02:44:34	10:44
6	12	3	9.8	27.3	27.3	2018-08-05 02:48:46	10:48
7	13	3	4.0	28.2	28.2	2018-08-05 02:50:24	10:50
8	14	3	2.0	27.7	27.7	2018-08-05 02:51:20	10:51
9	15	3	3.9	27.7	27.7	2018-08-05 02:52:01	10:52
10	16	3	8.0	28.3	28.3	2018-08-05 02:52:13	10:52
11	17	3	2.0	27.6	27.6	2018-08-05 02:53:43	10:53
12	18	3	14.3	27.8	27.8	2018-08-05 02:54:51	10:54
13	19	3	2.0	28.2	28.2	2018-08-05 02:56:16	10:56
14	20	3	12.0	28.9	28.9	2018-08-05 02:56:31	10:56
15	21	3	2.0	29.9	29.9	2018-08-05 03:01:46	11:1
16	22	3	11.8	29.8	29.8	2018-08-05 03:02:18	11:2
17	23	3	8.0	30.3	30.3	2018-08-05 03:03:40	11:3
18	24	3	2.0	29.9	29.9	2018-08-05 03:03:42	11:3
19	25	3	11.8	30.5	30.5	2018-08-05 03:07:10	11:7
20	26	3	2.0	30.6	30.6	2018-08-05 04:00:13	12:0

MongoDB

使用之前需要安装 pymongo

```
In [23]: import pymongo
```

```
In [24]: client = pymongo.MongoClient('127.0.0.1', 27017) # mongodb的默认端口号为 27017
```

```
In [25]: client
```

```
Out[25]: MongoClient(host=['127.0.0.1:27017'], document_class=dict, tz_aware=False, connect=True)
```

```
In [26]: db = 'local'
table = 'startup_log'
cursor = client[db][table].find()
```

```
In [27]: data = pd.DataFrame(list(cursor))
```

```
In [28]: data
```

```
Out[28]:
```

	_id	buildinfo	cmdLine	hostname
0	ziyichen-1528717249925	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen
1	ziyichen.local-1528802746726	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
2	ziyichen.local-1528813813064	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
3	ziyichen.local-1529591554637	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
4	ziyichen.local-1529596200879	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
5	ziyichendeMac.local-1529933646586	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichendeMac.l
6	ziyichendeMac.local-1530701229819	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichendeMac.l
7	ziyichendeMac.local-	{'version': '3.6.5',	{'config': '/usr/local/etc/mongod.conf',	ziyichendeMac.l

	1530806618388	'gitVersion': 'a20ecd3e3a...	'net'...	
8	ziyichen.local- 1530850909033	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
9	ziyichen.local- 1530887751415	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
10	ziyichen.local- 1531239840724	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
11	ziyichen.local- 1532019012709	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
12	ziyichen.local- 1532085990421	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen.local
13	ziyichen- 1532847648645	{'version': '3.6.5', 'gitVersion': 'a20ecd3e3a...	{'config': '/usr/local/etc/mongod.conf', 'net'...	ziyichen