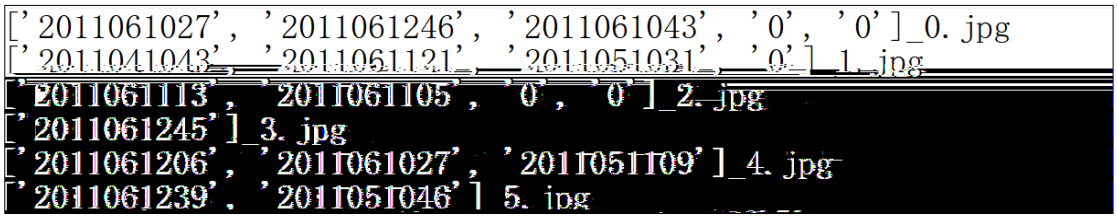
.....


['2011061027', '2011061246', '2011061043', '0', '0']_0.jpg
['2011041043', '2011061121', '2011051031', '0', '0']_1.jpg
['2011061113', '2011061105', '0', '0']_2.jpg
['2011061245']_3.jpg
['2011061206', '2011061027', '2011051109']_4.jpg
['2011061239', '2011051046']_5.jpg

E B

E B

E B J

DateTime	NY - grid	NY - EV	NY - solar	Austin - grid	Austin - EV	Austin - solar	Boulder - grid	Boulder - EV	Boulder - solar
2019-03-01 00:00:00	0.295	0.003	0.013999999999999999	0.76	0.0	0.0	0.62	0.004	0.0
0.7509999999999999	2019-03-01 00:15:00	0.46299999999999997	0.003	0.013999999999999999	0.0	0.77	0.004	0.0	0.789
2019-03-01 00:30:00	0.366	0.003	0.013999999999999999	0.36	0.0	0.004			

J

J J

E B

B B

B

B

E

E

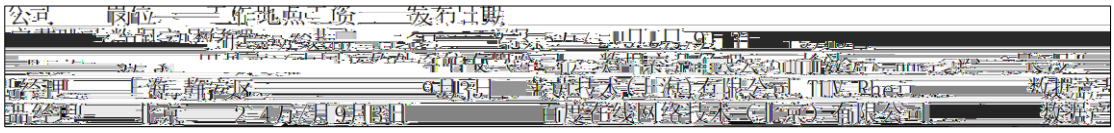
E

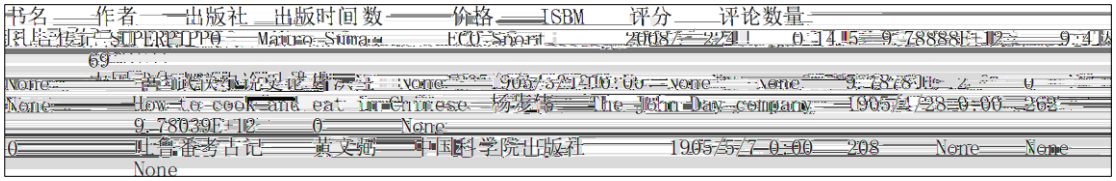
E

			E	

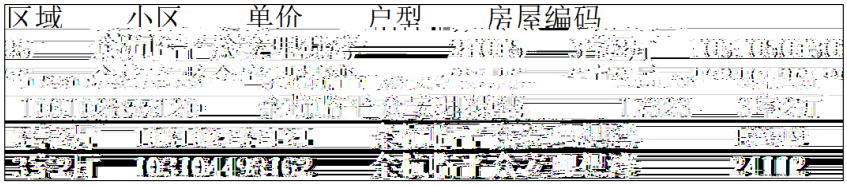
length	width
3	4
5	6
7	8
14	16
5	9
15	20
6	8

country	beer_servings	spirit_servings	wine_servings	total_litres_of_pure_alcohol	continent
Afghanistan	0	0	0	0	AS
Albania	89	132	54	4.9	EU
Algeria	25	0	14	0.7	AF
Andorra	245	138	312	12.4	EU




书名 作者 出版社 出版时间 数 价格 ISBM 评分 评论数量
None 吉里吉里2完全编程指南 None 2007-5-1 0.0 None None 5.828e-12 0
None How to cook and eat in Chinese 杨步伟 The John Day company 1905-12-8 0.00 263
9.78039E+12 0 None
0 北京考古记 黄文弼 中国科学院出版社 1905-7-1 0.00 208 None None
None

学校名称	录取批次
湘潭大学	本科一批
吉首大学	本科一批
湖南大学	本科一批
中南大学	本科一批


区域
小区
单价
户型
房屋编码

hadoop /usr/local

core-site.xml

/var/log/hadoop/tmp hadoop.tmp.dir value

hdfs://master:8020 fs.defaultFS value

hadoop-env.sh JAVA_HOME

cp mapred-site.xml.template mapred-site.xml

yarn mapreduce.framework.name value

yarn-site.xml ,

master yarn.resourcemanager.hostname value;

mapreduce_shuffle yarn.nodemanager.aux-services value

hdfs-site.xml

2 dfs.replication value

slaves slave1 slave2

iii. scp Hadoop

Hadoop

iv. /etc/profile JAVA_HOME HADOOP_HOME PATH source

v. NameNode

vi. Hadoop jps

vii.

i

ii

	I		I	
	J J			

Linux initnetwork

Linux wget -P /course/Hive/data
http://datasrc.tipdm.net:81/Hunan/module2/2-3/business_circle.csv
business_circle.csv Linux /course/Hive/data
hive Hadoop MySQL
hive hive Linux
$HADOOP_HOME/sbin/start-all.sh hadoop
hadoop MySQL hive hive
hive

Hive

I

Linux initnetwork

Linux wget -P /course/Hive/data http://datasrc.tipdm.net:81/Hunan/module2/2-4/business_circle.csv business_circle.csv Linux /course/Hive/data

hive Hadoop MySQL

hive hive Linux

$HADOOP_HOME/sbin/start-all.sh hadoop

hadoop MySQL hive hive

hive

business_circle.csv

Hive

Linux initnetwork

Linux wget -P /course/Hive/data http://datasrc.tipdm.net:81/Hunan/module2/2-5/business_circle.csv business_circle.csv Linux /course/Hive/data

hive Hadoop MySQL

hive hive Linux $HADOOP_HOME/sbin/start-all.sh hadoop hadoop MySQL hive hive hive

business_circle.csv

Hive

I

B

		B		
	BJ	B J		
	I J B	J E B		

http://datasrc.tipdm.net:81/Hunan/module2/2-7/student.csv student.csv
Linux /course/Hive/data
hive Hadoop MySQL
hive hive Linux
$HADOOP_HOME/sbin/start-all.sh hadoop
hadoop MySQL hive hive
hive

			BJ

Hive

I

Linux initnetwork

Linux wget -P /course/Hive/data http://datasrc.tipdm.net:81/Hunan/module2/2-8/emp.csv emp.csv Linux /course/Hive/data

hive Hadoop MySQL

hive hive Linux $HADOOP_HOME/sbin/start-all.sh hadoop hadoop MySQL hive hive hive

		B		
	BJ	B J		
	I J B	J E B		
	J	B J		
	B	J E B		

	J E	B BJ		

Hive

iii.

iv.

v.

vi.

I

Linux initnetwork

Linux wget -P /course/Hive/data http://datasrc.tipdm.net:81/Hunan/module2/2-9/emp.csv emp.csv Linux /course/Hive/data

hive Hadoop MySQL

hive hive Linux $HADOOP_HOME/sbin/start-all.sh hadoop hadoop MySQL hive hive hive

		B		
	BJ	B J		
	I J B	J E B		
	J	B J		
	B	J E B		
	J E	B BJ		

Hive

I

Linux initnetwork

Linux wget -P /course/Hive/data http://datasrc.tipdm.net:81/Hunan/module2/2-11/business_circle.csv business_circle.csv Linux /course/Hive/data

hive Hadoop MySQL hive hive Linux $HADOOP_HOME/sbin/start-all.sh hadoop hadoop MySQL hive hive hive

business_circle.csv

Hive

I

I

导航条
首页
产品中心
实验室建设
培训认证
校企合作
新闻中心
关于我们
大数据产品
人工智能产品
在线教育
教学资源
开源平台
其他
saas

etree

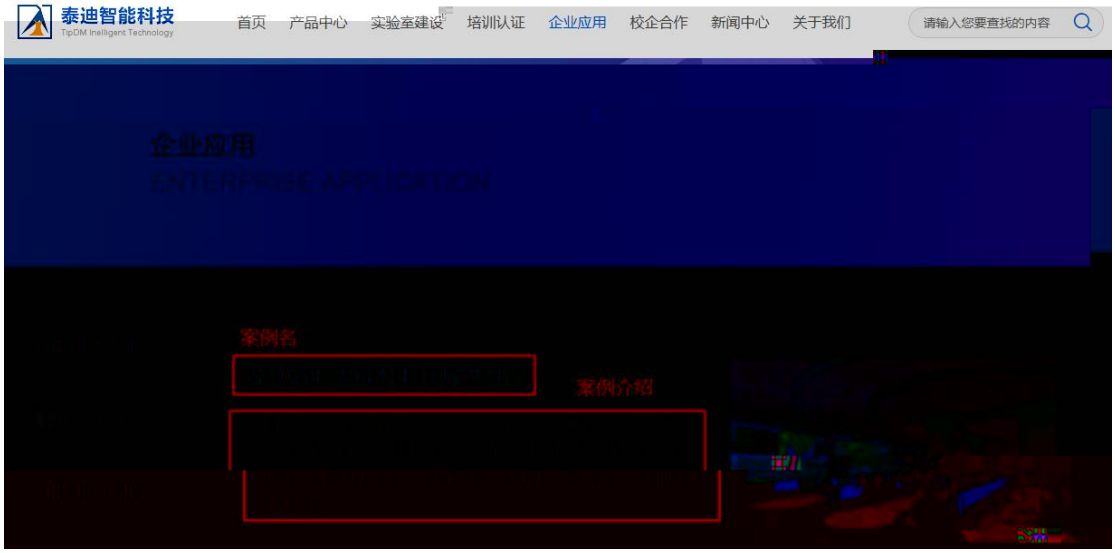
HTML

BeautifulSoup

BeautifulSoup BeautifulSoup

导航条
首页
产品中心
实验室建设
校企合作
新闻中心
关于我们
大数据产品
人工智能产品
在线教育
教学资源
开源平台
其他
saas

;



urls

4 urls

http://tipdm.com/qyyy/index_1.jhtml

requests BeautifulSoup

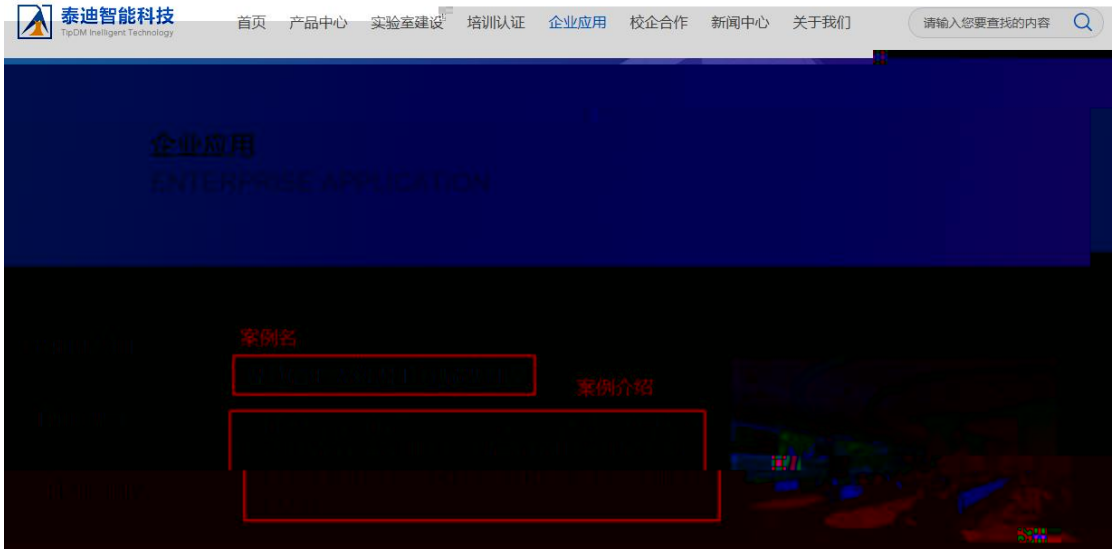
;

urls

requests

BeautifulSoup BeautifulSoup

BeautifulSoup



urls

4 urls

http://tipdm.com/qyyy/index_1.jhtml

urllib3 etree

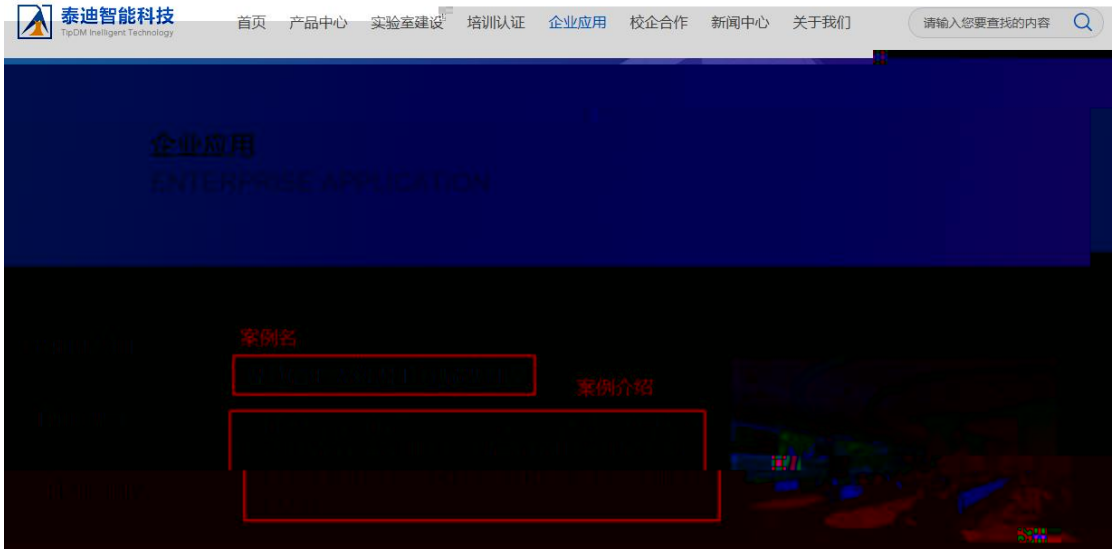
;

urls

urllib3

etree.HTML get html

html



urls

4 urls

http://tipdm.com/qyyy/index_1.jhtml

requests re

;

urls

requests

'<h1>( )</h1>'

' ( ) '

urls

4 urls

http://tipdm.com/xwzx/index_1.jhtml

urllib3 etree

;

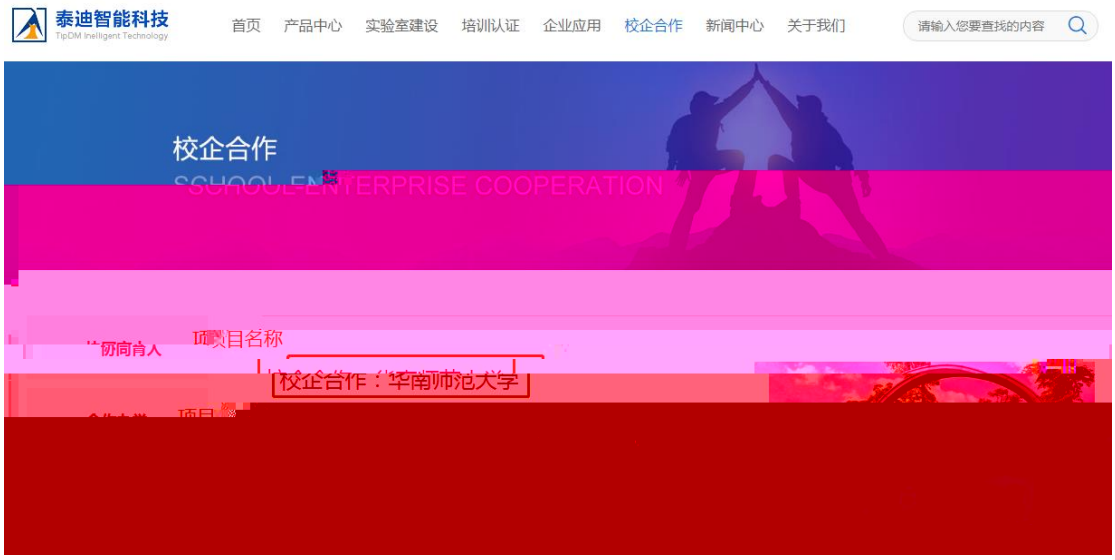
urls

urllib3

etree.HTML html

html

data



urls

4 urls

http://tipdm.com/xqhz/index_1.jhtml

requests BeautifulSoup

;

urls

url

BeautifulSoup BeautifulSoup

BeautifulSoup

data

urls

4 urls

http://tipdm.com/xqhz/index_1.jhtml

requests re

;

urls

url requests

'target="_blank">( )</h1'

' ( ) '

x

x

y

5

x
x
y

5

5

B

2017 6 5

x
x
y
zip

2017 6 5

B

(ii) “ ” “ ”

' '

(iii) “ ” “ ” “ ”

x
x
y

' '

(iii)

;

x
x
y

表 4-2 HR_comma_sep.csv 字段说明

' '

x
x
y
' '

表 4-3 movies.csv 字段说明

10

x
x
y
for

10

10

;

			I	

表 4-4 mum_baby_trade_history.csv 字段说明

;

;

x
x
y

(i) pandas matplotlib.pyplot plt.rcParams['font.sans-serif'] = 'WEnQuanYi Zen Hei' plt.rcParams['font.sans-serif'] = 'SimHei'

;

x

x

y

10

x

x

y

x

x

y

2000 data.iloc[0,5:]

2000

data.iloc[68, 2:5]

x
x
y

x

y