requests库

1. requests库的get()用法

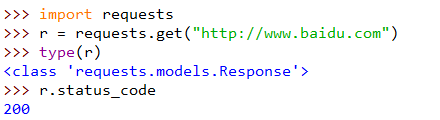
R = requests.get(url)

右边: 构造一个向服务器请求资源的**Requests**对象

左边：返回一个包含服务器资源的**Response**对象

**Response**对象包含爬虫返回的内容

1.1 **Response**对象

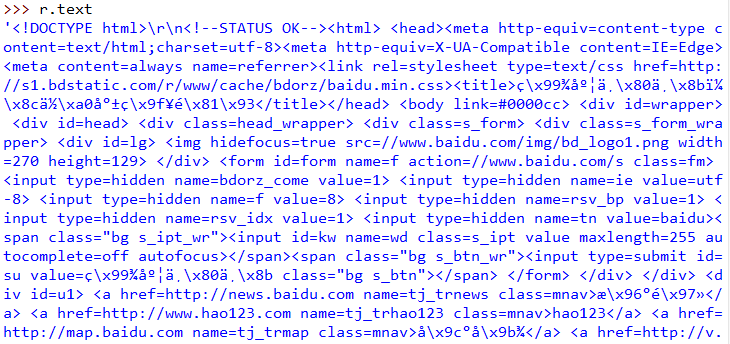


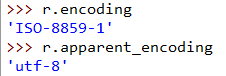
这里的r即为Response对象（参见type(r)）

status\_code为Response对象的一个属性，表示HTTP请求的返回状态，200表示连接成功，404表示失败

1.1.1 **Response**对象的属性

|  |  |
| --- | --- |
| 属性 | 说明 |
| r.status\_code | HTTP请求的返回状态，200表示连接成功，404表示失败 |
| r.text | HTTP相应内容的字符串形式，即url对应的页面内容 |
| r.encoding | 从HTTP header中猜测的响应内容编码方式 |
| r.apparent\_coding | 从内容中分析出的响应内容编码方式 |
| r.content | HTTP响应内容的二进制形式 |





注意：r.encoding只是从header中猜测响应内容的编码方式，r.text根据r.encoding显示网页内容，则此时可能出现不了中文

而r.apparent\_coding是根据网页内容分析出编码方式，所以将r.encoding改为’utf-8’后，可以显示出中文

