

# requests

## 1.基本使用

### 1. 文档:

官方文档

[http://cn.python-requests.org/zh\\_CN/latest/](http://cn.python-requests.org/zh_CN/latest/)

快速上手

[http://cn.python-requests.org/zh\\_CN/latest/user/quickstart.html](http://cn.python-requests.org/zh_CN/latest/user/quickstart.html)

### 2. 安装

```
pip install requests
```

### 3. response的属性以及类型

类型	: models.Response
r.text	: 获取网站源码
r.encoding	: 访问或定制编码方式
r.url	: 获取请求的url
r.content	: 响应的字节类型
r.status_code	: 响应的状态码
r.headers	: 响应的头信息

## 2.get请求

```
requests.get()
eg:
import requests
url = 'http://www.baidu.com/s?'
headers = {
    'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML,
        like Gecko) Chrome/65.0.3325.181 Safari/537.36'
}
data = {
    'wd': '北京'
}
response = requests.get(url, params=data, headers=headers)
```

定制参数

参数使用params传递

参数无需urlencode编码

不需要请求对象的定制

请求资源路径中? 可加可不加

## 3.post请求

```
requests.post()
百度翻译:
eg:
import requests
post_url = 'http://fanyi.baidu.com/sug'
headers={
    'User-Agent': 'Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36
    (KHTML, like Gecko) Chrome/68.0.3440.106 Safari/537.36'
}
data = {
    'kw': 'eye'
}
r = requests.post(url = post_url,headers=headers,data=data)
```

6: get和post区别?

- 1: get请求的参数名字是params post请求的参数名字是data
- 2: 请求资源路径后面可以不加?
- 3: 不需要手动编解码
- 4: 不需要做请求对象的定制

## 4.代理

7: proxy定制

在请求中设置proxies参数

参数类型是一个字典类型

```
eg:
import requests
url = 'http://www.baidu.com/s?'
headers = {
    'user-agent': 'Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML,
    like Gecko) Chrome/65.0.3325.181 Safari/537.36'
}
data = {
    'wd': 'ip'
}
proxy = {
    'http': '219.149.59.250:9797'
}
r = requests.get(url=url,params=data,headers=headers,proxies=proxy)
with open('proxy.html','w',encoding='utf-8') as fp:
    fp.write(r.text)
```

## 5.cookie定制

8: cookie定制

应用案例:

- (1) 古诗文网 (需要验证)
- (2) 云打码平台
  - 用户登陆    actionuser    action
  - 开发者登陆    actioncode    action

作业：国家统计局 (<http://www.stats.gov.cn/tjsj/tjbz/tjyqhdmhcxhfdm/2017/>) 共计68万条数