为什么反爬（了解）

服务器如何知道你是爬虫

通过请求的信息识别

1. UA
2. Referer：告诉服务器当前这个请求来自于哪里（一般是上一次浏览的网页跳转）
3. Cookie
   1. 登录
   2. 验证码

特定请求参数

1.Sign（比如百度翻译，现在要sign，你要翻译的东西经过当前页面的js算法得到sign=xxx，如何把要翻译的和sign发给服务器，服务器也会按照同js类似的算法算出一个值，如果两个值不等，百度服务器认为你是爬虫）

2.Token（令牌的意思，同样的）