

# 环境搭建

- 下载安装[anaconda-windows](#)  
如果网速慢可以使用[清华源](#)  
修改配置
- 配置环境变量(如果需要的化), 打开命令行, 输入python或者python3,没有报错就成功了  
如果报错了, 就把anaconda 安装目录加到windows的环境变量里面
- 下载安装[pycharm community](#)

## pycharm 打开项目

使用pycharm打开文件, 看一下pycharm上面提示, 选择安装就可以, 运行文件就可以抓数据, 最后数据保存在当前目录

抓取地址:

```
def fetch_data(page, key_word):  
    url = 'https://odin.re.taobao.com/m/Nwalltbuad?sbid=sem2_kgb_activity&ignore=CATID%2CRANKINF0%2CMATCHTYPE&pvid=TL  
    -41832&refpid=mm_26632258_3504122_32554087&clk1=abab6283306413775910d4b0b37ca047&idfa=%7Bidfa%7D&pid' \   
    '=430680_1006&keyword=' + key_word + '&count=60&offset=' + str(60 * page) + '&relacount=8&t=1535075213992' \   
    '&callback' \   
    '=mn17jsonp1535075213992'  
    r = session.get(url=url, headers=headers)  
    html = r.text  
    start = html.find('(')  
    return (json.loads(html[start + 1:-1]))['result']['item']
```

抓取地址

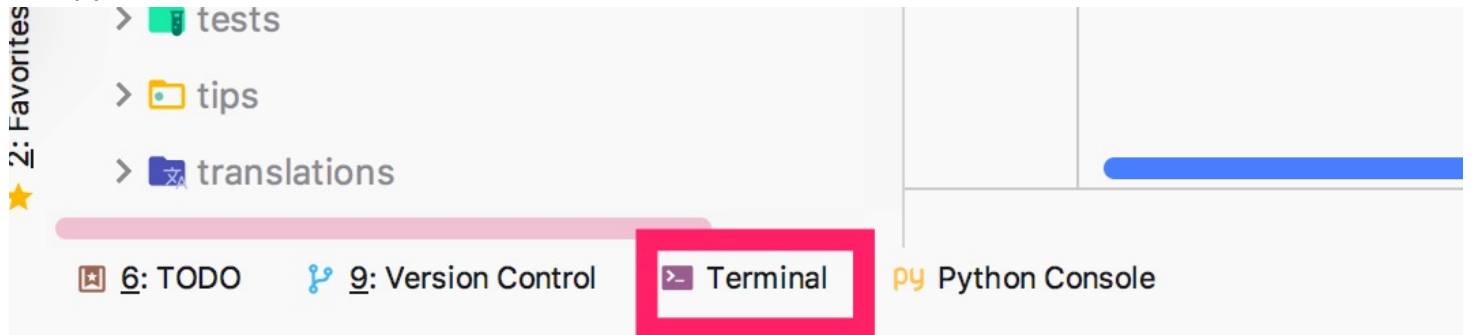
```
if __name__ == '__main__':  
    result = []  
    keyword = "pvdc保鲜膜".replace(" ", "+")  
    for i in range(10):  
        data = fetch_data(i+1, keyword)  
        result.extend(data)  
        time.sleep(random.choice([5, 6, 7, 8, 9, 10]))  
    save_to_excel(result, keyword)
```

搜索关键字

随机等待时间, 防止被封

手动安装依赖(上面没有成功)

打开pycharm的terminal:



运行:

```
pip install requests  
pip install openpyxl
```

## 数据处理可以参考

- [datacamp excel](#)
- [openpyxl](#)
- [openpyxl-说明](#)
- [fiddle 手机抓包](#)