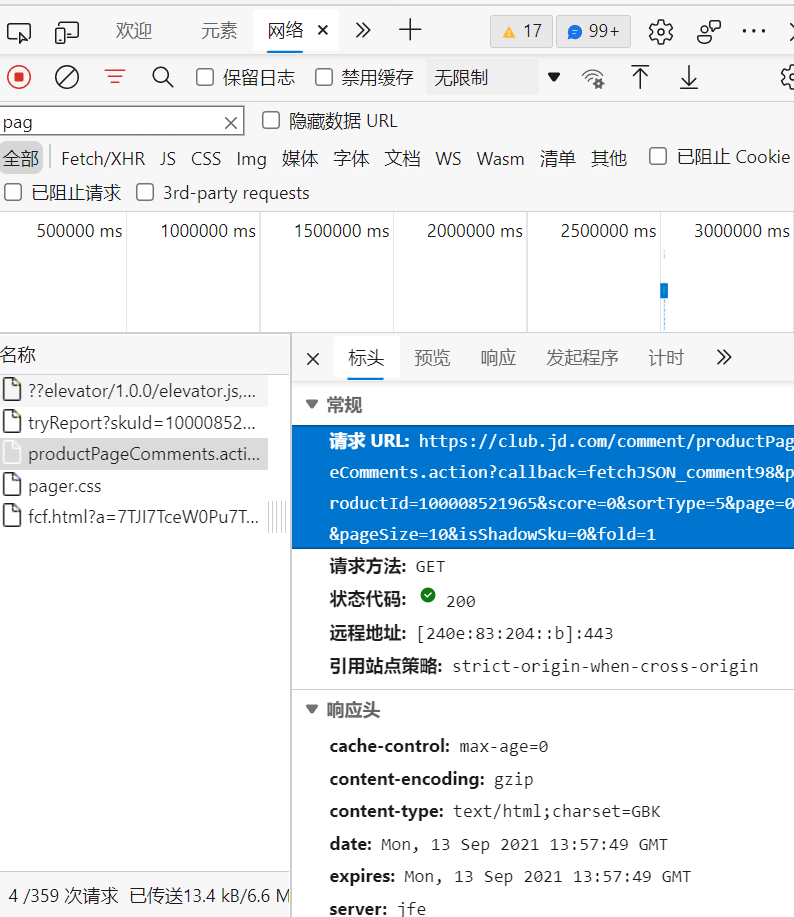
**Python爬取京东商品评论+制作关键字云+情感分析图**

一、分析网站

首先我们要了解一下，我们用requests包爬取的都只是写在HTML页面上的数据，但是呢，现在很多网站的数据都是通过js代码生成的，除非你找到它，不然是爬取不到的。

1. 首先我随意打开一个京东的商品的详情页面：<https://item.jd.com/100008521965.html>，然后点击检查，打开控制台，点击网络，再单击pagecomment过滤就发现评论是放再json文件的。

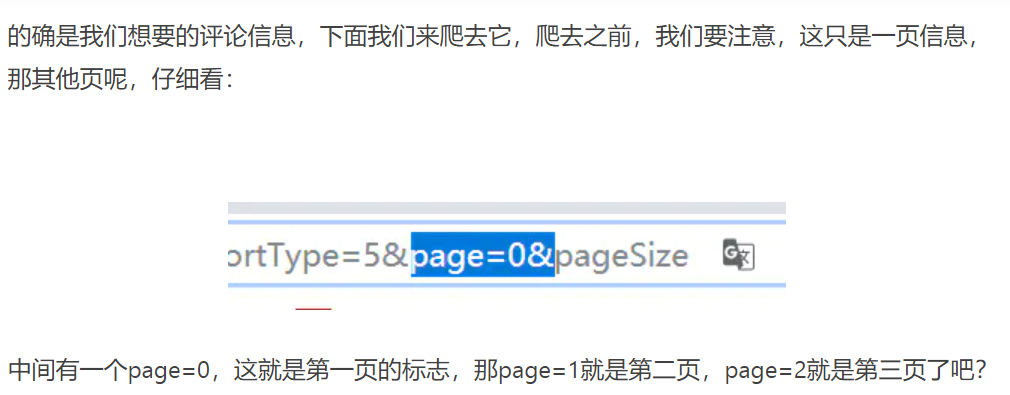
2.我们发现了productPageComment…，这也太像了吧，点开,header (双击也行)



点击预览：可以看到信息，比如每页显示多少评论，它的分数，

3.找到他的Request URL，复制下来，打开这个链接<https://club.jd.com/comment/productPageComments.action?callback=fetchJSON_comment98&productId=100008521965&score=0&sortType=5&page=0&pageSize=10&isShadowSku=0&fold=1>

4.的确是我们想要的评论信息，下面我们来爬去它，爬去之前，我们要注意，这只是一页信息，那其他页呢，仔细看：地址的变化。中间有一个page=0，这就是第一页的标志，那page=1就是第二页，page=2就是第三页了吧？试验一下：



全部评论的网址<https://club.jd.com/comment/productPageComments.action?callback=fetchJSON_comment98&productId=100008521965&score=0&sortType=5&page=0&pageSize=10&isShadowSku=0&fold=1>

好评：

<https://club.jd.com/comment/productPageComments.action?callback=fetchJSON_comment98&productId=100008521965&score=3&sortType=5&page=0&pageSize=10&isShadowSku=0&fold=1>

中评：

<https://club.jd.com/comment/productPageComments.action?callback=fetchJSON_comment98&productId=100008521965&score=2&sortType=5&page=0&pageSize=10&isShadowSku=0&fold=1>

差评：

<https://club.jd.com/comment/productPageComments.action?callback=fetchJSON_comment98&productId=100008521965&score=1&sortType=5&page=0&pageSize=10&isShadowSku=0&fold=1>

我们发现只要修改page 数字就可以了

二、开始数据爬取

1.准备需要的包

import json

import time

import random

import requests

import pandas as pd

import csv

import os

import re

二、数据爬取

1. requests基础请求地址