1. **从腾讯位置大数据爬取3月25日通过火车迁入北京的热度**

代码如下：

import requests

import json

import pandas as pd

import time

def get\_TecentData():

url='https://heat.qq.com/api/getLbsMigrateDataByBeijing.php?city=%E5%8C%97%E4%BA%AC&direction=0&type=2&date=2020-02-01'

response=requests.post(url)

datas=response.text

dictdatas=json.loads(datas)#dumps是将dict转化成str格式，loads是将str转化成dict格式

result1=pd.DataFrame.from\_dict(dictdatas,orient='index').T

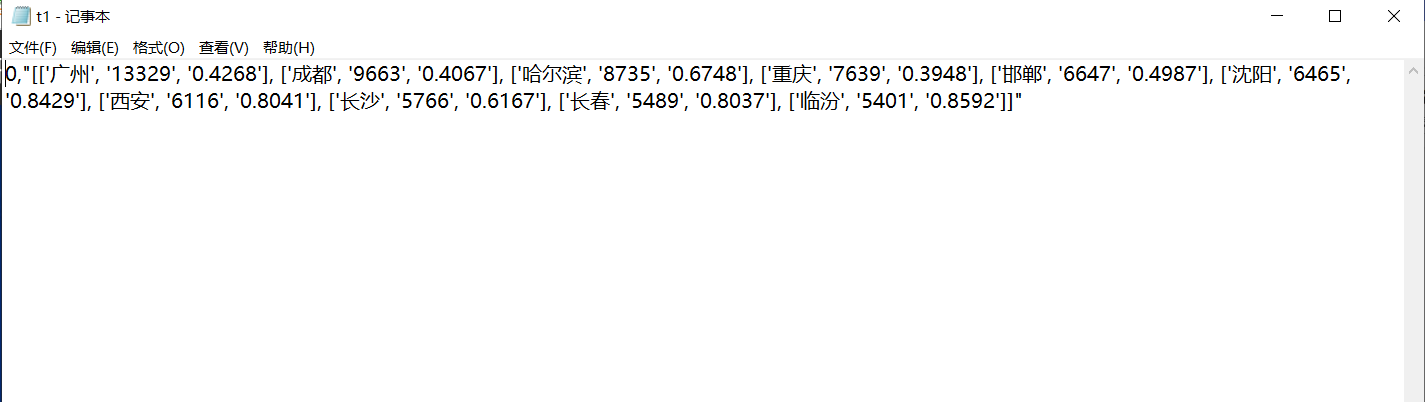
result1.dropna()

result1.to\_csv('t1.txt',mode='a',index = False,header=None)

if \_\_name\_\_ =='\_\_main\_\_':

get\_TecentData()

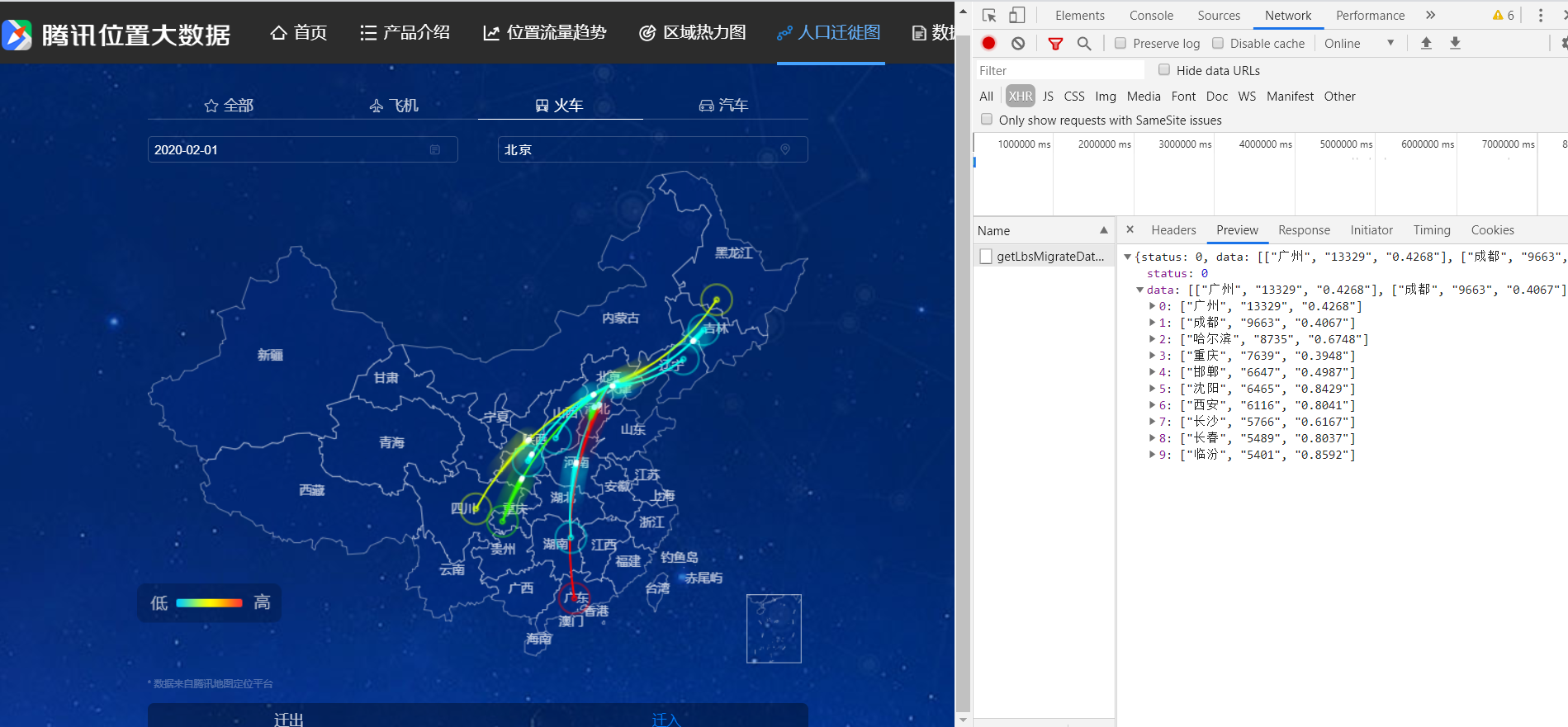
爬取结果：



改进：

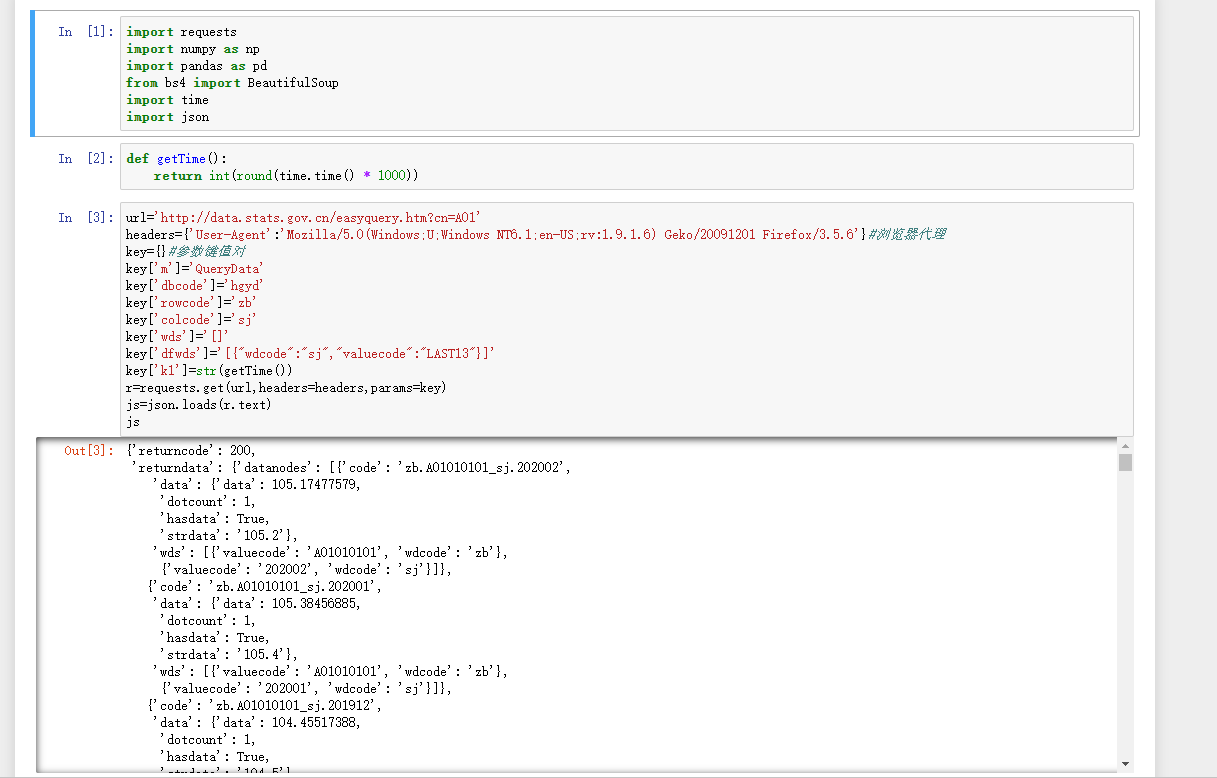
将结果导入Excel（正在进行中）

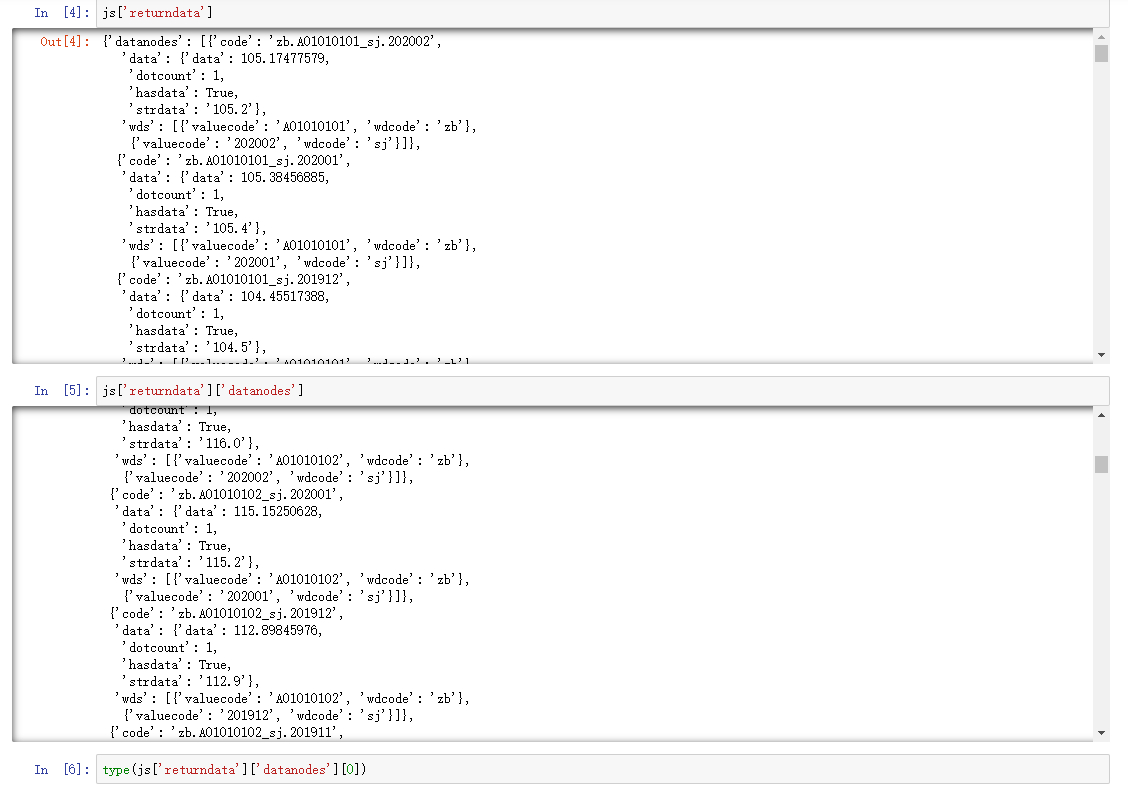
问题：



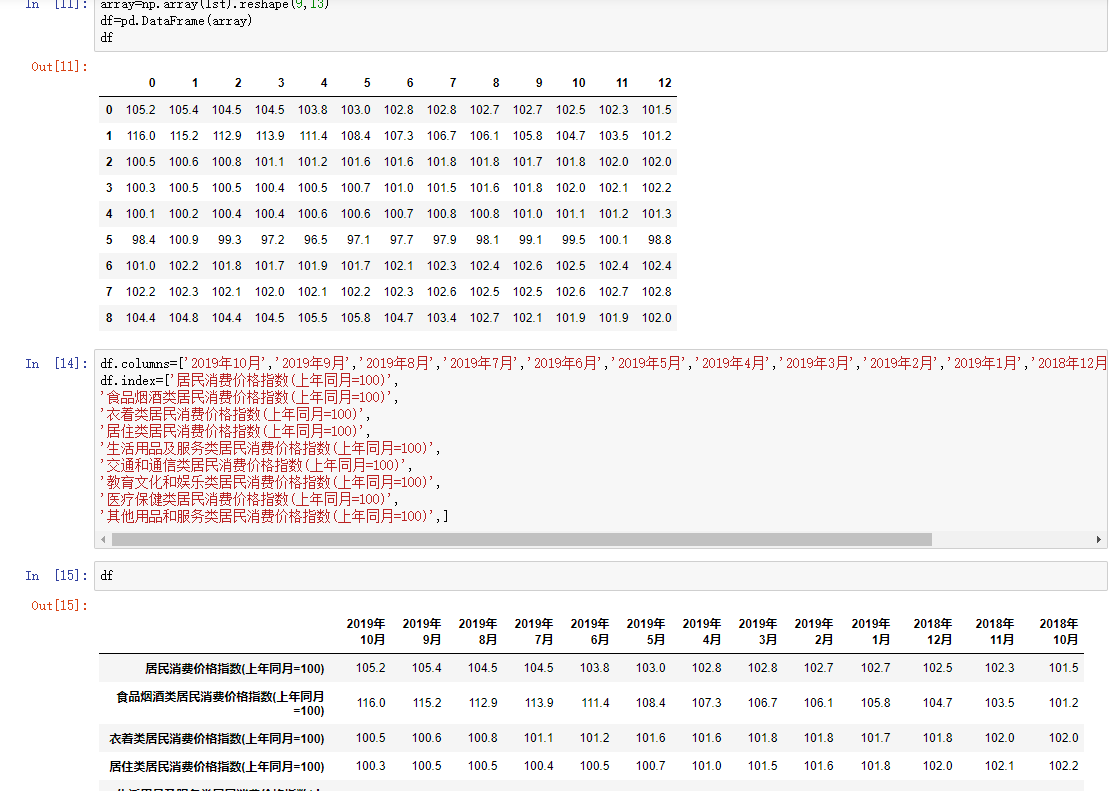
此网页只能对某一天进行爬取，要得到几个月的数据是从代码入手还是合并表格？

1. **从国家数据网站爬取数据**









以上为查询示例，问题在于：地区数据没有针对北京2020年疫情期间的月度详细数据。

1. 通过高德POI获取数据

from urllib.parse import quote  
from urllib import request  
import json  
import xlwt  
  
#TODO 替换为上面申请的密钥  
amap\_web\_key = '4dc929f465254fab4edc7acc30c34610'  
poi\_search\_url = "http://restapi.amap.com/v3/place/text"  
poi\_boundary\_url = "https://ditu.amap.com/detail/get/detail"  
#from transCoordinateSystem import gcj02\_to\_wgs84  
  
#TODO cityname为需要爬取的POI所属的城市名，nanning\_areas为城市下面的所有区，classes为多个分类名集合. (中文名或者代码都可以，代码详见高德地图的POI分类编码表)  
cityname = '北京'  
nanning\_areas = ['东城区','西城区','朝阳区','丰台区','石景山区','海淀区','门头沟区','房山区','通州区','顺义区','昌平区','大兴区','怀柔区','平谷区','密云区','延庆区']  
  
classes = ['高速收费站', '高速加油站服务区']  
  
  
# 根据城市名称和分类关键字获取poi数据  
def getpois(cityname, keywords):  
 i = 1  
 poilist = []  
 while True: # 使用while循环不断分页获取数据  
 result = getpoi\_page(cityname, keywords, i)  
 print(result)  
 result = json.loads(result) # 将字符串转换为json  
 if result['count'] == '0':  
 break  
 hand(poilist, result)  
 i = i + 1  
 return poilist  
  
  
# 数据写入excel  
def write\_to\_excel(poilist, cityname, classfield):  
 # 一个Workbook对象，这就相当于创建了一个Excel文件  
 book = xlwt.Workbook(encoding='utf-8', style\_compression=0)  
 sheet = book.add\_sheet(classfield, cell\_overwrite\_ok=True)  
  
 # 第一行(列标题)  
 sheet.write(0, 0, 'x')  
 sheet.write(0, 1, 'y')  
 sheet.write(0, 2, 'count')  
 sheet.write(0, 3, 'name')  
  
  
 for i in range(len(poilist)):  
 location = poilist[i]['location']  
 name = poilist[i]['name']  
 lng = str(location).split(",")[0]  
 lat = str(location).split(",")[1]  
  
 '''  
 result = gcj02\_to\_wgs84(float(lng), float(lat))  
  
 lng = result[0]  
 lat = result[1]  
 '''  
  
 # 每一行写入  
 sheet.write(i + 1, 0, lng)  
 sheet.write(i + 1, 1, lat)  
 sheet.write(i + 1, 2, 1)  
 sheet.write(i + 1, 3, name)  
  
  
  
 # 最后，将以上操作保存到指定的Excel文件中  
 book.save(r'' + cityname + "\_" + classfield + '.xls')  
  
  
# 将返回的poi数据装入集合返回  
def hand(poilist, result):  
 # result = json.loads(result) # 将字符串转换为json  
 pois = result['pois']  
 for i in range(len(pois)):  
 poilist.append(pois[i])  
  
  
# 单页获取pois  
def getpoi\_page(cityname, keywords, page):  
 req\_url = poi\_search\_url + "?key=" + amap\_web\_key + '&extensions=all&keywords=' + quote(  
 keywords) + '&city=' + quote(cityname) + '&citylimit=true' + '&offset=25' + '&page=' + str(  
 page) + '&output=json'  
 data = ''  
 with request.urlopen(req\_url) as f:  
 data = f.read()  
 data = data.decode('utf-8')  
 return data  
  
  
for clas in classes:  
 classes\_all\_pois = []  
 for area in nanning\_areas:  
 pois\_area = getpois(area, clas)  
 print('当前城区：' + str(area) + ', 分类：' + str(clas) + ", 总的有" + str(len(pois\_area)) + "条数据")  
 classes\_all\_pois.extend(pois\_area)  
 print("所有城区的数据汇总，总数为：" + str(len(classes\_all\_pois)))  
  
 write\_to\_excel(classes\_all\_pois, cityname, clas)  
 print('================分类：' + str(clas) + "写入成功")

能够爬取北京的高速收费站和加油站，问题在于高德地图提供的服务怎么爬取车流量？

