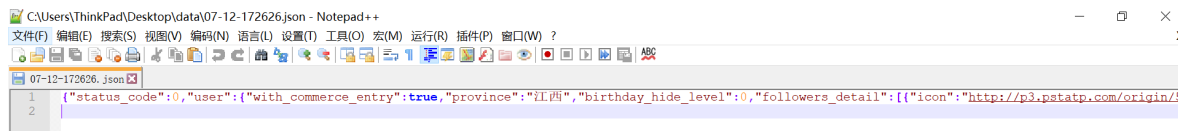


抖音大作业06——json文件转为csv

1 json文件的预览

1. 用fiddler爬取的json文件，使用notebook查看时，只能显示为一行

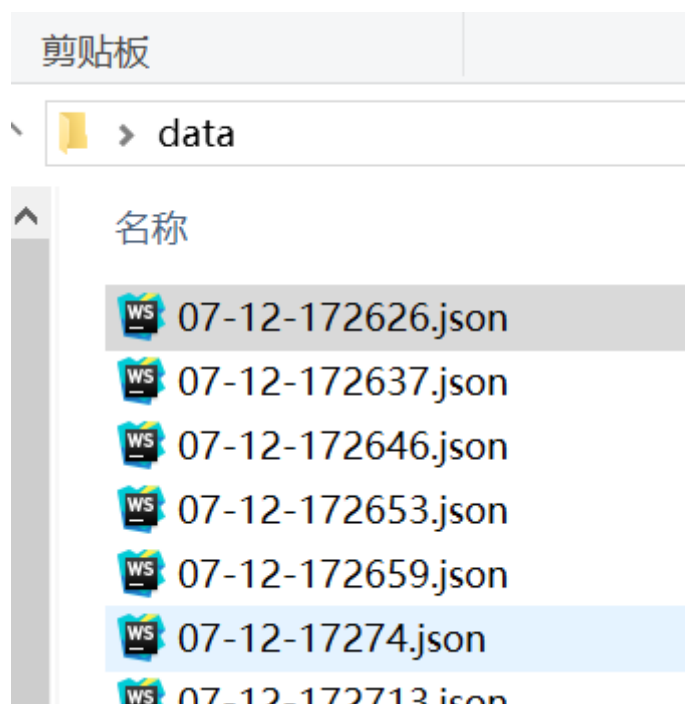


2. 因此可以进入网站https://www.bejson.com/explore/index_new/，复制json，即可
3. 可清晰观察格式化后的json文件结构



2 json转csv

1. 之前我们已经爬取了一些json文件



2. 我们可以在jupyter运行以下代码，将json运行行为python

```

# -*- coding:utf-8 -*-
import os
import json
import pandas as pd
import csv
import time

#文件夹所有json文件列表
json_list=os.listdir("C:/Users/ThinkPad/Desktop/data")

for jsons in json_list:
    with open("C:/Users/ThinkPad/Desktop/data/{}".format(jsons),'r',
encoding='UTF-8') as f:
        load_dict = json.load(f)
        s = json.dumps(load_dict,ensure_ascii=False)
        # 将 JSON 对象转换为 Python 字典
        params_json = json.loads(s)

        #字典的key
        name=params_json["user"]["nickname"]
        city=params_json["user"]["city"]
        follower_count=params_json["user"]["follower_count"]
        total_favorited=params_json["user"]["total_favorited"]

    now = time.strftime("%Y-%m-%d-%H_%M_%S",time.localtime(time.time()))
    filename="C:/Users/ThinkPad/Desktop/data1/"+now+".csv"

    #newline的作用是防止每次插入都有空行
    with open(filename, "a+", encoding='UTF-8',newline='') as csvfile:
        writer = csv.writer(csvfile)
        #以读的方式打开csv 用csv.reader方式判断是否存在标题。
        with open(filename, "r", encoding='UTF-8',newline='') as f:
            reader = csv.reader(f)
            if not [row for row in reader]:#空文件
                #先写入每一列的标题
                writer.writerow(["name","city","粉丝数","获赞总数"])
                #再写入每一列的内容
                writer.writerows([[name,city,follower_count,total_favorited]])
            else:#非空文件
                writer.writerows([[name,city,follower_count,total_favorited]])

```

3. 其中这一部分就是把json文件转换为对应结构的字典

```

load_dict = json.load(f)
s = json.dumps(load_dict,ensure_ascii=False)
# 将 JSON 对象转换为 Python 字典
params_json = json.loads(s)

```

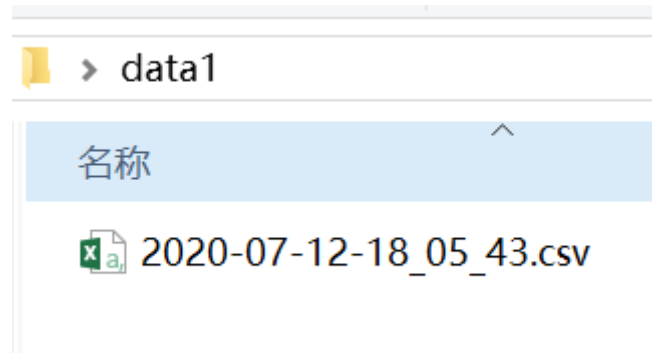
4. 这部分是根据key选择对应内容，以插入csv文件的行中

```

#字典的key
name=params_json["user"]["nickname"]
city=params_json["user"]["city"]
follower_count=params_json["user"]["follower_count"]
total_favorited=params_json["user"]["total_favorited"]

```

5. 运行此段代码可见文件已生成



6. 文件需要用notebook或记事本打开，用excel会乱码

