

江西財經大學

本科学生综合性实验报告

项目组长 张亮 学号 0194240

成员姓名 学号

专 业 软件工程 班级 软件 192 班

实验项目名称 基于 Python 爬虫的青年大学

习数据抓取及可视化分析

指导教师及职称 朱文强

开课学期 2021 至 2022 学年 1 学期

上课时间 2021 年 12 月 3 日

学生实验报告

学生姓名	张亮	学号	0194240	同组人：无
实验项目	基于 Python 爬虫的青年大学习数据 抓取及可视化分析			
<input type="checkbox"/> 必修 <input checked="" type="checkbox"/> 选修		<input type="checkbox"/> 演示性实验 <input type="checkbox"/> 验证性实验 <input type="checkbox"/> 操作性实验 <input checked="" type="checkbox"/> 综合性实验		
实验地点	W101		实验仪器台号	
指导教师	朱文强		实验日期及节次	2021—12—3 星期五 3-5 节

一、实验综述

1、实验目的及要求

实验目的：

梦想从学习开始，事业从实践起步。学习，是广大青年托举梦想、成就梦想的“奠基石”。中国特色社会主义进入新时代，广大青年必须不断学习，才能跟上时代前进 的脚步，才能融入民族复兴的伟业。共青团要站在理想信念的制高点，用习近平新时代中国特色社会主义思想构筑青年一代的强大精神支柱。党的十八大以来，习近平总书记高度重视青年的理论武装，多次指出共青团要做好对青年的思想引领工作，要善于使用青年容易接受且喜闻乐见的方式工作，用青年能学懂、易理解、学得会的方法阐释习近平新时代中国特色社会主义思想理论。在这一思想的引领下，“青年大学习”行动应运而生。

高等学校共青团组织是共青团的基础性组织，对于普及开展“青年大学习”活动，研究创新其活动内容和活动形式，发挥其活动作用，增强高等学校共青团组织的号召力、领导力、凝聚力具有至关重要的作用，但目前普遍存在着高校共青团员在“青年大学习”行动中的学习效果不佳的问题。“青年大学习”主要依托网络平台，开展一系列线上学习，线上学习的方式以观看教学视频和答题为主。在调研中发现，部分基层团组织检查学习效果以收取学习截图或通过检查学习完成后发朋友圈的方式进行。部分青年学生进行学习时，采用“挂时长”、“抄答案”等方式以应付检查，甚至都懒得应付，这就使得青年大学生的学习囫圇吞枣、疲于应付，基层团组织的检查工作流于表面，没有成效。

江西省团委采用的是云瓣科技大学习系统来对各个团委的完成率进行统计，并提供了后台系统方便各个团委查看完成情况，但并没有提供未完成同学的名单，这也使得在高校、学院层面难以对未完成学生、班级进行惩罚。

本实验尝试从学院的层面出发，设计制作一个 Python 脚本，利用 request、faker 模块模拟人工操作爬取云瓣科技系统后台 API 数据，使用 pandas、orjson 进行分析整理，openpyxl 导出学院完成情况、班级完成情况并使用 matplotlib 进行数据可视化，同时对比名单导出未完成同学数据，便于进一步的督促。

同时，项目使用 platform 进行设备判断，对一些无法通用的函数进行修改，能同时适配 macOS、Linux、Windows 三大主流操作系统；使用 configparser 读取 cfg 配置文件并引导用户进行配置；使用 tqdm 实现实时进度条；利用 pyinstaller 将项目打包为 exe 文件；项目同步开源至 [github 仓库](#)。

2、实验仪器、设备或软件

MacOS Monterey 12.2 Beta, MacBook Pro 2019, PyCharm 2021.3.1 (Professional Edition)

二、实验过程（实验步骤、记录、数据、分析）

1、所实现的项目内容

- ① 导出某期大学习各班完成率表格及统计图
- ② 导出总大学习完成情况表格及折线图
- ③ 导出某班每次大学习完成率情况表格及折线图
- ④ 导出最新一期未完成名单

⚠ 以上功能基于三级团组织实现（省属本科院校团委-江西财经大学团委-软件与物联网工程学院团委），理论上三级团组织完全适用，其他级别团组织自测。

2、实验过程

- ① 对项目进行配置
 - a. 修改默认配置文件 BigStudyConfigTemplate.cfg 中的账号密码为自己系统的账号密码（必填）
 - b. 根据默认配置文件 BigStudyConfigTemplate.cfg 中的提示，配置其他功能（影响功能 23）
 - c. 将 BigStudyConfigTemplate.cfg 修改为 BigStudyConfig.cfg
 - d. 将大学习名单（需要大学习的人的姓名、四级团支部名称）写入大学习名单 Template.xlsx 并改名为大学习名单.xlsx（影响功能 4）
 - e. pip install 以下依赖或者双击 exe 文件（需要 windows10+）
pip install -r requirements.txt

f. 运行 main.py 或者双击 exe 文件（需要 windows10+）

python3 main.py

② 使用 configparser 读取 cfg 配置文件并引导用户进行配置

```
if __name__ == '__main__':  
  
    print("程序初始化中。。")  
  
    readConfig()
```

readConfig()函数定义

```
def readConfig():  
  
    # 判断配置文件是否存在  
  
    if os.path.exists("BigStudyConfig.cfg") or os.path.exists("BigStudyConfigTemplate.cfg"):  
  
        if os.path.exists("BigStudyConfig.cfg"):  
  
            print("读取配置文件...")  
  
        else:  
  
            print("获取不到配置文件，请修改配置模板文件 BigStudyConfigTemplate.cfg 重命名为  
BigStudyConfig.cfg")  
  
            exit()  
  
        else:  
  
            print("配置文件和配置模板文件均已丢失，请从 github 重新获取")  
  
            print("")  
  
            exit()  
  
    # 生成 config 对象  
  
    conf = configparser.ConfigParser()  
  
    conf.sections()  
  
    # 用 config 对象读取配置文件
```

```

conf.read("BigStudyConfig.cfg")

# 以列表形式返回所有的 section

sections = conf.sections()

global account

global password

global IgnoreCourseBefore

global ShowCourseId

global Total

account = conf["DEFAULT"].get("Account", "")

password = conf["DEFAULT"].get("Password", "")

IgnoreCourseBefore = conf["DEFAULT"].get("IgnoreCourseBefore", "")

try:

    ShowCourseId = conf["DEFAULT"].getboolean("ShowCourseId", "False")

except ValueError:

    ShowCourseId = False

Total = conf["DEFAULT"].getint("Total", 0)

if len(account) == 0 or len(password) == 0:

    print("请配置必填项 Account Password")

    exit()

```

配置文件 BigStudyConfig.cfg（相关隐私信息已隐去）

```
[DEFAULT]
```

```
Account = #账号（必填）
```

```
Password = #密码（必填）
```

```
IgnoreCourseBefore = #C0038 选择菜单忽略指定 id 期数前的课程（选填，默认显示所有期数课程，影响功能 2、3）

ShowCourseId = #false（选填，默认菜单不显示课程 id）

Total = #12345 组织总人数（选填，功能 2 使用，总人数为计算完成率的分母，例如是总人数或者团员总数）
```

③ 检查之前的登录状态是否有效。如果有效则继续，否则尝试重新登陆，并将登录的 token 输出到 loginToken.data 文件中，便于下次使用

```
# 效验 token

def checkToken():

    print("效验登录状态...")

    global Token

    global headers

    data = readDataFromFile("loginToken", False)

    if not data:

        print("token 不存在，重新获取...")

        login()

        # checkToken()

    else:

        data = orjson.loads(data)

        Token = data["token"]

        headers["User-Agent"] = data["UA"]

        url = "https://jxtw.h5yunban.cn/jxtw-qndxx/cgi-bin/branch-api/info"

        values = {'accessToken': Token}

        status = sendGet(url, values)
```

```

# print(status)

global organizationInfo

organizationInfo = status

if status == "":

    login()

    # checkToken()

    return

print("效验成功...")

print("欢迎" + organizationInfo['branch'])

# time.sleep(3)

```

loginToken.data 文件

```

{"token":"F661C3DC-21D3-4647-A905-66620D7580FB",

"time":1640876236.030095,

"UA":"Mozilla/5.0 (iPod; U; CPU iPhone OS 4_2 like Mac OS X; ca-AD) AppleWebKit/534.40.6

(KHTML, like Gecko) Version/3.0.5 Mobile/8B115 Safari/6534.40.6"}

```

登录函数实现：

```

# 登录爬虫函数

def login():

    print("尝试登录...")

    global Token

    global headers

    fc = Factory.create()

    UA = fc.user_agent()

```

```

headers['User-Agent'] = UA

# 登录 API

url = "https://jxtw.h5yunban.cn/jxtw-qndxx/cgi-bin/login"

# application/json 参数

values = {'account': account, 'password': password}

Token = sendPostJson(url, values)['accessToken']

headers['User-Agent'] = UA

loginData = {'token': Token, 'time': time.time(), 'UA': UA}

loginData = orjson.dumps(loginData)

saveDataToFile("loginToken", loginData, True)

url = "https://jxtw.h5yunban.cn/jxtw-qndxx/cgi-bin/branch-api/info"

values = {'accessToken': Token}

status = sendGet(url, values)

global organizationInfo

organizationInfo = status

```

④ 导出某期大学习各班完成率表格及统计图

```

def getStudyInfo():

    global courseId

    global courseName

    url = "https://jxtw.h5yunban.cn/jxtw-qndxx/cgi-bin/branch-api/course/statis"

    values = {'course': courseId, 'nid': organizationInfo['nid'], 'accessToken': Token}

    result = sendGet(url, values)

    result = sorted(result, key=lambda info: info['title'])

```



```

df = pd.DataFrame(result)

# 数据清洗

df = df[~df['users'].isin([0])]

df = df.drop(['rate'], axis=1)

df.columns = ['团委 id', '团委名称', '团员人数', '已学习人数', '已学习人次', '完成率']

df['团员人数'].astype(int)

df['已学习人数'].astype(int)

df['已学习人次'].astype(int)

df['学习期数 old'] = courseName

df['完成率'].astype(float)

newLow = df.pop('学习期数 old')

df.insert(3, '大学习课程', newLow)

excel = pd.ExcelWriter(str(courseName + "大学习完成情况.xlsx"))

df.to_excel(excel, index=True, header=False)

excel.save()

excel.close()

mymovefile(str("./" + courseName + "大学习完成情况.xlsx"), "./处理结果/" + courseName + "
大学习完成情况.xlsx")

# for u in df['团委名称']:

#     u = str(u).replace("团支部", "")

for indexs in df.index:

    df.loc[indexs, '团委名称'] = df.loc[indexs, '团委名称'].replace("团支部", "")

# print(df)

```

```

print("已导出本期各班完成情况表 EXCEL 表到处理结果文件夹")

a = ""

for i in df['团委名称']:

    a += i

plt.rcParams['font.sans-serif'] = ['FangSong'] # 显示中、文

plt.rcParams['axes.unicode_minus'] = False # 显示负号

plt.figure(figsize=(df.shape[0], df.shape[0] / 2), dpi=100)

plt.bar(df['团委名称'], df['完成率'], color='#002EA6')

ax = plt.subplot(111)

plt.xticks(range(len(df['团委名称'])), df['团委名称'], rotation=45, fontsize=20)

plt.yticks(fontsize=80)

ax.set_title((courseName + "大学习完成情况", fontsize=80)

plt.xlabel("团委名称", fontsize=70)

plt.ylabel("完成率（单位：%）", fontsize=70)

plt.savefig(courseName + "大学习完成情况.png", bbox_inches='tight')

mymovefile(str("./" + courseName + "大学习完成情况.png"), "./处理结果/" + courseName + "
大学习完成情况.png")

print("已导出本期各班完成情况可视化统计图到处理结果文件夹")

```

⑤ 导出总大学习完成情况情况表格及折线图

```

elif int(fun) == 2:

    if Total <= 0:

        print("未配置总人数为 Total 或总人数 Total 不合法！")

        exit()

```

```
courseHis = []

courseList = getCourseInfo(True)

for i in courseList:

    courseHis.append({'大学习期数': i['title'], '完成率%': 100.0 * getLearnTime(i['id']) / Total})

courseHis.reverse()

df = pd.DataFrame(courseHis)

excel = pd.ExcelWriter("总大学习完成情况.xlsx")

df.to_excel(excel)

excel.save()

excel.close()

mymovefile("总大学习完成情况.xlsx", "./处理结果/总大学习完成情况.xlsx")

print("已导出总完成情况表 EXCEL 表到处理结果文件夹")

plt.rcParams['font.sans-serif'] = ['FangSong'] # 显示中、文

plt.rcParams['axes.unicode_minus'] = False # 显示负号

plt.figure(dpi=300)

plt.title(organizationInfo['branch'] + '总大学习完成情况', fontsize='20')

plt.xlabel("大学习期数", fontsize='16')

plt.xticks(range(len(df['大学习期数']), df['大学习期数'], rotation=45)

plt.ylabel("完成率（单位：%）", fontsize='16')

plt.plot(df['大学习期数'], df['完成率%'], color='#002EA6', linewidth=2, linestyle=':', marker='o')

# df.plot()

# plt.show()

outfile = organizationInfo['branch'] + '总大学习完成情况.png'
```

```

plt.savefig(outfile, bbox_inches='tight')

mymovefile(organizationInfo['branch'] + '总大学习完成情况.png', './处理结果/' +
organizationInfo['branch'] + '总大学习完成情况.png')

plt.close()

print("已导出总大学习完成情况可视化统计图到处理结果文件夹")

```

⑥ 导出某班每次大学习完成率情况表格及折线图

```

print("获取数据中")

courseHis = {}

courseList = getCourseInfo(True)

with tqdm(total=len(courseList), desc='数据统计', leave=True) as pbar:

    for i in courseList:

        res = getLearnTimeGroupByClass(i['id'])

        for ii in res:

            if ii.replace("团支部", "") not in courseHis.keys():

                courseHis[ii.replace("团支部", "")] = []

                courseHis[ii.replace("团支部", "")].append({i['title']: res[ii]})

        pbar.update(1)

noti = "请输入需要导出的班级的序号 (1~" + str(len(courseHis)) + "): "

for i in courseHis:

    courseHis[i].reverse()

classList = sorted(list(courseHis.keys()))

classList.reverse()

for i in range(1, len(classList) + 1):

```

```

        outMenu(str(i), classList[i - 1], len(noti.encode('GBK')) - 1)

print(noti)

c = input()

cleanWin()

data = courseHis[classList[int(c) - 1]]

datas = {'大学习期数': [], '完成率%': []}

# dataList = list(data.keys())

for i in data:

    # print(i)

    datas['大学习期数'].append(list(i.keys())[0])

    datas['完成率%'].append(i[list(i.keys())[0]])

df = pd.DataFrame(datas)

excel = pd.ExcelWriter(classList[int(c) - 1] + "总大学习完成情况.xlsx")

df.to_excel(excel)

excel.save()

excel.close()

mymovefile(classList[int(c) - 1] + "总大学习完成情况.xlsx", "./处理结果/" + classList[int(c) - 1] + "

总大学习完成情况.xlsx")

print("已导出" + classList[int(c) - 1] + "总完成情况表 EXCEL 表到处理结果文件夹")

plt.rcParams['font.sans-serif'] = ['FangSong'] # 显示中、文

plt.rcParams['axes.unicode_minus'] = False # 显示负号

plt.figure(dpi=300)

plt.title(classList[int(c) - 1] + '总大学习完成情况', fontsize='20')

```

```

plt.xlabel("大学习期数", fontsize='16')

plt.xticks(range(len(df['大学习期数']), df['大学习期数'], rotation=45))

plt.ylabel("完成率（单位：%）", fontsize='16')

plt.plot(df['大学习期数'], df['完成率%'], color='#002EA6', linewidth=2, linestyle='-', marker='o')

outfile = classList[int(c) - 1] + '总大学习完成情况.png'

plt.savefig(outfile, bbox_inches='tight')

mymovefile(classList[int(c) - 1] + '总大学习完成情况.png', "./处理结果/" + classList[int(c) - 1] + '总大学习完成情况.png')

plt.close()

print("已导出" + classList[int(c) - 1] + "总大学习完成情况可视化统计图到处理结果文件夹")

```

⑦ 导出最新一期未完成名单

```

elif int(fun) == 4:

    courseList = getCourseInfo(True)

    courseId = courseList[0]['id']

    df = pd.read_excel("大学习名单.xlsx")

    biaoshi = []

    for i in range(1, df.shape[0] + 1):

        biaoshi.append(df.loc[[i - 1], values[0]][0] + df.loc[[i - 1], values[0]][1])

    df.insert(df.shape[1], '标识符', biaoshi)

    dfJson = orjson.loads(df.to_json(orient='records'))

    datas = {}

    for i in dfJson:

        datas[i['标识符']] = {'姓名': i['姓名'], '所在团支部': i['所在团支部'], '完成情况': '未完成'}

```

```

url = "https://jxtw.h5yunban.cn/jxtw-qndxx/cgi-bin/branch-api/course/records"

values = {'course': courseId, 'nid': organizationInfo['branch'], 'accessToken': Token, 'pageSize':
'1000', 'pageNum': 1}

resData = {}

while True:

    result = sendGet(url, values)

    if values['pageNum'] != 1:

        resData.extend(result['list'])

    else:

        resData = result['list']

    if int(result['pagedInfo']['total']) > int(result['pagedInfo']['pageSize']) * int(
        result['pagedInfo']['pageNum']):

        values['pageNum'] = values['pageNum'] + 1

    else:

        break

errorTxt = ""

with tqdm(total=len(resData), desc='数据处理中', leave=True, as pbar:

    for i in resData:

        try:

            perData = orjson.loads(str(i).replace('\n', '\n'))

            datas[i['cardNo'] + i['branches'][-1].replace('团支部', '')]['完成情况'] = '已完成'

        except KeyError:

            errorTxt = errorTxt + (

```

```

        "姓名 " + i['cardNo'] + " 或者 " + i['branches'][-1].replace('团支部','') + "

团支部不存在或此人不属于此团支部，已跳过!\n")

        time.sleep(0.001)

        pbar.update(1)

    edf = pd.DataFrame(datas)

    edf = pd.DataFrame(edf.values.T, index=edf.columns, columns=edf.index)

    excel = pd.ExcelWriter("大学习未完成情况.xlsx")

    edf.to_excel(excel)

    excel.save()

    excel.close()

    mymovefile("大学习未完成情况.xlsx", "./处理结果/大学习未完成情况.xlsx")

    filename = r'./处理结果/大学习未完成情况.xlsx'

    wb = load_workbook(filename)

    ws = wb.active

    ws.delete_cols(1, # 删除第 13 列数据

    wb.save(filename)

    if errorTxt != "":

        print("已导出未完成情况表 EXCEL 表到处理结果文件夹，其中一些数据似乎有误：\n" +

errorTxt)

        print("按任意键退出！")

        c = input()

    else:

        print("已导出未完成情况表 EXCEL 表到处理结果文件夹")

```


三、结论

1. 实验结果

① 系统效验登录状态

token 存在时，效验 token 有效性，无效则重新登陆

```
/Library/Frameworks/Python.framework/Versions/3.10/bin/python3 /Users/zhangliang/PyCharmProject/BigStydyCompletionVisualization/main.py
程序初始化中。。
读取配置文件...
效验登录状态...
请求错误{"status":403,"message":"登录状态已过期，请重新登录","result":null,"debug":null}
尝试登录...
```

菜单显示：

```
欢迎软件与物联网工程学院团委
1          导出某期大学习各班完成率表格及统计图
2          导出总大学习完成情况情况表格及折线图
3          导出某班每次大学习完成率情况表格及折线图
4          导出最新一期未完成名单
5          退出程序
欢迎使用BigStydyCompletionVisualization，请选择功能(1~5):
```

② 导出某期大学习各班完成率表格及统计图

菜单下输入 1，进入功能 1：

```
运行: 启动 x
1、第十二季第十三期      C0051
2、第十二季第十二期      C0050
3、第十二季第十一期      C0049
4、第十二季第十期        C0048
5、第十二季第九期        C0047
6、“青年大学习”特辑      C0046
7、第十二季第八期        C0045
8、第十二季第七期        C0044
9、第十二季第六期        C0043
10、第十二季第五期       C0042
11、第十二季第四期       C0041
请输入需要进行操作的序号(1~11):
```

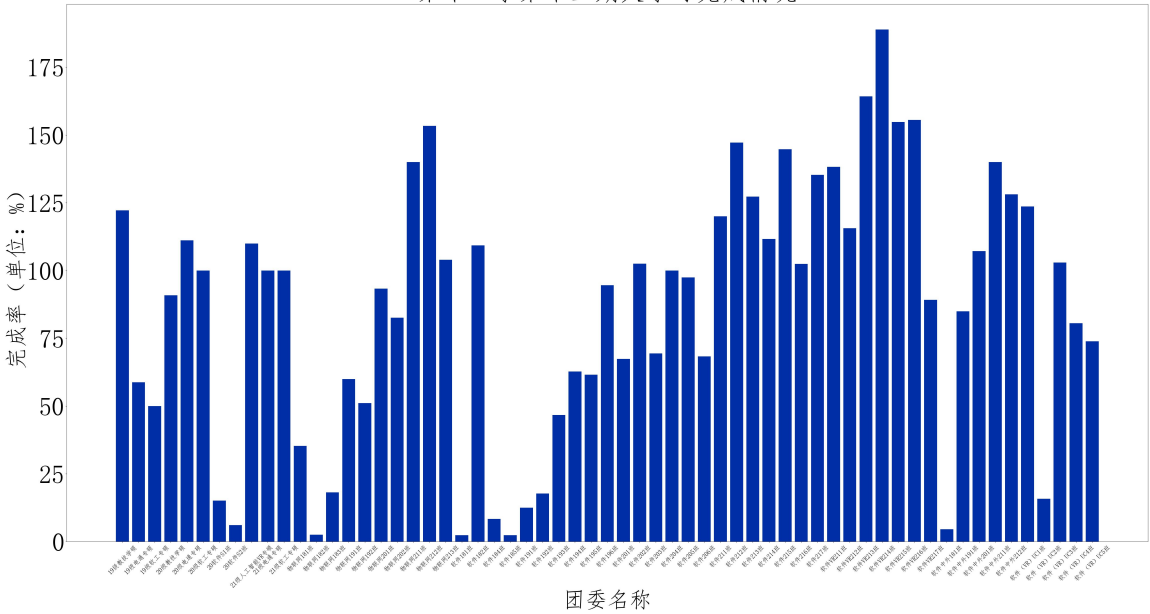
选中一期数据进行分析：

```
运行: 启动 x
已选中青年大学习的第十二季第十三期数据
已导出本期各班完成情况表EXCEL表到处理结果文件夹
已导出本期各班完成情况可视化统计图到处理结果文件夹
进程已结束,退出代码0
```

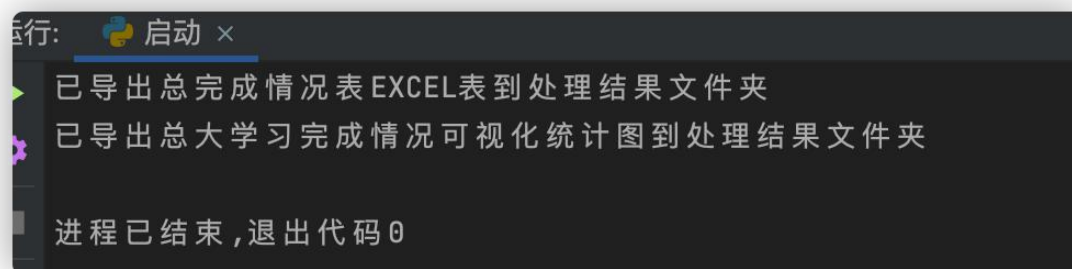
相关数据导出到处理结果文件夹下：

第十二季第十三期大学习完成情况.xlsx		使用 WPS Office 打开		
0	N00130004 19级教技专9	第十二季第11三期	26	122.22
1	N00130004 19级电通专17	第十二季第11三期	30	58.82
2	N00130004 19级软工专24	第十二季第11三期	26	50
3	N00130004 20级教技专11	第十二季第11三期	30	90.91
4	N00130004 20级电通专18	第十二季第11三期	49	111.11
5	N00130004 20级软工专31	第十二季第11三期	80	100
6	N00130004 20软件S1班66	第十二季第11三期	21	15.15
7	N00130004 20软件S2班66	第十二季第11三期	10	6.06
11	N00130004 21级人工智10	第十二季第11三期	24	110
12	N00130004 21级电通专16	第十二季第11三期	32	100
13	N00130004 21级软工专34	第十二季第11三期	92	100
16	N00130004 物联网181134	第十二季第11三期	23	35.29
17	N00130004 物联网182139	第十二季第11三期	2	2.56
18	N00130004 物联网183133	第十二季第11三期	13	18.18
19	N00130004 物联网191140	第十二季第11三期	57	60
20	N00130004 物联网192145	第十二季第11三期	47	51.11
21	N00130004 物联网201145	第十二季第11三期	96	93.33
22	N00130004 物联网202146	第十二季第11三期	96	82.61
23	N00130004 物联网211125	第十二季第11三期	95	140
24	N00130004 物联网212130	第十二季第11三期	121	153.33
25	N00130004 物联网213125	第十二季第11三期	73	104
26	N00130004 软件181班团支部	第十二季第11三期	4	2.33
27	N00130004 软件182班团支部	第十二季第11三期	140	109.3
29	N00130004 软件184班团支部	第十二季第11三期	10	8.33
30	N00130004 软件185班团支部	第十二季第11三期	2	2.33
31	N00130004 软件191班团支部	第十二季第11三期	10	12.5
32	N00130004 软件192班团支部	第十二季第11三期	18	17.78
33	N00130004 软件193班团支部	第十二季第11三期	49	46.67
34	N00130004 软件194班团支部	第十二季第11三期	54	62.79
35	N00130004 软件195班团支部	第十二季第11三期	67	61.54
36	N00130004 软件196班团支部	第十二季第11三期	81	94.59
37	N00130004 软件201班团支部	第十二季第11三期	61	67.44
38	N00130004 软件202班团支部	第十二季第11三期	101	102.5
39	N00130004 软件203班团支部	第十二季第11三期	55	69.44
40	N00130004 软件204班团支部	第十二季第11三期	94	100
41	N00130004 软件205班团支部	第十二季第11三期	110	97.5
42	N00130004 软件206班团支部	第十二季第11三期	76	68.29
43	N00130004 软件211班团支部	第十二季第11三期	135	120
44	N00130004 软件212班团支部	第十二季第11三期	163	147.22
45	N00130004 软件213班团支部	第十二季第11三期	147	127.27
46	N00130004 软件214班团支部	第十二季第11三期	115	111.63
47	N00130004 软件215班团支部	第十二季第11三期	168	144.74
48	N00130004 软件216班团支部	第十二季第11三期	122	102.44
49	N00130004 软件217班团支部	第十二季第11三期	131	135.29
50	N00130004 软件VR21134	第十二季第11三期	133	138.24
51	N00130004 软件VR21232	第十二季第11三期	97	115.63
52	N00130004 软件VR21328	第十二季第11三期	133	164.29
53	N00130004 软件VR21427	第十二季第11三期	128	188.89
54	N00130004 软件VR21531	第十二季第11三期	128	154.84
55	N00130004 软件VR21627	第十二季第11三期	111	155.56

第十二季第十三期大学习完成情况

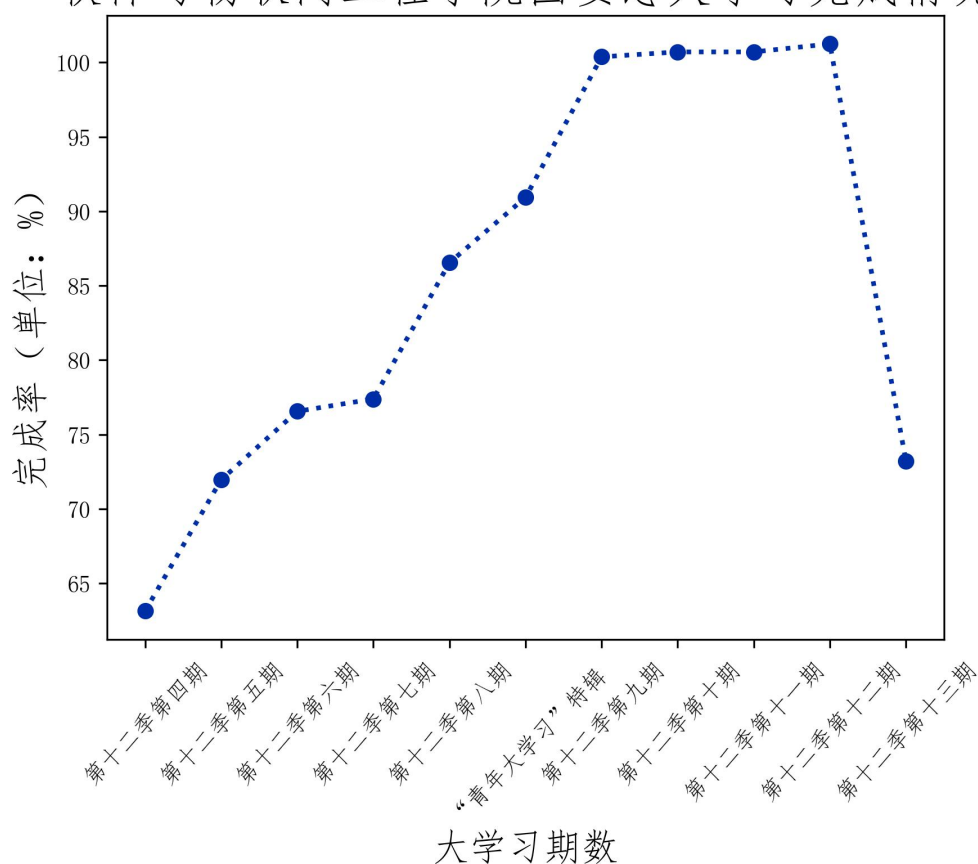


③ 导出总大学习完成情况情况表格及折线图
菜单栏选中功能 2，数据导出到处理结果文件夹：

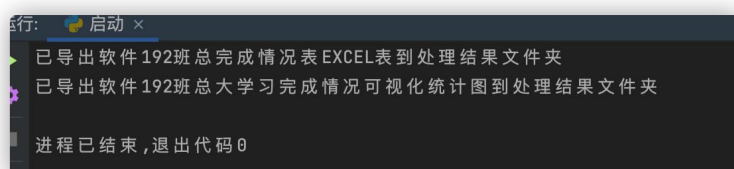
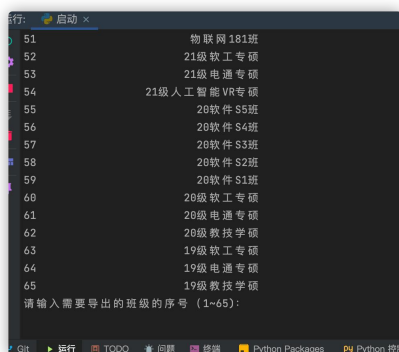
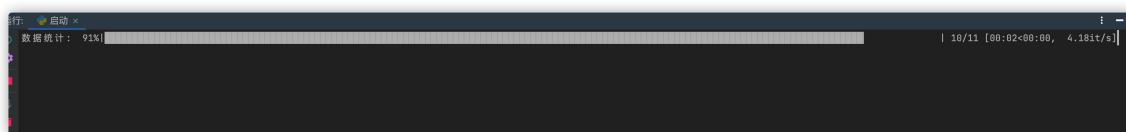


总大学习完成情况....			使用 WPS Office 打开
	大学习期数	完成率%	
0	第十二季第四期	63.13	
1	第十二季第五期	71.96	
2	第十二季第六期	76.55	
3	第十二季第七期	77.38	
4	第十二季第八期	86.54	
5	“青年大学习”特辑	90.93	
6	第十二季第九期	100.4	
7	第十二季第十期	100.7	
8	第十二季第十一期	100.7	
9	第十二季第十二期	101.2	
10	第十二季第十三期	73.2	

软件与物联网工程学院团委总大学习完成情况

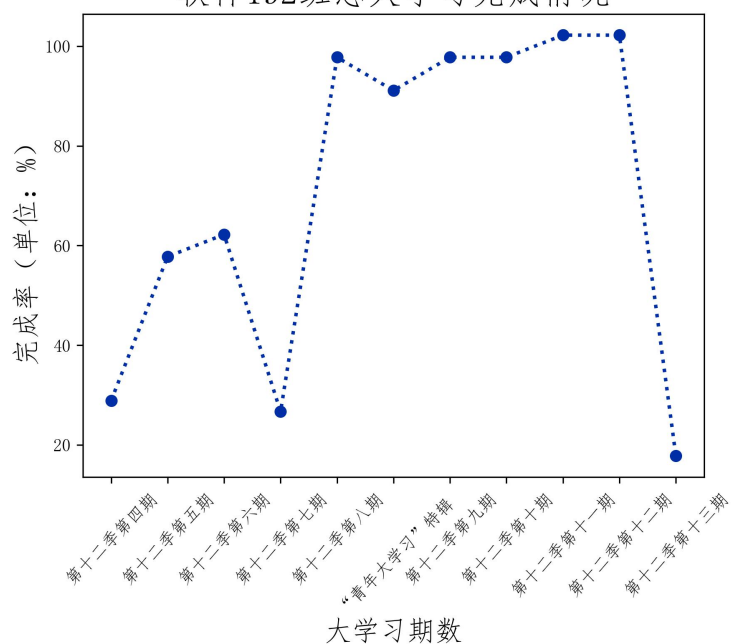


④ 导出某班每次大学习完成率情况表格及折线图
菜单栏选中菜单 3，选中一个班级进行分析如下：



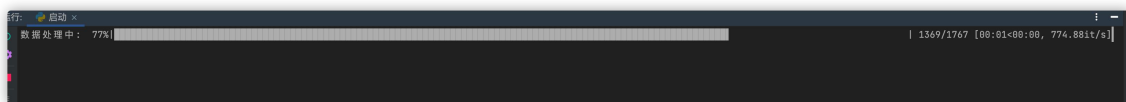
	大学习期数	完成率%
0	第十二季第四期	28.89
1	第十二季第五期	57.78
2	第十二季第六期	62.22
3	第十二季第七期	26.67
4	第十二季第八期	97.78
5	“青年大学习”特辑	91.11
6	第十二季第九期	97.78
7	第十二季第十期	97.78
8	第十二季第十一期	102.22
9	第十二季第十二期	102.22
10	第十二季第十三期	17.78

软件192班总大学习完成情况



⑤ 导出最新一期未完成名单
首先导入需要学习的同学名单及团支部 excel，并在菜单栏选中功能 4

	A	B	C	D
1	姓名	所在团支部		
2	陈鑫	软件181班		
3	宋梦琪	软件181班		
4	俞慧珍	软件181班		
5	漆园	软件181班		
6	龚晨艳	软件181班		
7	陈玉婷	软件181班		
8	蔡云帆	软件181班		
9	安雨珊	软件181班		
10	王婧	软件181班		
11	吴盈	软件181班		
12	胡伟清	软件181班		
13	刘泽源	软件181班		
14	陈晋	软件181班		
15	游文字	软件181班		
16	杨芳捷	软件181班		
17	陶之辰	软件181班		
18	柳加亮	软件181班		
19	胡建东	软件181班		
20	彭承龙	软件181班		
21	陈文录	软件181班		
22	夏振龙	软件181班		
23	胡超	软件181班		
24	黄泽龙	软件181班		
25	王威	软件181班		
26	程辉	软件181班		
27	江能宇	软件181班		
28	程华飞	软件181班		
29	谢凌峰	软件181班		



数据处理中: 100% | 1767/1767 [00:02:00:00, 785.531t/s]

已导出未完成情况表 EXCEL 表到处理结果文件夹，其中一些数据似乎有误：

姓名 宋昭顺 或者 软件 204 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 陈习京 0204476 或者 软件 中外 201 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 梅江文 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 戴飞龙 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 冯召先 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 张子豪 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 洪翻杰 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 梁天昊 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 周浩 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 董均坤 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

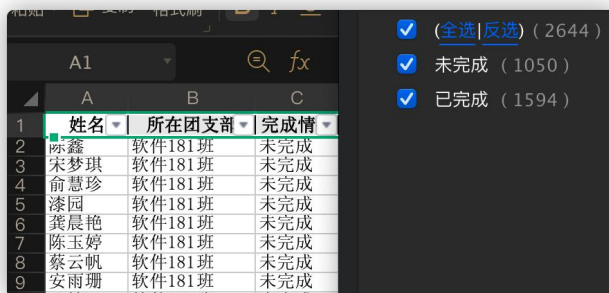
姓名 李文强 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 陶子昕 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 刘海旭 或者 物联网 212 班 团支部不存在或此人不属于此团支部，已跳过！

姓名 王瑞琪 或者 物联网 213 班 团支部不存在或此人不属于此团支部，已跳过！

	A	B	C	D	E
1	姓名	所在团支部	完成情况		
2	陈鑫	软件181班	未完成		
3	宋梦琪	软件181班	未完成		
4	俞慧珍	软件181班	未完成		
5	漆园	软件181班	未完成		
6	龚晨艳	软件181班	未完成		
7	陈玉婷	软件181班	未完成		
8	蔡云帆	软件181班	未完成		
9	安雨珊	软件181班	未完成		
10	王婧	软件181班	未完成		
11	吴盈	软件181班	未完成		
12	胡伟清	软件181班	未完成		
13	刘泽源	软件181班	未完成		
14	陈晋	软件181班	未完成		
15	游文字	软件181班	未完成		
16	杨芳捷	软件181班	未完成		
17	陶之辰	软件181班	未完成		
18	柳加亮	软件181班	未完成		
19	胡建东	软件181班	未完成		
20	彭承龙	软件181班	未完成		
21	陈文录	软件181班	未完成		
22	夏振龙	软件181班	未完成		
23	胡超	软件181班	未完成		
24	黄泽龙	软件181班	未完成		
25	王威	软件181班	未完成		
26	程辉	软件181班	未完成		
27	江能宇	软件181班	未完成		
28	程华飞	软件181班	未完成		
29	谢凌峰	软件181班	未完成		
30	熊露森	软件181班	未完成		



2. 分析与讨论

在整个分析过程中，综合运用到了 python 的基础知识，除此之外还有一些课外的扩展知识。在整个分析的过程中，我认识到了数据分析的重要性，它可以为人们的决策提供依据，对于进行数据分析的用户而言，信息的质量很大程度上依赖于其表达方式。将分析结果可视化，可以很好地帮助用户进行数据分析，借助可视化的手段，清晰有效地传达与沟通信息。

四、指导教师评语及成绩：

成绩：

指导教师签名：

批阅日期：