ITS概论作业一

王锦宏 19351125

本次作业的绝大部分操作均在Python下完成，题2画图使用Origin软件

测试运行环境：Python 3.8.3

需要安装numpy, pandas, matplotlib, datetime

Origin版本：Origin2020b(学习版)

**第一题**

此题第一部分为非常普通的数据提取，使用pandas的read\_csv读取，然后在pandas中进行数据筛选，分词，制作输出表格并将代码填入输出表格中即可，步骤如下：

1.将时间转化为标准格式，筛选所需要的六个设备CODE和预期的时间

2.创建DataFrame，遍历设备，载入初始化设置

3.遍历筛选后的数据，通过str.split()拆解得到所需要的数据并加和得到速度和时距总和

4.总和除以车辆数得到平均辆，更新表格并输出得到附件：ITS作业1.1.xls，即得到所需表格

代码如下：（见附件：ITS作业1.1.py）

import numpy as np

import pandas as pd

if \_\_name\_\_ == "\_\_main\_\_":

    a = pd.read\_csv("C:\\Users\\jhong\\Documents\\GitHub\\OpenITS-HeFei-Analyzation\\黄天路口视频交通流检测数据(2016.06.22~2016.06.30).csv")

    a['FTIME']=pd.to\_datetime(a['FTIME'])

    a['TTIME']=pd.to\_datetime(a['TTIME'])

    timecodeflitered=a.loc[(a['FTIME']>='2016/6/25 8:00')&(a['TTIME']<='2016/6/25 9:00')&(a['DEVICECODE'].isin(['hte001','hte002','hts001','htw001','htw002','htn001']))]

    flitered=timecodeflitered[['DEVICECODE','LANENUMBER','FLOW','SPEED','INTERVAL']]

    pathnum = [2,2,2,2,3,2]

    device = ['hte001','hte002','hts001','htw001','htw002','htn001']

    output=pd.DataFrame(columns=['路段','FLOW','SPEED','INTERVAL'])

    for i in range(0,6):

        for j in range(1,pathnum[i]+1):

            indname = (device[i]+'(车道'+str(j)+')')

            new=pd.DataFrame({'路段':indname,'FLOW':0,'SPEED':0,'INTERVAL':0},index=[1])

            output=output.append(new)

    output=output.set\_index('路段')

    dividerate=pd.DataFrame(columns=device,index=[1]).fillna(0)

    for index,row in flitered.iterrows():

        laneNum = row['LANENUMBER']

        FLOWS = row['FLOW'].split('\_',laneNum-1)

        SPEEDS = row['SPEED'].split('\_',laneNum-1)

        INTERVALS = row['INTERVAL'].split('\_',laneNum-1)

        dividerate.loc[1,row['DEVICECODE']] = dividerate.loc[1,row['DEVICECODE']]+1

        for i in range(0,laneNum):

            odname = (row['DEVICECODE']+'(车道'+str(i+1)+')')

            output.loc[odname].loc['FLOW']=int(output.loc[odname].loc['FLOW'])+int(FLOWS[i])

            output.loc[odname].loc['SPEED']=int(output.loc[odname].loc['SPEED'])+int(SPEEDS[i])

            output.loc[odname].loc['INTERVAL']=float(output.loc[odname].loc['INTERVAL'])+float(INTERVALS[i])

    for i in range(0,6):

        for j in range(1,pathnum[i]+1):

            indname = (device[i]+'(车道'+str(j)+')')

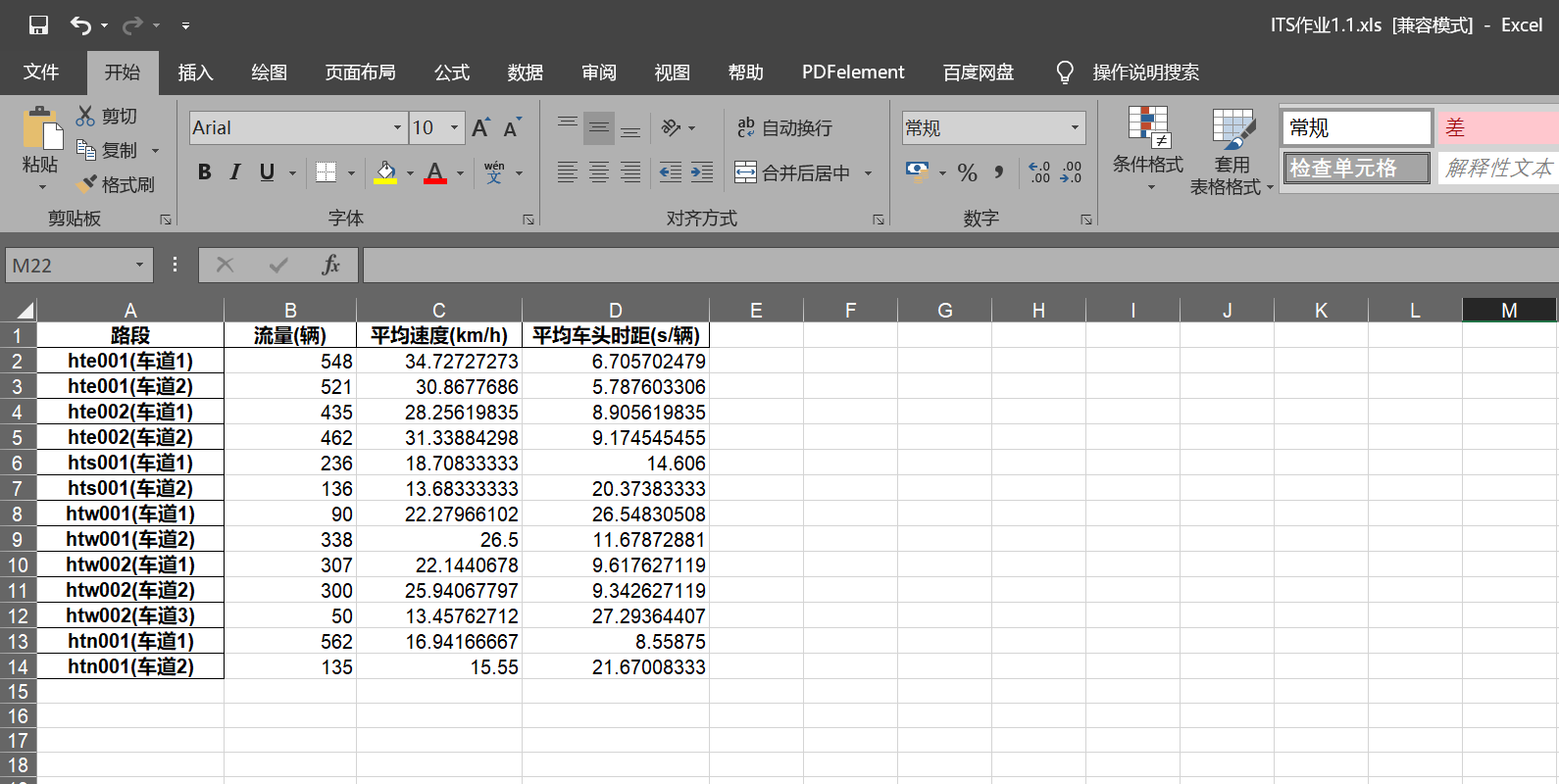
            output.loc[indname].loc['SPEED'] = output.loc[indname].loc['SPEED']/dividerate.loc[1,device[i]]

            output.loc[indname].loc['INTERVAL'] = output.loc[indname].loc['INTERVAL']/dividerate.loc[1,device[i]]

    output.columns = ['流量(辆)','平均速度(km/h)','平均车头时距(s/辆)']

    output.to\_excel("ITS作业1.1.xls")

得到的excel表格截图如下：(见附件"ITS作业1.1.xls")



|  |  |  |  |
| --- | --- | --- | --- |
| 路段 | 流量(辆) | 平均速度(km/h) | 平均车头时距(s/辆) |
| hte001(车道1) | 548 | 34.72727273 | 6.705702479 |
| hte001(车道2) | 521 | 30.8677686 | 5.787603306 |
| hte002(车道1) | 435 | 28.25619835 | 8.905619835 |
| hte002(车道2) | 462 | 31.33884298 | 9.174545455 |
| hts001(车道1) | 236 | 18.70833333 | 14.606 |
| hts001(车道2) | 136 | 13.68333333 | 20.37383333 |
| htw001(车道1) | 90 | 22.27966102 | 26.54830508 |
| htw001(车道2) | 338 | 26.5 | 11.67872881 |
| htw002(车道1) | 307 | 22.1440678 | 9.617627119 |
| htw002(车道2) | 300 | 25.94067797 | 9.342627119 |
| htw002(车道3) | 50 | 13.45762712 | 27.29364407 |
| htn001(车道1) | 562 | 16.94166667 | 8.55875 |
| htn001(车道2) | 135 | 15.55 | 21.67008333 |

此题第二部分为数据提取并制图，使用read\_csv读取后，然后在pandas中进行数据筛选，分词，得到流量——时间关系表，通过matplotlib制图即可，步骤如下：

import numpy as np

import pandas as pd

from matplotlib import pyplot as plt

if \_\_name\_\_ == "\_\_main\_\_":

    print('选择车道：hte001车道1')

    a = pd.read\_csv("C:\\Users\\jhong\\Documents\\GitHub\\OpenITS-HeFei-Analyzation\\黄天路口视频交通流检测数据(2016.06.22~2016.06.30).csv")

    a['FTIME']=pd.to\_datetime(a['FTIME'])

    a['TTIME']=pd.to\_datetime(a['TTIME'])

    timecodeflitered=a.loc[(a['FTIME']>='2016/6/25 8:00')&(a['TTIME']<='2016/6/25 9:00')&(a['DEVICECODE'].isin(['hte001']))]

    flitered=timecodeflitered[['DEVICECODE','FLOW','FTIME','TTIME']].reset\_index(drop=True)

    for i in range(0,(flitered.shape[0])):

        flitered.loc[i,'FLOW'] = (flitered.loc[i].loc['FLOW']).split('\_',1)[0]

    flitered=flitered[['FLOW','FTIME']]

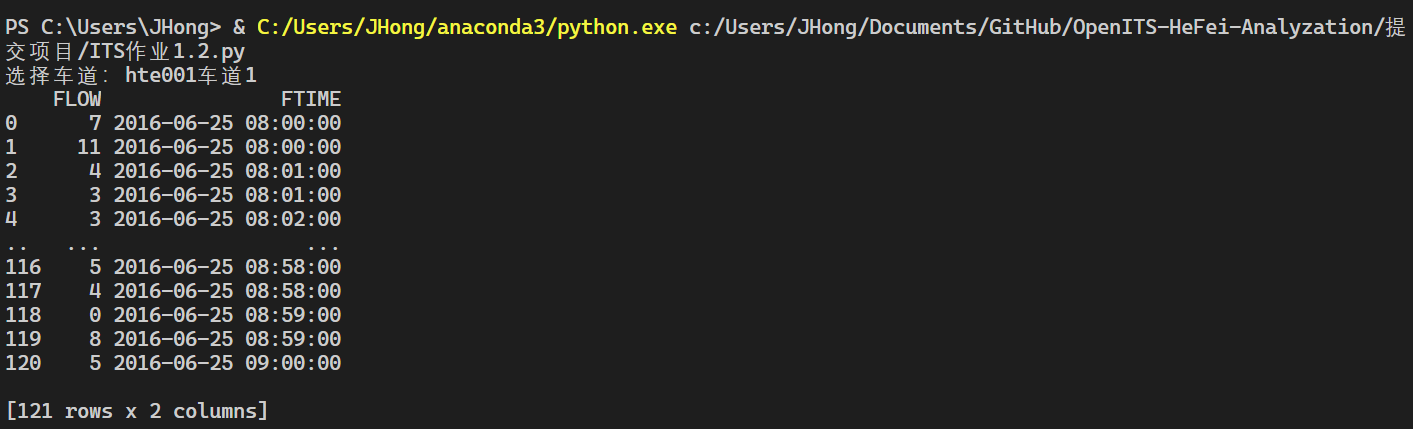
    print(flitered)

    flownum =  [int(x) for x in flitered['FLOW']]

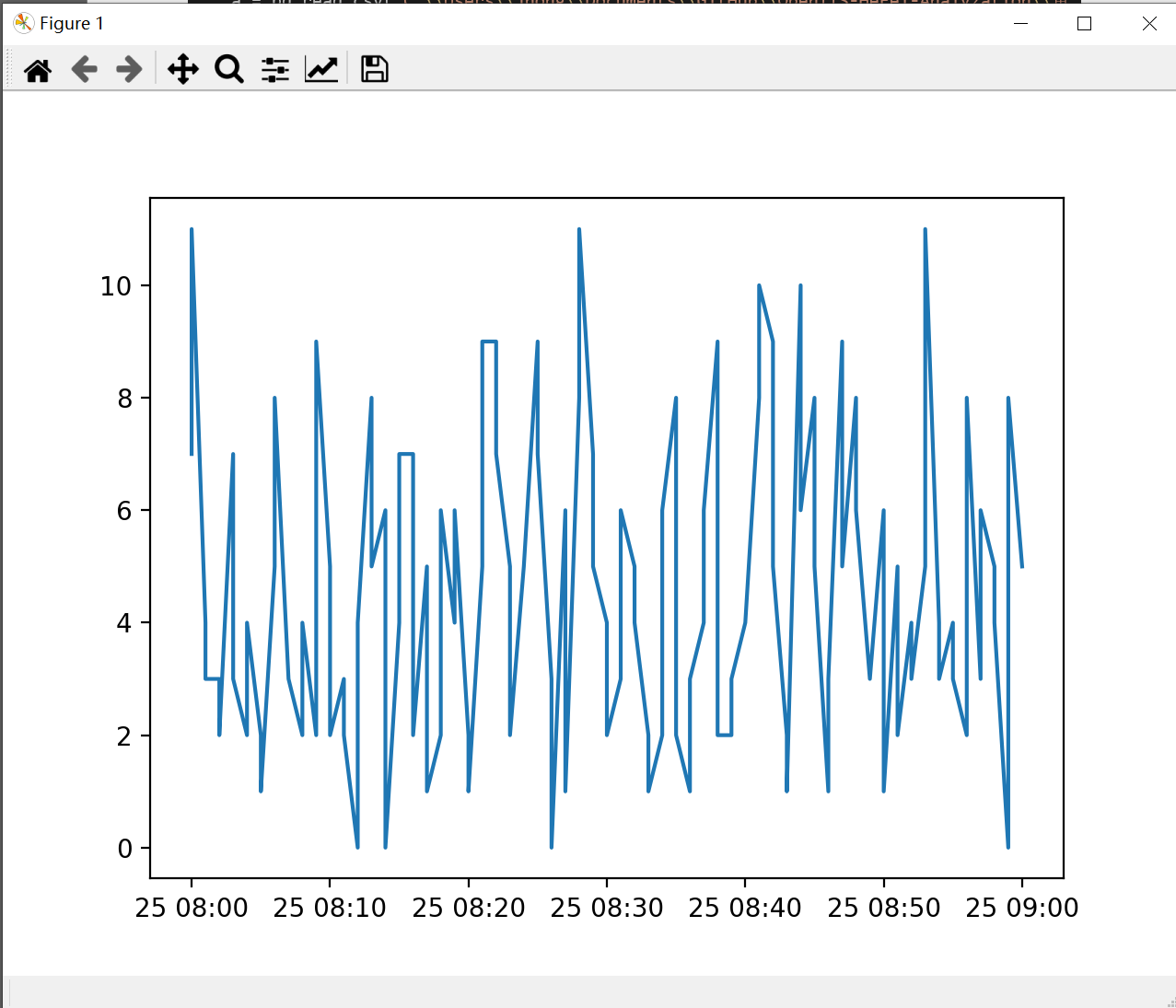
    plt.plot(flitered['FTIME'],flownum)

    plt.show()

代码行输出如下：



图像显示如下：



**第二题**

此题第一部分为在大量数据中分类并计算各个线圈的信息并制表，代码实现的思路如下：

1.读入数据，并将时间转化为标准格式，筛选黄天路口这一路口和所需要的时间范围。设置常量车长和线圈长度。

2.创建DataFrame，遍历线圈元数据表格，得到含有流量，总占用时间和总速度和的circuitaverinfo

3.使用新的list，计算占用时间比例和平均速度，并且根据LASTTIME创建时间切片，在循环时判断时间切片是否符合，并记录在此时间段中的信息

4.整理DataFrame信息并导出

代码如下：（见附件：ITS作业2.1.py）

import numpy as np

import pandas as pd

from datetime import datetime

if \_\_name\_\_ == "\_\_main\_\_":

    a = pd.read\_csv("C:\\Users\\jhong\\Documents\\GitHub\\OpenITS-HeFei-Analyzation\\示范区地磁检测数据(2016.06.22~2016.06.30).csv")

    a['INTIME']=pd.to\_datetime(a['INTIME'])

    a=a.loc[(a['INTIME']>='2016/6/25 8:00:00')&(a['INTIME']<='2016/6/25 9:00:00')&(a['DETECT\_ID'].isin(['0002']))]

    a=(a.sort\_values(by="INTIME"))[['DETECTOR\_NUMBER','OCC\_TIME','INTIME']].reset\_index(drop=True)

    car\_length = 3.5

    circuit\_length = 1.0

    circuitaverinfo=pd.DataFrame(columns=['FLOWNUM','OCCUPYTIME','SUMSPEED'],index=range(1,27)).fillna(0.0)

    for index,row in a.iterrows():

        circuitaverinfo.loc[row['DETECTOR\_NUMBER'],'FLOWNUM']=circuitaverinfo.loc[row['DETECTOR\_NUMBER'],'FLOWNUM']+1

        circuitaverinfo.loc[row['DETECTOR\_NUMBER'],'OCCUPYTIME']=circuitaverinfo.loc[row['DETECTOR\_NUMBER'],'OCCUPYTIME']+row['OCC\_TIME']

        circuitaverinfo.loc[row['DETECTOR\_NUMBER'],'SUMSPEED']=circuitaverinfo.loc[row['DETECTOR\_NUMBER'],'SUMSPEED']+float((car\_length+circuit\_length)/(row['OCC\_TIME']/1000))

    OCTIMERATE=[]

    AVESPEED=[]

    for index,row in circuitaverinfo.iterrows():

        OCTIMERATE.append(row['OCCUPYTIME']/(1000.00\*3600.00))

        AVESPEED.append(float(row['SUMSPEED']/row['FLOWNUM']))

    SUMDIS=[0.0 for x in range(1,27)]

    LASTTIME=[datetime.strptime('2016/6/25 8:00:00', '%Y/%m/%d %H:%M:%S') for x in range(1,27)]

    for index,row in a.iterrows():

        if (LASTTIME[row['DETECTOR\_NUMBER']-1]=='2016/6/25 8:00:00'):

            LASTTIME[row['DETECTOR\_NUMBER']-1]=row['INTIME']

        else:

            deltatime=(row['INTIME']-LASTTIME[row['DETECTOR\_NUMBER']-1]).total\_seconds()

            LASTTIME[row['DETECTOR\_NUMBER']-1]=row['INTIME']

            SUMDIS[row['DETECTOR\_NUMBER']-1]=SUMDIS[row['DETECTOR\_NUMBER']-1]+AVESPEED[row['DETECTOR\_NUMBER']-1]\*deltatime

    for index,row in circuitaverinfo.iterrows():

        SUMDIS[index-1]=SUMDIS[index-1]/row['FLOWNUM']

    circuitaverinfo.insert(1, 'OCCUPIED\_TIMERATE', OCTIMERATE)

    circuitaverinfo.insert(2, 'AVERAGESPEED', AVESPEED)

    circuitaverinfo.insert(3, 'AVERAGEHEADWAY', SUMDIS)

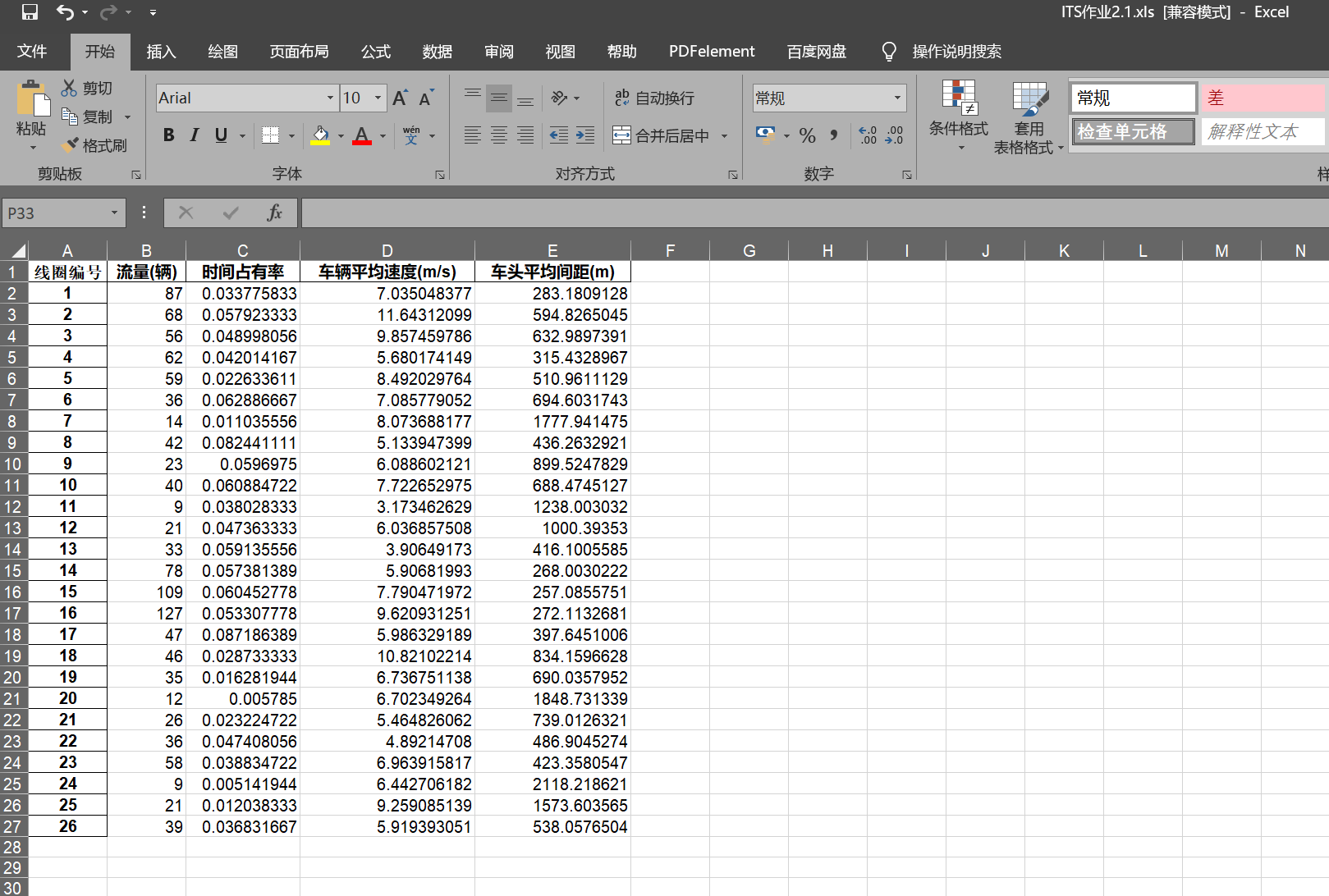
    circuitaverinfo=circuitaverinfo[['FLOWNUM','OCCUPIED\_TIMERATE','AVERAGESPEED','AVERAGEHEADWAY']]

    circuitaverinfo.columns = ['流量(辆)','时间占有率','车辆平均速度(m/s)','车头平均间距(m)']

    print(circuitaverinfo)

    circuitaverinfo.to\_excel("ITS作业2.1.xls")

即得到所需要的表格，截图如下：(注：首列”线圈编号”为手动添加)(见附件"ITS作业2.1.xls")



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 线圈编号 | 流量(辆) | 时间占有率 | 车辆平均速度(m/s) | 车头平均间距(m) |
| 1 | 87 | 0.033775833 | 7.035048377 | 283.1809128 |
| 2 | 68 | 0.057923333 | 11.64312099 | 594.8265045 |
| 3 | 56 | 0.048998056 | 9.857459786 | 632.9897391 |
| 4 | 62 | 0.042014167 | 5.680174149 | 315.4328967 |
| 5 | 59 | 0.022633611 | 8.492029764 | 510.9611129 |
| 6 | 36 | 0.062886667 | 7.085779052 | 694.6031743 |
| 7 | 14 | 0.011035556 | 8.073688177 | 1777.941475 |
| 8 | 42 | 0.082441111 | 5.133947399 | 436.2632921 |
| 9 | 23 | 0.0596975 | 6.088602121 | 899.5247829 |
| 10 | 40 | 0.060884722 | 7.722652975 | 688.4745127 |
| 11 | 9 | 0.038028333 | 3.173462629 | 1238.003032 |
| 12 | 21 | 0.047363333 | 6.036857508 | 1000.39353 |
| 13 | 33 | 0.059135556 | 3.90649173 | 416.1005585 |
| 14 | 78 | 0.057381389 | 5.90681993 | 268.0030222 |
| 15 | 109 | 0.060452778 | 7.790471972 | 257.0855751 |
| 16 | 127 | 0.053307778 | 9.620931251 | 272.1132681 |
| 17 | 47 | 0.087186389 | 5.986329189 | 397.6451006 |
| 18 | 46 | 0.028733333 | 10.82102214 | 834.1596628 |
| 19 | 35 | 0.016281944 | 6.736751138 | 690.0357952 |
| 20 | 12 | 0.005785 | 6.702349264 | 1848.731339 |
| 21 | 26 | 0.023224722 | 5.464826062 | 739.0126321 |
| 22 | 36 | 0.047408056 | 4.89214708 | 486.9045274 |
| 23 | 58 | 0.038834722 | 6.963915817 | 423.3580547 |
| 24 | 9 | 0.005141944 | 6.442706182 | 2118.218621 |
| 25 | 21 | 0.012038333 | 9.259085139 | 1573.603565 |
| 26 | 39 | 0.036831667 | 5.919393051 | 538.0576504 |

此题第二部分需要先整理数据，使用Python的代码思路如下：

1.导入数据，转换成标准时间，并筛选所需要的时间内1，14线圈的信息并排序

2.使用同2.1的方式创造的数据切片，创建新的表格，并遍历源数据完善新表格并输出。

密度为流量除以时间

代码如下：（见附件：ITS作业2.2.py）

import numpy as np

import pandas as pd

import datetime

import math

if \_\_name\_\_ == "\_\_main\_\_":

    print('选择车道:1,14')

    a = pd.read\_csv("C:\\Users\\jhong\\Documents\\GitHub\\OpenITS-HeFei-Analyzation\\示范区地磁检测数据(2016.06.22~2016.06.30).csv")

    a['INTIME']=pd.to\_datetime(a['INTIME'])

    a=a.loc[(a['INTIME']>='2016/6/25 8:00:00')&(a['INTIME']<='2016/6/25 9:00:00')&(a['DETECT\_ID'].isin(['0002'])&(a['DETECTOR\_NUMBER'].isin(['1','14'])))]

    a=(a.sort\_values(by="INTIME"))[['DETECTOR\_NUMBER','OCC\_TIME','INTIME']].reset\_index(drop=True)

    car\_length = 3.5

    circuit\_length = 1.0

    timelags=[] #用于形成数据切片

    for i in range(0,12):

        timelags.append(datetime.datetime.strptime('2016/6/25 8:00:00', '%Y/%m/%d %H:%M:%S')+i\*datetime.timedelta(minutes=5))

    b=pd.DataFrame(columns=['FLOW','SPEED','DENSITY'],index=[])

    b.insert(0,'时间段(起始)', timelags)

    b=b.set\_index('时间段(起始)').fillna(0.0)

    for index,row in a.iterrows():

        for i in timelags:

            if (row['INTIME']>i and row['INTIME']<=(i+datetime.timedelta(minutes=5))):

                b.loc[i,'FLOW']=b.loc[i,'FLOW']+1

                b.loc[i,'SPEED']=b.loc[i,'SPEED']+float((car\_length+circuit\_length)/(row['OCC\_TIME']/1000))

    flowsum = 0

    for index,row in b.iterrows():

        row['SPEED']=row['SPEED']/row['FLOW']

        row['FLOW']=math.ceil(row['FLOW']/2)

        flowsum = flowsum + row['FLOW']

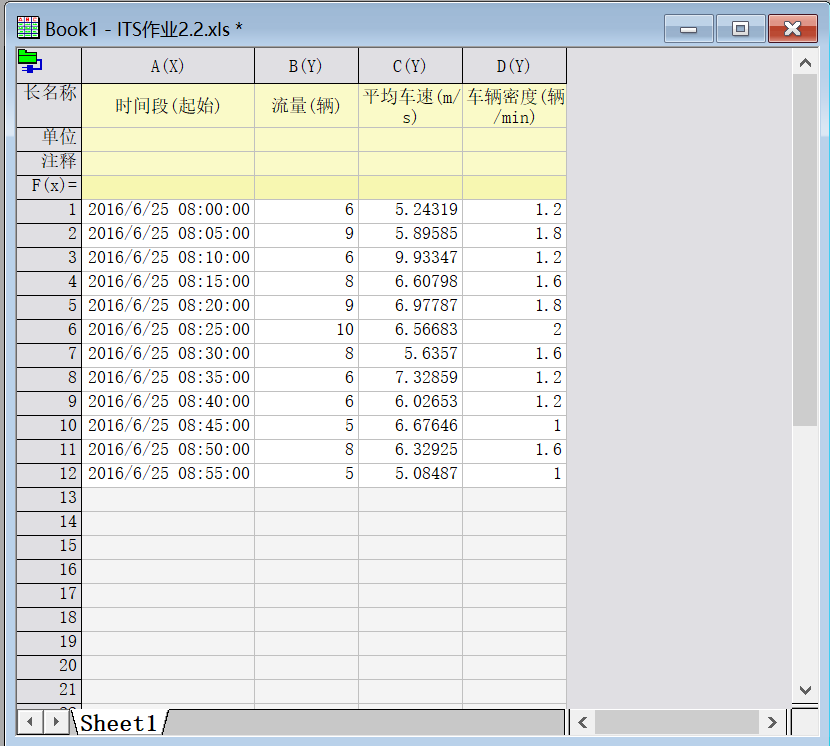
    for index,row in b.iterrows():

        row['DENSITY']=row['FLOW']/5

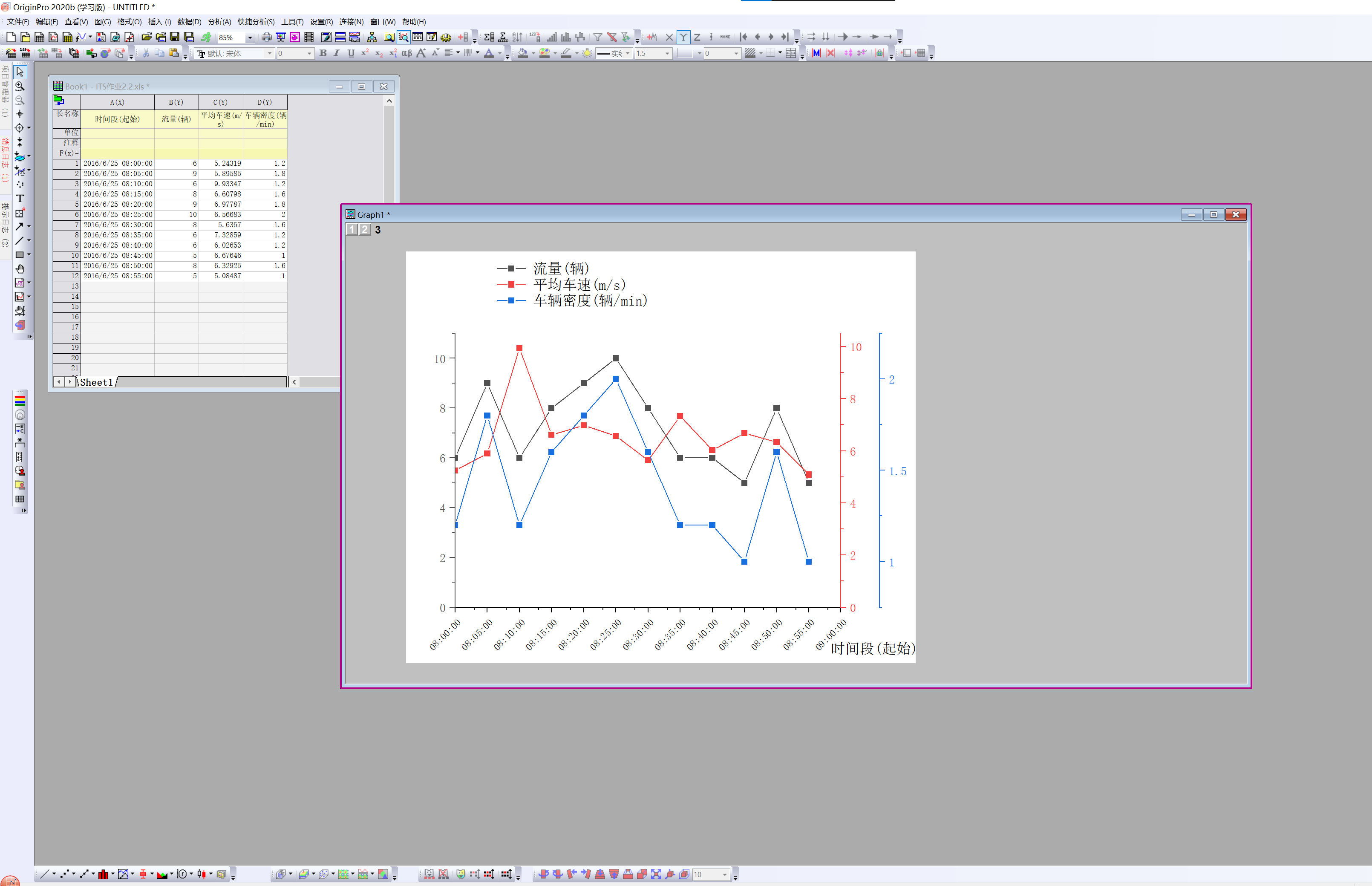
    b.columns = ['流量(辆)','平均车速(m/s)','车辆密度(辆/min)']

    b.to\_excel('ITS作业2.2.xls')

画图，将excel数据导入Origin软件中：



使用Origin 多y轴制图功能，调整线型、颜色、图例得到：





即得到所要图像，制图工程文件请见附件中的文件：ITS作业2.2 制图.opju