import pandas as pd

import os

import numpy as np

import matplotlib.pyplot as plt

os.chdir(r'G:\数据')

data= pd.read\_csv('task1\_X2.csv', engine = 'python', encoding = 'gbk')# 数据读取

index=['消费']

data\_cost=data.loc[data['Type'].isin(index)]#提取消费数据

#分组求刷卡次数

group=data\_cost[['CardNo','Major','Sex','Money']].groupby(by=['CardNo','Major','Sex'])

ddd=group.count()

print(ddd)

#转为表格形式

order=pd.DataFrame(ddd)

order.columns=['刷卡次数']

su=order.reset\_index()

su.head()

#求18级学生人均刷卡次数

sk=su['刷卡次数'].sum()

print('18级学生刷卡总次数为：',sk)

num=su['CardNo'].count()

print('18级学生人数为：',num)

zxc=sk/num

print('18级学生人均刷卡次数为：',zxc)

#18级学生本月人均消费金额

ttt=group.sum()

sss=pd.DataFrame(ttt)

sss.columns=['本月消费总金额']

sd=sss.reset\_index()

money\_all=sd['本月消费总金额'].sum()

print('18级学生本月消费总额为：',money\_all)

xcv=money\_all/num

print('18级学生本月人均消费金额为：',xcv)

#提取各专业数据

index=['18连锁经营']

data1=data\_cost.loc[data['Major'].isin(index)]

data1.shape

data1\_boy = data1.loc[(data1['Sex'].apply(lambda x: x=='男')), :]

print(data1\_boy.shape)

data1\_boy.to\_csv('2.csv',index = False, encoding = 'gbk')#保存文件

data1\_boymoney=data1\_boy['Money'].sum()

print('18连锁经营专业男生本月消费总额',data1\_boymoney)

data1\_girl = data1.loc[(data1['Sex'].apply(lambda x: x=='女')), :]#提取该专业女生数据

data1\_girlmoney=data1\_girl['Money'].sum()

print('18连锁经营专业女生本月消费总额',data1\_girlmoney)#计算消费总额

from collections import Counter

dict1=Counter(data1\_boy['CardNo'])

def returnSum(myDict):

sum = 0

for i in myDict:

sum = sum + myDict[i]

return sum

print("18连锁经营专业男生刷卡次数 :", returnSum(dict1))

print('18连锁经营专业男生人数为：',len(dict1))

data1\_boysk=returnSum(dict1)

data1\_boynum=len(dict1)

data1\_boymoneyavg=data1\_boymoney/data1\_boynum

data1\_boyskavg=data1\_boysk/data1\_boynum

print('18连锁经营专业男生本月人均消费金额为：',data1\_boymoneyavg)

print('18连锁经营专业男生本月人均刷卡次数为：',data1\_boyskavg)

a1=data1\_boymoneyavg

a2=data1\_boyskavg

from collections import Counter

dict2=Counter(data1\_girl['CardNo'])

print("18连锁经营专业女生刷卡次数 :", returnSum(dict2))

print('18连锁经营专业女生人数为：',len(dict2))

data1\_girlsk=returnSum(dict2)

data1\_girlnum=len(dict2)

data1\_girlmoneyavg=data1\_girlmoney/data1\_girlnum

data1\_girlskavg=data1\_girlsk/data1\_girlnum

print('18连锁经营专业女生本月人均消费金额为：',data1\_girlmoneyavg)

print('18连锁经营专业女生本月人均刷卡次数为：',data1\_girlskavg)

b1=data1\_girlmoneyavg

b2=data1\_girlskavg

index=['18国际金融']

data2=data\_cost.loc[data['Major'].isin(index)]

print(data2.shape)

data2\_boy = data2.loc[(data2['Sex'].apply(lambda x: x=='男')), :]

print(data2\_boy.shape)

data2\_boymoney=data2\_boy['Money'].sum()

print('18国际金融专业男生本月消费总额',data2\_boymoney)

from collections import Counter

dict3=Counter(data2\_boy['CardNo'])

print("18国际金融专业男生刷卡次数 :", returnSum(dict3))

print('18国际金融专业男生人数为：',len(dict3))

data2\_boysk=returnSum(dict3)

data2\_boynum=len(dict3)

data2\_boymoneyavg=data2\_boymoney/data2\_boynum

data2\_boyskavg=data2\_boysk/data2\_boynum

print('18国际金融专业男生本月人均消费金额为：',data2\_boymoneyavg)

print('18国际金融专业男生本月人均刷卡次数为：',data2\_boyskavg)

c1=data2\_boymoneyavg

c2=data2\_boyskavg

data2\_girl = data2.loc[(data2['Sex'].apply(lambda x: x=='女')), :]#提取该专业女生数据

print(data2\_girl.shape)

data2\_girlmoney=data2\_girl['Money'].sum()

print('18国际金融专业女生本月消费总额',data2\_girlmoney)#计算消费总额

from collections import Counter

dict4=Counter(data2\_girl['CardNo'])

print("18国际金融专业女生刷卡次数 :", returnSum(dict4))

print('18国际金融专业女生人数为：',len(dict4))

data2\_girlsk=returnSum(dict4)

data2\_girlnum=len(dict4)

data2\_girlmoneyavg=data2\_girlmoney/data2\_girlnum

data2\_girlskavg=data2\_girlsk/data2\_girlnum

print('18国际金融专业女生本月人均消费金额为：',data2\_girlmoneyavg)

print('18国际金融专业女生本月人均刷卡次数为：',data2\_girlskavg)

d1=data2\_girlmoneyavg

d2=data2\_girlskavg

index=['18会计']

data3=data\_cost.loc[data['Major'].isin(index)]

print(data3.shape)

data3\_boy = data3.loc[(data3['Sex'].apply(lambda x: x=='男')), :]

print(data3\_boy.shape)

data3\_boymoney=data3\_boy['Money'].sum()

print('18会计专业男生本月消费总额',data3\_boymoney)

from collections import Counter

dict5=Counter(data3\_boy['CardNo'])

print("18会计专业男生刷卡次数 :", returnSum(dict5))

print('18会计专业男生人数为：',len(dict5))

data3\_boysk=returnSum(dict5)

data3\_boynum=len(dict5)

data3\_boymoneyavg=data3\_boymoney/data3\_boynum

data3\_boyskavg=data3\_boysk/data3\_boynum

print('18会计专业男生本月人均消费金额为：',data3\_boymoneyavg)

print('18会计专业男生本月人均刷卡次数为：',data3\_boyskavg)

e1=data3\_boymoneyavg

e2=data3\_boyskavg

data3\_girl = data3.loc[(data3['Sex'].apply(lambda x: x=='女')), :]#提取该专业女生数据

print(data3\_girl.shape)

data3\_girlmoney=data3\_girl['Money'].sum()

print('18会计专业女生本月消费总额',data3\_girlmoney)#计算消费总额

from collections import Counter

dict6=Counter(data3\_girl['CardNo'])

print("18会计专业女生刷卡次数 :", returnSum(dict6))

print('18会计专业女生人数为：',len(dict6))

data3\_girlsk=returnSum(dict6)

data3\_girlnum=len(dict6)

data3\_girlmoneyavg=data3\_girlmoney/data3\_girlnum

data3\_girlskavg=data3\_girlsk/data3\_girlnum

print('18会计专业女生本月人均消费金额为：',data3\_girlmoneyavg)

print('18会计专业女生本月人均刷卡次数为：',data3\_girlskavg)

f1=data3\_girlmoneyavg

f2=data3\_girlskavg

index=['18环境艺术']

data4=data\_cost.loc[data['Major'].isin(index)]

print(data4.shape)

data4\_boy = data4.loc[(data4['Sex'].apply(lambda x: x=='男')), :]

print(data4\_boy.shape)

data4\_boymoney=data4\_boy['Money'].sum()

print('18环境艺术专业男生本月消费总额',data4\_boymoney)

from collections import Counter

dict7=Counter(data4\_boy['CardNo'])

print("18环境艺术专业男生刷卡次数 :", returnSum(dict7))

print('18环境艺术专业男生人数为：',len(dict7))

data4\_boysk=returnSum(dict7)

data4\_boynum=len(dict7)

data4\_boymoneyavg=data4\_boymoney/data4\_boynum

data4\_boyskavg=data4\_boysk/data4\_boynum

print('18环境艺术专业男生本月人均消费金额为：',data4\_boymoneyavg)

print('18环境艺术专业男生本月人均刷卡次数为：',data4\_boyskavg)

g1=data4\_boymoneyavg

g2=data4\_boyskavg

data4\_girl = data4.loc[(data4['Sex'].apply(lambda x: x=='女')), :]#提取该专业女生数据

print(data4\_girl.shape)

data4\_girlmoney=data4\_girl['Money'].sum()

print('18环境艺术专业女生本月消费总额',data4\_girlmoney)#计算消费总额

from collections import Counter

dict8=Counter(data4\_girl['CardNo'])

print("18环境艺术专业女生刷卡次数 :", returnSum(dict8))

print('18环境艺术专业女生人数为：',len(dict8))

data4\_girlsk=returnSum(dict8)

data4\_girlnum=len(dict8)

data4\_girlmoneyavg=data4\_girlmoney/data4\_girlnum

data4\_girlskavg=data4\_girlsk/data4\_girlnum

print('18环境艺术专业女生本月人均消费金额为：',data4\_girlmoneyavg)

print('18环境艺术专业女生本月人均刷卡次数为：',data4\_girlskavg)

h1=data4\_girlmoneyavg

h2=data4\_girlskavg

index=['18商务英语']

data5=data\_cost.loc[data['Major'].isin(index)]

print(data5.shape)

data5\_boy = data5.loc[(data5['Sex'].apply(lambda x: x=='男')), :]

print(data5\_boy.shape)

data5\_boymoney=data5\_boy['Money'].sum()

print('18商务英语专业男生本月消费总额',data5\_boymoney)

from collections import Counter

dict9=Counter(data5\_boy['CardNo'])

print("18商务英语专业男生刷卡次数 :", returnSum(dict9))

print('18商务英语专业男生人数为：',len(dict9))

data5\_boysk=returnSum(dict9)

data5\_boynum=len(dict9)

data5\_boymoneyavg=data5\_boymoney/data5\_boynum

data5\_boyskavg=data5\_boysk/data5\_boynum

print('18商务英语专业男生本月人均消费金额为：',data5\_boymoneyavg)

print('18商务英语专业男生本月人均刷卡次数为：',data5\_boyskavg)

l1=data5\_boymoneyavg

l2=data5\_boyskavg

data5\_girl = data5.loc[(data5['Sex'].apply(lambda x: x=='女')), :]#提取该专业女生数据

print(data5\_girl.shape)

data5\_girlmoney=data5\_girl['Money'].sum()

print('18商务英语专业女生本月消费总额',data5\_girlmoney)#计算消费总额

from collections import Counter

dict10=Counter(data5\_girl['CardNo'])

print("18商务英语专业女生刷卡次数 :", returnSum(dict10))

print('18商务英语专业女生人数为：',len(dict10))

data5\_girlsk=returnSum(dict10)

data5\_girlnum=len(dict10)

data5\_girlmoneyavg=data5\_girlmoney/data5\_girlnum

data5\_girlskavg=data5\_girlsk/data5\_girlnum

print('18商务英语专业女生本月人均消费金额为：',data5\_girlmoneyavg)

print('18商务英语专业女生本月人均刷卡次数为：',data5\_girlskavg)

m1=data5\_girlmoneyavg

m2=data5\_girlskavg

index=['18建筑工程']

data6=data\_cost.loc[data['Major'].isin(index)]

print(data6.shape)

data6\_boy = data6.loc[(data6['Sex'].apply(lambda x: x=='男')), :]

print(data6\_boy.shape)

data6\_boymoney=data6\_boy['Money'].sum()

print('18建筑工程专业男生本月消费总额',data6\_boymoney)

from collections import Counter

dict11=Counter(data6\_boy['CardNo'])

print("18建筑工程专业男生刷卡次数 :", returnSum(dict11))

print('18建筑工程专业男生人数为：',len(dict11))

data6\_boysk=returnSum(dict11)

data6\_boynum=len(dict11)

data6\_boymoneyavg=data6\_boymoney/data6\_boynum

data6\_boyskavg=data6\_boysk/data6\_boynum

print('18建筑工程专业男生本月人均消费金额为：',data6\_boymoneyavg)

print('18建筑工程专业男生本月人均刷卡次数为：',data6\_boyskavg)

n1=data6\_boymoneyavg

n2=data6\_boyskavg

data6\_girl = data6.loc[(data6['Sex'].apply(lambda x: x=='女')), :]#提取该专业女生数据

print(data6\_girl.shape)

data6\_girlmoney=data6\_girl['Money'].sum()

print('18建筑工程专业女生本月消费总额',data6\_girlmoney)#计算消费总额

from collections import Counter

dict12=Counter(data6\_girl['CardNo'])

print("18建筑工程专业女生刷卡次数 :", returnSum(dict12))

print('18建筑工程专业女生人数为：',len(dict12))

data6\_girlsk=returnSum(dict12)

data6\_girlnum=len(dict12)

data6\_girlmoneyavg=data6\_girlmoney/data6\_girlnum

data6\_girlskavg=data6\_girlsk/data6\_girlnum

print('18建筑工程专业女生本月人均消费金额为：',data6\_girlmoneyavg)

print('18建筑工程专业女生本月人均刷卡次数为：',data6\_girlskavg)

o1=data6\_girlmoneyavg

o2=data6\_girlskavg

plt.rcParams['font.sans-serif']='SimHei' # 显示中文

plt.rcParams['axes.unicode\_minus']=False # 显示负号

fig=plt.figure(figsize=(6,5))

x=['18连锁经营','18国际金融','18会计','18环境艺术','18商务英语','18建筑工程专业']

y1=[a1,c1,e1,g1,l1,n1]

plt.bar(x,y1)

plt.title('不同专业男生人均消费金额柱形图')

plt.savefig('不同专业男生人均消费金额柱形图.png')

plt.show()

plt.rcParams['font.sans-serif']='SimHei' # 显示中文

plt.rcParams['axes.unicode\_minus']=False # 显示负号

fig=plt.figure(figsize=(6,5))

x=['18连锁经营','18国际金融','18会计','18环境艺术','18商务英语','18建筑工程专业']

y2=[a2,c2,e2,g2,l2,n2]

plt.bar(x,y2)

plt.title('不同专业男生人均刷卡次数柱形图')

plt.savefig('不同专业男生人均刷卡次数柱形图.png')

plt.show()

plt.rcParams['font.sans-serif']='SimHei' # 显示中文

plt.rcParams['axes.unicode\_minus']=False # 显示负号

fig=plt.figure(figsize=(6,5))

x=['18连锁经营','18国际金融','18会计','18环境艺术','18商务英语','18建筑工程专业']

y1=[b1,d1,f1,h1,m1,o1]

plt.bar(x,y1)

plt.title('不同专业女生人均消费金额柱形图')

plt.savefig('不同专业女生人均消费金额柱形图.png')

plt.show()

plt.rcParams['font.sans-serif']='SimHei' # 显示中文

plt.rcParams['axes.unicode\_minus']=False # 显示负号

fig=plt.figure(figsize=(6,5))

x=['18连锁经营','18国际金融','18会计','18环境艺术','18商务英语','18建筑工程专业']

y2=[b2,d2,f2,h2,m2,o2]

plt.bar(x,y2)

plt.title('不同专业女生人均刷卡次数柱形图')

plt.savefig('不同专业女生人均刷卡次数柱形图.png')

plt.show()

plt.rcParams['font.sans-serif']='SimHei' # 显示中文

plt.rcParams['axes.unicode\_minus']=False # 显示负号

fig=plt.figure(figsize=(6,5))

x=['男','女']

y3=[n1,o1]

plt.bar(x,y3)

plt.title('18建筑工程专业不同性别人均消费柱形图')

plt.savefig('18建筑工程专业不同性别人均消费柱形图.png')

plt.show()

#并列柱形图

plt.rcParams['font.sans-serif']='SimHei' # 显示中文

plt.rcParams['axes.unicode\_minus']=False # 显示负号

a=[b1,d1,f1,h1,m1,o1]#各专业女生人均消费金额

b=[a1,c1,e1,g1,l1,n1]#各专业男生人均消费金额

name=['18连锁经营','18国际金融','18会计','18环境艺术','18商务英语','18建筑工程专业']

fig=plt.figure(figsize=(10,8))

x1=list(range(len(a)))

total\_width,n=0.8,2

width=total\_width/n#设置条形的宽度

plt.bar(x1,a,width=width,label='女生',fc='g')#'fc'设置条形颜色

for i in range(len(x1)):

x1[i] +=width

plt.bar(x1,b,width=width,label='男生',tick\_label=name,fc='b')

plt.legend()

plt.savefig('不同专业不同性别学生人均消费金额柱形图.png')

plt.show()

#并列柱形图

plt.rcParams['font.sans-serif']='SimHei' # 显示中文

plt.rcParams['axes.unicode\_minus']=False # 显示负号

c=[b2,d2,f2,h2,m2,o2]#各专业女生人均刷卡次数

d=[a2,c2,e2,g2,l2,n2]#各专业男生人均刷卡次数

name=['18连锁经营','18国际金融','18会计','18环境艺术','18商务英语','18建筑工程专业']

fig=plt.figure(figsize=(10,8))

x2=list(range(len(c)))

total\_width,n=0.8,2

width=total\_width/n#设置条形的宽度

plt.bar(x2,c,width=width,label='女生',fc='g')#'fc'设置条形颜色

for i in range(len(x2)):

x2[i] +=width

plt.bar(x2,d,width=width,label='男生',tick\_label=name,fc='b')

plt.legend()

plt.savefig('不同专业不同性别学生人均刷卡次数柱形图.png')

plt.show()